Received: 18 August 2025 | Revised: 8 June 2026 | Accepted: 23 August 2026 | Published online: 8 September 2026

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–9
DOI: 10.47852/bonviewAIA62027257

BON VIEW PUBLISHING

RESEARCH ARTICLE

MEAFNet: Multi-Expert Attention Fusion Network for Robust Grape Leaf Disease Classification

Hoang-Tu Vo[1,*], Nhon Nguyen Thien[1], Kheo Chau Mui[1], Phuc Pham Tien[1] and Huan Lam Le[1]

[1] *Information Technology Department, FPT University, Vietnam*

Abstract: Reliable detection of grape leaf diseases is critical for timely treatment and quality control in vineyard management. This paper presents MEAFNet, a novel Multi- Expert Attention Fusion Network designed to improve the reliability and accuracy of grape leaf disease classification. The proposed model integrates feature extraction mechanisms from the strengths of three pretrained convolutional neural networks such as MobileNetV2, EfficientNetB0, and DenseNet121 as parallel expert branches. A compact Transformer processes the three projected expert descriptors. Its self-attention operation allows each branch representation to be updated using evidence from the other two branches before classification. Experimental evaluations performed using a publicly available dataset of diseased grape leaves show that MEAFNet obtains an accuracy rate of 99.78% in classification within just 60 training epochs. The model achieves an F1-score of 1.000 on the Leaf Blight classes and maintains above 0.990 over all target classes. These results indicate the effectiveness and generalization capability of the proposed method in real-world agricultural environments. Overall, MEAFNet provides an effective framework for automated grape leaf disease classification and offers a promising foundation for intelligent disease monitoring in precision viticulture.

Keywords: deep learning, multi-expert network, attention mechanism, transfer learning, grape leaf disease

1. Introduction

Agriculture is essential for maintaining stable food supplies and security and economic stability across the globe [1–5]. In the last few years, the cultivation of grapes has gained increasing importance due to their commercial value and nutritional benefits [6–9]. However, grapevines are prone to various types of leaf diseases, which may result in major crop losses if not accurately and promptly detected [10, 11]. Field diagnosis commonly depends on visual inspection by trained personnel; this limits throughput and introduces observer-dependent variation. With recent advances in computer vision and deep learning, automated visual disease detection systems have become a promising solution in precision agriculture [12–16]. Visual recognition tasks, such as plant disease classification, have greatly benefited from the impressive performance of convolutional neural networks (CNNs). Nevertheless, A classifier trained with one CNN may be sensitive to acquisition changes, especially when lesions are partly occluded or visually close across classes among disease types. In order to resolve these problems, we propose a novel model named MEAFNet (Multi-Expert Attention Fusion Network), designed to effectively combine the complementary capabilities of several expert convolutional networks with attention-based modeling. Our architecture leverages a three-branch expert design based on MobileNetV2 [17], DenseNet121 [18], and EfficientNetB0 [19] to extract discriminative features from images of grape leaves. The extracted features are then fused through a self-attention mechanism [20] that learns inter-token dependencies and enhances feature representation. Furthermore, the inclusion of positional encoding and residual learning improves the model's robustness and discriminative capacity. Extensive experiments were conducted on a benchmark grape leaf disease dataset, where MEAFNet achieved superior classification performance relative to existing leading models. The results demonstrate that our method not only boosts precision but also improves generalization across unseen disease patterns and environmental conditions. This work offers the following key contributions:

1) This study introduces MEAFNet, a novel multi-expert attention-based architecture for grape leaf disease classification, which integrates multiple CNN backbones to extract diverse and complementary features.

2) We propose a token-wise self-attention fusion module with positional encoding to effectively capture inter-feature dependencies and enhance feature representation.

3) We show the efficiency of the proposed approach through extensive experiments on standard evaluation datasets, achieving superior accuracy and robustness in comparison with current state-of-the-art approaches.

The rest of this paper is organized into the following sections: Section 2 examines the related literature, while Section 3 introduces the dataset. Section 4 details the proposed MEAFNet architecture. Section 5 introduces evaluation metrics, results, discussion, and

***Corresponding author:** Hoang-Tu Vo, Information Technology Department, FPT University, Vietnam. Email: tuvh6@fe.edu.vn

comparisons, and Section 6 presents the conclusion along with prospective research directions.

2. Literature Review

Grape leaf disease recognition has been a widely researched area in agricultural informatics, leveraging advancements in image processing and deep learning to support early diagnosis and yield protection. Traditional approaches relied heavily on handcrafted features such as texture, color, and shape descriptors [21–25], which often lacked generalization across varying environmental conditions. With the rise of deep learning, CNNs have shown significant promise in learning hierarchical and spatially invariant features directly from images [14, 26–29]. Models such as MobileNetV2 [17], DenseNet121 [18], and EfficientNetB0 [19] have been successfully adapted for plant disease classification with encouraging results [30–34]. Recent studies have further introduced attention mechanisms into CNN architectures to improve the model's focus on lesion regions, leading to improved overall performance [35–38]. In the specific context of vine leaf disease detection, several structural improvements have been proposed in recent literature. For instance, studies [39, 40] employed a DenseNet-based approach to enhance feature reuse and improve classification accuracy. Meanwhile, other works such as [41, 42] applied Vision Transformers (ViT) to model long-range dependencies. However, training a ViT without substantial pretraining or data augmentation can be demanding in small agricultural datasets, both in sample efficiency and computation, making them less practical for standard agricultural applications. To balance feature diversity and redundancy, fusion-based models and multi-branch architectures utilizing dynamic weighting or ensemble learning have also emerged [43, 45]. Despite these continuous advancements, a critical research gap remains in effectively combining local feature extraction with global contextual awareness without introducing excessive noise. The limitation is not simply the depth of an individual CNN. A single feature extractor may underrepresent subtle lesions when background texture, illumination, or visually similar symptoms dominate the image. Existing multi-backbone systems partly address this issue, but many combine branch outputs through fixed concatenation or averaging. Such operations preserve diversity without determining which inter-branch relationships are useful for a particular image. This motivates a learnable fusion stage that compares the projected expert descriptors before classification and can suppress redundant evidence. In this work, we justify this need by proposing MEAFNet: a Multi-Expert Attention Fusion Network. By combining diverse CNN experts with a lightweight, token-wise Transformer attention fusion mechanism, our method successfully filters background noise and captures inter-branch dependencies, ultimately learning richer and more discriminative representations for robust disease classification.

3. Dataset

Experiments used the Grape Disease Dataset, which comprises 9027 RGB leaf images from four classes: Black Rot, ESCA, Healthy, and Leaf Blight. We separated the images into non-overlapping training, validation, and test subsets. The class labels provide the basis for evaluating four-way classification; they do not, by themselves, guarantee that the test set reflects the variability encountered in vineyards. To reduce potential bias caused by class imbalance, the samples were distributed in a relatively uniform manner across all subsets. Table 1 reports the distribution of images among different categories and data splits.

For compatibility with deep CNN architectures, all images were processed to have dimensions of 224×224. This dataset poses a considerable challenge because of substantial intra-class variability and variations in illumination, leaf orientation, and background complexity. Representative images from the dataset are presented in Figure 1.

4. Research Methodology

This section describes the architecture and main design considerations of MEAFNet for grape leaf disease classification. The framework employs MobileNetV2, EfficientNetB0, and DenseNet121 as parallel pretrained feature extractors, enabling the model to obtain complementary representations from the same input image. Because these CNN backbones produce feature vectors with different dimensions, their outputs are projected into a shared embedding space before fusion. The projected descriptors are then treated as expert tokens and processed by a compact Transformer encoder. Through self-attention, each token is updated using information obtained from the other expert branches, allowing dependencies among the CNN representations to be learned before classification. The resulting contextualized features are aggregated and passed to the classification head to predict the corresponding grape leaf disease category. The following subsections explain each component and its role in the complete framework.

4.1. Overview of MEAFNet architecture

Figure 2 illustrates the overall structure of MEAFNet. Multiple CNN backbones extract complementary features from each input image. A Transformer-based module subsequently integrates the resulting branch representations.

The MEAFNet framework is composed of four main components:

1) Multi-Expert Feature Extractors: Three different CNN backbones such as EfficientNetB0, MobileNetV2, and DenseNet121 are utilized to extract semantically rich and diverse features from the input images.

Table 1
Class-wise distribution of images in the vine leaf disease dataset

Class	Training	Validation	Testing	Total
Black Rot (BR)	1511	377	472	2360
ESCA (ES)	1536	384	480	2400
Healthy leaves (HL)	1354	338	423	2115
Leaf Blight (LB)	1378	344	430	2152
Total	5779	1443	1805	9027

Figure 1
Visual examples of images included in the dataset

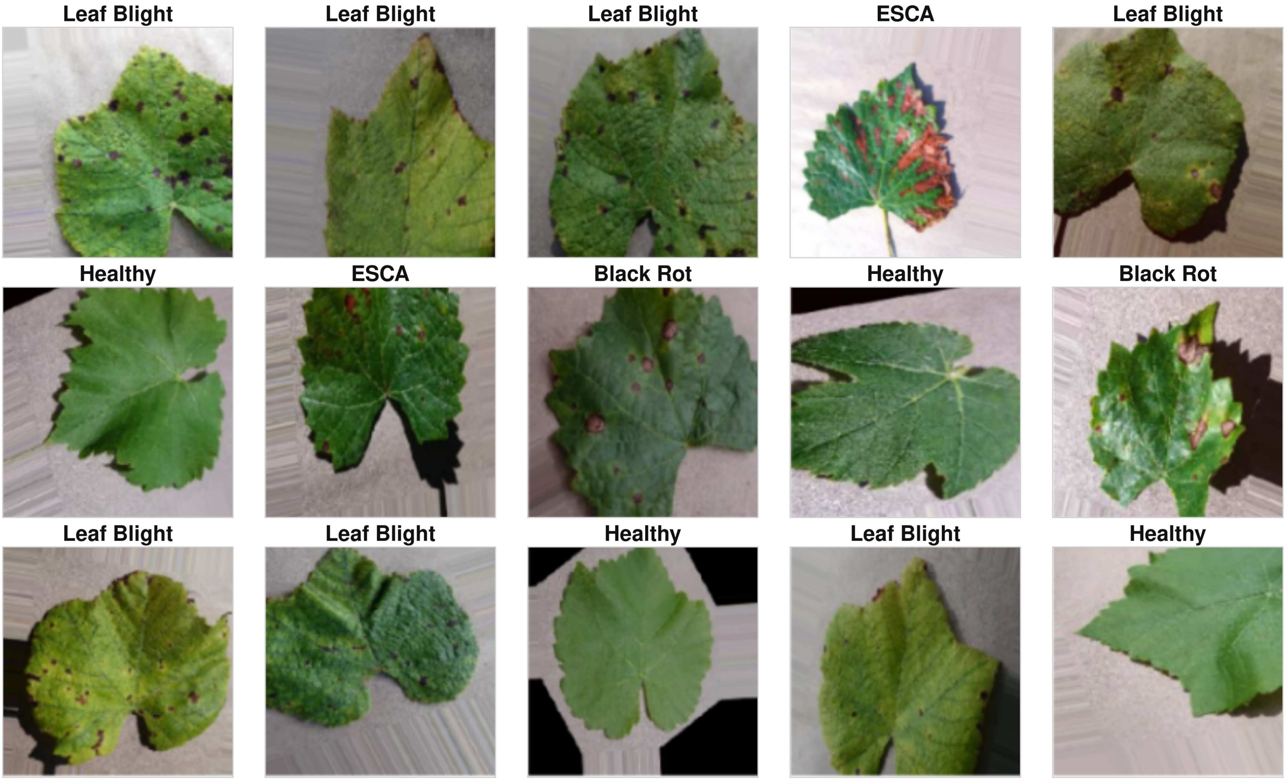
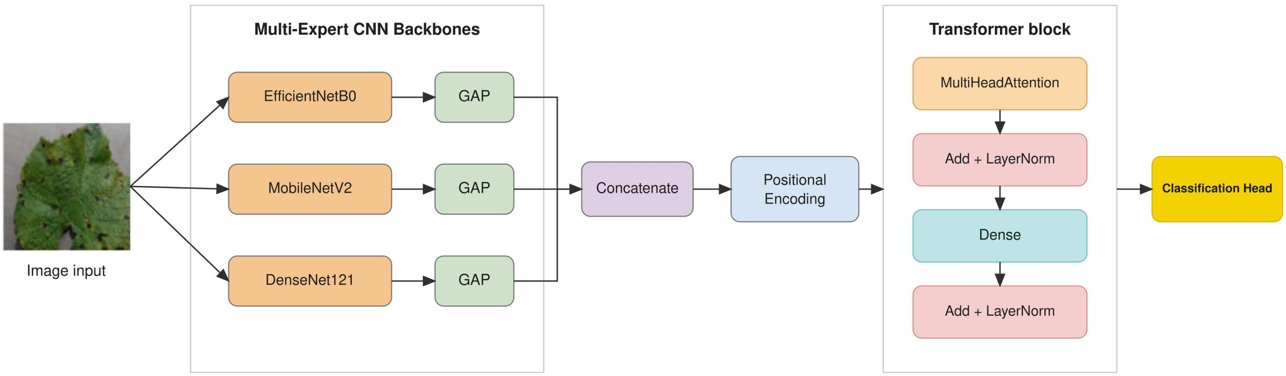


Figure 2
Overview of the proposed MEAFNet architecture integrating multiple expert CNNs with Transformer-based feature fusion



2) Tokenization and Positional Encoding: Feature vectors obtained from each expert are stacked and treated as sequential tokens. A positional encoding module is applied to inject position-aware information into the token sequence.

3) Transformer-Based Fusion: A lightweight Transformer encoder is employed to learn inter-token dependencies and globally integrate features across the experts.

4) Classification Head: The resulting fused feature representation is forwarded to a dense layer and softmax classifier to produce the final predictions.

4.2. Multi-expert feature extractors

Given an input RGB image with dimensions of 224 × 224 × 3, MEAFNet simultaneously feeds the input into three frozen pretrained CNNs: EfficientNetB0, MobileNetV2, and DenseNet121. We remove each backbone's native classifier and average the last convolutional feature maps spatially, yielding one descriptor per expert. These extracted features are next processed by a dense layer, followed by ReLU activation to map each to a common dimensionality $D = 256$:

$$F_i = \mathrm{ReLU}\left(\mathrm{Dense}_{256}\left(\mathrm{GAP}\left(CNN_i\left(x\right)\right)\right)\right)$$

with $i = 1, 2, 3$ and x corresponding to the input image.

4.3. Tokenization and positional encoding

The three feature vectors $\{F_1, F_2, F_3\}$ are stacked to form a token sequence:

$$T = \mathrm{Stack}\left(F_1, F_2, F_3\right) \in R^{3\times256}$$

To preserve positional relationships among the tokens, a learnable positional encoding is added:

$$T' = T + \text{Embedding}_{\text{pos}}(i), \quad i \in \{1, 2, 3\}$$

4.4. Transformer encoder for feature fusion

The enriched token sequence T' is passed through two consecutive Transformer encoder blocks, featuring multi-head attention, feed-forward components, and residual pathways combined with layer normalization:

$$Z = \text{TransformerEncoder}\left(T'\right)$$

The Transformer encoder enables global context modeling among expert features and captures complex inter-dependencies. In our implementation, each encoder block uses 4 attention heads and a feed-forward dimension of 128.

4.5. Classification head

The output token embeddings from the final Transformer layer are aggregated using Global Average Pooling (GAP) across the token dimension:

$$z = \text{GAP}(Z)$$

This stage is followed by a 128-unit dense layer (L2-regularized), a dropout layer for regularization, and a final softmax output layer for classification:

$$\hat{y} = \text{Softmax}\left(Dropout\left(Dense_{128}(z)\right)\right)$$

4.6. Loss function and optimization

To train the proposed MEAFNet model, we utilized a multi-class cross-entropy objective, widely adopted for multi-class prediction problems. The loss function is defined as follows:

$$\text{L}_{\text{CE}} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

where C indicates the total count of categories, y_i denotes the true class label, and $(\hat{y}_i)$ the output probability associated with class i.

To enhance the generalization across unseen data, we integrated an L2 regularization term into the total loss:

$$\text{L}_{\text{total}} = \text{L}_{\text{CE}} + \lambda \sum_{j=1}^{N} |w_j|^2$$

Here, w_j denotes the weights of the j-th layer, N is the number of trainable layers, and λ is the regularization coefficient set empirically to 1×10^{-4}.

5. Experimental Results

5.1. Evaluation metrics

To objectively assess the performance of the proposed model, we utilize a set of standard classification metrics. These metrics are particularly important in our context where class imbalance may exist, and a high accuracy alone may not reflect true performance across all categories. To compute these metrics, we use the following equations:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

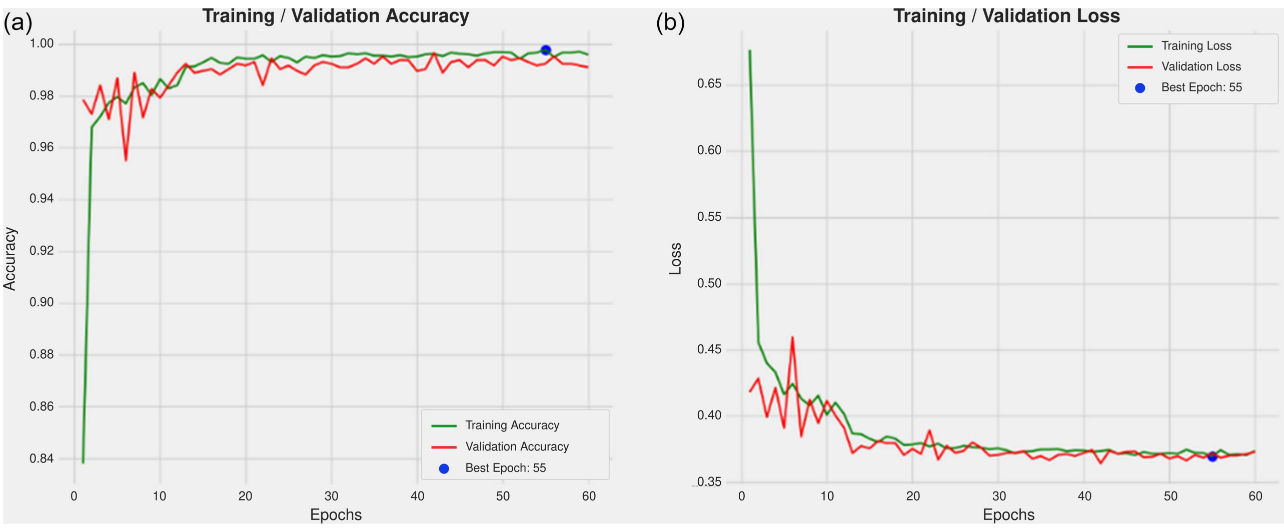
$$\text{F1\_Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Here, TP, FP, and FN refer to the counts of true positive, false positive, and false negative predictions, respectively.

5.2. Results and discussion

Figure 3 shows the training and validation accuracy and loss curves over 60 epochs. The proposed model demonstrates rapid convergence within the first 10 epochs, achieving over 98% accuracy on both training and validation datasets. The training accuracy gradually improves and stabilizes at nearly 100%,

Figure 3
(a) Accuracy and (b) loss curves for training and validation over 60 training epochs



while the validation accuracy consistently remains above 99% across epochs, indicating strong generalization capability. The loss curves in Figure 3(b) show a steady reduction in both training and validation losses, and they converge after approximately 20 epochs. Validation loss reached its minimum at epoch 55, and the weights from that epoch were retained for evaluation. Training and validation curves remained close over the later epochs, providing no obvious visual indication of severe overfitting. This observation should nevertheless be treated as descriptive evidence; performance on an independently collected dataset would be needed to assess distributional generalization.

Table 2 shows the classification effectiveness of the proposed architecture on the vine leaf disease dataset across four classes. The model attained exceptionally high accuracy, with an overall classification accuracy of 99.78%. Specifically, all per-class performance metrics exceed 99.5%, indicating robust and balanced performance. For instance, the Leaf Blight class obtained perfect scores of 1.0000 for all three metrics, while the Healthy class

Table 2
Classification performance summary of the proposed method

Disease class	Precision	Recall	F1-score	Support
BR	0.9979	0.9936	0.9958	472
ES	0.9958	0.9979	0.9969	480
HL	0.9976	1.0000	0.9988	423
LB	1.0000	1.0000	1.0000	430
Accuracy	–	–	0.9978	1805
Macro Avg	0.9978	0.9979	0.9979	1805
Weighted Avg	0.9978	0.9978	0.9978	1805

Figure 4
Confusion matrix of the developed grape leaf disease classification model, showing high accuracy across all categories

Table 3
Performance comparison of MEAFNet with existing approaches

Reference	Dataset source	Number of classes	Total images	Classification method	Classification accuracy (%)
[46]	Grapevine leaf dataset	4	3885	GoogLeNet-based CNN	94.05
[47]	Grape leaf dataset	4	9027	EfficientNet-B7 architecture	98.70
[48]	Grape leaf disease dataset	5	2500	MobileNetV2-based model	97.60
[49]	Grapevine disease dataset	4	62286	INSE-enhanced ResNet	99.47
[50]	Grape leaf image dataset	4	4062	ShuffleNet (V1/V2 variants)	99.14
[51]	PlantVillage (grape leaves)	4	4062	Support vector machine	92.00
[52]	AI Challenger 2018 (grape leaves)	7	2850	GrapeNet	86.29
[53]	Myanmar Grapevine Yard	6	6000	VGG16	98.40
[54]	Grape leaf image dataset	4	1619	UnitedModel architecture	98.57
[55]	Leaf GAN dataset	4	8124	Xception-based CNN	98.70
[56]	Grape leaf image dataset	4	14690	SwinGNet	99.03
[57]	Grape leaf image dataset	4	4062	GrapeLeafNet	99.56
[41]	Grapevine dataset	4	4062	DenseNet	98.67
This work	PlantVillage (grape leaves)	4	4062	MEAFNet (EfficientNetB0 MobileNetV2, DenseNet121 + Transformer)	99.84
This work	Grapevine leaf dataset	4	9027	MEAFNet (EfficientNetB0 MobileNetV2, DenseNet121 + Transformer)	99.78

Table 4
Ablation study comparing individual models with the proposed MEAFNet

Disease class	Precision	Recall	F1-score	Accuracy
MobileNetV2 (baseline)	0.9961	0.9963	0.9962	0.9961
EfficientNetB0 (baseline)	0.9925	0.9929	0.9927	0.9928
DenseNet121 (baseline)	0.9915	0.9909	0.9909	0.9911
MEAFNet (proposed)	0.9978	0.9979	0.9979	0.9978

achieved an F1-score of 0.9988, with perfect recall (1.0000) and a slightly lower precision (0.9976). Similarly, Black Rot and ESCA also recorded strong F1-scores of 0.9958 and 0.9969, respectively. Within the reported test split, the four classes were separated with few errors, including the disease categories that exhibit similar visual symptoms.

Figure 4 localizes the four test errors. Black Rot accounts for three errors (469 correct predictions among 472 samples), whereas ESCA has one error among 480 samples. All 423 Healthy images and all 430 Leaf Blight images were assigned to their recorded classes. Thus, 1801 of the 1805 test images were classified correctly. The matrix demonstrates strong performance on this split, but it does not by itself establish robustness to images collected under different vineyard conditions.

Table 3 presents a comparison of the proposed method and existing approaches for grapevine leaf disease classification. Various datasets with differing numbers of classes and image quantities were employed in prior studies. Among the listed studies, INSE-ResNet reported 99.47% accuracy using more than 62,000 images. MEAFNet obtained 99.78% on the 9027-image dataset and 99.84% on the PlantVillage grape subset. These figures place MEAFNet above the reported values in Table 3, but the comparison is descriptive because the studies used different datasets, class definitions, and evaluation protocols. A controlled reimplementation on a shared split would be required to claim method-level superiority.

Table 4 compares MEAFNet with each constituent backbone trained separately under the same dataset setting. MobileNetV2 was the strongest individual model at 99.61% accuracy, followed by EfficientNetB0 and DenseNet121. Fusion increased accuracy to 99.78%, a gain of 0.17 percentage points over MobileNetV2. The result supports the combined architecture, although a stricter component ablation such as concatenation without attention and attention with matched parameter counts is needed to isolate the contribution of the Transformer fusion block.

6. Conclusion

MEAFNet was developed to investigate whether features learned by different convolutional networks could be coordinated before classification. The model uses MobileNetV2, EfficientNetB0, and DenseNet121 as parallel feature extractors. Their outputs are projected into a common representation space and processed as expert descriptors by a Transformer encoder. The final prediction therefore reflects interactions among the three representations rather than a simple average of independently generated class scores. On the reported test split, MEAFNet achieved an accuracy of 99.78%, while Leaf Blight obtained an F1-score of 1.000. The ablation results also showed that the complete architecture outperformed each individual backbone under the same experimental setting. This finding suggests that the three CNNs provide partly complementary information for distinguishing grape leaf diseases. Nevertheless, the perfect result for Leaf Blight applies only to the available test samples. It should not be interpreted as evidence that the model will classify every field image correctly. The improvement also involves additional computational cost because all three CNN branches must process each input image. Classification accuracy alone is therefore insufficient for assessing practical deployment. Inference latency, memory consumption, throughput, and energy use should be measured on the intended hardware. Moreover, the benchmark dataset cannot fully represent vineyard conditions, where illumination, background vegetation, occlusion, camera distance, disease stage, and overlapping symptoms may vary considerably. Future work will examine a lightweight MEAFNet configuration for real-time edge applications. Pruning, quantization, knowledge distillation, and selective expert activation may reduce computational requirements while retaining useful diagnostic performance. Self-supervised learning will also be considered to exploit unlabeled vineyard images and reduce dependence on extensive manual annotation. Most importantly, the model should be evaluated on independently collected data and carefully constructed splits that prevent related images from appearing in both training and testing sets. Such validation is necessary before MEAFNet can be regarded as a reliable component of precision viticulture systems.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data supporting the results of this research are openly available in the Grape Disease Dataset at https://www.kaggle.com/datasets/pushpalama/grape-disease.

Author Contribution Statement

Hoang-Tu Vo: Conceptualization, Methodology, Writing – original draft, Supervision, Project administration. **Nhon Nguyen Thien:** Writing – review & editing, Visualization. **Kheo Chau Mui:** Software, Validation, Investigation. **Phuc Pham Tien:** Validation, Investigation, Writing – review & editing. **Huan Lam Le:** Formal analysis, Resources, Data curation.

References

[1] Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., . . . , & MacCarthy, D. S. (2023). Climate change impacts on crop yields. *Nature Reviews Earth & Environment*, *4*(12), 831–846. https://doi.org/10.1038/s43017-023-00491-0

[2] Fuglie, K., Morgan, S., & Jelliffe, J. (2024). *Global changes in agricultural production, productivity, and resource use over six decades*. U.S. Department of Agriculture, Economic Research Service. Retrieved from: https://www.ers.usda.gov/amber-waves/2024/september/global-changes-in-agricultural-production-productivity-and-resource-use-over-six-decades

[3] Sarma, P. K., Alam, M. J., Begum, I. A., & McKenzie, A. M. (2024). The effect of total factor productivity on the food security and livelihood vulnerability of farm households in Bangladesh. *Frontiers in Sustainable Food Systems*, *8*, 1395897. https://doi.org/10.3389/fsufs.2024.1395897

[4] Devaux, A., Goffart, J. P., Kromann, P., Andrade-Piedra, J., Polar, V., & Hareau, G. (2021). The potato of the future: Opportunities and challenges in sustainable agri-food systems. *Potato Research*, *64*(4), 681–720. https://doi.org/10.1007/s11540-021-09501-4

[5] Ajibade, S., Simon, B., Gulyas, M., & Balint, C. (2023). Sustainable intensification of agriculture as a tool to promote food security: A bibliometric analysis. *Frontiers in Sustainable Food Systems*, *7*, 1101528. https://doi.org/10.3389/fsufs.2023.1101528

[6] Maia, M., Cavaco, A. R., Laureano, G., Cunha, J., Eiras-Dias, J., Matos, A. R., . . . , & Figueiredo, A. (2021). More than just wine: The nutritional benefits of grapevine leaves. *Foods*, *10*(10), 2251. https://doi.org/10.3390/foods10102251

[7] Singh, J., Rasane, P., Kaur, R., Kaur, H., Garg, R., Kaur, S., . . . , & Mlcek, J. (2023). Valorization of grape (Vitis vinifera) leaves for bioactive compounds: Novel green extraction technologies and food-pharma applications. *Frontiers in Chemistry*, *11*, 1290619. https://doi.org/10.3389/fchem.2023.1290619

[8] de Almeida Sousa Cruz, M. A., de Barros Elias, M., Calina, D., Sharifi-Rad, J., & Teodoro, A. J. (2024). Insights into grape-derived health benefits: A comprehensive overview. *Food Production, Processing and Nutrition*, *6*(1), 91. https://doi.org/10.1186/s43014-024-00267-z

[9] Kellner, N., Antal, E., Szabó, A., & Matolcsi, R. (2022). The effect of black rot on grape berry composition. *Acta Alimentaria*, *51*(1), 126–133. https://doi.org/10.1556/066.2021.00195

[10] Dewasme, C., Mary, S., Darrieutort, G., Roby, J.-P., & Gambetta, G. A. (2022). Long-term esca monitoring reveals disease impacts on fruit yield and wine quality. *Plant Disease*, *106*(12), 3076–3082. https://doi.org/10.1094/PDIS-11-21-2454-RE

[11] Ouadi, L., Bruez, E., Bastien, S., Vallance, J., Lecomte, P., Domec, J.-C., & Rey, P. (2019). Ecophysiological impacts of Esca, a devastating grapevine trunk disease, on Vitis vinifera L. *PLOS ONE*, *14*(9), e0222586. https://doi.org/10.1371/journal.pone.0222586

[12] Aboelenin, S., Elbasheer, F. A., Eltoukhy, M. M., El-Hady, W. M., & Hosny, K. M. (2025). A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. *Complex & Intelligent Systems*, *11*(2), 142. https://doi.org/10.1007/s40747-024-01764-x

[13] Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., & Singh, S. (2022). Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. *Archives of Computational Methods in Engineering*, *29*(1), 641–677. https://doi.org/10.1007/s11831-021-09588-5

[14] Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based plant diseases detection using deep learning. *Materials Today: Proceedings*, *80*, 3500–3506. https://doi.org/10.1016/j.matpr.2021.07.281

[15] Qadri, S. A. A., Huang, N.-F., Wani, T. M., & Bhat, S. A. (2025). Advances and challenges in computer vision for image-based plant disease detection: A comprehensive survey of machine and deep learning approaches. *IEEE Transactions on Automation Science and Engineering*, *22*, 2639–2670. https://doi.org/10.1109/TASE.2024.3382731

[16] Upadhyay, A., Chandel, N. S., Singh, K. P., Chakraborty, S. K., Nandede, B. M., Kumar, M., . . . , & Elbeltagi, A. (2025). Deep learning and computer vision in plant disease detection: A comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review*, *58*(3), 92. https://doi.org/10.1007/s10462-024-11100-x

[17] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

[18] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708. https://doi.org/10.1109/CVPR.2017.243

[19] Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning*, *97*, 6105–6114.

[20] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . , & Polosukhin, I. (2017). Attention is all you need. In *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010. https://dl.acm.org/doi/10.5555/3295222.3295349

[21] Pacal, I., Kunduracioglu, I., Alma, M. H., Deveci, M., Kadry, S., Nedoma, J., . . . , & Martinek, R. (2024). A systematic review of deep learning techniques for plant diseases. *Artificial Intelligence Review*, *57*(11), 304. https://doi.org/10.1007/s10462-024-10944-7

[22] Ma, J., Jiang, X., Fan, A., Jiang, J., & Yan, J. (2021). Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision*, *129*(1), 23–79. https://doi.org/10.1007/s11263-020-01359-2

[23] Ahmad, N., Asif, H. M. S., Saleem, G., Younus, M. U., Anwar, S., & Anjum, M. R. (2021). Leaf image-based plant disease identification using color and texture features. *Wireless Personal Communications*, *121*(2), 1139–1168. https://doi.org/10.1007/s11277-021-09054-2

[24] S, S. P., P, K., & T A, S. L. (2023). Projection of plant leaf disease using support vector machine algorithm. In *2023 International Conference on Recent Advances in Science and Engineering Technology*, 1–6. https://doi.org/10.1109/ICRASET59632.2023.10419981

[25] Yang, C. (2021). Plant leaf recognition by integrating shape and texture features. *Pattern Recognition*, *112*, 107809. https://doi.org/10.1016/j.patcog.2020.107809

[26] Bedi, P., Gole, P., & Agarwal, S. K. (2021). 18 Using deep learning for image-based plant disease detection. In J. M. Chatterjee, A. Kumar, P. S. Rathore, & V. Jain (Eds.), *Internet of Things and Machine Learning in Agriculture* (pp. 369–402). De Gruyter. https://doi.org/10.1515/9783110691276-018

[27] Narmadha, R. P., Sengottaiyan, N., & Kavitha, R. J. (2022). Deep transfer learning based rice plant disease detection model. *Intelligent Automation & Soft Computing*, *31*(2), 1257–1271. https://doi.org/10.32604/iasc.2022.020679

[28] Sujatha, R., Krishnan, S., Chatterjee, J. M., & Gandomi, A. H. (2025). Advancing plant leaf disease detection integrating machine learning and deep learning. *Scientific Reports*, *15*(1), 11552. https://doi.org/10.1038/s41598-024-72197-2

[29] Sajitha, P., Andrushia, A. D., Anand, N., & Naser, M. Z. (2024). A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. *Journal of Industrial Information Integration*, *38*, 100572. https://doi.org/10.1016/j.jii.2024.100572

[30] Elfatimi, E., Eryigit, R., & Elfatimi, L. (2022). Beans leaf diseases classification using mobilenet models. *IEEE Access*, *10*, 9471–9482. https://doi.org/10.1109/ACCESS.2022.3142817

[31] Ashwinkumar, S., Rajagopal, S., Manimaran, V., & Jegajothi, B. (2022). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. *Materials Today: Proceedings*, *51*, 480–487. https://doi.org/10.1016/j.matpr.2021.05.584

[32] Balavani, K., Sriram, D., Shankar, M. B., & Charan, D. S. (2023). An optimized plant disease classification system based on resnet-50 architecture and transfer learning. In *2023 4th International Conference for Emerging Technology*, 1–5. https://doi.org/10.1109/INCET57972.2023.10170368

[33] Usha Ruby, A., George Chellin Chandran, J., Chaithanya, B. N., Swasthika Jain, T. J., & Patil, R. (2024). Wheat leaf disease classification using modified ResNet50 convolutional neural network model. *Multimedia Tools and Applications*, *83*(23), 62875–62893. https://doi.org/10.1007/s11042-023-18049-z

[34] Nandhini, S., & Ashokkumar, K. (2022). An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based Henry gas solubility optimization algorithm. *Neural Computing and Applications*, *34*(7), 5513–5534. https://doi.org/10.1007/s00521-021-06714-z

[35] Alirezazadeh, P., Schirrmann, M., & Stolzenburg, F. (2023). Improving deep learning-based plant disease classification with attention mechanism. *Gesunde Pflanzen*, *75*(1), 49–59. https://doi.org/10.1007/s10343-022-00796-y

[36] Zhao, Y., Sun, C., Xu, X., & Chen, J. (2022). RIC-Net: A plant disease classification model based on the fusion of Inception and residual structure and embedded attention mechanism. *Computers and Electronics in Agriculture*, *193*, 106644. https://doi.org/10.1016/j.compag.2021.106644

[37] Shu, X., Wang, J., Zhang, A., Shi, J., & Wu, X. J. (2024). CSCA U-Net: A channel and space compound attention CNN for medical image segmentation. *Artificial Intelligence in Medicine*, *150*, 102800. https://doi.org/10.1016/j.artmed.2024.102800

[38] Liu, Y., Gao, G., & Zhang, Z. (2022). Crop disease recognition based on modified light-weight CNN with attention mechanism. *IEEE Access*, *10*, 112066–112075. https://doi.org/10.1109/ACCESS.2022.3216285

[39] Madaan, V., & Sharma, N. (2024). Prediction of grapevine leaf identity with DenseNet121. In *2024 2nd International Conference on Emerging Trends in Engineering and Medical Sciences*, 127–131. https://doi.org/10.1109/ICETEMS64039.2024.10964988

[40] Bakr, M., Abdel-Gaber, S., Nasr, M., & Hazman, M. (2022). DenseNet based model for plant diseases diagnosis. *European*

Journal of Electrical Engineering and Computer Science, *6*(5), 1–9. https://doi.org/10.24018/ejece.2022.6.5.458

[41] Kunduracioglu, I., & Pacal, I. (2024). Advancements in deep learning for accurate classification of grape leaves and diagnosis of grape diseases. *Journal of Plant Diseases and Protection*, *131*(3), 1061–1080. https://doi.org/10.1007/s41348-024-00896-z

[42] Mamun, S. B., Ahad, M. T., Morshed, M. M., Hossain, N., & Emon, Y. R. (2024). Scratch vision transformer model for diagnosis grape leaf disease. In M. S. Kaiser, R. Singh, A. Bandyopadhyay, M. Mahmud, & K. Ray (Eds.), *Proceedings of the fifth international conference on trends in computational and cognitive engineering* (pp. 101–118). Springer, https://doi.org/10.1007/978-981-97-1923-5_8

[43] Liu, L., Qiao, S., Chang, J., Ding, W., Xu, C., Gu, J., . . . , & Qiao, H. (2024). A multi-scale feature fusion neural network for multi-class disease classification on the maize leaf images. *Heliyon*, *10*(7), e28264. https://doi.org/10.1016/j.heliyon.2024.e28264

[44] Lin, J., Zhang, X., Qin, Y., Yang, S., Wen, X., Cernava, T., . . . , & Chen, X. (2024). Local and global feature-aware dual-branch networks for plant disease recognition. *Plant Phenomics*, *6*, 0208. https://doi.org/10.34133/plantphenomics.0208

[45] Wang, W., He, J., & Liu, H. (2024). EMOST: A dual-branch hybrid network for medical image fusion via efficient model module and sparse transformer. *Computers in Biology and Medicine*, *179*, 108771. https://doi.org/10.1016/j.compbiomed.2024.108771

[46] Javidan, S. M., Banakar, A., Vakilian, K. A., & Ampatzidis, Y. (2023). Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. *Smart Agricultural Technology*, *3*, 100081. https://doi.org/10.1016/j.atech.2022.100081

[47] Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A., & Alabdali, A. M. (2022). Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction. *Sensors*, *22*(2), 575. https://doi.org/10.3390/s22020575

[48] Koklu, M., Unlersen, M. F., Ozkan, I. A., Aslan, M. F., & Sabanci, K. (2022). A CNN-SVM study based on selected deep features for grapevine leaves classification. *Measurement*, *188*, 110425. https://doi.org/10.1016/j.measurement.2021.110425

[49] Xie, X., Ma, Y., Liu, B., He, J., Li, S., & Wang, H. (2020). A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. *Frontiers in Plant Science*, *11*, 751. https://doi.org/10.3389/fpls.2020.00751

[50] Tang, Z., Yang, J., Li, Z., & Qi, F. (2020). Grape disease image classification based on lightweight convolution neural networks and channelwise attention. *Computers and Electronics in Agriculture*, *178*, 105735. https://doi.org/10.1016/j.compag.2020.105735

[51] Adeel, A., Khan, M. A., Sharif, M., Azam, F., Shah, J. H., Umer, T., & Wan, S. (2019). Diagnosis and recognition of grape leaf diseases: An automated system based on a novel saliency approach and canonical correlation analysis based multiple features fusion. *Sustainable Computing: Informatics and Systems*, *24*, 100349. https://doi.org/10.1016/j.suscom.2019.08.002

[52] Lin, J., Chen, X., Pan, R., Cao, T., Cai, J., Chen, Y., . . . , & Zhang, X. (2022). Grapenet: A lightweight convolutional neural network model for identification of grape leaf diseases. *Agriculture*, *12*(6), 887. https://doi.org/10.3390/agriculture12060887

[53] Thet, K. Z., Htwe, K. K., & Thein, M. M. (2020). Grape leaf diseases classification using convolutional neural network. In *2020 International Conference on Advanced Information Technologies*, 147–152. https://doi.org/10.1109/ICAIT51105.2020.9261801

[54] Ji, M., Zhang, L., & Wu, Q. (2020). Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks. *Information Processing in Agriculture*, *7*(3), 418–426. https://doi.org/10.1016/j.inpa.2019.10.003

[55] Liu, B., Tan, C., Li, S., He, J., & Wang, H. (2020). A data augmentation method based on generative adversarial networks for grape leaf disease identification. *IEEE Access*, *8*, 102188–102198. https://doi.org/10.1109/ACCESS.2020.2998839

[56] Nuthalapati, S. B., Mathew, M. P., K M, A., & A, S. K. (2025). Swingnet: A hybrid swin transform—GoogleNet framework for real-time grape leaf disease classification. *Procedia Computer Science*, *258*, 1629–1639. https://doi.org/10.1016/j.procs.2025.04.394

[57] Karthik, R., Menaka, R., Ompirakash, S., Murugan, P. B., Meenakashi, M., Lingaswamy, S., & Won, D. (2024). Grapeleafnet: A dual-track feature fusion network with inception-resnet and shuffle-transformer for accurate grape leaf disease identification. *IEEE Access*, *12*, 19612–19624. https://doi.org/10.1109/ACCESS.2024.3361044

How to Cite: Vo, H.-T., Thien, N. N., Mui, K. C., Tien, P. P., & Le, H. L. (2026). MEAFNet: Multi-Expert Attention Fusion Network for Robust Grape Leaf Disease Classification. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62027257