Systems and Computing
Volume 2, Issue 1, 105-114, 2026.
https://doi.org/10.64409/sycom.v2.i1.35

Research Article | Open Access

Evaluation of SE, ECA, and CA Modules in ResNet-50 for the Classification of Grape Leaf Diseases

Khedidja Mekhilef, Mounir Hemam

The Faculty of Science and Technology, Université Abbes Laghrour Khenchela, Algeria

Corresponding author: Khedidja Mekhilef (khedidja.mekhilef@univ-khenchela.dz)

Manuscript Review Record:

Submitted:
November 13, 2025

Accepted:
January 10, 2026

Published:
January 15, 2026

Cite This:

Khedidja Mekhlief. "Evaluation of SE, ECA, and CA Modules in ResNet-50 for the Classification of Grape Leaf Diseases". *Systems and Computing*. Volume 2, Issue 1, 105-114, 2026. https://doi.org/10.64409/sycom.v2.i1.35

Abstract- Context: Plant diseases pose a major threat to agricultural production, as they hinder crop growth and significantly reduce yields. In this context, the automatic detection of diseases using artificial intelligence techniques represents an effective solution. **Objective:** This work presents a comparative study on the integration of attention mechanisms into the ResNet-50 architecture, applied to the classification of grape leaf diseases. **Method:** Three attention modules were considered: SE (Squeeze-and-Excitation), ECA (Efficient Channel Attention), and CA (Coordinate Attention), each integrated separately into the model to evaluate its impact. **Results:** The performance of the different configurations was assessed by considering the quality of the results, the model complexity, and the processing speed. Experimental results show that the ECA module achieved the best performance with a reduced number of parameters, followed by SE and then CA. **Conclusion:** This study highlights the contribution of attention mechanisms in improving the effectiveness of CNN architectures for plant disease detection.

Keywords- Classification, Attention mechanisms, Agriculture, Plant diseases.

Acknowledgement

The authors are grateful to the anonymous referees for their valuable and helpful comments.

1. Introduction

Plants are exposed to various threats, including pathogens such as fungi, bacteria, and nematodes, as well as environmental factors like unsuitable soil pH, extreme temperatures, or humidity variations. These factors can cause plant diseases, affecting their growth and directly impacting agricultural yields [1]. Early detection of these diseases remains a major challenge, as manual methods are often ineffective on large scale farms. A late identification of symptoms can reduce production and threaten food security [2]. To address this issue, automated solutions based on artificial intelligence enable the rapid identification of symptoms on leaves, reducing labor requirements and improving diagnostic accuracy [3] [4]. Convolutional neural networks play a key role in this field, delivering competitive results due to their ability to automatically extract discriminative features. Pre-trained models such as ResNet, VGG, or EfficientNet have proven highly effective in plant disease classification, sometimes outperforming traditional methods like SVM [5] and KNN [6]. However, convolutional neural networks have a limitation in capturing long-range dependencies, which are essential for extracting global features crucial for accurate classification.

To address this issue, Hu et al. [7] proposed the Squeeze and Excitation attention mechanism, considered the first attention module introduced in computer vision. This mechanism enhances CNNs by emphasizing the most informative channels. Subsequently, Wang et al. proposed the Efficient Channel Attention [8] mechanism, arguing that the use of MLPs in SE increases model complexity. They introduced a lighter alternative based on 1D convolution to capture channel interdependencies while reducing computational cost without compromising performance. Among other notable mechanisms, Coordinate Attention [13] stands out by encoding both positional and channel information, thereby improving object localization while maintaining low complexity. The success of attention mechanisms when integrated with convolutional neural network architectures has led to their widespread adoption across various domains, including plant disease detection. These mechanisms have been incorporated into several recent architectures to enhance feature representation and improve overall performance.

In this work, we conduct a comparative study of several recent attention mechanisms, evaluating them in terms of performance, computational complexity, and inference speed. The objective is to determine the most suitable mechanism, particularly for real-time applications such as those used in smart agriculture.

2. Related work

Several studies have explored plant disease classification using machine learning and deep learning. For example, [10] used SVM to identify sugarcane diseases with 95% accuracy, while [11] employed ANN and KNN for cassava. Other works have combined K-means or GMS matrices with various classifiers. In deep learning, [12] proposed DenseNet77, [13] fused MobileNetV2 and NASNetMobile, and [14] investigated DCNs to enhance leaf disease detection.[15] proposed MFANet, a multi-scale feature attention network designed to enhance feature representation by integrating both spatial and channel attention mechanisms. The model was evaluated on the Cassava leaf disease dataset and achieved an accuracy of approximately 99%, demonstrating its effectiveness in capturing discriminative features across different scales. However, MFANet relies on a dedicated multi-scale attention architecture, which increases model complexity. In [16], the authors proposed HPDC-Net, a hybrid architecture composed of three Depthwise Separable Convolution blocks to maintain a lightweight design, a Dual-Path Adaptive Pooling block, and a Channel-Wise Attention Refinement Block (CARB). The propose model was trained and evaluated on three datasets for potato and tomato leaf disease classification, achieving accuracy scores exceeding 99% across all datasets while maintaining a minimal number of parameters. Although HPDC-Net demonstrates excellent performance with a compact architecture, it focuses on a task-specific network design. In contrast, our work explores the integration of different attention mechanisms into a widely used backbone (ResNet-50), enabling a fair comparison of their effectiveness and generalization across plant disease classification tasks.

3. Method

The Convolutional Neural Network (CNN) is a specialized architecture specifically designed for image processing. A CNN typically comprises a series of convolutional layers, which apply learnable filters to extract hierarchical features from input images, interleaved with pooling layers that reduce spatial dimensions and computational complexity, and non-linear activation functions that introduce non-linearity. These feature maps are ultimately processed by fully connected layers, which integrate the learned representations to perform high-level reasoning and classification. Since their emergence and the remarkable results achieved in computer vision, numerous architectural extensions of CNNs have been proposed, such as AlexNet [17], VGG16 [18], ResNet [19], and other more recent variants. To enhance performance, several strategies have been explored, including increasing the network's depth by stacking more layers, as well as leveraging multi-scale information through the concatenation of convolutions with different kernel sizes in parallel. More recently, the introduction of attention mechanisms has proven to be particularly effective, enabling a significant improvement in model performance. In order to demonstrate the importance of attention mechanisms in CNN architecture, a comparative study was conducted by integrating SE,CA, and ECA modules into the ResNet-50 architecture as shown in Figure 1. The latter represents an advanced version of the ResNet model, consisting of 50 layers, including convolutional and fully connected layers.

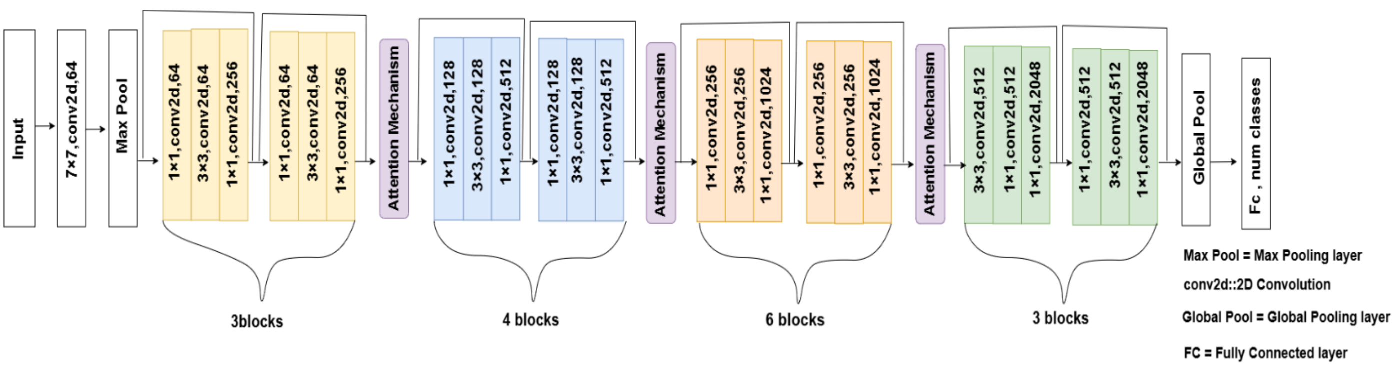

Figure 1. Proposed ResNet-50-Based Architecture with Integrated Attention Mechanisms.

3.1 Squeeze-and-Excitation module

Since the convolutional layer operates on a small local region of the image and each filter is applied independently, the interactions between channels remain implicit. For this reason, Hu et al. [7] suggested that learning could be improved by explicitly modeling inter-channel dependencies. They therefore proposed giving the network access to global information through two operations: Squeeze and Excitation. The Squeeze step is based on the idea of compressing the spatial information of each channel into a representative statistic. This is done using Global Average Pooling, as shown in the figure, which generates one average value per channel. The excitation step learns channel-wise importance weights through two fully connected layers separated by a ReLU activation, followed by a Sigmoid function to ensure normalized values between 0 and 1. These weights are then used to recalibrate the original feature maps via channel-wise multiplication, allowing the network to emphasize informative channels while suppressing less relevant ones. This mechanism introduces only a small computational overhead while significantly improving the representational power of the network. All the components described above are illustrated in Figure 2.

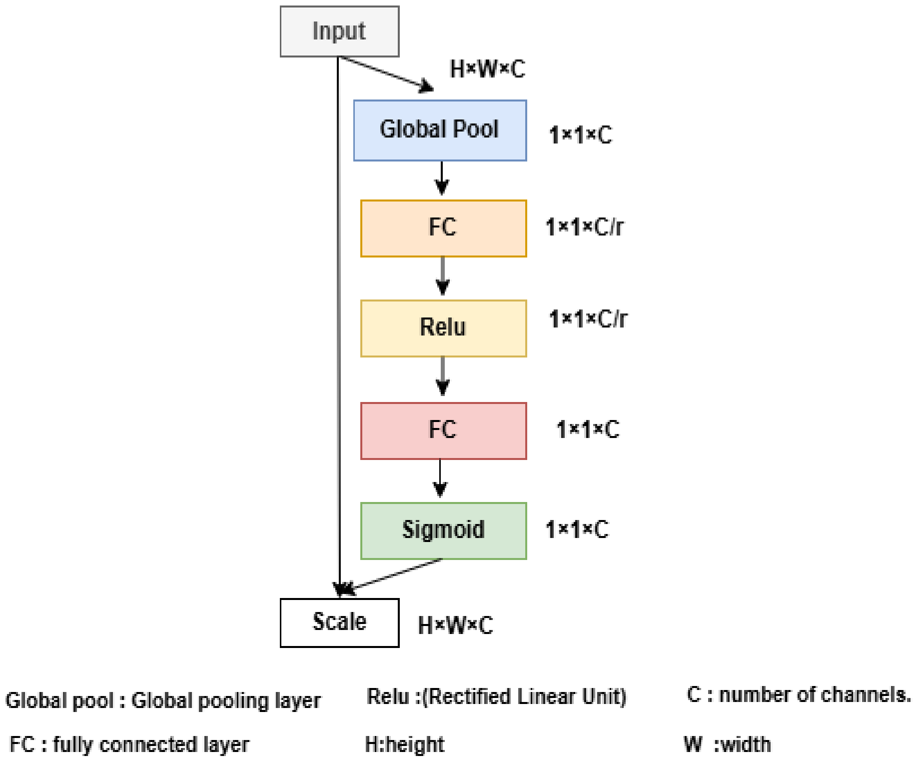

Figure 2. Squeeze-and-Excitation module.

3.2 Efficient Channel Attention

A drawback of the SE block is addressed by the ECA module: the channel-to-weight correlation is indirect because the fully connected layers in SE combine data from all channels. In order to get around this, ECA substitutes a 1D convolution with kernel size k, which records local interactions between nearby channels, for the two fully connected layers. Each channel is enhanced based on its local context by applying the derived channel-wise attention weights to the input feature maps through element-wise multiplication. This method maintains very low computational cost while enabling more accurate inter-channel attention. The ECA module is illustrated in Figure 3.

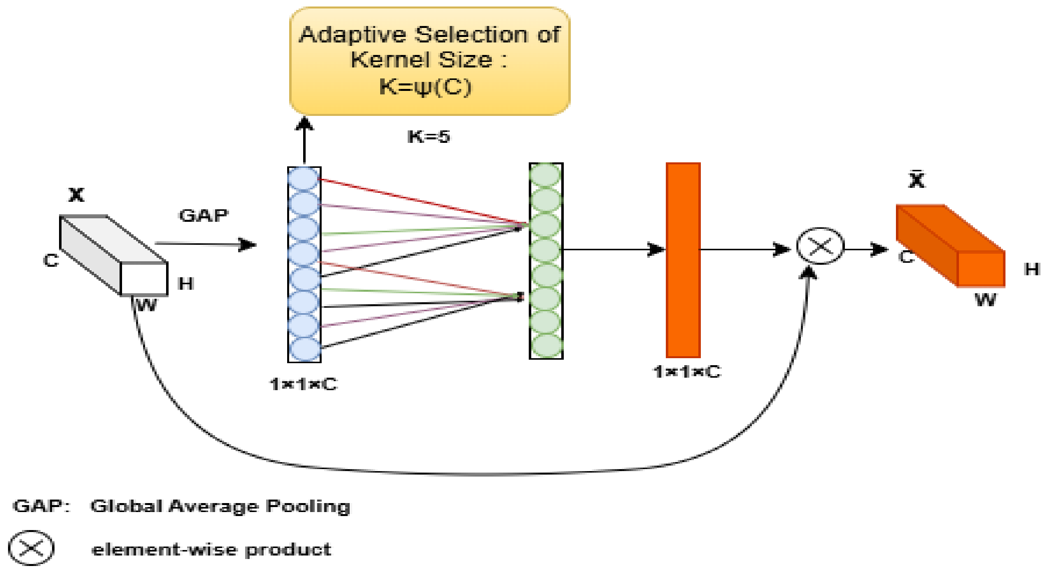

Figure 3. Efficient Channel Attention [8].

3.3 Coordinate Attention

In order to incorporate spatial information into channel attention, Coordinate Attention was proposed. Coordinate Attention uses two one-dimensional Global Average Pooling operations along the horizontal and vertical directions, in contrast to SE attention, which uses Global Average Pooling and compresses each channel into a single scalar, losing positional information. Long-range dependencies can be encoded in every spatial direction while maintaining accurate positioning information thanks to this approach. In order to describe inter-channel interactions, the obtained features are concatenated and run through convolutional layers. A sigmoid activation is then used to create direction-aware attention maps. The network may highlight discriminative regions while preserving spatial structure by applying these maps to the input feature maps through element-wise multiplication. This is especially useful for identifying elongated or organized patterns. The Coordinate Attention module is illustrated in Figure 4.

4. Results

4.1 Dataset

The experiments were conducted using the Grape Disease dataset, which is publicly available on Kaggle [20]. The dataset consists of four classes: Black Rot, ESCA, Healthy, and Leaf Blight. The dataset is already pre-split on Kaggle into training and testing subsets; therefore, this split was not performed by us. The training set contains 7,222 images, distributed as follows: Black Rot (1,888 images), ESCA (1,920 images), Healthy (1,692 images), and Leaf Blight (1,722 images). The test set consists of 1,805 images, including Black Rot (472 images), ESCA (480 images), Healthy (423 images), and Leaf Blight (430 images).

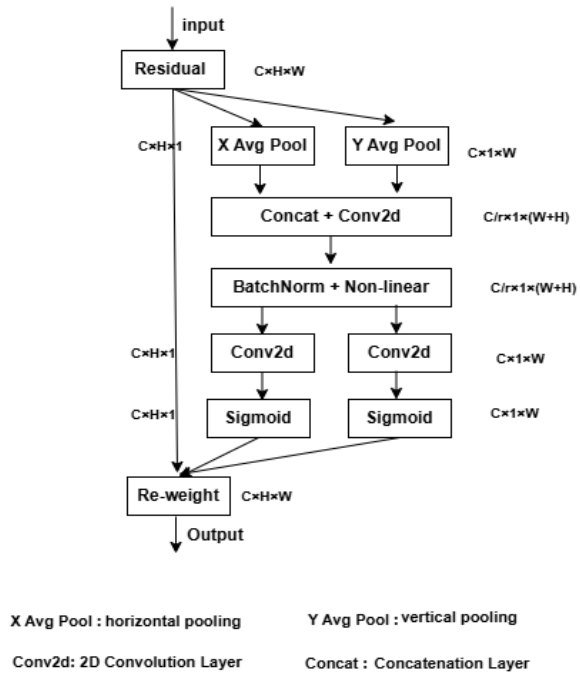

Figure 4. Coordinate Attention [9].

4.2 Optimization Parameters

For model development, the training set was further split into 80% for training and 20% for validation. All images were resized to a resolution of 224 × 224 pixels before training. A batch size of 32 was used, with the Adam optimizer and a learning rate of 0.001. The Cross-Entropy Loss function was employed to optimize model performance. All models were trained for 50 epochs. Since the main objective of this work is to fairly compare model performance, no data augmentation techniques were applied. The experimental results show no overfitting, as indicated by the consistent performance between the training and validation phases. All experiments were conducted on the Kaggle platform, using an NVIDIA Tesla P100 GPU to accelerate the training process.

4.3 Quantitative Results

To evaluate the performance of the Grape plant disease classification models, we used the following metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

5. Discussion

Table 1 presents the results obtained by ResNet-50 architecture combined with different attention mechanisms: Squeeze-and-Excitation (SE), Coordinate Attention (CA), and Efficient Channel Attention (ECA). The results show that ResNet-50-ECA achieves the best performance compared to the other approaches and the baseline ResNet-50, outperforming them in terms of accuracy, precision, recall, and F1-score. This demonstrates the effectiveness of replacing MLP layers with convolutions within the attention module. ResNet-50-SE ranks second, also demonstrating its capability to improve performance. It is followed by ResNet-50-CA in third place, which highlights its ability to incorporate spatial information through the combined use of horizontal and vertical global pooling for each channel. Finally, the baseline ResNet-50 confirms the added value introduced by integrating attention mechanisms. Furthermore, a comparison with other architectures, such as Vision Transformer (ViT), VGG19, and ResNet101 which represents a deeper version of ResNet-50 with a larger number of layers shows that the ResNet-50 model enhanced with the ECA module outperforms ResNet101 across all evaluation metrics, namely accuracy, recall, precision, and F1-score. These results highlight the effectiveness of attention mechanisms and confirm that their integration represents a promising approach for improving plant disease classification in future research.

As shown in Table 2 it illustrates the number of parameters for each approach as well as the number of images processed per second (IPS). The results indicate that architecture using Efficient Channel Attention (ECA) has the smallest number of parameters, confirming the effectiveness of using 1D convolutions to achieve good performance with reduced complexity, especially compared to SE. When comparing Coordinate Attention to SE, a reduction in the number of parameters is also observed, which may contribute to improved execution speed. In terms of images processed per second, ECA also achieves the highest score, followed by SE and then Coordinate Attention. Finally, the baseline ResNet-50 model shows the lowest performance in terms of both speedand model compactness.

Table 1 Comparison of ResNet-50 with different attention mechanisms and other architectures based on standard evaluation metrics.

	Accuracy	Recall	Precision	F1-Score
ResNet-50-CA	95.9	95.90	96.03	95.86
ResNet-50-SE	96.84	96.84	96.98	96.85
ResNet-50-ECA	**97.29**	**97.29**	**97.32**	**97.29**
ResNet-50	95.46	95.46	95.93	95.50
Vit	88.31	88.31	88.49	88.12
Vgg19	95.90	95.90	95.95	95.92
ResNet101	96.90	97.09	96.90	96.89

Table 2 Comparative table of ResNet-50 models with different attention mechanisms, showing the number of parameters and inference speed.

	parameters	IPS
ResNet-50-CA	23945842	57.52
ResNet-50-SE	24245386	64.28
ResNet-50-ECA	*23555027*	*69.6*
ResNet-50	23555018	54.85

The performance of ResNet-50 and its attention-based variations under various random seeds is shown in Table 3. The average accuracy of the baseline ResNet-50 is consistently improved by all attention strategies. ResNet-50-ECA outperforms the others in terms of both performance and resistance to random initialization, with the highest mean accuracy (98.15%) with the lowest standard deviation (0.15). While ResNet-50-SE exhibits comparatively greater sensitivity to seed selection, ResNet-50-CA likewise exhibits competitive accuracy (98.06%) with low variability. Overall, these findings support the efficacy of attention processes and show that ECA is the most reliable and successful module.

Table 3 Precision (%) of ResNet-50 and its variants with SE, ECA, and CA attention mechanisms, evaluated over three random seeds (Mean ± Std).

	Seed1	**Seed2**	**Seed3**	**Meen**	**Std**
ResNet-50-CA	97.92	97.92	98.35	98.06	0.2
ResNet-50-SE	97.92	97.20	96.51	97.21	0.58
ResNet-50-ECA	*98.28*	*98.23*	*97.95*	*98.15*	*0.15*
ResNet-50	96.17	97.91	96.89	96.99	0.71

Figure 5 presents the confusion matrices illustrating the classification performance of ResNet50 and its attention-enhanced variants (ResNet50+CA, ResNet50+SE, and ResNet50+ECA) on the test dataset. As shown in the sub-figures, the baseline ResNet50 model produces the highest number of misclassifications (80), followed by ResNet50+CA (74), ResNet50+SE (56), and ResNet50+ECA (51). These results demonstrate that the ECA module provides the best improvement in prediction accuracy compared to the other attention mechanisms. Overall, attention modules prove their effectiveness in enhancing the performance of the baseline architecture.

5.1 Limitations

A limitation of the current study is the placement of the attention modules within ResNet-50. While the integration of SE, ECA, and CA modules improved classification performance, the gains were sometimes modest. Placing attention modules at different layers or in multiple positions within the network may yield further performance improvements. Future work could explore systematic strategies for optimal module placement to fully exploit the benefits of attention mechanisms.

6. Conclusions

In summary, this study investigated the impact of integrating different attention mechanisms into the ResNet-50 architecture for the classification of grape diseases. By comparing the SE, ECA, and CA modules individually integrated into the model, we assessed their effect in terms of classification performance, computational complexity, and inference speed. Furthermore, we evaluated the robustness of these models by performing multiple runs with different random seeds. The results demonstrated that the ECA module consistently achieves the highest performance across seeds, followed closely by CA, while SE shows relatively higher variability. Overall, these findings confirm the relevance of attention mechanisms, particularly ECA, in enhancing both the accuracy and stability of convolutional neural networks for plant disease detection tasks.

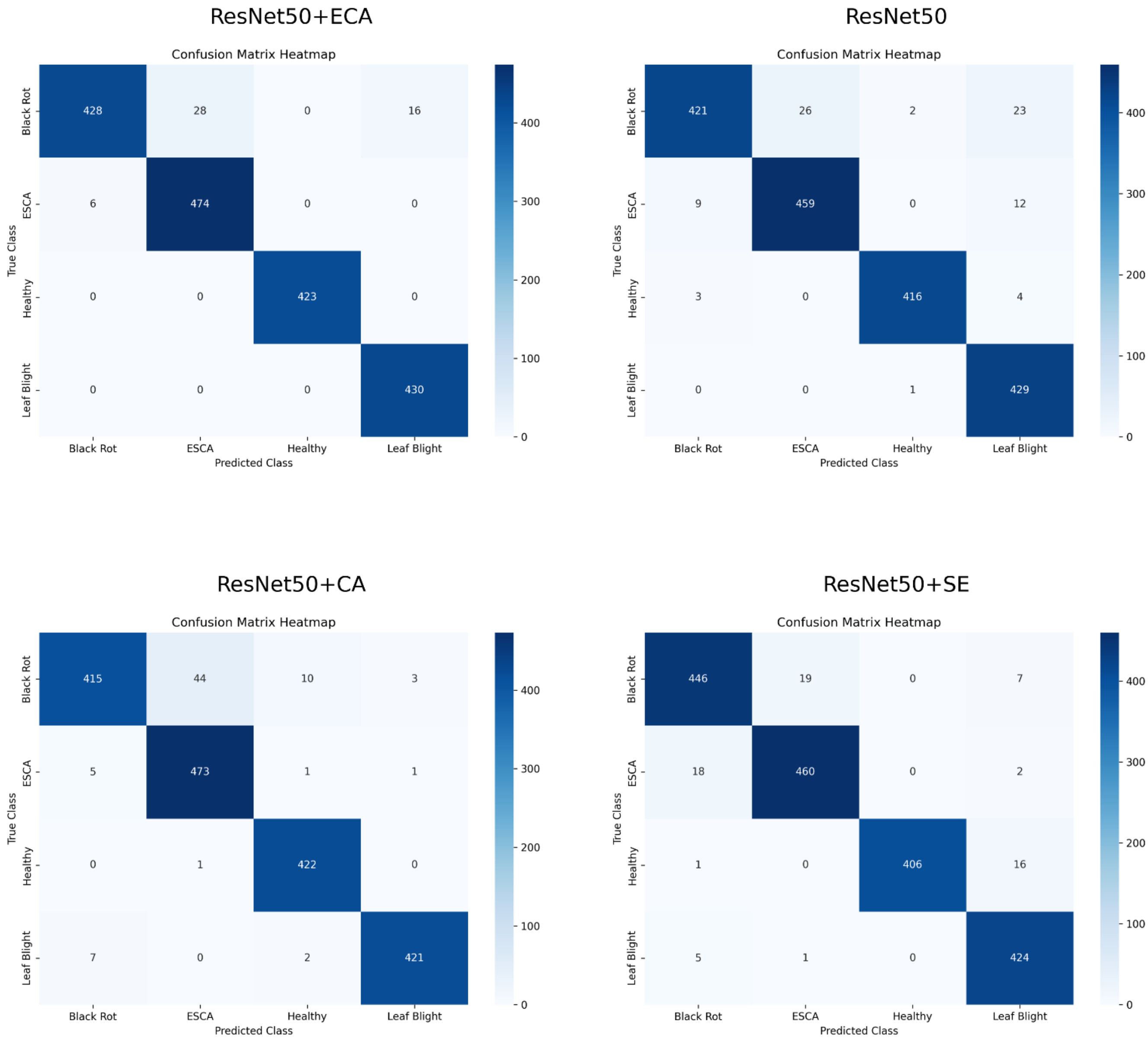

Figure 5 Confusion matrices illustrate the classification performance of ResNet-50 and its attention-enhanced variants, including ResNet-50+CA, ResNet-50+SE, and ResNet-50+ECA, on the test dataset.

References

[1] V. G. Krishnan, J. R. V. P. Deepa, P. V. Rao, V. Divya, and S. Kaviarasan, "An Automated Segmentation and Classification Model for Banana Leaf Disease Detection", *Journal of Applied Biology and Biotechnology*, vol. 10, no. 1, 2022, pp. 213–220. https://doi.org/10.7324/JABB.2021.100126

[2] W. B. Demilie, "Plant Disease Detection and Classification Techniques: A Comparative Study of the Performances," Journal of Big Data, vol. 11, no. 1, 2024, p. 5. https://doi.org/10.1186/s40537-023-00863-9

[3] M. Sardogan, A. Tuncer, and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", in 2018 3rd International Conference on Computer Science and Engineering (UBMK), IEEE, 2018, pp. 382–385. https://doi.org/10.1109/UBMK.2018.8566635

[4] E. Dhinesh and A. Jagan, "Detection of Leaf Disease Using Principal Component Analysis and Linear Support Vector n Machine," in 2019 11th International Conference on Advanced Computing (ICoAC), IEEE, 2019, pp. 350–355. https://dx.doi.org/10.1109/ICoAC48765.2019.246866

[5] K. Hilagavathi, K. Kavitha, R. Dhivya Praba, S. V. Arina, and R. C. Sahana, "Detection of Diseases in Sugarcane Using Image Processing Techniques", *Bioscience Biotechnology Research Communications*, Special Issue 11, 2020, pp. 109–115. https://doi.org/10.21786/bbrc/13.11/24

[6] O. Attallah, "Tomato Leaf Disease Classification via Compact Convolutional Neural Networks with Transfer Learning and Feature Selection", *Horticulturae*, vol. 9, no. 2, 2023, p. 149. https://doi.org/10.3390/horticulturae9020149

[7] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141. https://doi.org/10.48550/arXiv.1709.01507

[8] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 11534–11542. https://doi.org/10.1109/CVPR42600.2020.01155

[9] Q. Hou, D. Zhou, and J. Feng, "Coordinate Attention for Efficient Mobile Network Design," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 13713–13722. https://doi.org/10.48550/arXiv.2103.02907

[10] K. Thilagavathi, K. Kavitha, R. Dhivya Praba, S. V. Arina, and R. C. Sahana, "Detection of Diseases in Sugarcane Using Image Processing Techniques", *Bioscience Biotechnology Research Communications*, Special Issue 11, 2020, pp. 109–115. http://dx.doi.org/10.21786/bbrc/13.11/24

[11] R. Anand, S. Veni, and J. Aravinth, "An Application of Image Processing Techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method" in 2016 International Conference on Recent Trends in Information Technology (ICRTIT), IEEE, 2016, pp. 1–6. https://doi.org/10.1109/ICRTIT.2016.7569531

[12] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, "A Novel Deep Learning Method for Detection and Classification of Plant Diseases", *Complex & Intelligent Systems*, 2022, pp. 1–18. https://doi.org/10.1007/s40747-021-00536-1

[13] 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 Processing*, vol. 16, no. 3, 2022, pp. 913–925. https://doi.org/10.1049/ipr2.12397

[14] P. Nayar, S. Chhibber, and A. K. Dubey, "An Efficient Algorithm for Plant Disease Detection Using Deep Convolutional Networks" in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, 2022, pp. 156–160. https://doi.org/10.1109/CICN56167.2022.10008235

[15] S. Sharma and M. Vardhan, "MFANet: Multi-level Features and Enhancement Operations with the Attention Network for Plant Disease Classification", *Arab Journal of Science and Engineering*, 2025. https://doi.org/101007/s13369-025-10833-3

[16] M. Asghar, Z. F. Khan, M. Ramzan, M. A. Khan, J. Baili, Y. Zhang, Y. Nam, and Y. C. Nam, "A lightweight hybrid model for scalable and robust plant leaf disease classification", *Scientific Reports*, vol. 15, no. 1, 2025, p. 32353. https://doi.org/10.1038/s41-025-08788-4

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NeurIPS), vol. 25, 2012. https://doi.org/10.1145/3065386

[18] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition" arXiv preprint arXiv:1409.1556, 2014. https://doi.org/10.48550/arXiv.1409.1556

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

[20] Grapevine Disease Dataset https://www.kaggle.com/datasets/rm1000/grape-disease-dataset-original