Volume 11 Issue 1 Year 2026 | Page 193-204 ISSN: 2527-9866
Received: ***25-12-2025*** | *Revised:* ***12-01-2026*** | *Accepted:* ***27-01-2026***

Enhancing Generalization of Tomato Leaf Disease Classification via TDR-Model and Field-Conditioned Data Augmentation

Fernando Feliansyah[1], Ery Hartati[2]
[1,2] Multi Data Palembang University, Palembang, South Sumatra, Indonesia, 30113
e-mail: fernandofeliansyah_2226250051@mhs.mdp.ac.id[1], ery_hartati@mdp.ac.id[2]
*Correspondence: ery_hartati@mdp.ac.id

Abstract: *Tomato leaf diseases significantly affect agricultural productivity, particularly when detection systems are deployed under real-field conditions characterized by illumination variation, background clutter, and image noise. Although deep learning-based models have achieved high accuracy on laboratory datasets such as PlantVillage, their generalization performance often degrades when applied to real-world environments. This study proposes a lightweight CNN-based tomato leaf disease recognition model, referred to as the TDR-Model, combined with field-conditioned data augmentation strategies. The proposed model integrates MobileNetV3 with Convolutional Block Attention Module (CBAM) and Omni-Dimensional Dynamic Convolution (ODC) to enhance feature representation while maintaining computational efficiency. Field-conditioned augmentation using the Albumentations library to simulate real-world visual variations during training. The model is evaluated on Real-World tomato set consisting of 10 classes and 885 leaf images. Experimental results show that the proposed model achieves an overall test accuracy of 82.94%, with precision, recall, and F1-score of 85.06%, 83.04%, and 83.03%, respectively. Furthermore, the model requires only 3.47 million parameters, 0.23 GFLOPs, and an average inference time of 5.15 ms, making it suitable for real-time and resource-constrained agricultural applications. These results indicate that the proposed approach effectively balances accuracy and efficiency for practical tomato leaf disease detection.*

Keywords: *Data Augmentation, Image Classification MobileNetV3, PlantVillage, Tomato Disease* Recognition

1. Introduction

Agriculture plays a crucial role in sustaining food security in Indonesia, where tomato (Solanum lycopersicum L.) is one of the most economically important horticultural crops [1]. According to Statistics Indonesia, national tomato production decreased from approximately 1.16 million tons in 2022 to 1.14 million tons in 2023, indicating potential productivity challenges despite consistently high demand [2]. Therefore, early and accurate detection of tomato leaf diseases is essential to support sustainable agricultural production [3]. Currently, various computer vision techniques based on deep learning, particularly Convolutional Neural Networks (CNNs), have been widely adopted for automated plant disease classification [4]. In particular, lightweight CNN architectures such as MobileNet and EfficientNet have been proposed as alternatives to balance accuracy and computational efficiency [5]. MobileNetV3, in particular, incorporates depthwise separable convolution and squeeze-and-excitation mechanisms to achieve favorable performance–efficiency trade-offs [6]. Despite these advances, models trained on PlantVillage often exhibit poor generalization when applied to real-field conditions due to the dataset's controlled acquisition environment, characterized by uniform lighting, homogeneous backgrounds, and minimal noise [7]. In contrast, real-field agricultural images commonly suffer from illumination variation, shadows, motion blur, atmospheric effects, and sensor noise, leading to a significant generalization gap between laboratory-trained models and real-world deployment [8].

Several approaches have been proposed to improve model robustness, including architectural enhancements [9], architectural modification [10] and image preprocessing techniques such as dehazing or noise removal [11], [12]. However, these methods often emphasize model complexity while overlooking the importance of data distribution diversity during training. Furthermore, collecting large-scale real-field datasets is costly and labor-intensive, making it impractical for many agricultural research scenarios [13]. This study proposes a field-conditioned data augmentation strategy using the Albumentations library to simulate variations commonly observed in field environments, such as illumination changes, shadows, blur, and noise [14]. The augmented dataset is used to evaluate the TDR-Model, which integrates MobileNetV3 with CBAM and ODC to balance computational efficiency and adaptive feature representation [12].

While field-conditioned data augmentation increases input-level diversity, it does not explicitly guide the model to focus on disease-relevant regions or adapt feature extraction under complex backgrounds. Conversely, architectural modules such as CBAM and ODC enhance feature representation but remain sensitive to limited data diversity when trained solely on laboratory images. This study hypothesizes that combining simulated field-conditioned augmentation with adaptive architectural components yields complementary benefits for improving generalization under domain shift. The primary contributions of this study lies in systematically analyzing the extent to which simulated field-conditioned augmentation improves model generalization, while explicitly acknowledging the limitations of relying on synthetically augmented laboratory data.

2. Literature Review

Recent studies have demonstrated the effectiveness of deep learning approaches, particularly Convolutional Neural Networks (CNNs), in plant disease classification tasks. Early works using large-scale CNN architectures such as AlexNet, VGG, and ResNet achieved high accuracy on benchmark datasets like PlantVillage, benefiting from the dataset's controlled acquisition conditions and clean visual characteristics [4], [7], [15]. However, these architectures often involve high computational costs, limiting their deployment on mobile and edge-based agricultural systems. In study of [6], Howard *et al.* introducing MobileNetV3 which combines depthwise separable convolution with squeeze-and-excitation mechanisms to achieve a favorable balance between accuracy and computational efficiency. Several studies reported competitive performance of MobileNet-based models for plant disease classification while significantly reducing model size and inference latency [16], [17], [18]. Nevertheless, despite architectural efficiency, models trained on PlantVillage frequently exhibit degraded performance when exposed to real-field images due to domain shift caused by illumination variation, background clutter, and sensor noise [7].

Beyond backbone optimization, recent research has focused on enhancing feature representation through architectural augmentation [19]. Attention-based modules such as the CBAM have been shown to improve model performance by emphasizing informative spatial regions and discriminative channels [20]. Meanwhile, ODC introduces adaptive kernel weighting across multiple dimensions, enabling convolutional layers to dynamically adjust to varying input characteristics [21]. Another research direction aims to improve generalization through image preprocessing and data-level strategies. Techniques such as image dehazing, denoising, and enhancement have been proposed to mitigate visual degradation in plant disease images [12]. However, these methods typically address specific distortions and may not sufficiently capture the diversity of real-field conditions. Collecting large-scale field datasets remains a major challenge due to high labeling costs and environmental variability [13]. Consequently, data augmentation techniques have gained attention as a cost-effective alternative to enrich training distributions. Study of [14], presenting The Albumentations library that provides flexible and efficient augmentation pipelines that simulate illumination changes, shadows, blur, and noise commonly observed in outdoor agricultural environment.

Despite these advancements, existing studies often evaluate architectural improvements or augmentation strategies in isolation. Limited attention has been given to analyzing how field-conditioned augmentation interacts with adaptive architectural components such as CBAM and ODC, particularly within lightweight frameworks. This study addresses this gap by systematically evaluating the combined effect of simulated field-conditioned augmentation and the TDR-Model architecture on tomato leaf disease classification, with a focus on robustness under domain shift and computational efficiency.

3. Methods

A. Data Description

This study utilizes the PlantVillage dataset, a publicly available benchmark dataset widely used for plant disease classification. The dataset was originally introduced by Mohanty *et al* [22], and contains over 54.000 leaf images from 14 plant species and 38 disease types which captured under controlled laboratory conditions with uniform illumination and homogeneous backgrounds. In this work, only the tomato subset is used, consisting of 10 classes and 16.012 of total leaf images, 10 classes are separated into nine disease categories (e.g., Early Blight, Late Blight, Target Spot, Mosaic Virus, Bacterial Spot, Septoria Leaf Spot, Spider Mites Two Spotted, Leaf Mold, Yellow Leaf Curl Virus) and one healthy class which can be seen at Figure 1.

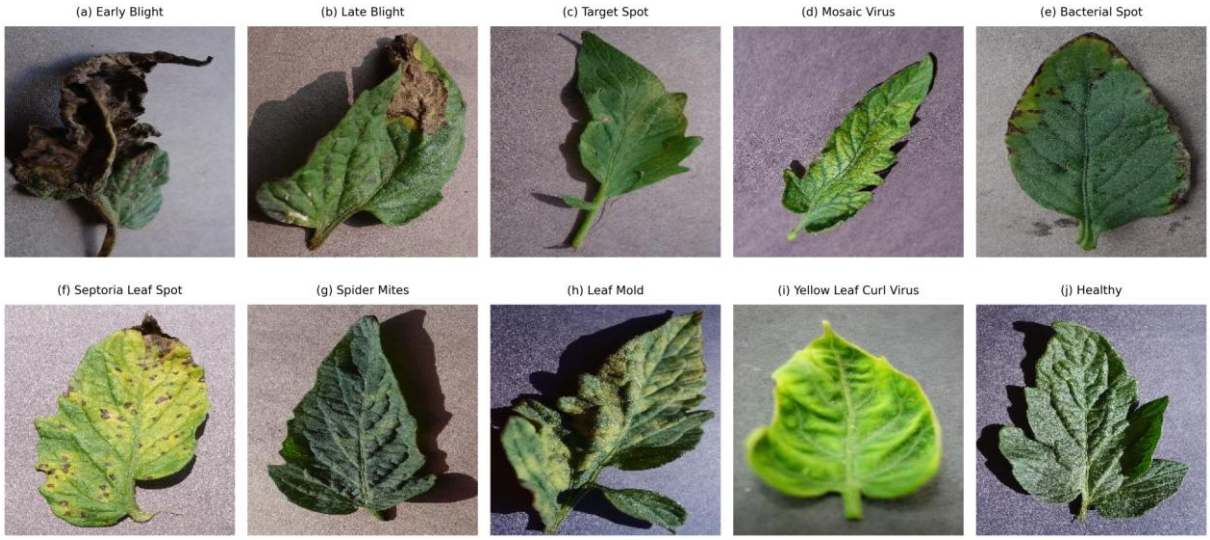

Figure 1. Image sample of tomato subset from PlantVillage

In addition to the PlantVillage dataset, a real-world tomato leaf image dataset was utilized exclusively for robustness evaluation. This dataset consist of 885 leaf images which curated and sourced from public repositories on PlantDoc Github [23] and Roboflow [24],[25], specifically focusing on samples with high environmental variability. The PlantDoc dataset does not include samples for the Target Spot and Two-Spotted Spider Mite classes, covering only 7 disease categories (e.g., Early Blight (66), Late Blight (117), Mosaic Virus (51), Bacterial Spot (85), Septoria Leaf Spot (76), Leaf Mold (71), Yellow Leaf Curl Virus(68)) and one healthy class (54) with per-sample class counts shown in parentheses. To ensure complete class coverage and maintain consistency with the PlantVillage label taxonomy, additional samples for the missing Target Spot (152) and Two-Spotted Spider Mite (145) classes were collected separately from public Roboflow repositories. This dataset consists of tomato leaf images captured under natural field conditions, exhibiting variations in illumination, background complexity, leaf orientation, and image quality which can be seen at Figure 2.

Figure 2. Sample of real-world tomato leaf images

Unlike PlantVillage images, these samples represent realistic deployment scenarios where environmental factors have non-homogeneous background. The real-world dataset was used only during the testing phase and was not included in the training and validation processes. By restricting the real-world dataset to the testing phase only, this study aims to verify the practical reliability of the model. This setup tests whether the model has learned discriminative disease features rather than simply memorizing the controlled laboratory environment of the PlantVillage dataset. All images are resized to 224 × 224 pixels to match the input requirements of MobileNetV3-based architectures. Pixel values are normalized to the range [0, 1]. The dataset from PlantVillage is split into 80% training, 10% validation, and 10% testing, ensuring class consistency across splits. To mitigate domain shift between laboratory images and real-field conditions, the training set is augmented using a field-conditioned data augmentation strategy, as described in Section 3.2.

B. Field-Conditioned Data Augmentation

Since collecting large-scale real-field agricultural datasets is costly and labor-intensive, this study adopts a simulated field-conditioned augmentation approach using the Albumentations library. The augmentation pipeline is designed to mimic common visual degradations observed in outdoor agricultural environments, including:

Table 1. Augmentation Parameter

	Augmentation Type	Parameters	Probability
1	RandomBrightnessContrass	brightness_limit=0.15, contrast_limit=0.1,	0.5
2	RandomShadow	shadow_roi=(0, 0.5, 1, 1), num_shadows_limit=1, shadow_dimension=3,	0.2
3	GaussNoise	mean=0, std_range=(0.05, 0.1)	0.25
4	MotionBlur	blur_limit=2	0.15
5	RandomFog	alpha_coef=0.03, fog_coef_range=(0.01, 0.05)	0.2
6	RandomRain	slant_range=(-10, 10), drop_length=20, drop_width=1, drop_color=(200, 200, 200), blur_value=2, brightness_coefficient=0.9	0.2
7	ImageCompression	quality_range=(70, 95)	0.3
8	HorizontalFlip	-	0.5
9	VerticalFlip	-	0.2
10	Rotate	limit=1	0.4

These transformations are applied only to the training set to improve robustness and generalization, while validation and testing sets remain unaltered to ensure fair evaluation. This approach aims to enrich the training data distribution without claiming equivalence to real captured field data.

C. TDR-Model

The proposed framework in this study is based on the TDR-Model, which integrates MobileNetV3 as the backbone with CBAM and ODC to enhance feature representation. MobileNetV3 employs depthwise separable convolution to reduce computational complexity. However, its static kernel structure limits adaptability to complex visual variations. In this study, ODC is inserted after the depthwise convolution layer within the bottleneck blocks, enabling dynamic kernel re-weighting across spatial, channel, and kernel dimensions.

Subsequently, CBAM is applied after ODC to refine feature maps through channel and spatial attention mechanisms, emphasizing disease-relevant regions while suppressing background noise. This sequential integration allows ODC to enhance adaptive feature extraction, followed by CBAM-based attention refinement. However, different from the original TDR-Model proposed by Zhang [12], which integrates image dehazing as a preprocessing step to mitigate atmospheric distortion, this study deliberately excludes dehazing operations. The proposed approach investigates whether architectural adaptation combined with field-conditioned data augmentation alone can improve generalization performance, without relying on additional image enhancement preprocessing. The structure of TDR-Model can be seen at Figure 2.

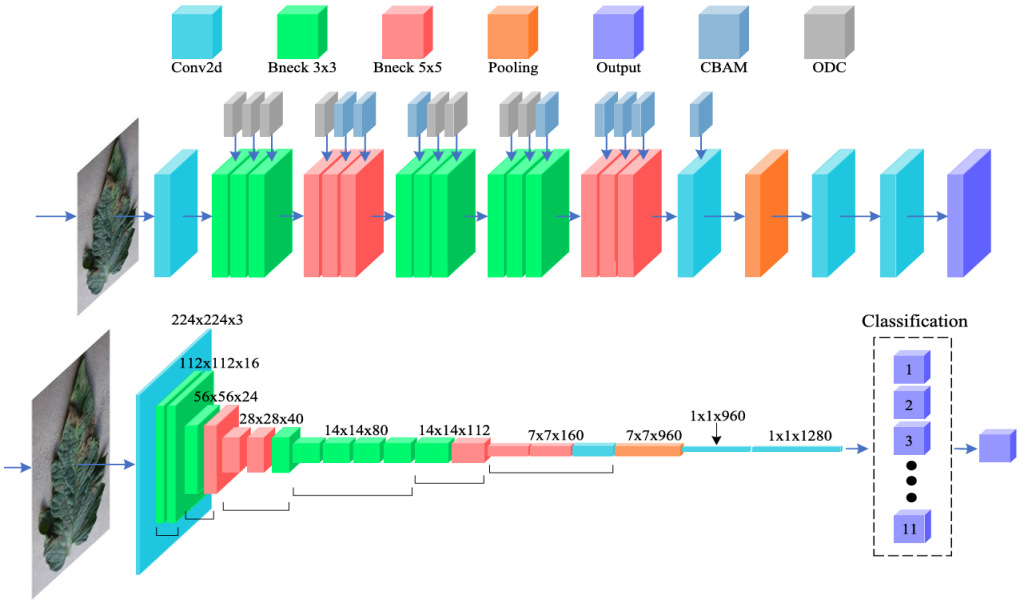

Figure 3. Structure of TDR-Model **[12]**

D. Training Strategy and Evaluation

The proposed model is trained using transfer learning with ImageNet-pretrained weights to accelerate convergence and improve feature generalization, particularly given the limited variability of laboratory-based datasets. The Adam optimizer is employed due to its adaptive learning rate mechanism which is well-suited for handling heterogeneous feature distributions in plant disease images, combined with a Cosine Annealing learning rate scheduler to stabilize convergence and avoid suboptimal local minima. The cross-entropy loss function is used for multi-class classification, aligning with the softmax-based output layer. The detailed training configuration and hyperparameter settings used in this study are summarized in Table 1.

Table 2. Hyperparameter settings

	Name	Value
1	Batch Size	32
2	Learning Rate	0.0001
3	Cosine-Annealing Rate Scheduler	0.000001
4	Epoch	100
5	Optimizer	Adam
6	$\beta 1$	0.9
7	$\beta 2$	0.999
8	Input Size	224x224x3

The proposed model performance is evaluated using overall accuracy, precision, recall, and F1-score, with particular emphasis on F1-score to account for potential class imbalance. In addition, computational efficiency is assessed in terms of model size and inference complexity to ensure suitability for deployment on resource-constrained devices. Model design stages can be seen at Figure 3. To analyze the contribution of the proposed method strategy, comparative experiments with the baseline of TDR-Model are conducted and discussed in the Results and Discussion section.

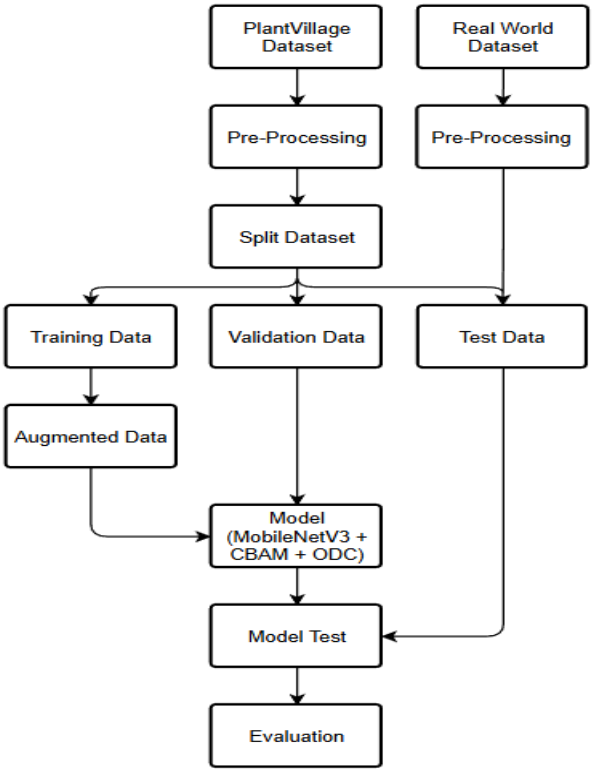

Figure 4. Model design stages

4. Results and Discussion

The proposed model was evaluated separately on two test sets: PlantVillage test set (PV-Test), representing in-domain laboratory conditions, and a real-world tomato leaf image test set (RW-Test), representing natural field conditions. During the training phase, eight model variants were evaluated to isolate the contribution of each component under both standard and augmented training settings. The baseline TDR model and the TDR model with augmentation exhibit stable convergence behavior. As summarized in Table 3, the baseline TDR-Model achieves a training accuracy of 99.45% and a validation accuracy of 99.06%, indicating strong fitting performance on

laboratory-condition data. In the other hand, the TDR-Model with augmentation attains a slightly higher training accuracy of 99.80% but achieves a lower maximum validation accuracy of 96.03%.

Table 3. Train and validation result

	Model	Training Accuracy (%)	Validation Accuracy (%)	Training Loss	Validation Loss
1	MobileNetV3	92.90	91.83	0.0002	0.0014
2	MobileNetV3 with Augmentation	94.02	90.37	0.0053	0.2622
3	MobileNetV3+ODC	91.64	90.95	0.0091	0.0144
4	MobileNetV3+ODC with Augmentation	94.47	90.71	0.0081	0.2652
5	MobileNetV3+CBAM	96.26	95.92	0.0094	0.0117
6	MobileNetV3+CBAM with Augmentation	97.44	93.78	0.0083	0.2297
7	TDR-Model	99.45	99.06	0.0132	0.0073
8	TDR-Model with Augmentation	99.80	96.03	0.0069	0.2714

This reduction in validation accuracy is accompanied by an increase in validation loss, suggesting that the augmented model is exposed to more challenging and diverse visual variations during training. Since the validation set remains composed of laboratory-condition images, the introduced distribution mismatch between augmented training data and validation data may lead to lower validation performance. Rather than indicating training instability, this behavior is consistent with the objective of field-conditioned augmentation, which aims to reduce overfitting to controlled laboratory data patterns. Consequently, training and validation metrics alone are insufficient to assess robustness under domain shift, and the impact of this training behavior is therefore examined through separate evaluations on the PlantVillage test set and Real-World test set.

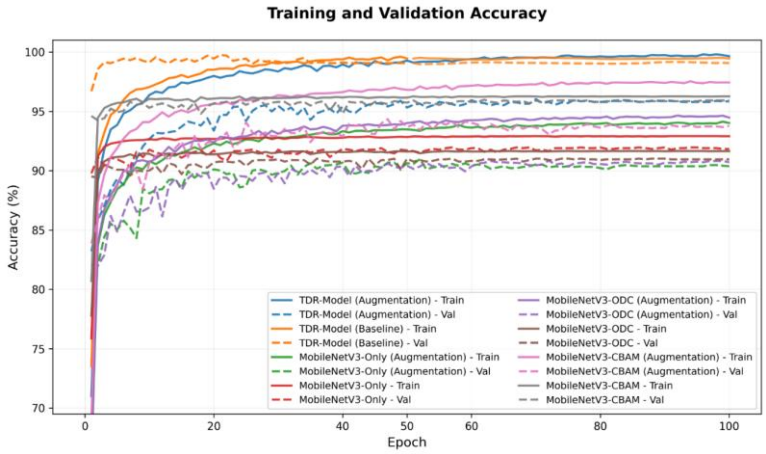

Figure 5. Training and validation accuracy

The corresponding accuracy curves shown in Figure 5 indicate that both the baseline TDR-Model and the TDR-Model with augmentation exhibit stable convergence behavior, with training accuracies increasing rapidly and stabilizing after approximately 30 - 40 epochs. However, a clear difference can be observed in the validation accuracy trends. The baseline model achieves a higher and more tightly aligned validation accuracy with its training performance, indicating tight fitting under laboratory-condition data. In contrast, the augmented model exhibits a lower but more stable validation accuracy, resulting in a wider train-validation gap. This behavior is consistent with the effect of field-conditioned augmentation in increasing training difficulty and acting as a regularization mechanism against overfitting to laboratory-condition patterns.

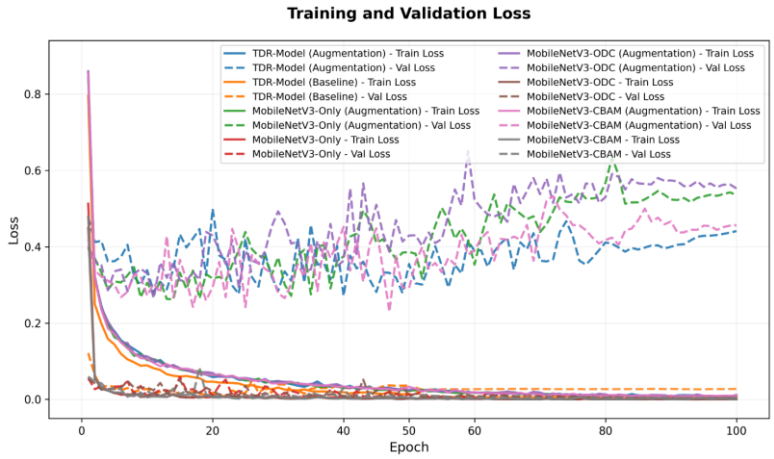

Figure 6. Training and validation loss

Figure 6 illustrates the training and validation loss across epochs for both the baseline TDR-Model and TDR-Model with augmentation. For both models, the training loss decreases rapidly during the initial epochs and gradually converges, indicating stable optimization behavior. The baseline TDR-Model exhibits low and stable validation loss throughout training, closely following the training loss, which suggests strong fitting to laboratory-condition data. In contrast, the TDR-Model with augmentation shows higher and more fluctuating validation loss, with a slight upward trend after reaching its minimum. This behavior reflects the increased training difficulty introduced by field-conditioned augmentation and indicates a regularization effect rather than optimization instability. After the training process converges and the validation performance reaches its optimal point, the proposed model is further evaluated separately on the PlantVillage test set and the Real-World tomato leaf image test set. The evaluation of the proposed model is conducted using accuracy, precision, recall, and F1-score to assess classification performance, alongside parameter count, GFLOPs, frames per second (FPs), and Inference time represents the duration the model needs to analyze a single image from the dataset.

Table 4. Test result on PV-Test

	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Parameter (M)	GFLOPs	FPs	Time (ms)
1	MobileNetV3	99.81	99.79	99.79	99.79	2.5883	0.2229	205.4	4.05
2	MobileNetV3 with Augmentation	99.56	99.50	99.60	99.55	2.5883	0.2229	166.9	4.98
3	MobileNetV3+ODC	99.81	99.77	99.83	99.80	3.9501	0.2343	174.8	5.34
4	MobileNetV3+ODC with Augmentation	99.38	99.42	99.47	99.43	3.9501	0.2343	207.6	4.67
5	MobileNetV3+CBAM	99.63	99.62	99.66	99.64	3.0691	0.2294	189.9	4.85
6	MobileNetV3+CBAM with Augmentation	99.63	99.64	99.67	99.65	3.0691	0.2294	228.2	4.73
7	TDR-Model	98.75	98.55	98.94	98.69	3.4691	0.2328	194.7	5.51
8	TDR-Model with Augmentation	99.69	99.44	99.72	99.58	3.4691	0.2328	198.1	5.38

As shown in Table 4, all model variants achieve high performance on the PlantVillage test set (PV-Test), reflecting strong in-domain classification capability under laboratory-controlled conditions. The baseline TDR-Model achieves an accuracy of 98.75%, while the TDR-Model with augmentation slightly improves the accuracy to 99.69%, indicating that the incorporation of field-conditioned augmentation does not degrade in-domain recognition performance. The consistently high precision, recall, and F1-score values further confirm stable class-wise prediction behavior on PlantVillage data. Additionally, the proposed TDR-Model maintains high computational efficiency, requiring only 3.47M parameters and 0.23 GFLOPs, with an average inference time of approximately 5.15 ms per image, making it suitable for deployment in resource-constrained and real-time agricultural applications.

Table 5. Test result on RW-Test

	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Parameter (M)	GFLOPs	FPs	Time (ms)
1	MobileNetV3	59.32	71.93	51.78	52.12	2.5883	0.2229	164.6	8.24
2	MobileNetV3 with Augmentation	63.05	73.84	59.37	61.93	2.5883	0.2229	153.4	5.28
3	MobileNetV3+ODC	60.68	69.72	53.52	54.34	3.9501	0.2343	188.1	4.21
4	MobileNetV3+ODC with Augmentation	63.73	74.10	60.33	62.57	3.9501	0.2343	174.4	3.22
5	MobileNetV3+CBAM	58.76	69.51	51.25	51.24	3.0691	0.2294	199.2	7.26
6	MobileNetV3+CBAM with Augmentation	63.39	73.10	59.35	61.90	3.0691	0.2294	133.8	8.18
7	TDR-Model	57.51	73.88	49.32	51.76	3.4691	0.2328	194.7	5.51
8	TDR-Model with Augmentation	82.94	85.06	83.04	83.03	3.4691	0.2328	230.7	5.15

Table 5 presents the performance comparison of different model variants evaluated on real-world tomato leaf images captured under field conditions. Unlike the PV-test results, all baseline models trained solely on laboratory-style data experience a substantial performance degradation, indicating a pronounced domain shift between PlantVillage and Real-World environments. Among the baseline models, MobileNetV3, MobileNetV3+ODC, MobileNetV3+CBAM, and the baseline TDR-Model achieve accuracies ranging from 57.51% to 60.68%. Although architectural enhancements such as ODC and CBAM contribute to moderate improvements when combined with augmentation, their individual impact remains limited compared to the proposed TDR-Model. This suggests that architectural modifications alone are insufficient to address severe domain shift without appropriate data-level regularization. To further analyze the classification behavior of the proposed model, a class-wise evaluation is conducted for the TDR-Model with augmentation using precision, recall, F1-score, and per-class accuracy, as presented in Table 6.

Table 6. Precision, Recall, and F1-Score of TDR-Model with Augmentation for each disease class

	Class	Precision (%)	Recall (%)	F1-Score (%)	Support
1	Bacterial Spot	53.42	91.76	67.53	85
2	Early Blight	76.71	84.85	80.58	66
3	Late Blight	98.91	77.78	87.08	117
4	Leaf Mold	65.52	84.51	73.62	71
5	Septoria Leaf Spot	92.86	85.53	89.04	76
6	Two-spotted Spider Mites	100	90.34	94.93	145
7	Target Spot	87.50	73.68	80.00	152
8	Yellow Leaf Curl Virus	93.85	89.71	91.73	68
9	Mosaic Virus	84.44	74.51	79.17	51
10	Healthy	97.67	77.78	86.60	54
	Average	85.06	83.04	83.03	

As shown in Table 6, the proposed model demonstrates consistent class-wise performance across multiple tomato leaf disease categories. High F1-scores are observed for visually distinctive diseases such as Two-spotted Spider Mites (94.93%), Yellow Leaf Curl Virus (91.73%), and Septoria Leaf Spot (89.04%), indicating strong discriminative capability under real-world conditions. However, certain classes exhibit an imbalance between precision and recall. For instance, Bacterial Spot achieves a high recall of 91.76% but a relatively low precision of 53.42%, suggesting that the model tends to over-predict this class when visual symptoms overlap with other foliar diseases as shown in Figure 7. Conversely, Late Blight shows very high precision (98.91%) but lower recall (77.78%), indicating conservative predictions that may miss some true samples. Despite these class-specific challenges, the overall macro-averaged precision, recall, and F1-score of 85.06%, 83.04%, and 83.03%, respectively, indicate balanced multi-class recognition performance. These results suggest that the proposed model maintains robust and stable classification behavior when deployed on real-world tomato leaf images characterized by significant visual variability.

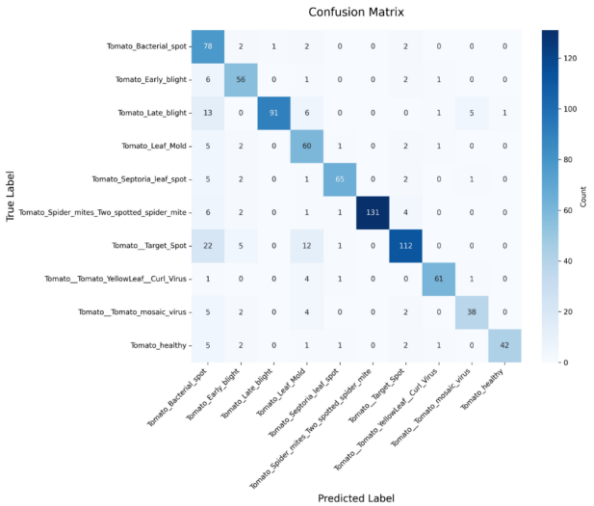

Figure 7. Confusion matrix of TDR-Model with Augmentation

As shown in Figure 7, the confusion matrix demonstrates strong diagonal dominance, indicating that majority of samples are correctly classified across all disease categories. This observation is consistent with the high overall accuracy and balanced per-class performance reported in Table 6. Several classes, such as Tomato_Spider_mites_Two_spotted_spider_mite, Tomato_Target_Spot, and Tomato_Late_blight, achieve high true positive rates, demonstrating the model's capability in learning discriminative visual features for most disease patterns. However, noticeable misclassifications are observed mainly in the Tomato_Target_Spot and Tomato_Late_blight classes. In particular, Tomato_Target_Spot samples are frequently confused with

Tomato_Bacterial_spot and Tomato_Leaf_Mold, while Tomato_Late_blight is occasionally misclassified as Tomato_Bacterial_spot and Tomato_Leaf_Mold. This confusion can be attributed to the visual similarity of disease symptoms, such as overlapping lesion textures and color variations, particularly under real-field imaging conditions. Overall, these evaluation confirms that the proposed model maintains robust classification performance across most disease categories, while the remaining misclassifications highlight visually challenging classes that may benefit from more discriminative feature learning and additional field-diverse training data.

5. Conclusions

This study proposed a lightweight CNN-based classification model for tomato leaf disease recognition using a TDR-Model combined with field-conditioned data augmentation strategies. Experimental results demonstrate that while all model variants achieve high accuracy on the PlantVillage test set, significant performance degradation occurs when evaluated on Real-World field images due to domain shift. By incorporating simulated field-conditioned augmentation, the proposed TDR-Model achieves a substantial improvement, attaining an accuracy of 82.94% on the RW-Test, compared to 57.51% achieved by the baseline TDR-Model trained without augmentation while maintaining identical architectural complexity, indicating that data-level regularization plays a critical role in mitigating domain shift, while architectural enhancements alone are insufficient. In addition, the model also demonstrates balanced classification performance, as reflected by precision (85.06%), recall (83.04%), and F1-score (83.03%), with high per-class accuracy for most disease categories. Analysis of the confusion matrix further confirms that the majority of samples are correctly classified, with misclassifications primarily occurring among visually similar disease classes such as Late Blight and other foliar diseases. In addition to its classification performance, the proposed model demonstrates high computational efficiency, requiring only 3.47 million parameters, 0.23 GFLOPs, and an average inference time of 5.15 ms, making it suitable for deployment in resource-constrained and real-time agricultural applications. Despite these promising results, performance degradation is observed in certain challenging classes, indicating that further improvements can be achieved. Future work may focus on enhancing feature discrimination for visually similar diseases through advanced attention mechanisms, larger and more diverse field datasets, and domain adaptation techniques to further improve robustness under certain real-world conditions.

References

[1] S. Pancono, N. Indroasyoko, and Asep Irfan Setiawan, “Pemantauan dan Deteksi Penyakit Daun Tomat Berbasis IoT dan CNN dengan Aplikasi Android,” *Indones. J. Comput. Sci.*, vol. 13, no. 3, pp. 4692–4709, Jun. 2024, doi: https://doi.org/10.33022/ijcs.v13i3.4083.

[2] B. P. Statistik, “Produksi tanaman sayuran, 2023,” BADAN PUSAT STATISTIK. [Online]. Available: https://www.bps.go.id/id/statistics-table/2/NjEjMg==/produksi-tanaman-sayuran.html

[3] M. H. Najim, S. K. Abdulateef, and A. H. Alasadi, “Early detection of tomato leaf diseases based on deep learning techniques,” *IAES Int. J. Artif. Intell.*, vol. 13, no. 1, p. 509, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp509-515.

[4] N. M. S and P. M. Singh, “Tomato leaves diseases classification using deep learning architecture,” *Math. Stat. Eng. Appl.*, vol. 72, no. 2, pp. 333–350, 2023, [Online]. Available: https://www.philstat.org/index.php/MSEA/article/view/2923

[5] Ellen Flores Mangaoang, “Analysis of Deep Learning Algorithms for Grape Leaf Disease Detection,” *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 33s, pp. 336–344, Apr. 2025, doi: 10.52783/jisem.v10i33s.5537.

[6] A. Howard *et al.*, “Searching for MobileNetV3,” in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2019, pp. 1314–1324. doi: 10.1109/ICCV.2019.00140.

[7] G. Fenu and F. M. Malloci, “Evaluating Impacts between Laboratory and Field-Collected Datasets for Plant Disease Classification,” *Agronomy*, vol. 12, no. 10, p. 2359, Sep. 2022, doi: 10.3390/agronomy12102359.

[8] M. Shafay *et al.*, "Recent advances in plant disease detection: challenges and opportunities," *Plant Methods*, vol. 21, no. 1, p. 140, Oct. 2025, doi: 10.1186/s13007-025-01450-0.
[9] A. Y. Ashurov *et al.*, "Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections," *Front. Plant Sci.*, vol. 15, no. January, pp. 1–16, Jan. 2025, doi: 10.3389/fpls.2024.1505857.
[10] F. Zubair, M. Saleh, Y. Akbari, and S. Al Maadeed, "A Robust Ensemble Model for Plant Disease Detection Using Deep Learning Architectures," *AgriEngineering*, vol. 7, no. 5, p. 159, May 2025, doi: 10.3390/agriengineering7050159.
[11] K. Zhou and D. Dera, "Robust Denoising and DenseNet Classification Framework for Plant Disease Detection," in *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, SCITEPRESS - Science and Technology Publications, 2024, pp. 166–174. doi: 10.5220/0012390400003660.
[12] Z. Zhang, T. Liu, J. Gao, M. Yang, W. Luo, and F. Lin, "TDR-Model: Tomato Disease Recognition Based on Image Dehazing and Improved MobileNetV3 Model," *IEEE Access*, vol. 13, no. December 2024, pp. 852–865, 2025, doi: 10.1109/ACCESS.2024.3522101.
[13] J. Li *et al.*, "Label-Efficient Learning in Agriculture: A Comprehensive Review," *Comput. Electron. Agric.*, vol. 215, p. 108412, May 2023, doi: 10.1016/j.compag.2023.108412.
[14] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, "Albumentations: fast and flexible image augmentations," *Information*, vol. 11, no. 2, p. 125, Sep. 2018, doi: 10.3390/info11020125.
[15] R. Chen, H. Qi, Y. Liang, and M. Yang, "Identification of plant leaf diseases by deep learning based on channel attention and channel pruning," *Front. Plant Sci.*, vol. 13, no. November, pp. 1–15, Nov. 2022, doi: 10.3389/fpls.2022.1023515.
[16] 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, p. 322, Jul. 2025, doi: 10.1007/s10462-025-11234-6.
[17] U. V. Nnamdi and V. Abolghasemi, "Optimised MobileNet for very lightweight and accurate plant leaf disease detection," *Sci. Rep.*, vol. 15, no. 1, p. 43690, Dec. 2025, doi: 10.1038/s41598-025-27393-z.
[18] C. Pal, S. Karmakar, I. Mukherjee, and P. P. Chakrabarti, "A lightweight and explainable CNN model for empowering plant disease diagnosis," *Sci. Rep.*, vol. 15, no. 1, p. 30720, Aug. 2025, doi: 10.1038/s41598-025-94083-1.
[19] M. K. A. Mazumder, M. M. Kabir, A. Rahman, M. Abdullah-Al-Jubair, and M. F. Mridha, "DenseNet201Plus: Cost-effective transfer-learning architecture for rapid leaf disease identification with attention mechanisms," *Heliyon*, vol. 10, no. 15, p. e35625, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35625.
[20] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," *Lect. NOTES Comput. Sci. (INCLUDING Subser. Lect. NOTES Artif. Intell. Lect. NOTES BIOINFORMATICS)*, vol. 11211 LNCS, pp. 3–19, Jul. 2018, doi: 10.1007/978-3-030-01234-2_1.
[21] C. Li, A. Zhou, and A. Yao, "Omni-Dimensional Dynamic Convolution," *ICLR 2022 - 10TH Int. Conf. Learn. Represent.*, pp. 1–20, Sep. 2022, doi: 10.48550/arXiv.2209.07947.
[22] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Front. Plant Sci.*, vol. 7, no. September, pp. 1–10, Sep. 2016, doi: 10.3389/fpls.2016.01419.
[23] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc," Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. [Online]. Available: https://github.com/pratikkayal/PlantDoc-Dataset
[24] Tomato Disease 2, "Tomato Target Spot Dataset," Roboflow. [Online]. Available: https://universe.roboflow.com/tomato-disease-2/tomato-target-spot-vs39n
[25] Thesis, "Tomato - Spider Mites - Two-Spotted Spider Mite Dataset," Roboflow. [Online]. Available: https://universe.roboflow.com/thesis-okplj/tomato-spider-mites-two-spotted-spider-mite