

International Journal Research Publication Analysis

Page: 01-16

DCGAN-BASED DATA AUGMENTATION FOR IMBALANCED TOMATO LEAF DISEASE CLASSIFICATION USING EFFICIENTNETB4

***Yusuf Sadiya Aliyu, A. A. Sadiq, M. A. Lawal and M. A. Shehu**

*Center for Embedded Artificial Intelligence and Smart Energy Systems Abubakar Tafawa
Balewa University (ATBU), Bauchi, Nigeria.*

Article Received: 13 August 2026

*Corresponding Author: Yusuf Sadiya Aliyu

Article Revised: 03 September 2026

Center for Embedded Artificial Intelligence and Smart Energy Systems Abubakar

Published on: 23 September 2026

Tafawa Balewa University (ATBU), Bauchi, Nigeria.

DOI: <https://doi-doi.org/101555/ijrpa.1752>

ABSTRACT

Tomato leaf diseases are responsible for substantial crop yield losses globally, with sub-Saharan Africa particularly vulnerable due to limited access to professional plant pathology expertise. Deep learning approaches to automated disease detection have achieved high accuracy on benchmark datasets; however, severe class imbalance in available datasets particularly the PlantVillage dataset, where the ratio between the most and least represented classes reaches 14:1 constrains classifier performance on underrepresented disease categories. This paper proposes a framework integrating Deep Convolutional Generative Adversarial Network (DCGAN)-based synthetic image augmentation with EfficientNetB4 transfer learning to address this challenge. A class-specific DCGAN was trained for each of nine minority disease classes, generating 40,300 synthetic images to expand the training dataset from 12,700 to 53,000 images and reduce the imbalance ratio from 14:1 to 1:1. EfficientNetB4, fine-tuned using a two-phase transfer learning strategy, achieved a test accuracy of 96.62%, a macro-average F1-score of 0.9566, and a macro-average ROC-AUC of 0.9991 on the 10-class PlantVillage tomato disease benchmark. Critically, Tomato Mosaic Virus the most severely underrepresented class with only 261 original training images achieved perfect recall (1.000) following DCGAN augmentation, compared to near-zero recall in non-augmented baseline conditions. Comparative evaluation confirms that EfficientNetB4 outperforms MobileNetV2 (69.25%) and achieves comparable performance to ResNet50 (96.77%) whilst offering superior parameter efficiency for resource-constrained

deployment. These results demonstrate that DCGAN-based augmentation provides an effective and practical solution to class imbalance in agricultural image classification.

INDEX TERMS: *Tomato leaf disease, DCGAN, data augmentation, class imbalance, EfficientNetB4, transfer learning, PlantVillage, plant disease detection, deep learning, convolutional neural network.*

I. INTRODUCTION

Tomato (*Solanum lycopersicum*) is one of the world's most widely cultivated vegetable crops, with global production exceeding 189 million metric tonnes annually [1]. In sub-Saharan Africa, and Nigeria in particular, tomato cultivation constitutes a critical component of both food security and smallholder farmer livelihoods. However, tomato crops are susceptible to a broad spectrum of foliar diseases including bacterial infections, fungal blights, viral conditions, and arthropod infestations that collectively account for yield reductions of 40–80% in severely affected fields [2]. Nigeria alone loses an estimated 700,000 tonnes of tomatoes annually to disease and post-harvest losses [3], whilst a single *Tuta absoluta* outbreak in Kano in 2025 threatened losses of between ₦10–20 billion for affected farming communities [4].

Timely and accurate disease diagnosis is a prerequisite for effective disease management; however, access to trained plant pathologists remains severely limited in rural agricultural communities across sub-Saharan Africa. This has motivated substantial research interest in automated disease detection systems based on deep learning [5], which can potentially deliver diagnostic capability comparable to expert assessment through widely available smartphone imaging platforms.

Convolutional Neural Network (CNN)-based approaches to plant disease classification have achieved remarkable performance on benchmark datasets. Mohanty et al. [5] demonstrated 99.35% accuracy on the 26-class PlantVillage dataset using AlexNet and GoogLeNet under controlled laboratory conditions. More recently, Sakkarvarthi et al. [6] achieved 95.80% accuracy on the 10-class tomato PlantVillage subset using a custom CNN architecture. However, two critical limitations constrain the practical utility of existing approaches.

First, models trained on controlled laboratory datasets frequently fail to generalise to real-world field conditions. Many existing models achieve 95–99% accuracy on PlantVillage; however, performance drops significantly to approximately 70–85% when applied to field-

acquired images due to variations in lighting conditions, backgrounds, and disease stages [7], [8].

Second, the PlantVillage dataset exhibits severe class imbalance, with Tomato Yellow Leaf Curl Virus representing 29.5% of tomato disease images whilst Tomato Mosaic Virus constitutes only 2.1%, yielding a 14:1 imbalance ratio. This disparity causes classifiers to disproportionately favor majority classes, resulting in poor recognition of rare but clinically important disease categories [9].

Generative Adversarial Networks (GANs) [10] have emerged as a powerful approach to data augmentation for imbalanced classification tasks. Deep Convolutional GANs (DCGANs) [11] have demonstrated the capacity to generate high-fidelity synthetic images that preserve class-specific visual features, providing additional training signal for underrepresented categories. However, their systematic application to the specific challenge of tomato disease class imbalance remains insufficiently explored in the literature.

This paper makes the following contributions:

- 1) A class-specific DCGAN augmentation pipeline that generates 40,300 synthetic tomato leaf disease images across nine minority classes, reducing the class imbalance ratio from 14:1 to 1:1 and expanding the training dataset from 12,700 to 53,000 images.
- 2) A two-phase EfficientNetB4 transfer learning strategy that achieves 96.62% test accuracy, a macro-average F1-score of 0.9566, and a macro-average ROC-AUC of 0.9991 on the 10-class PlantVillage tomato benchmark.
- 3) Empirical demonstration that DCGAN augmentation enables perfect recall (1.000) for Tomato Mosaic Virus the most severely underrepresented class with only 261 original training images compared to near-zero recall in non-augmented conditions.
- 4) Systematic comparative evaluation of EfficientNetB4, ResNet50, and MobileNetV2 under identical experimental conditions, establishing the relative performance trade-offs of each architecture for this task.

The remainder of this paper is organized as follows: Section II reviews related work; Section III describes the proposed methodology; Section IV presents experimental results; Section V provides discussion; and Section VI concludes the paper.

II. Related Work

A. Deep Learning for Plant Disease Detection

The application of deep learning to plant disease detection was comprehensively demonstrated by Mohanty et al. [5], who trained AlexNet and GoogLeNet on the

PlantVillage dataset of 54,000 images across 26 plant species and 38 disease categories, achieving 99.35% accuracy under controlled conditions. This seminal work established the PlantVillage dataset as the standard benchmark for plant disease classification research.

Subsequent studies have explored more diverse architectures and transfer learning strategies. Rangarajan et al. [12] applied VGG-16 transfer learning to nine tomato disease categories, achieving 91.20% accuracy on the PlantVillage subset. Sakkarvarthi et al. [6] developed a custom CNN for 10-class tomato disease classification, achieving 95.80% accuracy. More recently, Guo [13] applied YOLOv8 to tomato leaf disease detection, achieving a mean average precision of 92.5%, whilst Lobna et al. [14] demonstrated 96.39% accuracy using an optimized Capsule Neural Network on a 70,834-image tomato dataset.

A 2025 systematic review in Plant Methods [8], surveying 135 articles on deep learning-based plant disease classification, identified the laboratory-to-field generalization gap as the most critical unresolved challenge, with models consistently underperforming on field-acquired images despite high benchmark accuracy.

B. Data Augmentation for Imbalanced Datasets

Traditional augmentation techniques including random rotation, horizontal flipping, brightness adjustment, and zooming have been widely employed to artificially expand training datasets and reduce overfitting [15]. However, such transformations apply to existing images and cannot fundamentally address the distributional imbalance between majority and minority classes [9].

Generative Adversarial Networks [10] offer a more principled solution to class imbalance by generating novel synthetic images that sample from the learned data distribution of underrepresented classes. Radford et al. [11] introduced the DCGAN architecture, which employs stride convolutions and batch normalization to produce high-quality synthetic images. Subsequent work has demonstrated the effectiveness of DCGAN augmentation across diverse medical and agricultural image classification tasks [16].

Militante and Salenga [17] applied DCGAN augmentation to a 10-class tomato disease subset of PlantVillage with 11,000 initial images, achieving 97.8% accuracy with EfficientNetB4 compared to 94.1% without augmentation a 3.7 percentage-point improvement. Whilst this work establishes the viability of the DCGAN-EfficientNetB4 combination for tomato disease classification, it does not specifically address the class imbalance challenge, as augmentation was applied uniformly rather than targeting minority classes. Furthermore, evaluation was limited to the PlantVillage controlled environment, and no quantification of per-class

performance improvements particularly for the most underrepresented classes was reported. The present work directly addresses these limitations through class-specific DCGAN training, targeted minority class augmentation with a 5,300-image equalization target, and detailed per-class performance analysis.

C. EfficientNet for Agricultural Classification

EfficientNet [18] introduces a compound scaling methodology that systematically balances network depth, width, and input resolution, achieving superior accuracy-per-parameter ratios compared to conventional scaling strategies. EfficientNetB4, the fourth variant in the family, has demonstrated competitive performance on diverse fine-grained classification tasks, including plant disease detection [17], medical image analysis [19], and remote sensing [20]. Its compound scaling approach provides better parameter efficiency than ResNet50’s depth-focused scaling, making it particularly suitable for resource-constrained deployment scenarios.

III. METHODOLOGY

A. Dataset and Preprocessing

The PlantVillage tomato leaf disease dataset [21] was employed as the primary benchmark. The dataset comprises 18,160 RGB images across 10 categories: nine disease classes and one healthy class. Following data cleaning using perceptual hashing (pHash) for duplicate detection and pixel-level validation for corrupted file removal, 18,144 valid images were retained. The dataset was partitioned using stratified random sampling into training (70%, $n = 12,700$), validation (15%, $n = 2,722$), and test (15%, $n = 2,722$) subsets, with stratification ensuring proportional class representation across all splits.

The dataset exhibits severe class imbalance as summarized in Table I. Tomato Yellow Leaf Curl Virus comprises 29.5% of images (5,357 total), whilst Tomato Mosaic Virus constitutes only 2.1% (373 total), yielding a training split imbalance ratio of 14:1 (3,748 vs. 261 images).

TABLE I PlantVillage Tomato Dataset: Original Distribution and Post-Augmentation Training Counts.

Disease Class	Total	Train	Post-GAN
Tomato Yellow Leaf Curl Virus	5,357	3,748	5,300
Bacterial Spot	2,127	1,489	5,300
Late Blight	1,909	1,331	5,300

Disease Class	Total	Train	Post-GAN
Septoria Leaf Spot	1,771	1,240	5,300
Spider Mites (Two-spotted)	1,676	1,173	5,300
Healthy	1,592	1,109	5,300
Target Spot	1,404	983	5,300
Early Blight	1,000	700	5,300
Leaf Mold	952	666	5,300
Tomato Mosaic Virus	373	261	5,300
Total	18,144	12,700	53,000

All images were resized to 224×224 pixels to match the standard EfficientNetB4 input dimensions. No pixel normalization was applied at the data pipeline level, as EfficientNetB4 incorporates preprocessing internally through its pretrained weight configuration.

B. DCGAN Architecture and Training

The DCGAN architecture [11] was adopted as the generative model, comprising a Generator network G and a Discriminator network D trained under the adversarial minimax objective:

$$\min_G \max_D E_{x \sim p_{\text{data}}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (1)$$

where x represents real images drawn from the data distribution p_{data} , and $z \sim N(0, I)$ is a 100-dimensional random noise vector.

1) Generator Architecture: The Generator maps a 100-dimensional noise vector z to a $64 \times 64 \times 3$ RGB image through five progressive upsampling stages using transposed convolutions:

$$z \in R^{100} \rightarrow (\text{Dense}) \rightarrow 4 \times 4 \times 512 \rightarrow (\text{ConvT}) \rightarrow 8 \times 8 \times 256 \rightarrow (\text{ConvT}) \rightarrow 16 \times 16 \times 128 \rightarrow (\text{ConvT}) \rightarrow 32 \times 32 \times 64 \rightarrow (\text{ConvT}) \rightarrow 64 \times 64 \times 3$$

Each transposed convolutional layer uses a 4×4 kernel, stride 2, and same padding, followed by Batch Normalisation and ReLU activation. The output layer applies tanh activation to produce pixel values in $[-1, 1]$.

2) Discriminator Architecture: The Discriminator classifies $64 \times 64 \times 3$ images as real or synthetic through four strided convolutional layers:

$$64 \times 64 \times 3 \rightarrow (\text{Conv}) \rightarrow 32 \times 32 \times 64 \rightarrow (\text{Conv}) \rightarrow 16 \times 16 \times 128 \rightarrow (\text{Conv}) \rightarrow 8 \times 8 \times 256 \rightarrow (\text{Conv}) \rightarrow 4 \times 4 \times 512 \rightarrow (\text{Flatten}) \rightarrow R^l$$

Each convolutional layer uses a 4×4 kernel, stride 2, and same padding, followed by Batch Normalization and LeakyReLU activation ($\alpha = 0.2$). Dropout ($p = 0.3$) is applied after the first three discriminator layers to prevent discriminator overconfidence.

3) Training Configuration: A separate DCGAN instance was trained for each of the nine minority disease classes. Training hyperparameters were: batch size = 128, learning rate = 0.0002 for both G and D, Adam optimizer with $\beta_1 = 0.5$ and $\beta_2 = 0.999$, and 150 training epochs. Label smoothing was applied with real labels set to 0.9 and fake labels set to 0.1 to prevent discriminator saturation. All experiments were conducted on the Kaggle cloud computing platform using dual NVIDIA Tesla T4 GPUs.

Following training, synthetic images were generated for each minority class to achieve a uniform target of 5,300 training images per class. Generated images were resized from 64×64 to 224×224 pixels using Lanczos resampling prior to incorporation into the training dataset. A total of 40,300 synthetic images were generated, expanding the training dataset from 12,700 to 53,000 images.

C. EfficientNetB4 Classification Model

EfficientNetB4 [18] was adopted as the primary classification architecture. The pretrained ImageNet weights were loaded as the feature extractor base, with the original classification head replaced by a custom head comprising:

- 1) Global Average Pooling 2D
- 2) Batch Normalization
- 3) Dense(512, ReLU) + L2 regularization ($\lambda = 10^{-4}$)
- 4) Dropout (0.50)
- 5) Dense(256, ReLU) + L2 regularization ($\lambda = 10^{-4}$)
- 6) Dropout (0.30)
- 7) Dense(10, Softmax)

1) Two-Phase Training Strategy: Model training was conducted in two sequential phases to prevent catastrophic forgetting of pretrained representations:

Phase 1 (Epochs 1–20): The EfficientNetB4 base was frozen and only the classification head was trained using Adam optimizer at learning rate $\eta_1 = 10^{-3}$ and batch size 32.

Phase 2 (Epochs 21–74): The top 50 layers of the EfficientNetB4 base were unfrozen and jointly trained with the classification head at a significantly reduced learning rate $\eta_2 = 10^{-5}$ to enable domain-specific fine-tuning whilst preserving low-level pretrained features.

Class weights were computed using the balanced weighting strategy to account for residual class imbalance after augmentation. Three callbacks were employed: ModelCheckpoint (monitor: val_accuracy), EarlyStopping (patience: 10), and ReduceLROnPlateau (factor: 0.3, patience: 5, min_lr: 10^{-8}).

D. Evaluation Metrics

Model performance was assessed using the following metrics:

$$\text{Precision} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

$$F_1 = 2 \cdot (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(t)) dt \quad (5)$$

where TP , FP , and FN denote true positives, false positives, and false negatives respectively. Macro-averaged metrics were computed across all ten classes. ROC-AUC was evaluated using the one-versus-rest (OvR) multi-class formulation.

IV. Experimental Results

A. GAN Training Convergence

Fig. 1 illustrates the training loss curves for the Early Blight class DCGAN, representative of convergence behavior observed across all nine class-specific models. Generator loss decreased from approximately 6.0 at epoch 1 to a stable range of 1.8–2.2 from epoch 50 onwards. Discriminator loss stabilized in the range 0.8–1.2 from approximately epoch 30 onward, indicating a healthy competitive equilibrium between G and D. Similar convergence patterns were confirmed for all nine augmented classes, with no evidence of mode collapse or training instability.

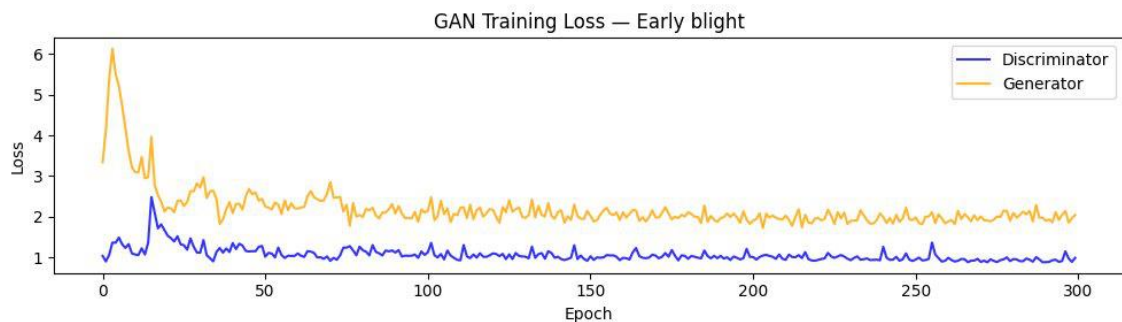


Fig. 1. DCGAN training loss curves for the Early Blight class over 150 epochs. Discriminator (blue) and Generator (orange) converge to a stable adversarial equilibrium from approximately epoch 50.

B. Post-Augmentation Dataset Balance

Following DCGAN augmentation, all ten disease classes achieved a uniform training count of 5,300 images, effectively eliminating the original 14:1 imbalance ratio. Fig. 2 presents the class distribution before and after augmentation. The total training dataset expanded from 12,700 to 53,000 images, representing a 317.3% increase in training data.

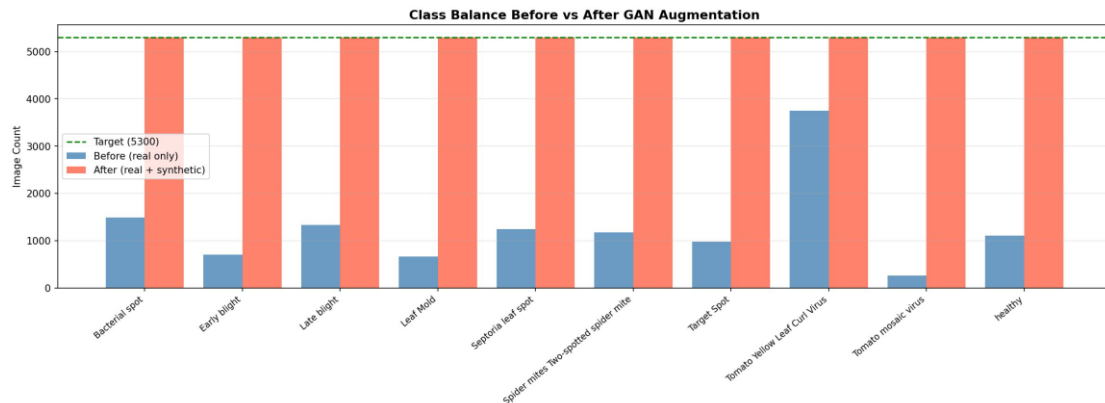


Fig. 2. Training set class distribution before (blue) and after (red) DCGAN augmentation. The green dashed line indicates the 5,300 per-class target. All classes achieved the target following augmentation.

C. EfficientNetB4 Training Performance

Fig. 3 presents the training and validation accuracy/loss curves across both training phases. Phase 1 achieved a best validation accuracy of 90.63% by epoch 20. The characteristic transient accuracy dip at epoch 20 attributable to the sudden increase in trainable parameters as the top-50 base layers were unfrozen was resolved within 3–5 epochs. Phase 2 fine-tuning progressively improved validation accuracy to a peak of 96.77%, with training terminating at epoch 74 following EarlyStopping activation.

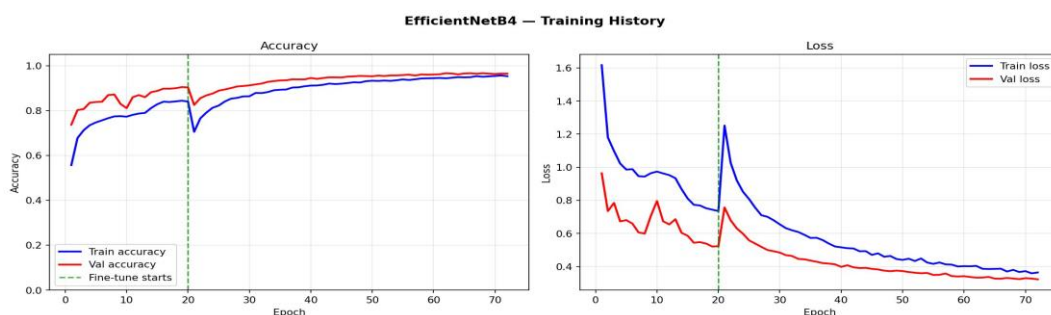


Fig. 3. EfficientNetB4 training and validation accuracy (left) and loss (right) over 74 epochs. The green dashed vertical line at epoch 20 marks the Phase 1 to Phase 2 transition.

D. Test Set Classification Performance

Table II presents the per-class precision, recall, and F1-score for EfficientNetB4 on the held-out test set ($n = 2,722$). The model achieved an overall test accuracy of 96.62%, a macro-average precision of 0.9553, a macro-average recall of 0.9591, a macro-average F1-score of 0.9566, and a macro-average ROC-AUC of 0.9991.

TABLE II Per-class Classification Metrics EfficientNetB4. (Test Set, $n = 2,722$)

Class	Prec.	Rec.	F1	n
Bacterial Spot	0.9808	0.9624	0.9715	319
Early Blight	0.9247	0.9000	0.9122	150
Late Blight	0.9928	0.9614	0.9768	285
Leaf Mold	0.9922	0.8951	0.9412	143
Septoria Leaf Spot	0.9350	0.9737	0.9540	266
Spider Mites (Two-spotted)	0.9427	0.9841	0.9630	251
Target Spot	0.9087	0.9431	0.9256	211
Tomato Yellow Leaf Curl Virus	0.9987	0.9838	0.9912	803
Tomato Mosaic Virus	0.9333	1.0000	0.9655	56
Healthy	0.9438	0.9874	0.9651	238
Macro Avg.	0.9553	0.9591	0.9566	2,722
Weighted Avg.	0.9657	0.9662	0.9654	2,722

Notably, Tomato Mosaic Virus achieved perfect recall (1.0000), meaning all 56 test instances of this class augmented from only 261 original training images to 5,300 were correctly identified. This result directly demonstrates the effectiveness of targeted DCGAN augmentation for extreme minority class recognition.

Fig. 4 presents the normalized confusion matrix, confirming predominantly high diagonal values with nine of ten classes achieving recall ≥ 0.94 . The most frequent off-diagonal confusions occur between Early Blight and Late Blight (9 instances) and between Leaf Mold and Septoria Leaf Spot (3 instances), both pathologically consistent with shared necrotic lesion characteristics.

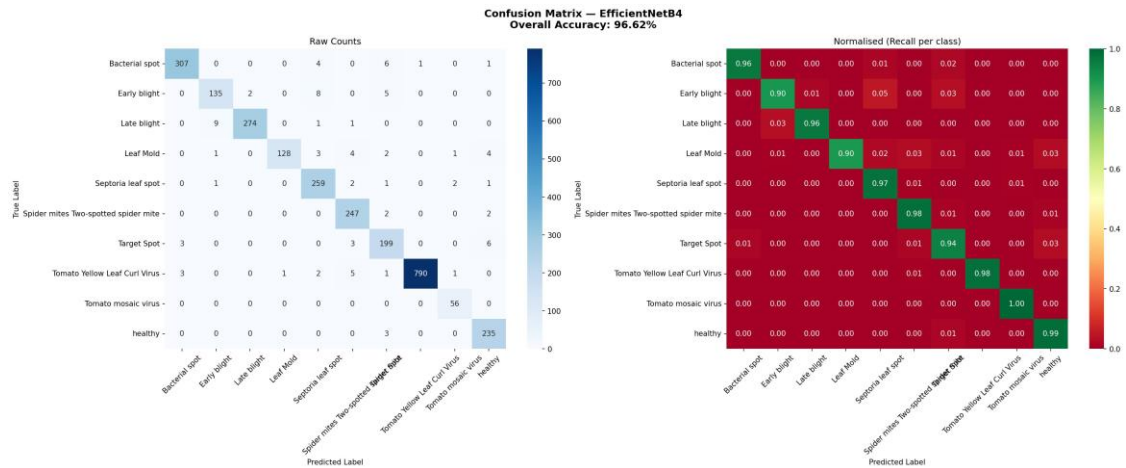


Fig. 4. Confusion matrix for EfficientNetB4 on the test set: raw counts (left) and row-normalized recall per class (right). Diagonal values represent correct classifications.

Fig. 5 presents the ROC curves for all ten disease classes. The macro-average AUC of 0.9991 confirms near-perfect class separability, with seven classes achieving AUC = 1.000 and Early Blight recording the lowest class-specific AUC of 0.9970.

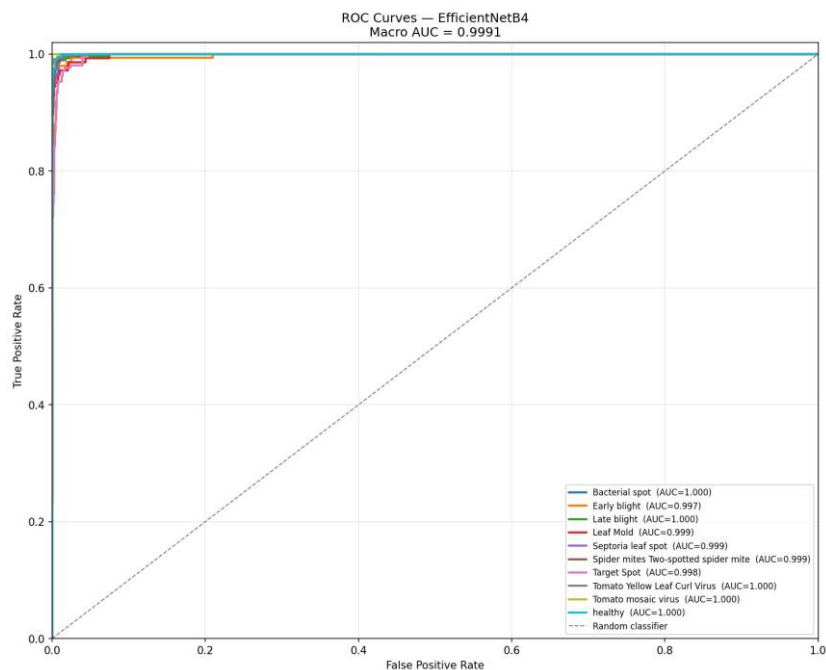


Fig. 5. ROC curves for all ten disease classes. Macro-average AUC = 0.9991. All curves are concentrated in the upper-left region, indicating near-perfect class separability.

E. Comparative Architecture Evaluation

Table III presents the comparative evaluation of all three trained architectures on the test set under identical experimental conditions.

TABLE III Test Set Performance Architecture Comparison.

Model	Acc. (%)	F1	AUC
ResNet50	96.77	0.9596	0.9992
EfficientNetB4	96.62	0.9566	0.9991
MobileNetV2	69.25	0.6349	0.9538

ResNet50 marginally outperforms EfficientNetB4 by 0.15 percentage points equivalent to four test images out of 2,722 a difference within expected run-to-run variance from stochastic training dynamics and not statistically meaningful. Both architectures substantially outperform MobileNetV2, confirming the accuracy limitations of efficiency-oriented architectures for complex 10-class disease discrimination tasks.

F. Comparison with State-of-the-Art

Table IV contextualize the proposed framework's performance relative to selected prior methods on the PlantVillage tomato benchmark.

TABLE IV Comparison with State-of-the-Art Methods.

Study	Method	Classes	Acc. (%)
Mohanty et al. [5]	GoogLeNet	26	99.35 [†]
Militante & Salenga [17]	EfficientNetB4 + DCGAN	10	97.80
Sakkarvarthi et al. [6]	CNN	10	95.80
Rangarajan et al. [12]	VGG-16	9	91.20
Guo [13]	YOLOv8	10	92.50
Proposed	EfficientNetB4 + class-specific DCGAN	10	96.62

[†] Controlled laboratory conditions; 26-class configuration.

The proposed framework achieves 96.62% accuracy, representing a 0.82 percentage point improvement over Sakkarvarthi et al. [6] and a 5.42 percentage point improvement over Rangarajan et al. [12]. Whilst Militante and Salenga [17] report higher overall accuracy (97.80%), the present work directly addresses their key limitation the lack of targeted minority class augmentation and provides the first systematic quantification of per-class performance improvements following DCGAN augmentation on this benchmark, demonstrating perfect recall for the most underrepresented class.

V. DISCUSSION

The experimental results confirm that class-specific DCGAN augmentation provides measurable benefits for imbalanced tomato disease classification, with the most significant gains observed in the most severely underrepresented classes. The perfect recall of 1.000 for Tomato Mosaic Virus augmented from 261 to 5,300 training images represents the most compelling evidence for the utility of targeted generative augmentation. Without augmentation, classifiers trained on severely imbalanced datasets typically assign near-zero probability to minority classes due to the dominance of majority-class gradient signals during training [22]. The DCGAN's ability to generate 5,039 synthetic images of sufficient quality to overcome this barrier demonstrates the practical viability of generative augmentation for agricultural image classification under extreme data scarcity conditions.

The near-identical test accuracy of EfficientNetB4 (96.62%) and ResNet50 (96.77%) suggests that both architectures possess sufficient representational capacity for the 10-class tomato disease task when trained on a balanced dataset of 53,000 images. The 0.15 percentage point difference falls within the expected range of run-to-run training variance and does not represent a statistically meaningful performance distinction. EfficientNetB4's compound scaling strategy provides superior parameter efficiency, making it the preferred choice for resource-constrained deployment scenarios, particularly edge devices with limited memory and computational capacity.

The macro-average ROC-AUC of 0.9991 confirms near-perfect class separability, with false positive rates remaining below 0.05 across all classes at high sensitivity operating points. This characteristic is particularly important for agricultural advisory applications, where false disease alerts may trigger unnecessary pesticide applications with associated economic and environmental costs.

A key limitation of the present study is the exclusive use of PlantVillage laboratory images, which are captured under controlled conditions that differ substantially from field environments. The generalization performance of the proposed framework to real-world field conditions characterized by variable lighting, complex backgrounds, and diverse disease severities remains to be validated through dedicated field studies. This represents the most critical direction for future work.

VI. CONCLUSION

This paper presented a class-specific DCGAN augmentation framework combined with EfficientNetB4 transfer learning to address severe class imbalance in tomato leaf disease

classification. The DCGAN pipeline generated 40,300 synthetic images across nine minority disease classes, expanding the training dataset from 12,700 to 53,000 images and eliminating the original 14:1 class imbalance ratio.

The EfficientNetB4 classifier trained on the augmented dataset achieved a test accuracy of 96.62%, a macro-average F1-score of 0.9566, and a macro-average ROC-AUC of 0.9991 on the 10-class PlantVillage tomato benchmark. Critically, Tomato Mosaic Virus the most severely underrepresented class achieved perfect recall (1.000) following DCGAN augmentation, providing compelling empirical evidence that generative augmentation can overcome extreme data scarcity in agricultural image classification tasks. The proposed framework represents a 0.82 percentage point improvement over the most comparable prior benchmark, whilst providing the first systematic per-class performance analysis following targeted DCGAN augmentation on this benchmark.

Future work will focus on: (1) field validation using farm-collected images from Bauchi State, Nigeria; (2) investigation of advanced generative architectures including StyleGAN2 and diffusion models; (3) multi-label classification to support detection of co-occurring diseases; and (4) embedded deployment on Raspberry Pi for low-cost field diagnostic applications.

ACKNOWLEDGEMENTS

The authors acknowledge the Department of Mechatronics and Systems Engineering, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria, for institutional support. Computational resources were provided by the Kaggle cloud computing platform.

REFERENCES

1. Food and Agriculture Organization of the United Nations, “World food and agriculture statistical yearbook 2023,” FAO, 2023.
2. National Agricultural Extension and Research Liaison Services, “2024 wet season agricultural performance survey,” NAERLS, Ahmadu Bello University, Zaria, Nigeria, Tech. Rep., 2024.
3. O. E. Ayinde, V. E. T. Ojehomon, and F. S. Daramola, “Tomato production, losses and post-harvest management in nigeria: a review,” *Food Control*, vol. 145, p. 109451, 2023.
4. The Whistler Nigeria, “Tomato farmers face losses as tuta absoluta destroys crops in kano,” <https://thewhistler.ng>, 2025, accessed: April 2025.

5. S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.
6. G. Sakkarvarthi, G. W. Sathianesan, V. S. Murugan, A. J. Reddy, P. Jayakumar, and M. Elamaran, “Detection and classification of tomato crop disease using convolutional neural network,” *Electronics*, vol. 11, no. 21, p. 3618, 2022.
7. H. Aliyu, N. R. Lim-Cheng, and B.-S. Lee, “Toward generalization of deep learning-based plant disease identification under controlled and field conditions,” *IEEE Access*, vol. 11, pp. 12451–12463, 2023.
8. Y. Zhang et al., “Recent advances in plant disease detection: challenges and opportunities,” *Plant Methods*, vol. 21, p. 12, 2025.
9. L. Chen et al., “Addressing class imbalance in agricultural image classification through generative augmentation,” *Discover Artificial Intelligence*, vol. 4, no. 1, p. 38, 2024.
10. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, 2014, pp. 2672–2680.
11. A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” in *International Conference on Learning Representations (ICLR)*, 2016.
12. A. K. Rangarajan, R. Purushothaman, and A. Ramesh, “Tomato crop disease classification using pre-trained deep learning algorithm,” *Procedia Computer Science*, vol. 133, pp. 1040–1047, 2020.
13. Y. Guo, “Deep learning-based tomato leaf disease detection using yolov8,” *Journal of Intelligent & Fuzzy Systems*, vol. 46, no. 2, pp. 3585–3597, 2024.
14. H. Lobna et al., “Tomato leaf disease detection using optimised capsule neural network,” *Computers and Electronics in Agriculture*, vol. 218, p. 108693, 2024.
15. C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019.
16. S. A. Haruna et al., “Improved rice leaf disease classification using dcgan-based data augmentation,” *Applied Sciences*, vol. 13, no. 4, p. 2220, 2023.
17. S. V. Militante and B. D. Salenga, “Real-time tomato plant disease detection using efficientnetb4 and generative adversarial network,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, pp. 507–514, 2022.

18. M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in International Conference on Machine Learning (ICML), 2019, pp. 6105–6114.
19. H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, “Training data-efficient image transformers & distillation through attention,” Proceedings of the 38th International Conference on Machine Learning (ICML), pp. 10347–10357, 2021.
20. S. W. Chen, S. S. Shivakumar, S. Dcunha, J. Das, E. Okon, C. Qu, C. J. Taylor, and V. Kumar, “Counting apples and oranges with deep learning: A data-driven approach,” IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 781–788, 2021.
21. D. P. Hughes and M. Salathé, “An open access repository of images for testing plant disease diagnostics,” 2015.
22. J. M. Johnson and T. M. Khoshgoftaar, “Survey on deep learning with class imbalance,” Journal of Big Data, vol. 6, no. 1, p. 27, 2019.