

Article

Improving Oriental Melon Leaf Disease Classification via DCGAN-Based Image Augmentation

Myeongyong Kang ¹, Niraj Tamrakar ²  and Hyeon Tae Kim ^{2,*} 

¹ Department of Smart Farm, Graduate School of Gyeongsang National University, Jinju 52828, Republic of Korea; kmy5077@naver.com

² Department of Bio-Systems Engineering, Gyeongsang National University (Institute of Smart Space Agriculture (ISSA)), Jinju 52828, Republic of Korea; niraj@gnu.ac.kr

* Correspondence: bioani@gnu.ac.kr; Tel.: +82-55-772-1896

Abstract

Deep learning-based plant disease classification models often suffer from performance degradation when training data are limited. Hence, generative models offer a promising solution for model performance in plant disease classification. In this work, images representing powdery mildew, downy mildew, and healthy plant leaves were generated using traditional augmentation methods as well as both DCGAN and a modified DCGAN featuring residual connection blocks with varied activation functions. Evaluation metrics IS and FID revealed that the modified DCGAN consistently produced generative images with strong class-distinctive features and greater overall diversity compared to basic GAN methods, with an IS increment of 7.9% to 11.54% and FID decrement of 6.6% to 7.8%. After selecting the best augmentation method, we input the generated images into the training sets for the classification models, AlexNet, VGG16, and Goog-LeNet, to measure improvements in disease recognition. All classifiers benefited from the augmented datasets, with the modified DCGAN-based augmentation yielding the highest precision, recall, and accuracy. GoogLeNet outperformed all classification models, with an overall precision, recall, and F1-Score value of 98%. Notably, this generative approach minimized errors between visually similar categories, such as powdery mildew and healthy samples, by capturing subtle morphological differences. The results confirm that class-aware generative augmentation can both expand the number of training images and preserve the critical features necessary for discrimination, significantly boosting model effectiveness. These advances show the practical potential of generative models not only to enrich datasets but also to improve the accuracy and robustness of plant disease detection for real-world agricultural scenarios.



Academic Editor: Aichen Wang

Received: 9 October 2025

Revised: 5 November 2025

Accepted: 6 November 2025

Published: 8 November 2025

Citation: Kang, M.; Tamrakar, N.; Kim, H.T. Improving Oriental Melon Leaf Disease Classification via DCGAN-Based Image Augmentation. *Agriculture* **2025**, *15*, 2324. <https://doi.org/10.3390/agriculture15222324>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: plant disease diagnosis; data augmentation; Generative Adversarial Networks; residual connection; activation function

1. Introduction

The Oriental melon (*Cucumis melo* L. var. *makuwa*) is a widely cultivated fruit in South and East Asia. According to the Korean Statistical Information Service (KOSIS), the total cultivation area in Korea has steadily increased from 3581 hectares in 2017 to 4580 hectares in 2023, with more than 95% of the produce being grown in greenhouses (<https://kosis.kr> accessed on 15 May 2025). However, maintaining stable fruit yield and quality remains a primary challenge for farmers. One of the major challenges in greenhouse

cultivation is the prevalence of plant diseases, particularly powdery mildew and downy mildew, which significantly impact fruit quality and yield [1]. One of the major issues in greenhouse cultivation is the prevalence of fungal diseases such as powdery mildew and downy mildew, which thrive under humid and controlled environments [2] and are difficult to detect in their early stages [3]. Traditional disease diagnosis relies on expert-based phenotyping, which is labor-intensive, time-consuming, and costly [4]. With advancements in computer vision and artificial intelligence (AI), deep learning-based plant disease detection has emerged as a promising alternative [5]. Deep convolutional neural networks (CNNs) have been extensively used to analyze plant leaf images and classify diseases with high accuracy [6]. However, CNN-based models require large and well-balanced datasets for effective training, which poses a significant challenge in real-world agricultural applications [7]. Publicly available datasets, such as Plant Village, often suffer from class imbalance, where healthy leaf samples outnumber diseased samples, leading to biased model performance [8]. To mitigate this issue, data augmentation techniques are commonly employed, including traditional image transformations such as rotation, scaling, and translation [9]. However, these techniques exhibit inherent limitations, as they lack feature diversity, fail to generate realistic disease patterns, and are unable to preserve critical morphological characteristics, such as vein-associated lesions or powdery textures, thereby increasing the risk of model overfitting [10].

To address these challenges, generative adversarial networks (GANs), introduced by Goodfellow et al. [11], have been increasingly used for data augmentation in plant disease classification [12]. GANs generate synthetic images by learning data distribution through adversarial training between a generator and a discriminator. Existing GAN-based approaches can be categorized into image-to-image translation and noise-to-image generation models. Image-to-image GANs, such as CycleGAN and LeafGAN, modify healthy leaf images by embedding disease symptoms [13,14]. However, these methods depend on paired or domain-matching datasets, which are not always available, and often produce repetitive results [15]. In contrast, deep convolutional GAN (DCGAN), introduced by Radford et al. [16], generates images directly from random noise, enabling the creation of novel and diverse disease features that were absent in the original dataset [17]. This noise-to-image approach enhances data diversity and generalization, leading to more robust deep learning models trained on generated images [18]. Although DCGANs have shown promising results in generating synthetic images for data augmentation, they commonly suffer from limited feature expressiveness, mode collapse, and difficulty in capturing complex structural details that are critical in modeling diverse and fine-grained disease symptoms [19].

To further improve model stability and feature representation, residual connections have been proven effective in enhancing gradient flow and maintaining feature consistency [20,21]. Accordingly, this study applies a simplified residual connection to the DCGAN generator to enhance training stability and structural realism in synthesized disease images. In addition to architectural improvements, other studies have explored various activation functions to improve gradient flow and enhance the non-linear representational capacity of GANs. The rectified linear unit (ReLU) became the standard due to its computational simplicity, but it often suffers from the dead neuron problem [22]. To address this, Maas et al. [23] introduced leaky ReLU, which allows a small gradient flow in the negative domain. Clevert et al. [24] later proposed the exponential linear unit (ELU), offering smoother optimization and faster convergence. More recently, Misra [25] proposed Mish, a self-regularized non-monotonic activation function that provides smoother and more continuous non-linearity, resulting in enhanced gradient propagation and expressive power. Despite these advances, few studies have systematically examined the

combined effects of residual connections and advanced activation functions such as Mish in GAN architectures.

To overcome the limitations of existing GAN-based approaches trained on controlled environment datasets, this study employs images collected from agricultural environments to enable the generative model to learn more realistic and diverse data distributions. This research was carried out in two phases. For the first phase, various data augmentation methods were explored to find the best augmentation method, and then, in the second phase, datasets were applied in classification models to determine performance improvement as well as to find the best-performing classification model.

2. Materials and Methods

2.1. Overview of the Research Methodology

The flow chart in Figure 1 shows the research methodology. Initially, plant leaf diseases and healthy image data were collected to prepare the dataset. The dataset was split into train and test. The naturally collected dataset has a limited number of images and higher class imbalances. The training data was augmented with different methods: traditional, DCGAN, and modified DCGAN. The dataset after augmentation was applied to different DL image classifiers to determine plant leaf disease classification. Finally, the performance of the different classifier models was compared to determine the best-performing augmentation method, as well as the best-performing classifier models among different DL classifiers.

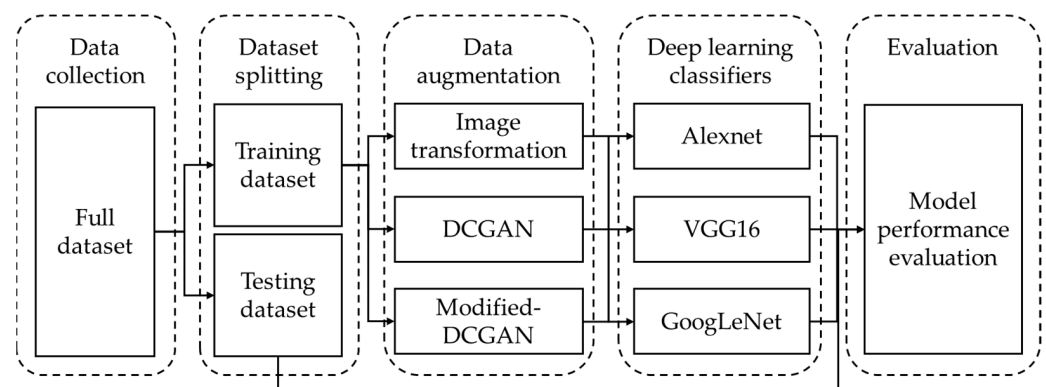


Figure 1. Flowchart of the research for improving the plant disease classification model using various methods of augmentation, namely image transformation (D1), base DCGAN (D2), and modified DCGAN augmentation (D3).

Figure 2 illustrates the selection process of the modified DCGAN algorithm for the generative image for the purpose of augmentation. The modified DCGAN was developed by modifying the base DCGAN model by adding a residual connection, with different activation functions (RELU, ELU, and Mish). The performance of the augmented datasets generated by these methods is evaluated using quantitative and qualitative metrics. The quantitative metrics used were inception score (IS), Fréchet inception distance (FID), and improved precision and recall (IPR). The qualitative metric used was t-distributed stochastic neighbor embedding (t-SNE). The metrics check the intuitive visual representation of how closely the generated images resemble real samples and check whether the synthetic data effectively captures the variability present in real plant disease images. This methodology enables a rigorous evaluation of modified DCGAN data augmentation techniques for synthetic plant disease image generation.

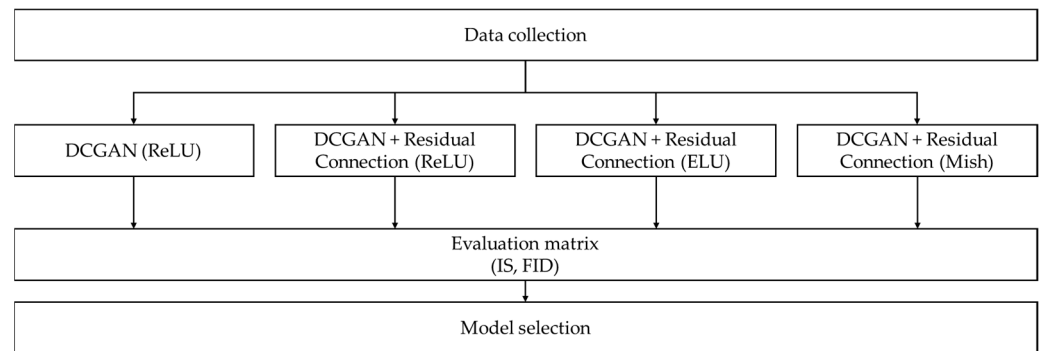


Figure 2. Comparison and mode selection of modified DCGAN using quantitative and qualitative evaluation methods for DL-based plant leaf disease classification.

2.2. Dataset

The oriental melon was cultivated in the greenhouse (Figure 3) located at 35°57' N, 128°19' E (Wolhang-myeon, Seongju-gun, Gyeongsangbuk-do, Republic of Korea). Inside the greenhouse, the average temperature was 29.1 °C, and the humidity was 78.0%. The oriental melon plants were grown without fungicide treatment from June to October 2022. To capture different variations in the leaf disease images, they were obtained on various natural lighting conditions, including morning, afternoon, and evening. The diseases were identified with the guidance of plant cultivation experts. The images were captured using an RGB camera (A32, Samsung electronics, Suwon, Gyeonggi-do, Republic of Korea). A total of 12,000 oriental melon leaf images were collected, comprising 4000 healthy, 4000 downy mildew, and 4000 powdery mildew samples. The dataset was divided into training and test sets at a ratio of 80:20. Data augmentation was applied only to the training set, where an additional 1600 generative images per class were generated using the proposed GAN-based method, corresponding to half of the number of real samples per class.

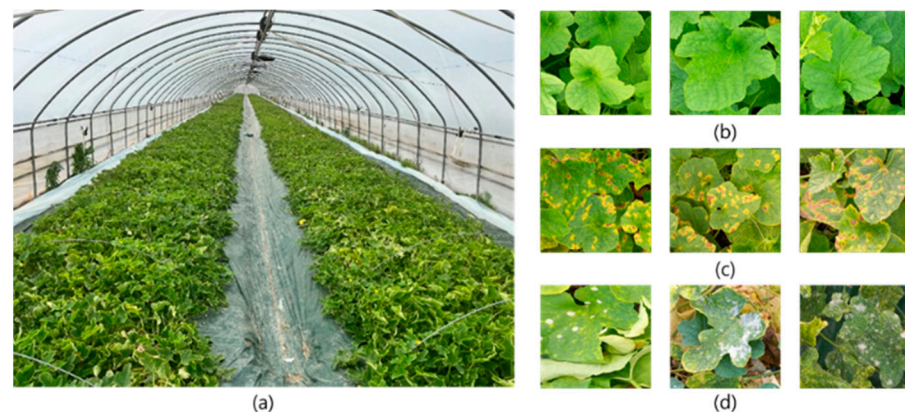


Figure 3. Experimental greenhouse growing oriental melon for plant leaf disease collection: (a) Greenhouse, (b) healthy leaves, (c) downy mildew infected leaves, and (d) powdery mildew infected leaves.

2.3. Data Augmentation

2.3.1. Traditional Data Augmentation

Traditional data augmentation utilizes various image transformations to increase the diversity of the healthy and diseased leaf images (Figure 4). The different types of transformation that have been applied to obtain augmented data involve operations such as crop and resize, horizontal and vertical flip, rotation, erasing, and noise injection using a Gaussian blur.

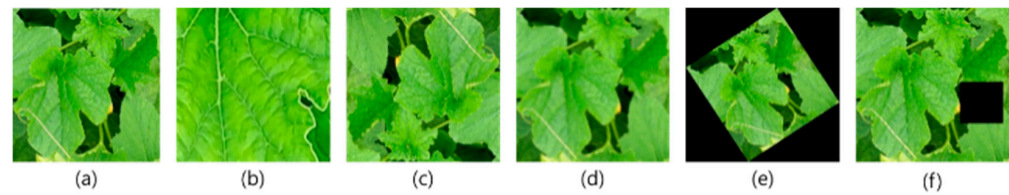


Figure 4. Traditional data augmentation sample images: (a) Original, (b) crop and resize, (c) horizontal and vertical flip, (d) gaussian blur, (e) rotation, and (f) erasing.

2.3.2. Generative Image Augmentation

DCGAN

Generative adversarial networks (GANs) consist of two neural network blocks, called a generator and a discriminator, that compete with each other (Figure 5). The generator creates synthetic images from random noise, while the discriminator attempts to distinguish between real and generated images. Through an adversarial training process, both networks improve simultaneously, allowing the generator to create better realistic images. A major strength of GANs is their ability to model data distributions implicitly, without relying on explicit probability models.

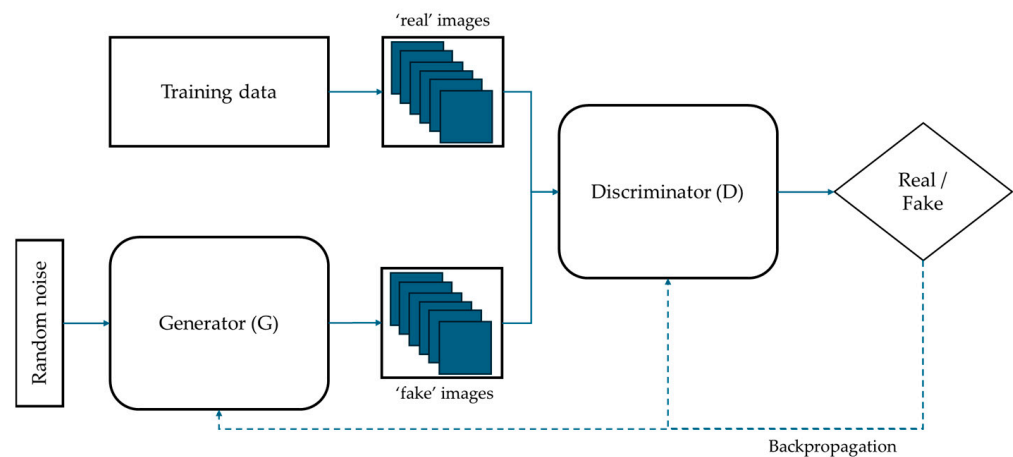


Figure 5. Overall structure of the generative adversarial network.

A deep convolutional generative adversarial network (DCGAN) is an extension of the basic GAN architecture that leverages convolutional neural networks (CNNs) to significantly improve the quality and stability of image generation (Figure 6). In DCGAN, both the generator and discriminator are constructed using CNN layers. The generator progressively upsamples a latent noise vector using transposed convolution layers to produce a full-resolution image, employing ReLU and tanh as activation functions. The discriminator processes input images through convolutional layers with LeakyReLU activations and outputs a probability indicating whether the image is real or generated. Training in DCGAN follows the standard adversarial framework, alternating between updating the discriminator to classify real and generated images correctly, training the generator to produce realistic images that can fool the discriminator. Due to its convolutional architecture and carefully designed layer configurations, DCGAN has demonstrated success in generating diverse and high-quality images across a wide range of domains.

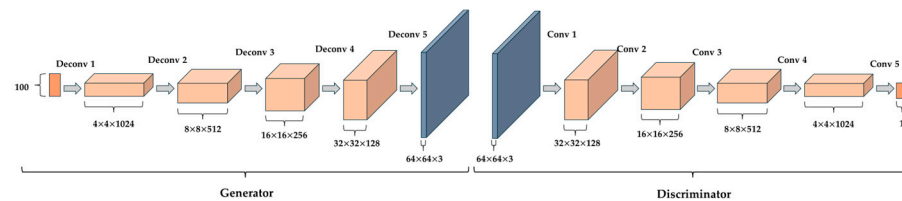


Figure 6. DCGAN architecture with a convolution block generator and discriminator.

Modified DCGAN

Modified DCGAN was achieved by modifying the base DCGAN model by adding a residual connection with various activation functions to achieve higher performance image generation with lesser computation requirements and simpler architecture modifications.

Residual Connection in DCGAN

Residual connections, also known as skip or shortcut connections, are a fundamental architectural component introduced to address the vanishing gradient problem, which hinders the training of deep neural networks. First introduced in ResNet (residual network), this technique facilitates smoother gradient propagation by allowing information to bypass certain layers, making it easier to train deeper networks efficiently. In conventional DCGANs, as the generator deepens, training may become more challenging, affecting the quality of generated images. To mitigate this issue, this study incorporates residual connections using 1×1 kernel convolutions in all layers of the generator (Figure 7). This design ensures a smoother information flow, helping the model retain important features throughout the generation process. Our aim in this is to prevent performance degradation in deeper networks and achieve more stable training.

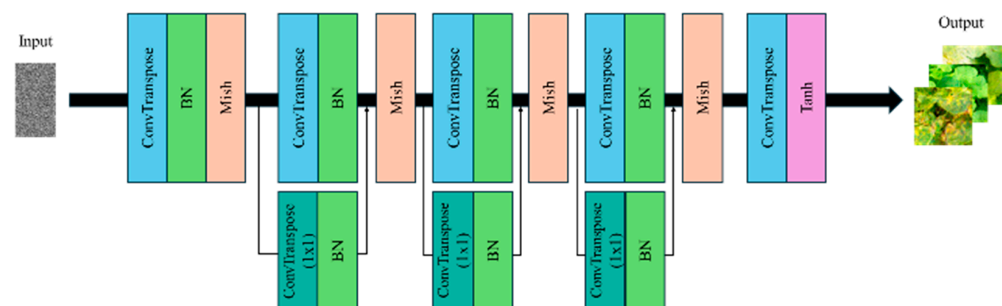


Figure 7. Detailed structure of the generator network used in the proposed model.

This lightweight residual mechanism does not alter the overall depth or structure of DCGAN but enables smoother training and improves the stability of the image generation process. Ultimately, this modification aims to enhance the quality and diversity of the synthesized plant disease images without increasing architectural complexity. Accordingly, Figure 8 shows the architecture of the discriminator network, which is the same as the conventional DCGAN.

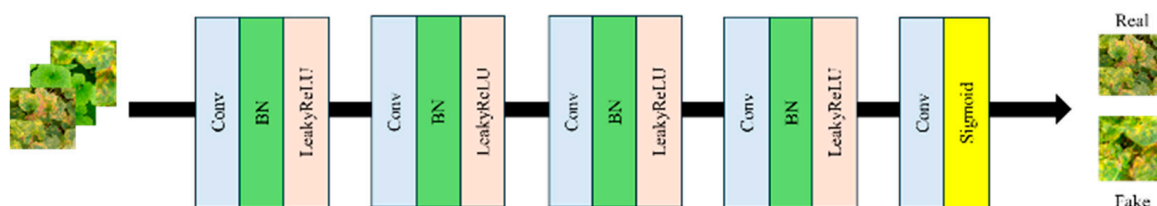


Figure 8. Detailed structure of the discriminator network used in the proposed model.

Activation Functions

The choice of activation function in the generator significantly influences the model's training efficiency and the quality of generated images. While ReLU is commonly used in DCGANs, this study compares three activation functions, namely ReLU, ELU, and Mish, to evaluate their impact on image generation.

ReLU is one of the most widely used activation functions, which sets negative values to zero while keeping positive values unchanged. It is computationally efficient and helps mitigate the vanishing gradient problem:

$$\text{ReLU}(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \quad (1)$$

ELU allows small negative outputs, unlike ReLU, maintaining a non-zero gradient for negative values. This property helps improve gradient flow and stabilizes training. Additionally, ELU's outputs are centered around zero, which can enhance learning efficiency:

$$\text{ELU}(x) = \begin{cases} a(e^x - 1), & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases} \quad (2)$$

Mish is a recently introduced activation function that is smooth and non-saturating, allowing for better gradient propagation. Compared to ReLU and ELU, Mish provides a more continuous activation curve, which can improve overall training dynamics and enhance the quality of generated images:

$$\text{Mish}(x) = x \cdot \tanh(\text{softplus}(x)) = x \cdot \tanh(\ln(1 + e^x)) \quad (3)$$

Three types of modified DCGAN were developed using a combination of residual connections with activation functions as follows:

- (a) DCGAN + Residual Connection (ReLU).
- (b) DCGAN + Residual Connection (ELU).
- (c) DCGAN + Residual Connection (Mish).

2.4. DL CNN Classification Models

To validate the performance of the proposed approach, we conducted a set of experiments using a real dataset combined with various augmented datasets. The DL CNN leaf disease classification models used are AlexNet, VGGNet, and GoogLeNet.

AlexNet was introduced in 2012 at the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) and marked a breakthrough in deep learning-based computer vision. The model consists of a total of eight layers: five convolutional layers and three fully connected layers. AlexNet uses data augmentation techniques like cropping and horizontal flipping to improve generalization, along with dropout as a regularization method to reduce overfitting. By leveraging powerful GPUs, AlexNet accelerates training and achieves significantly lower error rates than previous models, drawing widespread attention.

Developed by the Visual Geometry Group (VGG) in 2014, VGGNet was designed to explore the relationship between convolutional neural network (CNN) depth and performance [26]. VGGNet comes in 16-layer (VGG16) and 19-layer (VGG19) versions, and consistently uses small 3×3 filters throughout its layers. By deepening the network while maintaining a manageable number of parameters, VGGNet achieves substantial improvements in accuracy and generalization. Its deep structure allows it to capture fine image details effectively, and it has since become a foundational architecture in CNN research.

Introduced in the 2014 ILSVRC, GoogLeNet is based on the Inception architecture, which utilizes an Inception module to design an efficient network structure [27]. The Inception module combines filters of different sizes, including 1×1 , 3×3 , and 5×5 , within the same layer, allowing the network to capture features at various receptive fields simultaneously. Although it is a 22-layer network, GoogLeNet has a relatively small number of parameters, optimizing both computation and memory usage. This efficiency made GoogLeNet known for its improved speed and performance.

2.5. Evaluation Metric

2.5.1. DCGAN Evaluation Metric

Quantitative Metrics

The performance of a GAN is primarily determined by the quality and diversity of the images it generates. In general, the closer the generated images are to real images in terms of visual fidelity and variation, the better the model's performance. Generative performance of the GAN can be evaluated using two quantitative metrics: inception score (IS) and Fréchet inception distance (FID).

The inception score (IS), proposed by [28], combines two key criteria, fidelity and diversity, into a single metric. The IS captures the desirable properties of generated samples by measuring the average Kullback–Leibler (KL) divergence between the conditional label distribution and the marginal label distribution of generated images. A higher IS indicates a better model with sharper and more diverse samples.

The formula for IS is as follows:

$$IS = \exp\left(E_{x \sim p_g} D_{KL}(p(y|x) \| p(y))\right) \quad (4)$$

The $p(y|x)$ is the predicted label distribution for a generated image, and $p(y)$ is the overall distribution of labels across the dataset. The KL divergence measures the difference between the classifier's predictions for individual images and the overall distribution, promoting models that produce images with high label confidence and a broad variety of generated samples.

FID is a method for embedding generated samples in the feature space [29]. The mean and covariance between the generated samples and the real data are computed on the assumption that the generated samples follow a multivariate Gaussian distribution. The FID between the two Gaussian distributions is then computed to assess the quality of the generated samples. A lower FID score indicates a smaller discrepancy between the synthetic and real data distributions.

The formula for FID is as follows:

$$FID(r, g) = \|\mu_r - \mu_g\|_2^2 + \text{Tr}\left(\sum_r + \sum_g - 2\left(\sum_r \sum_g\right)^{\frac{1}{2}}\right) \quad (5)$$

The $(\mu_r, \sum_r)$ represent the mean and covariance of the real data distribution, while $(\mu_g, \sum_g)$ correspond to those of the generated data distribution.

Qualitative Metric

To qualitatively assess the distribution of the generated images, t-distributed stochastic neighbor embedding (t-SNE) is employed as a visualization technique [30]. t-SNE is a non-linear dimensionality reduction algorithm that projects high-dimensional data into a lower-dimensional space, typically two or three dimensions, while preserving the local relationships among data points. This allows for an intuitive visual representation of how closely the generated images resemble real samples and whether the synthetic data

effectively captures the variability present in real plant disease images. In this study, t-SNE is applied to the feature representations of real and augmented images. They are embedded into a shared two-dimensional space, enabling a qualitative comparison of their distributions.

First, the t-SNE algorithm begins by computing the pairwise similarities among all data points in the high-dimensional space. This is performed using a Gaussian kernel with a bandwidth σ . The similarity between two data points x_i and x_j is determined with the following formula:

$$p_{j|i} = \frac{e^{\frac{-\|x_i - x_j\|^2}{2\sigma_i^2}}}{\sum_{k \neq i}^N e^{\frac{-\|x_i - x_k\|^2}{2\sigma_i^2}}} \quad (6)$$

where $p_{j|i}$ denotes the conditional probability of selecting point i , assuming point j is the center of a Gaussian distribution in the high-dimensional space. The denominator ensures normalization so that the sum of conditional probabilities over all possible neighbors equals 1.

In the lower-dimensional space, the probability distribution is defined using a Student's t-distribution with one degree of freedom instead of a Gaussian distribution. This is performed to address the crowding problem, which occurs when mid-range distances in the high-dimensional space are overly compressed in the low-dimensional space. The similarity between embedded data points y_i and y_j in the lower-dimensional space is computed as follows:

$$q_{j|i} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{k \neq i}^N \left(1 + \|y_i - y_k\|^2\right)^{-1}} \quad (7)$$

where y_i and y_j are the corresponding embeddings of the high-dimensional data points in the lower-dimensional space. The Student's t-distribution ensures that mid-range distances are expanded, improving cluster separation. The distribution Q is normalized so that the sum of all probabilities equals 1.

The objective of t-SNE is to minimize the difference between the high-dimensional probability distribution P and the low-dimensional probability distribution Q . This difference is quantified using the KL divergence, which measures how much information is lost when Q is used to approximate P :

$$D_{KL}(P\|Q) = \sum_{j=1}^N P(x) \log\left(\frac{P(x)}{Q(x)}\right) \quad (8)$$

Since KL divergence is non-symmetric, it penalizes points that are close in high-dimensional space but far apart in low-dimensional space more than the reverse. This property ensures that local structures are preserved in the embedding.

To minimize the KL divergence, t-SNE employs stochastic gradient descent (SGD) to iteratively update the positions of points in the low-dimensional space. The gradient of the KL divergence with respect to each low-dimensional point y_i is given by the following:

$$y_i(t+1) = y_i(t) + \eta \sum_{j=1}^N (p_{ij} - q_{ij}) (y_i - y_j) \left(1 + \|y_i - y_j\|^2\right)^{-1} \quad (9)$$

where η is the learning rate. t represents the iteration number. The attractive force $(p_{ij} - q_{ij})$ moves similar points closer together, while the repulsive force expands dissimilar points. The term $\left(1 + \|y_i - y_j\|^2\right)^{-1}$ ensures that larger distances contribute

less to the gradient, helping stabilize the optimization process. This gradient descent step is repeated until the KL divergence reaches a minimum, resulting in an optimal low-dimensional representation.

This qualitative evaluation helps visualize the extent to which the modified DCGAN improves over the conventional DCGAN augmentation. Additionally, it serves as a complementary evaluation to quantitative metrics such as IS and FID, providing deeper insights into the model's ability to synthesize realistic plant disease images.

2.5.2. Classification Model Evaluation Matrix

The performance of the classification models was evaluated using four standard metrics: accuracy, precision, recall, and F1-score. Each row corresponds to an actual class, while each column represents a predicted class. Accuracy, precision, recall, and F1-score are defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (13)$$

For multi-class classification involving three categories, a confusion matrix was constructed to evaluate model performance. For each class, true positives (TP) represent correctly identified samples of that class, false positives (FP) denote samples from other classes misclassified as that class, false negatives (FN) refer to samples of that class misclassified into other categories, and true negatives (TN) correspond to all remaining correctly classified samples not belonging to that class.

3. Results

The results section compares the base DCGAN model with the modified DCGAN architecture. The performance evaluation will be based on quantitative and qualitative metrics to deduce the best modified DCGAN architecture among them all.

After finding the best modified DCGAN architecture, the output images from the dataset augmentation block of the traditional methods, DCGAN, and the modified DCGAN are inputted into various DL CNN disease classification models for differentiation of healthy, downy mildew, and powdery mildew. The performance of the classification models based on the various augmentations was evaluated and compared.

3.1. Quantitative Performance Comparison Between DCGAN and Modified DCGAN

To evaluate the impact of architectural modifications on generative models, we compared the conventional DCGAN with the modified DCGAN. Each model's performance was quantified using IS and FID. These metrics comprehensively evaluate image clarity, similarity, and representational completeness.

As summarized in Table 1, the IS results demonstrated considerable differences among model configurations for generating different disease images. The DCGAN achieved IS values of 1.56 for downy mildew, 1.77 for healthy leaves, and 2.31 for powdery mildew, whereas with residual connections and ReLU being added, the IS values improved to 1.60, 1.76, and 2.36, respectively, showing improvement. Further improvement was observed in the DCGAN model when residual connections and ELU were applied, with the highest gain observed in the powdery mildew. Finally, the model that combined residual connec-

tions and the Mish function showed the highest IS values for all classes, suggesting an improvement in the % value for classes with respect to conventional DCGAN. These results indicate that models with residual connections and Mish improved the clarity and diversity of different disease classes of generative images.

Table 1. Quantitative evaluation results for each GAN configuration based on inception score (IS) and Fréchet inception distance (FID) for three classes.

Model	Classes	IS	FID
DCGAN	Downy mildew	1.56	204.04
	Healthy	1.77	224.94
	Powdery mildew	2.31	169.98
DCGAN + Residual Connection (ReLU)	Downy mildew	1.60	196.05
	Healthy	1.76	231.13
	Powdery mildew	2.36	166.50
DCGAN + Residual Connection (ELU)	Downy mildew	1.56	191.23
	Healthy	1.83	230.24
	Powdery mildew	2.46	150.50
DCGAN + Residual Connection (Mish)	Downy mildew	1.74	187.99
	Healthy	1.91	208.60
	Powdery mildew	2.57	158.69

The FID results also supported the IS value, suggesting that the best performance of the DCGAN model is with the residual connection and Mish function. The DCGAN showed relatively high FID values of 204.04 for downy mildew, 224.94 for healthy, and 169.98 for powdery mildew, suggesting less similarity to real images. In comparison, the DCGAN model with residual connection and Mish achieved lower FID values by 7.87% for downy mildew, 7.26% for healthy leaves, and 6.64% for powdery mildew, respectively, suggesting higher resemblance with original disease classes for generative images.

Additionally, as per the IS value representing image similarity and diversity available in the generative images of different diseases, for the downy mildew class, a generative image model can be ranked as DCGAN + residual connection (Mish), DCGAN + residual connection (ReLU), and DCGAN + residual connection (ELU), whereas for healthy and powdery mildew classes, the generative image model can be ranked as DCGAN + residual connection (Mish), DCGAN + residual connection (ELU), and DCGAN + residual connection (ReLU).

Similarly, as per the FID value representing image distribution, for downy mildew and healthy classes, the generative image model can be ranked as DCGAN + residual connection (Mish), DCGAN + residual connection (ELU), and DCGAN + residual connection (ReLU); however, for powdery mildew, DCGAN + residual connection (ELU) performed best.

To complement the quantitative analysis, Figure 9 presents visual examples of generative leaf images generated by the DCGAN and modified DCGAN for three disease categories. The DCGAN with residual connection and Mish function visually produced samples with greater textural variation and morphological detail for different classes of disease, including downy mildew, healthy, and powdery mildew. These visual patterns generated for the different diseases resemble real images and are consistent with the numerical results, reinforcing the combination of residual learning and the Mish activation function.

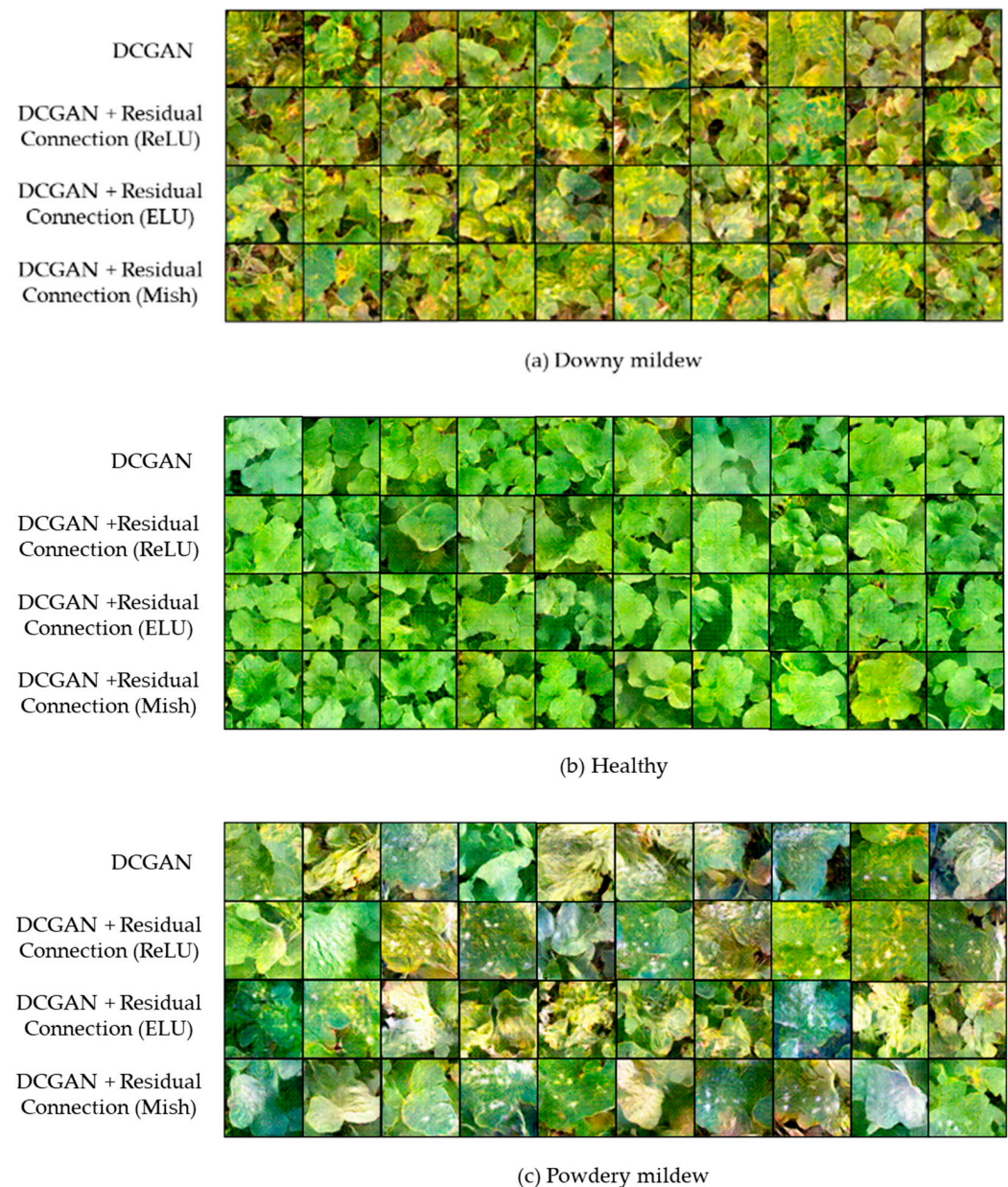


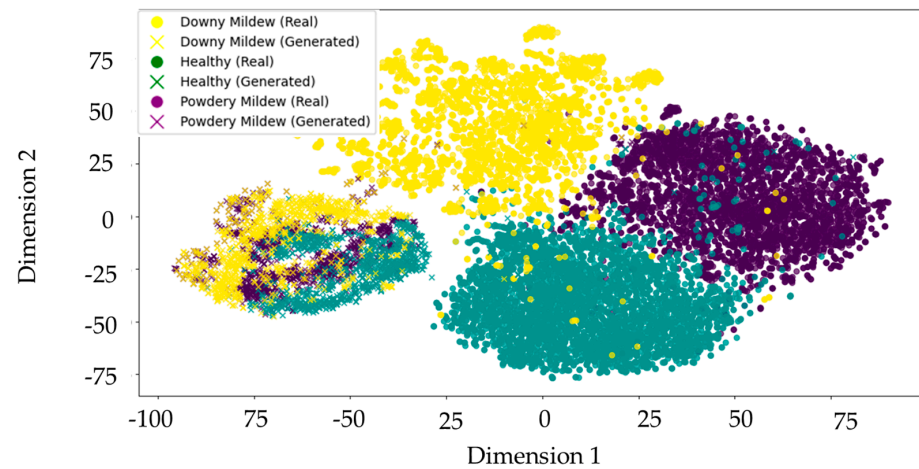
Figure 9. Image samples generated by various DCGAN models: (a) Downy mildew; (b) healthy; (c) and powdery mildew.

In summary, the IS value, FID, and visual representations of generative images for different classes of healthy and disease images suggested architecture enhancements that led to improvement. The model incorporating residual connections and the Mish activation function consistently outperformed DCGAN and all other modified DCGANs. Therefore, the disease classification model that includes the DCGAN with the residual connection and Mish activation function is chosen as the modified DCGAN augmentation.

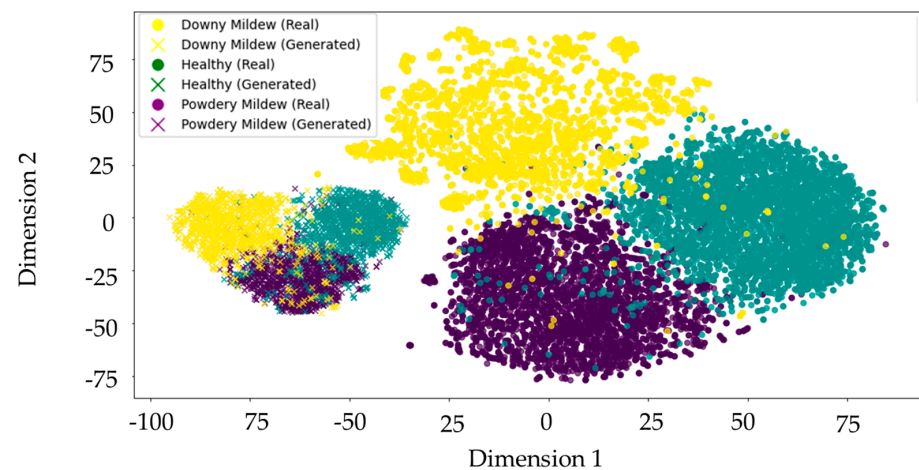
3.2. Qualitative Performance Comparison Between Traditional, DCGAN, and Modified DCGAN Augmentation

To evaluate how different data augmentation methods influence feature representation and variation of pattern representation with real images, we also conducted t-SNE visualization for each method. The three visualizations that correspond to traditional augmentation, DCGAN, and modified DCGAN (DCGAN with residual connection and Mish activation function) are shown in Figure 10. Each Figure illustrates the distribution

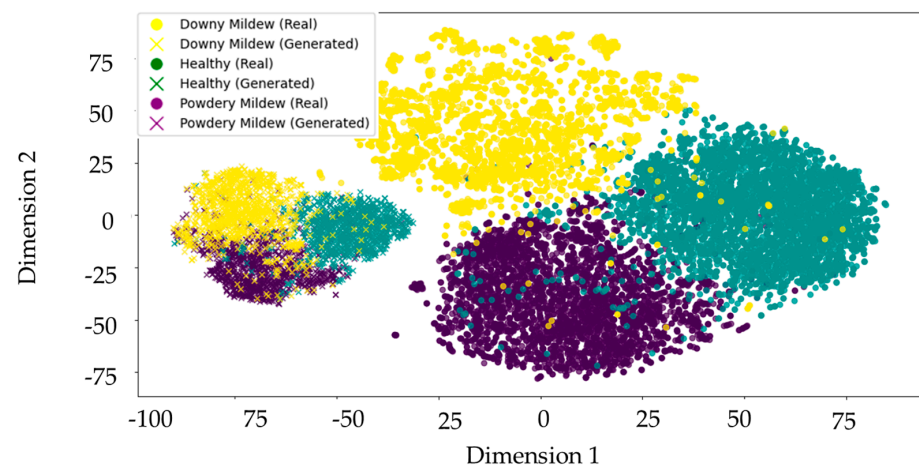
of real and generated images in the feature space for different augmentation methods, with yellow representing downy mildew, purple representing powdery mildew, and green representing healthy leaves. Real images are denoted by circles and generative images by crosses.



(a) Traditional augmentation



(b) DCGAN



(c) Modified DCGAN

Figure 10. t-SNE visualization results of different GAN models: (a) Traditional augmentation; (b) DCGAN augmentation; and (c) modified DCGAN.

A consistent pattern was observed across all three t-SNE visualizations for the generative and real images; they did not form overlapping clusters in feature space. The observed separation is due to image resolution, since t-SNE is sensitive to differences in input resolution. The resolution of the real images that were captured was very high, approximately from 1000×1000 to 1300×1300 pixels, whereas the generative image resolution was small, with 64×64 pixels, due to the higher computation requirement for generating high-resolution images. Furthermore, for traditional augmentation, the image size chosen was resized to 224×224 pixels, as per the classification model standard. Therefore, while analyzing distributional and representativeness, we focused our analysis on the internal clustering patterns among the generative images' classes and real images' distribution in feature space.

In the traditional augmentation setting (Figure 10a), generative images and various classes of images, including downy mildew, healthy leaves, and powdery mildew, seem to overlap in feature space with unknown distinction compared to the real images. Particularly, the generative image for the powdery mildew class showed significant overlap with the healthy class. This indicates that traditional augmentation was not sufficient to preserve class-specific characteristics and different feature spaces between the overlapping disease and healthy classes. The resulting clusters were less distinct, and class boundaries were poorly defined.

Although the overall distribution patterns of DCGAN and modified DCGAN were similar, the modified DCGAN produced slightly more dispersed clusters while reducing class interference. As shown in Figure 10b,c, the feature spaces of the DCGAN and modified DCGAN exhibit clear and distinct separation among the three classes.

In summary, GAN-based augmentation methods outperformed traditional augmentation by generating images that resembled real image disease-specific feature distributions. While both DCGAN and modified DCGAN improved the structure of generative images and the distinction between class features, modified DCGAN offered additional advantages by enhancing intra-class diversity, which can lead to improved generalization in classification models.

3.3. Disease Classification Model Performance on Augmented Datasets

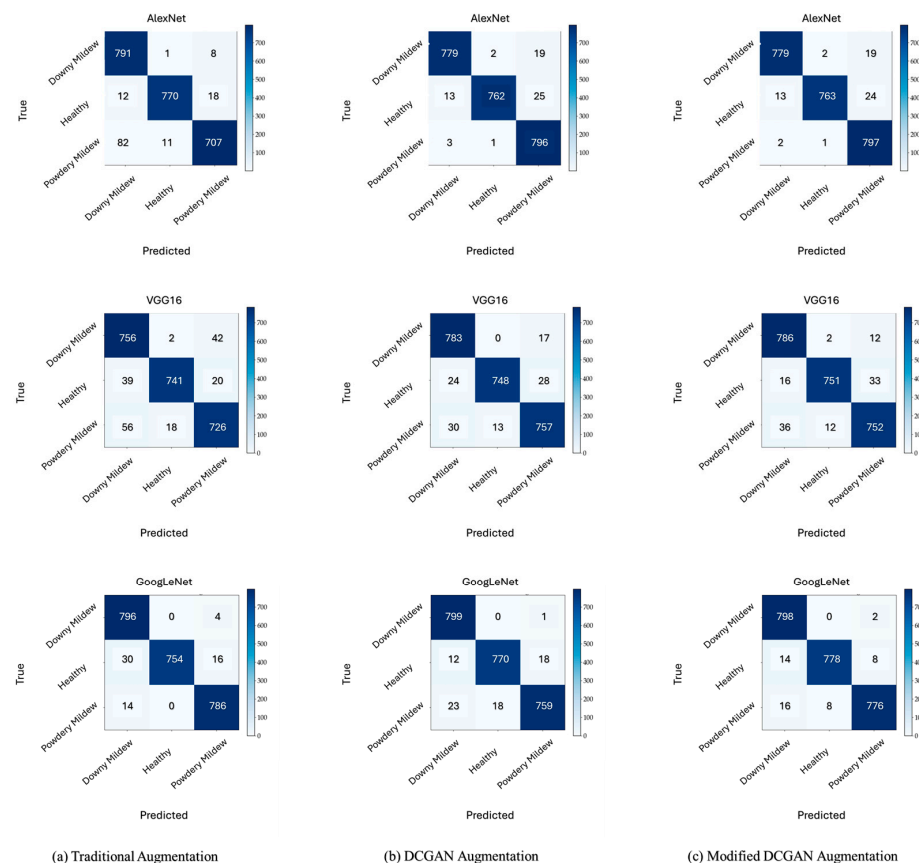
This section compares the effects of three data augmentation strategies on the classification performance of AlexNet, VGG16, and GoogLeNet. The augmentation methods applied include traditional augmentation, DCGAN, and modified DCGAN (DCGAN with residual connection and Mish function). Evaluation metrics used for the DL CNN classification model in this comparison include precision, recall, F1-score, and accuracy. To further validate the results, confusion matrices were analyzed for each combination of classification model and augmentation method.

Evaluation results were shown in Table 2 for AlexNet, VGG16, and GoogLeNet classification models; performance increased progressively when image augmentation changed from traditional augmentation to DCGAN, and it increased further when changed to modified DCGAN. For the AlexNet model, the modified DCGAN-based augmentation achieved the best overall performance, surpassing all other augmentation methods with a precision, recall, F1-score, and accuracy value of 0.975. For the VGG16 model, augmentation methods DCGAN and modified DCGAN have similar performance with an overall precision, recall, F1-score, and accuracy value of 0.954. However, the performance of DCGAN and modified DCGAN was significantly greater than traditional augmentation. Finally, for the GoogLeNet, the modified DCGAN outperformed all the augmentation methods by a slight increment of overall evaluation metrics over DCGAN and traditional augmentation methods.

Table 2. Classification performance metrics (precision, recall, F1-score, accuracy) for AlexNet, VGG16, and GoogLeNet under different augmentation methods.

Model	Aug.	Precision	Recall	F1-Score	Accuracy
AlexNet	Traditional augmentation	0.948	0.945	0.945	0.945
	DCGAN	0.975	0.974	0.974	0.974
	Modified DCGAN	0.975	0.975	0.975	0.975
VGG16	Traditional augmentation	0.928	0.926	0.926	0.930
	DCGAN	0.954	0.954	0.954	0.954
	Modified DCGAN	0.954	0.954	0.954	0.954
GoogLeNet	Traditional augmentation	0.973	0.973	0.973	0.973
	DCGAN	0.970	0.970	0.970	0.970
	Modified DCGAN	0.980	0.980	0.980	0.980

The classification model performance was further analyzed using a confusion matrix to determine class-wise prediction and misclassification patterns. The test image sample for each class was 800. According to Figure 11a, for the AlexNet model with different augmentation methods, the traditional augmentation performed well for the downy mildew and healthy classes; however, it showed the poorest performance for the powdery mildew class. In contrast, both the DCGAN and modified DCGAN methods achieved superior classification for powdery mildew, with fewer than three misclassifications. Furthermore, for Figure 11b, for the VGG16 model, modified DCGAN augmentation showed the best class-wise classification performance, with minimum misclassification among all augmentations. Furthermore, for Figure 11c, for the GoogLeNet model, the modified DCGAN overall performance for downy mildew and healthy class was the best; however, for powdery mildew, it slightly lags behind traditional augmentation.

**Figure 11.** Confusion matrix for different classification models with various augmentations.

When comparing the modified DCGAN augmentation's impact on the classification performance across different diseases and health classes, the GoogLeNet classifier demonstrated the best performance for the downy mildew and healthy classes; however, it showed relatively poor performance for the powdery mildew class.

These findings indicate that GAN-based data augmentation contributes meaningfully to the improvement of classification performance. In particular, modified DCGAN enhanced both the quality and diversity of synthetic images, leading to improved generalization overall for DL CNN-based classification models. The improvements observed for evaluation metrics and confusion matrix analysis validate the effectiveness of DCGAN and modified DCGAN generative augmentation approaches.

4. Discussion

The study evaluated different data augmentation strategies, traditional augmentation, DCGAN, and modified DCGAN, and their effects on generating images of healthy and diseased leaves. The two-phase analysis approach involved first selecting the best modified DCGAN augmentation model, and second, analyzing and comparing the improvement brought by different types of augmentation on various class prediction models. The result confirms the best generative augmentation method as the modified DCGAN with residual connection and Mish activation function with the highest IS (1.74 for downy mildew, 1.91 for healthy leaves, and 2.57 for powdery mildew) and decreasing FID value (187.99 for downy mildew, 208.6 for healthy leaves, and 158.69 for powdery mildew). This is due to the inclusion of the residual connection placed in the generative part of the modified DCGAN; the vanishing gradient loss was removed, allowing the model to generate different class images; and, with the Mish activation function, its self-regularized non-monotonic characteristics allowed for enhanced gradient propagation and expressive power, inducing greater resemblance and diversity in producing the generative image similar to real image classes.

Furthermore, t-SNE distribution for the modified DCGAN with residual connection and Mish function showed higher accuracy in representing disease-specific features and inter-class differences in feature space, confirming class difference between disease and healthy generative images.

Regarding the classification models under different augmentation scenarios, the classification model performance was compared with precision, recall, F1-score, and accuracy. The modified DCGAN showed an overall better performance for every classification model, where GoogLeNet had the highest classification results for all the augmentation methods, and is, therefore, the best classification model among them all. For augmentation performance comparison on the basis of classification, DCGAN and modified DCGAN perform almost the same, with a slight increment in the performance metrics for modified DCGAN. However, traditional augmentation has less classification performance. The GoogLeNet performed best, and was supported by its fast and efficient training time and moderate parameter size [31]. However, for AlexNet, although the training time was fast, it had a small parameter size, whereas for VGG16, it had a slow training time and a large parameter size, requiring higher computation time and memory [32,33].

Similarly, the results can be interpreted from the confusion matrix, showing that GoogLeNet performs the best, overall, for disease and healthy classification but struggles with powdery mildew across all augmentation methods. The VGG16 model, however, performs the worst across all classes and augmentation techniques, likely due to its larger parameter size and higher training data requirements, which may hinder effective feature learning in this context. Future work could extend this approach to various plant species,

disease types, and growth stages to evaluate its broader applicability and robustness across diverse agricultural conditions.

The study demonstrated that structure-aware GAN-based data augmentation can effectively improve the generation of different plant disease classes using real-world data. The findings highlight the potential of this approach to support accurate disease detection in agricultural applications. The key contribution of this study lies not in the simple application of existing GAN frameworks, but in the theoretical and architectural enhancement of the DCGAN structure through the integration of residual connections and the Mish activation function. This combination enables a more stable training process and richer feature representation that is specifically optimized for plant disease imagery. Moreover, the effectiveness of the proposed model allows for the overall improvement of the classification model performance, suggesting practical applicability.

5. Conclusions

This study evaluated the effectiveness of different data augmentation strategies, including traditional augmentation, DCGAN, and a modified DCGAN, for leaf disease classification. In particular, various modified DCGAN architectures incorporating residual connections and different activation functions were analyzed. Among these, the DCGAN model with residual connections and the Mish activation function demonstrated the best overall performance, showing an improvement in IS from 7.9 to 11.54% and a reduction in FID from 6.6 to 7.8% compared to the conventional DCGAN.

For the classification task, GoogLeNet combined with the modified DCGAN augmentation achieved the highest performance across all models, with average precision, recall, and F1-score values of approximately 98%. These results indicate that generative model-based data augmentation effectively mitigates the challenges associated with limited datasets in plant disease classification. By generating balanced and diverse generative images, the approach enhances the robustness and generalization ability of deep learning classifiers.

Furthermore, the findings emphasize the critical role of activation function selection in optimizing generative model performance, in comparison to the traditional augmentation methods, with Mish-based modified DCGAN augmentation performing best. Future research should focus on exploring alternative GAN architectures for better generative images and understanding the impact of different classification algorithm performance differences on different classes of disease to allow for a generalized classification model. Finally, integrating these augmentation methods into real-world agricultural systems and validating their effectiveness under practical field conditions will be essential to assess their applicability.

Overall, this study demonstrates the potential for application of the modified DCGAN for generative models to improve precision agriculture applications, particularly by properly detecting and classifying different plant diseases through enhanced data augmentation and deep learning strategies.

Author Contributions: Conceptualization, M.K.; Methodology, M.K.; Software, M.K.; Validation, N.T.; Formal analysis, M.K.; Resources, H.T.K.; Writing—original draft, M.K.; Writing—review & editing, N.T. and H.T.K.; Visualization, M.K.; Project administration, H.T.K.; Funding acquisition, H.T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Glocal University 30 Project Fund of Gyeongsang National University in 2025.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

- Gondi, S.P.; Ganiger, V.M.; Bhuvaneshwari, G.; Madalageri, M.B.; Kotikal, Y.K.; Manjunatha, G. Evaluation of Oriental Pickling Melon (*Cucumis melo* L. Var. Conomon) Genotypes for Pest and Disease Reaction under Northern Dry Zone of Karnataka. *J. Environ. Ecol.* **2016**, *7*, 29. [\[CrossRef\]](#)
- Aegerter, B.J.; Nuñez, J.J.; Davis, R.M. Environmental Factors Affecting Rose Downy Mildew and Development of a Forecasting Model for a Nursery Production System. *Plant Dis.* **2003**, *87*, 732–738. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pertot, I.; Kuflik, T.; Gordon, I.; Freeman, S.; Elad, Y. Identifier: A Web-Based Tool for Visual Plant Disease Identification, a Proof of Concept with a Case Study on Strawberry. *Comput. Electron. Agric.* **2012**, *84*, 144–154. [\[CrossRef\]](#)
- Mutka, A.M.; Bart, R.S. Image-Based Phenotyping of Plant Disease Symptoms. *Front. Plant Sci.* **2015**, *5*, 734. [\[CrossRef\]](#)
- Shoaib, M.; Shah, B.; El-Sappagh, S.; Ali, A.; Ullah, A.; Alenezi, F.; Gechev, T.; Hussain, T.; Ali, F. An Advanced Deep Learning Models-Based Plant Disease Detection: A Review of Recent Research. *Front. Plant Sci.* **2023**, *14*, 1158933. [\[CrossRef\]](#)
- Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant Disease Detection and Classification by Deep Learning. *Plants* **2019**, *8*, 468. [\[CrossRef\]](#)
- Miftahushudur, T.; Sahin, H.M.; Grieve, B.; Yin, H. A Survey of Methods for Addressing Imbalance Data Problems in Agriculture Applications. *Remote Sens.* **2025**, *17*, 454. [\[CrossRef\]](#)
- Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* **2021**, *8*, 53. [\[CrossRef\]](#) [\[PubMed\]](#)
- Khare, O.; Mane, S.; Kulkarni, H.; Barve, N. LeafNST: An Improved Data Augmentation Method for Classification of Plant Disease Using Object-Based Neural Style Transfer. *Discov. Artif. Intell.* **2024**, *4*, 50. [\[CrossRef\]](#)
- Antwi, K.; Bennin, K.E.; Pobi Asiedu, D.K.; Tekinerdogan, B. On the Application of Image Augmentation for Plant Disease Detection: A Systematic Literature Review. *Smart Agric. Technol.* **2024**, *9*, 100590. [\[CrossRef\]](#)
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *Commun. ACM* **2020**, *63*, 139–144. [\[CrossRef\]](#)
- Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. Generative Adversarial Networks (GANs) for Image Augmentation in Agriculture: A Systematic Review. *Comput. Electron. Agric.* **2022**, *200*, 107208. [\[CrossRef\]](#)
- Tian, Y.; Li, E.; Liang, Z.; Tan, M.; He, X. Diagnosis of Typical Apple Diseases: A Deep Learning Method Based on Multi-Scale Dense Classification Network. *Front. Plant Sci.* **2021**, *12*, 698474. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cap, Q.H.; Uga, H.; Kagiwada, S.; Iyatomi, H. LeafGAN: An Effective Data Augmentation Method for Practical Plant Disease Diagnosis. *IEEE Trans. Autom. Sci. Eng.* **2022**, *19*, 1258–1267. [\[CrossRef\]](#)
- Liu, D.; Cao, Y.; Yang, J.; Wei, J.; Zhang, J.; Rao, C.; Wu, B.; Zhang, D. SM-CycleGAN: Crop Image Data Enhancement Method Based on Self-Attention Mechanism CycleGAN. *Sci. Rep.* **2024**, *14*, 9277. [\[CrossRef\]](#)
- Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv* **2015**, arXiv:1511.06434.
- Bi, L.; Hu, G. Improving Image-Based Plant Disease Classification with Generative Adversarial Network Under Limited Training Set. *Front. Plant Sci.* **2020**, *11*, 583438. [\[CrossRef\]](#)
- Wu, Q.; Chen, Y.; Meng, J. DCGAN-Based Data Augmentation for Tomato Leaf Disease Identification. *IEEE Access* **2020**, *8*, 98716–98728. [\[CrossRef\]](#)
- Wang, Z.; She, Q.; Ward, T.E. Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy. *ACM Comput. Surv.* **2021**, *54*, 37. [\[CrossRef\]](#)
- Zhang, H.; Xu, T.; Li, H.; Zhang, S.; Wang, X.; Huang, X.; Metaxas, D. StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 5908–5916. [\[CrossRef\]](#)
- Brock, A.; Donahue, J.; Simonyan, K. Large Scale GAN Training for High Fidelity Natural Image Synthesis. In Proceedings of the 7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, 6–9 May 2019. [\[CrossRef\]](#)
- Nair, V.; Hinton, G.E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
- Maas, A.; Hannun, A.; Ng, A.Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013.
- Clevert, D.A.; Unterthiner, T.; Hochreiter, S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). In Proceedings of the 4th International Conference on Learning Representations, ICLR 2016—Conference Track Proceedings, San Juan, Puerto Rico, 2–4 May 2016.
- Misra, D. Mish: A Self Regularized Non-Monotonic Activation Function. *arXiv* **2019**, arXiv:1908.08681.
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*. [\[CrossRef\]](#)

27. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015.
28. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9. [[CrossRef](#)]
29. Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X.; Chen, X. Improved Techniques for Training GANs. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2016; Volume 29.
30. Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. *arXiv* **2017**, arXiv:1706.08500.
31. van der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
32. Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers. *Plants* **2020**, *9*, 1319. [[CrossRef](#)] [[PubMed](#)]
33. Tariq, M.; Ali, U.; Abbas, S.; Hassan, S.; Naqvi, R.A.; Khan, M.A.; Jeong, D. Corn Leaf Disease: Insightful Diagnosis Using VGG16 Empowered by Explainable AI. *Front. Plant Sci.* **2024**, *15*, 1402835. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.