A Comparative Review on GAN-Based Data Augmentation Techniques for Plant-Based Pest Detection

Dipanwita Chakraborty Bhattacharya, Md Tausif Mallick, Himadri Nath Saha and Amlan Chakrabarti

January 10, 2025

A comparative review on GAN-based data augmentation techniques for plant-based pest detection

Dipanwita Chakraborty Bhattacharya[1*], MD Tausif Mallick[2], Himadri Nath Saha[3], Amlan Chakrabarti[2]

dcb.wbes@gmail.com[1],tausif.realmadrid@gmail.com[2],contactathimadri@gmail.com[3*],acakcs@caluniv.ac.in[2]

[1]Department of Computer Science, Bethune College, University of Calcutta, Kolkata, India
[2]AKCSIT, University of Calcutta, Kolkata, India
[3]Department of Computer Science, SNEC, University of Calcutta, Kolkata, India

Abstract—The success of deep learning methods has led to the development of several applications for the automated identification of plant diseases and pest attacks. Nevertheless, these programs frequently experience overfitting, and when applied to test datasets from unfamiliar contexts, the diagnostic performance is significantly reduced. In this work, we present traditional CycleGAN, a unique image-to-image translation system with an attention mechanism of its own. CycleGAN is a data augmentation technology that improves the effectiveness of plant pest diagnosis by transforming a limited number of pest-damaged images into a broad variety of pest-lesioned images. In our work, CycleGAN outperformed other models in generating synthetic images of Sawfly pests. On the other hand, the Copy-Paste-Blend (CPB) approach has proven effective in seamlessly embedding pest masks into external leaf images. This technique blends pest masks at various scales with leaf backgrounds, resulting in synthetic images that appear more natural and realistic.

Index Terms—image-to-image translation, plant pest diagnosis, data augmentation, generative adversarial network, CycleGAN, Copy-Paste-Blend (CPB).

1. Introduction :

According to global development data, building healthy, sustainable, and inclusive food systems is crucial for reaching development goals worldwide. Agricultural growth is key to reducing poverty, securing food supplies, and feeding an anticipated population of 10 billion by 2050. Yet, various challenges threaten this progress, including disruptions from COVID-19, climate change, extreme weather, pests, conflicts, and surging food prices. These issues are worsening food insecurity and poverty, derailing efforts to end global hunger by 2030, and reversing key development achievements. According to the Food and Agriculture Organization (FAO), 20–40% of global crop yields are lost annually due to pests. Plant diseases alone cause $220 billion in economic losses each year, while invasive insects account for an additional $70 billion in damages worldwide. Early and accurate pest detection can prevent crop loss and improve agricultural productivity.

Plant pest surveillance is vital for the agricultural and biosecurity systems in Australia and New Zealand. It involves systematically checking for plant pests, which include harmful invertebrates and pathogens that can significantly impact food production, security, and ecosystems, leading to annual losses of 20–40%. Surveillance activities provide several key benefits: Early detection; identifying pests before they establish allows for effective eradication or management strategies; market access; surveillance data supports negotiations for domestic and international market access, with pest-free evidence reassuring importing countries; extent of incursions; delimiting surveillance helps determine the spread of newly introduced pests; monitoring levels; ongoing monitoring of existing pests is crucial for management and control. Deep learning plays a vital role in the early detection of plant pests. Deep learning contributes in several ways to early pest detection according to [13] and [14]: image recognition, data analysis, predictive modeling, automated monitoring, sensor data, integration with IoT, and many more. High-performance deep learning models, such as deep neural networks, require large amounts of high-quality data to achieve accuracy and robustness. Here, deep learning models like pest (object) detection models face challenges due to the scarcity of data in the agricultural pest dataset. These algorithms demand significant computing resources for model training and inference, especially when dealing with complex environments or numerous object classes. Accurate object detection requires detailed annotations (e.g., bounding boxes, class labels), which can be labor-intensive and expensive to produce, particularly for specific pest species or plant diseases.

In this regard, data augmentation is a vital tool to handle the data scarcity. According to [12], Data augmentation involves artificially creating new data from existing data to improve machine learning model training. Traditional data augmentation techniques, like cropping, flipping, and scaling, introduce only limited variation to datasets. Overfitting remains a significant issue, as these transformations don't introduce new object appearances. GAN-based techniques generate realistic synthetic images that increase both dataset size and diversity without requiring extensive manual labeling [3], [4], [5]. In this paper, a comparative review of different GAN-based augmentation techniques for pest detection has been delivered. This review will help to understand the performance of different GAN-based data augmentation models on different datasets of plant-based pest images in producing quality images.

2. Background and Related Work:

The Generative Adversarial Network (GAN), introduced by Ian Goodfellow et al. in 2014 [1] , is a type of deep learning framework that consists of two neural networks competing against each other in a game-like setting. Generator takes in random noise (typically a vector of latent variables) and generates synthetic data that mimics the real data distribution. The discriminator acts as a classifier, distinguishing between real data (from the actual dataset) and fake data (produced by the generator). The networks are trained simultaneously through a game-theoretic approach. The objective is for the generator to generate data so realistic that the discriminator can no longer reliably tell real from fake. This innovative architecture brought forth the concept of adversarial learning, which has since been broadly adopted and expanded across various domains, such as computer vision, medical imaging, and agriculture.

In the field of data augmentation, numerous GAN architectures have emerged, each with its own unique strengths and characteristics. In this list, the first name that comes is Deep Convolutional Generative Adversarial Networks (DCGANs) [2], a type of generative model that leverages deep convolutional neural networks (CNNs) to generate high-quality images. Introduced by Alec Radford et al. in 2015, DCGANs improve upon traditional GANs by replacing fully connected layers with convolutional layers, enabling the model to capture spatial hierarchies in images. This architecture typically includes techniques such as batch normalization and the use of ReLU and Leaky ReLU activation functions to stabilize training and enhance the quality of generated samples. DCGANs have found applications in various fields, including image synthesis, data augmentation, and unsupervised representation learning, making them a foundational model in the generative modeling landscape. Karam et al. applied DCGAN with the CPB (Copy-Paste-Blend) method [3] to develop their semi-automated data augmentation tool designed to aid in the detection of agricultural pests, specifically whiteflies. Notably, the use of GANs to enhance object diversity resulted in improved recall and precision metrics for lightweight detection models, such as YOLO and PestNet. Human reviewers assessed the realism of the generated data with varying outcomes, indicating that the tool is particularly effective in low-resolution contexts. CycleGAN is employed to enhance dataset diversity by generating synthetic diseased apple images and transferring disease traits to healthy ones. YOLOV3-Dense, an improved version of YOLO-V3 with DenseNet integration, optimizes feature extraction and improves lesion detection accuracy, outperforming Faster R-CNN and the original YOLO-V3 in tests.

CycleGAN, introduced by Jun-Yan Zhu et al. at ICCV 2017 [4], enables unpaired image-to-image translation by using two GANs in a cycle-consistent framework. This framework guarantees that an image translated from domain A to domain B can be reverted back to domain A, maintaining the essential characteristics of the original. In areas such as medical imaging or artistic style transfer, it can create multiple variations of images from one domain and apply them to another, expanding datasets even when ground truth pairs are unavailable. This not only improves the robustness of machine learning models but also reduces the dependency on large, curated datasets. In the paper, Tian et al. focus on detecting apple lesions, particularly anthracnose, using deep learning methods CycleGAN and YOLOV3-Dense

WGAN, proposed by Martin Arjovsky et al. at ICML 2017 [5], addresses the instability in GAN training by using the Wasserstein distance instead of traditional divergence metrics. Wasserstein distance in the context of GAN, measures how different the generated data distribution is from the real data distribution, offering smoother gradients for more stable training compared to traditional distance metrics. An enhanced generative adversarial network (GAN) model, named AWGAN, proposed by Xin et al. [6] gives notable performance for the data augmentation of plant diseased leaf images. The goal is to address the limitations of small datasets in deep learning-based plant disease recognition, which can cause overfitting and reduced model accuracy. AWGAN uses the Wasserstein GAN loss function for stable training and incorporates a self-attention layer to improve feature extraction, enhancing the model's ability to generate realistic images. The method is tested on corn leaf disease images from the PlantVillage dataset, generating around 30,000 images to augment the dataset. The augmented data is subsequently used to train models such as AlexNet, VGG16, and ResNet18. AWGAN outperforms other GAN-based augmentation techniques, achieving the highest recognition accuracy, with a 1-2% improvement. This method significantly enhances the training data, leading to better model generalization, especially in cases with limited sample sizes.

The paper presents Leaf GAN [7], a data augmentation technique utilizing generative adversarial networks (GANs) to tackle the issue of insufficient training data for grape leaf disease identification. Collecting sufficient disease images is labor-intensive, leading to overfitting in CNN-based models [8]. Leaf GAN generates synthetic images for four categories: black rot, Esca measles, leaf spot, and healthy leaves, using a generator with degressive channels and a discriminator enhanced by dense connectivity and instance normalization for effective lesion feature extraction. The model's stability is ensured through a deep regret gradient penalty technique. Leaf GAN produced 8,124 high-quality images from 4,062 original samples, outperforming DCGAN and WGAN in terms of Fréchet Inception Distance (FID). When the augmented dataset was used to train classifiers, models like Xception achieved 98.7% accuracy, demonstrating improved identification performance over traditional methods. This approach offers a practical solution to the data scarcity problem in agricultural disease detection, enabling better generalization and accuracy in automated grape disease diagnosis.

In this regard, most of the data augmentation models have been operated on the dataset for plant disease detection. Antwi et al. performed the comparison in their paper [9]. The novelty of the paper lies in its comprehensive comparison of traditional image augmentation methods and generative adversarial networks (GANs) specifically for plant disease detection (PDD). While previous studies have separately reviewed various data augmentation techniques, this paper uniquely provides a systematic analysis of how GANs stack up against basic augmentation methods in improving the performance of deep learning models like convolutional neural networks (CNNs) in PDD. The paper identifies specific GAN types, Techniques like DCGAN have proven effective in improving dataset quality and model accuracy, offering valuable insights for PDD applications. Furthermore, it addresses key challenges in applying GAN-augmented data to real-world farming situations, such as dataset imbalance and the complexities of generating realistic synthetic images. Pest detection and identification are crucial in the sense that they are tiny, numerous, and large amounts of data collection through imaging devices is much more difficult.

- This paper aims to study different data augmentation models suitable for plant-based pest image generation and their performance.
- In a comparative review of GAN-based models on the Sawfly dataset, key benefits emerge, particularly in optimizing performance for small, low-resolution datasets.
- Evaluating different GAN architectures (e.g., DCGAN, CycleGAN, WGAN, Hybrid GAN) can identify models best suited for generating high-quality, synthetic images that enhance data diversity and reduce overfitting risks, a common issue with limited datasets.
- This review offers insights into the trade-offs between performance and computational requirements by evaluating model stability, realism, and resource efficiency, which are crucial for choosing the most suitable models for limited agricultural datasets.
- These findings not only inform optimal GAN selection for Sawfly data but also serve as a resource for similar small-scale, low-resolution agricultural applications.

The task of plant-based pest detection is basically a pipeline of data augmentation to enhance the dataset and an object detection model. Object detection models are computationally intensive and require large datasets. Since in the applications of the agricultural domain, more lightweight and computationally less intensive models are desirable, to operate them on edge devices in the IoT framework. In addition to the data augmentation to overcome data scarcity to work with small datasets, a discussion based on lightweight object detection models like YOLOv7, YOLOv7-tiny, PestNet, and ImageNet [3], [10], [11] helps to understand the overall architecture of pest detection methods.

Organization of GAN-based Techniques for Detection

GAN Model	Description	Strengths	Limitations	Applications
DCGAN (Deep Convolutional GAN)	Generates small pest objects (e.g., whiteflies) to paste onto leaf backgrounds.	Increases object variety, enhancing model accuracy through diverse synthetic objects.	Limited in generating high-resolution images, reducing fine details crucial for pest detection.	Suitable for datasets focusing on small pest types with shape or texture variations.
CycleGAN	Performs style translation to convert healthy leaves into pest-infested ones.	Adds environmental variability (lighting, background changes) to improve dataset diversity.	Struggles with small object transformations due to lack of attention mechanisms.	Helps in augmenting datasets with different environmental settings to improve model robustness.
AWGAN (Attention Wasserstein GAN)	Uses self-attention to enhance feature extraction and generate high-quality pest images.	Better spatial awareness with stable training using Wasserstein loss. Excels with varied backgrounds.	May require more computational resources due to self-attention and advanced loss function.	Effective for detecting pests in complex backgrounds and cluttered environments.
Hybrid GAN Techniques (e.g., CPB + GAN)	Combines human labeling with GAN-generated objects using copy-paste-blend (CPB).	Produces realistic datasets by mixing real and synthetic data, improving model generalization.	Dependent on the quality of both labeled data and GAN-generated objects.	Best suited for large-scale datasets with widely distributed pests across multiple crops.

Table 1: Comparative study of different GAN models for Data Augmentation used in Agriculture

3. Structural Design of this Work:

This is a breakdown of this paper's structure:

- Section 2 outlines related works of the paper
- Section 4: Data Augmentation technique based on GAN for saw-fly pest images
- Section 5: proposed methodology for data augmentation.
- Section 6: provides the experiment of this work.
- Section 7: offers the results and discussion.
- This paper concludes in Section 8, providing future recommendations for improving the visualization and identification of plant pests.

In Figure-2, the organization of the work has been presented.

4. Dataset:

The problem of pest detection inherently faces the challenge of data scarcity. Collecting high-quality, labeled pest images is time-intensive and resource-demanding, especially for small pests like Sawfly, which require close-up imaging under controlled conditions. Additionally, the lightweight nature of the object detection models used in this study makes them suitable for deployment on edge devices such as smartphones and IoT-enabled agricultural sensors. These practical constraints further necessitate working with limited data, as such systems are typically optimized for low computational resources and small datasets.
In this study, we begin with a relatively small dataset of 100 images to evaluate how different GAN architectures perform in generating synthetic pest images under data-constrained scenarios. This approach helps in understanding the potential of GAN-based augmentation to alleviate the challenges posed by limited data availability. By creating diverse and realistic synthetic images, these models aim to enhance dataset variability and improve model training even with small sample sizes.

4.1. Dataset Arrangement:

At first, we gathered images of Sawfly pests to train and test our proposed classifier (a few sample images are present in Figure 1). The images were collected from different sources like laboratory data, field data, and several Weblinks, such as from the Plant-Village site. Next, we categorised the images into one class. Three hundred photos of maize leaves from PlantVillage data set B made up the experimental data set. The test and training data are separated by a ratio of 1:4. Table 2 displays the Sawfly data used in this experiment.

4.2. Dataset Labeling:

We re-shaped the collected mung bean images of our dataset to 160x160 to decrease the training time, which was spontaneously calculated in the Python platform by the defined written script, which used the context of the OpenCV.

Pest Name	Crop Name	No. of Images
Saw Fly	Cabbage	20
	Cauliflower	20
	Mustard	20
	Radish	20
	Turnip	20

Table 2: The Sawfly image dataset used in the experiment

Equipments	Specifications
Processor	Intel core i7, 7800X GPU 3.5 GHz
Memory	256 GB
Graphics	GeForce GTX 1080Ti

Table 3**:** Required Hardware setup for the experiments

Figure 1: Original image data (sample) of the Sawfly dataset

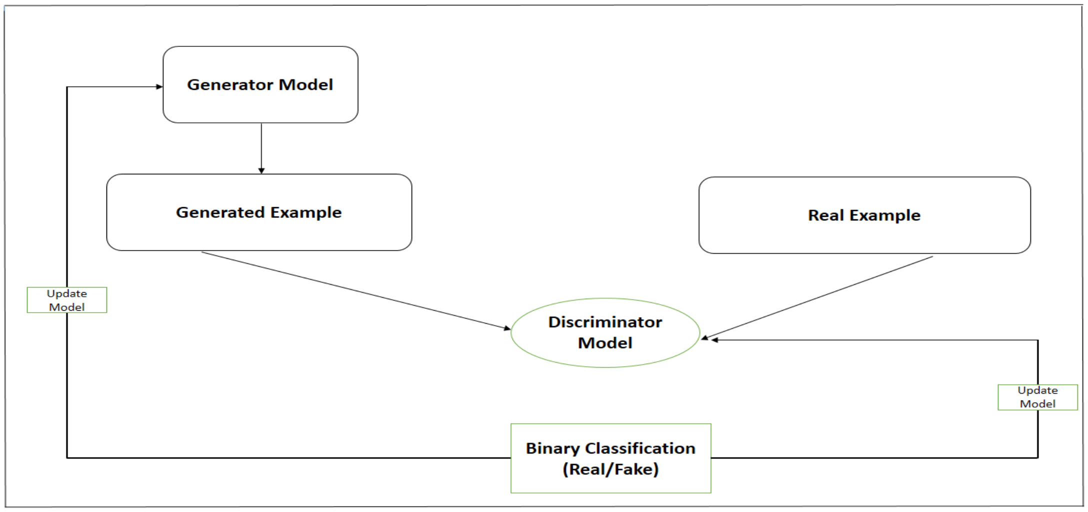

Figure 2: GAN based synthetic image generation technique for Data Augmentation

5. Proposed Methodology:

The workflow of the proposed model or the data augmentation pipeline is represented as follows-

- The pest images are extracted from the original images in the dataset using Bounding box method.
- Train CycleGAN with the extracted pest images to produce synthetic pest images.
- Prepare the pest masks from the pest images using binary thresholding method.
- Apply CPB to blend the pest masks into external leaf images to prepare the final synthetic image.
- Assemble the original images and synthetic images to get new dataset with adequate amount of data to be used for Deep Learning based object detection methods like YOLOv7 and more.

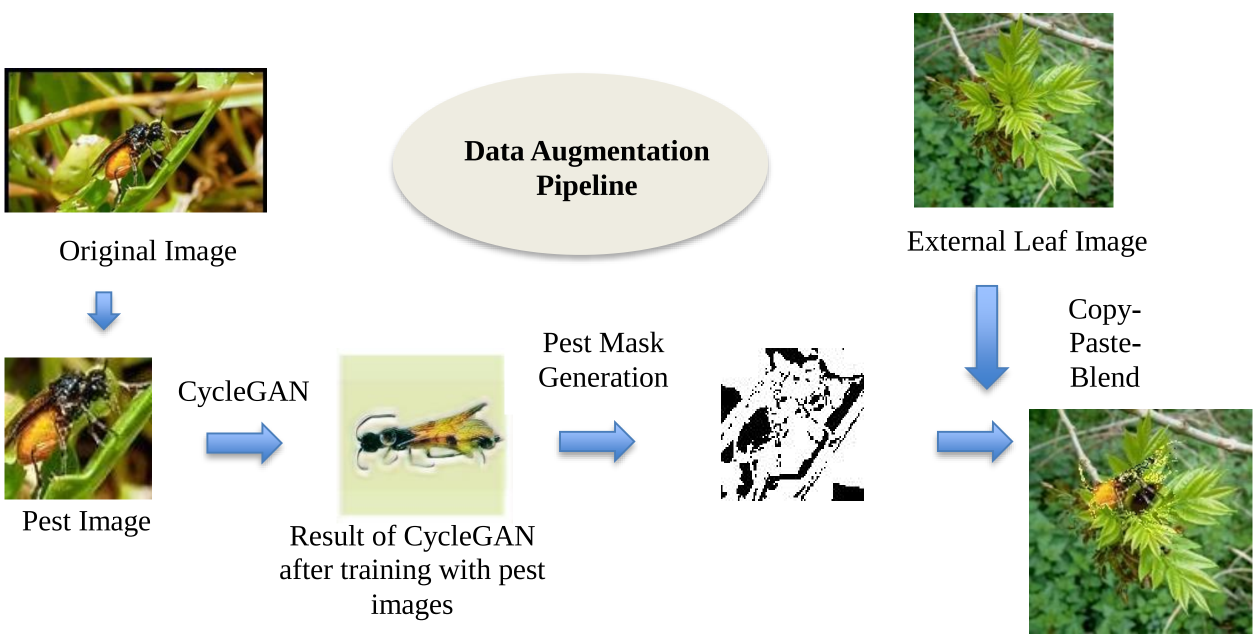

Figure 3: Proposed Data Augmentation Pipeline

6. Experiment:

6.1. Experimental Setup –

Here's a table summarizing the computational requirements for different GAN models based on their complexity and GPU requirements:

GAN Model	Computational Weight	Minimum Feasible GPU	Optimal GPU	Training Time (Small Dataset)	Memory Usage (VRAM)	Performance Observed	Recommended Task Type
CycleGAN	Moderate	NVIDIA GTX 1660 Ti (6GB)	NVIDIA RTX 3060 (12GB) or RTX 3080	~4–6 hours for 100 images (180×180)	~6GB–12GB	High-quality, diverse synthetic images suitable for augmentation.	Medium datasets, unpaired data augmentation.
WGAN	Heavy	NVIDIA GTX 1660 Ti (6GB)	NVIDIA RTX 3090 (24GB)	~12–15 hours for 100 images (180×180)	~8GB–24GB	Stable training, slower convergence; high-quality images.	High-resolution image synthesis, stable training.
DCGAN	Light	NVIDIA GTX 1050 Ti (4GB)	NVIDIA RTX 2060 (6GB)	~2–3 hours for 100 images (180×180)	~4GB–6GB	Moderate-quality images, faster training but limited diversity.	Small datasets, quick prototyping, low-resolution images.
CPB + GAN	Very Light	CPU (Intel Core i7) or NVIDIA GTX 1650	NVIDIA GTX 1660 Ti (6GB)	~1–2 hours for 100 images (180×180)	~4GB–6GB	Blends pest masks effectively but limited generalizability.	Small-scale blending tasks with minimal resource needs.

Table-4: Basic computational requirement for GAN based models

6.2. Our observations -

All the experiments were actually performed on a 12th Gen Core i7-12700H with 16 GB RAM. The resolution of the images used was 180X180. The generative models used in the experiments were trained using Keras and Tensorflow. According to the previous discussions four models DCGAN, CycleGAN, WGAN and CPB+GAN were applied on the SawFly dataset.

- Since the dataset is too small with only 100 images, DCGAN was trained with different setup with (Batch size=64, epochs=5000), (Batch size=64, epochs=200) and (Batch size=8, epochs=200), among these the first one gave better result.
- CycleGAN was trained using 100 epochs and got better results than DCGAN.
- Wasserstein GAN(WGAN) , trained with batch_size=16, epochs=1000, clip_value=0.01 did not show satisfactory results for the dataset used in the experiment.

- But surprisingly the CPB+GAN method was adopted by Karam et al. [7] to develop their webapp CPB Image Generation Tool at Humans and Machine Lab, American University of Beirut was generating repetitive images taking the images of our dataset as source images and the pest masks as mask images. Finally, inspired by [7] the Copy-Paste-Blend method was applied in our experiment on external leaf images with pest masks. Pest masks were generated using Binary thresolding applied on extracted pest images om original images from the dataset. The above mentioned pipeline gave a satisfactory result and can be used in future.

7. Result:

Figure 5: DCGAN result with batch size = 64, epoch = 5000

Figure 6: CycleGAN result with epoch = 100

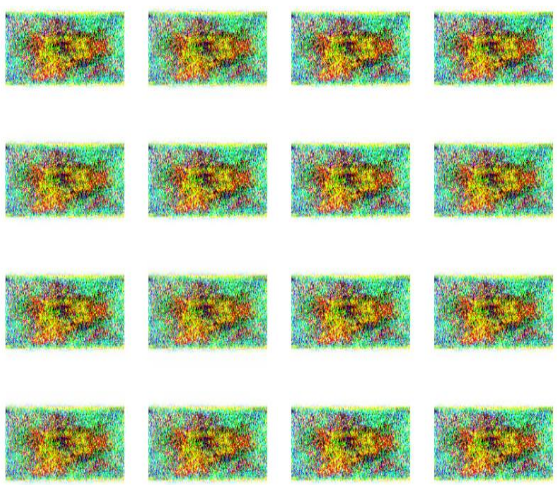

Figure 7: WGAN at epoch=400

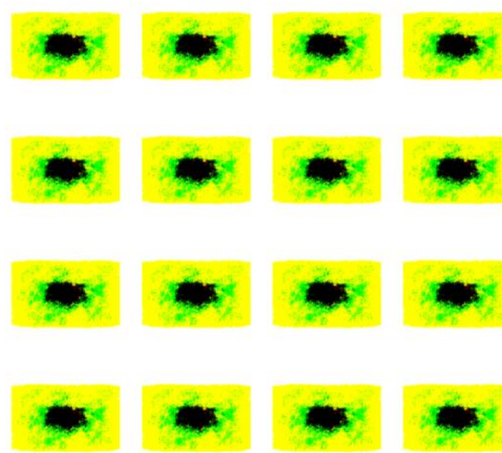

Figure 8: WGAN at epoch=1000

Figure 9: Before Copy-Paste-Blend Source leaf image

Figure 10: Copy-Paste-Blend output

Pest Image

Pest mask generated by binary thresholding

8. Conclusion & Future Scope-

In this study, we explored popular GAN models to enhance data augmentation, a key requirement when working with small datasets for plant pest detection. Future work will expand this study to larger and more diverse datasets, allowing for a comprehensive analysis of the generalizability and scalability of these GAN-based augmentation techniques. This progression will enable a more robust assessment of how such methods can be integrated into real-world agricultural systems for accurate and efficient pest detection.

Our primary goal was to identify a suitable model for augmenting images of small pests, specifically aiming at real-world scenarios where images captured on farms often make it challenging to distinguish smaller pests. However, for our experiment, we used a dataset containing zoomed-in images of sawflies, where pests were clearly visible, reducing the difficulty of identifying and extracting the objects in each image. This dataset was more applicable to data augmentation than direct object detection, as the extracted images were already optimized for isolating pests. Our findings indicate that CycleGAN outperformed other models in generating synthetic pest images for the dataset used. Besides CycleGAN, the Copy-Paste-Blend (CPB) technique demonstrated strong potential for seamlessly merging pest masks with leaf images. By blending pest masks of different scales with various leaf backgrounds, this method produces more natural-looking synthetic images. Moving forward, we aim to develop a more versatile model that can adapt to diverse datasets, learning patterns and scales to generate even more realistic synthetic images and can work for several types of pest. These improvements could refine data augmentation tools, leading to more accurate pest detection in real-world applications.

References :

1. Ali, F., Qayyum, H., & Iqbal, M. J. (2023). Faster-PestNet: A lightweight deep learning framework for crop pest detection and classification. *IEEE Access*, 11, 3317506. doi:10.1109/ACCESS.2023.3317506.
2. Antwi, K., Bennin, K. E., Asiedu, D. K. P., & Tekinerdogan, B. (2022). On the application of image augmentation for plant disease detection: A systematic literature review. *Computers and Electronics in Agriculture*, 201, 107341. https://doi.org/10.1016/j.compag.2022.107341.
3. Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pp. 214–223.
4. Cap, Q. H., Uga, H., Kagiwada, S., & Iyatomi, H. (2020). LeafGAN: An effective data augmentation method for practical plant disease diagnosis. *arXiv preprint*, arXiv:2002.10100.
5. Dai, M., Xin, M., & Liang, L. (2021). AWGAN: Empowering high-dimensional discriminator output for generative adversarial networks. *DeepAI*.
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 2672–2680.
7. Karam, C., Awad, M., Abou Jawdah, Y., Ezzeddine, N., & Fardoun, A. (2022). GAN-based semi-automated augmentation online tool for agricultural pest detection: A case study on whiteflies. *Frontiers in Plant Science*. doi:10.3389/fpls.2022.813050.
8. Liu, B., Tan, C., Li, S., He, J., & Wang, H. (2020). A data augmentation method based on generative adversarial networks for grape leaf disease identification. *IEEE Access*, 8, 226386–226394. doi:10.1109/ACCESS.2020.2998839.
9. Ming Yuan Xin, Ling Weay Ang, & Sellappan Palaniappan. (2023). A data augmented method for plant disease leaf image recognition based on enhanced GAN model network. *Journal of Informatics and Web Engineering*, 2(1), 1–12. doi:10.33093/jiwe.2023.2.1.1.
10. Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint*, arXiv:1511.06434.
11. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6, Article 60.
12. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2223–2232. doi:10.1109/ICCV.2017.244.
13. Arunabha M. Roy et. al. A Deep Learning Enabled Multi-Class Plant Disease Detection Model Based on Computer Vision, AI 2021, 2(3), 413-428; https://doi.org/10.3390/ai2030026
14. Liu, J., Wang, X. Plant diseases and pests detection based on deep learning: a review. Plant Methods 17, 22 (2021). https://doi.org/10.1186/s13007-021-00722-9