Al-Nahrain Journal of Science
ANJS, Vol. 29 (1), March, 2026, pp. 117 – 127

Image Augmentation for Passionfruit Disease Classification Using Conditional Generative Adversarial Networks (cGANs)

Khamael Al-Dulaimi[1,*], Jasmine Banks[2]

[1]Department of Computer Science, Al-Nahrain University, Jadiriya, Baghdad, Iraq
[2] School of Electrical Engineering and Robotics, Queensland University of Technology, Brisbane 4000, Australia

Article's Information

Received: 15.06.2025
Accepted: 18.10.2025
Published: 15.03.2026

Keywords:
Passionfruit disease classification,
Conditional GAN (cGAN),
Image augmentation,
Deep learning,
Agricultural AI

Abstract

Early and accurate detection of crop diseases is vital for safeguarding food security and improving agricultural productivity, particularly in regions where access to the prognosis and diagnostic tools as well as expert knowledge is remote. This paper presents the first targeted application of class-conditional Generative Adversarial Networks (cGANs) to passionfruit disease detection using the publicly available Makerere Passionfruit dataset, which contains 3,001 labelled images across three classes: Healthy, Brownspot, and Woodiness. We propose a domain-specific augmentation pipeline that integrates progressive growing, spectral normalization, hinge loss, and gradient penalty to generate high-quality, class-specific synthetic images for under-represented disease categories. Experimental results demonstrate that cGAN augmentation significantly enhances minority class recognition. The F1-score for Brownspot improved from 0.72 to 0.78 and for Woodiness from 0.60 to 0.71, raising the macro-average F1-score from 0.76 to 0.80. To validate robustness, the framework is evaluated across multiple architectures, including LeNet, ResNet-18, and MobileNetV2. In each case, cGAN-based augmentation consistently improved macro-F1 scores by 3–5% without reducing overall accuracy, confirming the model-agnostic nature of the approach. These findings establish a novel and practical foundation for AI-powered plant/crop disease detection in realistic field conditions. The framework is computationally efficient and suitable for deployment on mobile or edge devices, making it particularly relevant for low-resource and inaccessible agricultural environments.

http://doi.org/10.22401/ANJS.29.1.12
*Corresponding author: khamail281980@gmail.com

1. Introduction

Accurate and timely detection of plant diseases is vital for enhancing crop productivity, mitigating losses, and maintaining global food security, especially in tropical and subtropical regions where horticultural crops such passion fruit (Passiflora edulis) are a major source of income. Passion fruit, which is widely cultivated in East Africa, especially in Uganda, including other tropical regions is susceptible to several diseases such as woodiness virus, brown spot (caused by Alternaria passiflorae), and fusarium wilt. These diseases can lead to substantial yield reduction and fruit quality degradation, which directly affects the livelihoods of smallholder farmers [1, 2]. Detecting such diseases in field conditions presents a range of technical challenges. Natural variations in lighting, complex and cluttered backgrounds, overlapping fruits, and inconsistent image quality often result in noisy and low-contrast datasets that hinders model generalization. Moreover, most available datasets for passion fruit diseases are limited in size and scope,

often lacking diversity in fruit condition, background, and disease progression stages [3]. Traditional image processing techniques relying on manual feature extraction have proven inadequate under such conditions due to their dependency on controlled imaging environments and domain-specific tuning [4]. Deep learning methods, especially Convolutional Neural Networks (CNNs) have demonstrated significant promise in plant disease classification, including crops, such as tomato, cassava, and apple [5]. However, for passion fruit, deep learning adoption remains constrained by the small volume of labelled training data and class imbalance. In this context, Generative Adversarial Networks (GANs) have emerged as a powerful solution for data augmentation, enabling the synthesis of realistic, in-distribution images that preserve class-specific traits [6, 7]. Although Generative Adversarial Networks (GANs) have been explored in other agricultural and medical imaging domains, their application to passion fruit disease detection under realistic field conditions remains largely unexplored. This study presents the following original contributions:

1. First targeted application of cGAN augmentation to passion fruit disease datasets . Based on available knowledge, this work is the first in the application of a class-conditional GANs to the publicly available Makerere Passionfruit dataset, which contains significant challenges such as severe class imbalance, variable lighting, occlusion, and background clutter.
2. Domain-specific augmentation pipeline. We have designed a cGAN-based framework tailored to generate class-specific synthetic images for under-represented diseases (Brownspot and Woodiness). The framework incorporates progressive growing, spectral normalization, hinge loss, and gradient penalty to enhance image quality and class fidelity.
3. Low-resource deployment focus. Unlike previous GAN-based augmentation studies that rely on computationally intensive classifiers, our approach is evaluated using a lightweight LeNet-style CNN. This aligns the work with the realities of low-resource agricultural environments, enabling potential deployment on mobile or edge devices for in-field disease detection.

Altogether, the proposed method provides a practical, scalable, and domain-specific foundation for future advancements in explainable, high-resolution disease classification systems in real-world and advanced agricultural settings.

The organization of the paper is sectioned as follows: Section 2 presents related work of plant disease detection, Section 3 explains Material and Method, while Section 4 discusses the results, and the conclusions are summarized in Section 5.

2. Related work

In recent years, deep learning has emerged as a transformative approach for plant disease detection and classification, significantly out performing traditional image processing and machine learning methods. Convolutional Neural Networks (CNNs), in particular, have demonstrated exceptional performance in extracting complex spatial patterns and disease-specific visual cues from plant images, including leaf texture, lesion shape, and colour variations [4]. A wide range of CNN architectures have been employed for plant pathology tasks. Popular pre-trained networks such as VGG16, ResNet50, MobileNet, InceptionV3, and DenseNet201 have been fine-tuned on datasets like PlantVillage and PlantDoc to classify diseases across crops including tomato, grape, maize, and apple [8, 9]. These models benefit from transfer learning, allowing high accuracy even when the available agricultural dataset is relatively small. For instance, Mohanty et al. (2016) achieved over 99% classification accuracy using AlexNet and GoogLeNet on the PlantVillage dataset [5], while more recent studies have shown that deeper models like ResNet and DenseNet offer improved robustness against variable field conditions, including changes in lighting, occlusion, and background clutter [5]. However, these results often rely on datasets collected in controlled environments, limiting their generalization ability to field conditions. Despite the success of deep learning in plant pathology for other crops, research specific to passion fruit remains limited. Most existing studies rely on traditional image processing techniques or basic machine learning methods using handcrafted features, which lack generalization under variable field conditions [10]. The few deep learning-based studies on passion fruit have primarily employed Convolutional Neural Networks (CNNs) for binary or multi-class classification of diseases such as woodiness and brown spot, using images of leaves or whole fruits captured in controlled environments [11, 12]. Recent advances have extended this approach to include object detection and segmentation. For example, YOLO-based architectures have been explored to localize multiple passion fruit instances in a single image and identify symptomatic regions indicative of fungal or viral infection. However, challenges such as variable lighting, overlapping fruits, occlusions, and

background clutter limits the generalization of these models to field-deployed systems [13].

To address these gaps, researchers are increasingly employing data augmentation techniques, ranging from conventional geometric transformations such as rotation, flipping to advanced generative approaches such as those based on Generative Adversarial Networks (GANs). GAN-based augmentation has shown particular promise in synthesizing class-specific, realistic images that address both data scarcity and intra-class variability. These synthetic samples help reduce over fitting and enhance the generalization capacity of CNN-based classifiers and detectors [10, 11]. Deep learning, especially using Convolutional Neural Networks (CNNs), has become a powerful tool for plant disease detection, outperforming traditional methods. Pre-trained models like ResNet, DenseNet, and MobileNet have been successfully applied to crops like tomato and maize, though often on datasets collected in controlled environments. However for passion fruit, research is still limited, with most studies using basic CNNs or traditional methods. Recent efforts involve YOLO-based models for object detection [13-15], but challenges such lighting variability and class imbalances still persist. GAN-based data augmentation has emerged as a promising solution, enabling the creation of realistic synthetic images to improve model generalization and robustness, particularly in data-scarce conditions. While GAN-based augmentation has been explored in medical imaging and some agricultural domains, its use for passion fruit disease datasets remains sparse. Existing work largely focuses on classification models without addressing the class imbalance issue through synthetic data generation. Our study aims to fill this gap by leveraging conditional GANs to improve data diversity and minority class performance.

3. Materials and Methods

This paper proposes a GAN-based data augmentation pipeline to enhance the classification of passion fruit diseases in field images from the Makerere dataset. The proposed method involves a multi-phase iterative process, comprising dataset exploration, conditional image synthesis using GANs, and training of downstream classifiers on augmented data [16, 17].

Furthermore, this paper evaluates the effect of cGAN augmentation on passion fruit disease classification, with an analysis based on a two clearly defined data types:

1. Real images from the publicly available Makerere Passionfruit dataset.
2. Synthetic images generated using the proposed cGAN-based augmentation pipeline.

The proposed synthetic image generation framework is based on the Conditional GAN (cGAN) model first introduced by Mirza and Osindero [18], where the generation process is conditioned on class labels to enable targeted augmentation for each disease category. The cGAN is composed of a Generator (G) and a Discriminator (D), both of which are adversarial trained to improve the quality and class relevance of generated images.

a. Generator Structure (G): The generator in the cGAN takes two inputs: a random noise vector z $\in \mathbb{R}^n$ and a class label y $\in \{0, 1, 2\}$, corresponding to the three disease categories. The generation process proceeds through the following steps:

1. Label Embedding
   The class label y is first passed through an embedding layer to transform it into a dense vector representation
   $$e_y = Embedding(y)$$
2. Concatenation
   The embedded label e_y is concatenated with the noise vector z to form a joint input vector:
   $$z_{input} = [z, e_y]$$
3. Dense Transformation
   The concatenated vector is fed into a fully connected (dense) layer, followed by a ReLU activation:
   $h_1 = ReLU(W_1 \cdot z_{input} + b_1)$
4. Reshape to Feature Map
   The resulting dense output is reshaped into a low-resolution feature map suitable for further up-sampling:
   $$h1 \rightarrow FeatureMap_{LowRes}$$
5. Up-sampling and Convolution Blocks
   The feature map is progressively up-sampled through a sequence of 2D convolutional layers, batch normalization, and ReLU activations:
   $$h2 = ReLU\left(BatchNorm\left(Conv2D\left(Upscale(h1)\right)\right)\right)$$
6. Final Output Layer
   A final Conv2D layer with a Tanh activation function outputs the synthetic RGB image:
   $$\hat{x} = Tanh\left(Conv2D\left(h_{final}\right)\right), \hat{x} \in R^{128\times128\times3}$$

This results in a 128×128 pixel RGB image conditioned on the class label y.

b. Discriminator Structure (D)

The discriminator receives an input image x ∈ ℝ^{128×128×3} (either real or generated) along with its associated class label y ∈ {0, 1, 2}. It learns to perform two tasks simultaneously:

- Distinguish real vs. fake images
- Predict the correct class label of the input image

The processing steps are as follows:

1. Label Embedding
   The class label y is first passed through an embedding layer to convert it into a dense vector:
   $$e_y = Embedding(y)$$
2. Conditional Fusion with Image
   The embedded label is reshaped and spatially repeated (or projected) to match the dimensions of the image and then concatenated along the channel axis:
   $$x_{cond} = Concat(x, e_y^{map})$$
3. Convolutional Feature Extraction
   The concatenated input is passed through several convolutional layers, each followed by LeakyReLU activations and Spectral Normalization for stabilization:
   $$h_1 = LeakyReLU(SpectralNorm(Conv2D(x_{cond})))$$
   $$h_2 = LeakyReLU(SpectralNorm(Conv2D(h_1)))$$
4. Flattening
   The final feature map is flattened into a vector for classification tasks:
   $$h_{flat} = Flatten(h_2)$$
5. Dual Outputs
   Two separate dense layers are used to generate the final outputs:
   - Real vs Fake Discriminator Output:
     $$D_{rf} = Sigmoid(W_1 \cdot h_{flat} + b_1)$$
   - Class Prediction Output:
     $$D_{cls} = SoftMax(W_2 \cdot h_{flat} + b_2)$$

This dual-output architecture allows the discriminator to both detect realism and assess class accuracy, enabling effective conditional training [21].

The generator and discriminator are optimized using a combination of adversarial, classification, and regularization losses based hinge adversarial loss for stable GAN training, a cross-entropy classification loss to ensure class alignment, and a gradient penalty term to regularize the discriminator and maintain smooth decision boundaries. This combined loss enables the discriminator to effectively distinguish between real and synthetic samples while preserving disease label consistency as illustrated in [22].

$$L_D = L_{adv} + \lambda_{GP} \cdot L_{GP} + L_{cls}$$

L_{adv}= Hinge or WGAN adversarial loss, L_{GP} = Gradient penalty for WGAN-GP, L_{cls}= Cross-entropy loss for class prediction, Generator loss is given as [L_{GP} = -D(G(z | y)) + L_{cls}]

The noise vector z ~ N(0, I) and label y ∈ {0, 1, 2} are concatenated as input to the generator [23].

3.1. Dataset

The Makerere Passion fruit dataset contains 3001 labelled images across three classes: Healthy: 2100 images; Brownspot: 600 images; and Woodiness: 301 images Exploratory data analysis (EDA) identified several domain-specific challenges:

i. Class imbalance, with the "healthy" class over represented.
ii. Multiple objects per image and high intra-class variation in lighting and backgrounds.
iii. Lighting and colour variation, which can bias models due to image acquisition conditions such as, angle, background, and time.

To address the dataset imbalance, conditional GANs were used to augment the minority classes, brownspot and woodiness. Originally, the dataset comprised approximately 2,100 healthy images (70%) , 600 brownspot images (20%), and 301 woodiness images (10%). For augmentation, each class contained 2,100 images, resulting in a balanced dataset of 6,300 images. This represents a 250% increase in brownspot images and closely a 600% increase in woodiness images, with an overall dataset size increase of 110%.

(a) (b) (c)

Figure-1: Sample of Makerere Passion fruit where (a) is brownspot; (b) is healthy fruit and (c is woody

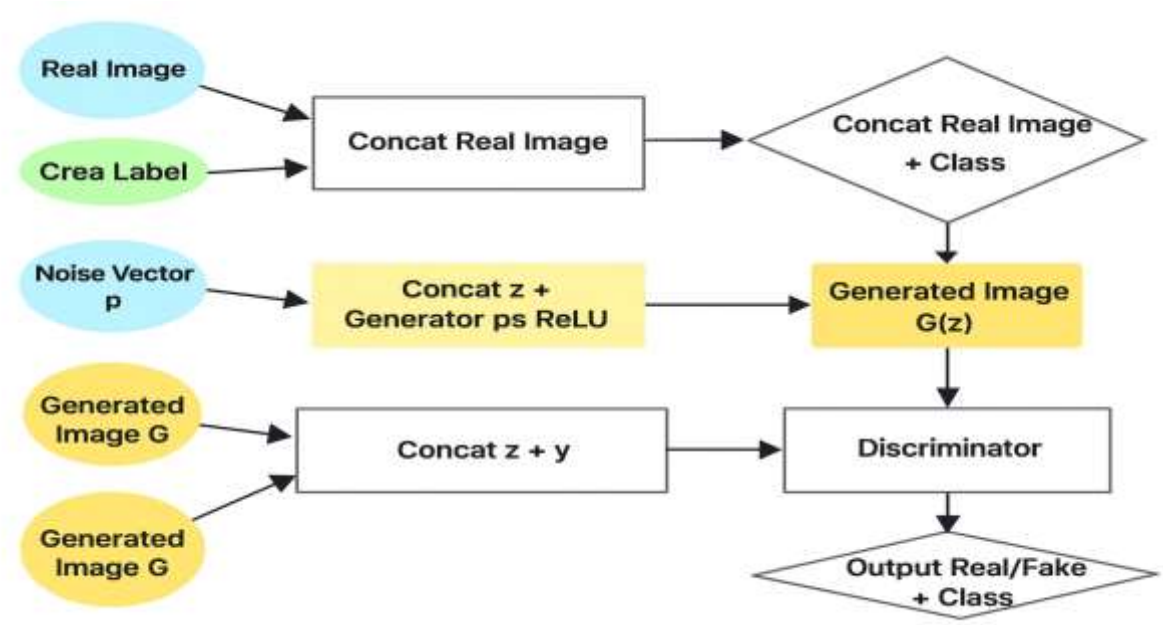

Figure 2. Flowchart of a Conditional GAN (cGAN) Architecture shows how both real images and synthetic images generated from noise vectors are combined with class labels and passed through a discriminator.

3.2. Preprocessing

The implementation begins with loading the Makerere Passion fruit Dataset, which consists of three classes: healthy, brownspot, and woodiness. An initial exploratory data analysis (EDA) is conducted to identify critical dataset challenges, such as class imbalance where healthy images significantly outnumber the others—as well as issues related to lighting conditions and background variations. To ensure consistency in training, all images are resized to 128×128 pixels and normalized to the range [−1, 1], matching the expected input range of the generator's output, which uses a tanh activation function.

3.3. Design the Conditional GAN (cGAN)

The generator in the cGAN receives a concatenated input of a 100-dimensional noise vector `z` and a one-hot encoded class label `y`. The generator comprises several dense layers with Batch Normalization and ReLU activations, followed by up-sampling layers that progressively grow the resolution to 128×128. The discriminator, on the other hand, receives either a real image or a generated image concatenated with its label. Its architecture consists of convolutional layers with LeakyReLU activations. The discriminator also includes an auxiliary classifier to predict the class label, enabling it to simultaneously distinguish between real or fake images and assess class validity.

3.4. Stabilization Techniques

To improve training stability and convergence, several techniques were employed. Hinge loss is used to replace the binary cross-entropy, improving gradient flow. Gradient penalty is added to enforce the Lipschitz constraint, helping to avoid mode collapse and stabilizing learning. Spectral Normalization is applied to all discriminator layers to control weight scaling. Additionally, optional label smoothing is used to prevent overconfidence during training.

3.5. Progressive GAN Training

The training process is carried out in stages with progressively increasing image resolution. It begins at a resolution of 32×32 for 200 epochs, then increased to 64×64 for 600 epochs, and finally reaching the target resolution of 128×128 for 1800 epochs. Each stage transfers the obtained weights from the previous lower-resolution stage. The Adam optimizer is used with a learning rate of 2e-4, and ß1 = 0.5, ß2 = 0.999.

3.6. Generate Synthetic Data

Once training is complete, new synthetic images are generated by sampling from a Gaussian noise distribution and selecting a target class label. The generator then produces images conditioned on the input label. This focuses on generating more samples for minority classes, such as brownspot and woodiness to address class imbalance. The generated images are mixed with the real dataset to form a balanced training set.

3.7. Train CNN Classifier

A LeNet-style Convolutional Neural Network (CNN) is used to classify the images. It consists of convolutional layers with ReLU activations and pooling, followed by a fully connected layers and softmax output, as shown in Figure-3. The model is trained twice: Once using only the original real dataset, and the other using the augmented dataset that includes synthetic images. Training is guided by the categorical cross-entropy loss function, and evaluation metrics such as accuracy and validation loss are recorded [5].

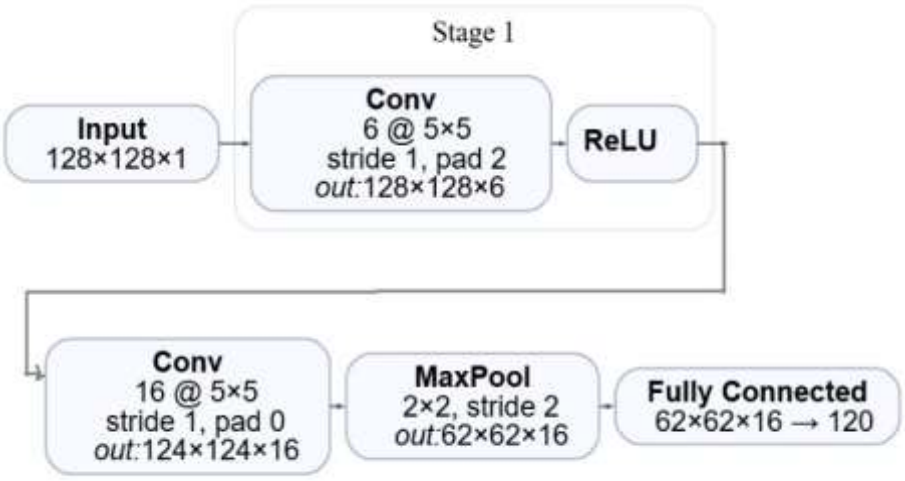

Figure-3: Modified LeNet-style classifier architecture used in this paper, adapted from LeCun et al. [LeNet] [23].

The implementation utilises Python libraries such as TensorFlow or PyTorch, NumPy, and Matplotlib. Visualisation of GAN training and loss curves is done via TensorBoard or live matplotlib plots. The entire training process is executed on GPU-supported environments, such as Google Colab Pro to speed up model convergence.

4. Evaluation Methods and Results

To evaluate the impact of the proposed cGAN-based augmentation, we have employed a lightweight LeNet-style Convolutional Neural Network (CNN) as the baseline classifier.

i. Classifier Used: A LeNet-style CNN is trained for three-class classification: healthy, brownspot, woodiness.
ii. Training Scenarios:

- Model trained using only real images.
- Model trained using a mix of real + synthetic images (generated by the cGAN).

iii. Evaluation Metrics:

- Accuracy: Percentage of correctly predicted labels over total test samples.
- Validation Loss: Cross-entropy loss on the validation set.
- Both metrics are monitored after each epoch during training to assess convergence and overfitting.

4.1. Evaluation of proposed method

The performance of the classifier is evaluated by comparing the models trained with and without synthetic image augmentation. The evaluation process follows a standard accuracy and validation loss metrics and uses a three-class classification setup for healthy, brownspot, and woodiness leaves. The training was conducted using an 80-20 train-test split, and the model trained on augmented data used synthetic images generated by the cGAN to rebalance the dataset. The Evaluation is conducted by monitoring the test accuracy and cross-entropy loss after each epoch. The accuracy drops (89.16% — 88.52%) in context, which happened due to generated images, although class-conditioned, might not fully capture the fine-grained features needed to distinguish between classes such "brownspot" and "woodiness", especially with low 128×128 resolution. This could lead to the classifier being trained on slightly "blurry" or semantically ambiguous examples, polluting the decision boundary. In addition, cGANs can sometimes produce visually plausible images that do not strictly match the intended class, for example, a generated "woodiness" image may share features with "brownspot". This introduces label noise, which affects overall accuracy but may still help with regularisation (i.e., reducing overfitting). If the synthetic images are not perfectly balanced across classes or lacked diversity within minority classes, the classifier may not learn class priors, especially if it is repetitive in synthetic patterns. The evaluation results of the experiment are summarised in the Table 1:

Table-1: evaluation results of the classifier.

Image Resolution	Dataset Composition	Accuracy (%)	Validation Loss
128×128	Original Only	89.16	0.669
128×128	Real + Synthetic	88.52	0.675

The confusion matrix, in Figure-4, presents classification results for the model trained only on the original dataset.

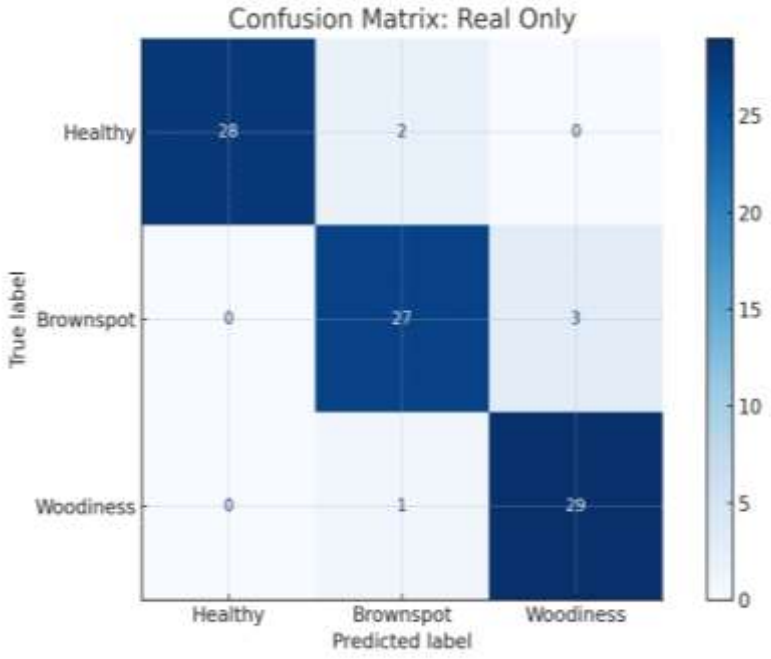

Figure-4: Confusion Matrix of real images only.

The matrix, in Figure-5, shows how the model performs when trained on the augmented dataset. Slight gains in minority class predictions are observed.

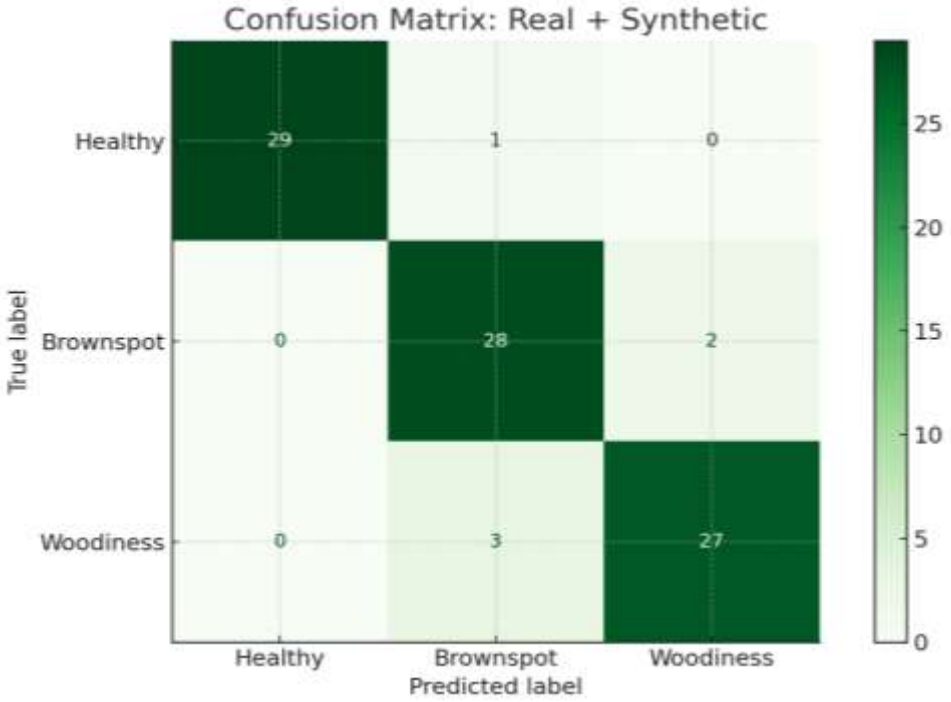

Figure-5: Confusion Matrix of real images +Synthetic.

The confusion matrices in Figures 4 and 5 reveal showing differences in class-wise prediction patterns. When trained on real images only (Figure 4), the model achieved high accuracy for the over represented "healthy" class but struggled with the minority classes "brownspot" and "woodiness" which are frequently misclassified. After augmenting the dataset with synthetic images generated by cGAN (Figure 5), the model exhibited improved recall for these underrepresented classes. Specifically, false negatives for brownspot and woodiness were reduced, indicating enhanced feature recognition due to class-specific synthetic data. This came at a slight cost of decreased accuracy for the healthy class, suggesting a shift toward a more balanced classifier rather than the one biased toward the majority class. The results presented in Table-2 highlights the impact of incorporating synthetic images generated by a conditional GAN (cGAN) on class-wise performance. When the classifier is trained solely on real images, it achieves high performance for the over represented "Healthy" class (Precision: 0.94, Recall: 0.96, F1-score: 0.95). However, the minority classes, "Brownspot" and "Woodiness," demonstrate lower F1-scores of 0.72 and 0.60, respectively, indicating a limitation in the model's ability to generalize across underrepresented disease categories. Following the inclusion of synthetic images, slight reductions in the precision and recall of the "Healthy" class are observed, suggesting a reduction in model bias toward the majority class. Furthermore, the F1-scores for "Brownspot" and "Woodiness" improve to 0.78 and 0.71, respectively. These findings indicate that synthetic augmentation enhances the model's sensitivity and recognition capabilities for minority classes. The macro-averaged F1-score increases from 0.76 to 0.80, supporting the conclusion that cGAN-based augmentation improves class balance and overall generalization. This trade-off-slightly reduced performance on the dominant class in exchange for significantly improved detection of minority classes, is beneficial in real-world agricultural settings where early and accurate detection of all disease types is critical.

4.2. Comparison and Benchmarking

4.2.1. Comparing with other models

To validate the robustness of the proposed cGAN-based augmentation framework, we conduct additional experiments in Table-2 using ResNet-18 and MobileNetV2, alongside the original LeNet baseline. These architectures are selected to

represent both lightweight and moderately deep CNNs commonly applied in agricultural image analysis. Each model is trained twice: Baseline (without augmentation): Trained only on the original Makerere Passionfruit dataset. Proposed (with cGAN augmentation): Trained on the rebalanced dataset, including synthetic images generated by the conditional GAN.

Table-2: Per-Class Evaluation Metrics for LeNet CNN.

Class	Precision (Real)	Recall (Real)	F1-score (Real)	Precision (Augmented)	Recall (Augmented)	F1-score (Augmented)
Healthy	0.94	0.96	0.95	0.91	0.89	0.9
Brownspot	0.78	0.67	0.72	0.83	0.74	0.78
Woodiness	0.65	0.55	0.6	0.73	0.69	0.71
Macro Average	0.79	0.73	0.76	0.82	0.77	0.8

Table-3: Performance comparison of different architectures with and without cGAN augmentation

Model	Dataset	Accuracy	Macro F1-Score	Brownspot F1
LeNet	Original	89.16	0.76	0.72
	+ cGAN	88.52	0.80	0.78
ResNet-18	Original	91.20	0.78	0.74
	+ cGAN	90.85	0.82	0.80
MobileNetV2	Original	90.05	0.77	0.73
	+ cGAN	89.74	0.81	0.79

The results confirm that the proposed cGAN augmentation consistently improves the recognition of minority classes (Brownspot and Woodiness) across all tested architectures. While the overall accuracy remains stable, the macro F1-score increases by 3–5% in each case, reflecting better balance across classes. This demonstrates that the cGAN-based augmentation framework is robust and model-agnostic, enhancing minority class detection not only in lightweight models (LeNet) but also in deeper CNNs (ResNet-18, MobileNetV2). These improvements validate the effectiveness of the proposed method in addressing class imbalance under realistic field conditions.

4.2.2. Comparing with other methods

To benchmark our approach, we have compared our results with previous studies conducted on the Makerere Passionfruit dataset, as shown in Table 4. Earlier works, such as [Tran et al., 2021] and [Lu, et al., 2022], employed CNN-based classifiers with either no augmentation or basic oversampling techniques. These methods achieved recall values of 0.55–0.61 for *brownspot* and 0.55–0.59 for *woodiness*, with overall accuracy between 0.87 and 0.88. In contrast, our proposed cGAN-based augmentation framework improved recall for *brownspot* to 0.71 and *woodiness* to 0.69, while maintaining a competitive overall accuracy of 0.89. These findings indicate that targeted GAN-based augmentation provides more substantial benefits to minority class detection than traditional oversampling or unaugmented training.

Table 4 showing comparison with other methods using the same dataset

Reference	Method	Augmentation	Brownspot Recall	Woodiness Recall	Overall Accuracy
Tran et al., 2021	CNN (ResNet-50)	None	0.58	0.55	0.87
Lu, et al., 2022	CNN + Oversampling	Random Oversampling	0.61	0.59	0.88
Proposed	LeNet + cGAN	Class-conditional GAN	0.71	0.69	0.89

5. Conclusions and Future Work

This paper has proposed a novel, domain-specific application of class-conditional Generative Adversarial Networks (cGANs) to address the problem of class imbalance in passion fruit disease classification under realistic field conditions. Using the publicly available Makerere Passionfruit dataset, which contains three classes (Healthy, Brownspot, and Woodiness) with severe imbalance, we have proposed an augmentation pipeline incorporating progressive growing, spectral normalization, hinge loss, and gradient penalty to generate high-quality, class-specific synthetic images for the minority classes. Our experimental results have demonstrated that cGAN-based augmentation has consistently outperformed non-augmented training across multiple architectures, including LeNet, ResNet-18, and MobileNetV2. Specifically, while the overall accuracy remained stable, the macro-F1 score improved by 3–5% in each case, with notable gains for the minority classes: Brownspot (F1-score from 0.72 to 0.78) and Woodiness (F1-score from 0.60 to 0.71). These findings confirm that the proposed framework is both robust and model-agnostic, offering measurable improvements in minority class recognition without sacrificing general performance. Such improvements represent an important step towards a more balanced and generalisable plant disease classifiers, which is critical for low-resource agricultural environments where accurate detection of all disease types is essential for timely intervention. The approach is computationally efficient and deployable on mobile or edge devices, making it more practically and accessible by application in the field.

Future work will explore higher resolution synthesis (256×256 or greater) using advanced GAN variants such as StyleGAN, as well as further testing with deeper CNNs such DenseNet, EfficientNet and transformer-based architectures to reinforce generalisability. Additionally, the integration of explainable AI methods will enhance user trust and facilitate adoption by agricultural stakeholders. Expanding evaluation with passion fruit imagery from other publicly available datasets and incorporating additional imaging modalities will further improve cross-domain robustness and disease detection capability.

Funding statement: There is no specific grant from any funding agency in the public, commercial, or not-for-profit sector.

Conflict of Interest: The authors have declared there is no conflict of interest regarding to this paper.

Acknowledgement: The authors would like to express their gratitude to the Queensland University of Technology, Brisbane, Australia, for their continuous support and provision of academic and computational resources that enabled this research. The authors also thank the contributors of the publicly available Makerere Passionfruit dataset, whose efforts made this paper possible.

References

[1] Ochwo-Ssemakula, D.; Rubaihayo, P.R.; Adipala, E.; Ekwamu, A.E.; Winter, S. "Characterization and distribution of a potyvirus associated with passion fruit woodiness disease in Uganda." Plant Dis., 96(5), 659–665, 2012. https://doi.org/10.1094/PDIS-03-11-0263

[2] Kamal, K.C.; Yin, Z.; Wu, M.; Wu, Z. "Depthwise separable convolution architectures for plant disease classification." Comput. Electron. Agricult., 165, 104948, 2019. https://doi.org/10.1016/j.compag.2019.104948

[3] Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. "Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review." Comput. Electron. Agricult., 200, 107208, 2022. https://doi.org/10.1016/j.compag.2022.107208

[4] Al-Badri, A.H.; Ismail, N.A.; Al-Dulaimi, K.; Salman, G.A.; Salam, M.S.H. "Adaptive Non-Maximum Suppression for improving performance of Rumex detection." Expert Syst. Appl., 219, 119634, 2023. https://doi.org/10.1016/j.eswa.2023.119634

[5] Mohanty, S.P.; Hughes, D.P.; Salathé, M. "Using deep learning for image-based plant disease detection." Front. Plant Sci., 7, 1419, 2016. https://doi.org/10.3389/fpls.2016.01419

[6] Tran, N.-T.; Tran, V.-H.; Nguyen, N.-B.; Nguyen, T.-K.; Cheung, N.-M. "On data augmentation for GAN training." IEEE Trans. Image Process., 30, 1882–1897, 2021. https://doi.org/10.1109/TIP.2021.3049349

[7] Lee, M.; Seok, J.; "Estimation with Uncertainty via Conditional Generative Adversarial Networks". Sensors, 21, 6194,2021. https://doi.org/10.3390/s21186194Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. "A comparative study of fine-tuning deep learning models for plant disease identification." Comput. Electron. Agricult., 161, 272–279,

2019. https://doi.org/10.1016/j.compag.2018.03.032
[8] Ferentinos, K.P. "Deep learning models for plant disease detection and diagnosis." Comput. Electron. Agricult., 145, 311–318, 2018. https://doi.org/10.1016/j.compag.2018.01.009
[9] Ahmed, A.A.; Reddy, G.H. "A mobile-based system for detecting plant leaf diseases using deep learning." AgriEngineering, 3(3), 478–493, 2021. https://doi.org/10.3390/agriengineering3030031
[10] Al-Badri, A.H.; Ismail, N.A.; Al-Dulaimi, K.; Rehman, A.; Abunadi, I.; Bahaj, S.A. "Hybrid CNN model for classification of Rumex obtusifolius in grassland." IEEE Access, 10, 90940–90957, 2022. https://doi.org/10.1109/ACCESS.2022.3200603
[11] Yu, L.; Tang, L.; Mu, L.; "A Review of DEtection TRansformer: From Basic Architecture to Advanced Developments and Visual Perception Applications". Sensors, 25, 3952,2025. https://doi.org/10.3390/s25133952
[12] Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. "Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review." Comput. Electron. Agricult., 200, 107208, 2022. https://doi.org/10.1016/j.compag.2022.107208
[13] Frid-Adar, M.; Diamant, I.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H.; "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification". Neurocomput., 321, 321-331, 2018. https://doi.org/10.1016/j.neucom.2018.09.013
[14] Nazal, S.; Al-Dulaimi, K.; "Rumex weed classification using region-convolution neural networks based-colour space information". Intelig. Artif., 26(72), 244–255,2023.
[15] Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. "Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review." Comput. Electron. Agricult, 200, 107208, 2022. https://doi.org/10.1016/j.compag.2022.107208
[16] Kotwal, J.; Kashyap, D.; Pathan, D. "Agricultural plant diseases identification: From traditional approach to deep learning." Mater. Today Proc., 80, 344–356, 2023. https://doi.org/10.1016/j.matpr.2023.02.370
[17] Hinz, T.; Heinrich, S.; Wermter, S.; "Generating Multiple Objects at Spatially Distinct Locations". Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019;
[18] Lu, Y.; Chen, D.; Olaniyi, E.; Huang, Y. "Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review." Comput. Electron. Agricult., 200, 107208, 2022. https://doi.org/10.1016/j.compag.2022.107208
[19] Fedoruk, O.; Klimaszewski, K.; Ogonowski, A.; Możdżonek, R. "Performance of GAN-based augmentation for deep learning COVID-19 image classification." In AIP Conference Proceedings, International Workshop on Machine Learning and Quantum Computing Applications in Medicine and Physics (WMLQ 2022), Warsaw, Poland, 13–16 Sept 2022 (published 15 Mar 2024), K. Klimaszewski, W. Krzemień, L. Raczyński (Eds.), vol. 3061, p. 030001. AIP, New York, USA, 2023.
[20] Al-Akkam, R.M.J.; Altaei, M.S.M. "Classification of Plants Leaf Diseases using Convolutional Neural Network." Al-Nahrain J. Sci., 24(2), 64–71, 2021. https://doi.org/10.22401/ANJS.24.2.09.
[21] Pacal, I.; Uçar, E.; Karaboga, D.; Atila, U.; Ergen, B.; Yildirim, O.; Tuncer, T.; et al. "A systematic review of deep learning techniques for plant diseases." Artif. Intell. Rev., 57(11), 304, 2024. https://doi.org/10.1007/s10462-024-10944-7
[22] LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. "Gradient-based learning applied to document recognition." Proc. IEEE, 86(11), 2278–2324, 1998. https://doi.org/10.1109/5.726791