Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A ROBUST GAN FRAMEWORK WITH PERCEPTUAL MULTI-LOSS LEARNING FOR APPLE LEAF IMAGE ENHANCEMENT

SHIVANI GUPTA[1*], POOJA RANI[2]

[1]Research Scholar, MM Institute of Computer Technology and Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala-133207, Haryana, India.

[2] MM Institute of Computer Technology and Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala-133207, Haryana, India.

[1]E-mail: shivanigupta.sh@gmail.com, [2]E-mail: poojasachdeva1886@gmail.com

Corresponding Author E-mail: shivanigupta.sh@gmail.com

ABSTRACT

High-quality images are essential for improving the accuracy of plant disease identification. Images captured in agricultural fields under natural conditions are often affected by various types of noise due to illumination changes, environmental disturbances and sensor limitations. This noise reduces the effectiveness of deep learning models. In this study, a perceptual GAN-based framework is developed to enhance the quality of apple leaf images obtained from PlantVillage dataset. In this proposed framework, adversarial loss, L1 loss, perceptual loss using a pre-trained VGG16 network, and structural similarity (SSIM) loss are combined so that both visual quality and structural information can be preserved. An encoder–residual–decoder architecture is used for the generator, while the PatchGAN is adopted as discriminator to enforce local texture realism. To make the training data more realistic, Gaussian, salt-and-pepper, and speckle noise were added during training. The performance of the proposed approach is compared with the Gaussian filter, Median filter, DnCNN, and UNet, and the best results are obtained in terms of PSNR and SSIM. The proposed perceptual multi-loss GAN achieved an average PSNR of 31.29 dB and an SSIM of 0.82 on images affected by Gaussian, salt-and-pepper, and speckle noise. The enhanced images exhibit improved structural clarity and are suitable for preprocessing in automated plant disease detection systems.

Keywords: *Apple Leaf Image Enhancement, Image Denoising, Generative Adversarial Network, Perceptual Loss, Deep Learning*

1. INTRODUCTION

Agriculture plays an important role in the economic development of any country.[1]. In agriculture, apples are the most common fruit, but various crop diseases are affecting agricultural growth and leading to substantial economic losses for both farmers and the Apple industry [2]. So, early detection is very important for plant disease detection. For this, some automated techniques help farmers at the early stage of disease. To detect plant disease, it is essential to have a high-quality image if the image contains noise, the model will be unable to identify the disease accurately [3]. Sometimes, under certain weather conditions, capturing a photo in real-world environments introduces noise that reduces image quality [4]. Therefore, to effectively train a model, it becomes essential to preprocess the image so that the model can produce accurate results [5].

Some filters, like Gaussian and median, are commonly used to remove noise from an image. Although these methods are simple, they are not enough to clean the image [6]. All these traditional methods frequently lose the important texture information and structural details, which are very important for identifying diseases on plant leaves [7],[8]. For mapping between noisy and clean images, Convolutional neural networks have achieved superior effectiveness in image restoration task with the development of deep learning. All techniques, such as DnCNN and U-Net, perform better than traditional filters, but they mainly focus on pixel-level error [9],[10]. Due to this, the output images can become too smooth and may lose

important details. For image enhancement and restoration, the most effective approach, which is Generative Adversarial Network (GAN), is widely used. A generator and a discriminator are trained at the same time during training, where the generator generates the synthetic images and the discriminator verifies whether the images created by generated are real or not. This process helps in producing more detailed and realistic results. As a result, GAN-based techniques have been widely used in several studies, such as low-light enhancement, image denoising, and super-resolution [11]. Maintaining structural similarity between the original image and the enhanced image using adversarial loss may not be enough [12].

To address this issue, adversarial loss, perceptual loss and structural similarity loss have all been combined in this research. Instead of comparing only pixel values, perceptual loss evaluates images using high-level features extracted by a pre-trained neural network such as VGG16. This enables the model to capture more meaningful details and produce results that appear more natural and visually appealing [13]. Perceptual loss, which is calculated using a pre-trained Convolutional network like VGG16, focuses on high-level feature representations rather than individual pixels. It preserves the features and texture for better visual quality of the image. Similarly, the SSIM loss also helps to maintain structure by measuring the similarity in edges, shapes, brightness and contrast. By using the combination of these loss functions, it produces enhanced images that are more realistic and accurate, which is very important to detect diseases accurately. The PlantVillage dataset is most commonly adopted in agricultural domains to detect diseases in plants.[14]. Although the dataset already contains cleaned images, in real-life images, they may be noisy and of low quality. So, preprocessing the image before classification is very important for an automated disease detection system. Although image enhancement techniques have already advanced, there is still limited research on a multi-loss GAN-based framework to enhance the image quality of apple plant leaf images under different noise conditions [15] [16].

This study aims to enhance the visual quality of apple leaf images using a multi-loss perceptual Generative Adversarial Network (GAN). The specific objectives are to evaluate the proposed framework for restoring apple leaf images affected by different types of noise, including Gaussian, salt-and-pepper, and speckle noise and to compare its performance with traditional filtering methods, including Gaussian and median filters as well as deep learning methods such as DnCNN and U-Net. Performance is evaluated using PSNR and SSIM.

1.1 Research Contribution

In this work, a multi-loss perceptual GAN-based image enhancement framework has been proposed to enhance the visual quality of Apple leaf images affected by noise. The proposed framework combines Adversarial Loss, L1 loss, a perceptual loss from a pre-trained VGG16 network, and SSIM loss to improve both pixel accuracy and perceptual accuracy. Several types of noise, which includes Gaussian noise, Speckle noise, Salt and Pepper noise, are used to represent the real-world agricultural imaging conditions. A class-wise performance evaluation is conducted on four apple leaf disease categories: apple scab, black rot, cedar apple rust, and healthy. Experimental comparisons with traditional filters (Gaussian and Median) and deep learning methods (DnCNN and UNet) demonstrate that the proposed framework achieves higher PSNR and SSIM values, demonstrating improved denoising performance. The ablation study further investigates the contribution of individual loss components to image restoration performance. The distinctive contribution of this work is the systematic evaluation of the combined multi-loss framework across different noise types and apple leaf categories together with baseline comparison and loss-component analysis.

2. LITERATURE REVIEW

J. Wu et al. [17] presented Plant PDN that learned the density gradient of the point cloud to solve the complex problem of the plant point clouds. The main parts of PDN included three units of point density feature, the UOF computations, and the density gradient (DG) measurement. PDN was checked on several plant species' point clouds with different noise levels to gauge its efficacy. The best denoising performance was obtained with this method, which outperformed state-of-the-art baseline methods by 7.6% to 19.3% under various Gaussian noise levels. For most noise classes, these denoising techniques outpaced its counterparts. Furthermore, this approach processed point clouds with low noise levels 0.5 times quicker than the traditional approaches, and point clouds with high noise levels 8.6 times faster.

F. Aziz, et al. [18] used the Deep Back-Projection Network (DBPN) method to enhance the image quality of plant images before classification. These

enhanced images were then provided to a Convolutional Neural Network (CNN), resulting in improved accuracy to identify different plant pests. The dataset consisted of 3,150 images representing nine distinct pest species. A four-layer convolutional neural network (CNN) combined with hyperparameter tuning was utilized for classification, resulting in an overall accuracy of 94%. Even though the mites and sawflies recall values were bit lower, the results stated that the model performed well across most categories of pest. The classification accuracy of image quality was increased by 10% when the researchers applied DBPN to increase the quality of image so that the enhanced images permitted the model to recognise the salient features impressively.

Z. Yang et al. [19] proposed an algorithm named as TFEGAN, which was designed to produce the images of infections in maize plants. TFEGAN introduced a Two-Way Extraction Module (TWEM) to enhance the model capabilities and effectively grasp features of maize leaf diseases. This framework improves low-light and noisy images of diseased maize leaves using pre-trained Generative Adversarial Networks (GANs). The low-light enhancement module used self-regularization and a dual discriminator to enhance the image quality while preserving the original visual details. The image denoising component combined an environmental encoder with a coordinated attention mechanism to detect noisy area and their surrounding details. This improved the image quality and helped the discriminator to evaluate them more accurately. The results showed that TFEGAN enhanced the quality of maize leaf images affected by noise and poor lighting, suggesting that the framework can be effectively applied in real-world scenarios.

N. S. Rani et al. [20] introduced the PLM-RES U-Net approach for improving the image quality of palm leaves. This lightweight architecture consists of encoding and decoding blocks connected through skip links. By preserving the text strokes and removing degradations, the model successfully improved the clarity of palm leaves. For evaluation, two palm leaves containing diverse textural backgrounds and different types of degradation were selected. PLM-Res-U-Net was trained for 50 epochs from scratch using three different sampling strategies with a learning rate of 1×10^{-8}. Two distinct benchmark datasets were also used to evaluate the efficacy of cutting-edge deep learning models, such as ResUnet, U-Net++, Pspnet, and Segnet. According to the investigation, the results produced by the suggested approach obtained a Dice score of 0.986, indicating both computational efficacy and generalizability.

Y.-C. Chen et al. [21] presented a multi-pooling and multi-scale ConvNet to detect illnesses of mango leaves from real-world photos. To maximize its performance, the suggested model underwent several operations. On its own, the model first attained a 95% accuracy rate. The use of the refined and the TL methods has made a huge difference; accuracy grew to 98.5%. Even with transfer learning or fine-tuning, the t-SNE outcomes of the classification were the ones used to evaluate the quality of the classifications. Hotspots of the pictures representing a class with CAM heat maps were also a means to check if the model paid attention to the right regions for disease detection. The analysis provided insightful details on the efficacy of the model in the case of an infection on the mango leaf.

V. G. Bhujade et al. [22] presented OptCF (Optimization Assisted Cascaded Filtering) to pre-process imagery. In the design of the new OptCFA approach, multilayered criteria for the GABF (Gaussian Amended Bilateral Filter) findings were calculated. The filter values for the first noise removal phase were optimised using the Amended Pelican Optimisation (AmPel). In the second stage of noise reduction an Extended Savitzky-Golay Filter (E_SGF) was applied to remove the noise from image. The proposed method used different techniques into one single approach. The performance of this method was compared with already existing several pre-processing filters.

V.-D. Hoang et al. [23] introduced a novel approach that integrated image filtering as a pre-processing step with Vision Transformer (ViT) to improve the learning process for identifying plant insect pests and diseases. The first step was Neural network-based image filtering in the suggested approach. The findings were then sent via a ViT module to generate feature maps, which were subsequently input into a multi-head classification network for final prediction. The proposed strategy used image filtering before the Vision Transformer (ViT) as a pre-processing step to enhance important features, instead of directly feeding the raw input images into the model. In addition, the proposed approach also employed element-wise amplification in the frequency domain rather than using spatial domain convolutional operations to save processing time. Testing findings indicated that the image filtering step did not substantially

increase the number of trainable parameters or the overall training time as compared to directly applying the Vision Transformer (ViT).

W. Chang et al. [24] developed an iterative ConvNet framework that was developed into a deep network using regularised Noise2Noise unsupervised learning technique. Using the half-quadratic splitting and gradient descent techniques, the iterative framework theoretically resolved the denoising issue while taking advantage of CNN's versatility and ability to exploit picture features. Furthermore, the regularised Noise2Noise introduced a constraint that measured the difference between the input data and its corresponding ground truth label. Three well-known denoising techniques were contrasted with the iterative framework using dynamic plant PET pictures. According to experimental data, the iterative approach significantly increased denoising efficiency, beating all examined approaches with a mean of 34.93 SNR and 62.51 CNR.

H. Jin et al. [25] developed a U-Net based generator by combining convolutions with residual blocks and reorganising techniques to counteract the loss of texture detail during image synthesis. To preserve more scale texture information, the generator simultaneously employed the concatenation (concat) method. The authors proposed a discriminator based on convolution blocks and capsule architecture to generate structurally complete grape images and to reduce mismatch between petiole and leaf regions. Convolutional operations were employed to extract general characteristics, and the capsule design was used to encode spatial data as well as the probability of spots. The authors compared the performance of GrapeGAN with WGAN and DCGAN by conducting experiments on same dataset. The experimental results indicated that GrapeGAN outperformed the other models, obtaining an FID value of 5.495 and a NIMA score of 4.937 ± 1.515.

S. O. Oppong et al. [26] developed an image pre-processing technique to enhance particular aspects of an image that were crucial for subsequent processing, including object recognition or edge detection. In image processing applications, preprocessing was important because it eliminated noise and highlighted or suppressed particular attributes. The proposed approach used CLAHE and homomorphic filtering as pre-processing techniques to improve the quality of medicinal plant leaf images. In order to separate the image from its background, this study also employed the OTSU global thresholding technique. The performance of the suggested approach was measured using performance metrics for both image enhancement and segmentation. In terms of image enhancement metrics as well as image segmentation metrics, the MATLAB experiments demonstrated that the suggested approach performed better than Contrast Limited Adaptive Histogram Equalization.

3. RESEARCH METHODOLOGY

This study proposes a multi-loss perceptual Generative Adversarial Network (GAN)-based framework for enhancing the visual quality of apple leaf images. An experimental approach was adopted to evaluate the proposed multi-loss perceptual GAN for apple leaf images enhancement. The study involved generating noisy images from input images. The dataset was derived from the PlantVillage repository, which includes four categories: Apple Scab, Black Rot, Cedar Apple Rust and Healthy. In this framework, the generator follows an encoder–residual–decoder architecture comprising six residual blocks with 3×3 convolution kernels and 2 stride, while the discriminator uses a PatchGAN structure of 70×70 and five convolutional blocks. Adversarial loss is based on training, a Perceptual loss based on VGG16, and an SSIM structural loss is used to balance denoising and texture preservation. Images were resized to 256×256 resolution. The complete image dataset was separated into three distinct groups in which 70% of images were used for training, 15% used for validation, and the remaining 15% for testing. To simulate realistic conditions, synthetic Gaussian Noise (σ = 15,25), salt-and-pepper noise (density = 0.05), and speckle noise were added to clean images before training. Comparative experiments and an ablation study were used to examine the performance of the proposed approach and the contribution of its loss components.

Figure 1 illustrates the methodology to preprocess the apple plant image using a Multi-loss GAN-based framework.

Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

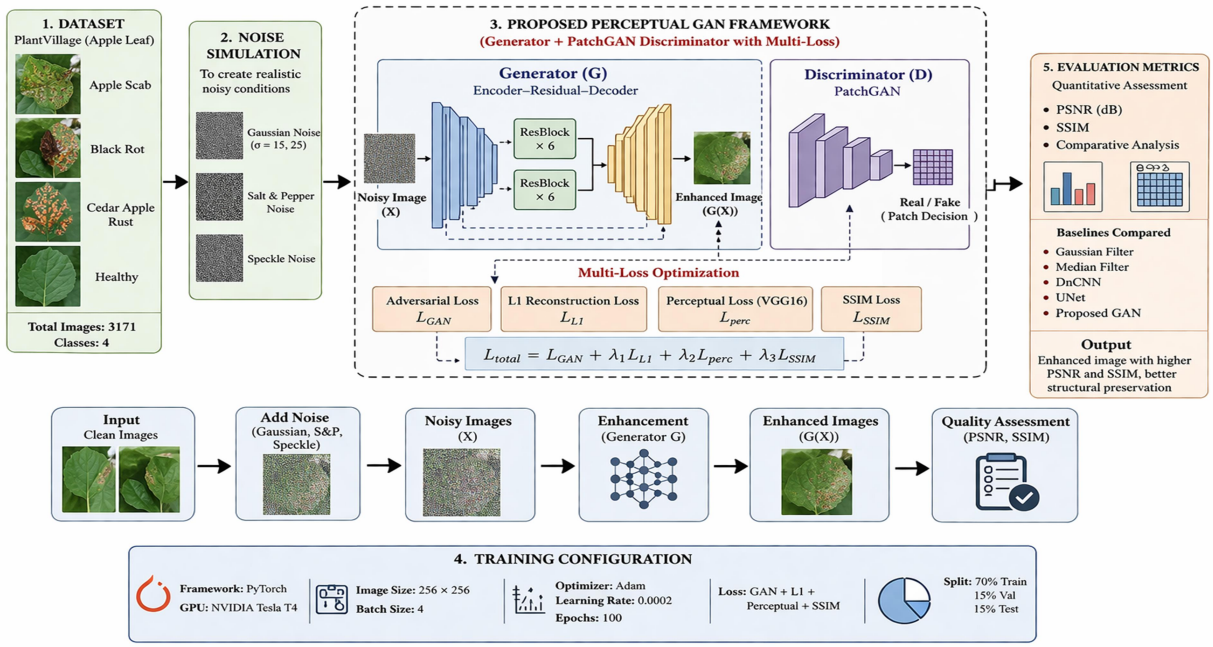


Figure 1. Methodology

The dataset description is mentioned below in Table 1:

Table 1. Dataset Description

Dataset	Classes	Images	Train	Val	Test
PlantVillage	Four	3171	70%	15%	15%

Algorithm: Proposed Pseudo-code

Algorithm: Multi-Loss GAN Image Enhancement

Input: Noisy plant images
Output: Enhanced images
Step 1: Data pre-processing
Load dataset
Resize images to 256×256
Add synthetic noise
Step 2: Model Initialization
Initialize Generator G
Initialize Discriminator D
Step 3: Training Process
for epoch = 1 to N do
Generate image G(x)
Compute adversarial loss
Compute L1 loss
Compute perceptual loss
Compute SSIM loss
Compute total loss
Update Generator
Update Discriminator
End

Output: High-resolution enhanced image.

3.1. Data Augmentation

During the training of a machine learning model, the main objective is to adjust its parameters so that input data can be accurately mapped to the expected outputs, while minimizing the overall loss function. Contemporary neural networks often contain millions of trainable parameters, which require large amounts of data to achieve reliable performance. Rather than collecting entirely new datasets, data augmentation is commonly used to address this challenge. Data augmentation involves generating modified versions of existing images through simple transformations such as flipping, rotation, scaling and shifting. When the existing dataset is limited, these transformations can be applied to all images, whereas for larger datasets, augmentation is typically performed dynamically on randomly selected samples during training. In addition to expanding the effective size of the dataset, data augmentation helps to prevent the model from learning irrelevant or misleading patterns, thereby improving its overall ability to perform well on unseen data.

3.1.1. Flip

Flipping an image along the horizontal and vertical axes produces new variations of the original image.

3.1.2 Random crop

During image cropping, random portions of the original images were selected and resized to match the dimensions of the input images.

3.2. Network Design Framework

3.2.1. Basic network layers

A Convolutional neural network's basic building blocks are the convolutional layers. Creating complete connections to every neuron is impractical while working with data that is high-dimensional. CNNs analyze images through identically sized filters, focusing on square patches. These filters go over the entire picture, creating the final activation maps. To calculate their outputs, neurons of the Convolutional Layer attach to particular areas of the input. Then they apply the filter weights to the corresponding input region by combining the values based on the filter size.

3.2.2. Rectified linear unit activation function

Neural network definitions usually favour non-linear activation functions. The non-linear activation process that the ReLU (Rectified Linear Unit) layer provides is shown in Figure 2.

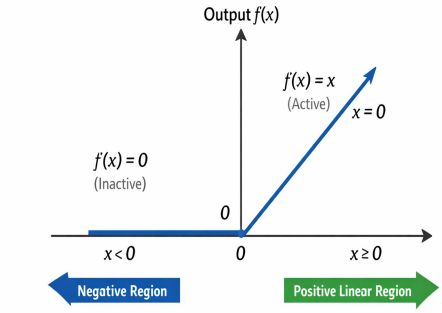


Figure 2. ReLU Activation

$$f(x) = max(0, x)$$

- The ReLU activation function uses zero as the threshold, producing an output of zero for negative inputs while preserving positive values unchanged.
- In the proposed generator, each convolutional layer is followed by batch normalization and a ReLU activation function, whereas the PatchGAN discriminator employs LeakyReLU activation. These activation functions enable the network to learn complex image features and improve training stability.

3.2.3. Batch normalization

In contrast to handmade features, neural networks may learn features on their own; nevertheless, data needs to be normalized before training. Input data are usually normalized because this helps the model to learn quickly during training. However, the distribution of input for each layer may change as model parameters are adjusted during training, which may slow down the training. To overcome this problem, normalisation is introduced into the network so that the model can learn quickly. This problem has been addressed by introducing batch normalisation, which normalises each mini-batch during training. It has been successful in training deep learning models. It helps maintain more stable feature distributions and supports affective training of deep neural networks. In the proposed network, Batch normalisation layers are used in the convolutional blocks of both generator and discriminator.

3.3. Other Parameters

The two main factors are the activation function used and network depth, which influence the training and performance of a neural network. These factors affect how efficiently the model reaches an optimal solution. Deep neural networks usually require more training time and computational power for training. So, their performance can be improved by selecting appropriate values for learning rate, weight initialization and optimizer.

3.3.1. Optimization

An objective function is used to measure how efficiently model performs. During training an optimization technique is used to minimize or maximize this function. The optimizer updates the network weights and bias parameters to minimize the difference between the predicted output and the actual output. In this study, an Adam optimizer is used because it is efficient, widely adopted and helps the model to train faster with stable convergence. An initial learning rate of 0.0002 is used and gradually reduced during training using a polynomial decay schedule.

3.4. Loss Function

The primary objective of AI-based model is to find suitable value for each neuron during the training to make accurate predictions. This is achieved by optimizing an objective function that reduces the mismatch between the predicted output and the expected result. The objective of the

training process is to reduce the value of the loss function so that the error becomes as small as possible. A well-designed loss function is generally expected to produce reliable results on different datasets. However, its overall performance is affected by several factors, including the characteristics of the data, the optimization strategy adopted during training, and the learning algorithm used in the model. Creating a connection between two training specimen sets is the primary purpose here: X and Y.$G: X \rightarrow Y$ and $F: Y \rightarrow X$ are two functions involved in this. D_Y and D_X, the corresponding discriminators, accompany these functions $X \sim Pdata(X) and\, Y \sim Pdata(Y)$ are the definitions of the data distributions.

3.4.1. Adversarial loss for G is calculated as shown in equation 1:

$$L_{\{GAN\}(G,D_Y,X,Y)} = E_{\{y \sim p_{\{data\}(y)}\}}[\log D_{Y(y)}] + E_{\{x \sim p_{\{data\}(x)}\}}[\log(1 - D_{Y(G(x))})] \quad (1)$$

G and D, respectively minimize and maximize loss. Therefore, the main goal is to optimize $min_G max_{D_Y} L_{GAN}$. A similar strategy is also used to obtain the mapping function F.

L1 Loss is defined as in equation 2:

$$LL1 = || G(x) - y || 1 \quad (2)$$

Perceptual Loss is computed as in equation 3:

$$Lperc = || \phi(G(x)) - \phi(y) || 2 \quad (3)$$

SSIM Loss is expressed as in equation 4:

$$LSSIM = 1 - SSIM(G(x), y) \quad (4)$$

Total Loss function is formulated as in equation 5:

$$L = LGAN + \lambda 1 LL1 + \lambda 2 Lperc + \lambda 3 LSSIM \quad (5)$$

The weighting coefficients were empirically selected as λ1=100, λ2=10 and λ3=5 to balance reconstruction accuracy, perceptual quality and structural similarity during optimization.

3.4.2. Perceptual similarity loss

In image generation tasks, even minor variations between two perceptually similar images can result in a high loss if a per-pixel loss function is used. Perceptual loss generates images with better visual quality than pixel-wise loss. This mechanism utilises the VGG16 architecture, previously trained on the ImageNet dataset, to retrieve key features. The VGG network extracts meaningful visual features from the original image and uses them during image generation. This helps the enhanced image closely match the original image in terms of preserving the structure and visual quality.

3.4.3 DnCNN baseline

DnCNN, proposed by Zhang et al., was adopted as a baseline denoising framework for evaluating the proposed model. The network follows a strategy of residual-learning denoising beginning with a convolutional layer containing 64 filters of size 3x3 which is followed by the ReLU activation function. The implementation used here consists of 15 convolutional [Conv (64, 3×3) + BatchNorm + ReLU] blocks, and a final Conv (3, 3×3) layer that predicts the noise residual; the denoised output is obtained as Input − Residual. The network has approximately 560K parameters and is trained with a plain MSE (pixel-wise) loss, without adversarial, perceptual, or SSIM terms.

3.4.4 UNet baseline

UNet is included as a second deep-learning baseline, using a standard encoder–decoder structure with skip connections. The implementation uses 3 downsampling/upsampling stages with a base filter width of 32 (32→64→128→256 channels), giving a bottleneck of 256 channels at 32×32 resolution. Like DnCNN, it is trained with a plain MSE loss only.

Both baselines were trained on the same noise-augmented data Gaussian, Salt-and-Pepper, and Speckle noise as the proposed model, for a fixed number of epochs.

4. RESULT AND DISCUSSION

4.1. Techniques and Tools

The proposed framework was implemented using a deep learning–based adversarial architecture comprising a generator–

discriminator network for image enhancement. A PatchGAN discriminator and a VGG16-based perceptual loss module were employed to improve texture realism and structural preservation. The model was optimized using the Adam optimizer with learning rate, and performance is evaluated using PSNR and SSIM metrics to quantify denoising and perceptual quality. Data augmentation techniques with flipping and spatial transformation were used to enhance the model generalization capability. The proposed framework was analyzed using PSNR and SSIM metrics.

The model configuration is presented in Table 2:

Table 2. Configuration Table

Parameter	Value
Epochs	100
Batch size	4
Learning Rate	0.0002
Optimizer	Adam
Early Stopping	Patience = 5 epochs, min delta = 0.001
Image Size	256×256
Framework	TensorFlow/Keras (Tf 2.20.0)
Loss Weights	Adversarial = 1, L1 = 100, Perceptual = 10, SSIM = 5
GPU	NVIDIA T4 × 2 (Kaggle)

4.1.1. Generator

The proposed framework uses an encoder–residual–decoder architecture with a convolutional neural network (ConvNet) which takes an input images of size 256 × 256 × 3. Initially, the input image passes through a convolutional layer with a stride of 2 to extract low-level features while reducing the image resolution. These features are further processed through several convolutional layers for downsampling and then through multiple residual blocks to learn important image features. The decoder then reconstructs the enhanced image using transposed convolution layers for upsampling. Finally, a convolution layer produces the enhanced output images.

4.1.2. Discriminator

The discriminator's function is a classification network that checks whether the images are real or not, which is generated by generator. During training, it receives both real and generated images and classifies whether each input is genuine or fake. Based on the feedback provided by the discriminator the generator learns to create the better images that look more like real images. In this study, the PatchGAN architecture is used for discriminator, which evaluates small patches on image instead of the whole image. This helps the model to preserve fine texture and structural details, producing more realistic images. Several Conv-Batch Norm-LeakyReLU blocks are used in the proposed network architecture. The discriminator follows a PatchGAN architecture composed of convolution-batchnorm-LeakyReLU blocks. Instead of classifying the whole image, PatchGAN evaluates local patches, which improves texture realism. The discriminator is trained with real images and the synthetic images generated by the generator, as shown in Figure 3.

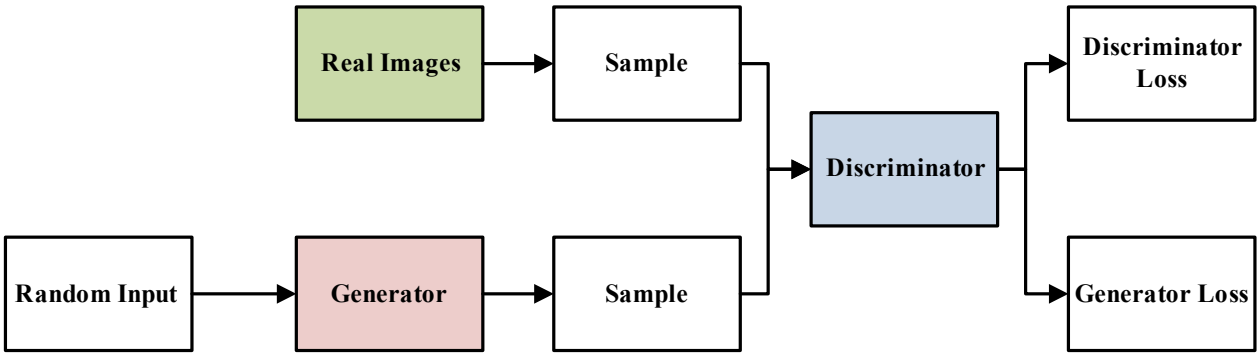


Figure 3. Representation of Generative Adversarial Network

4.1.3. Training Details

The deep learning (DL) packages in Python 3 are used to replicate the method that is being discussed. The proposed model was developed using TensorFlow/Keras and trained on the Kaggle platform with NVIDIA T4×2 GPUs.The MirroredStrategy framework was employed with mixed-precision (float16) training for efficient use of GPU and to accelerate the training. The model was trained for a maximum of 100 epochs using the Adam optimizer. The learning rate was 0.0002, and a batch size of 4 was used during training. An early

stopping strategy with a patience of 5 epochs was applied based on PSNR and terminated the training automatically when no further improvement was observed. It reduces overfitting while avoiding unnecessary computation.

4.2. Results

The proposed multi-loss GAN framework demonstrates superior performance compared with traditional filters and deep learning baselines. Under Gaussian noise (σ = 15,25), Salt-and-Pepper, and Speckle noise conditions, the proposed model achieves an average PSNR of 31.29 dB and SSIM of 0.82, which is higher than traditional filtering methods, DnCNN, and UNet baselines evaluated under the same protocol. Higher PSNR value indicates that the model removes noise more effectively, while a higher SSIM shows that image structure is preserved more accurately. The experimental results for different noise types are presented below in Table 3.

Table 3. Comparative Analysis of PSNR and SSIM Values Across Different Noise Types

Noise Type	GAN PSNR	GAN SSIM	DnCNN PSNR	DnCNN SSIM	UNet PSNR	UNet SSIM
Gaussian σ=15	31.15	0.8401	18.04	0.4221	24.86	0.6519
Gaussian σ=25	28.50	0.7157	16.85	0.2880	24.63	0.6284
Salt & Pepper	30.67	0.8588	16.07	0.2989	24.59	0.6276
Speckle	31.09	0.8407	18.28	0.4902	24.83	0.6505

Figure 4 illustrates a qualitative comparison of the image enhancement capabilities under various noise conditions, including Gaussian(σ = 15,25), Salt-and-Pepper, and Speckle noise. The enhanced images show that the proposed GAN-based framework effectively removes the noise while maintaining the original texture, edges, and structural characteristics of the apple leaf images.



Figure 4. Gan Image Enhancement-All Noise Types

4.2.1. Evaluation metrics

The proposed approach enhances the image quality. A variety of metrics are used to evaluate the image quality. To determine the quality of the generated image, two metrics are used. Below is an explanation of the PSNR and SSIM:

a). Peak Signal-to-Noise Ratio (PSNR): It can be described as the ratio of the highest signal strength to the power of the noise that degrades image quality. The way an image is represented is greatly impacted by this occurrence. The dynamic range of the various signals found in the applications is represented by the PSNR values using the logarithmic decibel scale.
The PSNR values are described by the MSV and are shown as follows:

$$PSNR = 10.log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20.log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) = 20.log_{10}(MAX_I) - 10.log_{10}(MSE) \quad (7)$$

b). Structural Similarity Index Ratio (SSIM): To determine the SSIM value of two windows x and y of the same size in N×N, the following formula can be used:

$$SSIM(x,y) = \frac{\left(\left(2\mu_{x\mu_y} + c_1\right)\left(2\sigma_{\{xy\}} + c_2\right)\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \quad (8)$$

Where the average value of x and y is represented byμ_x and μ_y. Also, the variances are denoted by σ_x^2 and σ_y^2 and the covariance of x and y is represented byσ_{xy}.

Table 4 compares the proposed framework across different classes of apple disease. Table 5 shows the average performance across all classes. Table 6 illustrates the contribution of each loss component to denoising performance and structural preservation. The results analysis is described below:

Table 4. Classwise PSNR/SSIM Evaluation with proposed framework

Class	Gaussian PSNR	Gaussian SSIM	Median PSNR	Median SSIM	DnCNN PSNR	DnCNN SSIM	UNet PSNR	UNet SSIM	GAN PSNR	GAN SSIM
Apple Scab	27.700	0.6454	26.673	0.5579	17.253	0.3622	23.892	0.6023	28.862	0.7539
Black Rot	23.807	0.5300	23.155	0.4162	17.373	0.5117	22.226	0.5012	25.945	0.7595
Cedar Apple Rust	30.454	0.7058	31.234	0.7246	17.451	0.2960	24.891	0.6494	32.394	0.8259
Healthy	27.673	0.6361	26.892	0.5651	17.418	0.3987	24.531	0.5902	29.244	0.7699

Figure 5 presents a classwise PSNR comparison across different categories of apple diseases including apple scab,black rot,cedar apple rust and healthy.The results shows that the proposed model consistently perform well in comparison of Gaussian filter, Median filter, DnCNN and UNet architectures by achieving the highest PSNR values across all the categories.

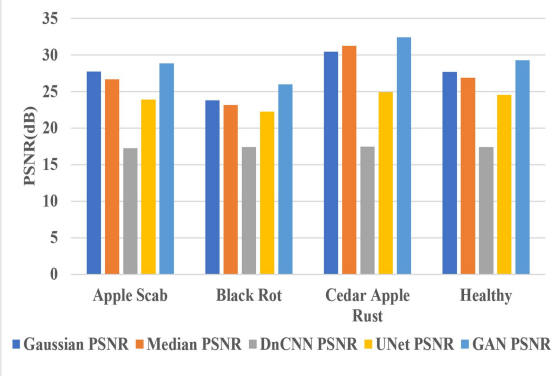


Figure 5. Class-wise PSNR Comparison- GAN vs DnCNN vs UNet vs Traditional Methods

Figure 6 shows that the proposed model achieves the highest SSIM value to preserve the structural similarity and texture details for all classes. These results demonstrate that the proposed model performs better than the traditional methods.

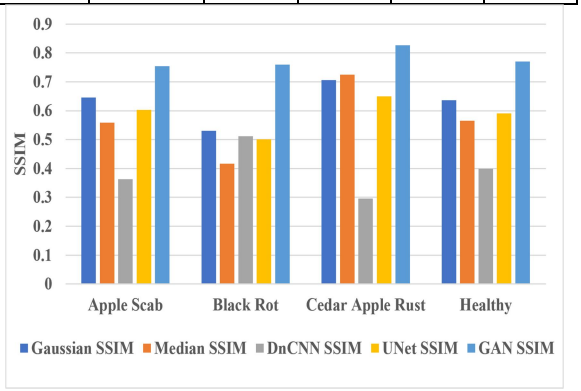


Figure 6. Class-wise SSIM comparison - GAN vs DnCNN vs UNet vs Traditional Methods

4.2.2. Result discussion

Table 4. presents the performance of the Gaussian Filter, Median Filter, DnCNN, UNet and the proposed framework. The comparison is based on Peak Signal-to-Noise Ratio (PSNR) and structural similarity index measure (SSIM) values for four categories: Apple scab, Black rot, Cedar apple rust and Healthy.

a). Apple Scab

The proposed framework achieved the highest PSNR value of 28.86, while the Gaussian filter achieved 27.70 and median filter showed results of

26.67). This result demonstrates that the proposed model performs well compared to other methods.

- The proposed framework achieved an SSIM value of 0.75, which is higher than the Gaussian filter (0.64) and median filter (0.56). This result shows that the model preserves the image structure more accurately while reducing noise.

b). Black Rot

- The proposed framework shows better performance than the Gaussian filter (23.80) and the median filter (23.16) in terms of noise suppression and image clarity, and achieves the highest PSNR value of 28.86.
- However, among the three approaches, the proposed approach has the best SSIM value (0.76), which shows a balance between structural similarity preservation and noise reduction. The SSIM value of the Gaussian filter is (0.53) and that of the median filter is (0.42).

c). Cedar Apple Rust

- For cedar apple images, the proposed framework achieved better noise reduction in comparison to the Gaussian filter and median filter, with a PSNR value of 32.39 compared to 30.45 and 31.23, respectively.
- However, the SSIM score of the proposed framework (0.83) in comparison with the Gaussian filter (0.71) and the median filter (0.72) indicates that some structural details were overlooked for better noise reduction.

d). Healthy

- The proposed framework achieved the highest PSNR value (29.24) compared to the Gaussian filter (27.67) and the median filter (26.89). This result shows the better performance of proposed model compared to traditional approaches.
- The proposed framework achieved a higher SSIM of 0.77 as compared to the Gaussian filter (0.64) and the median filter (0.57). This result shows better preservation of image quality and structure.

Class-wise results indicate that the proposed framework works best for the cedar apple rust, with a PSNR of 32.39 dB and SSIM of 0.83. The proposed multi-loss GAN framework exhibits better performance than traditional filtering methods and deep learning models. The combination of the perceptual loss and the SSIM loss helps to effectively remove noise while preserving important structural information. At same time, the adversarial loss improves texture quality and makes the image clearer. The results show that the proposed method achieves the best PSNR and SSIM values for all apple disease classes, which proves its effectiveness as a preprocessing step for automated plant disease detection systems.

4.2.3. Ablation study

This study provides a comprehensive quantitative evaluation of the proposed model in comparison with deep learning models such as DnCNN, UNet and traditional filters. The empirical results demonstrate that the proposed multi-loss GAN achieves the highest average performance of PSNR 31.29 dB and SSIM value of 0.82 across all evaluated classes. The higher PSNR and SSIM values indicate improved noise removal while preserving important structural details and edge information in apple plant leaf images.

Table 5. Average Performance Across All Classes

Method	PSNR (dB)	SSIM
Gaussian Filter	29.61	0.6958
Median Filter	29.86	0.6794
DnCNN	17.48	0.3399
UNet	24.95	0.6519
Proposed multi-Loss GAN	31.29	0.8184

Higher SSIM values indicate better structural preservation. The proposed model achieves both the highest PSNR and the highest SSIM among all compared methods in every class, indicating that in this framework, improved noise suppression does not come at the cost of structural fidelity.

The graphical representation of average performance with PSNR (dB) and SSIM across all class as shown in Figure 7 and Figure 8.

Figure 7 demonstrates that the proposed multi-loss GAN achieves the highest average PSNR value among traditional filtering techniques and deep learning models, including DnCNN and UNet, indicating its ability to generate enhanced images with lower noise levels.

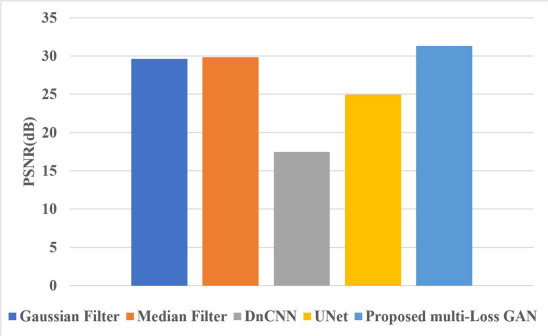


Figure 7. Average PSNR Performance Across All Classes

Figure 8 demonstrates that the proposed multi-loss GAN achieves the highest structural similarity score as compared to traditional filters and deep learning baselines.

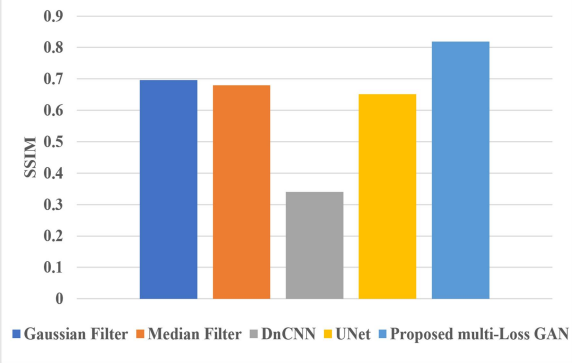


Figure 8. Average SSIM Performance Across All Classes

Table 6. Study of Loss Components

Model Configuration	PSNR (dB)	SSIM
GAN Only (Adversarial)	11.54	0.0500
GAN + L1	27.60	0.6952
GAN + L1 + Perceptual (VGG16)	27.55	0.7618
Proposed Model (GAN + L1 + Perceptual + SSIM)	29.99	0.8048

Table 6 summarizes the results of the ablation study for different loss functions.

To investigate the effect of each loss function on image enhancement, the performance was evaluated through an ablation study. The ablation experiments were conducted under a fixed 15-epoch training protocol to isolate the contribution of individual loss components, whereas the final proposed model was trained using the complete training strategy with early stopping and PSNR-based model selection. The absolute performance values of the ablation models should not be directly compared with the final performance reported for the fully trained proposed model. The proposed model, which combines adversarial, L1, perceptual and SSIM losses, achieves the highest PSNR of 29.99 dB and SSIM of 0.8048. These results indicate that combining these loss functions plays an important role in image enhancement and preserving the structural quality of images.

Figure 9 and Figure 10 compare the PSNR and SSIM values, highlighting the contribution of each loss function to the overall performance of the proposed model.

Figure 9 illustrates the effect of different framework configurations on model performance. It can be observed that the integration of GAN, L1 loss, and VGG16-based perceptual loss and SSIM-based perceptual loss significantly improves image quality in overall PSNR values while preserving important structural features, highlighting the contribution of each component.

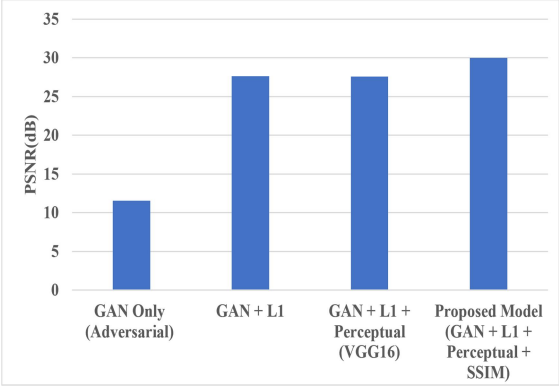


Figure 9. PSNR Performance across Each Loss Component

Figure 10 compares the ablation study results focused on the Structural Similarity Index (SSIM) across different configurations of the network. It is observed that the GAN model using only adversarial network achieves the lowest SSIM score, while the L1 loss, VGG16-based perceptual loss, and the final SSIM loss component stepwise enhance the structural fidelity of the reconstructed images. The proposed model achieves the highest SSIM value, demonstrating the effectiveness of combining these loss functions for preserving image structure.

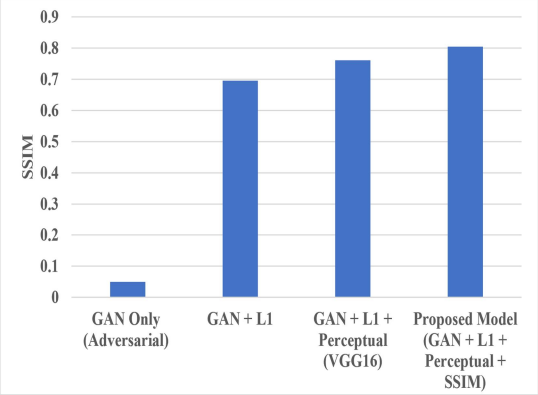


Figure 10. SSIM Performance across Each Loss Component

4.2.4 Statistical significance testing

To confirm that the PSNR improvements of the proposed framework over each baseline are statistically meaningful and not due to chance, a Wilcoxon signed-rank test was conducted on paired PSNR and SSIM scores from the same 30 noisy test images used in Table 6. As shown in Table 7 and Table 8, the proposed framework significantly outperforms all four baselines ($p < 0.05$ in every comparison)

Table 7. PSNR Statistical Testing

Comparison	PSNR Gain	W-Stat	p-value	Significant (p<0.05)
Proposed GAN vs Gaussian Filter	+1.68	55.0	< 0.001	Yes
Proposed GAN vs Median Filter	+1.43	47.0	< 0.001	Yes
Proposed GAN vs DnCNN	+13.81	0.0	< 0.001	Yes
Proposed GAN vs UNet	+6.34	0.0	< 0.001	Yes

Table 8. SSIM Statistical Testing

Comparison	SSIM Gain	W-Stat	p-value	Significant (p<0.05)
Proposed GAN vs Gaussian Filter	0.1226	4	< 0.001	Yes
Proposed GAN vs Median Filter				
Proposed GAN vs DnCNN	0.4785	0.0	< 0.001	Yes
Proposed GAN vs UNet	0.1665	0.0	< 0.001	Yes

5. CONCLUSION

This research proposes a perceptual GAN-based image enhancement framework to improve the quality of apple plant leaf images for automated plant disease detection. The proposed approach combines the adversarial loss, L1 reconstruction loss, perceptual loss derived from a pre-trained network VGG16, and structural similarity loss to maintain the balance between effectively removing image noise and better structural preservation. By combining adversarial, L1, perceptual, and SSIM losses, the framework removes noise more effectively and preserves important details better than traditional filters and other deep learning methods. Results on the PlantVillage dataset show that the proposed framework achieves better PSNR and SSIM when compared with the traditional baselines, highlighting its effectiveness as a preprocessing step for plant disease detection. The proposed framework achieved a PSNR of 31.29 dB and an SSIM of 0.82, indicating improved noise reduction and structure preservation. The results further showed that restoration performance decreased as Gaussian noise severity increased, with PSNR declining from 31.15 dB at σ =15 to 28.50 dB at σ =25, while SSIM decreased from 0.8401 to 0.7157.

These improvements were further confirmed to be statistically significant using the Wilcoxon signed-rank test ($p < 0.05$) compared with all evaluated baselines. Additionally, the analysis of class-wise categories of different diseases shows the strongest performance on the Cedar Apple Rust disease class. The experimental findings address the study objectives by evaluating the proposed framework under different noise conditions, comparing it with baseline methods using PSNR and SSIM, and also examining its performance across different apple categories. However, this study has some limitations. The evaluation was conducted using the PlantVillage dataset with synthetically added noise, and the proposed framework has not yet been tested on real-world apple leaf images which is captured

under different field conditions. Therefore, its performance under practical agricultural conditions remains to be investigated. Future work will focus on evaluating the proposed framework using real-world agricultural images and exploring its implementation in real-time applications. The proposed approach can also be applied to other crop datasets, providing an effective framework for automated crop monitoring.

Conflict of Interest: There is no conflict of Interest.

REFERENCES:

[1]. V. Sangeetha and D.J., "Enhancing Leaf Image Quality Using Deep Learning Denoising Techniques," *Nanotechnology Perceptions*, Vol. 20, No. S2, March 2024, pp. 308-316.

[2]. S. G and S. M. Vennila, "A Sequential Method for Enhancing the Image Quality of Rice Leaf Diseases by Employing Various Preprocessing Techniques," *proceedings of 2023 9th International Conference on Smart Structures and Systems (ICSSS)*, CHENNAI (India), 2023, pp. 1-7.

[3]. P.K. Shukla, K. Seethalakshmi, P. Hema and J. C. Musale, "An Effective Approach for Image Denoising Using Wavelet Transform Involving Deep Learning Techniques," *Proceedings of 2023 4th International Conference on Smart Electronics and Communication (ICOSEC),* Trichy (India), 2023, pp. 1381-1386.

[4]. S. k. Veldurthi and K. Samalla, "Image denoising analysis by Deep Learning Algorithms," *Proceedings of 2022 6th International Conference on Electronics, Communication and Aerospace Technology*, Coimbatore (India), December 1-3, 2022, pp. 1307-1311.

[5]. P. Limshuebchuey, R. Duangsoithong, and M. Saejia, "Comparison of Image Denoising using Traditional Filter and Deep Learning Methods," *17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Phuket (Thailand), June 24-27, 2020, pp. 193-196.

[6]. S. Ghose, N. Singh and P. Singh, "Image Denoising using Deep Learning: Convolutional Neural Network," *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence),* Noida (India), January 29-31, 2020, pp. 511-517.

[7]. P. Maharjan, L. Li, Z. Li, N. Xu, C. Ma, and Y. Li, "Improving Extreme Low-Light Image Denoising via Residual Learning," *Proceedings of 2019 IEEE International Conference on Multimedia and Expo (ICME)*, Shanghai (China, 2019, pp. 916-921.

[8]. S. Kumari and K. Kumar, "Analysis of Image Enhancement and Deep Learning Techniques," *Proceedings of 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi (India), 2023, pp. 1365-1370.

[9]. L. Ding and W. Du, "Application Analysis of Image Enhancement Method in Deep Learning Image Recognition Scene," *Proceedings of 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA),* Coimbatore, India, 2021, pp. 1459-1462.

[10]. C. Kang and S. Kang, "Self-Supervised Denoising Image Filter Based on Recursive Deep Neural Network Structure," *Sensors*, vol. 21, No. 23, November 2021, pp. 7827.

[11]. E. Ilesanmi and T. O. Ilesanmi, "Methods for image denoising using convolutional neural network: a review," *Complex & Intelligent Systems,* Vol. 7, No. 5, October 2021, pp. 2179-2198.

[12]. H. Patel, S. Patel, and A. Patel, "Analysis of Various Image Preprocessing Techniques for Denoising of Flower Images," *International Journal of Computer Sciences and Engineering,* Vol. 6, No. 5, May 2018, pp. 1111–1117.

[13]. N. A and B. N, "Impact of Image Pre-Processing in Image Quality and Segmentation for Banana Leaf Disease Prediction," *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)*, Theni, (India), 2023, pp. 1302- 1307.

[14]. P. Khandelwal, A. Shukla and M. Sain, "A Survey on Automated Disease Diagnosis and Classification of Herbal Plants Using Digital Image Processing," *Proceedings of 2023 6th International Conference on Information Systems and Computer Networks (ISCON),* Mathura (India), 2023, pp. 1-5.

[15]. A. Devaraj, K. Rathan, S. Jaahnavi and K. Indira, "Identification of Plant Disease using Image Processing Technique," *Proceedings of 2019 International Conference on Communication and Signal Processing*

(ICCSP), Chennai (India), 2019, pp. 0749-0753.

[16]. B. Deepa, "A Pre-Processing Approach for Accurate Identification of Plant Diseases in leaves," *Proceedings of 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT),* Mysuru, India, 2018, pp. 249-252.

[17]. J. Wu, L. Xiang, H. You, L. Tang, and J. Gai, "Plant-Denoising-Net (PDN): A plant point cloud denoising network based on density gradient field learning," *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 210, April. 2024 pp. 282–299.

[18]. F. Aziz, D. U. E. Saputri, S. Hadianti, D. N. Sulistyowati, S. Fauziah and N. A. Mayangky, "Image Enhancement with DBPN for Plant Pest Classification," *Proceedings of 2024 International Conference on Information Technology Research and Innovation (ICITRI),* Jakarta (Indonesia), 2024, pp. 208-211.

[19]. Z. Yang, S. Fang, and H. Huang, "Maize leaf disease image enhancement algorithm using TFEGAN," *Crop Protection*, Vol. 182, May 2024, pp. 106734.

[20]. N. S. Rani, T. M. Akhilesh, B. J. Bipin Nair, K. S. Koushik, and E. B. Smith, "PLM-Res-U-Net: A light weight binarization model for enhancement of multi-textured palm leaf manuscript images," *Digital Applications in Archaeology and Cultural Heritage*, Vol. 34, September 2024, pp. e00360.

[21]. Y.-C. Chen, J.-C. Wang, M.-H. Lee, A.-C. Liu, and J.-A. Jiang, "Enhanced detection of mango leaf diseases in field environments using MSMP-CNN and transfer learning," *Computers and Electronics in Agriculture,* Vol. 227, December. 2024, pp. 109636.

[22]. V. G. Bhujade, V. Sambhe, and B. Banerjee, "Digital image noise removal towards soybean and cotton plant disease using image processing filters," *Expert systems with applications*, Vol. 246, July 2024, pp. 123031.

[23]. V. -D. Hoang and T. -A. Pham, "Fusion of ViT Technique and Image Filtering in Deep Learning for Plant Pests and Diseases Recognition," *Proceedings of 2023 International Conference on System Science and Engineering (ICSSE), Ho Chi Minh* (Vietnam), 2023, pp. 438-443.

[24]. W. Chang, N. D'Ascenzo, E. Antonecchia, F. Zhou and Q. Xie, "A Convolutional Neural Network Based Iterative Framework for Plant PET Image Denoising," *Proceedings of 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* (Italy), 2022, pp. 1-5.

[25]. H. Jin, Y. Li, J. Qi, J. Feng, D. Tian, and W. Mu, "GrapeGAN: Unsupervised image enhancement for improved grape leaf disease recognition," *Computers and Electronics in Agriculture*, Vol. 198, July 2022, pp. 107055.

[26]. S. O. Oppong, F. Twum, J. B. Hayfron-Acquah and Y. M. Missah, "Hybrid Approach to Medicinal Plant Leaf Image Enhancement Using Contrast Limited Adaptive Histogram Equalization and Homomorphic Filter," *Proceedings of 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET),* Cape Town (South Africa), 2021, pp. 1-6.