

Research Article

A Custom Backbone UNet Framework with DCGAN Augmentation for Efficient Segmentation of Leaf Spot Diseases in Jasmine Plant

Shwetha V. D, Arnav Bhagwat, Vijaya Laxmi D, and Sakshi Shrivastava

Department of Electrical and Electronics Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, Karnataka, India

Correspondence should be addressed to Vijaya Laxmi; vijaya.laxmi@manipal.edu

Received 28 August 2023; Revised 8 February 2024; Accepted 12 February 2024; Published 27 February 2024

Academic Editor: De Rosal Ignatius Moses Setiadi

Copyright © 2024 Shwetha V. et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Leaf blight spot disease, caused by bacteria and fungi, poses a considerable threat to commercial plants, manifesting as yellow to brown color spots on the leaves and potentially leading to plant mortality and reduced agricultural productivity. The susceptibility of jasmine plants to this disease emphasizes the necessity for effective detection methods. In this study, we harness the power of a deep convolutional generative adversarial network (DCGAN) to generate a dataset of jasmine plant leaf disease images. Leveraging the capabilities of DCGAN, we curate a dataset comprising 10,000 images with two distinct classes specifically designed for segmentation applications. To evaluate the effectiveness of DCGAN-based generation, we propose and assess a novel loss function. For accurate segmentation of the leaf disease, we utilize a UNet architecture with a custom backbone based on the MobileNetV4 CNN. The proposed segmentation model yields an average pixel accuracy of 0.91 and an mIoU (mean intersection over union) of 0.95. Furthermore, we explore different UNet-based segmentation approaches and evaluate the performance of various backbones to assess their effectiveness. By leveraging deep learning techniques, including DCGAN for dataset generation and the UNet framework for precise segmentation, we significantly contribute to the development of effective methods for detecting and segmenting leaf diseases in jasmine plants.

1. Introduction

In recent times, there has been a notable rise in the occurrence of plant diseases caused by microorganisms such as bacteria, fungi, and viruses [1] in plants, animals, and humans. These infections present a significant threat to plants throughout different stages of agricultural production, ultimately resulting in reduced plant yield [2, 3]. The consequences of these diseases have far-reaching implications for human dependence on agriculture, encompassing vital necessities such as food, shelter, and clothing. This is especially notable in low-income countries [4, 5]. Jasmine plants, commonly cultivated in coastal regions of Southeast Asia [6], are known to be vulnerable to a range of leaf diseases, including Alternaria leaf blight spot [7]. This disease exhibits initial signs characterized by yellow patches

with dark brown stains surrounded by yellow rings [8]. As the disease progresses, the spots grow larger, spreading across a significant portion of the leaves, eventually leading to blight. Notably, concentric rings can be observed within the lesions, and the disease also affects the stem, petiole, and flowers [9]. The final two stages are critical for disease detection, with the transition from yellow to brown in leaves referred to as the "brown stage." Subsequently, the "final stage" involves the maximum coverage of brown spots on the leaf, leading to the plant's fatality. Identifying these crucial stages is essential, as timely action can be taken to address the issue effectively [10]. Various datasets, including cassava, tomato, cotton, and tobacco, have been utilized to report several CNN-based approaches for detecting plant leaf diseases [11-16]. Nevertheless, the research on jasmine plant leaf spot disease detection remains limited. The scarcity

of suitable datasets poses a significant challenge in the development of CNN-based detection algorithms capable of detecting various stages of jasmine plant leaf spot disease. In the past, several segmentation and morphological methods have been reported for grapes and other leaves [17–19]. However, there is a need for a semantic segmentation method specifically designed to extract the leaf spot features of jasmine plants.

The contribution of this work is as follows:

- (1) This study introduces a novel leaf image augmentation strategy employing DCGAN, resulting in the generation of an expanded dataset with 10,000 synthetic jasmine plant images. Diverging from conventional methods, our approach exhibits superior scalability and image quality. Comparative analyses underscore the effectiveness of our DCGAN-based augmentation, positioning it as an advanced and impactful contribution in dataset expansion techniques.
- (2) Our proposed methodology for identifying the "brown stage" and "final stage" of leaf spot disease in jasmine leaves introduces an original approach using UNet-based semantic segmentation, specifically ResUNet with a custom CNN backbone. Outperforming traditional methods, our approach achieves heightened accuracy and efficiency. Comparative evaluations highlight its superiority in disease stage recognition, marking it as a significant advancement over current identification techniques.
- (3) This research explores various semantic segmentation techniques and pretrained CNN backbones for leaf spot identification. The proposed model, boasting an mIoU of 0.95, surpasses alternative segmentation methods, providing a more precise and reliable classification of disease stages. Comparative assessments underscore the effectiveness of our model in capturing nuanced details, establishing it as a leading solution in the field of leaf spot identification.

Section 2 discusses recent works proposed for leaf disease detection in the literature. Sections 3 and 4 introduce the proposed model and present the experimentation results. Finally, Section 5 provides the study's conclusion.

2. Related Works

In recent years, remarkable progress has been achieved in detecting diseases from leaf images. These approaches can be broadly classified into two main groups: traditional detection methods and deep learning-based detection methods. In addition, this section will delve into various augmentation techniques employed to expand the dataset.

2.1. Traditional Detection Methods. A leaf stage recognition system was developed by incorporating K-mean clustering approaches [20] to focus on specific areas that play a crucial role in leaf disease detection. Geetha et al. [21] proposed four

preprocessing steps to reduce noise in the leaf image dataset. Furthermore, Annabel et al. [22] utilized traditional detection techniques, including the K-nearest neighbor (KNN) algorithm, to classify plant leaves based on morphological features such as color, intensity, and size. For color analysis, Narmadha and Arulvadivu [23] reported the conversion of primary leaf colors into LAB color space and employed clustering algorithms. In the work of Gupta et al. [24], an automated strategic removal of the background was performed, and the desired diseased portion was extracted for mildew disease detection from cherry leaves. In addition, Kurmi and Gangwar [25] employed color transformation for seed region identification in leaf analysis. Literature [26] describes several methods used in precision agriculture. However, achieving high classification accuracy in leaf spot detection has proven to be a challenge for most machinelearning approaches. In this context, various literature studies have explored deep learning methods for leaf morphology identification, which will be discussed in the following subsection.

2.2. Deep Learning-Based Detection Methods. The detection of tomato plant disease through deep learning-based segmentation has been previously explored in the works of Shoaib et al. [27] and Agarwal et al. [14]. Another study by Xie et al. [28] proposes a technique utilizing a fully convolutional neural network (FCN) for the segmentation of maize leaf disease. Prior studies have presented deep neural network-based classification models for plant diseases. Hridoy et al. [29] employed a deep neural network approach to identify betel leaf diseases. Kaur et al. [30] introduced a semiautomatic CNN model for soybean leaf disease classification. Haridasan et al. [31] developed a CNN-based detection model for paddy leaf diseases. Furthermore, Alsabai et al. [32] proposed a hybrid deep learning approach, incorporating improved Salp swarm optimization, for the multiclass detection of grape diseases. Shoaib et al. [33] focused on addressing the challenge of accurately identifying diseased spots amidst complex field conditions. They trained their proposed system using a dataset comprising crop leaf images with both healthy and diseased sections. The algorithm's performance was evaluated using metrics like accuracy and intersectional union ratio (IoU) to segment lesion regions from the images precisely. In a different context, Lin et al. [34] propose a semantic segmentation model that employs convolutional neural networks (CNNs) to recognize and segment powdery mildew in individual pixel-level images of cucumber leaf. Their approach achieved a joint intersection ratio score of 79.54% and a dice accuracy of 81.5% based on 20 test images. Finally, Soliman et al. [35] presented work that proposed employing deep learning techniques to detect plant lesions by extracting hidden patterns from plant leaf disease. Despite the availability of plant disease datasets such as the PlantVillage dataset [36], the AgriVision collection [37], and the Plant Disease Identification dataset, implementing CNN-based detection algorithms requires large datasets. Past research such as Kumar et al. [38], Sladojevic et al. [39], and He et al. [40]

examined the significant consequences of crop diseases on food security and economic losses in India's agriculturereliant rural regions. It underscored the requirement for innovative computer vision methods to autonomously identify and categorize these diseases, with studies showing diverse approaches and notable successes, especially in deep learning-based techniques. The following subsection will discuss various augmentation techniques used to increase the plant disease dataset.

2.3. Various Augmentation Techniques. Rotate, flip, shift, and scale techniques were employed to augment the leaf dataset [41, 42]. In addition, a combination of rotation and shift was explored to increase the dataset further [43]. By utilizing GAN-based augmentation, the dataset's enlargement resulted in a 20% increase in classification accuracy [44]. Another study employing a detection framework saw an improvement of 7.4% in classification accuracy [45]. Data augmentation is of utmost importance in efficiently enhancing the dataset for detection and classification approaches. A novel augmentation method will be detailed in the next section.

3. Methodology

This research focuses on enhancing the detection of disease spots on Jasmine leaves, particularly brown-stage and finalstage spots that are challenging to identify accurately. To overcome limited data, a GAN-based augmentation model is employed to expand the leaf dataset used for segmentation. The study explores the effectiveness of UNet, WUNet, U2Net, and ResUNet architectures in this context while also investigating different segmentation backbones to optimize the detection performance, as shown in Figure 1.

3.1. Dataset. In this study, Figure 2 presents image samples depicting different stages of diseased leaves. The dataset for these images was collaboratively developed with experts from Krishi Vigyan Kendra, Karnataka, India, who utilized digital cameras to capture a total of 1000 images. These images cover four stages of Alternaria leaf blight spot disease, including 450 images for the brown stage, as illustrated in Figure 2(a), where the blight spot covers a quarter of the leaf, and 550 images for the final stage, as depicted in Figure 2(b), which covers a larger area of the leaf with blight spots. To enhance the dataset, generative advisory networkbased augmentation techniques were employed. It is worth mentioning that the early stage of leaf spot disease was not considered in this study. Instead, the focus was on the later stages of brown stage and final stage, which are crucial for understanding disease progression. Further details regarding the application of augmentation techniques can be found in the subsequent section.

3.2. Data Augmentation Using DCGAN. Ian Goodfellow and his colleagues pioneered the creation of DCGAN (deep convolutional generative adversarial network) in 2015 [46].

The DCGAN's conditional input allows the generator to produce synthetic samples based on specified conditions. Convolutional neural networks (CNNs) are widely adopted in GANs, particularly for image processing, delivering remarkable results in various computer vision tasks. The generator takes a compressed representation of the training image set, consisting of 1000 images, and generates new images with a resolution of 256×256 pixels in RGB format. A 100-dimensional vector with random values between 0 and 1 augments the input image generation process. To achieve the desired resolution for generated images, the generator incorporates convolutional transpose layers, while the discriminator relies on two convolutional layers with 256 neurons each and LeakyReLU activation. The training process utilizes the SGD optimizer and focuses on minimizing the Ladv loss. The aim is to prevent the discriminator from accurately distinguishing fake images. During training, the GAN model aims for a Frechet inception distance (FID) score below 15 as a performance measure. Training involves 200 epochs with a batch size of 32, and the similarity between generated and template images is evaluated using the structural similarity index (SSIM) and signal-to-noise ratio (SNR) metrics. The overall methodology is illustrated in Figure 3.

3.3. Proposed Segmentation Model for Jasmine Plant Leaf Disease Detection. Segmentation of images is a crucial aspect of computer vision, wherein an image is divided into different regions and assigned specific class labels to create a map that provides information about each pixel of the image. A custom backbone based on MobileNetV4 is integrated into the UNet-based architectures to detect critical types of leaf spot diseases in jasmine plants. Integrating MobileNetV4 into various UNet frameworks, including UNet, WUNet, U2Net, and ResUNet, involves utilizing it as the encoder component. It replaces conventional convolutional layers, seamlessly integrating its efficient multiscale feature extraction capabilities. This reduces model parameters significantly compared to traditional architectures while preserving the decoder's precision, thereby reducing computational load. Our choice of these semantic segmentation models was driven by specific strengths: UNet's efficiency in preserving structural elements, U2Net's lightweight design for real-time segmentation without compromising precision, WUNet's adaptability to resource constraints, and ResUNet's balance between accuracy and efficiency. We conducted experiments to determine the optimal model for jasmine leaf disease detection. Mobile-NetV4 is a big architecture. Its implementation at the encoder part is shown in Figure 4 and detail information is provided in Table 1. The proposed CNN network uses a novel computation technique called depthwise separable convolution, which bears similarities to traditional convolution but involves a two-stage calculation process. Unlike the conventional approach, where a single convolutional calculation is performed per layer, depthwise separable convolution divides the process into two phases. The first stage encompasses a separate convolution operation with



FIGURE 1: Overall methodology for jasmine leaf disease segmentation.



FIGURE 2: Jasmine plant leaf spot disease dataset.

a 3×3 kernel for each input channel, followed by batch normalization and activation. This phase is referred to as depthwise convolution. The second stage involves further processing the output channels from the depthwise convolution using a 1×1 pointwise convolution. This pointwise convolution is applied across all depthwise convolution output channels. Overall, depthwise separable convolution significantly enhances computational efficiency by reducing the computational load. Table 1 presents a comprehensive description, providing details about convolution layers 1 and 2. For clarity, in this study, we denote the depthwise convolution layer as "conv_dw" and the pointwise convolution layer as "conv_pw." This process is repeated for layers 3 to 6, and the final convolutional layer is identified as "layer 7." Notably, Table 1 showcases the parameter reduction at each sequential layer, highlighting the achieved computational efficiency.

UNet is an encoder-decoder model comprising two distinct networks, namely, the contraction network and the expansion network. The contraction network, referred to as the encoder, is responsible for extracting pertinent features from the leaf image [47]. On the other hand, the expansion



FIGURE 3: Data augmentation using DCGAN approach.



FIGURE 4: Flowchart used for the leaf spot stage detection in jasmine plant.

TABLE 1: Overview of the network framework, featuring Mobile-NetV4 large as the backbone architecture.

Name	Layer	Feature maps	Parameter	
Input_1	InputLayer	3	0	
conv1	Conv2D	32	864	
conv1_bn	BatchNormalization	32	128	
conv1_relu	ReLU	32	0	
conv_dw_1	DepthwiseConv2D	32	288	
conv_dw_1_bn	BatchNormalization	32	128	
conv_dw_1_relu	ReLU	32	0	
conv_pw_1	Conv2D	64	2048	
conv_pw_1_bn	BatchNormalization	64	256	
conv_pw_1_relu	ReLU	64	0	
conv_pad_2	ZeroPadding2D	64	0	
conv_dw_2_bn	BatchNormalization	64	256	
conv_dw_2_relu	ReLU	64	0	
conv_pw_2	Conv2D	128	8192	
conv_pw_2_bn	BatchNormalization	128	512	
conv_pw_2_relu	(ReLU)	128	0	
conv_dw_3	DepthwiseConv2D	128	1152	
conv_dw_3_bn	BatchNormalization	128	512	
conv_dw_3_relu	ReLU	128	0	
conv_pw_7	Conv2D	128	11384	
conv_pw_7_bn	BatchNormalization	128	512	
conv_dw_7	DepthwiseConv2D	1024	9216	
conv_dw_7_bn	BatchNormalization	1024	4096	
conv_dw_7_relu	ReLU	1024	0	
conv_pw_7	Conv2D	1024	18576	
sequential_1	Sequential	128	19728	
dense_1	Dense	5	645	

network, known as the decoder, reconstructs the segmentation map using the encoded features obtained from the encoder [48]. The earlier proposed UNet model is designed with four blocks to extract spatial features from the image. Each block consists of two convolution layers with ReLU activation and a max-pooling layer, downsampling the input by a factor of 2 [49]. The proposed UNet model extends beyond the four blocks and includes additional convolution layers activated by the leaky ReLU activation function. These enhancements, along with the custom backbone, contribute to capturing low-level features essential for the leaf spot disease model. The overall network architecture is shown in Figure 5.

3.3.1. Comparison of Different UNet-Based Segmentation Approaches. In this research, we investigate and compare several UNet-based segmentation architectures, each offering distinctive design characteristics and advantages for leaf spot detection tasks. The UNet architecture features a symmetric encoder-decoder design, skillfully utilizing skip connections to concatenate feature maps from the encoder with corresponding decoder layers. This approach effectively preserves high-resolution information during the decoding process. WUNet, an extension of UNet, is commonly referred to as wide UNet [50]. It enhances the architecture by widening the convolutional layers with an increased number of channels. This design choice significantly improves the model's capture of contextual information, potentially leading to enhanced segmentation performance. U2Net,

a recent and specialized architecture, is purposefully tailored for salient object detection. Inspired by UNet, U2Net incorporates several improvements, including additional branches and attention mechanisms. These attention modules are crucial in highlighting salient features, rendering U2Net highly suitable for tasks requiring precise boundary detection. On the other hand, ResUNet, also known as residual UNet, is a variant of UNet that integrates residual connections derived from the ResNet architecture. By leveraging these residual connections, the model efficiently facilitates gradient flow during training, enabling the effective training of deeper architectures. This capability makes ResUNet [51] particularly well-suited for handling more complex segmentation tasks. Through a comprehensive evaluation and comparison of these U-Net-based models, we aim to gain valuable insights into their individual performance, strengths, and suitability for a diverse range of semantic segmentation challenges.

3.3.2. Assessing the Different Backbone Architectures for Segmentation Models. To assess the semantic functionality, all the models are trained with various pretrained networks, such as ResNet, EfficientNet, VGG16, and VGG19, as backbones. In addition, a custom backbone is utilized for the evaluation. The backbone models are employed in the encoder part of the various semantic segmentation models such as UNet, WUNet, U2Net, and ResUNet. Initially, a semantic segmentation model with baseline backbones is assessed. Subsequently, one-by-one, pretrained models and custom backbone CNN networks were used to assess the model. For this study, UNet with skip connection is employed as the segmentation model.

3.3.3. Steps Used for Leaf Spot Disease Detection Using the Custom Backbone UNet Framework

- (1) Prepare the leaf dataset using DCGAN augmentation
- (2) Train the chosen segmentation model with input images and corresponding masks, considering performance metrics like mIoU, Dice, and pixel accuracy calculated using equations (1)–(4)
- (3) Train and fine-tune the segmentation model parameters.
- (4) Iteratively train the model until achieving a satisfactory training and validation accuracy curve, otherwise, repeat Step 3
- (5) Deploy the model for testing on a real image test set
- (6) Output the segmentation results to identify the brown stage and final stage of the leaf

The overall flowchart of the leaf spot disease detection is illustrated in Figure 4.

3.4. Training Details. The segmentation task involves using an augmentation model to generate a total of 5000 images for each type. During the training process, the loss function L_{adv} of the DCGAN is fine-tuned. The GAN model is trained



FIGURE 5: Proposed UNet segmentation model with a custom backbone.

for 100 epochs, with 5000 iterations in total. To initiate the process, the initial value of the loss function's λ is set at 0.1 [52]. As iterative training progresses, the lambda value is updated to 0.01. Following the iterative training process, these augmented images, along with their corresponding masks, are passed to the segmentation block. The study employs various models, including UNet, WUNet, U2Net, and ResUNet, all utilizing a 3×3 kernel. Each model undergoes 300 training epochs.

3.5. Hyperparameter Tuning of Segmentation Models. The UNet, WUNet, U2Net, and ResUNet models were trained using various backbones, each with a batch size of 32, over 300 epochs. Here, the batch size determines how many samples are processed before updating the model's parameters, while epochs represent the number of complete passes over the training data. The learning rate, a critical hyperparameter, was set to 0.0001 to balance learning speed and convergence. The Adam optimization method was used for model compilation, activating all convolutional layers with the ReLU activation function using a 3×3 kernel. An early-stop mechanism based on validation performance was implemented during training to prevent overfitting. In our proposed segmentation model with a custom backbone, the Adam optimizer was utilized with a learning rate of 0.001 and a batch size of 32. The model underwent training for 100 epochs with the ReLU activation function, following an iterative process to determine the optimal parameter settings.

3.6. Evaluation Metrics. To assess the efficacy of the DCGAN augmentation method, we analyze the similarity between the synthesized images and the template images. This evaluation

utilizes well-established similarity metrics, including the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [53]. These metrics offer valuable insights into the degree of resemblance between the generated images and the target images, enabling a thorough evaluation of DCGAN augmentation model performance. The segmentation tasks are evaluated based on the calculated metrics, which are determined by the following equations:

pixel accuracy =
$$\frac{TP + TN}{TP + TN + FP + FN}$$
, (1)

$$IoU = \frac{TP}{TP + FP + FN},$$
 (2)

$$m\text{IoU} = \frac{1}{N_{\text{class}}} \sum_{i=1}^{N_{\text{class}}} \frac{\text{TP}(i)}{\text{TP}(i) + \text{FP}(i) + \text{FN}(i)},$$
 (3)

$$Dice = \frac{2TP}{2TP + FP + FN}.$$
 (4)

4. Result and Discussion

The FID score serves as a widely adopted metric for assessing the fidelity of generated images in relation to real images from a given dataset. In this context, the objective is to train the DCGAN in a manner that ensures the FID score remains stable and lies within the specified range of 13 to 15, as depicted in Figure 6. Sustaining the FID score within this designated range signifies that the generated images closely mirror the characteristics of the real images in the dataset, showcasing a notable level of visual quality and diversity.



FIGURE 6: FID score obtained for DCGAN plant leaf image generation.



FIGURE 7: Sample images obtained using DCGAN image augmentation on brown-stage images.



FIGURE 8: Sample images obtained using DCGAN image augmentation on final-stage images.



FIGURE 9: Segmentation results with the different semantic segmentations with the proposed custom backbone.

This consistency in the FID score reflects the success of the training process in achieving realistic and diverse image generation.

In this study, we shared the outcomes of our assessment of image generation models. The images generated during the brown stage garnered an SSIM score of 0.84 ± 0.02 and a PSNR score of 23 ± 1.7 . In contrast, images produced during the final stage achieved an SSIM score of 0.90 ± 0.172 and a PSNR score of 25 ± 2.4 . The SSIM score serves as an indicator of the structural similarity between the generated



FIGURE 10: Training accuracy of the segmentation models (a) ResUNet, (b) WUNet, (c) U2Net, and (d) UNet with various backbones.

images and real images, while the PSNR score reflects the level of fidelity and noise present in the generated images. Higher SSIM scores suggest a closer resemblance to real images, whereas higher PSNR scores indicate enhanced image quality. Upon comparing both sets of generated images, it was observed that final-stage images attained superior scores in both SSIM and PSNR metrics. This implies that they exhibit better similarity and higher quality when compared to real images. These findings offer valuable insights into the performance of our image generation models and underscore the superior capabilities of the final stage in producing high-quality images. Figure 7 visually presents brown-stage-generated images obtained using DCGAN, showcasing the visual aspects of our research outcomes.

In Figure 8, the final-stage-generated images produced by DCGAN are displayed. The analysis reveals that the brown-stage-generated images outperform the final-stage images both qualitatively and quantitatively.

Figure 9 shows the segmentation results of four models: UNet, WUNet, U2Net, and ResNet, each equipped with a custom backbone. The comparison indicates that the UNet model with the custom backbone outperforms the WUNet, U2Net, and ResUNet models with the same custom backbone regarding segmentation performance.

In Figure 9, we present the training accuracy of several segmentation models, each integrating unique pretrained

CNN networks in conjunction with our novel custom backbone. Notably, the UNet segmentation with the custom backbone emerges as particularly effective for detecting leaf spot diseases. The training accuracies depicted in Figures 10(b) and 10(c) display variations throughout the epochs. Concurrently, Figure 10(a) illustrates the performance of ResUNet, which exhibits similar fluctuations. However, a comparative analysis in Figure 10(d), representing the proposed UNet with the custom backbone, reveals that the latter demonstrates superior performance. This suggests that our innovative custom backbone enhances the UNet segmentation model's efficacy in comparison to other configurations, underscoring its potential for accurate and robust leaf spot disease detection. Further details and insights into these results will be discussed in the subsequent sections.

In Table 2, performance metrics, namely, mean of IOU referred as mIoU and Dice coefficient referred as Dice, are shown for the two-stage leaf disease classification employing various backbone CNN networks. The results demonstrate that the proposed custom backbone combined with the UNet semantic segmentation yields superior outcomes. This innovative framework successfully extracts the low-level features for leaf spot disease detection, enhancing the accuracy of the classification process.

Table 3 presents the outcomes of the evaluation conducted to determine the most suitable backbone for the

Backbones	Stage	mIoU			Dice				
		UNet	WUNet	U2Net	ResUNet	UNet	WUNet	U2Net	ResUNet
ResNet	Brown stage	0.68	0.61	0.62	0.72	0.73	0.71	0.68	0.69
	Final stage	0.78	0.72	0.72	0.73	0.65	0.68	0.71	0.70
EfficientNet	Brown stage	0.86	0.72	0.88	0.82	0.77	0.75	0.72	0.71
	Final stage	0.81	0.68	0.62	0.82	0.78	0.76	0.78	0.74
VGG16	Brown stage	0.83	0.84	0.85	0.79	0.75	0.76	0.77	0.74
	Final stage	0.85	0.81	0.80	0.82	0.84	0.77	0.79	0.78
VGG19	Brown stage	0.85	0.82	0.87	0.78	0.72	0.76	0.77	0.78
	Final stage	0.86	0.81	0.85	0.75	0.79	0.75	0.76	0.79
Custombackbone	Brown stage	0.88	0.86	0.84	0.92	0.85	0.87	0.85	0.95
	Final stage	0.85	0.85	0.87	0.91	0.84	0.85	0.89	0.96

TABLE 2: Performance metrics for various segmentation models with various backbones.

Bold values represent the best performance model.

TABLE 3: Selection of ResUNet for custom backbone MobileNetV4.

Segmentation models	mIoU	Dice	Pixel accuracy
UNet	0.88	0.85	0.82
WUNet	0.86	0.87	0.85
U2Net	0.87	0.85	0.88
ResUNet	0.91	0.96	0.95

TABLE 4: Selection of MobileNetV4 as custom backbone for the ResUNet segmentation model.

Backbone	mIoU	Dice	Pixel accuracy
ResNet	0.72	0.69	0.71
EfficientNet	0.82	0.72	0.73
VGG16	0.79	0.78	0.77
VGG19	0.78	0.77	0.71
MobileNetV4	0.91	0.96	0.95

ResUNet model, focusing on the overall segmentation process and considering various performance metrics. In this analysis, ResUNet exhibited robust performance, achieving an impressive mIoU of 0.91, a Dice coefficient of 0.96, and a pixel accuracy of 0.95. These metrics collectively gauge the model's efficacy in accurately segmenting images. Moving on to the evaluation of ResUNet with MobileNetV4 as its backbone, as illustrated in Table 4, the segmentation model demonstrated exceptional performance. Mobile-NetV4 outperformed other backbones, securing the highest scores across all evaluated metrics: an mIoU of 0.91, a Dice coefficient of 0.96, and a pixel accuracy of 0.95. These results underscore the notable enhancement in segmentation capabilities achieved by coupling ResUNet with MobileNetV4. Contrastingly, when EfficientNet served as the backbone, there was a slight decline in performance, reflected in an mIoU of 0.82, a Dice coefficient of 0.72, and a pixel accuracy of 0.73. Similarly, ResNet, as a backbone, exhibited the lowest performance among the configurations assessed, with an mIoU of 0.72, a Dice coefficient of 0.69, and a pixel accuracy of 0.71. These nuanced findings underscore the importance of carefully selecting a compatible backbone for the ResUNet segmentation model. MobileNetV4 emerges as the optimal choice, demonstrating superior segmentation accuracy across multiple performance metrics. This detailed analysis provides comprehensive insights into the

comparative performance of different backbones, facilitating informed decisions about model architecture.

Figure 11 offers a visual depiction of pixel accuracy metrics derived from the comprehensive evaluation of four distinct models: UNet, WUNet, U2Net, and ResUNet. It is noteworthy that each of these models is configured with its unique backbone architecture. Notably, our proposed UNet for semantic segmentation stands out for its remarkable performance, a feat amplified by the integration of a custom backbone. In the specific case of the custom backbone working in tandem with ResUNet, the results are particularly impressive, achieving the highest pixel accuracy recorded at an exceptional 0.98. This underscores the effectiveness of the custom backbone in enhancing the segmentation capabilities of ResUNet. To delve deeper into the comparative analysis of pixel accuracy metrics among these models, UNet demonstrated a pixel accuracy of 0.90, WUNet recorded a pixel accuracy of 0.85, and U2Net yielded a pixel accuracy of 0.87. These individual outcomes emphasize the superior performance of our proposed ResUNet with a custom backbone, especially when contrasted with other segmentation models explored in this study that employed diverse backbone configurations. This detailed examination of pixel accuracy metrics not only highlights the exemplary performance of the proposed ResUNet but also provides valuable insights into the relative strengths of each model.



FIGURE 11: Pixel accuracy metrics of four different models UNet, WUNet, U2Net, and ResUNet, each utilizing distinct backbones.



Confusion Matrix of various segmentation model for ResNet, EfficientNet, VGG16,

FIGURE 12: Confusion matrix of models UNet, WUNet, U2Net, and ResUNet, each utilizing distinct backbones for the prediction of brown and final stages of leaf disease.

The integration of a custom backbone, particularly in conjunction with ResUNet, emerges as a pivotal factor in achieving outstanding pixel accuracy.

Figure 12 provides a detailed view of the confusion matrix associated with four distinct models: UNet, WUNet, U2Net, and ResUNet. Each of these models utilizes diverse backbones to predict both the brown and final stages of leaf disease. A standout observation is the remarkable performance achieved by our proposed backbone in conjunction with ResUNet, resulting in an impressive prediction accuracy of 95%.

This outstanding accuracy underscores the efficacy of our proposed backbone when integrated with ResUNet, showcasing its capability to accurately predict both brown and final stages of leaf disease. The synergy between the custom backbone and ResUNet evidently contributes to superior predictive outcomes.

In conclusion, the results presented in Figure 12 affirm the excellence of our proposed approach. 95% prediction accuracy demonstrates the practical success of our model in effectively handling the complexity of leaf disease prediction. This achievement not only highlights the advancements made in the field but also serves as a testament to the potential impact of innovative backbone configurations in enhancing the overall performance of segmentation models. The combination of a well-designed backbone with ResUNet stands out as a key factor in achieving this commendable accuracy.

5. Conclusion

In conclusion, this paper introduces a groundbreaking segmentation approach for effectively detecting leaf spot disease. The study employs various baseline models (UNet, WUNet, U2Net, and ResUNet), each integrated with distinct pretrained CNN network backbones in the encoder path, leading to significant improvements in segmentation efficiency. One of the key contributions of this research is the proposal of a custom backbone specifically tailored for UNet segmentation, which demonstrated exceptional accuracy in precisely delineating spots associated with both brown-stage and final-stage leaf spot diseases. In addition, the study explores the efficacy of DCGAN-based augmentation, a semantic and efficient process that successfully generates 10,000 images (5,000 images for each type). This augmentation technique significantly enriches the dataset, resulting in notable performance enhancements for the segmentation models. Specifically, our proposed DCGAN augmentation achieved an impressive SSIM score of 0.90 ± 0.172 and a PSNR score of 25 ± 2.4 . The proposed approach exhibits remarkable potential in advancing leaf spot disease detection and has practical implications for agricultural research and applications. The study's promising results underscore the importance of employing efficient segmentation techniques and augmentations to elevate the accuracy and reliability of disease classification processes. Furthermore, the integration of the custom backbone has proven to be particularly beneficial, enabling the detection model to capture low-level features of brown spots of varying sizes with an impressive

mIoU of 0.95. This customized backbone can be implemented in lightweight networks suitable for mobile-based applications. Despite the heightened computational complexity and extended training time associated with the larger and more diverse dataset, our deployed segmentation model exhibited improved efficiency. Initially, segmentation results were suboptimal, with mIoU and Dice scores falling below the 0.5 range. However, substantial enhancements were observed post-augmentation, with reported mIoU reaching 0.91 and Dice reaching 0.96, underscoring the effectiveness of DCGAN augmentation in refining segmentation accuracy and consistency. In addition, our proposed segmentation model effectively handled the larger and more complex dataset, achieving a pixel accuracy of 0.96 and achieving efficient segmentation. It is worth noting that the augmentation process was introduced to address data-related efficiency challenges, yet our segmentation model proved capable of managing these complexities adeptly. Overall, the findings of this research open up new avenues for more effective leaf spot disease detection, offering valuable insights into the application of efficient segmentation methods and augmentations in the field of agriculture. With continued development and implementation, the proposed approach has the potential to make a significant impact on crop disease management and contribute to the advancement of agricultural practices.

Data Availability

The datasets generated during and/or analyzed during the current study can be publicly available and accessible.

Ethical Approval

This article contains no studies with human participants or animals performed by any of the authors.

Conflicts of Interest

All the authors declare that they have no conflicts of interest.

Acknowledgments

Open-access publication funding will be provided by the Manipal Academy of Higher Education, Manipal.

References

- P. A. Nazarov, D. N. Baleev, M. I. Ivanova, L. M. Sokolova, and M. V. Karakozova, "Infectious plant diseases: etiology, current status, problems and prospects in plant protection," *Acta Naturae*, vol. 12, no. 3, p. 46, 2020.
- [2] S. Rathi, H. McFeeters, R. L. McFeeters, and M. R. Davis, "Purification and phytotoxic analysis of botrytis cinerea virulence factors: new avenues for crop protection," *Agriculture*, vol. 2, no. 3, pp. 154–164, 2012.
- [3] N. Bodenhausen, M. W. Horton, and J. Bergelson, "Bacterial communities associated with the leaves and the roots of arabidopsis thaliana," *PLoS One*, vol. 8, no. 2, Article ID e56329, 2013.

- [4] H. El-Ramady, P. Hajdú, G. Törős et al., "Plant nutrition for human health: a pictorial review on plant bioactive compounds for sustainable agriculture," *Sustainability*, vol. 14, no. 14, p. 8329, 2022.
- [5] O. Calicioglu, A. Flammini, S. Bracco, L. Bellù, and R. Sims, "The future challenges of food and agriculture: an integrated analysis of trends and solutions," *Sustainability*, vol. 11, no. 1, p. 222, 2019.
- [6] T. P. Dodiya, G. D. Patel, N. K. Patel, D. A. Patel, and A. D. Gadhiya, "Economics of different intercropping systems in jasmine (jasminum sambac l.) var. baramasi," *Advances in Life Sciences*, vol. 10, pp. 6896–6898, 2016.
- [7] K. Hemanandhini, A. Muthukumar, A. K. Reetha, R. Udhayakumar, and R. Logeshwari, "Eefect of different media on growth and cultural characteristics of alternaria jasmini causing jasmine leaf blight," *Plant Archives*, vol. 19, no. 2, pp. 2220–2224, 2019.
- [8] I. Kamenova, S. Adkins, and D. Achor, "Identification of Tomato Mosaic Virus Infection in Jasmine," 2006, https:// www.actahort.org/books/722/722_34.htm.
- [9] M. Nivedha, E. G. Ebenezar, K. Kalpana, and A. Kumar, "In vitro antifungal evaluation of various plant extracts against leaf blight disease of jasminum grandiflorum caused by alternaria alternata (fr.) keissler," *Journal of Pharmacognosy* and Phytochemistry, vol. 8, no. 3, pp. 2143–2147, 2019.
- [10] R. Sanoubar and L. Barbanti, "Fungal diseases on tomato plant under greenhouse condition," *European Journal of Biological Research*, vol. 7, no. 4, pp. 299–308, 2017.
- [11] H. Durmuş, E. O. Güneş, and M. Kırcı, "Disease detection on the leaves of the tomato plants by using deep learning," in *Proceedings of the 2017 6th International conference on agrogeoinformatics*, IEEE, Fairfax, VA, USA, August 2017.
- [12] A. A. Habibollah, "Deep residual learning for tomato plant leaf disease identification," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 24, 2017.
- [13] S. Zhang, W. Huang, and C. Zhang, "Three-channel convolutional neural networks for vegetable leaf disease recognition," *Cognitive Systems Research*, vol. 53, pp. 31–41, 2019.
- [14] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: tomato leaf disease detection using convolution neural network," *Procedia Computer Science*, vol. 167, pp. 293–301, 2020.
- [15] J. Guo, B. Yue, G. Xu, Z. Yang, and J.-M. Wei, "An enhanced convolutional neural network model for answer selection," in *Proceedings of the 26th international conference on world wide web companion*, pp. 789-790, Perth Australia, April 2017.
- [16] M. Shobana, S. Vaishnavi, C. Gokul Prasad et al., "Plant disease detection using convolution neural network," in *Proceedings of the 2022 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–5, IEEE, Coimbatore, India, January 2022.
- [17] K. Khan, R. U. Khan, W. Albattah, and A. M. Qamar, "Endto-end semantic leaf segmentation framework for plants disease classification," *Complexity*, vol. 2022, Article ID 1168700, 11 pages, 2022.
- [18] M. Dong, H. Yu, L. Zhang et al., "Measurement method of plant phenotypic parameters based on image deep learning," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7664045, 9 pages, 2022.
- [19] A. S. Ansari, M. Jawarneh, M. Ritonga et al., "Improved support vector machine and image processing enabled methodology for detection and classification of grape leaf disease," *Journal of Food Quality*, vol. 2022, Article ID 9502475, 6 pages, 2022.

- [20] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, R. Pramodhini, and R. Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, 2022.
- [21] G. Geetha, S. Samundeswari, G. Saranya, K. Meenakshi, and M. Nithya, "Plant leaf disease classification and detection system using machine learning," *Journal of Physics: Conference Series*, vol. 1712, Article ID 012012, 2020.
- [22] L. S. P. Annabel, T. Annapoorani, and P. Deepalakshmi, "Machine learning for plant leaf disease detection and classification-a review," in *Proceedings of the 2019 international conference on communication and signal processing (ICCSP)*, IEEE, Chennai, India, April 2019.
- [23] R. P. Narmadha and G. Arulvadivu, "Detection and measurement of paddy leaf disease symptoms using image processing," in *Proceedings of the 2017 International Conference* on Computer Communication and Informatics (ICCCI), pp. 1–4, IEEE, Coimbatore, India, January 2017.
- [24] V. Gupta, N. Sengar, M. K. Dutta, C. M. Travieso, and J. B. Alonso, "Automated segmentation of powdery mildew disease from cherry leaves using image processing," in *Proceedings of the 2017 International Conference and Workshop on Bioinspired Intelligence (IWOBI)*, IEEE, Funchal, Portugal, July 2017.
- [25] Y. Kurmi and S. Gangwar, "A leaf image localization based algorithm for different crops disease classification," *Information Processing in Agriculture*, vol. 9, no. 3, pp. 456–474, 2022.
- [26] G. S. Nagaraja, K. Vanishree, and F. Azam, "Novel framework for secure data aggregation in precision agriculture with extensive energy efficiency," *Journal of Computer Networks* and Communications, vol. 2023, Article ID 5926294, 11 pages, 2023.
- [27] M. Shoaib, T. Hussain, B. Shah et al., "Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease," *Frontiers in Plant Science*, vol. 13, Article ID 1031748, 2022.
- [28] X. Xie, Y. Ma, B. Liu, J. He, S. Li, and H. Wang, "A deeplearning-basedreal-time detector for grape leaf diseases using improved convolutional neural networks," *Frontiers of Plant Science*, vol. 11, p. 751, 2020.
- [29] R. H. Hridoy, M. Tarek Habib, M. Sadekur Rahman, and M. S. Uddin, "Deep neural networks-based recognition of betel plant diseases by leaf image classification," in *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN*, pp. 227–241, Springer, Singapore, 2022.
- [30] S. Kaur, S. Pandey, and S. Goel, "Semi-automatic leaf disease detection and classification system for soybean culture," *IET Image Processing*, vol. 12, no. 6, pp. 1038–1048, 2018.
- [31] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," *Environmental Monitoring and Assessment*, vol. 195, no. 1, p. 120, 2023.
- [32] S. Alsubai, A. K. Dutta, A. H. Alkhayyat, M. M. Jaber, A. H. Abbas, and A. Kumar, "Hybrid deep learning with improved salp swarm optimization based multi-class grape disease classification model," *Computers & Electrical Engineering*, vol. 108, Article ID 108733, 2023.
- [33] M. Shoaib, B. Shah, S. Ei-Sappagh et al., "An advanced deep learning models-based plant disease detection: a review of recent research," *Frontiers in Plant Science*, vol. 14, Article ID 1158933, 2023.

- [34] K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, "Deep learningbased segmentation and quantification of cucumber powdery mildew using convolutional neural network," *Frontiers of Plant Science*, vol. 10, no. 155, 2019.
- [35] M. M. Soliman, M. H. Kamal, M. A. E.-M. Nashed, Y. M. Mostafa, B. S. Chawky, and D. Khattab, "Violence recognition from videos using deep learning techniques," in *Proceedings of the 2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS)*, pp. 80–85, IEEE, Cairo, Egypt, December 2019.
- [36] D. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," 2015, https://arxiv.org/abs/1511.08060.
- [37] M. T. Chiu, X. Xu, Y. Wei et al., "Agriculture-vision: a large aerial image database for agricultural pattern analysis," in *Proceedings of the IEEE/CVF Conference on Computer Vision* and Pattern Recognition, pp. 2828–2838, Seattle, WA, USA, June 2020.
- [38] R. Kumar, A. Chug, A. P. Singh, and D. Singh, "A systematic analysis of machine learning and deep learning based approaches for plant leaf disease classification: a review," *Journal* of Sensors, vol. 2022, Article ID 3287561, 13 pages, 2022.
- [39] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 11 pages, 2016.
- [40] Y. He, Q. Gao, and Z. Ma, "A crop leaf disease image recognition method based on bilinear residual networks," *Mathematical Problems in Engineering*, vol. 2022, Article ID 2948506, 15 pages, 2022.
- [41] X. Wenchao and Y. Zhi, "Research on strawberry disease diagnosis based on improved residual network recognition model," *Mathematical Problems in Engineering*, vol. 2022, Article ID 6431942, 13 pages, 2022.
- [42] J. A. Pandian, G. Geetharamani, and B. Annette, "Data augmentation on plant leaf disease image dataset using image manipulation and deep learning techniques," in *Proceedings of the 2019 IEEE 9th international conference on advanced computing (IACC)*, pp. 199–204, IEEE, Tiruchirappalli, India, December 2019.
- [43] P. Enkvetchakul and O. Surinta, "Effective data augmentation and training techniques for improving deep learning in plant leaf disease recognition," *Applied Science and Engineering Progress*, vol. 15, no. 3, p. 3810, 2021.
- [44] L. Bi and G. Hu, "Improving image-based plant disease classification with generative adversarial network under limited training set," *Frontiers of Plant Science*, vol. 11, Article ID 583438, 2020.
- [45] Q. H. Cap, H. Uga, S. Kagiwada, and H. Iyatomi, "Leafgan: an effective data augmentation method for practical plant disease diagnosis," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 2, pp. 1258–1267, 2022.
- [46] I. Goodfellow, "Nips 2016 tutorial: generative adversarial networks," 2016, https://arxiv.org/abs/1701.00160.
- [47] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, pp. 234–241, Springer, Munich, Germany, October 2015.
- [48] X.-X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-net-based medical image segmentation," *Journal of Healthcare Engineering*, vol. 2022, Article ID 4189781, 16 pages, 2022.

- [49] C. Etmann, R. Ke, and C.-B. Schönlieb, "iunets: fully invertible u-nets with learnable up-and downsampling," 2020, https:// arxiv.org/abs/2005.05220.
- [50] Z. Zhou, S. Rahman, M. Mahfuzur, N. Tajbakhsh, and J. Liang, "U.++: A nested u-net architecture for medical image segmentation," in Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, pp. 3–11, Springer, Granada, Spain, September 2018.
- [51] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 94–114, 2020.
- [52] B. Liu, J. Lv, X. Fan, J. Luo, and T. Zou, "Application of an improved dcgan for image generation," *Mobile Information Systems*, vol. 2022, Article ID 9005552, 14 pages, 2022.
- [53] B. Ehteshami Bejnordi, M. Veta, P. Johannes van Diest et al., "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," *JAMA*, vol. 318, no. 22, pp. 2199–2210, 2017.