

RESEARCH

Open Access



Enhanced fungal disease detection in brinjal leaves using deep semantic segmentation with U2-Net

Ruchita R Patil¹ , Shwetha V^{1*} and Annamalai Muthusamy²

*Correspondence:

Shwetha V

shwetha.v@manipal.edu

¹Manipal Institute of Technology,
Manipal Academy of Higher
Education, Manipal, India

²Department of Plant Sciences,
Manipal School of Life Sciences,
Manipal Academy of Higher
Education, Manipal,
Karnataka 576104, India

Abstract

Accurate identification of plant diseases is paramount for ensuring optimal crop health and minimizing yield losses. This study focuses on *Solanum melongena* L. (eggplant), a widely cultivated vegetable highly susceptible to fungal infections. We employ a semantic segmentation approach to detect leaf symptoms and damage, utilizing deep learning models. A comparative analysis is conducted on three convolutional neural network architectures: U2-Net, U-Net, and WU-Net, for image segmentation tasks. Each model is trained on a dataset comprising 500 augmented images derived from an initial set of 200 images, with a resolution of $256 \times 256 \times 3$. Model performance is evaluated based on pixel accuracy, Dice coefficient, and Intersection over Union. Experimental results demonstrate that U2-Net exhibits outstanding performance, particularly in capturing fine-grained details, attributed to its deeper architecture and enhanced feature extraction capabilities. The proposed U2-Net model achieved a training accuracy of 96% and test accuracy of 93%, demonstrating its effectiveness in precise symptom detection and plant disease classification. This study contributes to leaf disease detection, facilitating timely intervention and targeted disease management in agriculture.

Keywords U2-Net, U-Net, WU-Net, Segmentation, Classification, Fungal infection, Controlled, Disease leaf

1 Introduction

The identification of plant diseases is essential for minimizing yield loss, preserving and maintaining the quality of plant products. Timely detection of plant diseases serves as the initial step in controlling and preventing disease outbreaks. The study of plant diseases focuses on analysing morphological symptoms that appear on plants [1, 2]. Plant pathogens are the organisms that infect the plants, causing disease by invading their tissues and extracting essential nutrients. These pathogens include fungi, virus, bacteria, nematodes and other harmful microbes. Among the pathogens, the fungal pathogens cause substantial yield reductions and pose severe challenges to global food security [3]. Leaf spot disease was prevalent in the Mattu Gulla, a unique brinjal variety and different



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

type of fungal disease was recorded in fields during 2018–2019 and 2019–2020 [4]. A predominant leaf spot disease was caused by fungal genera such as *Alternaria*, *Cercospora*, and *Septoria* was observed in eggplant [5]. Pathogens infiltrate plants through wounds, natural openings, or by being transmitted and injected by vectors like insects. They invade plant tissues by producing enzymes that degrade cell structures, leading to cell death. Additionally, pathogens release various bioactive compounds, including toxins that directly kill cells and growth regulators that trigger abnormal growths such as galls [6, 7]. Measuring the severity of the plant disease is crucial for assessing infection levels, predicting crop yield and determining appropriate treatments. Timely and accurate evaluation of disease severity can help reduce these losses [8]. Conventionally, trained experts assess disease severity through visual examination of plant tissues. However, the high cost and inefficiency associated with manual assessment present challenges to the preferment of modern agriculture [9]. Even skilled plant pathologist and experienced agronomists find it difficult to accurately diagnose the plant diseases due to their intricacy, vast number of cultivated crops and their ongoing phytopathological issues, this leads to inappropriate conclusions and incorrect treatments [10]. With advancement in technology, deep learning techniques have become more effective tools for identifying and evaluating plant disease severity through image analysis, demonstrating significant potential in accurately categorizing different levels of disease severity [11]. There are different performance metrics, such as accuracy, precision, recall and F1 score are used to evaluate the efficacy of various strategies and architectures. Building strong plant disease detection models requires precise evaluation using these measures, which will ultimately increase agricultural output and food security [12–14]. Despite this, there are certain challenges need to be addressed and considered which are restricting the effectiveness and accurate performance of these models for crop disease diagnosis. The main issues affecting automatic plant disease detection and classification includes unnecessary noise, many non-essential elements in the image background, uneven and non-uniformity in training dataset, different types of diseases with similar patterns and symptoms which can be difficult to distinguish between disease type, capturing the images at controlled environmental conditions and size of the dataset. Therefore, it is essential to develop new methods and approaches to tackle the existing issues and have betterment [15, 16]. Semantic segmentation has been extensively employed in various medical applications and has significantly improved diagnostic precision and treatment planning [17–23]. Alongside such advances, recent studies also highlight that accurate leaf identification is often a necessary step prior to disease detection; for instance, a quantum classifier was employed to recognize structural features of different leaf species with a 97.3% accuracy, underscoring the importance of precise leaf recognition [24]. The novelty of the present study lies in its application of semantic segmentation for detecting morphological patterns in leaves.

Image segmentation is a critical task in computer vision, enabling the partitioning of images into meaningful regions for applications ranging from medical imaging to autonomous driving. Convolutional neural networks (CNNs) have significantly advanced this field, with architectures like U-Net setting benchmarks for segmentation tasks. Recent developments, such as U2-Net and WU-Net, have introduced modifications aimed at enhancing segmentation performance [25, 26]. The primary focus of this study is the semantic segmentation of fungal-infected regions in brinjal leaves using U²-Net. While

image-level classification results are also reported, they are derived from the segmentation outputs and serve only to validate the practical utility of the proposed model.

The contribution of the study is as follows.

- In this study, we focus on eggplant (*Solanum melongena* L.), a widely consumed vegetable that is highly susceptible to fungal diseases.
- We employ semantic segmentation to detect leaf damage, aiming to enhance the disease identification and improve crop health.
- The proposed U2-Net model scored training accuracy of 96% and 93% for test accuracy.

2 Literature review

A thorough review from the previous studies reveals that traditional methods were employed to detect the plant diseases as reported by Khakimov et al., include optical examination, microscope based visualisation and mycological analysis [27]. According to Ray et al., these techniques, often considered as the ‘gold standard,’ where they depend on symptom observation, culture-based methods, and morphological and biochemical identification [28]. Patel et al. addressed that these methods were time consuming, labour intensive and requires more time to estimate the severity of the disease [29]. To overcome these limitations, Khakimov et al. reported advanced techniques, such as molecular-genetic methods, mass spectrometry and immunodiagnosics. Serological assays and DNA-based methods are widely used molecular biology techniques for plant disease detection. Martinelli et al. noted that serological assays, such as ELISA approach is more reliable and provide rapid and specific pathogen detection, On the other hand, quantitative PCR (qPCR) is widely available and effective method [30]. Aslam et al. and Hariharan & Prasannath, according to their analysis they mentioned that polymerase chain reaction (PCR) and its variants, including nested PCR, multiplex PCR, and quantitative PCR, are widely used for detecting plant pathogens [31, 32]. Despite the fact that these methodologies are trusted for developing effective management programs for various plant diseases, Gharbi et al. mentioned that their effectiveness may be limited by factors such as the indeterminate spread of the disease, the presence of asymptomatic infections, and the absence of verified sampling protocols [33].

To overcome the limitations of traditional plant disease detection and improve spatial identification of diseased leaf regions, several segmentation-based deep learning methods have recently emerged. (Lin et al., 2019) developed a CNN-based semantic segmentation model to identify and quantify powdery mildew on cucumber leaves at the pixel level. The model achieved 96.08%-pixel accuracy, 72.11% IoU, and 83.45% Dice score, outperforming traditional methods like K-means and Random Forest, and offering a valuable tool for assessing disease severity in breeding programs [34]. Deng et al. proposed MC-UNet, a tomato leaf disease segmentation model integrating a Multi-scale Convolution Module and Cross-layer Attention Fusion to enhance edge detection, achieving 91.32% accuracy on a custom dataset [35]. Addressing geometric variation in persimmon leaf lesions, Jia et al. enhanced UNet with deformable convolution and a self-attention mechanism, achieving 89.18% mPA and 83.58% mIoU [36]. For small spot segmentation on potato leaves, Fu et al. introduced RS-UNet using ResNet50 and SE attention, resulting in 79.8% mIoU and 88.86% Dice score [37]. Similarly, Wang et al. developed FFAE-UNet for pear leaf disease segmentation by incorporating an attention

guidance module (AGM) and a feature enhancement supplementation module (FESM), achieving 86.60% mIoU, 92.58% Dice, and 91.85% MPA, outperforming existing methods [38]. Shwetha et al. employed a UNet model with a MobileNetV4 backbone was used for identification and segmentation of leaf blight spots, achieving 0.91 average pixel accuracy and 0.95 mIoU [39]. Building on segmentation approaches, a MobileNetV3-based classifier was later developed by Shwetha et al. for leaf blight spot disease, incorporating depthwise convolutions, max-pooling layers, and data augmentation via Conditional GAN. This classifier achieved 97% training accuracy and 94–96% test accuracy under varied real-world conditions, demonstrating robust and efficient disease identification [40].

Although segmentation-based models focus on identifying diseased regions at the pixel level, deep learning classification models aim to directly categorize leaf images. In this context, EfficientNet-based deep learning model have been proposed by Atila et al. for plant leaf disease classification, achieving up to 99.97% accuracy on the PlantVillage dataset, performing better than the other state-of-the-art architectures [41]. El-Shahat et al. introduced a hybrid MobileNet-DNN model (HMobileNetDNN) for detecting rice plant diseases. When tested on the Paddy Doctor dataset, the model achieved 0.918 accuracy, 0.918 precision, 0.907 recall, 0.912 F1 score, and 0.949 AUC, indicating its strong capability in disease detection [42]. These studies provide the foundational basis for this research. The models employed are highly reliable, robust, and accurate in plant disease detection, demonstrating their potential to advance the field.

3 Methodology

The input images in the given diagram are leaf images, which serve as the primary data for the segmentation and classification process. These images likely include both healthy (control) leaves and diseased leaves with visible symptoms such as discoloration, necrosis, or other signs of infection. The images are processed to extract relevant features for segmentation and classification into control or diseased categories. Figure 1 depicts the workflow for the proposed framework.

Training the U2-Net model for 50 epochs on a dataset of 500 augmented images (originally 200 images) with a resolution of $256 \times 256 \times 3$ involves several key steps:

3.1 Dataset

Figure 2 illustrates the image samples of control and fungal-infected diseased leaves used in this study. The dataset for these images was developed in collaboration with an

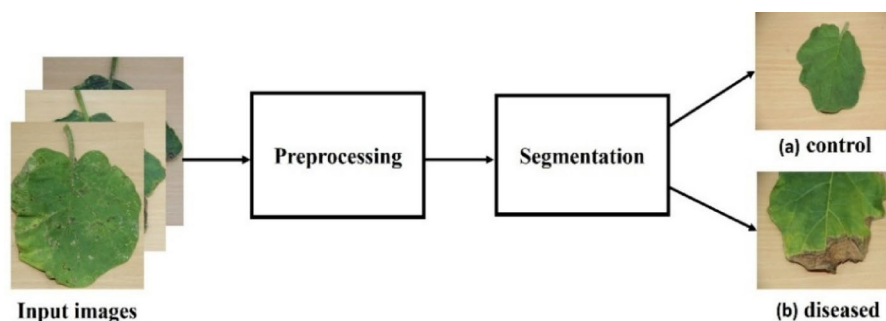


Fig. 1 The overall workflow of the proposed framework for detecting and segmenting diseased regions in eggplant leaves, highlighting preprocessing, segmentation using U²-Net, and classification stages

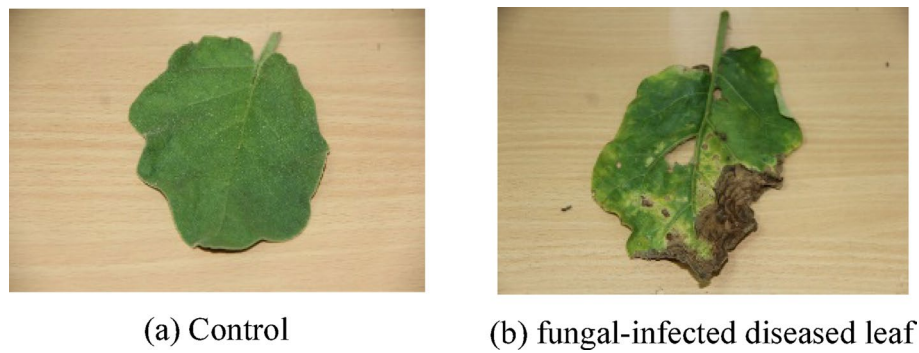


Fig. 2 **a** Control and **b** fungal-infected diseased leaf

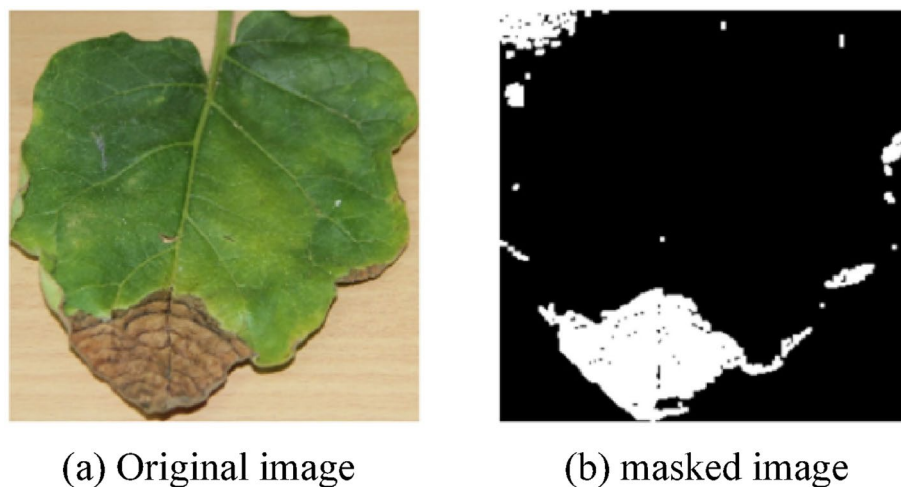


Fig. 3 **a** Original image of an eggplant leaf showing visible brown lesion symptoms, and **b** the corresponding binary ground truth mask highlighting the infected lesion area used for segmentation training

expert from the Manipal School of Life Sciences, Manipal, Karnataka, India. A digital camera was used to systematically capture the images for further analysis. To ensure the image quality and consistency, standard lighting conditions and uniform background were maintained during the image capture.

3.2 Lesion mask generation

To facilitate segmentation model training, each original leaf image was paired with a binary ground truth mask highlighting the diseased region. The lesion masks help the model learn to localize infection-affected areas. An example of an original leaf image and its corresponding lesion mask is depicted in Fig. 3.

3.3 Preprocessing

The dataset comprised 200 original images with corresponding segmentation masks, including 50 healthy and 150 diseased leaves. Data augmentation including rotations, flips, scaling, and elastic deformations expanded the dataset to 500 images. All images were resized to 256×256 pixels with three color channels. From the augmented dataset, an 80:20 train/test split was applied, resulting in 400 training images and 100 test images. During training, 20 percent of the training set (80 images) was held out as a validation

set for learning-rate scheduling and checkpointing, yielding final counts of 320 training, 80 validation, and 100 test images.

3.4 Model architecture

U-Net is a widely adopted architecture for image segmentation that features a symmetric encoder-decoder structure with skip connections, facilitating the transfer of spatial information between corresponding layers. An evolution of U-Net, U2-Net incorporates a two-level nested U-structure with Residual U-blocks (RSU) at each stage, enabling effective multi-scale feature extraction without significant computational overhead. WU-Net, a variant of U-Net designed to enhance edge information extraction, employs a deeper network structure with two independent paths for feature extraction, where the feature maps from both paths are concatenated for classification, ultimately improving the accuracy of edge information extraction in high-resolution images.

The architectural structure of U²-Net used for segmentation is illustrated in Fig. 4.

The figure provides an overview of the U2-Net architecture used for brinjal leaf disease segmentation, while the detailed per-layer model configuration, output shapes, and parameter counts are provided in Table 1.

3.4.1 Implementation details

PyTorch: Construct the U2-Net model using the torch.nn module. Implement custom datasets and dataloaders to manage image-mask pairs efficiently. Employ the torchvision.transforms module for on-the-fly data augmentation during training.

3.4.2 Proposed U2-Net

To adapt the U2-Net model for binary segmentation of diseased and healthy leaves with an input size of 256×256 pixels, we modified the architecture as outlined in Table 1.

1. Input layer:

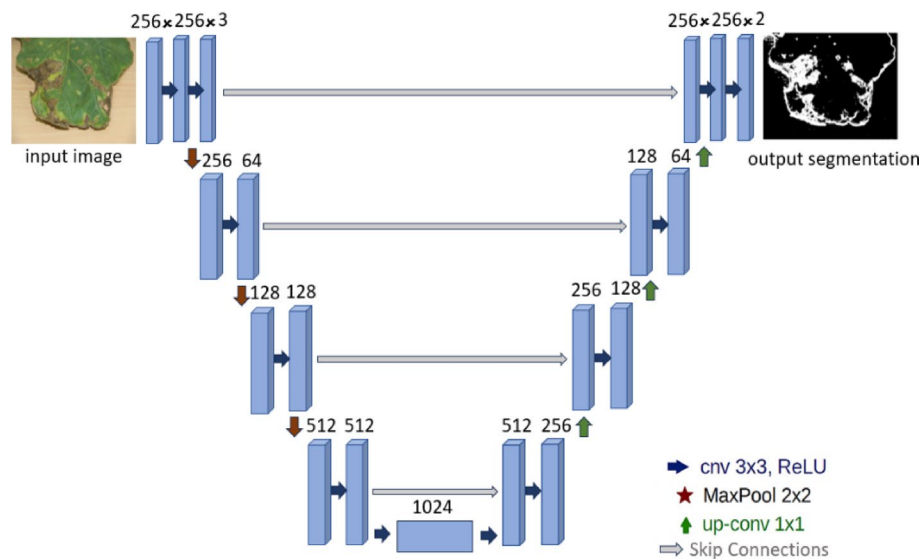


Fig. 4 The architecture of the U²-Net model used for brinjal leaf disease segmentation, highlighting its nested U-shaped design and multiscale Residual U-blocks (RSUs)

Table 1 Represents the model summary of the proposed framework

Layer name	Layer type	Output shape	Parameters
Input	Input layer	(256,256,3)	0
Conv1	Conv2D+ ReLU	(256,256,64)	1,792
Conv2	Conv2D+ ReLU	(256,256,64)	36,928
Pool1	MaxPooling2D	(128,128,64)	0
Conv3	Conv2D+ ReLU	(128,128,128)	73,856
Conv4	Conv2D+ ReLU	(128,128,128)	147,584
Pool2	MaxPooling2D	(64,64,128)	0
Conv5	Conv2D+ ReLU	(64,64,256)	295,168
Conv6	Conv2D+ ReLU	(64,64,256)	590,080
Pool3	MaxPooling2D	(32,32,256)	0
Conv7	Conv2D+ ReLU	(32,32,512)	1,180,160
Conv8	Conv2D+ ReLU	(32,32,512)	2,359,808
Pool4	MaxPooling2D	(16,16,512)	0
Conv9	Conv2D+ ReLU	(16,16,1024)	4,719,616
Conv10	Conv2D+ ReLU	(16,16,1024)	9,438,208
upConv1	UpConv2D	(32,32,512)	2,097,664
Concat1	Concatenate	(32,32,1024)	0
Conv11	Conv2D+ ReLU	(32,32,512)	4,719,104
Conv12	Conv2D+ ReLU	(32,32,512)	2,359,808
UpConv2	UpConv2D	(64,64,256)	524,544
Concat2	Concatenate	(64,64,256)	0
Conv13	Conv2D+ ReLU	(64,64,256)	1,179,904
Conv14	Conv2D+ ReLU	(64,64,256)	590,080
UpConv3	UpConv2D	(128,128,128)	131,200
Concat3	Concatenate	(128,128,256)	0
Conv15	Conv2D+ ReLU	(128,128,128)	295,040
Conv16	Conv2D+ ReLU	(128,128,128)	147,584
UpConv4	UpConv2D	(256,256,64)	32,832
Concat4	Concatenate	(256,256,128)	0
Conv17	Conv2D+ ReLU	(256,256,64)	73,792
Conv18	Conv2D+ ReLU	(256,256,64)	36,928
Output	Conv2d+ Softmax	(256,256,2)	130
Metric			Value
Total parameters			~ 17.8 million
Trainable parameters			~ 17.8 million
Non-trainable parameters			0
FLOPs (approx.)			~ 32.4 GFLOPs

Dimensions: Adjust the input layer to accept images of size 256×256 pixels with 3 color channels (RGB).

2. Output layer:

Classes: Configure the final layer to produce two channels corresponding to the two classes: diseased and healthy.

Activation function: Apply the softmax activation function to obtain class probabilities for each pixel.

The proposed U2-Net architecture consists of approximately 17.8 million parameters and an estimated 32.4 GFLOPs for an input size of $256 \times 256 \times 3$. Despite its nested structure, the model remains computationally feasible for deployment on modern GPUs and high-performance edge devices.

3.5 Training process

Each model was trained for 50 epochs using the Adam optimizer. The learning rate was set to 0.0001 and adjusted dynamically using the ReduceLROnPlateau scheduler to ensure better convergence. A batch size of 16 was used during training. To handle class imbalance and improve segmentation performance, the Binary Cross Entropy loss function was employed. Additionally, model checkpoints were saved periodically to retain the best-performing models and avoid data loss, as shown in Table 2.

The U2-Net, U-Net, and WU-Net models were trained using the following hyperparameters:

Data augmentation techniques were employed to enhance model robustness and generalization, such as:

- Rotation ($\pm 30^\circ$)
- Horizontal and vertical flips
- Zoom range: 0.8–1.2
- Elastic deformations
- Brightness/Contrast adjustment

3.5.1 Problem formulation and loss function

The fungal disease detection task was formulated as a binary semantic segmentation problem, where each pixel of an input image $I \in \mathbb{R}^{H \times W \times 3}$ is classified into one of two categories: healthy (0) or diseased (1). The corresponding ground-truth mask is denoted as given in the Eq. (1).

$$M \in \{0, 1\}^{H \times W} \quad (1)$$

Let p_i represent the predicted probability of pixel i belonging to the diseased class, and $y_i \in \{0, 1\}$ denote the true label. To train the segmentation model effectively, the following loss functions were employed:

1. Binary cross-entropy (BCE) loss

Binary Cross-Entropy (BCE) is a commonly used loss function for pixel-level binary classification. It penalizes incorrect predictions at the pixel level and is defined in Eq. (2).

$$L_{BCE} = -\frac{1}{N} \sum_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (2)$$

where $N = H \times W$ is the total number of pixels in an image.

2. Dice loss

Since diseased regions typically occupy few pixels (much smaller portion of the image) compared to healthy ones, class imbalance becomes a challenge. To address this, Dice

Table 2 Summary of key hyperparameters for model training

Hyperparameters	Value/Description
Optimizer	Adam
Learning rate	0.0001 (with ReduceLROnPlateau scheduler)
Batch size	16
Epochs	50
Loss function	Binary cross entropy

Loss was applied, which directly maximizes the spatial overlap between predictions and ground truth. It is formulated in Eq. (3).

$$L_{Dice} = 1 - 2 \sum i y i p_i + \varepsilon / \sum i y i + \sum i p_i \quad (3)$$

where ε is a small constant to avoid division by zero.

3. Final loss function

The overall objective combined BCE and Dice Loss in a weighted form, as represented in Eq. (4)

$$L_{final} = \alpha L_{BCE} + \beta L_{Dice} \quad (4)$$

with weights $\alpha = 0.5$ and $\beta = 0.5$.

This joint formulation balances pixel-wise accuracy with region-level overlap, enabling the U²-Net model to achieve precise segmentation and capture fine-grained lesion boundaries more effectively.

3.6 Evaluation metrics

The models were evaluated using the following metrics:

- Pixel accuracy: The ratio of correctly classified pixels to the total number of pixels.
- Dice coefficient: Measures the overlap between the predicted segmentation and the ground truth.
- Intersection over union (IoU): Calculates the ratio of the intersection to the union of the predicted and true masks.

3.7 Training environment

Model training was conducted on a high-performance computing setup to ensure efficient execution and reproducibility. The environment included an NVIDIA Tesla T4 GPU with 16 GB VRAM, an Intel Xeon Silver 4210R CPU, and 32 GB RAM. All models were implemented using PyTorch version 1.13.1. On average, the training time per epoch varied slightly depending on the model architecture: approximately 2.5 min for U2-Net, 1.8 min for U-Net, and 2.2 min for WU-Net. Each model was trained for 50 epochs, resulting in a total training time of approximately 2 h. This setup provided an optimal balance between computational efficiency and training accuracy, as demonstrated in Table 3.

3.7.1 Reproducibility and statistical reporting

To assess the stability of each model, we trained every architecture five times using different random seeds while keeping the hyperparameters constant. All reported

Table 3 Training environment and model runtime summary

Component	Specification
GPU	NVIDIA Tesla T4 (16 GB VRAM)
CPU	Intel xeon silver 4210R
RAM	32 GB
Framework	PyTorch 1.13.1
Time per epoch	~2.5 min for U2-Net, ~1.8 min for U-Net, ~2.2 min for WU-Net
Total training time	~2 h (50 epochs)

segmentation and classification metrics (Pixel Accuracy, IoU and Dice) are presented as mean $\pm$ standard deviation across these independent runs. Where appropriate, pairwise statistical tests were performed to ensure that the observed differences were statistically significant.

3.8 Post-processing

- Thresholding: Apply thresholding to the model's output probabilities to obtain binary segmentation masks.
- Morphological operations: Utilize operations such as erosion and dilation to refine segmentation results, removing noise and closing gaps.

3.9 Image-level classification from segmentation masks

In addition to segmentation, U²-Net outputs were used to derive image-level classification results. Specifically, the binary masks generated by the segmentation model were analyzed to determine whether an image belonged to the healthy or diseased category. Images with lesion pixels exceeding a predefined threshold were classified as diseased, while the absence of such lesions were classified as healthy. These binary labels were then compared with the ground truth image labels to compute the classification accuracy. This post-processing step provided a practical validation of the segmentation model's effectiveness for image-level disease diagnosis. The procedure was repeated over five independent runs, and all metrics are reported as mean $\pm$ standard deviation to account for variability.

3.9.1 Comparison with U-Net and WU-Net

After training, the U2-Net model's performance was evaluated against U-Net and WU-Net architectures. Studies have shown that U2-Net often outperforms traditional U-Net models in segmentation tasks due to its deeper architecture and efficient feature extraction capabilities. For instance, in calving front segmentation, U2-Net demonstrated more effective classification compared to fully convolutional networks (FCN) and U-Net models.

4 Results and discussion

The comparative analysis revealed that all three models performed effectively in segmenting images; however, notable differences were observed:

- U2-Net: Demonstrated better performance in capturing fine-grained details, likely due to its deeper architecture and efficient feature extraction capabilities.
- U-Net: Provided robust segmentation results but struggled with accurately capturing complex edge information in high-resolution images.
- WU-Net: Aimed to enhance edge information extraction through a deeper network structure with two independent paths for feature extraction.

Figure 5 presents the training and validation accuracy graphs. Figure 5a illustrates that U2-Net demonstrates more stable and consistent learning behavior in detecting fungal patterns on leaves, compared to U-Net in Fig. 5b and c depicts an underfitting scenario, offering insights into model performance and supporting the selection of a suitable architecture for the segmentation task.

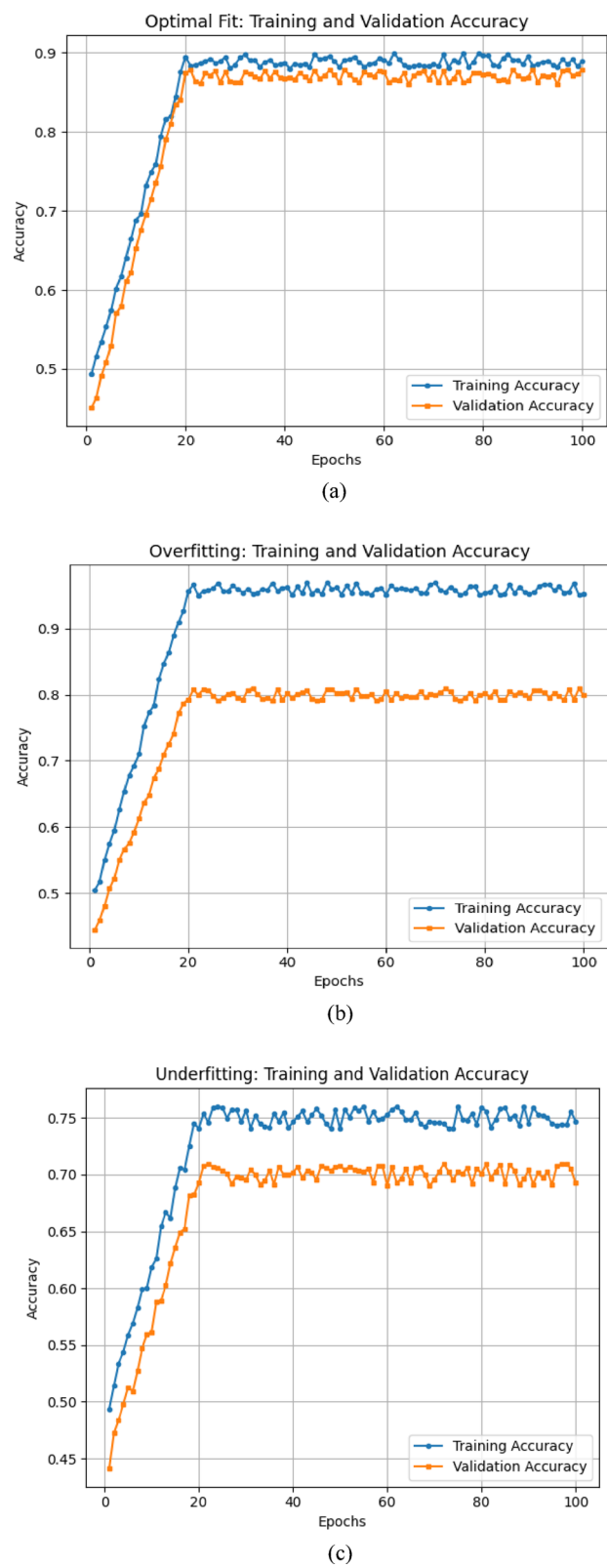


Fig. 5 Training and validation performance curves of different models: **a** U2-Net displays stable learning with minimal gap between training and validation, indicating an optimal fit; **b** U-Net exhibits signs of overfitting; and **c** WU-Net shows underfitting, with consistently low performance across both sets

To assess the segmentation quality of deep learning models, three widely adopted metrics were considered: Pixel accuracy, intersection over union (IoU), and the Dice Coefficient. Each metric evaluates different aspects of the model's ability to correctly delineate segmented regions [33].

Pixel accuracy: Pixel accuracy reflects the percentage of pixels in the input image that are correctly predicted by the model, regardless of class, as shown in the Eq. (5).

$$\text{PixelAccuracy} = \frac{\text{TotalCorrectlyPredictedPixels}}{\text{TotalPixelsintheImage}} \quad (5)$$

Intersection over union (IoU): IoU evaluates how well the predicted segmentation overlaps with the actual target area. It is the ratio of the intersection (common area) to the union (combined area) of the predicted and ground truth regions, as represented in the Eq. (6).

$$\text{IoU} = \frac{\text{TruePositive (TP)}}{\text{TP} + \text{FalsePositive (FP)} + \text{FalseNegative (FN)}} \quad (6)$$

Dice coefficient: The dice coefficient measures the similarity between the predicted and actual segmented regions. It places more emphasis on correctly predicted pixels and is particularly useful in handling class imbalance, as expressed in Eq. (7).

$$\text{DiceScore} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (7)$$

Here, TP (true positives) refers to correctly identified diseased cases, FP (false positives) denotes instances incorrectly classified as diseased, and FN (false negatives) indicates actual diseased cases that the model failed to detect.

Table 4 illustrates the performance scores of three segmentation models, including U2Net, UNet, and WUNet, for leaf disease detection. The results derived from experimental evaluations, are reported using standard segmentation metrics (Pixel Accuracy, IoU, Dice) and are expressed in terms of mean $\pm$ standard deviation over five independent runs to reflect model's accuracy and quality.

Figure 6 presents a bar plot comparing the performance of U2-Net, U-Net, and WU-Net based on three key metrics: pixel accuracy, intersection over union (IoU), and Dice Coefficient. The results indicate that U2-Net outperforms the other models across all metrics. Among these, U2-Net achieves the highest training accuracy of 96%.

Figure 7 presents the confusion matrices illustrating the performance of the models. Figure 7a shows that the U2-Net model achieved an accuracy of 93%, while Fig. 7b indicates that the U-Net model attained 89% accuracy. In comparison, Fig. 7c demonstrates that the WU-Net model achieved an accuracy of 80%.

Table 5 presents the confusion matrix components—true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), for each segmentation model.

Table 4 Segmentation performance metrics for each model, represented as mean $\pm$ standard deviation

Model	Pixel accuracy	IoU score	Dice score
U2-Net	0.98 $\pm$ 0.01	0.95 $\pm$ 0.01	0.97 $\pm$ 0.01
U-Net	0.75 $\pm$ 0.02	0.78 $\pm$ 0.02	0.80 $\pm$ 0.02
WU-Net	0.70 $\pm$ 0.03	0.73 $\pm$ 0.03	0.75 $\pm$ 0.03

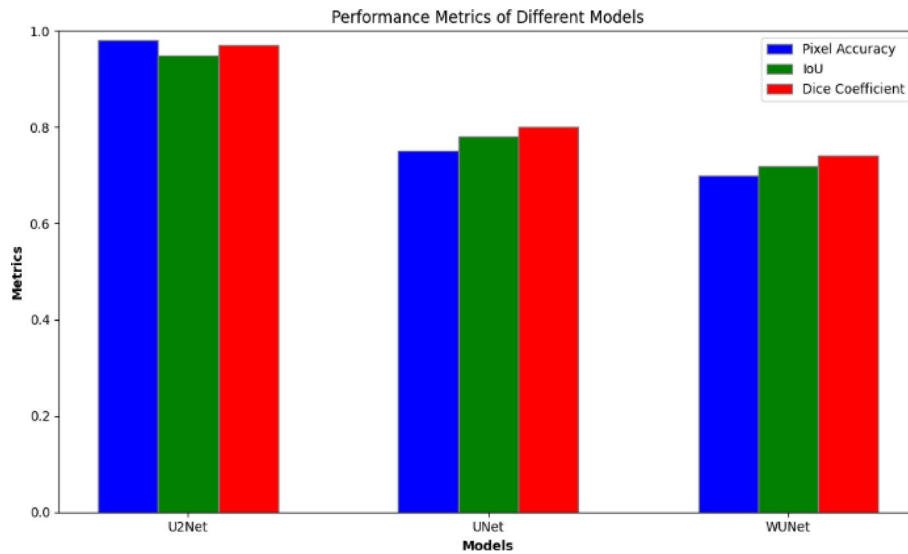


Fig. 6 Comparative analysis of U2-Net, U-Net, and WU-Net across key segmentation metrics-Pixel Accuracy, IoU, and Dice. U2-Net achieved the highest values, indicating its enhanced capability to capture lesion boundaries accurately

These values support the reported accuracy metrics and provide deeper insight into the classification behavior and reliability of each architecture.

These values were derived from a test set of 100 images. The high TP and TN rates for U2-Net validate its effective segmentation and classification performance.

Figure 8 illustrates the different stages of the image segmentation process. Figure 8a represents the input image, Fig. 8b shows the segmented output generated by the proposed model, and Fig. 8c presents the post-processed result after applying morphological closing and opening operations.

Figure 9a shows the original eggplant leaf image with visible symptoms of infection. Figure 9b displays the corresponding binary segmentation mask used to localize the diseased region. Figure 9c illustrates the extracted brown lesion area obtained by applying the mask to the original image, effectively isolating the infected region for further analysis.

Figure 10 presents the ROC curve analysis, demonstrating that all classifiers performed effectively, with AUC values ranging from 0.95 to 1.0. The U2-Net model, utilized for classification, exhibited better segmentation capabilities, leading to improved accuracy. These results indicate that while U-Net remains a strong baseline for segmentation tasks, U2-Net provides enhanced performance for applications requiring detailed segmentation. Although WU-Net's approach to edge information extraction offers potential advantages, its effectiveness may depend on the complexity of the images.

In this study, we compared several state-of-the-art CNN architectures for leaf disease classification (Fig. 11), focusing on performance metrics such as accuracy, training time, parameter count, and implementation platforms. Among the models evaluated, EfficientNetB0 achieved the highest classification accuracy at 94.3% while maintaining a relatively low number of parameters (5.3 million), making it highly suitable for mobile or edge device deployment. DenseNet121 also performed well, achieving 93% accuracy with fewer parameters compared to older models like VGG16, which, despite its historical popularity, is significantly heavier with 138 million parameters and offers lower

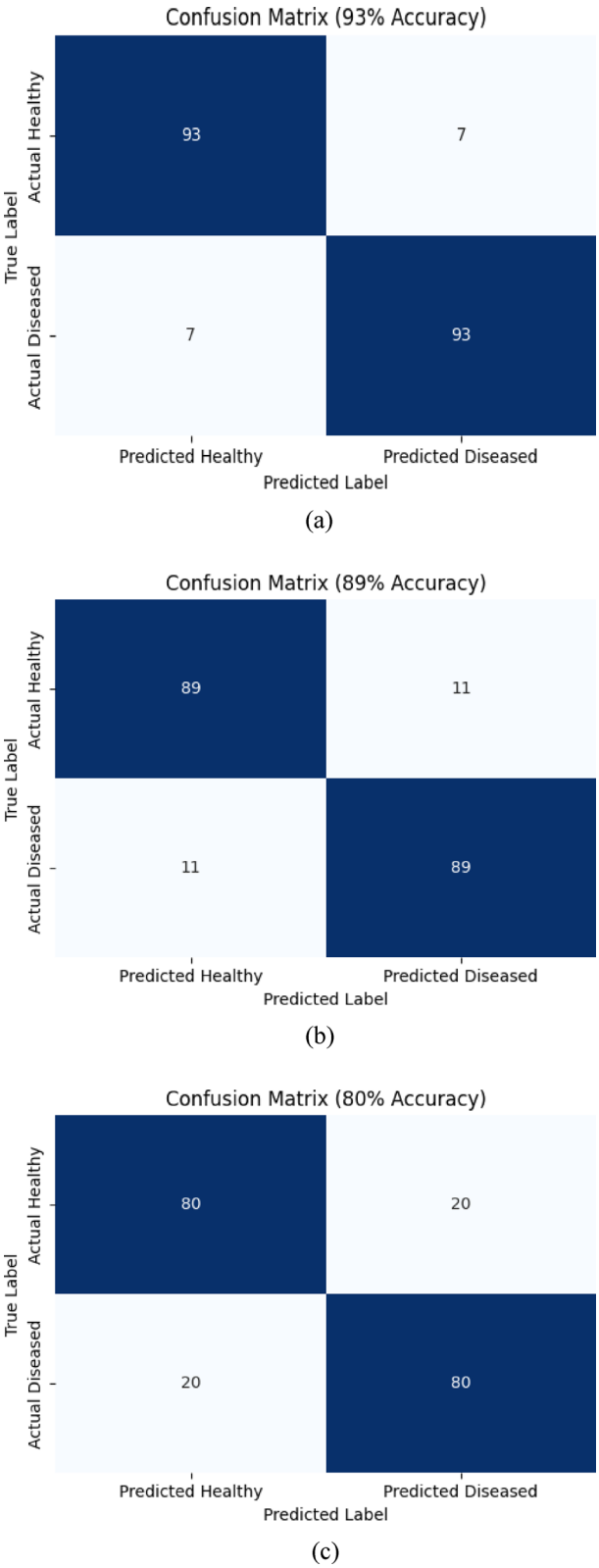


Fig. 7 Confusion matrices illustrating model classification performance: **a** U2-Net with 93% accuracy, **b** U-Net with 89%, and **c** WU-Net with 80%. These matrices highlight true vs. false predictions across lesion classes

Table 5 Quantification summary of confusion matrix and accuracy per model

Model	TP	TN	FP	FN	Accuracy
U ² -Net	93	93	7	7	93%
U-Net	89	89	11	11	89%
WU-Net	80	80	20	20	80%

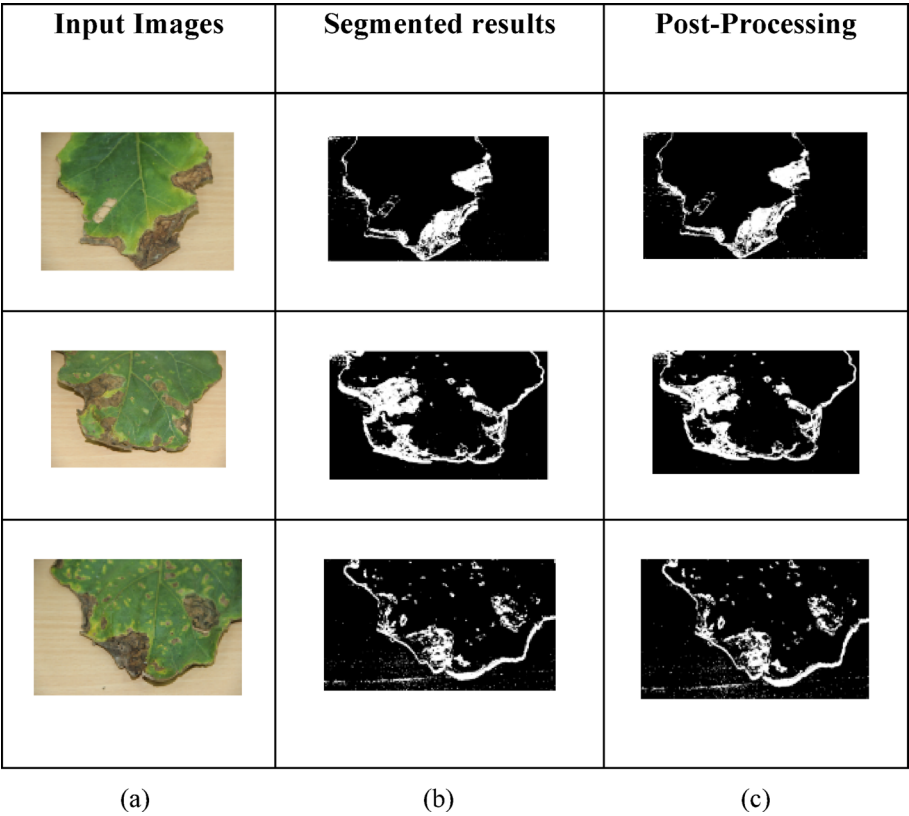


Fig. 8 **a** original eggplant leaf image, **b** segmented lesion region output by the U2-Net model, and **c** post-processed mask after applying morphological operations to refine lesion boundaries

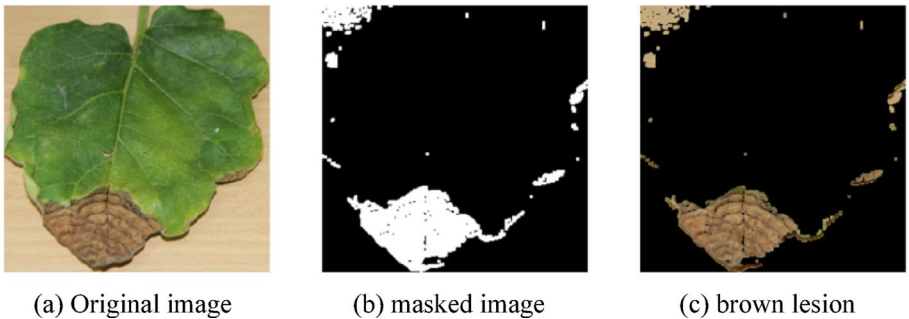


Fig. 9 **a** Original eggplant diseased leaf image, **b** corresponding binary mask showing the diseased area, and **c** extracted brown lesion area

efficiency. MobileNetV2 proved to be the lightest model (3.5 million parameters), offering faster training and compatibility with TensorFlow Lite, albeit with a slightly reduced accuracy of 89.8%. Training times for all models ranged between 70 and 120 min, depending on model complexity, with 35–50 epochs used across the experiments. The

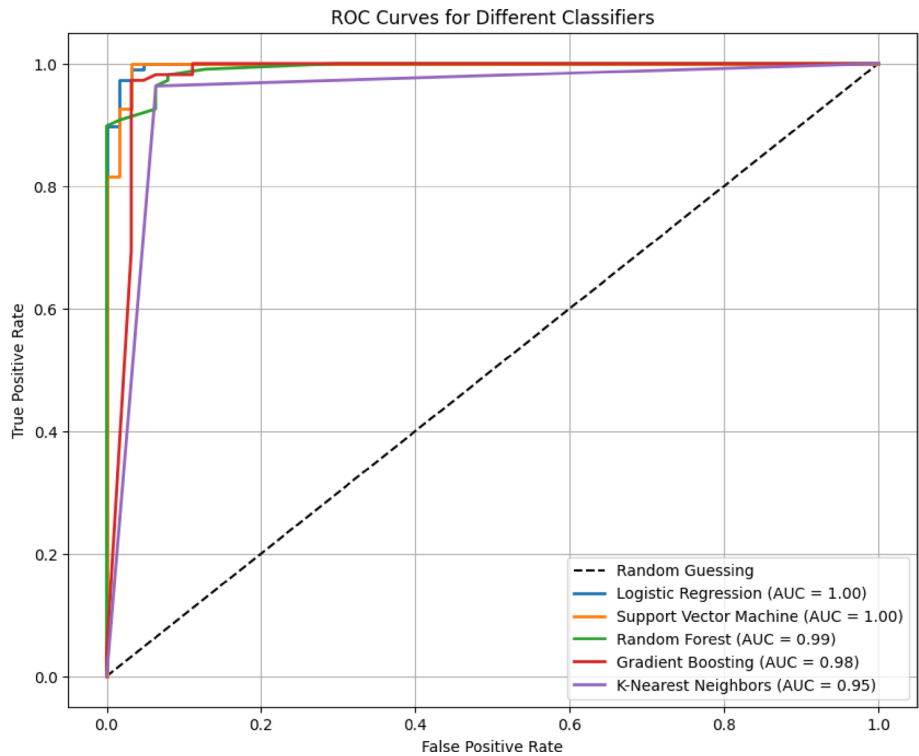


Fig. 10 ROC curve analysis for classification performance across models, with AUC values between 0.95 and 1.0. U2-Net demonstrated improved segmentation, outperforming U-Net and WU-Net in handling complex lesion patterns

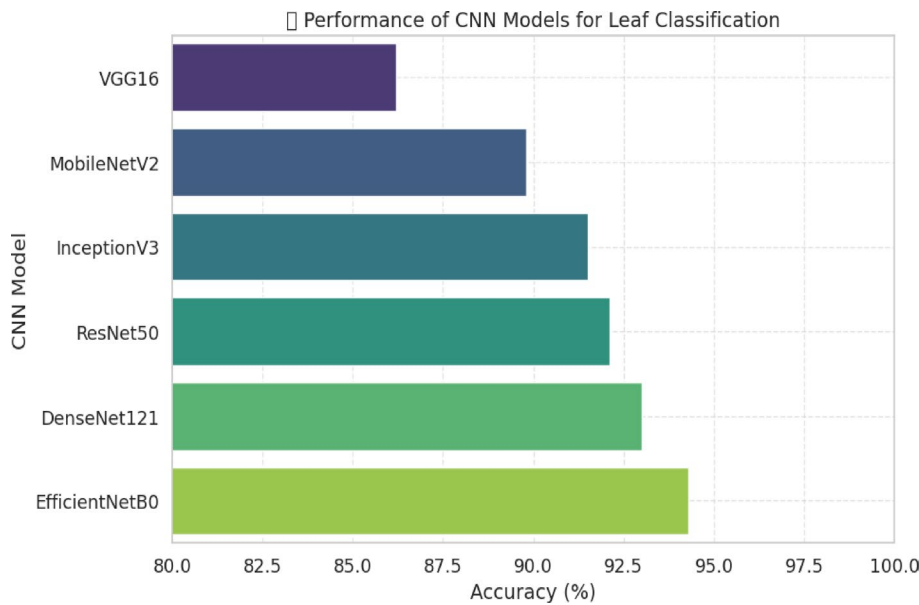


Fig. 11 Comparative performance of six CNN architectures (VGG16, MobileNetV2, InceptionV3, ResNet50, DenseNet121, and EfficientNetB0) for leaf disease classification based on accuracy. EfficientNetB0 achieved the highest accuracy (~94%), indicating its effective performance in identifying leaf diseases

Table 6 Performance comparison of CNN architectures used for brinjal leaf disease classification. Metrics include classification accuracy, number of parameters, training time, number of epochs, and the platform used. EfficientNetB0 achieved the highest accuracy (94.3%) with low model size and training time, indicating suitability for edge deployment

Model	Accuracy (%)	Parameters (M)	Training time (min)	Epochs	Platform
VGG16	87.5	138	120	50	Keras
MobileNetV2	89.8	3.5	70	35	TensorFlow Lite
InceptionV3	90.1	23.9	90	40	Keras
ResNet50	92.5	25.6	100	50	PyTorch
DenseNet121	93.0	7.9	105	50	PyTorch
EfficientNetB0	94.3	5.3	90	40	TensorFlow Lite

Table 7 Comparison of model size, training and inference times, coverage of leaf varieties, and leaf surface employed

Model	Parameters (M)	Training time (50 epochs)	Inference time (per image)	Leaf variety coverage	Leaf surface
U ² -Net	44.3	~2 h	0.3 s	Multiple cultivars	Upper side only
U-Net	31.0	~1.5 h	0.5 s	Multiple cultivars	Upper side only
WU-Net	36.5	~1.8 h	0.4 s	Multiple cultivars	Upper side only
Classifier A (baseline CNN)	5.2	~45 min	0.1 s	Limited (subset only)	Upper side only
Classifier B (ResNet-18)	11.7	~1.2 h	0.15 s	Limited (subset only)	Upper side only

models were implemented using platforms such as PyTorch and Keras. This comparison highlights that models like EfficientNetB0 and MobileNetV2 are optimal choices for real-time, low-resource agricultural applications, while ResNet50 and DenseNet121 are more suited for high-accuracy laboratory environments. Such evaluations help guide the selection of deep learning architectures based on specific application constraints and performance goals.

To complement the visual comparison in Fig. 11, a detailed quantitative summary of each CNN model's performance, complexity, and training requirements is presented in Table 6.

This table summarizes the CNN architectures used for brinjal leaf disease classification. EfficientNetB0 outperformed others in terms of accuracy (94.3%) with relatively low parameter count (5.3 M) and training time, making it optimal for mobile and edge deployment scenarios.

Table 7 represents a comparative analysis of various models used for leaf disease detection, focusing on their computational requirements and the extent of dataset coverage. U²-Net, model, which features a size of 44.3 M, demonstrates a training duration of 2 h across 50 epochs and an inference time of 0.3 s per image, reflecting its effectiveness in modeling. Conversely, the classification-oriented baselines, specifically Classifier A and ResNet-18, exhibit a markedly reduced number of parameters along with accelerated inference speeds; however, they are constrained to a more limited array of cultivars. Importantly, it is noteworthy that all models examined exclusively analysed the adaxial (upper) surface of leaves, thereby underscoring a prevalent limitation associated with surface variability.

4.1 Ablation study

We conducted ablation studies to evaluate the individual contribution of three components:

Table 8 Ablation study results showing the impact of key components in the U2-Net architecture

Model variant	IoU (%)	Dice (%)	Accuracy (%)
U ² -Net (full)	88.5	91.6	93.0
Without RSU blocks	80.2	85.7	88.5
Without data augmentation	76.8	82.0	84.1
Without post-processing	82.4	87.6	89.2

1. Data augmentation
2. Post-processing (morphological operations)
3. U2-net structural components (RSU blocks)

Each variant was tested under controlled conditions using the same dataset and training configuration. The results are summarized in Table 8.

5 Limitations and future work

This study has certain limitations that should be acknowledged. The dataset was constructed from a limited number of original images (200). Although extensive data augmentation was applied, the small base set restricts diversity in terms of lighting, leaf variety, and field conditions, as a result, the generalizability of the models to broader agricultural scenarios may be limited. Practical deployment will require a larger dataset spanning multiple locations and seasons, along with additional validation strategies such as k-fold cross-validation. Detection is currently limited to the upper (sun-exposed) leaf surface and specific cultivars, so expanding coverage to include different leaf surfaces, additional cultivars, and varied field conditions will be essential for real-world applicability. Future work will focus on addressing these limitations to improve robustness, scalability, and generalizability.

6 Conclusion

This study highlights the importance of selecting an appropriate model architecture for image segmentation tasks in plant disease detection. Among the evaluated models, U²-Net consistently outperformed U-Net and WU-Net, achieving the highest segmentation accuracy (Pixel Accuracy=0.98, IoU=0.95, Dice=0.97) and demonstrating advanced ability to capture fine-grained lesion details. With a training accuracy of 96%, test accuracy of 93%, and AUC values ranging from 0.95 to 1.0, U²-Net proved most effective for brinjal leaf fungal disease segmentation and classification, confirming its practical utility in disease diagnosis and management. U²-Net, a robust model, provided the best balance between accuracy and detail preservation while maintaining relatively low computational cost, making it well-suited for pixel-level segmentation in agricultural environments. These findings underscore the model's value in enabling timely intervention and targeted disease management within precision agriculture. Future research may focus on hybrid or attention-based architectures, advanced post-processing methods, larger datasets, and real-time detection strategies to further enhance segmentation accuracy, robustness, and real-world applicability.

Author contributions

Ruchita R Patil wrote the main manuscript. Shwetha V and A Muthusamy critically reviewed the manuscript.

Funding

Open access funding provided by Manipal Academy of Higher Education, Manipal. This work is not supported by any research grants from private or public funding agencies. Open-access publication funding will be provided by the

Manipal Academy of Higher Education, Manipal. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

The datasets generated during and/or analyzed during the current study may be available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 9 April 2025 / Accepted: 24 March 2026

Published online: 03 April 2026

References

1. Khirade SD, Patil AB. Plant disease detection using image processing. In: 2015 international conference on computing communication control and automation, 2015. pp. 768–771. <https://doi.org/10.1109/ICCUBEA.2015.153>.
2. Mwangi RW, Mustafa M, Charles K, Wagara IW, Kappel N. Selected emerging and reemerging plant pathogens affecting the food basket: a threat to food security. *J Agric Food Res.* 2023;14:100827. <https://doi.org/10.1016/j.jafr.2023.100827>.
3. Seybold H, Demetrowitsch TJ, Hassani MA, et al. A fungal pathogen induces systemic susceptibility and systemic shifts in wheat metabolome and microbiome composition. *Nat Commun.* 2020;11:1910. <https://doi.org/10.1038/s41467-020-15633-x>.
4. Kaniyassery A, Goyal A, Thorat SA, Rao MR, Chandrashekar HK, Murali TS, et al. Association of meteorological variables with leaf spot and fruit rot disease incidence in eggplant and YOLOv8-based disease classification. *Ecol Inform.* 2024;83:102809. <https://doi.org/10.1016/j.ecoinf.2024.102809>.
5. Kaniyassery A, Sathish SB, Thorat SA, et al. Deciphering the dynamics and trophic mode distribution of the leaf spot-associated fungal community of eggplant (*Solanum melongena* L.). *Phytopathol Res.* 2024;6:59. <https://doi.org/10.1186/s42483-024-00277-2>.
6. Agrios GN. Plant pathogens and disease: general introduction. In: Schaechter M, editor. *Encyclopedia of microbiology*, 3rd ed.. Academic Press; 2009. pp. 613–646. <https://doi.org/10.1016/B978-012373944-5.00344-8>.
7. Al-Ani LKT, Furtado EL. The effect of incompatible plant pathogens on the host plant. In: Sharma V, Salwan R, Al-Ani LKT, editors. *Molecular aspects of plant beneficial microbes in agriculture*. Academic Press; 2020. pp. 47–57. <https://doi.org/10.1016/B978-0-12-818469-1.00004-3>.
8. Bock CH, Poole GH, Parker PE, Gottwald TR. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Crit Rev Plant Sci.* 2010;29(2):59–107. <https://doi.org/10.1080/07352681003617285>.
9. Mutka AM, Bart RS. Image-based phenotyping of plant disease symptoms. *Front Plant Sci.* 2015;5:734. <https://doi.org/10.3389/fpls.2014.00734>.
10. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. *Comput Electron Agric.* 2018;145:311–8. <https://doi.org/10.1016/j.compag.2018.01.009>.
11. Wang G, Sun Y, Wang J. Automatic image-based plant disease severity estimation using deep learning. *Comput Intell Neurosci.* 2017;2017:2917536. <https://doi.org/10.1155/2017/2917536>.
12. Anim-Ayeko AO, Schillaci C, Lipani A. Automatic blight disease detection in potato (*Solanum tuberosum* L.) and tomato (*Solanum lycopersicum* L. 1753) plants using deep learning. *Smart Agric Technol.* 2023;4:100178. <https://doi.org/10.1016/j.atech.2023.100178>.
13. Dogra R, et al. Deep learning model for detection of brown spot rice leaf disease with smart agriculture. *Comput Electr Eng.* 2023;109(Part A):108659. <https://doi.org/10.1016/j.compeleceng.2023.108659>.
14. Sangeetha R, Logeshwaran J, Rocher J, Lloret J. An improved agro deep learning model for detection of Panama wilt disease in banana leaves. *AgriEngineering.* 2023;5(2):660–79. <https://doi.org/10.3390/agriengineering5020042>.
15. Barbedo JGA. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst Eng.* 2016;144:52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>.
16. Wani JA, Sharma S, Muzamil M, et al. Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: methodologies, applications, and challenges. *Arch Comput Methods Eng.* 2022;29:641–77. <https://doi.org/10.1007/s11831-021-09588-5>.
17. Istead L, Istead J, Pocol A, Kaplan CS. A simple, stroke-based method for gesture drawing. *Virtual Real Intell Hardw.* 2022;4(5):381–99. <https://doi.org/10.1016/j.vrih.2022.08.004>.
18. Zhu Q, Li J, Yuan F, Gan Q. Multiscale temporal network for continuous sign language recognition. *J Electron Imaging.* 2024;33:023059. <https://doi.org/10.1117/1.JEI.33.2.023059>.
19. Nazir A, et al. ECSU-Net: an embedded clustering sliced U-Net coupled with fusing strategy for efficient intervertebral disc segmentation and classification. *IEEE Trans Image Process.* 2021;31:880–93. <https://doi.org/10.1109/TIP.2021.3136619>.
20. Nazir A, et al. OFF-eNET: An optimally fused fully end-to-end network for automatic dense volumetric 3D intracranial blood vessels segmentation. *IEEE Trans Image Process.* 2020;29:7192–202. <https://doi.org/10.1109/TIP.2020.2988741>.
21. Ali R, et al. Optic disk and cup segmentation through fuzzy broad learning system for glaucoma screening. *IEEE Trans Ind Inform.* 2020;17(4):2476–87. <https://doi.org/10.1109/TII.2020.3000204>.

22. Guerroudjia MA, et al. A 3D visualization-based augmented reality application for brain tumor segmentation. *Comput Anim Virtual Worlds*. 2023;35:e2223. <https://doi.org/10.1002/cav.2223>.
23. Kumar AK, Mai NN, Kumar A, et al. Quantum classifier for recognition and identification of leaf profile features. *Eur Phys J D*. 2022;76:110. <https://doi.org/10.1140/epjd/s10053-022-00429-z>.
24. Xu Y, Liu S. Botanical-based simulation of color change in fruit ripening: taking tomato as an example. *Comput Anim Virtual Worlds*. 2024;35(1):e2225. <https://doi.org/10.1002/cav.2225>.
25. Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M. U2-Net: going deeper with nested U-structure for salient object detection. *Pattern Recogn*. 2020;106:107404. <https://doi.org/10.1016/j.patcog.2020.107404>.
26. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: *MICCAI 2015*. Cham: Springer; 2015. pp. 234–241. <https://doi.org/10.48550/arXiv.1505.04597>.
27. Khakimov A, Salakhutdinov I, Omolikhov A, Utaganov S. Traditional and current-prospective methods of agricultural plant diseases detection: a review. *IOP Conf Ser Earth Environ Sci*. 2022;951(1):012002. <https://doi.org/10.1088/1755-1315/951/1/012002>.
28. Ray M, et al. Fungal disease detection in plants: traditional assays, novel diagnostic techniques and biosensors. *Biosens Bioelectron*. 2017;87:708–23. <https://doi.org/10.1016/j.bios.2016.09.032>.
29. Patel R, et al. Plant pathogenicity and associated/related detection systems. A review. *Talanta*. 2023;251:123808. <https://doi.org/10.1016/j.talanta.2022.123808>.
30. Martinelli F, et al. Advanced methods of plant disease detection: a review. *Agron Sustain Dev*. 2015;35:1–25. <https://doi.org/10.1007/s13593-014-0246-1>.
31. Hariharan G, Prasannath K. Recent advances in molecular diagnostics of fungal plant pathogens: a mini review. *Front Cell Infect Microbiol*. 2021;10:600234. <https://doi.org/10.3389/fcimb.2020.600234>.
32. Aslam S, et al. Recent advances in molecular techniques for the identification of phytopathogenic fungi—a mini review. *J Plant Interact*. 2017;12(1):493–504. <https://doi.org/10.1080/17429145.2017.1397205>.
33. Gharbi Y, et al. Technological advancement in the detection and identification of plant pathogens. In: *Emerging technologies in agriculture and food science*. Bentham Science Publishers; 2020. pp. 64–90. <https://doi.org/10.2174/97898114700041200101>.
34. Lin K, Gong L, Huang Y, Liu C, Pan J. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. *Front Plant Sci*. 2019;10:155. <https://doi.org/10.3389/fpls.2019.00155>.
35. Deng Y, Xi H, Zhou G, Chen A, Wang Y, Li L, et al. An effective image-based tomato leaf disease segmentation method using MC-UNet. *Plant Phenomics*. 2023;5:0049. <https://doi.org/10.34133/plantphenomics.0049>.
36. Jia Z, Shi A, Xie G, Mu S. Image segmentation of persimmon leaf diseases based on UNet. In: *7th Int. Conf. Intell. Comput. Signal Process. (ICSP 2022)*. IEEE; 2022. pp. 2036–2039.
37. Fu J, Zhao Y, Wu G. Potato leaf disease segmentation method based on improved UNet. *Appl Sci*. 2023;13(20):11179. <https://doi.org/10.3390/app132011179>.
38. Wang W, Ding J, Shu X, Xu W, Wu Y. FFAE-UNet: an efficient pear leaf disease segmentation network based on u-shaped architecture. *Sensors*. 2025;25(6):1751. <https://doi.org/10.3390/s25061751>.
39. Shwetha V, Bhagwat A, Laxmi V, Shrivastava S. A custom backbone UNet framework with DCGAN augmentation for efficient segmentation of leaf spot diseases in jasmine plant. *J Comput Netw Commun*. 2024. <https://doi.org/10.1155/2024/5057538>.
40. Shwetha V, Bhagwat A, Laxmi V. LeafSpotNet: a deep learning framework for detecting leaf spot disease in jasmine plants. *Artif Intell Agric*. 2024. <https://doi.org/10.1016/j.iaia.2024.02.002>.
41. Atila Ü, Uçar M, Akyol K, Uçar E. Plant leaf disease classification using EfficientNet deep learning model. *Ecol Inform*. 2021;61:101182. <https://doi.org/10.1016/j.ecoinf.2021.101182>.
42. El-Shahat D, Tolba A, Kumar AK. An improved deep learning model for detecting rice diseases. *Optim Agric*. 2024;1:1–10. <https://doi.org/10.61356/j.oia.2024.1194>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.