

# Comparative Analysis of Segmentation Models to Detect Leaf Diseases in Tomato Plant

**Prabhjot Kaur**

Chitkara University Institute of Engineering and Technology, Chitkara University

**Shilpi Harnal**

Chitkara University Institute of Engineering and Technology, Chitkara University

**Vinay Gautam**

Chitkara University Institute of Engineering and Technology, Chitkara University

**Mukund Pratap Singh**

Bennett University

**Santar Pal Singh** (✉ [spsingh78@gmail.com](mailto:spsingh78@gmail.com))

Rashtrakavi Ramdhari Singh Dinkar College of Engineering

---

## Research Article

**Keywords:** Mask R-CNN, Semantic segmentation, Tomato leaf disease, Classification

**Posted Date:** August 1st, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1893425/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

In agriculture around 22% of crop yield loss is due to living and non-living organisms such as biotic and abiotic stress/disease. The early-stage diagnosis of these stresses is an important issue for farmers through naked eyes. Using computer vision technologies can detect the pattern and clustering of diseases at an early stage. However, in the last few years, computer vision-based deep learning technology is helpful for the diagnosis of biotic stress (single biotic and multi biotic) in tomato plant leaves. In this work, the PlantVillage dataset is gathered for the segmentation of object detection. The labeled, enhanced and augmented data has been used for training the model. The proposed hybrid Deep Segmentation Convolutional Neural Network (Hybrid-DSCNN) model has been segmenting the diseased objects in the tomato plant. This Hybrid-DSCNN is assembled using U-Net and Seg-Net pre-trained models with instance segmentation for better detection of objects. The result of the proposed Hybrid DSCNN model is compared with other modified U-Net, M-SegNet, and modified U-SegNet in terms of Accuracy, Precision, Recall, Intersection over Union (IoU), and mean Intersection over Union (mIoU). The proposed model processed 1004 images in 30ns, which is better than other compared models. The accuracy achieved using the proposed model is 98.24%, which is far better than other modified segmentation models. In this paper, the semantic segmented data has been recognized for the single and multiple leaf diseases for identification and classification.

## 1. Introduction

For the vast majority of countries, agriculture and allied industries are their principal source of revenue. A nation's basic and vital needs include the preservation and safety of agricultural supplies. Malnutrition is a constant problem in developing countries like 'India,' and it is closely linked to nutrition security. Agriculture in any country is reliant on the number and quality of its goods, particularly plants. Diseases on plant leaves play a key role in agriculture production, having a considerable impact on crop growth output, resulting in economic losses, environmental losses, and social losses [1]. With output quantities of 21.2 million metric tonnes in 2019–2020 and 19 million metric tonnes in 2018-19, tomatoes, India's second most widely cultivated produce after potatoes, provide a reasonable view on the country's agricultural post-harvest challenges. Tomato diseases are frequent, and they are caused by a range of factors like as weather, temperature, biotic stress, and nutrition. There are up to 15 different types of tomato leaf diseases, according to current statistics, all of which have had a substantial influence on tomato output and quality, as well as inflicting significant economic losses [2]. As a result, tomato leaf disease prevention and treatment are vitally crucial in tomato production. Manually identifying sick symptoms on leaves appears to be a difficult but manageable task. Even the most experienced plant pathologists and agronomists have trouble detecting specific single or multiple diseases. As a result, erroneous assumptions and conclusions are drawn. In plant and agricultural development, the use of plant features to identify physiological changes in response to biotic and abiotic difficulties such as a lack of water, nutrients, poor illumination, bacteria, fungi, and other variables is critical. Plants that have been starved of nutrients will become sick. The nutritional requirements of plants are divided into two

categories: macronutrients and micronutrients. Macronutrients include calcium, nitrogen, sulphur, potassium, phosphorus, and magnesium, among other nutrients. Plants only need trace levels of micronutrients including copper, zinc, chlorine, and iron. The potassium, nitrogen, calcium, and phosphorus deficiencies were determined by counting the leaves, measuring the plant's height, and measuring the thickness and broadness of the leaves. As a result, it will have an effect on the production and quality of the fruit.

A plant may contract a different type of infection from a variety of lesions depending on the source. Some of the most important difficulties that tomato growers face include pests and illnesses that destroy tomato trees, causing losses and lower yield. Experts are usually in charge of determining whether or not chemicals are required for production. Crop protection technologies including artificial intelligence and smartphone apps for disease categorization have become more automated in recent years [3]. As a result, a number of researchers have turned to image processing techniques to assist farmers in completing this difficult task. Tomato leaf disease detection processes now involve automated image processing, as well as expert querying of tomato disease maps. The manual identification of lesions is a time-honored procedure that is costly, time-consuming, and unreliable. Although, thanks to mapping technology, it is now possible to identify related illnesses that have a significant impact on tomato leaf growth [4]. The most modern computer-aided recognition method has yet to be extensively implemented due to its extremely low accuracy, negative environmental impact, and slow recognition speed. As a result, it's vital to develop a tomato disease detection technology that can make quick and accurate decisions about tomato leaf diseases. Because tomato diseases are confined to the leaves, leaf diagnostics can assist growers in determining whether or not to spray their crops. It's worth noting that, thus far, significant progress in diagnosing diseases using leaf features has been made [5].

Among the many ways, Deep Learning (DL) techniques based on Artificial Intelligence (AI) and utilising "Convolutional Neural Network (CNN)" fared well in image identification tests. CNN is a good tool for automatically detecting and classifying diseases/lesions and pests, according to [6]. However, there are challenges with automatic detection and image classification in the field. The intricacy of the background, which matches the subject of the lesion itself, settings such as brightness changes, the angle at which the shots were taken, and the equipment used to capture the photos are all concerns [7] related to the sort of diseased images [8]. [9] has stressed the necessity of finding lesions rather than just classifying them. If the categorization is used alone, the results will be unsatisfactory until it is paired with the location of the lesion. The accuracy in tomato crop leaves was examined using non-segmented and segmented images [10, 11], and it was discovered that the network achieved an accuracy of 98.6% using segmented photos, but only 42.3% using non-segmented images. Photographs from the internet were used to create the test dataset. The segmentation of disease lesions in leaf images has a direct impact on the recognition effect of crop illnesses and the accuracy of a quantitative assessment of the severity of crop diseases. How to section damaged leaves of crops with great efficiency and good quality is a research focus. In the last two decades, traditional image processing techniques such as edge detection, colour space transformation, feature space transformation, and others have been used to extract and recognise lesions.

The creation of a method for identifying tomato leaf disease automatically has been tough. It's worth noting that both diagnosis and detection require processes that could make pinpointing the exact area within the tomato plant where symptoms appear extremely difficult. The capturing parameters are difficult to control, which makes image prediction and disease detection more difficult [12]. Furthermore, the symptoms caused by several diseases may appear to be physically identical, and discrimination methods may be based on little distinctions. In recent years, advances in computer vision have resulted in a new way for accurately diagnosing tomato blemishes [13]. This research examines the various types of biotic and abiotic plant leaf diseases that have been discovered and classified using deep learning technologies at an early stage. Many writers have studied the classification of single biotic and multi-biotic leaf diseases using pre-trained DL models. The authors employed classification models to detect the disease's spots after the initial phase of image processing, which included pre-processing the dataset. The lesions boundaries were not identified accurately with the classification model till now. Thus, this work has proposed the Hybrid-DSCNN model which segments the lesion objects after pre-processing of unstructured data for better accuracy. The proposed approach i.e., Hybrid DSCNN has used the instance and semantic segmentation to detect the foreground object from the background as shown in Fig. 1.

As a result, this work attempts to increase segmentation accuracy with a contribution of:

1. The proposed Hybrid-DSCNN model is used for the segmentation of single biotic and multi biotic leaf disease in tomato plants with the Single and Multi-Biotic Leaf Diseases (SMBLD) algorithm.
2. Hybrid-DSCNN has used a mixed approach to the U-Net and SegNet CNN models by replacing 7 Conv layers of U-Net and 13 Conv layers of SegNet CNN models in downsampling. This work has used the pooling indices (features' vector) from the encoder feature and transfers them for mapping to the corresponding up-sampling layer.
3. The proposed Hybrid-DSCNN model has converted the 5 Conv layers to 2 Conv layers with  $2 \times 2$  kernel size using a skip connection in each layer. Such skip connection method is applied to extract features' information from lower layers.
4. This work has improved expandability as well as general applicability as it is based on pixel-wise remarks using complete deep learning segmentation networks.

Content described in the remaining paper is: Section 2 discusses the literature study of segmentation; Followed section 3 is vital because it defines the method and material of the segmentation used in this work; Section 4 discusses the different evaluation metrics; resulting in part of segmented lesions and comparison of hybrid model with other models is discussed in Section 5; Section 6 discuss the conclusion part.

## 2. Literature Study

Some researchers used information technology to diagnose and identify plant diseases as early as the 1990s. In 1999, some of the authors employed genetic algorithms to identify the condition by

establishing identification factors based on some characteristics such as spectrum reflection and form [14]. In 2007, the authors suggested some machine learning models after the use of neural network models and suggest some parameters for support vector machines and genetic algorithms [15]. In 2011, researchers diagnosed the cotton leaf features based on color and spots which are then combined and discriminated with original images [16]. Using divergence approaches, the authors suggested different feature approaches such as “CYMK color feature”, “GA feature”, and “color and texture”, and then used a hybrid model for the classification task [17]. Using the Inception V3 network in 2017, the authors identified two insect pests and three cassava lesions using a neural network. “Brown leaf spot”, “red mite”, “green mite”, “cassava brown streak disease”, and “cassava mosaic disease” were all recognized at a rate of 98%, 96%, 95%, 98%, and 95%, respectively [18]. In 2018 and 2019, authors suggested different deep learning models for the recognition and identification of a disease with an accuracy rate of 93.4% and 96.64% using the 9-layer model [19–20]. Some researchers provided a classification of symptoms identified in the leaves, with an accuracy rate of more than 90%. Characterization of some lesions problem has also been proposed in other research. Some authors have [21–22] created an application that uses two processes to diagnose illnesses in wheat leaves: localization and classification of the lesion. A technique for detecting illnesses in cucumber leaves was developed by [23]. Their proposed method combines leaf lesion segmentation with CNN-based categorization of segmented regions [24]. To construct a classification system for wheat leaf diseases in smartphone photographs, statistical inference, and segmented images are combined for the detection of disease by some researchers. In [25] suggested a Deep Convolutional Neural Network (DCNN) for the categorization of olive peacock spots and illnesses. The DCNN model is utilized to create the VGG16 and VGG19 architectures. The “Stochastic Gradient Descent” and “RMS Prop optimization algorithm” were used to increase the network’s performance.

In [26], the identification of the tomato plant has been carried out with different CNN frameworks. The approach to strength and system calculation for the identification of a virus at a growing stage is easily identified. For the identification of lesion parts under a clear environment and uneven color analysis, the author in [27] proposed a color fusion balancing method for the detection of lesion leaf. Steps involved firstly, the color transformation method is used. Secondly, for the separation of the unwanted lesion part from the rest of the leaf part the threshold technique is applied. Finally, with the cross-validation method, the classification of images is done with an accuracy of 93.12%. In [28] some machine learning models like SVM, Decision Tree, etc. are used with deep learning models such as CNN for the recognition of the disease in plant leaves. For better results, some deep learning models (VGG16 and VGG19) work with machine learning feature extractors such as K-NN to get an accuracy of 94.9% and 93.5% at an early stage of leaf lesion growth. The main challenging task in the area of agriculture is disease detection which can’t be identified accurately with naked eyes. For the diagnosis and categorization of leaf disease, many ML and DL approaches have been defined in the field of image processing.

In [29] different residual network architectures are defined with the feature extractor model CNN for extracting features of the infected leaf. The mechanism achieved an accuracy of 98% in 5-fold cross-validation. To control as well as detect disease in plant leaves different Machine Learning techniques

such as “Decision tree”, “Naive Bayes”, “SVM”, “KNN”, and “Random Forest” are employed for detection and classification. With these aforesaid models, the highest accuracy with the random forest is 79.23% [30]. The usage of pesticides such as bactericides and fungicides to combat diseases is rapidly increasing, which is having severe consequences for the agricultural environment. The huge occurrence of illnesses in yields has an impact on the amount and quality of yield. The author considers various convolutional neural network models for disease diagnostics to detect disease at early stages [31]. The classification with feature extractor and parameter tuning will give the best accuracy range of 10–15% with Inception V3. In [32] the author suggested the majority of the network structure is made up of residual blocks and attention extraction modules. The model is capable of accurately extracting intricate characteristics of a variety of diseases. According to thorough comparative testing findings, the tomato's average accuracy was 96.81%. After increasing the image's brightness and contrast, the author used segmentation techniques like bipolar thresholding to segment the image data [33]. They combine image processing and soft computing techniques to detect rice plants infected with various diseases. When used correctly, the technique's underlying concept was strong. However, the average identification accuracy across the four datasets was just 82%, indicating that more work is needed.

Following a review of the literature, the authors were more focused on the single disease, single class, and multi-class dataset. But a lesser amount of work has been done on multi diseases dataset. Also, the discussion related to segmentation and classification has been done. All focus is done on the complete leaf area, not on a diseased area only. There is also a need to develop and implement a more generalized plant segmentation method that can be used in both controlled and outdoor environments. The proposed approach i.e., Hybrid DSCNN has used the instance and semantic segmentation to detect the foreground object from the background with single and multi-diseases.

### **3. Material and Method**

This segment explains an overview of the dataset and approach used to detect single biotic and multi biotic stress in tomato plant pictures. Firstly, discuss the dataset utilized in our studies, then will go over the image pre-processing stage briefly. Later, will talk about the collection of attributes retrieved using feature extraction algorithms and classify the stress.

#### **3.1 Data Collection**

The tomato leaf images used for evaluating the performances are downloaded from Plant Village [34] (accessed on 20th Jan 2022). The total number of image datasets obtained from a database with symptoms is summarized in Table 1. The dataset contains single disease and multi diseases leaf images of biotic stress. Moisture, insects, temperature, nutrients, fungus, bacteria, and water influence disease patterns on tomato leaves. In real-time activities, the data collection method is crucial since erroneous data in a dataset can sabotage the experimental outcome. As a result, during the data gathering operation, the common rule should be stated and followed. The entire dataset is partitioned into two datasets: training, a testing ratio of 80:20.

Table 1  
Number of Tomato Leaf Disease Dataset (TLDD)

Disease Class Name	Dataset	Symptoms	Caused due to
Target_Spot	100	Turn from green to red color	Fungus
Leaf_mold	100	Turns into greenish-yellow spots	Fungus
Bacterial_spot	100	Yellow-green to dark brownish-red color	Bacterium
Early_blight	80	Dark round brown red spots	Fungus
Healthy	110	Green color	--
Mosaic_virus	100	Currly, reduce in size	Virus
Septoria_spot	100	Spot become yellow, dark gray from the center and brown from the sides	Fungus
Early Blight and Anthracnose	105	Turn from green to red color and Dark round brown-red spots	Fungus
Late Blight and Powdery mildew	107	Turns into greenish-yellow spots	Fungus
Septoria Spot and Early Blight	102	Spot become yellow, dark gray from center and brown from sides, Dark round brown-red spots	Fungus and Bacterium
Total	1004		

Figure 2 shows some sample images of tomato leaves with different types of disease. Table 1 shows the detail corresponding to these images.

## 3.2 Pre-processing

After examining the photos in the dataset, it has been discovered that the color distribution in the photographs of the leaves has the greatest impact on the final results. Due to the vast range of shot angles and viewpoints from which the photographs were obtained, every spot figureproperty that can influence categorization outcomes have already been altered. As a result, the focus of our future research is to transform the color histograms to generate synthetic images to supplement the training set. Various histogram transforms, like histogramequalization, have been highlighted for improving picture identification in a variety of domains, likeX-ray, satellite, or thermalpictures, because of better variation as well as feature retrieval in low-grade shots. Histogram equalization is mainly used in a picturehavingfewerdivergence areas since it expands the dynamic range of the image. Histogram

equalization techniques like “Adaptive Histogram Equalisation”, “Contrast Limiting Adaptive Histogram Equalisation” (CLAHE), and “Multi-peak Histogram Equalisation” (MPHE) have been used for picture initialization in multiple picture identification applications.

In this study, the authors present a new parameterized image histogram modification method to create pictures for data modification applications. Pictures having changed color-value division can be added to increase the system’s ability to recognize significant color-value division-based characteristics that are not much vulnerable to picture smudging. Furthermore, biological neural networks have a transition that is identical to histogram equalization, making this method of image identification with artificial neural networks a persuasive argument. The suggested method is based on mixing orthogonal functions with color space values-based probability distribution functions (PDFs). Orthogonal functions are a set of fundamental functions that can be used to represent any function in a space of functions. In this case, the Chebyshev orthogonal functions are applied. As far as we know, these functions have never been used for histogram equalization for the problem to process or classify images. “Chebyshev polynomials” have an orthogonal set of coefficients and are a type of ultraspherical polynomial.

The Chebyshev polynomials equation is represented as:

$$P_{x+1}(a) = 2aP_x(a) - P_{x-1}(a) \text{ Eq. (1)}$$

Where  $P_0(a) = 1$  and  $P_1(a) = a$

By multiplying the Chebyshev polynomial in Eq. 1 by the required PDF value and scaling the resultant function until the sum equals 1, the transformation is completed. After this, the modification is applied to histograms of distinct RGB networks in the initial picture. It’s worth mentioning that for different channels, various orthogonal roles might be employed. Then, for each pixel of the image, apply a transform to create a new image.

### 3.3 Object Detection Technique

For the detection of particular objects, the instance segmentation-based Mask R-CNN is used which detects equal intensity objects with better segmentation results in comparison with Faster R-CNN. The network is a progression of “Faster R-CNN” and “Mask Prediction” for the detection of an object [35]. Object detection and semantic segmentation, often known as pixel-level segmentation, are two elements of instance segmentation. Mask R-CNN is utilized for object recognition, and semantic segmentation is done with a fully connected network (FCN) applied to each RoI shown in Fig. 3. The “Region Proposal Network (RPN)” compiled characteristics that are generated by object regions. The proposed regions from the RPN network are then put into the “RoI Align network,” which resizes them. Finally, the result was sent to Fully Connected Network to be classified. This scaled output is also fed into a network that segments the data and creates masks for the things found.

### 3.4 Leaf Diseases Segmentation Network



A convolution Neural Network is the most popular Deep Learning (DL) model for extracting information from photos (CNN). The CNN is effective for the segmentation and classification of certain objects. CNN is the basis for the majority of the networks. The CNN network is made up of several layers, including a convolutional layer, a pooling layer, and a fully connected layer. One of these layers is in charge of the convolutional operator rather than the matrix multiplication result. CNN's operations are based on this operation. Using photos with some filters, the convolutions serve as feature extractors. The filter weights are also optimized during training. The pooling layer's operation of transforming an image's extracted characteristics in such a way that local picture features are kept while other unnecessary information is down-sampled. The Activation function, Sigmoid, and Rectified Linear Unit (ReLU) are some of the functions utilized in CNN. CNN has achieved success in industries and image-related fields as a promising result. Following AlexNet [36], several networks such as GoogleNet, ResNet, SqueezeNet, VGG, and InceptionNet were built based on the CNN network. Because of their ability to extract features, these networks are utilized as the primary support in image processing networks. According to earlier research, the network can bring out superior visual qualities by using a larger convolutional layer. When the number of convolutional layers is few, the picture features obtained are simple. The success of neural networks in a classification task has raised the segmentation challenge. The different CNN networks used for segmentation tasks are explained below in this work.

### 3.4.1 U-Net Model

[37] Proposes the U-Net model, which is based on the VGG network. The U-Net performs better in semantic segmentation and was originally designed for biomedical image segmentation applications. It has two stages: contraction and expansion, as seen in Fig. 4. Because of two primary reasons, this network is often used for segmentation tasks: it performed well with tiny training samples and with data training from beginning to end. The contraction routes of convolutional neural networks are comparable to those of contraction paths. The encoder/contraction network path is made up of two un-padded convolution layers of  $3 \times 3$  with an activation function called ReLU and the max-pool function of  $2 \times 2$  with a stride of 2 in downsampling.

With each downsampling level, the number of feature channels doubles (encoder step). A tiny feature map can capture a large image context using max pooling. At each stage of the decoder network, the number of feature channels is reduced by half. The expansion path is created by combining features maps  $2 \times 2$  from both the contraction and expansion sides, followed by a convolutional layer  $3 \times 3$  and an activation function. At the price of location data, the decoder aids in the collection of more intricate characteristics.  $1 \times 1$  Convolutions are employed in the final layer to convert each 64-dimensional feature vector to the desired number of classes [37].

### 3.4.2 SegNet Model

The Seg-Net network [38] in Fig. 5 is another segmentation network that is encoder-decoder architecture with a total of 26 Convolutional Layers. In the VGG16 network, the contraction path has 13 Convo layers, and the expansion path has the same. The two layers of fully connected (FC) are employed between the

encoder and decoder networks. Rectified Linear Unit is used to create a collection of feature maps (ReLU). In downsampling, this combination is used to a 1024 filter size. After each layer, a stride of 2 is applied to the feature map for downsampling. The number of filter banks/feature channels is doubled during the downsampling process. Every encoder layer has a decoder layer that corresponds to it.

## 3.5 Proposed methodology

The Hybrid Deep Segmentation Convolutional Neural Network Model (Hybrid-DSCNN) network was proposed for the segmentation of leaf diseases in tomato plants in this study. The model's primary skeleton was inspired by the Seg-Net model. The original SegNet design had longer pooling times. Tiny targeted diseases with minor symptoms will diminish or disappear. As a result, the proposed approach originally reduced the number of pooling layers. On the other hand, to improve the detection rate of minor diseases, the model might be reduced in size and the loss of small targets reduced while maintaining accuracy. This work used a hybrid technique of the U-Net and SegNet CNN models, substituting 7 Conv layers of the U-Net and employing 13 Conv layers of the SegNet CNN model. The maximum pool size was  $2 \times 2$  and  $3 \times 3$ , which did not employ the SegNet completely linked layer in the up-sampling layer. The collection of feature maps has been moved to appropriate up-sampling layers using the down-sampling pooling indices. This model swiftly constructed the object's frame and accurately segmented the single disease and multi-disease objects. Also, create a better size of the feature vector for high-density pixels. The Hybrid-DSCNN model has used three sub-variants shown in Fig. 6.

In the basic U-Net model, information from the same level layers is mapped using Skip Connection (SC), resulting in a network of feature maps. Finally, batch normalization was used for each convolutional layer to ensure data stability [37]. This paper presented a Hybrid-DSCNN model based on SegNet and the Skip Connection (SC) mechanism. Table 2 lists the symbolic limits used in the model architectural diagram, with "Convolution Layer (CL)" having a Kernel size of 3 and a padding layer of 0 at the feature map's edges and the activation function name ReLu. With the same size output image, the feature space of a given convolution is down-sampled from 64 to 512 and up-sampled from 512 to 64. Furthermore, the activation for binary issues used the sigmoid function, with the data set to 0–1. The whole architecture is presented in the training section.

Table 2  
Parameters Description.

Code	Description	Parameter Description
CL	“Convolutional Layer”	Kernel Size: (3,3) Activation function: ReLU
BN	“Batch Normalization Layer”	
MP	“Max Pooling Layer”	
UP	“Up Sampling Layer”	
DS	“Down Sampling Layer”	
Output	“Convolutional Layer”	Kernel Size: (1,1)
⊕	“Skip Connection Layer”	

- **3-Layer U-SegNet (3Layer-USN)**

Figure 7 depicts the 3Layer-USN, which was developed utilizing the U-Net Skip Connection (SC) [30] and the 3-Layer SegNet network. After SegNet pooling, an SC connection is used with the lowering of two layers. The purpose of this model is to improve detection capabilities by increasing the image's information during upsampling. Using a max pool of 2, the batch normalization with 64 feature space is down-sampled to 128 and finally to 256. The result is up-sampled with the same image size using a feature space of 256 to 128 and then 64 using a  $2 \times 2$  layer of up-sampling.

- **2-Layer U-SegNet (2Layer-USN)**

As indicated in the network model in the preceding section, the number of convolutions in the Convolutional layer was limited, and the number of pooling layers was lowered while the skip connection was kept. When utilizing a two-layer pooling degree, the purpose was to investigate if there was still some segmentation potential for multi-disease diagnosis. Figure 8 indicates that the 2Layer-USN achieves superior segmentation results than the 3Layer-USN. Three convolution feature space layers were employed, with batch normalizing from 64 to 128 in the first layer and 256 to 512 in the second layer, with a maximum pool of  $2 \times 2$ . The down-sampled output is up-sampled with a filter of 2 and skips connection for improved segmentation.

- **2-Layer-Convo-U-SegNet (2Layer-Convo-USN)**

Figure 9 depicts the extra convolutional layer in the 2Layer-USN model's bottom encoding stage, which was used to extract feature map information from lower-level input in the hopes of enhancing recognition efficiency. In comparison to other models, the extra microscopic particles of illnesses are segmented properly with this extra convolution of 256 feature space. Using a skip connection, the up-sample  $2 \times 2$  localises the pixels from down-sampled to the corresponding layer.

The proposed Hybrid-DSCNN model has been used for the segmentation of single and multi-diseased objects in tomato leaves. The model is subdivided into three different models based upon the size of Convo layers, max-pool  $3 \times 3$ , and  $2 \times 2$  strides. In Hybrid-DSCNN the 2Layer-Convo-USN sub-model achieved better accuracy than the other sub-models with an observation of Convo layers.

This work has implemented the single and multi-disease in tomato leaves. The complete description is given in SMBLD Algorithm. A few different abbreviations used in the algorithm are given in Table 3.

Table 3  
Abbreviation used in SMBLD algorithm.

Variable Name	Expansion of variable
$T_{RGB}, T_{RGB} (Enhance)$	Tomato Input image, Enhance image
$Sb, Sb (n)$	Segmentation Boundary, Segmentation Boundary of n number of object
$F_{V1}, F_{V2}, SG_{BW}$	The feature vector of background, foreground, Semantic segmented image
$M_p (indices), e^{z_i} \text{ and } e^{z_j}$	Max pool indices, exponent variable of softmax.
CE, and BCE	Cross-Entropy and Binary Cross-Entropy.
$T (Mask),$ $sT (Mask), mT (Mask)$	Masked Object, Masked Object for single leaf, Masked Object for multi leaves.

---

**Algorithm 1:** Hybrid-DSCNN model -Single and Multi-Biotic Leaf Diseases (SMBLD)

---

**Input:**  $T_{RGB}, G_{BW}, A$ .**Output:**  $F_{V1}$  and  $F_{V2}$  as a feature vector for instance segmentation image.

---

1. Take input  $T_{RGB}$  as  $256 \times 256 \times 3$  and pre-processed using resize, enhancement of image as  $224 \times 224 \times 3$  in RGB color image.
  2. Achieve the  $T_{RGB}(Enhance)$  image from  $T_{RGB}$  image.
  3. Set  $B_L$  as the boundary of the leaf image.
  4. Draw segmentation boundary ( $Sb$ ), from ( $T_{RGB}(Enhance)$ ) and achieve  $Sb(n)$ , where ( $n = 0, 1, 2, 3, 4, \dots, n - 1$ ) is the number of tomato leaves for the boundary vegetation.
  5. Create the mask as  $T(Mask)$  from  $Sb(n)$  images.
  6. Differentiates  $T(Mask)$  and  $mT(Mask)$  from  $T(Mask)$  for single disease leaf and multi diseases leaf, respectively by following steps from 7 to 16.
  7. If  $mT(Mask) =$  high – density pixel
  8. then
  9. put vegetation pixel in  $F_{V1}$ .
  10. Print “Multi-diseases leaf”
  11. Elseif  $sT(Mask) =$  high – density pixel
  12. then
  13. put vegetation pixel in  $F_{V2}$ .
  14. Print “Single-diseases leaf”
  15. else
  16. Print “Healthy leaf”.
  17. Set the 2700 threshold value of the pixel feature vector  $F_{V1}$  and  $F_{V2}$  and avoid over-segmentation.
  18. Set C36 Convo and collect feature map  $F_{V1}$  and  $F_{V2}$ .
  19. Transfer the  $F_{V1}$  and  $F_{V2}$  to corresponding up-sampling Convo layer
  20. For the final pixel-wise classification and validation add  $P = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}}$  as Softmax layer, where  $e^{z_i}$  and  $e^{z_j}$  are two exponential variables.
  21. For validation, calculate the values of significant parameters such as CE, BCE, IOU, and mIoU from P and compare them to basic modified U-Net and M Seg-Net, as well as modified U-SegNet segmentation models.
  22. Return  $SG_{BW}$  and significant value.
  23. Stop.
- 

This work has been taken Take input  $T_{RGB}$  and pre-processed using resize, and enhancement of the image. Furthermore, set the  $B_L$  is the boundary of leaf image and segmentation boundary ( $Sb$ ), from ( $T_{RGB}(Enhance)$ ) and achieve  $Sb(n)$  where ( $n = 0, 1, 2, 3, 4, \dots, n - 1$ ) is the number of tomato leaves for the boundary vegetation. For the discrimination of single and multi-object diseases, use  $sT(Mask)$  and  $mT(Mask)$  from  $Sb(n)$  image and achieved segmented parts using feature vectors ( $F_{V1}$  and  $F_{V2}$ ). Then, collect high-density pixels from feature vectors and implement the Hybrid-DSCNN model. As a result, the SMBLD algorithm returns  $SG_{BW}$  and significant value. Next, the values of the significant classification parameters such as CE, BCE, IOU, and mIoU are compared with earlier proposed segmentation models by various authors for validation. The complete description is given in Fig. 10.

## 4. Performance Metric

In this study, a total of nine assessment criteria are utilized to analyze and compare the performance of the Hybrid-DSCNN model with different proposed models. “Precision”, “Recall”, “Dice coefficient”, “F1-

Score”, “Average Precision”, “Mean Average Precision”, “Intersection over Union (IoU)”, “Mean Intersection over Union (mIoU)” and “Accuracy” [39] are the metrics utilized for the analysis.

## 4.1 Quantitative Analysis

For the analysis of data using digital information is done with some criteria called quantitative analysis. Analysis of data is expressed in different terms, such as; “True Positive ( $T_p$ )”, “False Positive ( $F_p$ )”, “True Negative ( $T_n$ )”, “False Negative ( $F_n$ )”.  $T_p$  represents the actual true value i.e., the masked infected area is segmented as the actual infected area,  $F_p$  predicts the tomato leaf disease as non-damaged or healthy pixels,  $F_n$  the false negative is tomato leaf diseased pixels that are present in the actual truth area of the leaf but not identified by an algorithm, and  $T_n$  represents the non-infected part as actual and as non-infected.

$$\text{Precision} = \frac{T_p}{T_p + F_p} \text{ Eq. (2)}$$

$$\text{Recall} = \frac{T_p}{T_p + F_n} \text{ Eq. (3)}$$

Equation (2) shows the ratio of predicted actual positive values with the total value calculated from the given image dataset is called “Precision”. “Recall value” is the calculation of positive value w.r.t. the predicted value of complete one class which is shown in Eq. (3).

$$\text{F1 - Score} = 2 \times \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \text{ Eq. (4)}$$

For detecting the performance of infected regions in a dataset, the mean of recall and precision gives the output as F1-Score shown in Eq. (4).

For every single image dataset, the “Average Precision (AP)” is calculated as the ratio of the number of actual correct data values to the sum of all pixel numbers the ground-truth value as is shown in Eq. (5).

$$\text{Average Precision (AP)} = \frac{T_p}{T_p + F_n + F_p} \text{ Eq. (5)}$$

When a dataset has many classes (as in this study), for evaluating the performance of a model a single number is essential which is called “mAP” [46]. The “Mean Average Precision (mAP)” is calculated in Eq. (6) as the average of all classes given in the dataset.

$$\text{mAP} = \frac{1}{M} \sum_{x=1}^M \text{AP}_x \text{ Eq. (6)}$$

The  $\sum_{x=1}^M \text{AP}_x$  represents the average precision sum of all images in the tomato leaf disease dataset.

Accuracy is one of the parameters for evaluating the classification model. Informally, it is a degree of the way the neural network changed into positive detection and the way by which it becomes calculated is presented in Eq. (7).

$$\text{Accuracy} = \frac{\sum T_p + \sum T_n}{\sum T_p + \sum T_n + \sum F_p + \sum F_n} \times 100 \text{ Eq. (7)}$$

Formally, the total true values both positive and negative divided by the total true and false values based on the detection result. With these parameters, the analysis is also done based on segmentation results such as IoU and mIoU.

## 4.2 Segmentation Analysis

The segmentation analysis is based upon two basic parameters Intersection over Union (IoU) Eq. (8) and mean intersection over union (mIoU) Eq. (9). The IoU is the statistic for determining how accurate an item detector is on a given data. The mIoU is a typical semantic image segmentation assessment metric that computes the averaging of IoU for each semantic over all classes.

$$\text{IoU} = \frac{\sum T_p}{\sum T_p + \sum F_p + \sum F_n} \times 100 \text{ Eq. (8)}$$

$$mIoU = \frac{1}{N_{class}} \sum_i IoU \text{ Eq. (9)}$$

The  $\frac{1}{N_{class}}$  represents a total of all classes. The mean of all classes over total values corresponding to given class values is represented in terms of mean intersection over the union. Based on these evaluation parameters the result of single biotic diseases and multi biotic diseases are shown in the result section.

## 5. Result And Discussion

### 5.1 Result of single biotic disease in tomato plant leaf

Only the fine-tuning set is used in the training process. In 100 cycles, it was clear that the training set's accuracy and loss were different from the validation set. The losses of validation and training sets were gradually lowered as the number of training iterations increased, and the accuracy was gradually increased shown in Fig. 11. The number of samples for training was sufficient. Before applying the semantic segmentation technique to a specific disease in the tomato leaf dataset, the Mask R-CNN extracts feature maps from the backbone ResNet50. The data is sent to the Region Proposal Network by the feature extractor (RPN). Images of various sizes can be sent to the RPN. The outputs included some polygon object recommendations. To produce a proposal network, the final shared layer slides a little network over the convolutional feature map output. The sliding window of the higher dimension is communicated to the lower dimension via RPN. The Region of Interest (ROI) pooling layer received RPN's offers. From a variety of input sizes, ROI layers generate fixed-size feature maps. The layer's settings influence the fixed dimension of the ROI pooling output. Finally, this feature was input into the box-classification layer (Classifier) and the box-regression layer (Box-Regression) (regressor). The classifier's regressor was an updated bounding box, and the regressor categorized the objects.

Figure 12 shows how the high-density pixels obtained from the last layer of the Mask R-CNN are input into the Hybrid DSCNN, a semantic segmentation model. The Hybrid DSCNN model distinguishes illness spots with varying color intensities for better segmentation and classification. The Adam optimizer was used to set the network's parameters, which had a fixed learning frequency of 0.0001, a feature vector size of 2700 px, and compactness of 25 px. For the various forms of augmentation processes used in all models, the batch size is 32, and the parameter values are rotation range 90° and 180°, width range 0.2, zoom range 0.2, and height shift 0.2.

$Ann \times n$  spatial window convolutional feature map is used as the input for this small network. At each sliding-window position, many region ideas are projected simultaneously. The sliding window of the higher dimension is communicated to the lower dimension via RPN. The ROI pooling layer receives RPN's offers. From a variety of input sizes, ROI pooling layers generate fixed-size feature maps. The layer's settings influence the fixed dimension of the ROI pooling output. The Hybrid DSCNN model is then used to segment the high-density pixels, which has a higher classification and accuracy rate.

## 5.2 Result of multibiotic diseases in tomato plant leaf

The Tomato Leaf Disease Dataset (TLDD) contains images that were scaled from  $256 \times 256 \times 3$  pixels to  $224 \times 224 \times 3$  pixels and image brightness was increased using various pre-processing techniques. The 880 photos were divided into 8:2 ratios and the Hybrid-DSCNN model was used to train the model. First, the instance segmentation technique employed Mask R-CNN to generate the object's boundary boxes for the identification of multi-biotic tomato leaf illnesses, as indicated in the methodology section. The high-density pixels were collected into the feature vector, which was achieved using background pixel discrimination. In the feature vector, a comparable sort of high-density pixel is clustered together. Second, semantic segmentation is implemented to improve object identification and classification of numerous diseases in the tomato plant. For the discrimination of many illnesses in a particular leaf, the feature vectors vary the intensity value of pixels using a color feature chart [30]. For the various forms of augmentation techniques used in all models, the batch size is 32, and the parameter values for rotation range are 90° and 180°, width range of 0.2, zoom range of 0.2, and height shift of 0.2. In 100 cycles, it was clear that the training set's accuracy and loss were different from the validation set. As the training progressed, the gap between the two increasingly widened. The accuracy did not improve, and the loss of the validation set did not decrease as shown in Fig. 13. This meant the network becomes overfit and the training samples were insufficient. As a result, data augmentation was required.

The factor that can affect the metrics is that the data is labeled by hand, which could lead to errors because the delimitation of the different lesions areas was not well enhanced, given how difficult it is to distinguish between these two separate lesion locations and the healthy leaf tissue area is shown in Fig. 14.

## 5.3 Comparison with earlier proposed Segmentation models



The state-of-art various image segmentation methods such as modified U-Net [40], modified U-SegNet [41], and M-SegNet [42] was used for the segmentation of the tomato leaf diseases based on the dataset of single and multi-diseases leaf images of biotic stress. Table 3 and Fig. 15 show the results in terms of evaluation metrics. As seen in Table 4, the proposed modified U-Net shows a comparatively low result among all other segmentation models. It could be seen from the output that the modified U-Net could not segment the objects with better precision and the recall rate is very low. The multi-biotic diseases were not identified accurately with modified U-Net in many areas. Hence, the overlapped diseases were not detected and segmented efficiently. In comparison with this, the M-SegNet improves the IoU and the pixel-wise accuracy for the segmentation. The pixel-wise object detection properties were segmenting the objects in a very precise manner using M-SegNet; this increases the amount of segmentation information.

Table 4  
Comparison of different segmentation methods with the Hybrid DSCNN method

Method Name	Dataset Name	Accuracy	Precision	Recall	IoU	Segment Time (ms)
Modified U-Net [40]	TLDD	93.41	84.04	75.15	79.96	67.80
M-SegNet [42]		94.91	89.99	72.35	80.13	82.52
Modified U-SegNet [41]		97.88	93.98	89.67	91.13	71.30
Proposed Hybrid DSCNN		98.24	92.83	94.36	92.91	51.71

Using skip connection in modified U-Net, the network discards some of the information that comes from the downsampling layer towards the up-sampling layer. This lack of information results in bad output. The encode section was streamlined, and the U-decoding Net's section was simplified as well.

The four sections of the up-sample were reduced to three. Because fewer abstract features were extracted, in the decoding stage, redundant convolutional layers were not required to produce abstract feature layers. As a result, following up-sampling, decrease the two subsequent convolutional layers to one. To increase the segmentation accuracy of details on the images, the input layer and the final convolutional layer were concatenated at the same time. The more pooling layers in M-SegNet transfer the segmented data using pooling indices from the down-sampling layer to up-sampling for the correlation of data. Therefore, the segmentation performance becomes better than the modified U-Net. The M-SegNet architecture consists of an attention-based mechanism with a multi-scale network at the down-sampling side and supervision at the up-sampling side with different kernel sizes. By merging local characteristics with their respective global dependencies, the global attention method is used to gather rich contextual information in the up-sampling stage. In addition, in both the contraction and expansion modules, multiscale convolutional kernels of various sizes were employed to extract several semantic features from the dataset.

The result with modified U-SegNet is better as compared with modified U-Net and M-SegNet in Fig. 16. The pixel-wise object detection using modified U-SegNet shows a higher accuracy rate. The proposed squeeze U-SegNet is designed using multi-scale and multi-global modules. At each encoding layer, the multi-scale feature characteristics encode the local and global contexts. Furthermore, the extraneous data and concentrated information were filtered out using a new global module at the encoder-decoder side for the segmentation task. When the whole image data is used as an input to sequence the network, the model loses the irrelevant local information. To tackle this problem, propose a distribution of each input data slice, which is used for training the model and achieving better segmentation accuracy. Using these semantic models has some limitations were there. The traditional usage of encoder-decoder techniques, for instance, results in the extraction of comparable low-level features numerous times, resulting in redundant information utilization. Furthermore, each semantic class is likely to be connected with non-accurate discriminative feature representations as a result of inefficient modeling of long-range dependencies, resulting in low segmentation accuracy shown in Fig. 16. The proposed Hybrid DSCNN model refines the segmentation and feature extraction representational power of segmentation models with better accuracy.

The edges of the diseased area are sharper in M-SegNet and modified U-Net than Hybrid DSCNN, as seen in Fig. 17 based on multi-disease detection. Because some pixels can end up being erroneous, these rounded edges could be a role in M-SegNet and modified U-Net somewhat lower performance when compared to the modified U-SegNet model. The factor that has affected the metrics' value is due to unclear boundaries of the lesions' area and the labeling of lesion regions. When looking at the data, it's evident that the different frameworks of the Hybrid DSCNN model worked well and correctly detected the lesions in the vast majority of cases.

The results are more precise, with just minor categorization errors. The results obtained exhibited more inaccuracies as the foliage density increased and the influence of light increased. Despite successfully detecting part of the lesion, the network misclassified other parts of the image as a lesion with a lower confidence rate when detecting multi-diseases as shown in Fig. 17. With this segmentation task, the tiny lesion areas with two different types are segmented in different colors. The images under different environmental conditions are segmented and the model discerns all of the multiple objects very clearly.

## 6. Conclusion

Due to several connected aspects such as ambient lighting, complicated backdrop, variety of symptom features, and more, the segmentation and detection of single biotic and multi-biotic lesion spots of tomato leaves is a highly challenging task. This research presented a Hybrid-DSCNN model with the addition of an instance segmentation step to deal with the backdrop complexity of real-world images. A data augmentation strategy was devised to tackle the difficulties in data labeling and the scarcity of training data. With the analysis of the experiment, it is found that:

1. The result showed that the Hybrid-DSCNN model achieved an accuracy of 98.24% which is far better than other modified segmentation models.
2. The hybrid model can segment the lesion objects. The segmentation parameter IoU reached 92.91% with a time utilization of 51.71ms which is less than compared with other models.
3. Combining the foreground and background data sets could substantially enrich the data collection. It has the potential to reduce overfitting caused by a lack of training samples.
4. The proposed method could accurately segment the multi and single objects in any environmental condition.
5. The segmentation of higher pixel ranges is accurate with the Hybrid-DSCNN model, but the segmentation of lower pixel ranges is inaccurate.

One of the current approach's limitations in terms of semantic segmentation is the inability to categorize areas when lesions overlap. The other factor is the small number of image datasets in this hybrid model. Other image segmentation techniques can be investigated in the future for better identification, categorization, and detection of tiny biotic and abiotic lesions. Different size datasets with different morphological properties can be used for more accuracy.

## Declarations

**Ethical Approval:** Not applicable

**Author Contributions:** **PK.:** Conceptualization, data collection and Methodology, Writing, **SH.:**Methodology, **VG:** Conceptualization and Supervision, **MPS:** Validation and supervision, **SPS:** validation. All authors have read and agreed to the published version of the manuscript.

**Funding Information:** Not applicable

**Data Availability Statement:** Data and source codes are available from the authors upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

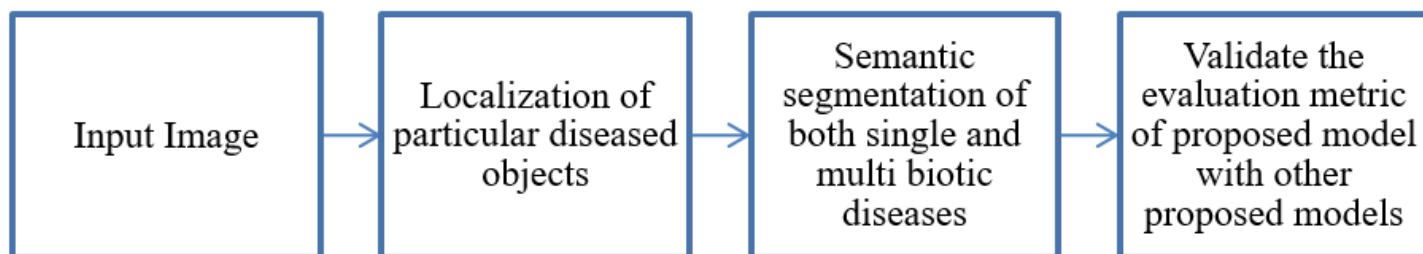
1. Yadav, S., Sengar, N., Singh, A., Singh, A., and Dutta, M. K. (2021). Identification of disease using deep learning and evaluation of bacteriosis in peach leaf. *Ecological Informatics*, 61, 101247.
2. Khan, S., and Narvekar, M. (2020). Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment. *Journal of King Saud University-Computer and Information Sciences*.
3. Esgario, J.G., Krohling, R.A., Ventura, J.A., 2020. Deep learning for classification and severity estimation of coffee leaf biotic stress. *Comput. Electron. Agric.* 169, 105162.

4. Agarwal, M., Gupta, S. K., and Biswas, K. K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics and Systems*, 28, 100407.
5. Chen, X., Zhou, G., Chen, A., Yi, J., Zhang, W., and Hu, Y. (2020). Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet. *Computers and Electronics in Agriculture*, 178, 105730.
6. Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* 153, 46–53.
7. Barbedo, J.G.A., 2016. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosyst. Eng.* 144, 52–60.
8. Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A., and Alabdali, A. M. (2022). Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction. *Sensors*, 22(2), 575.
9. Boulent, J., Foucher, S., Theau, J., St-Charles, P-L., 2019. Convolutional neural networks for the automatic identification of plant diseases. *Front. Plant Sci.* 10, 941.
10. Kaur, P., and Gautam, V. (2021). Research patterns and trends in classification of biotic and abiotic stress in plant leaf. *Materials Today: Proceedings*, 45, 4377-4382.
11. Sharma, P., Berwal, Y.P.S., Ghai, W., 2019. Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. *Information Processing in Agriculture*.
12. Salih, T. A. (2020). Deep Learning Convolution Neural Network to Detect and Classify Tomato Plant Leaf Diseases. *Open Access Library Journal*, 7(05), 1.
13. Thangaraj, R., Anandamurugan, S., and Kaliappan, V. K. (2021). Automated tomato leaf disease classification using transfer learning-based deep convolution neural network. *Journal of Plant Diseases and Protection*, 128(1), 73-86.
14. Sasaki, Y., Okamoto, T., IMOU, K., and TORII, T. (1999). Automatic diagnosis of plant disease recognition between healthy and diseased leaf. *Journal of the Japanese Society of Agricultural Machinery*, 61(2), 119-126.
15. Sammany, M., and Medhat, T. (2007, June). Dimensionality reduction using rough set approach for two neural networks-based applications. In *International Conference on Rough Sets and Intelligent Systems Paradigms* (pp. 639-647). Springer, Berlin, Heidelberg.
16. Gulhane, V. A., and Gurjar, A. A. (2011). Detection of diseases on cotton leaves and its possible diagnosis. *International Journal of Image Processing (IJIP)*, 5(5), 590-598.
17. Revathi, P., and Hemalatha, M. (2014). Cotton leaf spot diseases detection utilizing feature selection with skew divergence method. *International Journal of scientific engineering and technology*, 3(1), 22-30.
18. Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., and Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in plant science*, 8, 1852.

19. Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., and Sun, Z. (2018). A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Computers and electronics in agriculture*, 154, 18-24.
20. Geetharamani, G., and Pandian, A. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers and Electrical Engineering*, 76, 323-338.
21. Lu, J., Hu, J., Zhao, G., Mei, F., Zhang, C., 2017. An in-field automatic wheat disease diagnosis system. *Comput. Electron. Agric.* 142, 369–379.
22. Kaur, P., and Gautam, V. (2021). Plant Biotic Disease Identification and Classification Based on Leaf Image: A Review. In *Proceedings of 3rd International Conference on Computing Informatics and Networks: ICCIN 2020* (pp. 597-610). Springer Singapore.
23. Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., Sun, Z., 2018. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Comput. Electron. Agric.* 154, 18–24.
24. Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A.D., Ortiz-Barredo, A., 2017. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Comput. Electron. Agric.* 138, 200–209.
25. Uguz, S., and Uysal, N. (2021). Classification of olive leaf diseases using deep convolutional neural networks. *Neural Computing and Applications*, 33(9), 4133-4149.
26. Mkonyi, L., Rubanga, D., Richard, M., Zekeya, N., Sawahiko, S., Maiseli, B., and Machuve, D. (2020). Early identification of *Tuta absoluta* in tomato plants using deep learning. *Scientific African*, 10, e00590.
27. Khan, S., and Narvekar, M. (2020). Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment. *Journal of King Saud University-Computer and Information Sciences*.
28. Agarwal, M., Gupta, S. K., and Biswas, K. K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics and Systems*, 28, 100407.
29. Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., and Menaka, R. (2020). Attention embedded residual CNN for disease detection in tomato leaves. *Applied Soft Computing*, 86, 105933.
30. Panigrahi, K. P., Das, H., Sahoo, A. K., and Moharana, S. C. (2020). Maize leaf disease detection and classification using machine learning algorithms. In *Progress in Computing, Analytics and Networking* (pp. 659-669). Springer, Singapore.
31. Ahmad, I., Hamid, M., Yousaf, S., Shah, S. T., and Ahmad, M. O. (2020). Optimizing pretrained convolutional neural networks for tomato leaf disease detection. *Complexity*, 2020.
32. Zhao, S., Peng, Y., Liu, J., and Wu, S. (2021). Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module. *Agriculture*, 11(7), 651.
33. Phadikar, S., and Sil, J. (2008, December). Rice disease identification using pattern recognition techniques. In *2008 11th International Conference on Computer and Information Technology* (pp. 420-423). IEEE.

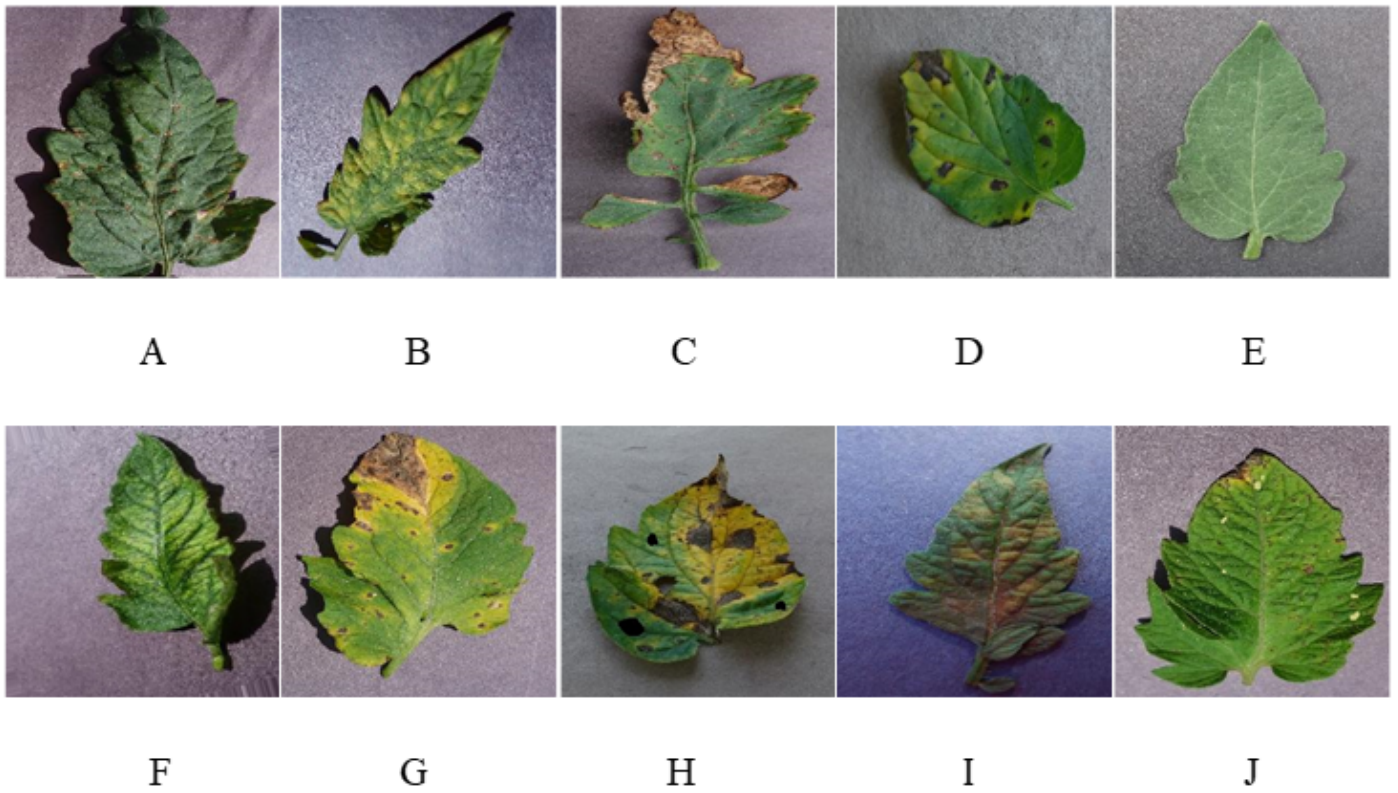
34. Hughes, D.P.; Salathe, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv 2015, arXiv:1511.08060.
35. He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. arXiv e-prints, arXiv:1703.06870.
36. He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.
37. Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. arXiv e-prints, arXiv:1505.04597.
38. Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 2481–2495.
39. Mishra, A. M., Harnal, S., Mohiuddin, K., Gautam, V., Nasr, O. A., Goyal, N., and Singh, A. (2022). A Deep Learning-Based Novel Approach for Weed Growth Estimation. intelligent automation and soft computing, 31(2), 1157-1172.
40. Zou, K., Chen, X., Wang, Y., Zhang, C., and Zhang, F. (2021). A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field. Computers and Electronics in Agriculture, 187, 106242.
41. Dayananda, C., Choi, J. Y., and Lee, B. (2021). Multi-Scale Squeeze U-SegNet with Multi Global Attention for Brain MRI Segmentation. Sensors, 21(10), 3363.
42. Yamanakkanavar, N., and Lee, B. (2021). A novel M-SegNet with global attention CNN architecture for automatic segmentation of brain MRI. Computers in Biology and Medicine, 136, 104761.

## Figures



**Figure 1**

Complete Flow of the work.



**Figure 2**

Sample tomato leaf images. (A) Bacterial\_spot, (B) Leaf\_mold, (C) Septoria\_leaf\_spot, (D) Target\_spot, (E) Tomato\_mosaic\_virus, (F) Healthy, (G) Early\_blight, (H) Early Blight and Anthracnose, (I) Late Blight and Powdery mildew, (J) Septoria Spot and Early Blight.

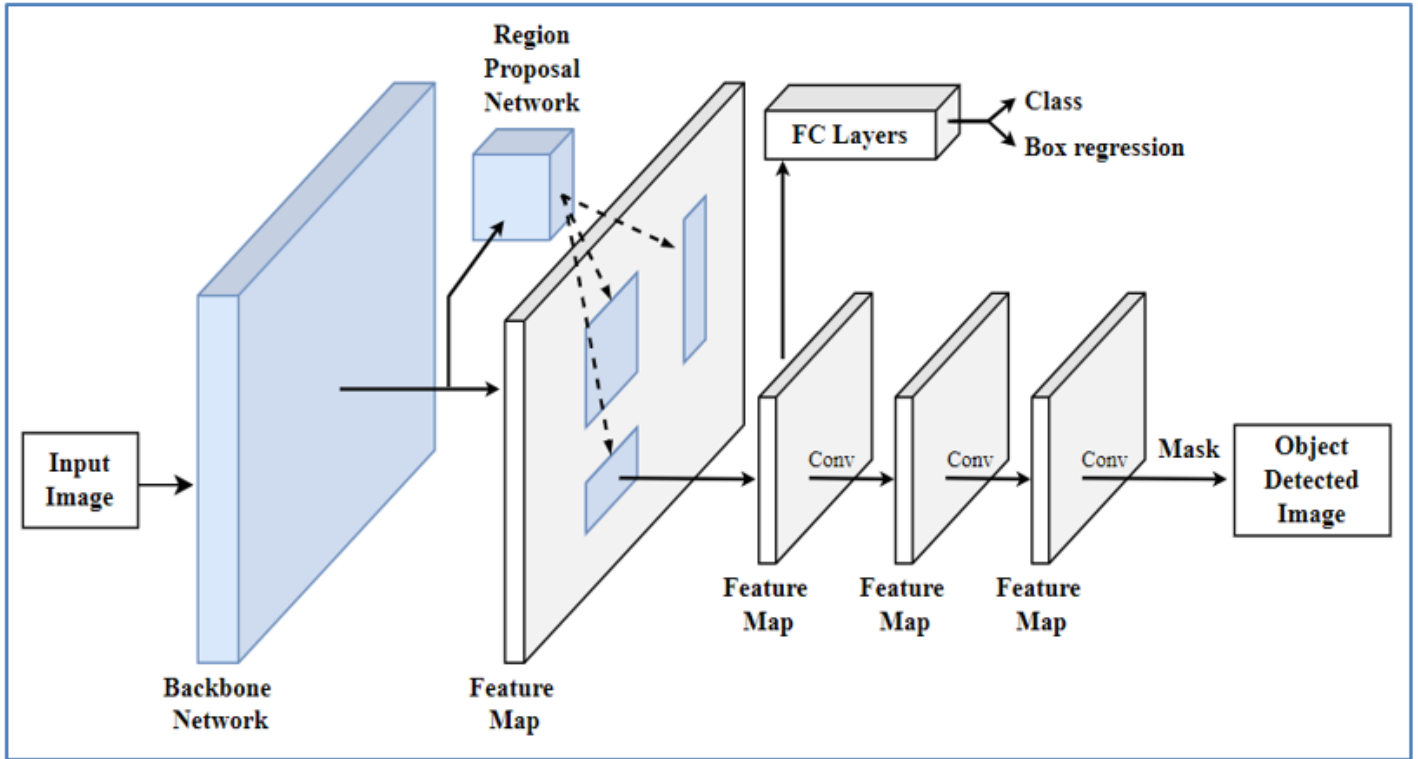


Figure 3

Mask R-CNN architecture.

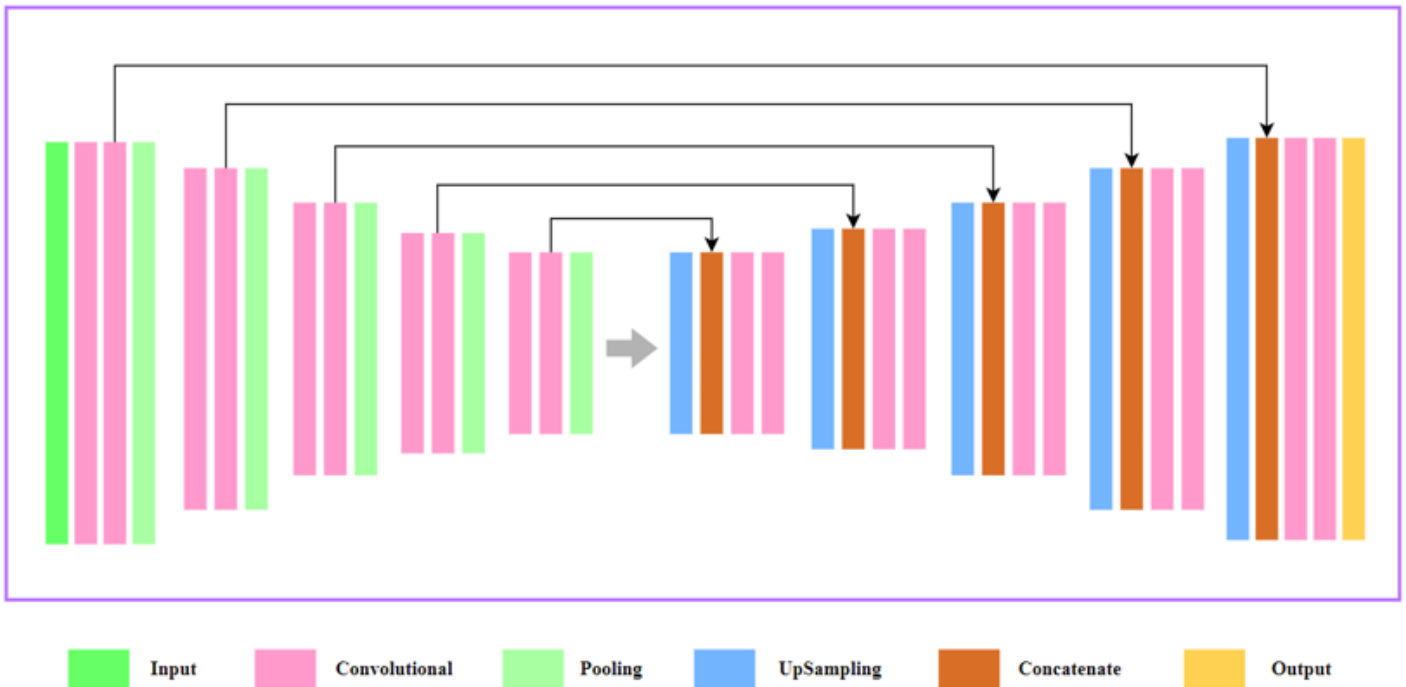


Figure 4



U-Net network architecture.

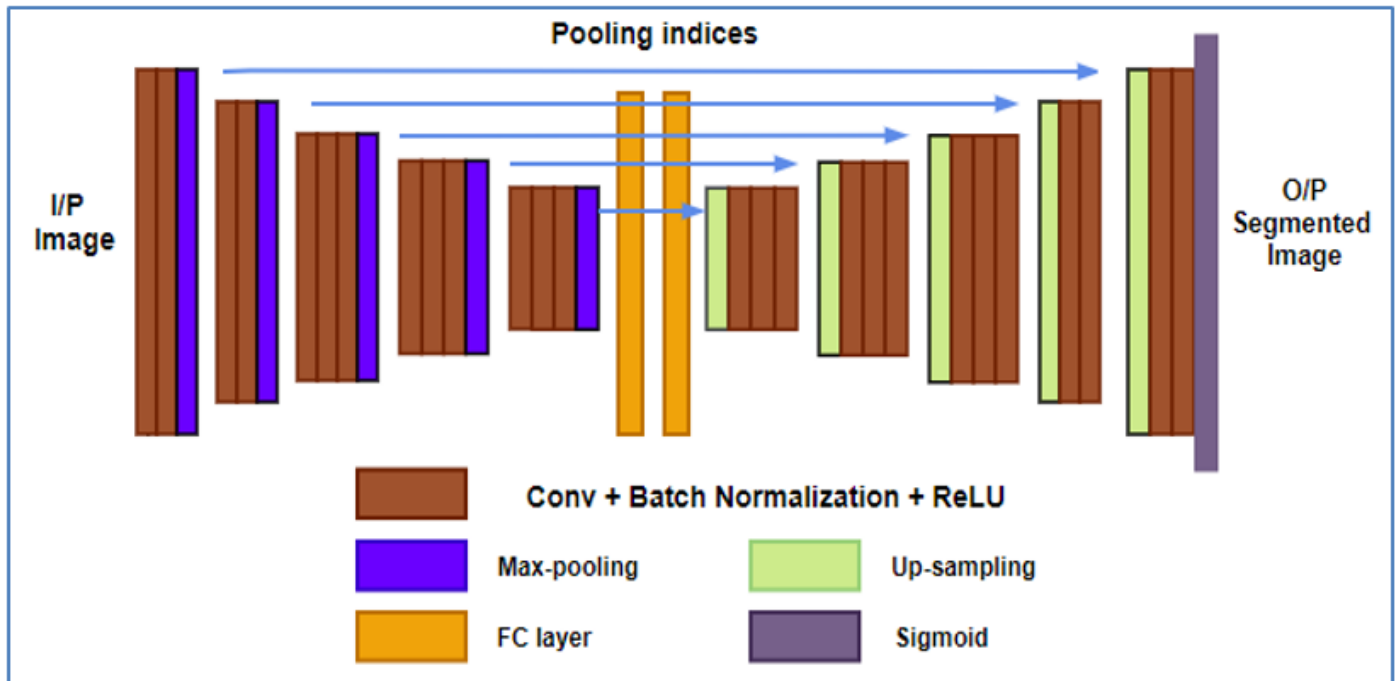


Figure 5

Seg-Net Model.

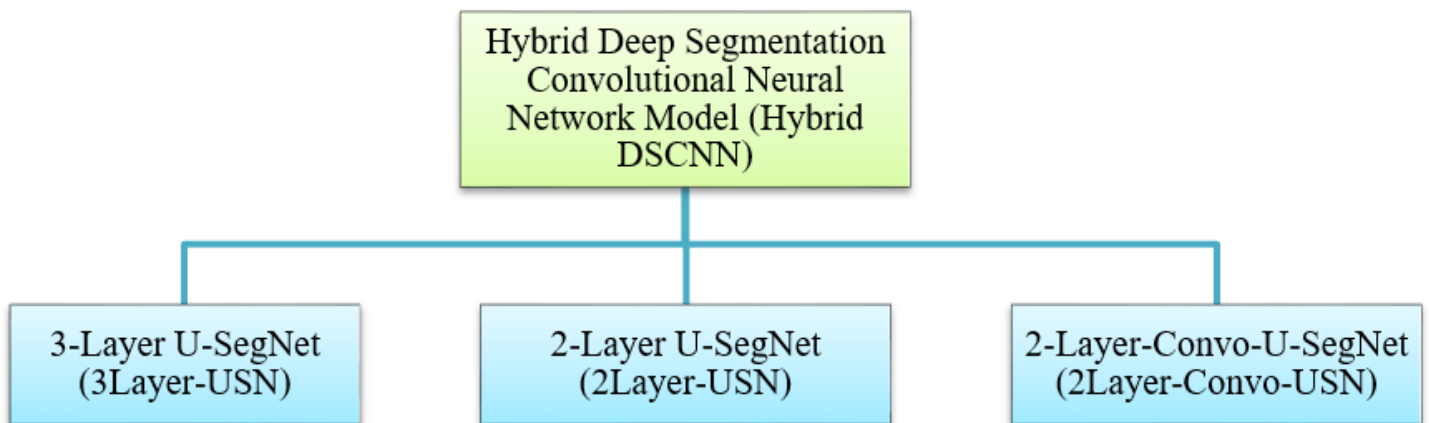


Figure 6

Different types of Hybrid Deep Segmentation Convolutional Neural Network models.

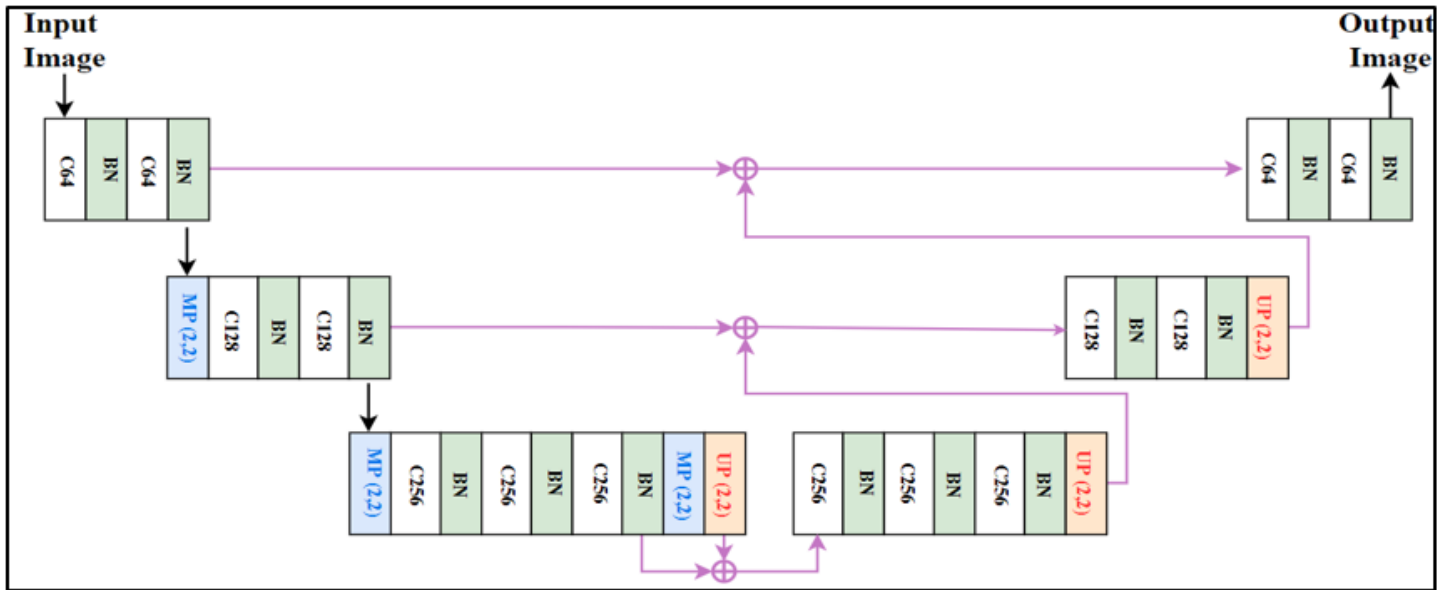


Figure 7

Architecture of 3Layer-USN using Skip Connection.

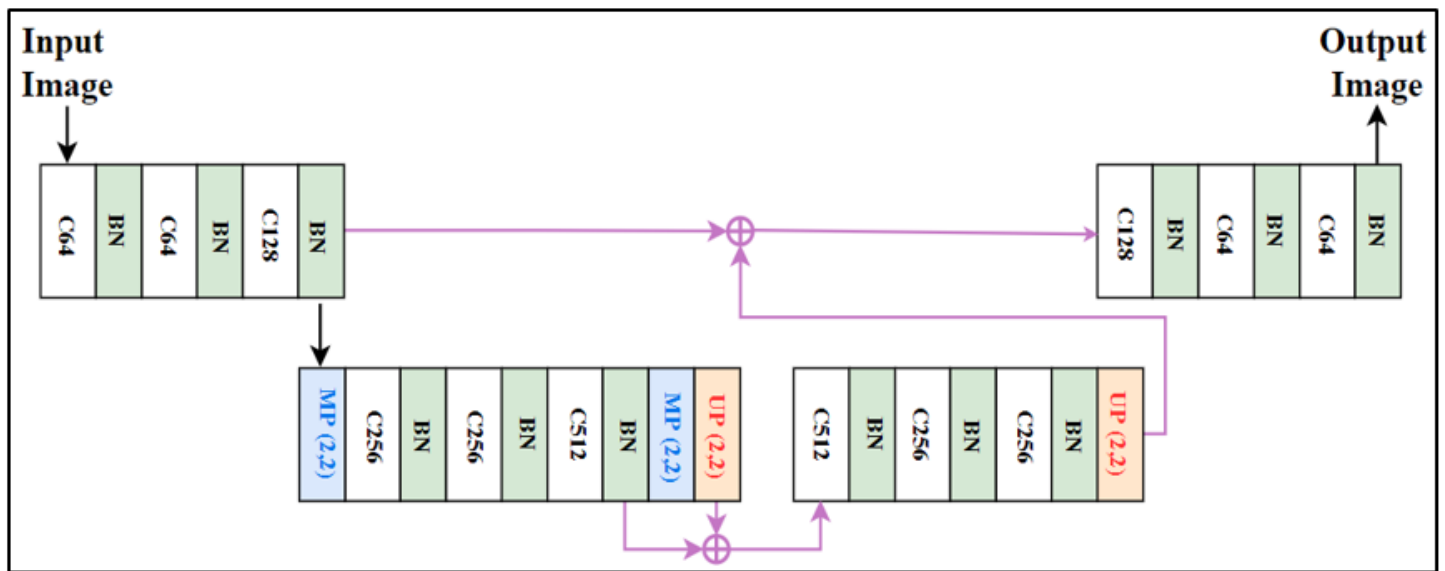


Figure 8

2Layer-USN model with skip connection.

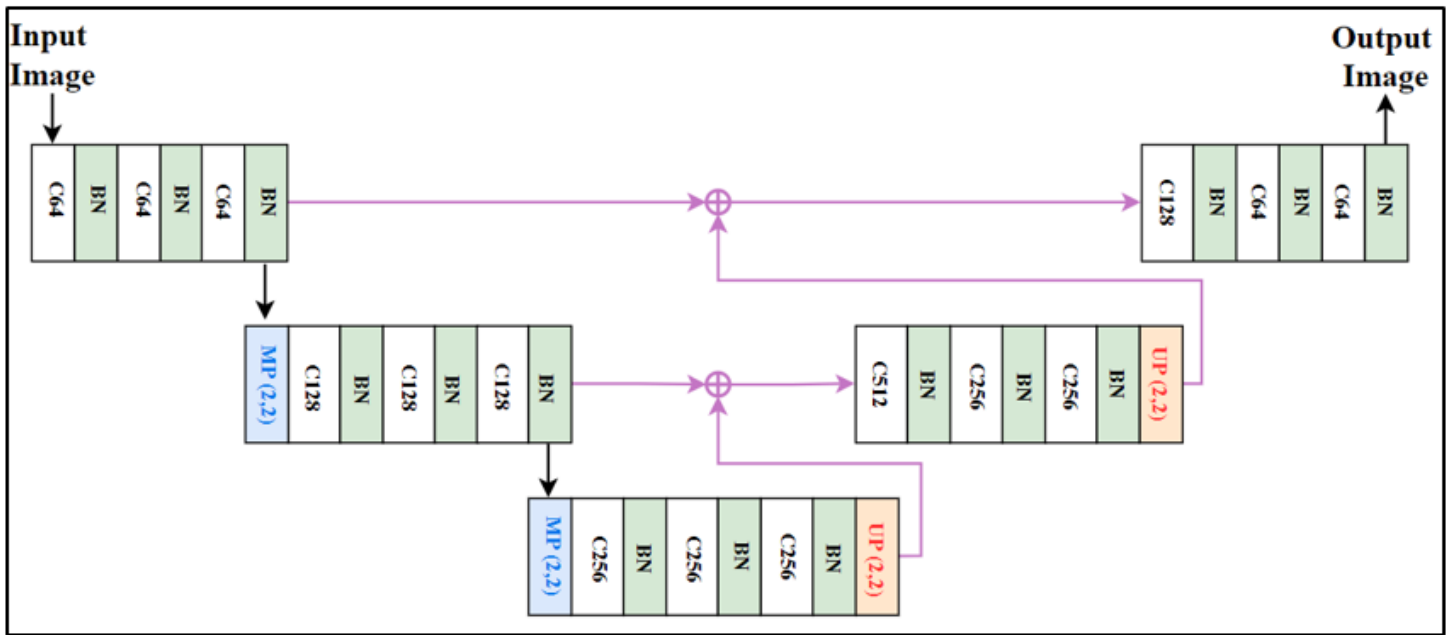
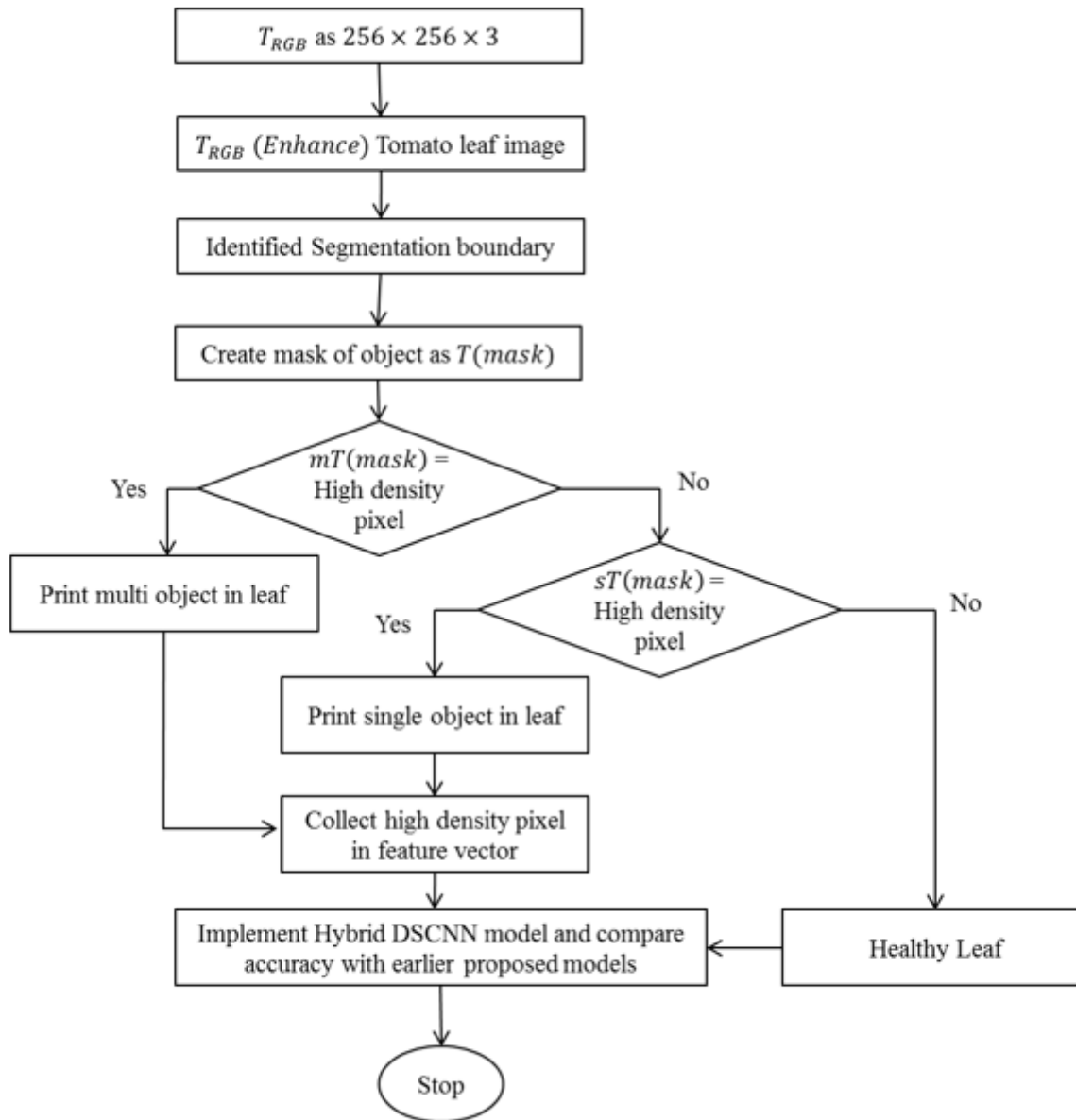


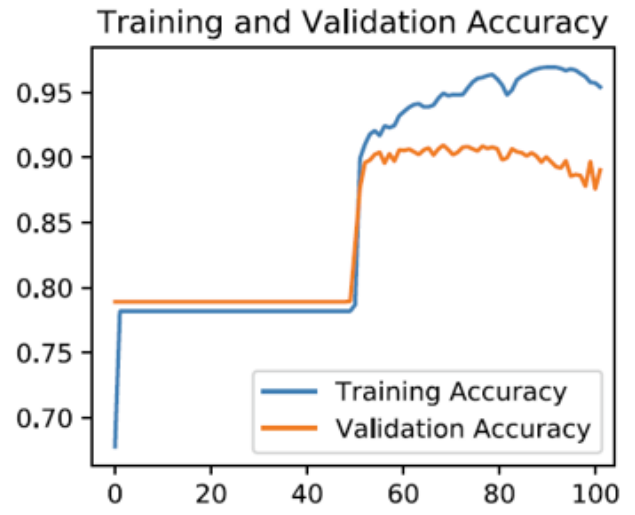
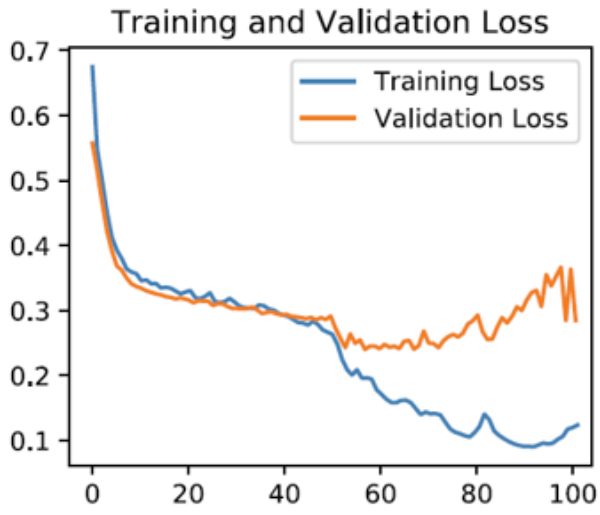
Figure 9

2Layer-Convo-USN model architecture using skip connection.



**Figure 10**

Data flow of the proposed model



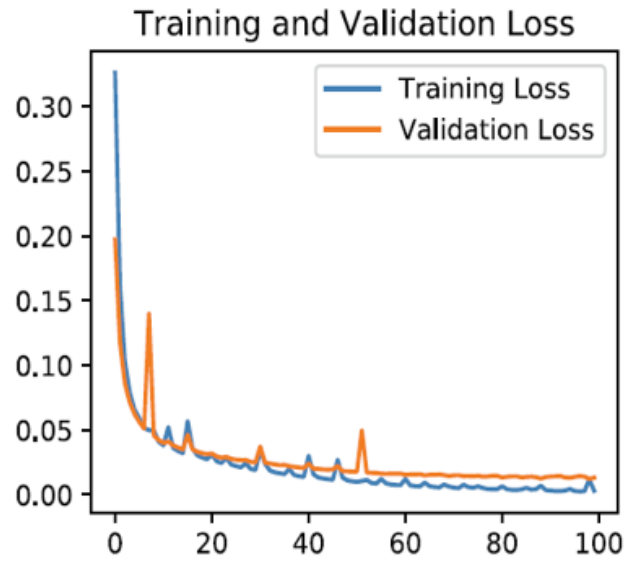
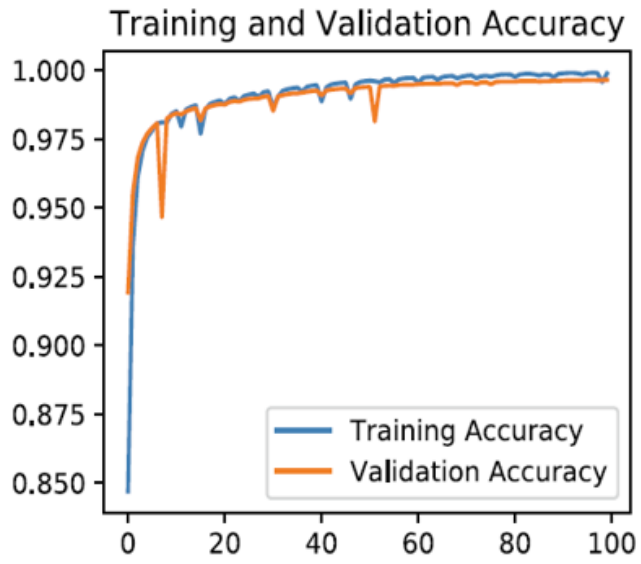
**Figure 11**

Loss and Accuracy graph for single biotic disease detection.



**Figure 12**

Detection result using Mask R-CNN.



**Figure 13**

Loss and Accuracy graph for multi-biotic diseases detection.

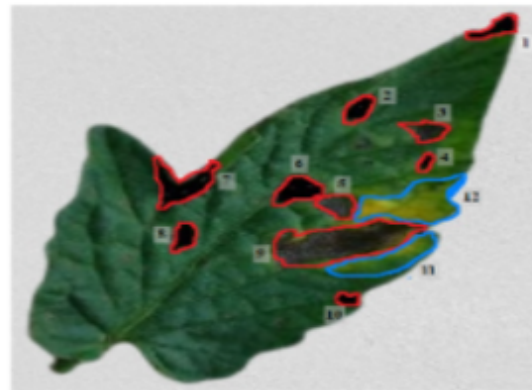
## Input Image



## Output Image



■ Early Blight: 1.5% ■ Anthracnose: 0.7%



■ Late Blight: 2.9% ■ Powdery mildew: 0.7%

Figure 14

Severity estimation using Mask R-CNN for multi diseases.

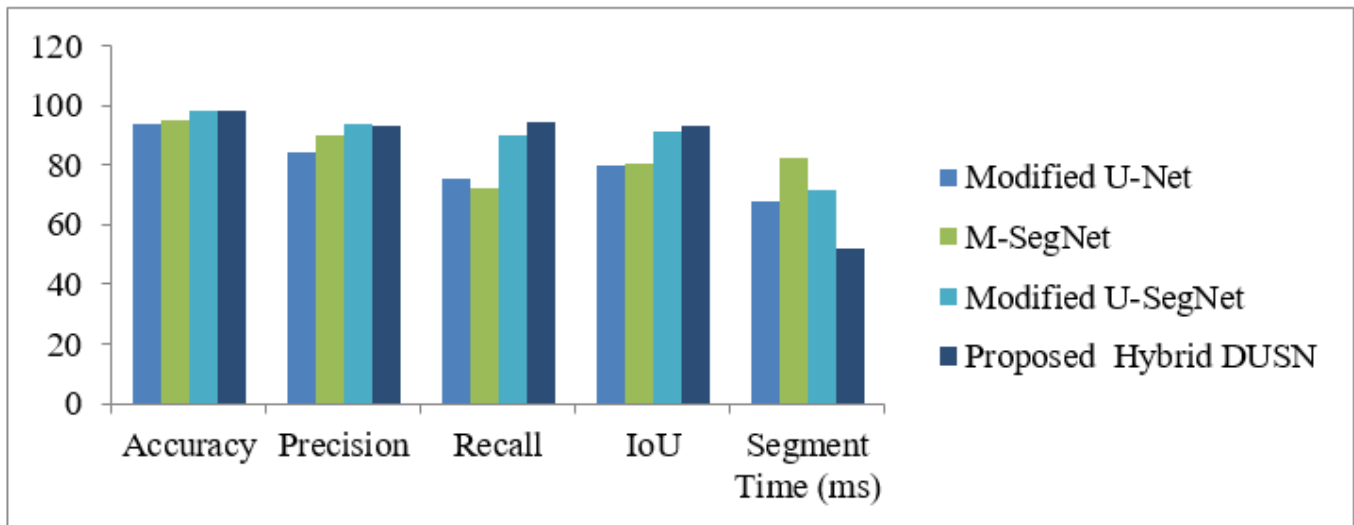


Figure 15

Models evaluation comparison.

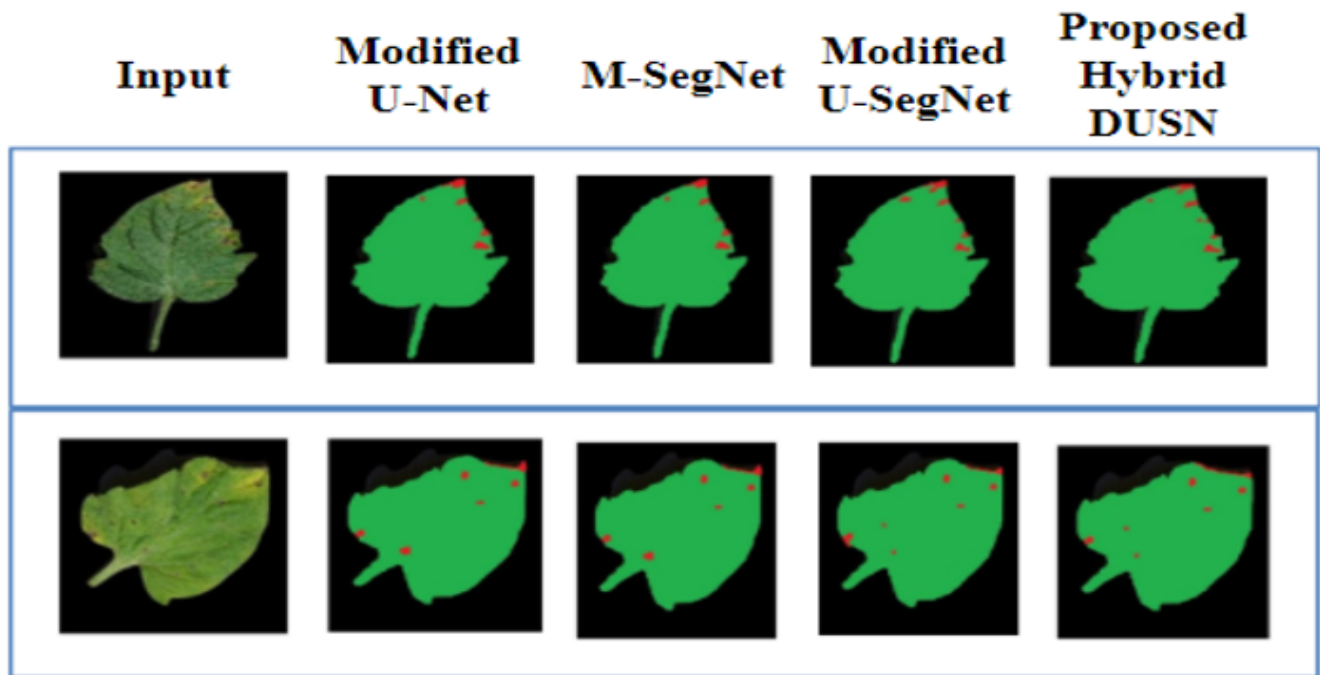


Figure 16

Semantic segmentation output of single disease shown by different models.



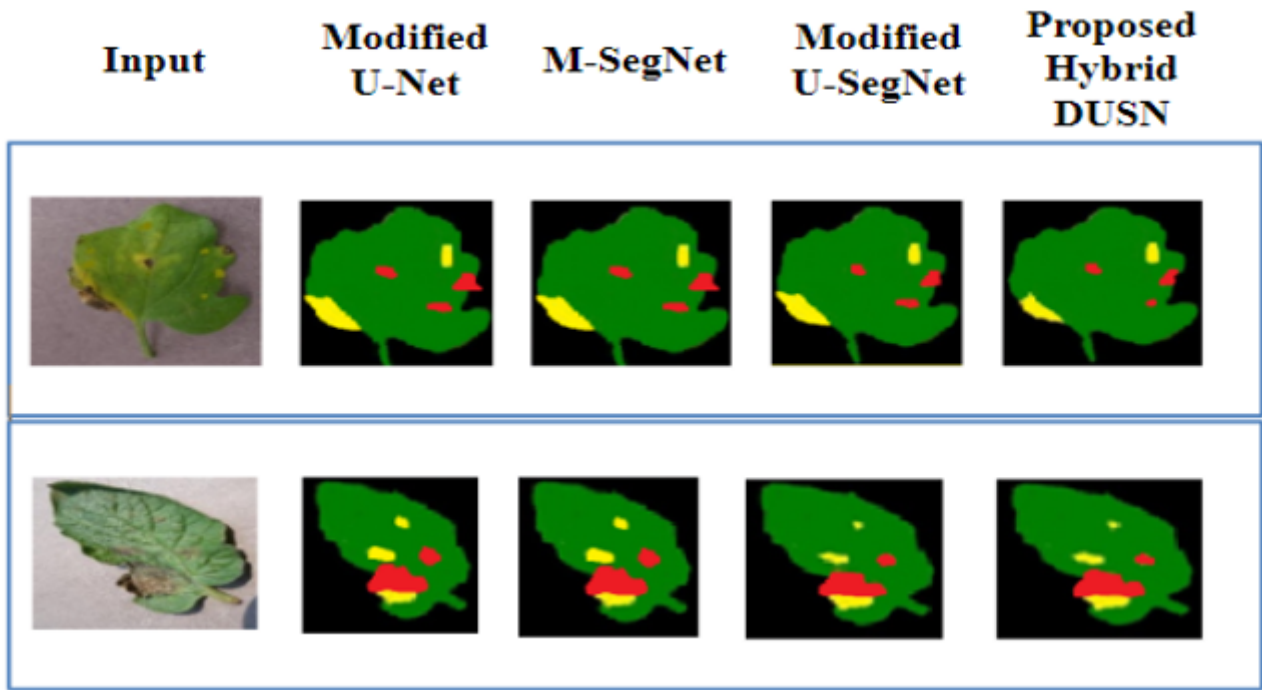


Figure 17

Segmentation output of multi diseases shown by different models.