International Journal of Computers and Informatics

Journal Homepage: https://www.ijci.zu.edu.eg

Int. j. Comp. Info. Vol. 3 (2024) 1–31

Paper Type: Review Article

Artificial Intelligence Techniques for Early Plant Leaf Disease Detection: A Comprehensive Review

Mahmoud H. Alnamoly [1,*], Anar A. Hady [2], Sherine M. Abd El-Kader [2] and Ibrahim El-henawy [1]

[1] Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Egypt.
Emails: mahhassan@zu.edu.eg; mahmoudalnamoly2014@gmail.com; henawy2000@yahoo.com.
[2] Computers and Systems Department, Electronics Research Institute, Cairo 12611, Egypt.
Emails: anar_abdelhady@eri.sci.eg; sherine@eri.sci.eg.

Received: 01 Jan 2024 **Revised**: 15 Mar 2024 **Accepted**: 07 Apr 2024 **Published**: 10 Apr 2024

Abstract

The agricultural area is still a major provider to many countries' economy, but diseases that continuously infect plants represent continuous threats to agriculture and cause massive losses to the country's economy and it is a main threat to the food security. There are a large range of diseases that attack the plants which have its symptoms and pathogen categories. These diseases control the yield and quality of plants and early detection can reduce disease severity and protect crops. Traditionally, farmers use their own experience or hire experts for identification of diseases in their crops but this method requires a thorough knowledge of plant pathogens and take a lot of time and is also prone to being mistaken with a high error rate. Now a day's artificial intelligent techniques are introduced widely to identify plants diseases in a short time and with low error rate and high accuracy. This paper reviews the problem of plant diseases detection and recognition, it overviews the main diseases which can infect the plants, divides them into categories and discusses the symptoms of each disease. After that, it overviews more than fifty current states of the arts related to plants diseases detection using machine learning and image processing algorithms, deep learning algorithms and Internet of Things, comparing between a lot of them in aspects of accuracy and time complexity. Finally, it overviews the main challenges facing researchers in this aspect, and based on these challenges, we propose a plant diseases detection framework using deep learning algorithms.

Keywords: Artificial Intelligence; Deep Learning; Convolution Neural Network; Plant Leaf Diseases; Image Processing; Transfer Learning.

1 | Introduction

Agriculture has become one of the most important areas of research due to its great impact on production and economy, it is one of the main elements for the continuous development of the majority of countries in the world. To meet this urgent need, it is crucial to enhance agricultural yields and safeguard crops [1] by continuously monitoring plants as it is very important to detect plant diseases at their earlier stages to avoid huge losses in our production [2]. Before using artificial intelligence in the agricultural sector, the traditional ways to monitor plants and detect any disease were completely dependent on visual methods with the help of experts as the farmers used their own experience or hired experts to detect the symptoms in the plant and

Corresponding Author: mahhassan@zu.edu.eg

then based on it the disease could be identified, and the needed protective action can be taken. However, these earlier traditional techniques need a lot of money and time and lack accuracy as they require experts to manually monitor the crop fields. Moreover, experts are not always available in poor and isolated areas [3, 4]. Nowadays due to the continuous development of artificial intelligence techniques through the past few years, the process of automatic plant disease detection has become more flexible and they are commonly used in smart farming. These AI systems do not require any human intervention and are fast, inexpensive, and more accurate than the traditional. Smart farming techniques are applied to a lot of agricultural applications [5] such as weed detection, plant disease early detection, fruit grading, land cover classification, etc. In this respect, these automatic plant disease detection models are implemented using the algorithms of machine learning for disease classification and detection, but it achieve a low performance or accuracy due to the complexity of preprocessing of images and feature extraction steps and it isn't suitable for real-time detection scenarios. In this regard, to enhance the accuracy of plant disease detection, deep learning algorithms [6] are widely used in automated agricultural technology including fruit and crop classification, crop detection, and image segmentation and it made a major advance in the computer vision field. One of the highly used algorithms in this respect is the Convolution Neural Network (CNN) algorithm. CNN can avoid complex preprocessing of traditional Machine learning algorithms by automatically extracting features directly from the input images [7].

The overall contribution of this paper is summarized as follows:

- This article has discussed and summarized fifty states of the arts for the early detection of plant diseases based on machine learning, deep learning algorithms, and the Internet of Things. A comparative study was conducted between some of these studies in terms of algorithm performance and model complexity.
- We explored and summarized the challenges that researchers faced in the current studies.
- Based on these challenges, we proposed a deep learning model based on convolutional neural network algorithms and transfer learning algorithms to detect plant diseases and determine the percentage of disease infection based on applied image segmentation techniques.

This paper is organized as follows. Section 2 summaries the different diseases for different plants, their symptoms, their pathogen category, and the reasons behind these diseases. In Section 3, we discuss the role of artificial intelligence techniques (i.e., machine learning, deep learning, and IoT) in this regard by overviewing more than fifty related states of the art. In Section 4, we review a comparative study between some of the most important state of art that have been discussed in the previous section and others. In Section 5, we discuss the main challenges that faced the authors in their methodologies. In Section 6, we present a proposed framework for automated plant disease detection based on the current state of the arts and the challenges that face them. Finally, Section 7 concludes this paper where conclusions and future directions are displayed.

2 | Crop Diseases Classification and Symptoms

Yields are highly prone to different diseases due to a huge number of pathogens in the environment around them. In this section, we will discuss these diseases' pathogens and their symptoms. Plants and their products are affected by two main pathogens which are classified into two categories, biotic and abiotic pathogens [8, 9]. Abiotic pathogens are those that have been produced from surrounding conditions such as pesticides, spring frost, burning chemicals, weather conditions, hail, etc. and they are not as contagious and dangerous as living organisms and cannot be transmitted. In contrast, biotic pathogens are those that have arisen from living organisms such as fungi, bacteria, and viruses. Plant diseases are categorized under three classes' fungal diseases, bacterial diseases, and viral diseases with a few common forms shown in Figure 1. Some plant diseases and their symptoms and pathogen categories are shown in Table 1.

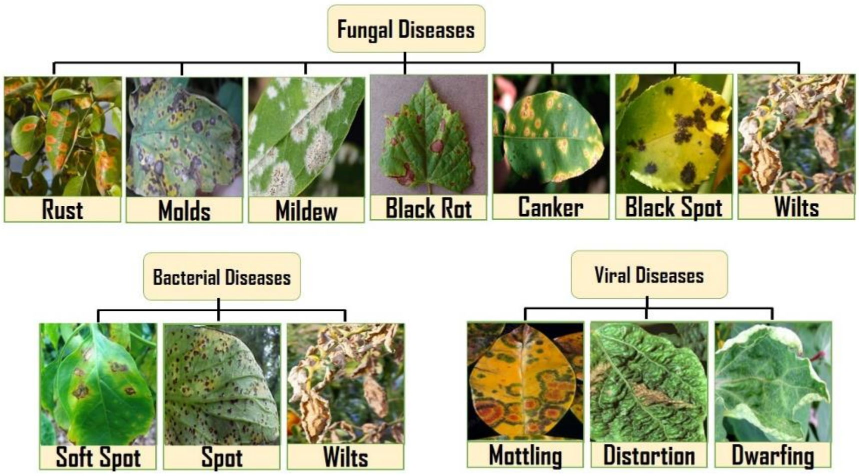

Figure 1. The different classes of plant diseases caused by different biotic pathogens.

Table 1. Different categories of plant diseases caused by biotic pathogens and their Symptoms.

Pathogen Category	Disease	Plant leaf	Symptoms
Fungal	Powdery mildew	Tomato	**Crinkled and yellow leaf**
	Early blight	Tomato	**A dark ring spot surrounded by yellow**
	Late blight	Tomato	**A dark ring spot grows fast**
	Brown spot	Rice	**Gray white center**
	Canker	Citrus	**Brown spots on leaves surrounded by a yellow aura, often with an oily or water-soaked appearance.**
	Stalk rot	Maize	**The leaves of infected plants wither and become dry and appear grayish-green.**
	Downey mildew	Watermelon	**Yellow to white spots**
Bacterial	Bacterial blight	Wheat	**Yellowish green aura, small spots wet with water then it expands and appears as dry dead tissue.**
	Foliar leaf	Cotton	**Light yellow spot with dark brown margins, a reddish lesion will appear during the early stages**
Viral	Areolate mildew	Cotton	**Tanned brown spots andsmall lesions appear on the leaves in the lower dome late in the growing season.**
	Yellow curl	**Tomato**	**Submerged in water surrounded by a yellow aura, a small leaf that turns yellow, also curls upward toward the middle of the leaf.**

The main reason for leaf diseases is the fungus such as downy mildew, anthracnose, and powdery mildew. Old lower leaves with gray-green patches or leaves that have been wet with water are where they first emerge, then these spots get darker and develop fungus as the parasite progresses. A virus disease is the most difficult disease to recognize and diagnose among all plant diseases. In addition, because there is no recognizable indicator that can be continuously observed, these symptoms are frequently misinterpreted for indications of nutritional shortage or damage. Regular virus disease carriers include cucumber-crawling insects, leafhoppers, aphids, and whiteflies. A bacterial disease infects a lot of plants, for example, vegetables can contract serious infections from pathogens. Instead of entering the vegetation directly, it does it through holes or other damages in the crop. Various infections, insects, and agricultural tools cause crop injury when picking and trimming are being done. The top three forms that are most commonly considered for identification and classification are mildew, rust, and spots (i.e., fungi or bacteria). Plant disease symptoms are easily visible in different portions of a plant, leaves and plants are extremely prone to diseases so they are found to be the most commonly observed part for detecting an infection. Because of the similarities in disease symptoms, the texture and the shape of the leaves are used to identify the type of disease in the plant [10, 11].

3 | Plants' Diseases Detection using Artificial Intelligence Techniques

Nowadays the recent improvements in the accessibility of data sources, processing, and algorithms have enabled Artificial Intelligence (AI) to start delivering on its promise of providing genuine value and the process of plant disease detection has become easier and faster than before as it recognizes the disease in minimum time and with low error rate and a high accuracy. Given that farmers spend a lot of time observing and assessing their crops, smart agriculture has become vital. AI and IoT techniques offer accurate monitoring remotely with economical and wise crop management [12]. Smart farming is a new concept that aims to increase the productivity and effectiveness of agriculture by using advanced information technologies. Using the most recent developments in automation, artificial intelligence, and networking; farmers are better able to keep an eye on every step of the process and apply exact treatments selected by machines with superhuman accuracy. Engineers, data scientists, and farmers are still developing methods to optimize the human labor needed in agriculture. Smart farming develops into a learning system that gets smarter every day as vital information resources get better [13]. With the fast development of smart farming, plant disease detection becomes digitalized and data-driven, enabling

intelligent decision support, clever analysis, and strategic planning using one or more algorithms of artificial intelligence such as deep leaning algorithms which is a subset of machine learning algorithms as shown in Figure 2. One of the major causes of productivity losses in agriculture is plant diseases. It is essential to keep an eye on the crops' condition and to stop the spread of sickness. Different plants have different ways of preventing plant diseases and diagnosing diseases [14]. Identification of diseases is carried out on plant leaf images taken either using satellites, drones, or using a camera.

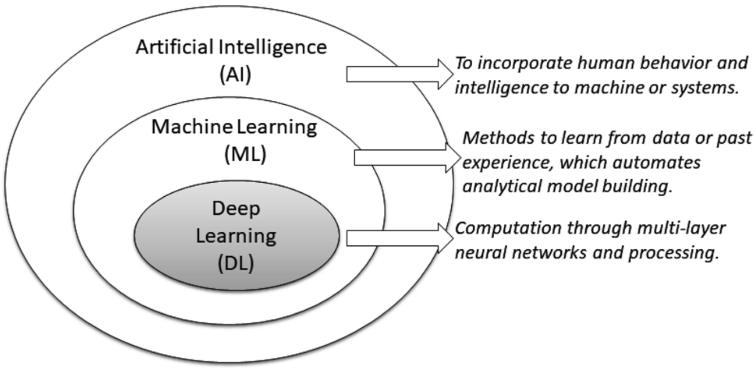

Figure 2. An illustration of the position of deep learning (DL) and machine learning (ML) comparing with artificial intelligence (AI).

There are many plant disease datasets that we can use but the most used dataset by the researchers is the Plant-Village Dataset [15]. It consists of 54306 images of plant leaves taken in a controlled setting which are divided into 38 different classes for 14 different plant species (i.e., Tomato, Cherry, Squash, Apple, Orange, Potato, Strawberry, Raspberry, Grape, Blueberry, Bell Pepper, Soybean, Peach, Corn). These 38 classes are divided as follows, 12 different healthy species, 17 fungal diseases, 4 bacterial diseases, 2 viral infections, 2 mould diseases, and 1 mite disease. The dataset has a large variety as a result of the use of conventional digital cameras outside, in various weather situations, and from a variety of sources. In the following sections, we will discuss the various methodologies that have been used to detect the diseases of different plants based on artificial intelligence technologies in detail.

3.1 | Plants Diseases Detection Based on Machine Learning Algorithms

A Machine Learning (ML) algorithm [16] such as Support Vector Machines (SVM), Random Forests, k-Nearest Neighbors (k-NN), Decision Trees, Fully Connected Neural Networks (FCNN), etc. is a computer procedure that makes use of input data to complete an intended task without being explicitly programmed to do so. The "learning" component of machine learning is this training. The training does not have to be restricted to a first adaption over a set period. Here, our training data is the leaves images of the plant. In machine learning algorithms, features are extracted from images manually after applying some image pre-processing steps as crop images often have noise, therefore processing the image requires removing these contaminants [17]. Figure 3 illustrates how machine learning algorithms apply image classification in four stages. The first stage is pre-processing, which primarily includes fine-tuning the image, eliminating unwanted noise, and enhancing some features using contrast enhancement for further processing, like a Gaussian function that produces some soft blur in the image [18]. After that, the ML algorithm separates the image from its background and segments the Area of Interest (AoI) to highlight the key features that make it simpler to identify objects and boundaries in the acquired image [19], this procedure is called image segmentation. The next phase, feature extraction [20], is when an image's information and features are revealed. As we previously stated, the shape, texture, and color of the leaf are the typical features that will be utilized to detect crop disease. These extracted features are considered the strongest features which will be converted to a vector of features fed to the classifier algorithm to predict the correct class. The last stage is image classification [21], which uses the feature map that was extracted in the previous stage to classify the images. The best classifier depends on the nature of the problem being solved. Prediction or classification is divided into two steps: first, the dataset is divided into a training dataset and a test or validation dataset; second, our model is trained on the images in the training dataset, and finally, its performance or accuracy is evaluated using the images in the test dataset.

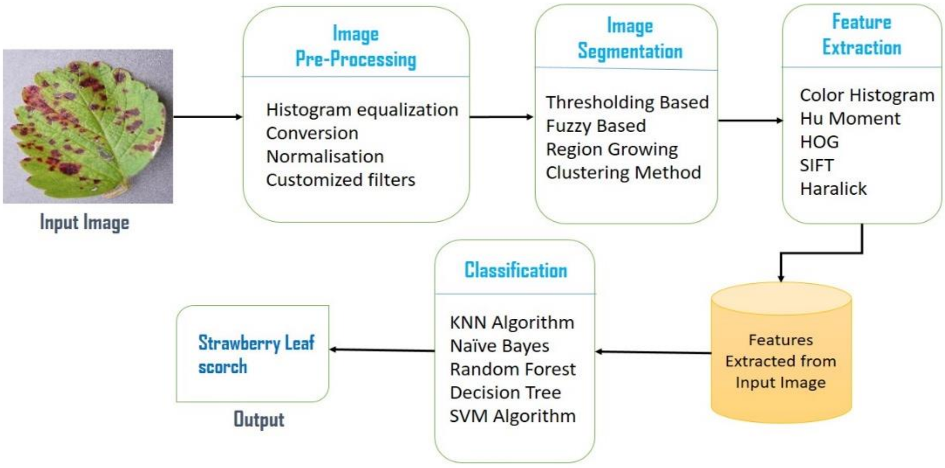

Figure 3. The main image classification workflow of machine learning (ML) Algorithms.

In the next section, we will review the state of the art in this regard.

Singh et al., 2022 [22] proposed a novel algorithm for the segmentation of diseased parts in apple leaf images. This segmentation algorithm firstly enhanced the plant's leaves using the Brightness-Preserving Dynamic Fuzzy Histogram Equalization technique, after that it extracted the diseased part on the apple leaf which was classified after that using K-nearest neighbor classifier and achieved about 96.4% accuracy.

Kianat et al., 2021 [23] suggested a new hybrid method based on feature fusion & selection approaches to classify the diseases of cucumber plants. They first, applied data augmentation techniques to increase the number of trainable images, after that they applied feature extraction, fusion, and selection approaches. They used a proposed method called the Manhattan Distance Controlled Entropy (MDcE) technique to extract the most significant features, then they used a proposed method called the Probability Distribution-based

Entropy (PDbE) technique to reduce the number of extracted features. The final phase was classifying the most significant features using a variety of classifiers. The accuracy of this hybrid framework was 93.5% on average.

Almadhor et al., 2021 [24] proposed a new framework for guava plant disease classification based on a high-resolution image dataset of guava leaves with about (18 MP) resolution. They applied this framework as follows: In the phase of segmentation, color difference image segmentation methodology is used for separating the infected regions on the image, after that in the phase of feature extraction, RGB & HSV color and Local Binary Patterns (LBP) histograms have been used. Finally, in the phase of classification, they utilized advanced machine-learning classifiers for disease recognition i.e., KNN, Fine Cubic SVM, Complex Tree, Boosted Tree, and Bagged Tree. Bagged Tree had the best recognition performance, which was 99%.

Pothen et al., 2020 [25] proposed a new method for rice leaf disease recognition. They first, segmented images using Otsu's segmentation method, and after that, they extracted the main features from the images using Histogram of Oriented Gradients (HOG) & Local Binary Patterns (LBP) histograms. Finally, they applied a support vector machine algorithm to classify the extracted features. The best performance they achieved through using a support vector machine with a Histogram of Oriented Gradients which is a histogram with a polynomial kernel function. Their method's accuracy was about 94.6%.

Mugithe et al., 2020 [26] proposed a new alerting system using a buzzer to alert the farmers in case there is any disease on the plant to take care of it at an early stage to prevent or minimize the disease from spreading. In the case of healthy plants, the buzzer doesn't turn on and remains silent. They applied a k-means clustering algorithm for image segmentation, and then from these segmented images they extracted the features that are used to detect the type of disease. This alerting system achieved an accuracy of 95.1613%.

Al-bayat et al 2020 [27] proposed a plant disease detection system based on deep neural network architecture to detect diseases on the apple leaves. The first step of their model is a preprocessing step in which they applied some image enhancement and region of interest segmentation. After that, they applied the Speeded Up Robust Feature (SURF) for feature extraction and the Grasshopper Optimization Algorithm (GOA) for feature selection or optimization. Thanks to these methods, their model's accuracy became 98.28%.

Kumar et al 2020 [28] proposed a framework for the identification of various plant leaf diseases depending on machine learning algorithms and image processing methods that use the K-mean segmentation algorithm and a multi-class support vector machine algorithm. It was carried out in four steps. RGB color is converted to HSI color in step 1, in step 2 they applied a k-means clustering algorithm for image segmentation. In step 3 they extracted features from the images such as color, texture, and shape features and finally, multiclass SVM is applied to these features to detect the type of these diseases. The accuracy of these suggested approaches was 95.7%.

Khan et .al 2020 [29] suggested a plant disease identification approach based on machine learning and image processing algorithms. They first, applied a k-means clustering algorithm to isolate the region of interest (diseased portion) from the whole plant leaf image, this step is called image segmentation. Nine features were extracted from the segmented RGB image, these features are variance, skewness, RMS root mean square, kurtosis, standard deviation, inverse difference, mean, smoothness, and entropy. Four features were extracted using grayscale images such as energy, homogeneity, correlation, and contrast. Finally, these feature vectors were fed to the multiclass Support Vector Machine (SVM) classifier algorithm for classification purposes. The accuracy of the proposed approach was 92.8571% using a dataset that consisted of 148 images with 5 types of leaf diseases.

Hossain et al., 2019 [30] suggested a way to identify the diseases of five different types of plants. Their proposed methodology pipeline consisted of, converting RGB image to LAB image, after that they applied the k-nearest neighbor classifier algorithm for image segmentation. After that applied two processes, one for increasing the width of maximum areas and removing the unwise bad noises from the images called dilation and the other called erosion for reducing the width of the smallest region. After that, texture features (energy, homogeneity, correlation, and contrast) were extracted using gray-level co-occurrence matrices, and color features were extracted from the segmented RGB disease image. Finally, classification was done using the k-nearest neighbor classifier algorithm with a performance of 96.76%.

Rauf et al., 2019 [31] proposed a method for classifying citrus diseases. They first, applied some filters such as the top-hat filter and the Gaussian function to improve the contrast of the infected region. Secondly, to segment the enhanced images, weighted segmentation, and a saliency map were applied. The segmented images were then passed to a feature extractor algorithm, which extracted texture, geometric, and color features. Next, the features selector algorithm used PCA, skewness, and entropy approaches to select from these features. The classification algorithm is then used to assign each instance of an image to the appropriate disease class. Ali, et al 2017 [32] suggested an approach to identify and classify the major citrus diseases. Their proposed method consists of the following steps: Step 1, using histogram equalization to improve image contrast and applying color space transformation technique (converts RGB color space to LAB color space as in RGB color space some information cannot be visible). Step 2, used a color difference-based algorithm called Delta E (ΔE) for image segmentation purposes. Step 3, applied local binary patterns for textural features extraction and RGB, HSV histogram for color features extraction purposes. In the last step, they applied each of the Boosted tree, Fine KNN, Bagged tree ensemble, and Cubic SVM classification algorithms on the extracted features vector and they used the principal components analysis algorithm to reduce the extracted features dimension finally, they found that Bagged tree ensemble classifier is the best between the others with 99.9% accuracy. The dataset consisted of 199 citrus images.

The success of machine learning-based solutions and related applications largely depends on both data and learning algorithms. Machine learning models may lose their usefulness or perform less accurately if the training data are unsuitable for learning, such as unrepresentative, low-quality, irrelevant, or insufficient features in quantitative terms. Therefore, handling the various learning algorithms and correctly processing the data is crucial for a machine learning-based solution and it may achieve a low performance due to the complex preprocessing of the image and feature extraction steps and it is not suitable for real-life detection scenarios [33].

3.2 | Plants Diseases Detection Based on Deep Learning Algorithms

The deep Learning (DL) algorithm is one of the algorithms of machine learning that are based on artificial neural networks. To train the model on specific data, deep learning uses a computational architecture consisting of some layers such as the input layer, hidden layers, and output layers. Deep learning algorithms outperform traditional machine learning algorithms in the model accuracy especially when using a huge dataset [34, 35]. Figure 4 illustrates how deep learning generally performs better than machine learning when taking the growing amount of data into account. However, it might change based on the study design and data characteristics. Computer vision technologies based on Deep learning algorithms are widely used in automated agricultural technology including fruit and crop classification, water management, crop diseases detection, seed quality, yield estimation, soil analysis, and others and it made a major advance in the field of smart agriculture [36].

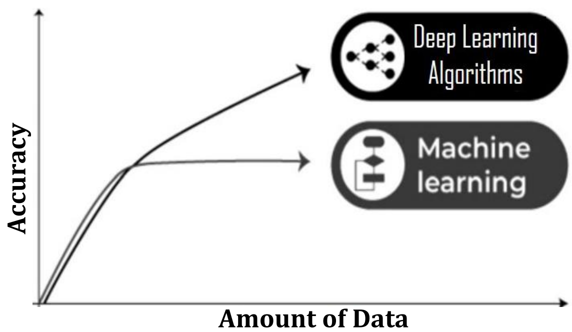

Figure 4. The accuracy of machine learning vs deep learning Algorithms based on the amount of data

The main difference between ML and DL algorithms is that deep learning algorithms apply an automatic approach to extract features from the images but in the case of machine learning algorithms, this operation is done manually. Convolutional Neural Networks, Long Short-Term Memory, Multi-Layer Perceptron, Recurrent Neural Networks, and Long Short-Term Memory Recurrent Neural Networks are the most popular used deep learning algorithms [37]. One of the most usable algorithms in the case of image classification is the Convolutional Neural Network algorithm (CNN). The convolutional Neural Network (CNN) algorithm is a feed-forward Neural Network that is inspired from artificial neural networks, and trained in a supervised manner [38, 39]. It consists of convolutional layers and pooling layers for feature extraction and reduction and fully connected layers or dense layers for classification. CNN algorithms make use of the two-Dimensional (2D) structure of the input data so that they are used in various fields such as image processing and classification, natural language processing, image and video recognition, medical image analysis, etc. CNN is thought to be more powerful than traditional ANNs despite having a higher computational cost because it automatically detects the key traits without any human interaction. In the area, complex deep learning models built on CNN, such as AlexNet [40], Xception [41], Inception [42], Visual Geometry Group (VGG) [43], ResNet [44], etc. can be applied.

Figure 5 illustrates the two halves of convolutional neural networks: a convolutional half or feature extraction half and a classifier half. The convolutional half contains some of the layers that are employed to extract and discover the main features of the image, such as convolutional layers with filters or kernels, activation layers after each conv layer, and pooling layers. In convolutional layers, we applied some filters sometimes called kernel with a specific size (i.e., 3x3) to find the presence of specific features such as edges, and textures which will be helpful in image classification.

The feature map produced using convolutional layers may contain negative values and an activation function called rectified linear unit (ReLU) is applied to it to get rid of these negative values, the formula of ReLU is $F(x)= \max (0, x)$. The next layer, the pooling layer, aims to reduce the perimeter of the feature maps in which it selects the strongest features and it in return decreases the number of parameters also which leads to a decrease in the number of computations in the network. Images are classified in the classifier half which is made up of dense layers and dropout layers using the features generated by the convolutional half.

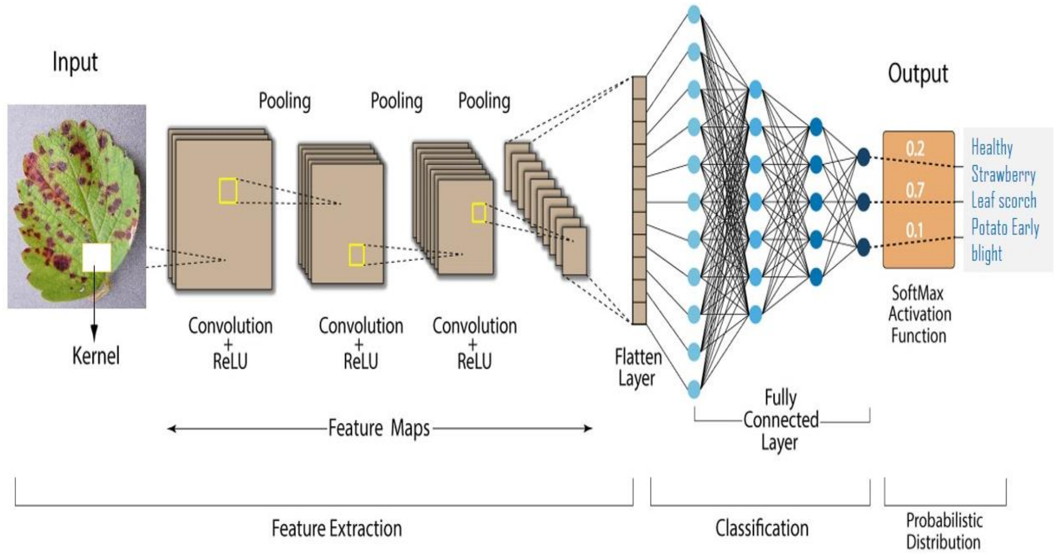

Figure 5. The architecture of convolutional neural network for multi-class image classification

The fully connected layer responsible for the classification of the produced features is also called the dense layer. Each of the dense layers consists of several neurons fully connected, their number in the input dense layer depends on the features map but their numbers in hidden dense layers vary depending on you also their numbers on the final dense layer depend on the number of classes. The activation function in a fully connected layer differs from the other used in the convolutional layer, here we use the sigmoid or softmax activation function. In the case of multiclass classification purposes, Softmax is the suitable choice as it gives decimal probabilities for each class but in the case of binary classification, we use the sigmoid activation function as shown in Figure 6. CNN uses a layer called dropout layer after dense layer which helps prevent the problem of overfitting.

Any CNN model's accuracy or performance depends on the used hyper-parameters such as learning rate momentum, number of epochs, number of maximum polling layers, initialization of the network weights, dropout rate, batch size, activation function, number of convolution layers, etc. so that, the hyperparameter tuning process is considered as one of the most important steps in the process of improving model accuracy. The genetic algorithm, grey wolf, PSO, and other meta-heuristic optimization techniques can all be used to accomplish this manually or automatically [45-47].

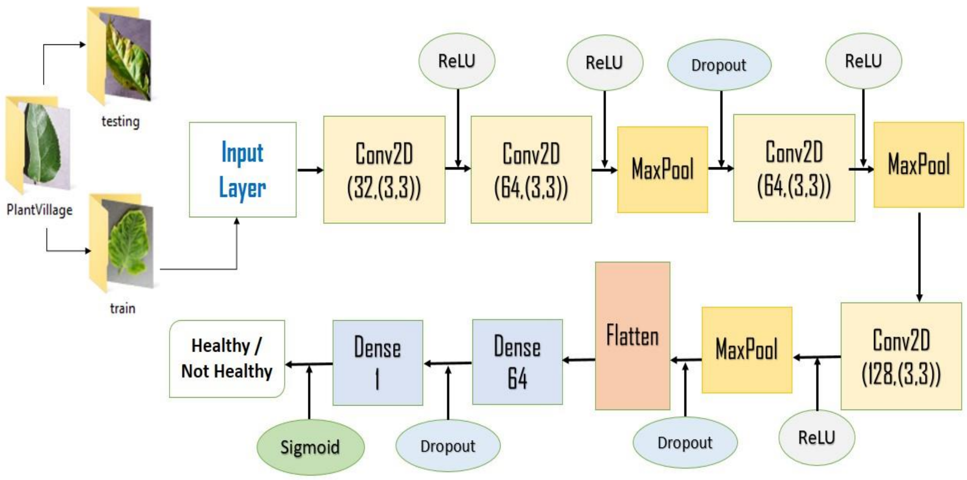

Figure 6. The architecture of convolutional neural network for binary image classification.

In the next section, we will review the state of the art in this regard.

Ahmed et.al (2022) [48] proposed a transfer learning-based strategy to identify diseases in tomato leaves using the PlantVillage dataset. This method's initial step involves enhancing the leaf images with illumination correction using a powerful preprocessing technique to improve classification. The illumination issue that persisted in the dataset has been resolved thanks to the use of the adaptive contrast enhancement technique. Following that, a hybrid model made up of a classifier network and a pre-trained MobileNetV2 architecture was used for feature extraction. To prevent data leakage and address the problem of class imbalance, they applied a runtime augmentation instead of applying traditional data augmentation. Their model achieved an accuracy of 99.30% and over 2.4 million training parameters with a model size of about 9.60 MB.

Schuler et.al (2022) [49] Proposed a feeding deep convolutional neural networks CIE Lab color space (L* a* b*) as opposed to RGB color space. They updated the Inception V3 design to have two branches, one for chromatic data (AB channels) and the other for achromatic data (L channel). The updated Inception V3 model's first three convolutional layers were focused on learning chromatic and achromatic characteristics from the CIE Color space. They used CIE Lab color space since RGB channels are highly correlated, working in an uncorrelated color space (CIE) that avoids excessive filter weights, and the partitioning that has occurred reduces the number of trainable parameters. Their model performed better than traditional one-branch RGB color space images. They tested their model on 2 datasets, the Cropped PlantDoc dataset with an accuracy of 76.91% and the PlantVillage dataset with an accuracy of 99.48% with over 5 million trainable parameters.

Amreen et.al (2021)[50] suggested a deep-learning approach for tomato plants diseases detection. They created a synthetic image using Conditional Generative Adversarial Network (C-GAN) as a data augmentation action to increase the size of their dataset and to prevent their model from the problem of overfitting. They applied the transfer learning technique by using one of the well-known deep learning models called the

DenseNet121 model to classify their dataset which consisted of ten different types of real and synthetic tomato disease images. The accuracy of this model, which can be divided into five classes, seven classes, and ten classes, was tested on the PlantVillage dataset and was 99.51%, 98.65%, and 97.11%, respectively. The over-fitting issue is avoided by the suggested data augmentation technique (C-GAN), which also increases the generalizability of the network. There are 1,735,904 trainable parameters in this model.

Trivedi et al., 2021 [51] proposed a CNN classification model for tomato diseases. The images were subjected to some preprocessing, as well as image segmentation. Secondly, varying hyper-parameters of the CNN model were applied to process the images in a further way. Finally, the proposed model extracted additional characteristics from the images, such as texture, color, edges, etc., which will then be classified. The suggested model was performed with a 98.49% accuracy utilizing 1,422,542 training parameters after being tested and trained on a dataset with 3000 photos of tomato leaves.

Alhaj et al., 2021 [52] employed transfer learning to identify tomato leaf disease using the InceptionV3 model. It said that Transfer learning cuts down on execution time and that it obtained a 99.8% accuracy rate, but it made no mention of the time complexity or the amount of training parameters that were used. The finished model was made available as a cloud-based web application.

Punam et al., 2021 [53] proposed a new hybrid model for identifying bacterial spot disease in peach plants. They used Convolutional Auto-encoder (CAE) networks with Convolutional Neural Network (CNN). The purpose of the CAE network was to reduce the number of features on the final features map which is used for prediction or classification. They claimed that by using the CAE network, the number of training parameters was decreased and in return prediction time was decreased without a weighty decrease in the classification performance. Their CNN model was trained and tested on the publically available dataset PlantVillage with a training accuracy of about 99.35% and testing accuracy of about 98.38%. This proposed model used only 9,914 training parameters.

Sun et al., 2021 [54] developed a mobile application called Mobile End AppleNet-based SSD algorithm (MEAN-SSD) based on GoogLeNet's Inception module for the detection of the diseases of five common apple leaf diseases on the apple plants in a real-time scenario. It consisted of a part for extracting the disease features from the plant called MEAN block and a second part for locating the disease spots called Single Shot Multi-Box Detector (SSD). Their model updated the inception model by changing each of the 3*3 conv filters with their MEAN block. Their model depended on the SSD algorithm. They first applied a data augmentation technique on their owned dataset "AppleDisease5" which consisted of 2,230 images, the total number of images on the new dataset after applying augmentation is 26,767 images. This proposed model MEAN-SSD which was implemented using the keras framework enhanced the accuracy of the model in real-time scenarios. With 2,152,744 parameters, this model achieved about 97.07% accuracy.

Akram et al., 2020 [55] suggested a deep convolutional neural network model based on well-known models such as AlexNet and VGG for detection and classifying various types of fruit leaves. This technique is called the transfer learning technique in which they used a pre-trained deep model like AlexNet and VGG-s to extract the deep features. For the aim of selecting features, they also presented a multi-level fusion process based on an entropy-controlled threshold value that was computed by averaging the selected features. A multi-support vector machine approach was used to classify the final features that were chosen. Five distinct diseases were identified in apple images from a plantvillage dataset using this algorithm with an accuracy of 97.8%.

Sanga et al., 2020 [56] developed a smartphone application based on deep learning models for early detection of banana diseases. Their model used a pre-trained model such as VGG-16, ResNet-152, ResNet-18, ResNet-50, and InceptionV3 to train and test their dataset and finally, they compared them to choose the best one for deploying in their application. The best accuracy after being applied to the dataset which consisted of 3000 images of banana leaves was 99.2% achieved by Resnet152 while Inceptionv3 achieved an accuracy of 95.41%. Their model used 60 million training parameters.

Chohan et al., 2020 [57] suggested a convolutional neural network model based on pre-trained models such as VGG-19 and InceptionV3 as a plant disease detector. Before training their model, they first increased the number of images in their dataset PlantVillage by applying a data augmentation technique. After that, they employed each of InceptionV3 and VGG-19 on the new augmented dataset and they found that the VGG-19 model outperformed the InceptionV3 as it achieved a testing accuracy of 98.3% and training accuracy of 95% using about 143 million training parameters.

Mohameth et al., 2020 [58] proposed a hybrid convolutional neural network model since they used 3 well-known pre-trained models i.e. GoogLeNet, VGG-16, and ResNet 50 for the features extraction phase, and after that, in the phase of classification, they employed 2 machine learning classification algorithm i.e. k-Nearest Neighbour and Support Vector Machine (SVM) classifiers. Their model used the open publically PlantVillage dataset and its best accuracy was achieved using ResNet-50 in the feature extraction phase with SVM in the classification phase. Their model achieved an accuracy of 98% and used approximately 25 million training parameters.

Singh et al. 2020 [59] created a new plant diseases dataset called PlantDoc which consisted of 2598 images for 17 different types of diseases for 13 plant types. PlantDoc required 300 human hours of work to be collected. They think their dataset can lower the entry cost for computer vision algorithms used to detect plant diseases and therefore, they employed three well-known models on their dataset for classification and for showing the performance of their PlantDoc dataset. According to their findings, modeling with their dataset can improve classification accuracy by up to 31%. They claimed that it may be possible to increase the dataset's utility by using image segmentation techniques to extract leave spots from the images, so they used a Faster Region-based Convolutional Neural Network (R-CNN). They found that the InceptionResnetV2 model with Faster R-CNN achieved a performance of 70.53%.

Tiwari et al., 2020 [60] proposed a new hybrid convolutional neural network for potato plant disease detection from the PlantVillage dataset. They used and employed three well-known models i.e. InceptionV3 VGG-16, and VGG-19 in the phase of feature extraction and after that, they used various types of machine learning classifiers such as SVM, Logistic Regression, k-Nearest Neighbor classifiers, and Neural Networks for classification. After applying these models with these classifiers, they were obvious that using a Logistic Regression classifier after the pre-trained VGG-19 achieved the best accuracy compared with the others with an accuracy of 97.8%. Their model used about 143 million training parameters.

Khamparia et al., 2020 [61] proposed a hybrid model that consisted of a Convolutional Auto-encoder and a convolutional neural network to detect the disease in three plants tomatoes, potatoes, and maize. They used a convolutional auto-encoder to reduce the number of features on the final features map to reduce the number of trainable parameters. Their model achieved a training accuracy of about 100% but on the other side, the testing accuracy was about 86.78% in their model over-fitted the training data. They trained their model on a dataset consisting of 900 images (tomatoes, potatoes, and maize) divided into 6 classes and finally, they said that their model had a 97.50% accuracy rate and used about 3.3 million training parameters.

Mohit et al., 2020 [62] developed a convolutional neural network model that consisted of three convolutional layers, three max-pooling layers, and two fully connected layers to identify the disease of tomato plants. Before training their model, they used a data augmentation technique to balance the data in each class as the class's images weren't balanced. After that, they trained and tested their model using 17500 images from the PlantVillage dataset which was divided into 10000 images for training, 7000 for validation, and 500 for testing. Their model achieved a testing accuracy of 91.2% with a model size of about 1.5 MB and about 208802 training parameters.

Shradha et al., 2020 [63] employed three well-known convolutional neural networks i.e. SqueezeNet, Inception V3, and AlexNet to evaluate the severity of the late blight disease on tomato plants using feature extraction and transfer learning techniques. To evaluate the severity of the disease, they separated the images into the PlantVillage dataset based on their disease stage. They selected 355 colored images in their early stage, 347 colored images in their middle stage, and 382 colored images in their end stage. They applied the model based on only transfer learning and by using a multiclass support vector machine as a classifier algorithm after extracting the features from images. AlexNet outperformed the other two networks in both approaches, with accuracy rates of 89.69% and 93.4%, respectively. Their model used about 61 million training parameters and the size of the model is about 227 MB.

Sachin et al., 2020 [64] proposed a deep convolutional neural network based on well-known pre-trained models i.e. AlexNet and GoogleNet for the identification of three different types of soybean diseases i.e. FLS, bacterial blight, and brown spot. They employed the three models on a dataset consisting of 649 disease images and 550 health images after updating the three models by changing their entire layers and tuning their hyper-parameters such as the bias learning rate, max epoch, and mini-batch size. AlexNet achieved an accuracy of 96.25% and on the other side, GoogleNet achieved 98.75% of accuracy.

Chen et al., 2019 [65] suggested an improved deep convolutional neural network called LeafNet which improved the AlexNet model for disease detection on tea plant leaves. Their modified model consisted of various feature extractor filters with different sizes to extract the features automatically form the tea images. Their dataset consisted of 3810 tea leaf images divided into 7 classes for seven diseases, but they applied some data augmentation techniques to increase the number of images in their dataset and finally, there were 7905 images in their database. LeafNet achieved an accuracy of 90.16% and to compare their model with others they employed each of support vector machine classifier algorithm and multi-layer perceptron classifier algorithm on their dataset. In this case, they used a dense scale-invariant feature transform method for feature extraction and constructed a bag of visual words ready for classification by SVM and MLP classifiers. SVM achieved an average classification accuracy of 60.62% and 70.77% for the MLP classifier. Their modified model outperformed the other two classifiers SVM and MLP.

Arsenovic et al. 2019 [66] introduced a new dataset of 79265 images which were taken at different times of the day in various weather conditions. They introduced this new dataset as the datasets currently in use were taken in a controlled environment, and the models applied to these datasets may be misleading when used on real field data. They also increased the number of images in their dataset using the traditional data augmentation method and by applying a novel approach to data augmentation based on generative adversarial networks (GANs). They proposed a novel two-stage neural network called PlantDiseaseNet for plant disease detection and classification. The first stage is responsible for plant leaf detection and the other is responsible for classification. The accuracy of their model was 93.67%.

Jiang et al., 2019 [67] proposed a modified convolutional neural network called INAR-SSD for real-time detection of 5 common types of apple leaf diseases. Their model is based on the GoogLeNet Inception module for disease classification and a single-shot multi-box detector with Rainbow concatenation for disease object detection. They trained their model on a dataset called the Apple Leaf Disease Dataset (ALDD) which consisted of 26,377 augmented images of apple-diseased leaves collected under real field conditions. Their model can detect the same disease of different sizes in the same leaves diseased images. According to the experimental findings, the INAR-SSD model achieved 78.80% mAP of detection performance on ALDD with 23.13 FPS of detection speed.

(SINGH et al., 2019) [68] proposed a modified Multilayer Convolutional Neural Network based on the well-known AlexNet module for the classification of Anthracnose fungal disease on mango leaves. Their model consisted of 6 of 3×3 convolutional layers, 3 of 2×2 max-pooling layers, and 2 dense layers. They used ReLU as an activation function after each Conv layer and a Softmax activation function after thelast layer for multiclass classification purposes. They trained and tested their model on a dataset consisting of 2200 images (i.e., 1070 images collected in real-time and 1130 images from the plantVillage dataset). They first applied some pre-processing steps through employed histogram equalization methods for contrast images and central square crop method for image rescaling. This model achieved a performance of 97.13%.

Karthik et al., 2019 [69] proposed 2 deep learning approaches for detecting the type of three diseases on the tomato leaves plants namely leaf mold, late blight, and early blight. For better classification, they need to learn or extract the significant features so that they apply residual learning CNN. To improve the accuracy of residual CNN and specifically learn significant feature maps, they employed an attention mechanism on top of it as the attention mechanism gives more weightage to the significant features for accurate classification. This was the first attempt to implement an attention-based residual CNN. Their model was trained on 95999 augmented images from the PlantVillage dataset after applying a central zoom and random crop & zoom techniques to focus only on the leaf and not the background information. Their model achieved an accuracy of 98% on the validation sets in the 5-fold cross-validation and used around 600K training parameters.

Elhassouny et al., 2019 [70] developed a smart mobile application embedded with a deep convolutional neural network based on a well-known model called MobileNet to identify the most ten common types of tomato leaf diseases. The model accepted colored images of size 224 × 224 as input and it consisted of a set of 3×3 Depthwise Separable convolutions layers for reducing the number of computations followed by batch normalization and ReLU activation function and ended by average Pooling layer, dense layer with softmax function for 10 diseases classes. They trained their model on a PlantVillage dataset consisting of 7176 images of tomato leaves and achieved an accuracy of 90.3%.

Ashqar et al., 2019 [71] proposed a convolutional neural network model for tomato plant disease detection. It consisted of 4 convolutional layers followed by a max-pooling layer after each of them with a RELU activation function and two dense layers. In this work, the model was trained using both full-color and grayscale images. The total number of training parameters used for full-color images was 3,601,478 while those used for grayscale images were 1,994,374. They trained their model on a PlantVillage dataset consisting of 9000 images of tomato leaves and achieved an accuracy of 99.84%.

Costa et .al 2019 [72] proposed a modified deep learning model based on well-known InceptionV3 and CNN using a Hierarchical approach to identify 16 types of diseases in tomatoes, apples, and peaches. They trained their model on a PlantVillage dataset consisting of 24,000 images of tomato, apple, and peach leaves and achieved an accuracy of 97.74%.

Zhang et al 2019 [73] proposed a novel deep learning model based on the main problem of the well-known deep learning model AlexNet which has a large number of training parameters. Their model is called global pooling dilated convolutional neural network (GPDCNN) and it is used for cucumber plant disease identification. It consisted of a dilated convolution layer which extended the receptive field without losing resolution and a global pooling layer which was replaced with the dense layer in CNNs to reduce the number of parameters used. GPDCNN outperformed the classical convolutional neural network as it didn't increase the complexity of computations if the receptive field of the convolution increased and it didn't lose the discriminant formation if we replaced the dense layer with the global pooling layer. 94.65% accuracy was attained by the suggested model with 6.2 hours of training time and 3.58 seconds of testing time.

Ferentinos et al 2018 [74] employed various deep learning models for identifying and detecting the disease type on various types of plants using a dataset consisting of 87848 of health and infected images taken in both lab settings and real production fields. In the end, VGG outperformed the others with a performance of 99.5% using 138 million training parameters, beating the well-known models such as AlexNetOWTBn, GoogLeNet, AlexNet, VGG, and Overfeat models.

Liu et .al 2017 [75] applied various types of classifier algorithms for identifying and detecting 4 types of apple leaf diseases using a dataset consisting of 1053 images of apple leaves. They applied a support vector machine algorithm and a back propagation neural network for classification purposes. They also applied some well-known deep learning models such as ResNet-20, VGGNet-16, GoogLeNet, and AlexNet, and finally, they observed that the improved AlexNet achieved the best accuracy with 97.62%. In the case of detection, they applied both a "you only look once" algorithm for object detection, a one-shot detection algorithm, and a faster area-based convolutional neural network.

3.3 | Plant Diseases Detection Based on IoT

The Internet of Things is an extensive and open network of smart devices that can self-organize, share knowledge and resources, and act in response to situations and environmental changes [76]. Agricultural capacity has been increased by the Internet of Things (IoT) as their apps provide farmers with many services such as keeping them connected to the latest crops and informing them of weather information. IoT allows farmers to monitor their crops remotely [77]. Because they identify crop diseases early on before they spread to the crop, IoT applications have a significant impact on raising production and reducing crop losses. Farmers can also prepare for the upcoming season. Early detection of plant diseases is done in smart farms based on the concept of IoT as shown in Figure 7.

Multiple sensors measuring environmental characteristics, plant canopy, and leaf indices derived from remote sensing imaging, and IoT sensors can also be offered to the agriculture sector. Data fusion techniques are needed to compile such sorts of data given the range of retrieved data in order to better understand crop growth circumstances and the emergence of disease signs. In addition, machine learning-based data fusion has made great progress, and when it is applied to agricultural data, it will have an important impact on the field of plant protection, especially in the areas of disease detection and early detection of diseases. In order to achieve this, a variety of multi-sensor and remote sensing-based fusion techniques have been applied in agriculture [78, 79]. The most important aspects of theInternet of Things are sensors and actuators[80]. To collect and detect an event or make any changes in the surrounding environment IoT uses sensors that transform each of movement, thepresence of chemical substances, air humidity, temperature, and many other physical phenomena into their corresponding electrical impulses which can be usefully interpreted later. On the other hand, actuators deal opposite of sensors as they apply some actions or changes to the surrounding environment such as changing the position or angle of other devices, opening and closing valves, emitting sounds or light, and many other actions. It is done by interpreting the current electrical impulses and converting them into a mechanical movement so that we can consider actors as a mover. The Internet of Things has put sensors and applied some actuators to create an intelligent model that needs information from the surrounding environment with the help of cloud and theInternet [81].

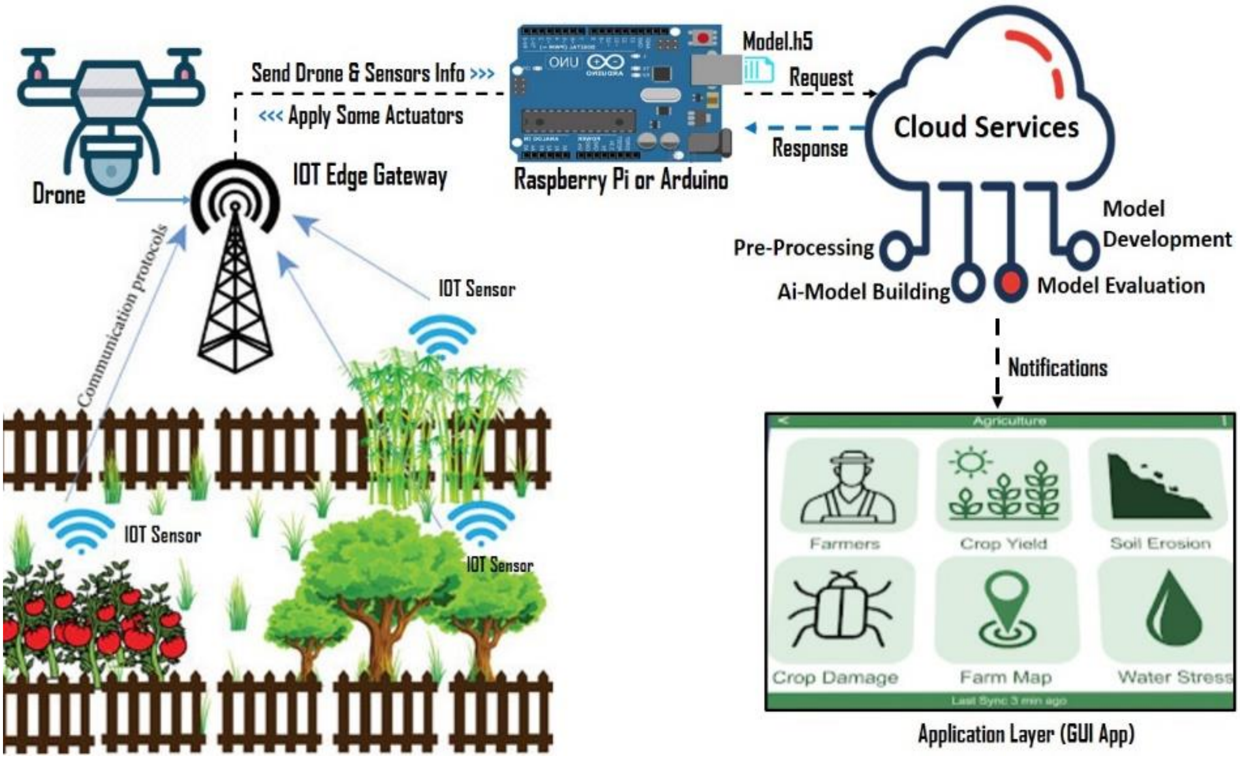

Figure 7. The main architecture of IoT agriculture's system.

Early detection of plant diseases is done in smart farms based on the concept of IoT as shown in Fig.6, first by collecting the required data from the farm using different devices where we can collect images of plant leaves using satellites or drones [82] or any digital cameras, and the physical phenomena such as temperature, movement and light, the presence of chemicals, air humidity, and many more that can be collected using IoT sensors. Then, this data will be transmitted to a Raspberry Pi or Arduino [26] where we can upload our training model to it for plant disease prediction with the help of cloud services. Finally, after pre-processing the images captured by the drone and classifying it with the model and other information sent by the sensors, the IoT system can take an action or trigger based on this data for example, if there are any diseases on the plants the system will alert the farmer with this through a buzzer which will ring. The state of the arts in this regard will be discussed in the next section.

M. Mishra et al. 2021 [83] developed an IoT system for monitoring crops and plant disease early detection. They put nodes over the environment to capture images of plant leaves and a sink node to receive information from the model used to help in monitoring the crops. They first applied some pre-processing steps on the collected images through median filter, then they applied a segmentation process on the images and extracted the segment-level features and the pixel-level features. They proposed a classifier algorithm based on arider neural network optimizer called sine cosine algorithm to make sure that the weights are chosen in an optimal way in the neural network. It finds out if the plant is suffering from any disease as early as possible which makes the loss of the crops small. This model didn't detect the type of disease as it detected only if the plant is infected or not. The proposed model achieved disease detection accuracy of 91.56 %.

Chen et al. 2021 [84] proposed the use of long short-term memory with a bi-directional recurrent neural network and using the various climate variables to predict the occurrence of the disease on the cotton plant. It is considered as a time series prediction problem whereas the long-term dependencies on past and future consecutive data contexts captured by Bi-LSTM network is used to deal with the issue. They observed that the prediction of cotton disease occurrence is performed well by applying Bi-LSTM in which it achieved 87.84% of accuracy with Area under the Curve of 0.95.

Nawaz et .al 2020 [85] proposed an inexpensive control framework based on internet of things for identifying the plants diseases using each of temperature, shade, humidity and etc. collected using a Wireless Sensor Network (WSN) from the farm. Their framework was deployed to check whether the plant is healthy or infected, identifying the closeness of disease in the plant and tracking the development of plant using some sensors and Arduino programming. Their model uses the help of cloud services to finally decide whether the plant is healthy or infected. Farmers with limited resources can buy it and use it to prevent the spread of diseases. But on the other side, because diseased plants' leaves cannot be classified using this method, the disease types are yet unknown.

Mugithe et al., 2020 [26] developed a plant leaves diseases identification and notification system based on internet of things which notifies the farmer if immediate action is required to stop the disease from spreading in the field. They used six-step image Processing (IP) approaches in their work. First, a webcam that was connected to a Raspberry Pi was used to capture live images of the leaves. Then, they pre-processed the images and segmented them and finally clustered those using the k-means clustering algorithm after which features (such as perimeter and light intensity) were extracted from the dataset. Finally, classification of the leaf diseases was performed using these derived features. When a disease is found, a buzzer sounds and an alert is produced so that the farmer may act quickly. Be aware that this system functions in two different ways, namely in real time and on the GUI. One of the advantages of this system is using a warning system to identify diseases, the accuracy of the disease results in the graphical interface was 95.16313%.

Chen et .al 2019 [86] implemented an IoT platform called RiceTalk to predict the disease on the rice plant. They depended on non-image IoT devices where their sensors collected these non-images data which are analyzed and trained after that using RiceTalk in a real time. They also depended on the basic image-based plant disease detection methods. The AI model is managed and controlled like other IoT devices, which is the magic of RiceTalk. Their model applied the real time prediction process with a lower cost. A new features extraction algorithm is proposed by them based on a novel spore germination mechanism. According to this study, the relative humidity min/max, range, and average measurements as well as temperature average have a big impact on predicting rice blasts. The average humidity measurement may be wrong for prediction in case the humidity variance (min/max) is high. RiceTalk achieved a prediction accuracy of 89.4%.

Devi et .al 2019 [87] proposed a quick and effective disease detection system that uses the Internet of Things (IoT) to identify and classify sigatoka diseases in hill banana plants and the bunchy top of banana disease. The proposed disease detection system processes plant images and extracts texture information using image processing and IoT. At the monitoring location, the agricultural experts classify the GLCM characteristics using the Random Forest Classification (RFC) technique in order to offer solutions. In addition to pathogens, climate changes also cause plant diseases. Therefore, it is essential to keep an eye on the agricultural field's environmental factors, such as the temperature and moisture of the soil. Temperature, humidity, and soil moisture sensors connected to the Raspberry Pi3 are used to measure these parameters. Their system achieved an accuracy close to 99%, the performance findings demonstrate that for the hill banana dataset, RFC-GLCM based classification outperforms other methods.

Win et al., 2018 [88] proposed two prototypes for rice plants diseases detection based on internet of things. The first one is a mobile app that classifies diseases on rice plants, without the help of agronomists, a farmer can quickly and easily detect diseases on rice plants with this straightforward application. The second one is anIoT system which keeps tabs on rice fields' temperatures, atmospheric pressure, water levels, and levels of sunlight. This environmental data is collected and monitored remotely from anywhere using only internet connection and low requirements functionality. They gathered six different types of rice images for the smartphone app's creation, including images of healthy rice, brown spot, mice attacks, bacterial leaf blight and rice blast. They used Arduino nano which is directly connected to the Raspberry Pi for solenoid valve turn on/off purpose every 30 minutes in order to ensure effective power usage and a lengthy system runtime. Transfer Learning and Deep Learning models were employed to categorize diseases. The benefit of this method is that they created a straightforward Android app to track a SensorTag's temperature and battery level. Consequently, it is simple to monitor the farm sensor data reading on smartphones or computers. On the other hand, they placed 8 SensorTags in various places. Less than 50 metres separated them from the Raspberry Pi. A different kind of communication system must be taken into consideration in place of the Bluetooth technology used by the SensorTags to monitor a large number of rice fields. SensorTag has a problem with its coil battery. Normally, the batteries keep the SensorTag active for at least a year. The SensorTag's LEDs flicker and the batteries only last a few weeks due to firmware level updates implemented to make it easier to advertise constantly. They frequently had to swap out the button batteries as a result.

Aher et al., 2018 [89] proposed a smartphone based on an internet of things used to gather data from various areas inside a farm. The farmers will have access to this data through the cloud service. Through a mobile application, this data can be accessed. The smartphone app will not only offer data in a graphical format, but it will also offer a wide range of helpful services for farmers. This research focuses on agricultural remote monitoring systems along with various farmer-friendly applications. The primary goal is to gather readings from several nodes and assist farmers in managing various operations wirelessly while offering a smart agricultural field for knowledgeable farmers.

Thorat et al., 2017 [90] proposed a system for detecting and identification of plant leaves diseases which is used insmart farming. They used each of a Raspberry PI module, Wireless Sensor Network and a camera. In addition, Computer Vision (CV) methods like segmentation, feature extraction, and masking were applied to find leaf diseases. It should be emphasized that they send and receive data via the Apache server. The proposed technology enables agricultural monitoring from a distance. As a result, the identification of many leaf diseases was successful. On the other hand, because the system power supply is restricted, the malfunction may cause the whole process to stop, which is inconvenient. In addition, images taken during the day may be affected by many factors such as reflections, sunlight and etc. making it difficult for the camera to discern the colors of leaves during the day or to capture them clearly at night.

Patil et al., 2016 [91] proposed a monitoring system based on hidden Markov model for grape diseases early detection in which it sends an SMS to a farmer as a notification if there is any disease detected on the grape. Moisture, Zig-Bee are usedfor wireless data transmission, a leaf wetness sensor, temperature, relative humidity and etc. are used in this system. Their system achieved an accuracy of 90.9%.

4 | Comparative Study

In the previous sections, we discussed the state of the arts of plants diseases identification and detection approaches depending on each of machine learning, deep learning algorithms and IOT. But most of the emphasis has been placed on states of the art based on deep learning algorithms because most of the researchers focus on deep learning algorithms; because the classification algorithms of machine learning extract the features manually but on the other side, deep learning algorithms apply this step automatically and the accuracy of the model is better than machine learning algorithms. The most suitable deep learning algorithm for image classification purpose is Convolutional Neural Networks algorithm and thus, a comparative study will be done in this section among various recent states of the art for plant disease classification depending on deep learning algorithms especially Convolutional Neural Networks algorithm.

We noticed that a lot of models achieved good accuracy but a lot of them use a hug number of trainable parameters and this made running time increase also. For example, Ahmed et.al 2022 [48] achieved 99.30% accuracy on tomato plants diseases using PlantVillage dataset and on the other side, Ashqar et al., 2019 [67] achieved 99.84% accuracy on tomato plants diseases using the same dataset but the difference here is that Ashqar et al., 2019 used 3.6 million trainable parameters and Ahmed et.al 2022 used only 2.4 million trainable parameters and this made the running time of the first study better than the second one with a tiny less accuracy. Mohit et al., 2020 [62] achieved 91.2% accuracy with 208,802 trainable parameters and this made it run faster than the other two studies but with less accuracy. In this regard, this leads us to the tradeoff between accuracy and number of trainable parameters the model uses. We applied a comparative study between various recent states of the arts for plant leaf disease classification depending on the type of deep learning model architecture, where improved/modified models, cascade/hybrid models and well-known models represent the three classes of deep learning structures as shown in Figure 8. The comparison is done in the terms of the used model architecture (MA), used model (Model), used plant (plant), number of used images (NI), image size (IS), number of classes (NC), dataset distribution (DS), accuracy (Acc), F1-Score (F1S), Precision (Pre), Recall (Rec), no. of trainable parameters (NTP) per millions, model size (MS) per megabytes, Training time (TRT) per seconds as shown in Table 2.

There are some popular CNN models but the most well-known ones are VGG-16 [92], ResNet-50 [44] , AlexNet [40] , DenseNet-121 [93] , Inception-v4 [94] , Inception ResNet-v2 [94] , MobileNet [95] , Xception [41] , OverFeat [96] and ZFNet [97]. To achieve better accuracy to detect and classify the plant leaves disease, improved or modified versions of well-known deep learning architectures are proposed by some researchers such as reduced and modified MobileNet [98] which is inspired by the MobileNet model [95], improved GoogLeNet [99], Cifar-10 [99], LeafNet [67], a Multilayer Convolutional Neural Network (MLCNN) [68] which is inspired by the AlexNet model[40], and improved GoogLeNet [99], which is inspired by the well-known GoogLeNet model [42]. The last category of deep learning architectures is the cascaded/hybrid variants of deep learning architectures, such as a hybrid deep learning architecture of the well-known AlexNet with VGG models as proposed in [108] and a cascaded version of the well-known AlexNet with GoogLeNet models as explained in [75]. The study showed that the dataset most often used by researchers is the PlantVillage dataset and the plant most often used by researchers is tomato plant as shown in Table 2. Figure 9 shows a comparison between five Convolutional Neural Network models on tomato plants diseases on a plantvillage dataset, we have noticed that four of them have trainable parameters of over a million and only one of them has a small number of trainable parameters around 208,802 parameters and this is a good number but on the other side with less accuracy. We can improve the other four models by trying to reduce the number of training parameters which in turn will reduce the training time of the model and the model size. As a future work, we can propose a new model with a smaller number of trainable parameters and good accuracy compared to these five models and this new model must have trainable parameters less than 1.4 million and performance better than 98.49% as shown in Figure 10.

Table 2. Comparative study between various recent states of the arts for different plant leaves disease.

MA	Model (Ref.)	plant	Dataset				Model Performance (%)				Model Complexity		
			NI	IS	NC	DS	Acc	F1S	Pre	Rec	NTP	MS	TRT
Well-known	AlexNet [63]	tomato	1909	227	10	80:20	93.4	93	92.99	93.02	61	227	–
	MobileNetV2[48]	tomato	18160	256	10	60:20:20	99.3	99.12	99.18	99.07	–	9.60	–
	EfficientNetB5[100]	tomato	11000	200	10	10 K-Fold	99.07	99.5	99.5	99.5	28.8	–	–
	DenseNet121 [50]	tomato	16012	224	10	60:10:30	97.11	97	97	97	1.74	–	–
	SqueezeNet [101]	Banana	937	–	4	–	96.25	96.17	96.53	96.25	0.737	4.78	–
Improved / Modified	CNN [62]	tomato	17500	224	10	57:40:3	91.2	91	91	91	0.21	1.5	–
	CNN [51]	tomato	3000	256	10	–	98.49	–	–	–	1.42	22.5	–
	CNN [102]	tomato	3000	128	10	70:30	88.17	99	99	99	1.06	–	–
	CNN [103]	tomato	18160	224	10	5 K-Fold	96.87	0.97	0.97	0.968	0.49	6	5820
	Attention embedded Residual CNN model [69]	tomato	95999	256	10	5 K-Fold	98.0	–	–	–	0.60	–	36000
	Convolutional Autoencoder network [53]	peach	4457	256	2	–	98.38	98.36	98.0	98.72	0.009	–	–
	Modified MCNN based on Inception V3[49]	Vary species of crops	–	224	38	60:20:20	99.48	0.992	–	–	5	–	–
Cascaded / Hybrid	Vgg16 with svm [58]	14 crops	54000	–	36	–	97.82	96.42	–	–	138	–	–
	ResNet50 with MobileNet [104]	olive	5400	224	4	80:20	97.08	96.86	97.61	97.11	–	–	1138
	MobileNetV2 with U-Net [105]	Guava	1316	416	5	73:18:9	83.40	–	73.3	73.1	–	–	–
	EfficientNet B7 [106]	grape	9027	224	4	80:20	98.7	94.0	95.0	22.0	66	–	–
	DenseNet with improved U-Net & ROI extraction algorithm [107]	Rice	2988	128	4	5 K-Fold	96.0	–	–	–	18.5	–	–

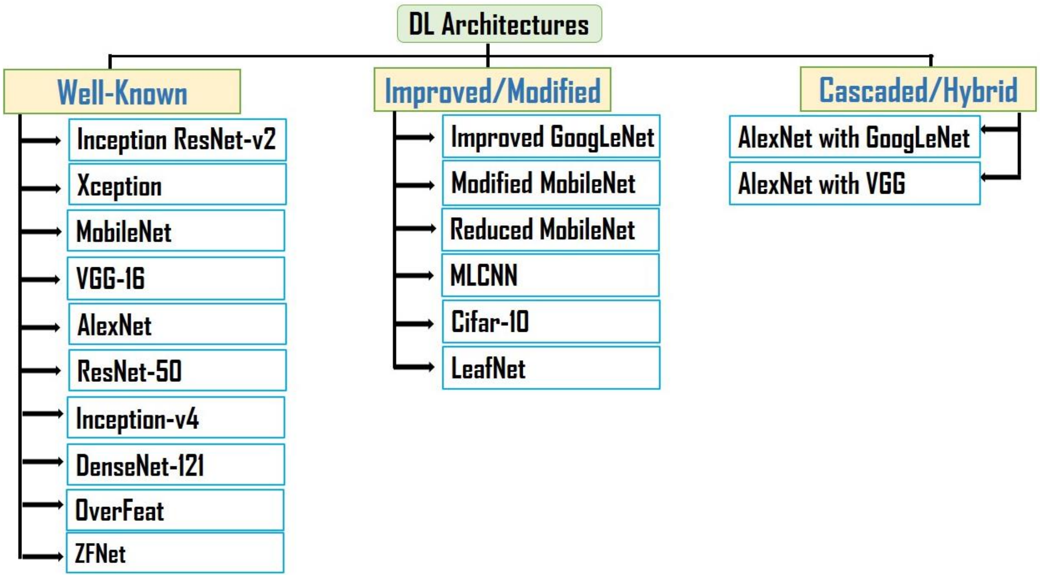

Figure 8. The three main deep learning architectures categories classification models.

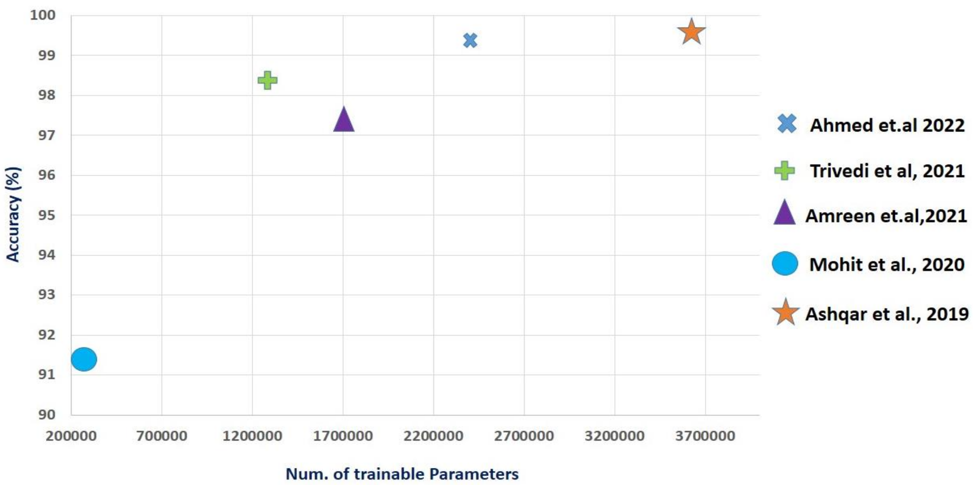

Figure 9. Accuracy vs num. of trainable parameters for five tomato plant disease detection models.

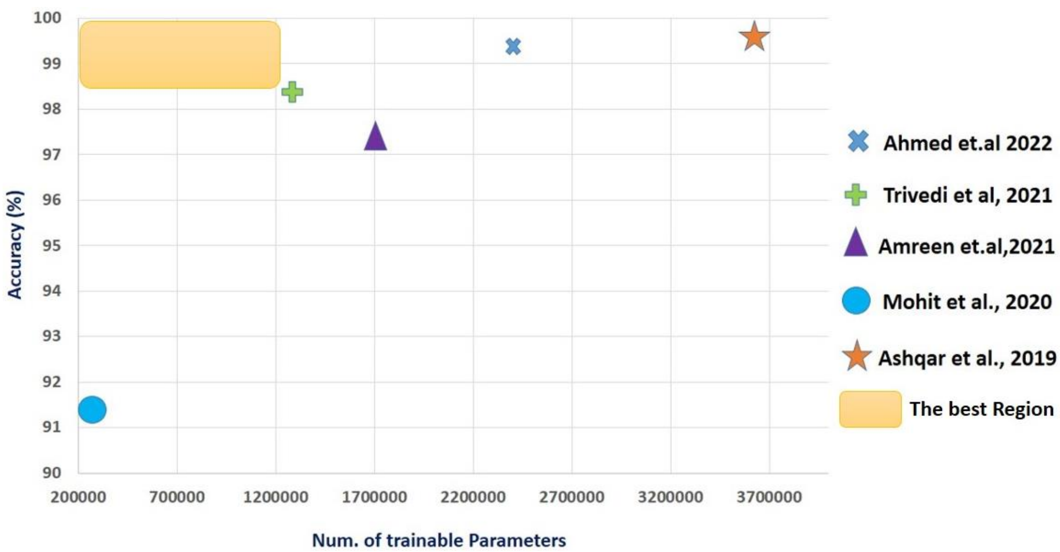

Figure 10. The desirable region for any new model based on above five-tomato plant disease detection models.

5 | Challenges

The detection and identification of plant leaf diseases is mainly affected by various challenges affecting the performance and accuracy of AI technologies. In this section, we will summarize and discuss the challenges researchers face and possible solutions to each. Figure 11 shows some of the challenges that faced the researchers discussed in this paper with corresponding solutions.

5.1 | Insufficient Datasets

There are some researchers who have trained their models using their own datasets which they have collected in real time but the main issue here is related to the small size of their dataset for training therefore most of the researchers used open and available datasets since they are larger in size. The most well-known open and available dataset for plant leaves diseases images is the plant-village dataset which consists of 54,305 images of 38 different infected and healthy images for 14 plant species. But on the other hand, compared to other datasets such as the ImageNet dataset that consists of more than 14 million images, the plantvillage dataset is still a small dataset for training a deep learning model. This problem is the main challenge for plant disease detection models in terms of scale and diversity[109]. There are many solutions to overcome this problem such as citizen science [110], data sharing , transfer learning[111] , few-shot learning [112], and data augmentation approaches [113]. There are other techniques used to generate a new synthetic images besides real images on the dataset such as Generative Adversarial Networks (GANs) [114] and Variational Automatic Encoder (VAE) [115].

5.2 | Imbalanced Data

An uneven distribution of classes within a dataset is typically the cause of data imbalance. But to avoid being distracted by any issues during training time and to be fully focused, the dataset has to be cleaned up before starting in training or we can ignore their unbalanced nature. The distribution between classes is skewed and unbalanced in real-world settings, and ranges from somewhat biased to severely unbalanced [116]. Machine learning classification algorithms are based on having an equal number of images in each class, and this is a challenge for predictive modeling and may require specific strategies as the minority class couldn't be classified well enough like the others with more images i.e., majority class. We can solve this problem using resampling technique which includes under-sampling (removing samples from the majority class) and/or over-sampling (increasing the number of examples from the minority class).

5.3 | Image Preprocessing, Segmentation and Symptom Discrimination

Before training the algorithm on images, there are some pre-processing steps done because there is some noise on the images and we need to enhance some features of the image as symptoms usually do not have clear boundaries; they gradually disappear in normal tissues, making it difficult to distinguish between healthy and diseased parts. The accuracy of the threshold and the features retrieved are certainly affected by this. However, some information may be lost due to increased compression. Large lesions may not be much affected by this, while small symptoms may be greatly affected. Therefore, if the symptoms are small, compression should be reduced or even avoided. Image segmentation is done after the preprocessing step to isolate infected parts from the whole leaf image and locate the regions of interest on the image but there are some issues that happen when applying this segmentation such as that the complex background of images could render the process of segmentation of the region of interest and also leaf images could be overlapped with others or other parts of the same image. There are some solutions for these issues [117], since strong features will usually be focused more sharply than the rest of the image, we may establish a measure of the sharpness of their main features. So that applying some enhancements on the images on the edges will help greatly in the segmentation process.

5.4 | Small-size Lesions in Early Identification

Plant leaves images may have a large and small sizes of lesion (ROI). In case of large sized lesions, the algorithm detection may be wrong in many cases due to the problem of noise on large sized complex backgrounds especially on low resolution images. On the other side, in case of early detection of plant diseases small sized lesions may be ignored during down-sampling operations on pooling layers in the phase of features extraction. The main technique used in this case for solving the problem of discarding small sized lesions or features is called attention-based mechanisms. Finding a region of interest rapidly and ignoring unnecessary data is the core function of the attention mechanism to achieve more early accurate detection. Using the plantVillage dataset, experiments by Karthik et al. [69] on the residual network with attention mechanism produced an overall accuracy of 98%.

5.5 | Overfitting and Underfitting Problem

Underfitting and overfitting are two issues that can affect the training process and reduce the performance of the deep learning model during the training stage [118]. The underfitting problem happens when the model can't be able to identify the basic features in the images frequently and this indicates that there is a high bias and low variance and the model doesn't fit the trained data. This problem usually happens when there aren't enough data to train or when using nonlinear data to create a linear model. On the other side, overfitting problem occurred when the model has been trained using a lot of data and has learned from the dataset's inaccurate and noisy data inputs. The model then fails to appropriately identify the data as a result of the noise and excessive detail. It merely indicates minimal bias and great variance when it occurs. Nonlinear and nonparametric techniques produce overfitting because these learning algorithms have more freedom to construct a synthetic model. This issue was raised by Ahmad et al. [119] who found that their model, which uses effective convolutional neural networks, has a tendency to overfit when training in the early epochs. The model may become overfitted if it is trained over a larger number of epochs, and it may become underfitted if it is trained over fewer epochs. Here, the Keras library's Early Stopping approach is applied. When the validation process shows that the model's performance has reached saturation, training is terminated. The DL-based classifier additionally incorporates the ModelCheckpoint callback to save the top-performing model at the end of each epoch.

5.6 | Detection and Recognition Speed Problem

Deep learning algorithms achieve better results than most conventional techniques, but they also have more computing complexity. To achieve a high detection and identification accuracy, the model must fully learn all the strong image features and this leads to high computational complexity which in return leads to slow detection speed and in the same time not meeting the real time requirements. Usually, less computation operations need to be done in order to guarantee detection speed but this also leads to insufficient training and missed detection. Therefore, we need to develop an algorithm that is fast and at the same time still have accurate accuracy. Today, the majority of plant disease and pest detection techniques focuses on accurate recognition. One of the main suitable approaches that achieve a good speed with good accuracy is a deep separable convolution network. Kamal et al. [98] proposed a deep separable convolution network model for plant leaf disease detection which showed that the separable convolution network achieved a successful trade-off between detection speed and accuracy, making it suitable for real-time identification of crop diseases.

Another approach that can be used to address this issue without significantly lowering classification accuracy is the Convolutional AutoEncoder network (CAE) which is used to reduce the number of features on the final features map used for prediction or classification after that which in return reduces the number of training parameters. Punam et al. 2021 [51] proposed a novel hybrid model that uses a CAE with a CNN for detecting the peach diseases using only 9,914 training parameters with an accuracy of about 98.38%.

Figure 11. Some of examples on the challenges faced the previous state of arts with their solutions.

5.7 | Other Challenges

Some plant species have similar leaf shapes, despite the fact that some species of plants may be recognized by their leaves. Furthermore, symptoms may not always appear in areas that are straightforward to access; in fact, they may sometimes be hidden by leaves or other obstructions, or diseases may show themselves as symptoms on stems, fruits, or even flowers. Unfortunately, researchers haven't given the latter issue enough of their attention. Most researchers focus on developing plant disease detection based on the plant's upper leaves surface only. Another challenge that researchers faced is that a leaf of plant can be infected by various diseases at the same time in which the leaf consists of hybrid symptoms which may be challenging to recognize. The methods must rely on extremely slight changes to distinguish between distinct diseases whose symptoms may look physically similar.

6 | Proposed Methodology

Based on the previous states of the art that have been discussed in this paper and based on the challenges faced the researchers we propose a new method divided into three stages. In the first stage, we enhanced images using some preprocessing steps and after that we segmented image using k-means clustering algorithm. The second stage contains a two-block diagram, the first one depends on building the model using a convolutional neural network with a hyper-parameter tuning with the help of one of the meta-heuristic optimization algorithms and reducing the number of extracted features using a modified PSO algorithm, the second block diagram depends on the transfer learning technique [111]. The disease classification is applied in the last stage. More through description is provided on the following subsections.

6.1 | Image Pre-processing & Segmentation Phase

As we mentioned before, symptoms usually do not have clear boundaries; they gradually disappear in normal tissues, making it difficult to distinguish between healthy and infected parts. The accuracy of the model and retrieved features is definitely impacted by this. Small sized lesions may be ignored during down-sampling operations on pooling layers in the phase of features extraction. So that, image preprocessing is one of the most important steps to enhance the features on the leaf images with illumination correction before feeding it to the model to learn from it for improved classification.

The entire procedures that are done to enhance the images before moving on to the training phase as shown in Figure 12 is firstly, convert RGB image to HSV (Hue, Saturation, Value) color space image which is more suitable for image segmentation purpose and it more better than other color models [120], after that we split the image to 3 channels (Hue, Saturation, Value) as shown in Figure 13.

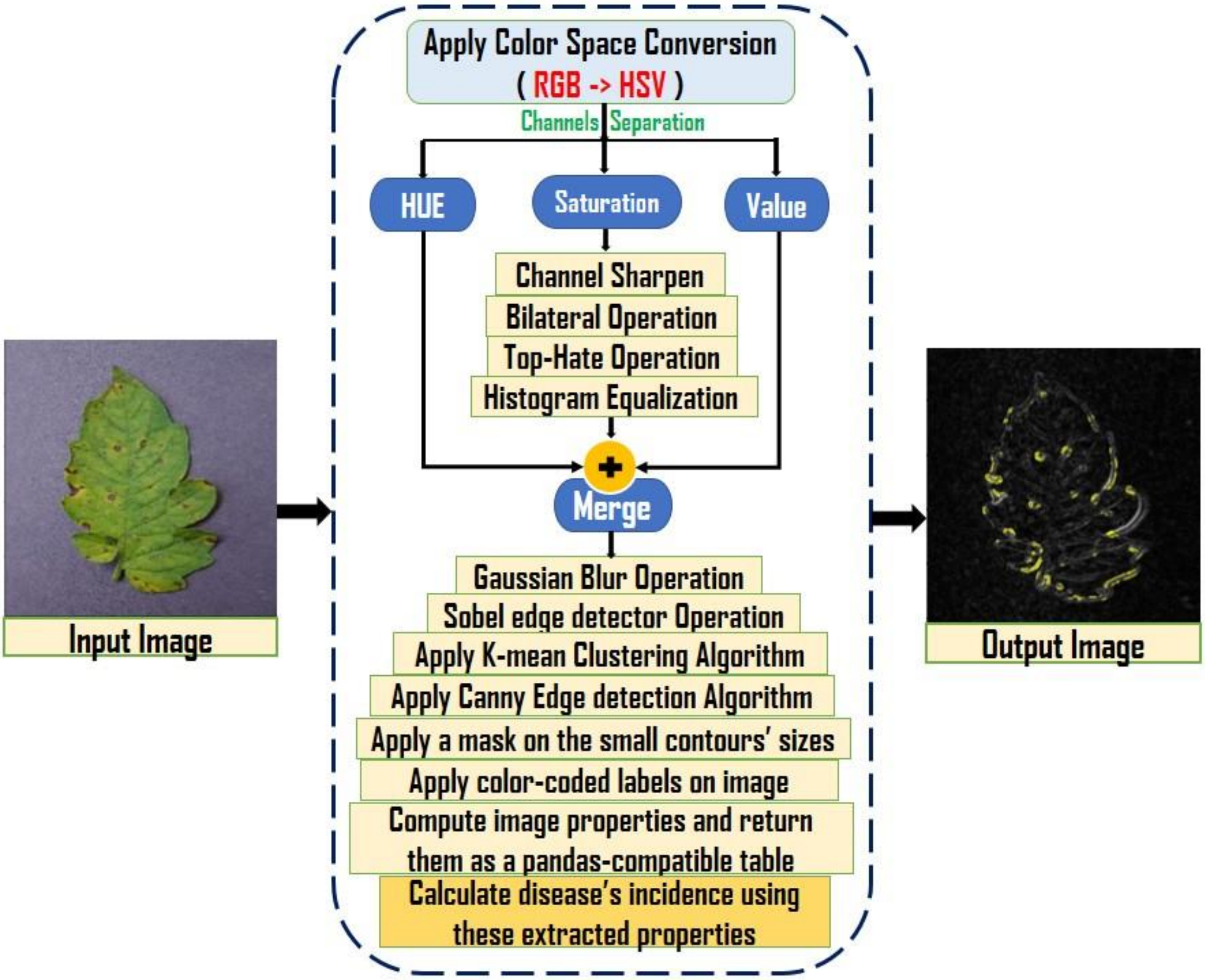

Figure 12. The stage of image pre-processing and image segmentation on our proposed methodology which is done to enhance the images before moving on to the training phase.

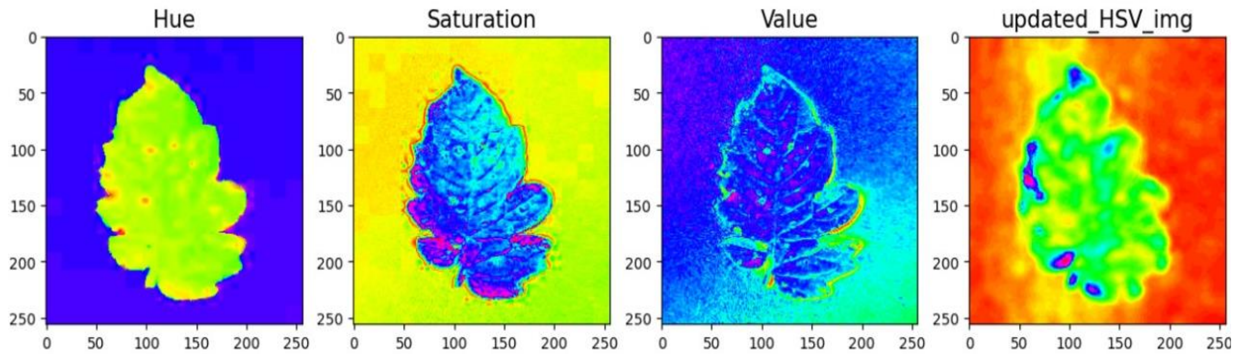

Figure 13. H, S, V separated channels and the updated image after merging them again.

Images were taken in various lighting conditions as some of them have a bright spots or regions and the others have dark regions which in return creates an imbalanced histogram. One of the most used techniques for correcting this imbalanced histogram and enhancing the details on the images is the histogram equalization techniques such as the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique which applies a local contrast enhancement in order to get around the limitations of global techniques. Figure 14 shows the enhanced image after applying the CLAHE algorithm on it, we are obvious that the region of lesions became more visible and contrasted. We applied sharpen kernel and Bilateral operation with a top-hat operation for eliminating unwanted noise and improve the infected regions' contrast. Each of these operations were applied to the HSV image saturation channel. Finally, we combined the adjusted saturation channel with the other two channels (Hue, Value). After that, we applied a Gaussian-blur Operation with a discrete differentiation operator called Sobel Operator to create an image with focus on the edges.

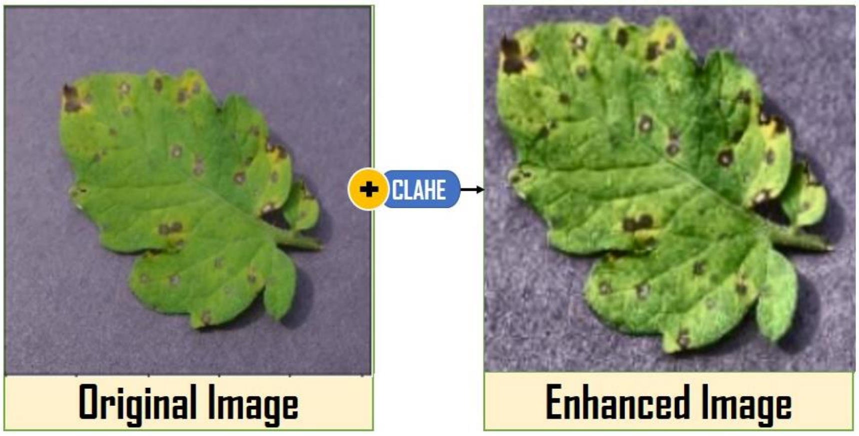

Figure 14. Enhanced tomato plant leaf image after applying CLAHE algorithm on it using python code.

To segment our image, we applied K-means clustering algorithm with 10 attempts and k=8, after that we applied canny edge detection algorithm and then contours with different sizes were founded, saved in a sorted list and the small contours (infected regions) were masked based on a threshold as shown in Figure 15.

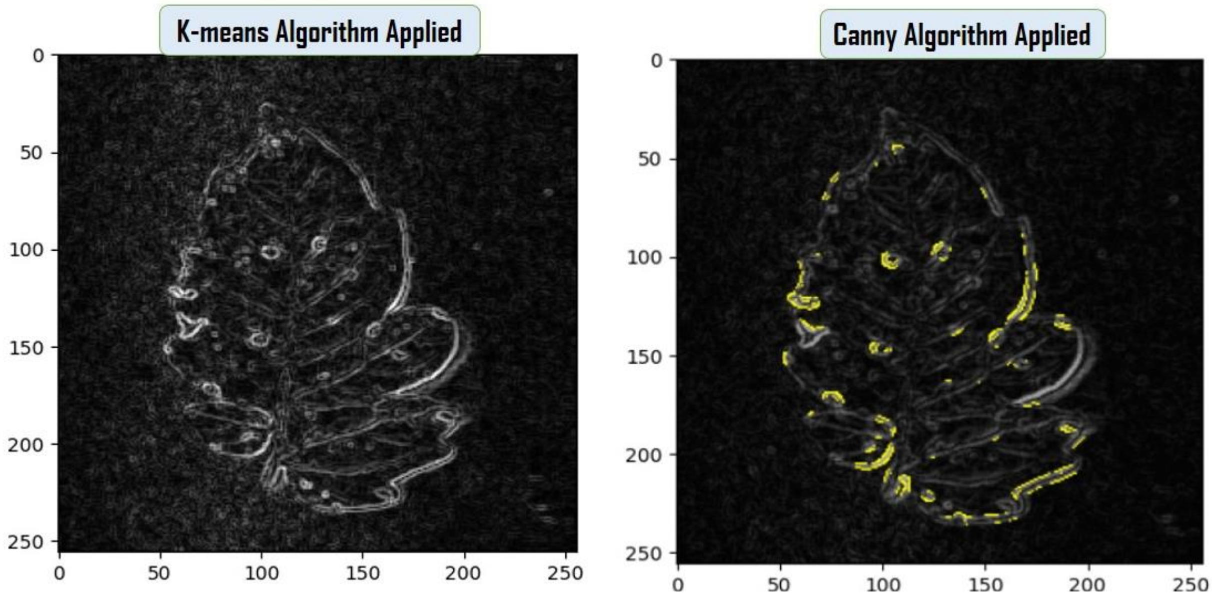

Figure 15. Segmentation phase using K-means clustering algorithm with 10 attempts and canny edge detection algorithm

We after that applied a color-coded labels on the image and computed image properties which stored as a pandas-compatible table as shown in Figure 16. These extracted properties are used after that to calculate the disease's incidence on these plant's leaf using Eq. (1).

$$\boldsymbol{Disease\ Incidence}(\%) = (The\ total\ area\ of\ the\ infected\ colored\ coded\ regions\ ) / (The\ total\ area\ of\ the\ plant'\ s\ leaf) * 100 \tag{1}$$

	label	area	mean_intensity-0	mean_intensity-1	mean_intensity-2	solidity
0	1	8	62.750000	134.125000	137.875000	0.571429
1	2	1	76.000000	152.000000	151.000000	1.000000
2	3	2	67.000000	156.500000	137.500000	1.000000
3	4	13	85.692308	126.076923	110.538462	1.000000
4	5	5	96.400000	124.600000	104.200000	0.625000

Figure 16. Colored coded labels image with the properties of the first five labels which stored as a pandas-compatible table

6.2 | Building the Model Phase

We propose two models, the first one as shown in Figure 17 is a hybrid convolutional neural network inspired by Inception V3 architecture which consists of five convolutional layers and two max-pooling layers. In this model, hyper-parameters such as number of layers, loss function, optimizer, activation function at intermediate layers, number of epochs, activation function at output layer and batch size, etc. are tuning using one of the meta-heuristic optimization algorithms such as genetic algorithm, practical swarm optimization, gray wolf optimizer, etc. To select the strongest features from features maps we use a modified meta-heuristic algorithm based on practical swarm optimization algorithm which reduces the number of features which in return reduces the number of trained parameters. The second proposed model depend on a transfer learning technique such as Efficient Net B4, MobileNetV2, etc. which we use them as features extraction as shown in Figure 18.

6.3 | Classification Phase

Now we have a features map with the strongest features ready for classification. On model 2, we use one of the machine learning classification algorithms such as Bagged Tree Algorithm, Random Forest Algorithm, etc. to detect the type of the disease. Beside to a new method which will be overviewed on our future work which classify these features in less time and more accurate classification in which it will achieve the trade-off between model performance and model complexity. Finally, we compare the performance matrixes of the two models and choose the best or make ensemble model using them. Finally, the Block diagram of our proposed methodology is shown in Figure 18.

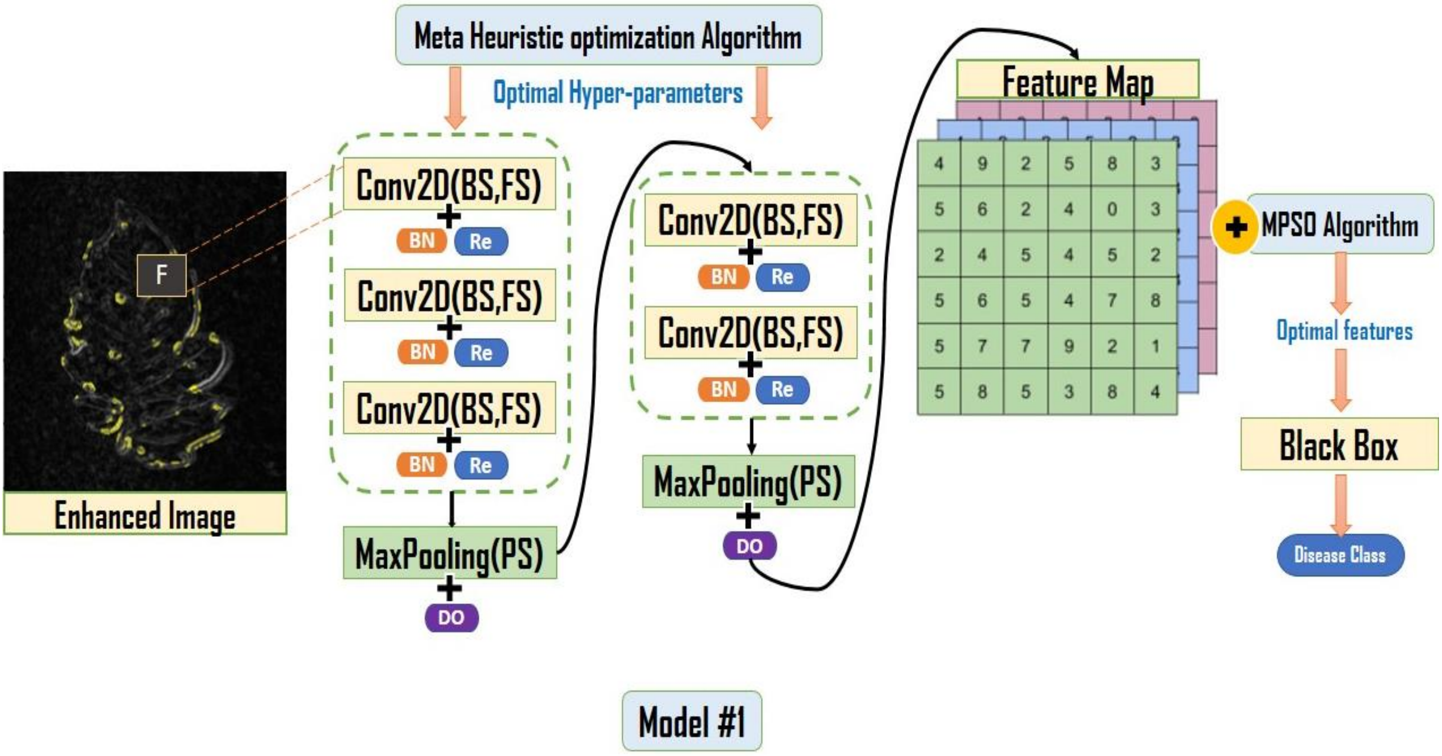

Figure 17. Model 1 which inspired from inception V3 model which depending on fetching the best hyper parameters using one of meta-heuristic optimization algorithm and finding the strongest features

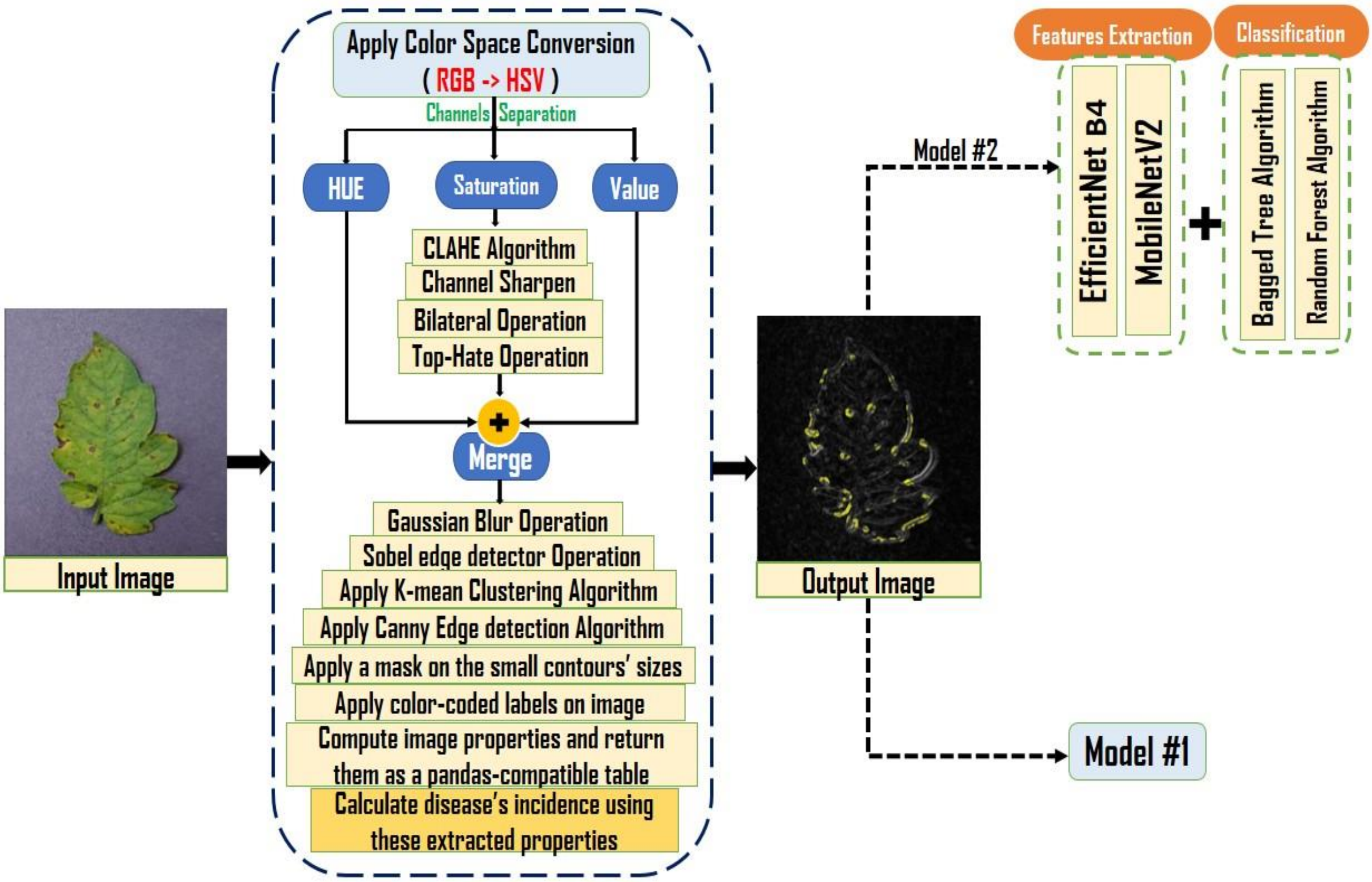

Figure 18. The framework of our proposed tomato plant leaf disease classification model from the pre-processing and segmentation stage towards feature extraction and classification phase.

7 | Conclusion and Future work

In this work, various plant disease detection methods that have been developed by many researchers in recent years such as machine learning, deep learning, and the Internet of Things are reviewed and summarized. We are obvious that each of the k-means classification algorithms, k-nearest neighbor algorithm, support vector machines algorithm, artificial neural networks, and convolutional neural networks algorithm are the most used classification algorithms in the current studies. The histogram equalization method is the one that researchers use most frequently to increase contrast in images during the preprocessing stage. For image segmentation and extracting the region of interest on the leaves images researchers primarily utilize some of the fuzzy-based algorithms discussed in this paper. For feature extraction such as texture, shape, and color researchers primarily use gray-level co-occurrence matrix, local binary patterns, and histogram of oriented gradients, etc. In the classification phase, researchers use a variety of classification algorithms based on machine learning algorithms such as artificial neural networks, decision tree classifiers, naive Bayes classifiers, support vector machines classifiers, random forests, etc. but the most used one is the support vector machine algorithm. Researchers use SVM primarily because it provides excellent classification performance by nonlinearly transforming the input feature vector into a high dimensional space where it can be easily separated. But in case of extreme noise in the data, SVM is not the appropriate choice for classification purposes [121].

Our proposed methodology concentrates on disease detection speed and accuracy. There was a trade-off between speed and accuracy in the last studies but we tried to reach for the most suitable hyper-parameters which make the model achieve good accuracy besides trying to reduce the number of trainable parameters. We suggested that using modified meta-heuristic optimization algorithms depends on a practical swarm optimization algorithm to reduce the features extracted. Finally, it is an appropriate choice to build platforms such as mobile applications, and web applications integrated with deep convolutional neural networks to achieve good accuracy with a good detection speed. Additionally, due to its critical importance, we will focus more on identifying diseases in various sites on plants and trees, such as fruits, blooms, and stems as future direction.

Acknowledgments

Acknowledgments enable you to thank all those who have helped in carrying out the research. Careful thought needs to be given concerning those whose help should be acknowledged and in what order. The general advice is to express your appreciation concisely and to avoid strong emotive language.

Author Contribution

Research plan execution, **Mahmoud H. Alnamoly, Anar A. Hady**; collect data, **Mahmoud H. Alnamoly**; Validation, **Mahmoud H. Alnamoly**; Writing – original draft, **Mahmoud H. Alnamoly**; Review & editing, **Anar A. Hady**; Supervision, **Anar A. Hady, Sherine M. Abd El-Kader** and **Ibrahim El-henawy**; Overall research plan, **Sherine M. Abd El-Kader** and **Ibrahim El-henawy**

Funding

This research has no funding source.

Data Availability

The dataset used in this study is available at Kaggle on the following website: https://www.kaggle.com/datasets/cookiefinder/tomato-disease-multiple-sources. Researchers interested in accessing the dataset can do so by visiting the provided link.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Ethical Approval

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

References

[1] Zhou, R., et al., Disease detection of Cercospora Leaf Spot in sugar beet by robust template matching. Computers and electronics in agriculture, 2014. 108: p. 58-70.

[2] Liu, L., et al., A disease index for efficiently detecting wheat fusarium head blight using sentinel-2 multispectral imagery. IEEE Access, 2020. 8: p. 52181-52191.

[3] Prasad, R., K.R. Ranjan, and A. Sinha, AMRAPALIKA: An expert system for the diagnosis of pests, diseases, and disorders in Indian mango. Knowledge-based systems, 2006. 19(1): p. 9-21.

[4] Singh, V. and A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture, 2017. 4(1): p. 41-49.

[5] Idoje, G., T. Dagiuklas, and M. Iqbal, Survey for smart farming technologies: Challenges and issues. Computers & Electrical Engineering, 2021. 92: p. 107104.

[6] Zhu, N., et al., Deep learning for smart agriculture: Concepts, tools, applications, and opportunities. International Journal of Agricultural and Biological Engineering, 2018. 11(4): p. 32-44.

[7] Alzubaidi, L., et al., Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 2021. 8(1): p. 1-74.

[8] Kaur, M. and R. Bhatia. Leaf disease detection and classification: A comprehensive survey. in Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019), NITTTR Chandigarh, India. 2020. Springer.

[9] Kaur, S., S. Pandey, and S. Goel, Plants disease identification and classification through leaf images: A survey. Archives of Computational Methods in Engineering, 2019. 26(2): p. 507-530.

[10] Chouhan, S.S., U.P. Singh, and S. Jain, Applications of computer vision in plant pathology: a survey. Archives of computational methods in engineering, 2020. 27(2): p. 611-632.

[11] Orchi, H., M. Sadik, and M. Khaldoun, On using artificial intelligence and the internet of things for crop disease detection: A contemporary survey. Agriculture, 2021. 12(1): p. 9.

[12] Ayaz, M., et al., Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. IEEE access, 2019. 7: p. 129551-129583.

[13] Ünal, Z., Smart farming becomes even smarter with deep learning—a bibliographical analysis. IEEE Access, 2020. 8: p. 105587-105609.

[14] Guo, Y., et al., Plant disease identification based on deep learning algorithm in smart farming. Discrete Dynamics in Nature and Society, 2020. 2020.

[15] Hughes, D. and M. Salathé, An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv preprint arXiv, 2015. 1312.6229.

[16] El Naqa, I. and M.J. Murphy, What is machine learning?, in machine learning in radiation oncology. 2015, Springer. p. 3-11.

[17] Fiona, J.R. and J. Anitha. Automated detection of plant diseases and crop analysis in agriculture using image processing techniques: A survey. in 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). 2019. IEEE.

[18] Perumal, S. and T. Velmurugan, Preprocessing by contrast enhancement techniques for medical images. International Journal of Pure and Applied Mathematics, 2018. 118(18): p. 3681-3688.

[19] Iqbal, M.A. and K.H. Talukder. Detection of potato disease using image segmentation and machine learning. in 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). 2020. IEEE.

[20] Srivastava, A.R. and M. Venkatesan, Tea leaf disease prediction using texture-based image processing, in Emerging Research in Data Engineering Systems and Computer Communications. 2020, Springer. p. 17-25.

[21] Whitmire, C.D., et al., Using machine learning and feature selection for alfalfa yield prediction. AI, 2021. 2(1): p. 71-88.

[22] Singh, S., et al., Extraction of multiple diseases in apple leaf using machine learning. International Journal of Image and Graphics, 2022. 22(03): p. 2140009.

[23] Kianat, J., et al., A joint framework of feature reduction and robust feature selection for cucumber leaf diseases recognition. Optik, 2021. 240: p. 166566.

[24] Almadhor, A., et al., AI-driven framework for recognition of guava plant diseases through machine learning from DSLR camera sensor based high resolution imagery. Sensors, 2021. 21(11): p. 3830.

[25] Pothen, M.E. and M.L. Pai. Detection of rice leaf diseases using image processing. in 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). 2020. IEEE.

[26] Mugithe, P.K., et al. Image processing technique for automatic detection of plant diseases and alerting system in agricultural farms. in 2020 International Conference on Communication and Signal Processing (ICCSP). 2020. IEEE.

[27] Al-bayati, J.S.H. and B.B. Üstündağ, Evolutionary feature optimization for plant leaf disease detection by deep neural networks. International Journal of Computational Intelligence Systems, 2020. 13(1): p. 12.

[28] Kumar, D.A., P.S. Chakravarthi, and K.S. Babu. Multiclass Support Vector Machine based Plant Leaf Diseases Identification from Color, Texture and Shape Features. in 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT). 2020. IEEE.

[29] Khan, M.A., Detection and classification of plant diseases using image processing and multiclass support vector machine. Int. J. Comput. Trends Technol, 2020. 68(4): p. 5-11.
[30] Hossain, E., M.F. Hossain, and M.A. Rahaman. A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier. in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). 2019. IEEE.
[31] Rauf, H.T., et al., A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning. Data in brief, 2019. 26: p. 104340.
[32] Ali, H., et al., Symptom based automated detection of citrus diseases using color histogram and textural descriptors. Computers and Electronics in agriculture, 2017. 138: p. 92-104.
[33] Rananavare, L.B. and S. Chitnis. Crop Yield Prediction Using Temporal Data. in 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). 2022. IEEE.
[34] Sarker, I.H., et al., Cybersecurity data science: an overview from machine learning perspective. Journal of Big data, 2020. 7(1): p. 1-29.
[35] Xin, Y., et al., Machine learning and deep learning methods for cybersecurity. Ieee access, 2018. 6: p. 35365-35381.
[36] Tian, H., et al., Computer vision technology in agricultural automation—A review. Information Processing in Agriculture, 2020. 7(1): p. 1-19.
[37] Sarker, I.H., Deep cybersecurity: a comprehensive overview from neural network and deep learning perspective. SN Computer Science, 2021. 2(3): p. 1-16.
[38] Chen, L., et al., Review of image classification algorithms based on convolutional neural networks. Remote Sensing, 2021. 13(22): p. 4712.
[39] Naranjo-Torres, J., et al., A review of convolutional neural network applied to fruit image processing. Applied Sciences, 2020. 10(10): p. 3443.
[40] Krizhevsky, A., I. Sutskever, and G.E. Hinton, Imagenet classification with deep convolutional neural networks. Communications of the ACM, 2017. 60(6): p. 84-90.
[41] Chollet, F. Xception: Deep learning with depthwise separable convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[42] Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
[43] He, K., et al., Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence, 2015. 37(9): p. 1904-1916.
[44] He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[45] Loussaief, S. and A. Abdelkrim, Convolutional neural network hyper-parameters optimization based on genetic algorithms. International Journal of Advanced Computer Science and Applications, 2018. 9(10).
[46] Mohakud, R. and R. Dash, Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection. Journal of King Saud University-Computer and Information Sciences, 2022. 34(8): p. 6280-6291.
[47] Wang, Y., H. Zhang, and G. Zhang, cPSO-CNN: An efficient PSO-based algorithm for fine-tuning hyper-parameters of convolutional neural networks. Swarm and Evolutionary Computation, 2019. 49: p. 114-123.
[48] Ahmed, S., et al., Less is more: lighter and faster deep neural architecture for tomato leaf disease classification. IEEE Access, 2022. 10: p. 68868-68884.
[49] Schuler, J.P.S., et al. Color-Aware Two-Branch DCNN for Efficient Plant Disease Classification. in MENDEL. 2022.
[50] Abbas, A., et al., Tomato plant disease detection using transfer learning with C-GAN synthetic images. Computers and Electronics in Agriculture, 2021. 187: p. 106279.
[51] Trivedi, N.K., et al., Early detection and classification of tomato leaf disease using high-performance deep neural network. Sensors, 2021. 21(23): p. 7987.
[52] Alhaj Ali, A., et al. Tomato Leaf Diseases Detection Using Deep Learning. in Proceedings of the Computational Methods in Systems and Software. 2021. Springer.
[53] Bedi, P. and P. Gole, Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network. Artificial Intelligence in Agriculture, 2021. 5: p. 90-101.
[54] Sun, H., et al., MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks. Computers and Electronics in Agriculture, 2021. 189: p. 106379.
[55] Akram, T., M. Sharif, and T. Saba, Fruits diseases classification: exploiting a hierarchical framework for deep features fusion and selection. Multimedia Tools and Applications, 2020. 79(35): p. 25763-25783.
[56] Sanga, S., et al., Mobile-based deep learning models for banana diseases detection. arXiv preprint arXiv:2004.03718, 2020.
[57] Chohan, M., et al., Plant disease detection using deep learning. International Journal of Recent Technology and Engineering, 2020. 9(1): p. 909-914.
[58] Mohameth, F., C. Bingcai, and K.A. Sada, Plant disease detection with deep learning and feature extraction using plant village. Journal of Computer and Communications, 2020. 8(6): p. 10-22.
[59] Singh, D., et al., PlantDoc: a dataset for visual plant disease detection, in Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. 2020. p. 249-253.

[60] Tiwari, D., et al. Potato leaf diseases detection using deep learning. in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). 2020. IEEE.
[61] Khamparia, A., et al., Seasonal crops disease prediction and classification using deep convolutional encoder network. Circuits, Systems, and Signal Processing, 2020. 39(2): p. 818-836.
[62] Agarwal, M., et al., ToLeD: Tomato leaf disease detection using convolution neural network. Procedia Computer Science, 2020. 167: p. 293-301.
[63] Verma, S., A. Chug, and A.P. Singh, Application of convolutional neural networks for evaluation of disease severity in tomato plant. Journal of Discrete Mathematical Sciences and Cryptography, 2020. 23(1): p. 273-282.
[64] Jadhav, S.B., V.R. Udupi, and S.B. Patil, Identification of plant diseases using convolutional neural networks. International Journal of Information Technology, 2021. 13(6): p. 2461-2470.
[65] Chen, J., Q. Liu, and L. Gao, Visual tea leaf disease recognition using a convolutional neural network model. Symmetry, 2019. 11(3): p. 343.
[66] Arsenovic, M., et al., Solving current limitations of deep learning based approaches for plant disease detection. Symmetry, 2019. 11(7): p. 939.
[67] Jiang, P., et al., Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access, 2019. 7: p. 59069-59080.
[68] Singh, U.P., et al., Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. IEEE Access, 2019. 7: p. 43721-43729.
[69] Karthik, R., et al., Attention embedded residual CNN for disease detection in tomato leaves. Applied Soft Computing, 2020. 86: p. 105933.
[70] Elhassouny, A. and F. Smarandache. Smart mobile application to recognize tomato leaf diseases using Convolutional Neural Networks. in 2019 International Conference of Computer Science and Renewable Energies (ICCSRE). 2019. IEEE.
[71] Ashqar, B.A. and S.S. Abu-Naser, Image-based tomato leaves diseases detection using deep learning. 2018.
[72] Costa, J., C. Silva, and B. Ribeiro. Hierarchical deep learning approach for plant disease detection. in Iberian Conference on Pattern Recognition and Image Analysis. 2019. Springer.
[73] Zhang, S., et al., Cucumber leaf disease identification with global pooling dilated convolutional neural network. Computers and Electronics in Agriculture, 2019. 162: p. 422-430.
[74] Ferentinos, K.P., Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture, 2018. 145: p. 311-318.
[75] Liu, B., et al., Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 2017. 10(1): p. 11.
[76] Anar A. Hady, B.M.M.E.-B., Sherine M. Abd El-Kader, Environmental Monitoring for Cultivated Lands: Designing and Implementing IoT based Complete Prototype Model. Journal of Computer Science, 2022. 18.
[77] Kitpo, N. and M. Inoue. Early rice disease detection and position mapping system using drone and IoT architecture. in 2018 12th South East Asian Technical University Consortium (SEATUC). 2018. IEEE.
[78] Li, W., et al., A survey on multisensor fusion and consensus filtering for sensor networks. Discrete Dynamics in Nature and Society, 2015. 2015.
[79] Liao, W., J. Chanussot, and W. Philips, Remote sensing data fusion: Guided filter-based hyperspectral pansharpening and graph-based feature-level fusion, in Mathematical Models for Remote Sensing Image Processing. 2018, Springer. p. 243-275.
[80] Ead, W.M. and M.M. Abbassy. IOT Based on Plant Diseases Detection and Classification. in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). 2021. IEEE.
[81] Mouha, R.A., Internet of Things (IoT). Journal of Data Analysis and Information Processing, 2021. 9(2): p. 77-101.
[82] Jung, J., et al., The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. Current Opinion in Biotechnology, 2021. 70: p. 15-22.
[83] Mishra, M., P. Choudhury, and B. Pati, Modified ride-NN optimizer for the IoT based plant disease detection. Journal of Ambient Intelligence and Humanized Computing, 2021. 12(1): p. 691-703.
[84] Chen, P., et al., Occurrence prediction of cotton pests and diseases by bidirectional long short-term memory networks with climate and atmosphere circulation. Computers and Electronics in Agriculture, 2020. 176: p. 105612.
[85] Nawaz, M.A., et al., Plant disease detection using internet of thing (IoT). International Journal of Advanced Computer Science and Applications, 2020. 11(1).
[86] Chen, W.-L., et al., RiceTalk: Rice blast detection using Internet of Things and artificial intelligence technologies. IEEE Internet of Things Journal, 2019. 7(2): p. 1001-1010.
[87] Devi, R.D., et al. IoT enabled efficient detection and classification of plant diseases for agricultural applications. in 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). 2019. IEEE.
[88] Win, T.T., AI and IoT methods for plant disease detection in Myanmar. Kobe Institute of Computing: Kobe, Tokyo, 2018.
[89] Aher, A., et al., Smart agriculture using clustering and IOT. International Research Journal of Engineering and Technology (IRJET), 2018. 5(03): p. 2395-0056.
[90] Thorat, A., S. Kumari, and N.D. Valakunde. An IoT based smart solution for leaf disease detection. in 2017 international conference on big data, IoT and data science (BID). 2017. IEEE.
[91] Patil, S.S. and S.A. Thorat. Early detection of grapes diseases using machine learning and IoT. in 2016 second international conference on Cognitive Computing and Information Processing (CCIP). 2016. IEEE.

[92] Simonyan, K. and A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
[93] Huang, G., et al. Densely connected convolutional networks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[94] Szegedy, C., et al. Inception-v4, inception-resnet and the impact of residual connections on learning. in Thirty-first AAAI conference on artificial intelligence. 2017.
[95] Howard, A.G., et al., Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017.
[96] Sermanet, P., et al., Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229, 2013.
[97] Zeiler, M.D. and R. Fergus. Visualizing and understanding convolutional networks. in European conference on computer vision. 2014. Springer.
[98] Kamal, K., et al., Depthwise separable convolution architectures for plant disease classification. Computers and Electronics in Agriculture, 2019. 165: p. 104948.
[99] Zhang, X., et al., Identification of maize leaf diseases using improved deep convolutional neural networks. Ieee Access, 2018. 6: p. 30370-30377.
[100] Bhandari, M., et al., Botanicx-ai: Identification of tomato leaf diseases using an explanation-driven deep-learning model. Journal of Imaging, 2023. 9(2): p. 53.
[101] Bhuiyan, M.A.B., et al., BananaSqueezeNet: A very fast, lightweight convolutional neural network for the diagnosis of three prominent banana leaf diseases. Smart Agricultural Technology, 2023. 4: p. 100214.
[102] Sakkarvarthi, G., et al., Detection and classification of tomato crop disease using convolutional neural network. Electronics, 2022. 11(21): p. 3618.
[103] Ulutaş, H. and V. Aslantaş, Design of Efficient Methods for the Detection of Tomato Leaf Disease Utilizing Proposed Ensemble CNN Model. Electronics, 2023. 12(4): p. 827.
[104] Ksibi, A., et al., MobiRes-net: a hybrid deep learning model for detecting and classifying olive leaf diseases. Applied Sciences, 2022. 12(20): p. 10278.
[105] Rashid, J., et al., Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique. Computers, Materials & Continua, 2023. 74(1).
[106] Kaur, P., et al., Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction. Sensors, 2022. 22(2): p. 575.
[107] Liu, W., L. Yu, and J. Luo, A hybrid attention-enhanced DenseNet neural network model based on improved U-Net for rice leaf disease identification. Frontiers in Plant Science, 2022. 13: p. 922809.
[108] Chavan, T.R. and A.V. Nandedkar, AgroAVNET for crops and weeds classification: A step forward in automatic farming. Computers and Electronics in Agriculture, 2018. 154: p. 361-372.
[109] Barbedo, J.G., Factors influencing the use of deep learning for plant disease recognition. Biosystems engineering, 2018. 172: p. 84-91.
[110] Brown, N., et al., The role of passive surveillance and citizen science in plant health. CABI Agriculture and Bioscience, 2020. 1(1): p. 1-16.
[111] Hagos, M.T. and S. Kant, Transfer learning based detection of diabetic retinopathy from small dataset. arXiv preprint arXiv:1905.07203, 2019.
[112] Wang, Y. and Q. Yao, Few-shot learning: A survey. 2019.
[113] Shorten, C. and T.M. Khoshgoftaar, A survey on image data augmentation for deep learning. Journal of big data, 2019. 6(1): p. 1-48.
[114] Goodfellow, I., et al., Generative adversarial networks. Communications of the ACM, 2020. 63(11): p. 139-144.
[115] Pu, Y., et al., Variational autoencoder for deep learning of images, labels and captions. Advances in neural information processing systems, 2016. 29.
[116] Wang, G., Y. Sun, and J. Wang, Automatic image-based plant disease severity estimation using deep learning. Computational intelligence and neuroscience, 2017. 2017.
[117] Barbedo, J.G.A., A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems engineering, 2016. 144: p. 52-60.
[118] Gavrilov, A.D., et al., Preventing model overfitting and underfitting in convolutional neural networks. International Journal of Software Science and Computational Intelligence (IJSSCI), 2018. 10(4): p. 19-28.
[119] Ahmad, M., et al., Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning. IEEE Access, 2021. 9: p. 140565-140580.
[120] Hema, D. and D.S. Kannan, Interactive color image segmentation using HSV color space. Sci. Technol. J, 2019. 7(1): p. 37-41.
[121] Pisner, D.A. and D.M. Schnyer, Support vector machine, in Machine learning. 2020, Elsevier. p. 101-121.