Review

On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey

Houda Orchi, Mohamed Sadik and Mohammed Khaldoun

Review

On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey

Houda Orchi *, Mohamed Sadik and Mohammed Khaldoun

Networking Embedded Systems and Telecommunications (NEST) Research Group, Engineering Research Laboratory (LRI), Department of Electrical Engineering, National Higher School of Electricity and Mechanics (ENSEM), Hassan II University of Casablanca, Casablanca 8118, Morocco; m.sadik@ensem.ac.ma (M.S.); m.khaldoun@ensem.ac.ma (M.K.)
* Correspondence: houda.orchi@ensem.ac.ma

Abstract: The agricultural sector remains a key contributor to the Moroccan economy, representing about 15% of gross domestic product (GDP). Disease attacks are constant threats to agriculture and cause heavy losses in the country's economy. Therefore, early detection can mitigate the severity of diseases and protect crops. However, manual disease identification is both time-consuming and error prone, and requires a thorough knowledge of plant pathogens. Instead, automated methods save both time and effort. This paper presents a contemporary overview of research undertaken over the past decade in the field of disease identification of different crops using machine learning, deep learning, image processing techniques, the Internet of Things, and hyperspectral image analysis. Additionally, a comparative study of several techniques applied to crop disease detection was carried out. Furthermore, this paper discusses the different challenges to be overcome and possible solutions. Then, several suggestions to address these challenges are provided. Finally, this research provides a future perspective that promises to be a highly useful and valuable resource for researchers working in the field of crop disease detection.

Keywords: machine learning; deep learning; image processing; hyperspectral image analysis

1. Introduction

Agriculture is the mainstay of many countries. Due to population growth, the demand for food is steadily increasing. To satisfy this pressing need, it is necessary to increase agricultural productivity and protect crops. Nevertheless, crops are highly prone to different diseases due to a large number of pathogens present in their environment. Some of these disease pathogens are virus organisms, whereas others are fungal or bacterial [1]. Crop diseases can reduce productivity by 10% to 95% [2], resulting in a significant decrease in the quantity and quality of agricultural production. Therefore, early identification of diseases is crucial to avoid huge losses and reduce the excessive use of pesticides, which can harm human health and the environment. In most cases, and especially in developing countries and small farms, farmers identify crop diseases with the naked eye based on visual symptoms. This is a tedious task that requires expertise in plant pathology and excessive treatment time [3]. Moreover, if the field is attacked by a rare disease, farmers seek expert advice to obtain an accurate and efficient diagnosis, which obviously generates additional treatment costs [4]. Thus, this method of visual observation is not practical and feasible for large farms and may even yield erroneous predictions due to biased decisions [5]. The restrictions of the traditional approach have motivated researchers to develop technological proposals for the early identification of crop diseases in an accurate, fast, and reliable manner, and in order to meet the increasing demands of consumers and alleviate the environmental impact of chemical inputs on the environment and health. In this regard, several methods [6–9] have been proposed to automate the process of disease detection. These methods for automatic recognition of crop diseases are divided into two groups, direct and indirect methods [10].

Direct methods comprise molecular [11] and serological techniques [12,13] that provide accurate and direct detection of the pathogens triggering the disease, although these techniques require a significant amount of time for the collection, processing, and analysis of the collected samples. By comparison, optical imaging techniques [14,15] are among the indirect methods that are able to identify diseases and predict the health of the crop through different parameters such as morphological change and transpiration rate. Fluorescence and hyperspectral imaging [16] are some of the most widely used indirect methods for early disease identification. Although hyperspectral images are a valuable source of data and contain more information than ordinary photos [17], hyperspectral devices are very expensive, bulky, and difficult to obtain for low-income farmers. Alternatively, other types of digital cameras are available at reasonable prices in electronics stores. As a result, most of the automatic identification processes considered so far are focused on visible domain images, which enables the use of very accurate and fast algorithms. Hence, this review focuses on various approaches based on image processing techniques and spectroscopy for automatic crop disease detection using numerous approaches and algorithms using deep and machine learning, fuzzy logic, and transfer learning.

The main objectives of this paper are, first, to identify the major issues that have not yet been properly explored in previous studies on the automation of the disease recognition process; and, second, to highlight future directions that may help circumvent these challenges. The upcoming sections are structured in the following order. Section 2 provides an insight into the current state of the art in disease recognition. Then, Section 3 is devoted to the comparative study of the various techniques used, identifying their advantages and drawbacks, followed by Section 4, in which the results are discussed and analyzed. In Section 5, the gaps in the existing literature are addressed. These shortcomings constitute possible avenues to explore in future research, which is addressed in Section 6. Eventually, the conclusion is drawn in Section 7.

2. Background

Manual identification of crop diseases is both fastidious and inaccurate, meaning it is only feasible in small farms [5]. In contrast, automatic disease detection is significantly more accurate and takes less time and labor [18]. As a result, numerous studies [19–22] have been conducted and are discussed in detail below. This section provides a review of different techniques applied in the identification of crop diseases, presents the taxonomy of various crop diseases, and describes the concept of image processing and machine learning. It also demonstrates the application of hyperspectral imagery, the Internet of Things, and deep and transfer learning in the field of disease recognition.

2.1. Taxonomy of Crop Diseases and Their Symptoms

The leaves of crops are highly prone to diseases, which are a natural phenomenon [23]. However, if corrective measures are not taken at the right time to stop the spread of the disease, it leads to a significant reduction in the quality and quantity of agricultural products [24]. Crops are affected by various pathogens [1] such as viruses, bacteria, fungi, and deficiencies. Thus, the pathogens responsible for the disease are classified into two categories [25]: autotrophs, which thrive on living tissue, or saprophytes, which dwell on dead tissue. The symptoms of the disease adversely affect the development and growth of crops and are easily visible. Leaf discoloration is the first symptom of disease in plants. In addition, the shape and texture of the leaves are highly useful in detecting various diseases. Therefore, different diseases, such as mildew, rust, and powdery mildew, can be detected by processing images of the leaves [26,27].

The following is a brief description of the three common types of plant diseases [28] that are illustrated in Figure 1 and described in Table 1:

- *Virus diseases* [1]: Among all plant diseases, those caused by infection are difficult to identify and diagnose; moreover, these symptoms are mistaken for signs of nutritional deficiency or injury, as there is no preconceived indicator that can be constantly

monitored. Whiteflies, leafhoppers, aphids, and cucumber crawling insects are regular carriers of virus diseases.

- *Fungal diseases* [1]: Foliar diseases are caused by a fungus, such as downy mildew, anthracnose, and powdery mildew. It initially appears on old lower leaves, which have gray-green spots or are soaked in water. As the parasite matures, these spots darken and cause fungus to grow on them.
- *Bacterial diseases* [1]: Pathogens cause serious diseases in vegetables. They do not directly enter the vegetation, but rather through injuries or apertures in the crop. Crop injuries result from various pathogens, insects, and agricultural implements during tasks such as picking and pruning.
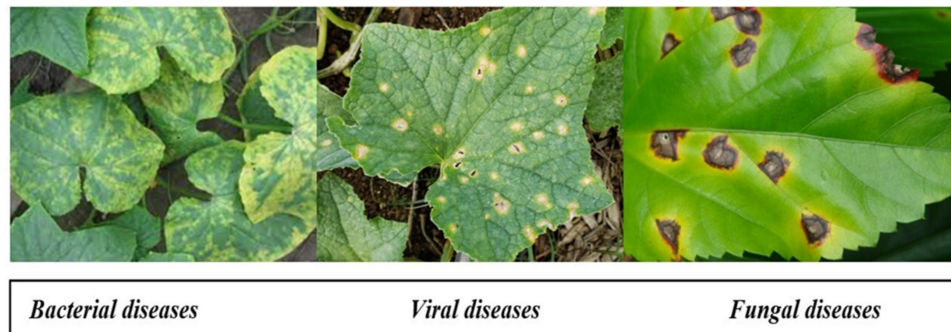

Figure 1. Different types of pathogens: viruses, fungi, and bacteria.

Table 1. Classification of some leaf diseases with their symptoms.

Plant Leaf	Diseases	Symptoms	Pathogen Category
Rice	- Brown spot/*Bipolaris oryzae* - Blast leaf/*Pyricularia oryzae Cavara*	- Whitish-gray center - An irregular dark brown	- Fungi
Cotton	- Faliar leaf/*Stemphylium solani* - Areolate mildew/*Cercospora* - Leaf spot/*Alternaria spot* - Bacterial blight/*Xanthomonas campestris*	- Spot of light-yellow color with dark brown margins - Tanned brown spot - Circular dark brown leaf spots to black - Halo yellowish green	- Fungi, bacterial, virus - Fungi - Fungi - Bacterial
Citrus	- Melanose/*Diaporthe citri* - Greasy spot/*Amycosphaerella africana* - Canker/*Xanthomonas citri* subsp	- The leaf becomes rough to the touch - Blister yellowish-brown - Includes flattened, swollen, cracked, round to irregular sunken	- Fungi - Fungi - Fungi, bacterial
Tomato	- Early blight/*Alternaria tomatophila* - Late blight/*Phytophthora infestans* - Powdery mildew/*Leveillula taurica* - Yellow curl/*tomato infectious chlorosis virus*	- Dark ring spot around it yellow - The dark spot is growing rapidly - Curly and yellowish leaf - Soaked in the water ringed by a yellow halo	- Fungi - Fungi - Fungi - Virus
Maize	- Stalk rot/*Erwinia carotovora*	- Yellowing of dull green leaves and the lower parts of the stem	- Fungi
Wheat	- Rust/*Puccinia triticina Erikss.* - Powdery mildew/*Blumeria graminis* - Bacterial blight/*Pseudomonas syringae*	- Pale leaves spots - While gray or brown spot - Halo yellowish green	- Fungi - Fungi - Bacterial
Watermelon	- Anthracnose/*Colletotrichum obiculare* - Downey mildew/*Pseudoperonospora cubensis*	- Irregular yellow patches - Yellow to white spots	- Fungi

2.2. Application of Machine Learning and Image Processing in Disease Identification

Foliar images are an excellent and rich source of data on plant pathology and morphological behavior; thus, these data must be thoroughly extracted and analyzed. Image processing plays [28] a crucial role in the diagnosis and analysis of leaf diseases. The procedure adopted in this leaf disease identification process is illustrated in Figure 2, showing

an insight into the different techniques employed by the authors to detect the disease by means of image processing and artificial intelligence.
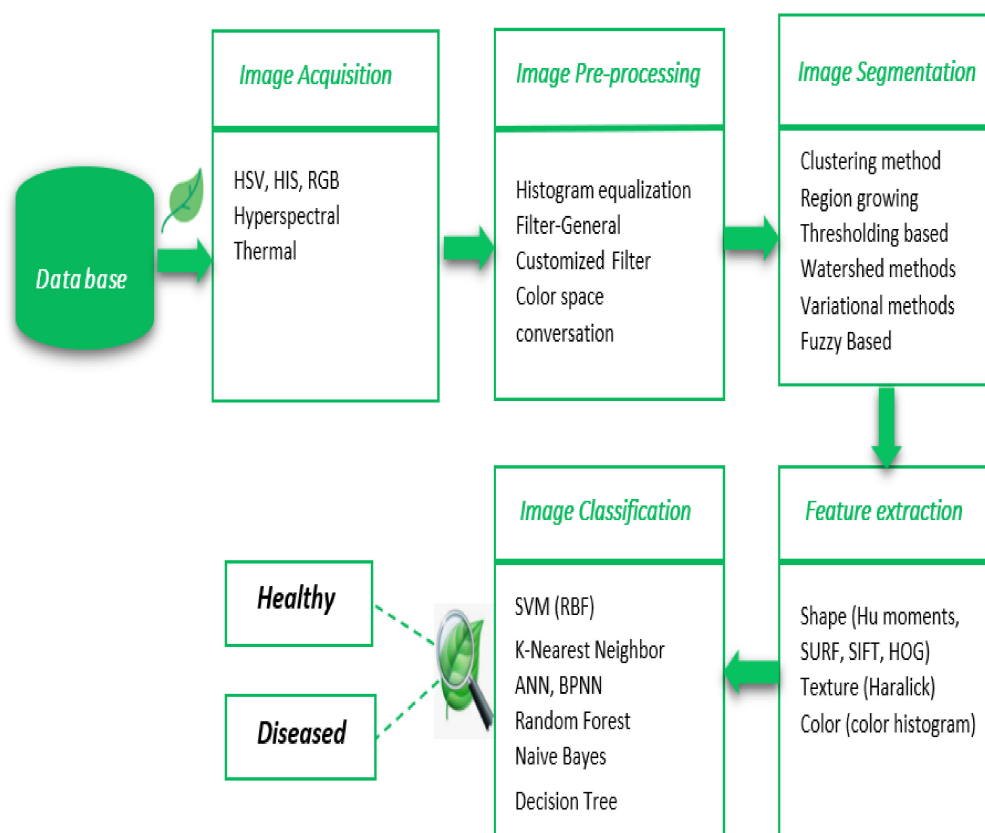

Figure 2. Different approaches for the identification of leaf diseases.

The primary step [29] in identifying diseases is the acquisition of images. In most cases, images can be fetched either from a digital camera or an imaging system. As raw images tend to contain noise, removing these impurities is required. As a result, the second step is known as image pre-processing, and involves the removal of unwanted distortions, in addition to contrast enhancement, to clarify and brighten the image features. For example, a Gaussian function that creates soft blur is commonly used to lessen the noise in the image. Subsequently, image segmentation [30] is the third step in which the image is segmented from its background, whereas the region of interest (ROI) is partitioned to emphasize the prominent features. The fourth step is feature extraction [31], which unveils the information and details of an image. As a side note, the leaf features usually include shape, texture, and color, which are used to diagnose the crop. Thus, these chosen features form an input feature vector which is then fed into the classifier. Using this vector, it is possible to discriminate one class of objects from another. The final step is classification [32]. Note that the choice of a suitable classifier depends on the specific problem. The classifier's aim is to recognize the images by sorting them into several predefined classes based on the resulting feature vector obtained in the fourth step. For this purpose, the classification task contains two phases, namely, training and testing. The training operation trains the classifier on a training dataset; thus, the greater the number of training sets, the better the accuracy obtained. It should be noted that the result, which is the crop's healthy state or diseased state associated with the species name, must be achieved as swiftly as possible.

2.3. Application of Deep and Transfer Learning in Disease Recognition

Over the past decade, deep learning [33–35] and transfer learning [19,36] applications in agriculture have gained widespread success and yielded highly promising outcomes due to their capability to reliably learn and discern visual features. Numerous intriguing stud-

ies [33,35,37–40] have been published on the employment of these promising approaches for identifying diseases. Of particular note, the use of transfer learning is a trend that is becoming increasingly popular and is widely used by researchers [41,42]. Furthermore, transfer learning is not a sole technique, but rather a set of fine-tuned techniques, which enables the development of highly accurate models on a more restrictive specialized dataset, such as those for plant diseases. Mohanty et al. [43] showed that the fine-tuning approach is far better than a CNN model that is trained from scratch. Another model is the Neural Network (NN), which is broadly employed and recommended to analyze hyperspectral data for the premature detection of diseases. Its basic mechanism was inspired by the human nervous system, and it possesses specific capabilities such as learning and generalization that aid in crop disease diagnosis. In contrast to other machine learning methods, it has a more accurate diagnostic capability because it is better able to combine training sets. Another similar comparative study was undertaken by Zhu et al. [44], wherein back-propagation neural networks (BPNNs) were tested with the support vector machine (SVM), random forest (RF), latent Dirichlet allocation (LDA), extreme learning machine (ELM), LS-SVM, and partial least squares discrimination analysis (PLS-DA) for pre-symptomatic detection and classification of tobacco mosaic virus (TMV) disease with the use of hyperspectral imaging. Similarly, Zhu et al. [45] studied the feasibility of hyperspectral imaging as a non-invasive technique for early detection of TMV disease with machine learning classifiers and the variable selection technique. The results revealed that the back-propagation neural network model (BPNN) achieved 95% accuracy, whereas the chemometric models achieved an accuracy of 80%. It is worth mentioning that it is possible to implement pattern identification methods such as the random forest and support vector machine by utilizing a new pattern recognition technique, named the Artificial Intelligent Nose. Cui et al. [46] provided a review of different invasive and non-invasive techniques, including their advantages and drawbacks, in which the authors noted that the smart nose is a non-invasive and fast method for plant disease diagnosis. In essence, neural networks ensure the highest quality and unaltered spectral information for hyperspectral data analysis. The most well-known study of the ANN spawned the concept of deep learning, which has recently become popular in farming applications. Deep learning has received increasing and widespread interest from many researchers, particularly since 2018, as shown in Table 2. Researchers have made remarkable progress in crop image classification; some of the most typical and representative models are the convolutional neural network (CNN), auto-encoder (AE) recurrent neural network (RNN), and restricted Boltzmann machine (RBM). Many fascinating studies have been published on deep learning for crop disease classification and detection. Among these works, that of Ma et al. [47] presented a deep convolutional neural network (DCNN) model able to detect over four types of cucumber disease. In a comparison with other traditional methods, such as the support vector machine, naive Bayes, and AlexNet, the DCNN was capable of identifying the different cucumber diseases with very high accuracy of up to 93.41%. Similarly, Tran et al. [48] offered a monitoring system for tomato growth and to maximize tomato yield. This system was able to classify nutritional deficiencies and diseases during growth. Thus, agricultural experts evaluate the symptoms based on the results to protect tomato crops. In a similar manner, to effectively monitor apple tree growth at each stage and estimate the yield, Tian et al. [49] deployed a dense YOLOV3 model that utilizes techniques for data augmentation to prevent overfitting. Their approach was found to be valid and applicable to apple orchards, although their study included wavy lights, interlaced fruit, and complex backgrounds.

Table 2. A brief summary of different research on transfer and deep learning since 2018 for identifying crop diseases.

| Year | Authors | Model | Dataset | | | | Accuracy |
			Crop Name	Nb of Classes	Nb of Images	Name of Dataset	
2020	Singh et al. [37]	MobileNet, R-CNN	13 types	27	2598	PlantVillage	70.53%
2020	Al-bayati et al. [38]	DNN, SURF, GOA	Apple	6	2539	PlantVillage	98.28%
2019	Arsenovic et al. [39]	CNN-Multichannel	12 species of crops	42	79,265	PlantVillage	93.67%
2019	Costa et al. [40]	InceptionV3 and CNN using a Hierarchical Approach	Apple, Tomato, peach	16	24,000	PlantVillage	97.74%
2019	Geetharamani et al. [35]	9-layer deep CNN	14 species of crops	39	61,486	Leaf disease dataset	96.46%
2018	De Luna et al. [50]	CNN, Faster R-CNN	Tomato	4	4923	Own	91.67%
2018	Ferentinos et al. [33]	Overfeat, VGG16, AlexNet	25 species of crops	58	87,848	Open Dataset	99.53%

2.4. Hyperspectral Imaging Applied to Disease Recognition

The hyperspectral imagery method has been strongly developed during the past two decades [51], and used to identify abiotic and biotic stresses in cultivated plants [52]. Hyperspectral imaging is a technique combining spectroscopy and imagery, making it possible to simultaneously obtain the spatial and spectral information of an object. Disease infection causes changes in plant biochemical and biophysical properties, such as transpiration rate, tissue structure, water, and pigment content. These changes can then alter plant spectral properties, intercellular space, and water content [53]. However, the hyperspectral system is able to capture these spectral features. Zhu et al. [44,45] conducted a similar study to detect TSWV infection growth in tobacco, in which the authors reported that hyperspectral reflectance was gathered in the visible and near-infrared range to distinguish between healthy and infected TSWV tobacco leaves using statistical analysis methods. Primarily, the TSWV presence was identified at 14 DPI. Moreover, Zhu et al. [45] demonstrated that hyperspectral imaging is able to detect the tobacco mosaic virus (TMV) infection before showing any symptoms, while utilizing SPA for the selection of the effective EW wavelength and, most significantly, for identifying various diseases. Due to the huge number of spectral values that are highly correlated in the hyperspectral dataset, high dimensionality and multi-collinearity frequently appear in hyperspectral data [54,55]. Accordingly, the selection of EWs is crucial for hyperspectral analysis in order to lessen the computational complexity, increase the efficiency of using hyperspectral data, and reduce the computational complexity. Thus, to address this multi-collinearity issue, a variety of approaches and methods have been presented, such as the successive projection algorithm (SPA) [45,56], partial least squares regression (PLSR) models [54] and genetic algorithms (GAs).

2.5. Application of IoT in the Field of Leaf Disease Recognition

The Internet of Things (IoT) has improved agricultural capabilities. IoT applications can help farmers at any time during their farming activities and keep them updated with the latest crop and weather information to remotely monitor their fields. By means of IoT applications [57–59], farmers can make plans for the next season's harvest. Furthermore, they can detect crop diseases at an early stage to curb the spread of disease and save their yield. Agricultural IoT apps clearly play a major role in increasing agricultural production and decreasing crop losses due to diseases. In this context, a large amount of research has been conducted to identify diseases, as shown in Table 3. Truong et al. [60] devised an IoT-based system made up of various devices that is able to deliver real-time environmental information and send it to the cloud to be stored. These environmental data are processed and scrutinized to predict weather conditions by means of the SVM algorithm deployed in the cloud in order to detect crop fungal diseases. In addition, better results have been achieved when the Internet of Things and image processing have been combined and implemented in the area of disease recognition. Krishna et al. [58] implemented an IoT

system featuring SMS alerts that enables automatic disease detection and pesticide spraying using the NodeMCU.

Table 3. Summary of the literature survey on Internet of Things systems.

Researchers	Detection Techniques and Algorithms	Parameter Evaluation
M. Mishra et al. [39] 2021	An IoT-based automated plant disease monitoring and detection system, using the median filter and a modified optimizer called the SCA-based RideNN Cycling Neural Network	The RideNN model based on SCA achieved accuracy of 91.56%
Devi et al. [37] 2019	IoT system using GLCM, RFC, and k-means clustering	The overall accuracy of disease detection and classification based on RFC-GLCM was almost 99.99%
Krishna et al. [38] 2019	IoT system using SVM and k-means clustering	An immediate SMS alert to the farmer
Chen et al. [36] 2019	RiceTalk platform using an AI model and IoT devices	Net prediction accuracy was 89.4%
Win et al. [41] 2018	IoT-based remote rice monitoring IoT system using deep learning and transfer learning	Real-time monitoring of environmental parameters
Truong et al. [40] 2017	IoT system using SVM	Real time analysis

3. Comparison of Various Crop Disease Detection Techniques

The primary objective of this section is to provide an overview of research carried out during the past decade for identifying crop diseases. Table 4 provides an outline of several methodologies adopted by researchers in the field of crop disease using machine learning, image processing, the Internet of Things, transfer learning, and deep learning techniques. It also indicates the limitations and gaps that need to be filled to help develop an automatic, efficient, accurate, and faster system in the future. Thus, in the conducted research, the authors conclude that deep learning provides accurate and highly promising results compared to other classification and detection methods. Additionally, the use of preprocessing techniques significantly improves segmentation accuracy. The k-means algorithm is the most widely and commonly used technique [29,58,61–63] for segmenting diseased leaves and classifying crop diseases. In practice, no generalizable algorithm is able to solve all issues, so choosing a suitable learning algorithm for a specific problem is a crucial step for the model efficiency. Note that the extracted texture features are the most relevant and most useful for representing the disease-affected regions in the images, which are then employed to train the support vector machine (SVM) and neural network (NN) classifier. It is further emphasized that these texture features are arithmetical parameters that are automatically calculated by means of the gray level co-occurrence matrix (GLCM) [64,65], as stated below:

1. *ASM:* The second angular momentum that stands for the total sum of squares in the GLCM.

$$Energy = \sum_{i,j=0}^{N-1} (P_{i,j})^2 \tag{1}$$

2. *Contrast:* Denotes the sum of the difference in local intensity, where $i \neq j$.

$$Contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{2}$$

3. *Entropy:* The quantity of image information necessary for the compression.

$$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{i,j})P_{i,j} \tag{3}$$

4. *Correlation:* Refers to the linear dependence of the adjacent pixels' gray levels.

$$Correlation = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu)(j-\mu)/\sigma^2 \tag{4}$$

where $P_{i,j}$ is the i, j component of the GLCM normalized symmetric matrix and N denotes the number of gray levels. σ^2 is the intensity variation of all pixels as given below:

$$\sigma^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu)^2 \tag{5}$$

Moreover, the average of GLCM is μ given by:

$$\mu = \sum_{i,j=0}^{N-1} iP_{i,j} \tag{6}$$

5. *Homogeneity feature:* Represents the homogeneity of the voxel pairs of the gray level, and is equal to 1 for a diagonal GLCM.

$$Homogeneity = \sum_{i,j=0}^{N-1} P_{i,j}/1 + (i-j)^2 \tag{7}$$

Ultimately, the models for identifying and classifying crop diseases were evaluated by means of various metrics, which were specific to the model used in each study, such as sensitivity, precision (P), recall *(R)*, quality measure *(QM)*, and F1-score. The statistical evaluation measures used to analyze the quantitative performance of crop disease detection models with deep and transfer learning can be calculated as follows:

$$Precision = \frac{TP}{(TP + FP)} \tag{8}$$

where *Precision (P)* is the fraction of true positives (*TP*) to the total amount of relevant results, that is, the sum of *TP* and false positives (*FP*). For multi-class classification problems, P is averaged across the classes.

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{9}$$

Sensitivity/Recall (R) is the fraction of *TP* to the total amount of *TP* and false negatives (*FN*). For multi-class classification problems, R obtains the average of all classes.

$$Specificity = \frac{TN}{(TN + FP)} \tag{10}$$

Specificity is the proportion of true negative (*TN*) samples to all healthy samples (true negatives and false positives). This measure is utilized to evaluate the performance of a proposed model in forecasting true negatives.

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{11}$$

Accuracy is the proportion of correctly classified samples to the total number of classified samples. This measure is employed to assess the overall performance of a suggested model.

$$F1\_score = \frac{2 \times (Sensitivity \times Precision)}{(Sensitivity + Precision)} \tag{12}$$

F1-score is the harmonic average of both precision and recall. For multi-class classification problems, *F1* is averaged across all classes, where:

TP: represents the number of true positive image samples that are perfectly identified as infected.
FP: is the number of false-positive image samples that are incorrectly classified as infected.
TN: is the number of true-negative image samples that are correctly classified as healthy.
FN: is the number of false-negative image samples that are incorrectly identified as uninfected.

As revealed in the previous studies [66], it was found that accuracy was the most widely adopted metric; it was widely used in 72% of the articles reviewed, followed by the confusion matrix, then precision, recall, and the F1 measures. Some research has examined the root mean square error (RMSE), mean absolute error, R-squared, and mean squared error (MSE), among others. The authors note that it is difficult, if not impossible, to make a comparison across papers because different metrics were used for different tasks, different models, datasets, and parameters; in addition, different crops and diseases were analyzed under different conditions. Furthermore, it is very important to examine whether the researchers tested their implementation on an identical dataset (e.g., by splitting the dataset into training and validation sets), or whether they used different datasets to test their solution.

Table 4. Related work in the area of crop disease identification.

Title	Methodology	Advantages	Disadvantages
Tomato plant disease detection using transfer learning with C-GAN synthetic images (Abbas et al., 2021) [19]	This research paper provides a DL-based method for detecting tomato disease that uses the C-GAN to generate the synthetic images of tomato leaves for data augmentation purposes. Then, a pre-trained DenseNet121 model is fine-tuned on synthetic and real images to classify tomato leaves images into ten disease categories.	- The proposed method reached an accuracy of up to 99.51%, 98.65% and 97.11% in classifying tomato leaf images to the categories of 5, 7 and 10, respectively. It is shown that this method outperforms the current methodologies. - The C-GAN prevents overfitting and enhances the network generalization	- The tomato disease detection was executed on the leaves, but other plant areas such as stems, and branches must be highly involved. - The study was only done on tomato plant disease.
MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks (Sun et al., 2021) [21]	- This paper proposes a lightweight CNN detection model suitable for mobile device deployment, namely MEAN-SSD, to detect apple leaf diseases in real time. Data annotation and augmentation techniques were used to generate 26,767 disease spot images for training by collecting 2230 original images with simple backgrounds from the laboratory and complex backgrounds images collected from the orchard. - The model is capable of automatically extracting the features of five common disease spots from apple leaves.	- The results showed that the MEAN-SSD model is able to detect apple diseases accurately by reaching 83.12 mAP and a speed of 12.53 FPS. - The MEAN block is used as a basic module to boost the detection speed and shrink the model's size.	- Disease detection was only devoted to apple leaf diseases and more specifically to 5 types of disease spots, such as Brown spot, grey spot, Mosaic, Alternaria blotch, and Rust.
Detection of oil palm leaf disease based on color histogram and supervised classifier (Hamdani et al., 2021) [20]	- A new method for detecting oil palm leaf disease is proposed in this paper to discriminate between two leaf classes: healthy and diseased. Then, feature extraction is carried out in the RGB (R, G, and B), LAB (a and b), HSI (H and S), and HSV (H and S) color spaces by splitting the histogram of the 8-bin color channel. This is further performed on the segmented leaf regions resulting from the k-means clustering. A total of 41 selected features are produced using PCA and subsequently fed into the ANN classifier.	- The classification results have shown that the proposed method performs satisfactorily, as evidenced by the high specificity, sensitivity, and accuracy values, which reach 100%, 99.3%, and 99.67% respectively. - The applied method produces a smaller number of features with discriminatory and more powerful characteristics.	- According to the classification outcomes, an error occurred due to a leaf being misclassified as a healthy one. - The study was only conducted on the oil palm leaf disease.
Detection of Rice Leaf Diseases Using Image Processing (Pothen et al., 2020) [22]	The proposed system identifies three diseases (bacterial leaf blight, leaf smut, and brown spot) that affect rice plant leaves using IP and ML techniques. This system, in turn, helps farmers save their crops at an early stage. As a first step in the process sequence, the images are collected and further pre-processed to ensure that the image features are upgraded and undesired distortions are eliminated, followed by segmenting the images through the Otsu thresholding algorithm. Using the segmented area, a range of features are extracted using the LBP and HOG. Then, these obtained features are classified using the SVM and reached 94.6%.	- SVM + HOG with polynomial kernel function can be used to detect other plant diseases. - The proposed work is relevant, offering better precision (of 94.6%) compared to other work.	- One downside of the LBP fundamental operator is its inability to capture certain prevailing features.
Image Processing Technologies for Automatic Detection of Plant Disease and Alerting System in Agricultural Farms (Mugithe et al., 2020) [7]	The authors developed a system able to detect leaf diseases and alert the farmer in case of the need to promptly act to circumvent the spread of the disease in the field. They employed IP techniques that entail six steps. Firstly, the leaf images were taken in real time via a webcam connected to RaspberryPi. Secondly, the images were pre-processed, segmented, and clustered using the k-means clustering algorithm, then features (e.g., perimeter and light intensity) were retrieved from the images. Lastly, these extracted features were evaluated to classify the leaf diseases. Once the disease is detected, the buzzer rings and an alert is generated so that the farmer can intervene promptly. Note that this system works in two ways, namely in the GUI and in real time.	- The results of the disease Alternaria Alternata achieved precision of 95.16313% in the graphical interface. - The use of a warning system when detecting a disease.	- The authors did not address any other disease and not even the reported results are obvious.

Table 4. *Cont.*

Title	Methodology	Advantages	Disadvantages
Plant Disease Detection Using Internet of Thing (IoT) (Usman et al., 2020) [67]	This paper presents the innovation of the IoT in agricultural infection and insect pest control. Data on insects and diseases are collected using a WSN, so an IoT-based control framework was proposed to obtain horticultural data from a farm with levels of trees and three frameworks. First, a computation framework was deployed to deem whether the plant is healthy or affected and second, an automated framework determined the disease closeness in the plants, and a mechanized framework was then set up to recognize the diseases through humidity, temperature, and shade sensors. Thus, through its sensors, the plants' progress is registered and then dissected using Arduino programming. Then, this collected information is transmitted to the cloud by WIFI to be processed and analyzed. Eventually, this information is compared to the whole data to determine if the studied plant is healthy or affected.	- The proposed IoT-based model is low-cost. - Low-income farmers can purchase it and take benefit of it for curtailing the disease spread.	- The authors restricted their model to only the three following parameters: temperature, moisture, and leaf shade. Another constraint is that the evaluated features for the considered parameters are not accurate. Moreover, a range of features was taken from these parameters that may fluctuate unexpectedly depending on the environmental conditions. - The leaves of diseased plants are not classified, so the disease types are still not able to be known.
Detection of Plant Leaf Disease using Digital Image Processing (Mojjada et al., 2020) [29]	This article focuses on the early identification of plant leaf diseases by image analysis. Thus, automated disease detection reduces the work for monitoring agricultural sites. The identification of diseases is carried out through various IP techniques and ML; in particular, a genetic optimization algorithm was used after image segmentation by k-means to obtain optimized results and they also exploited SVM for disease classification.	- The algorithm used was tested with an accuracy of 75% in five classes of infected leaf images identified for corn, tomato, bell pepper, peach, and grape.	- Very common methods were used in the paper.
Precision Method for Pest Detection in Plants using the Clustering Algorithm in Image Processing (Reddy et al., 2020) [68]	This work presents an accurate method for detecting pests in plants using the k-means clustering algorithm. Disease recognition involves steps such as image acquisition, image preprocessing, segmentation, and classification. First, the RGB leaf images were converted to HSV for partitioning, and then the median filter and boundary detection algorithm were both applied during the pre-processing step to suppress the clamor. Finally, the k-means clustering was employed to cluster the images.	- This paper provides an efficient and accurate framework for the detection of affected images. - The k-means clustering provides high accuracies compared to other methods and takes less time for the processing.	- Very typical and old techniques were considered in this paper.
Detection and Classification of Plant Diseases Using Image Processing and Multiclass Support Vector Machine (Khan et al., 2020) [69]	The authors described a framework for plant disease using ML and IP techniques. First, the suggested algorithm is applied to a 148-image dataset which contains 5 types of leaves diseases, namely, Alternaria, fire blight, Anthracnose, and Cercospora leaf spot, and the plant images were split up into two sets. A training set is composed of 73 images and a testing set is composed of 75 images. Then, image segmentation is performed to isolate the pathogenic parts of the leaf. Then, 13 texture features were extracted from the image, of which, nine features (standard deviation, variance, mean, entropy, smoothness, skewness, RMS root mean square, inverse difference, and kurtosis) are calculated using the assigned segment in RGB color space. The other four features (homogeneity, contrast, energy, and correlation) are determined from a grayscale image. Finally, the healthy and diseased leaves are classified based on the feature vector extracted using SVM.	- The obtaining results for plant disease detection showed that the proposed method yields a highly accurate rate of up to 92.8571%.	- The intervention of the operator is crucial to select the segment affected by the disease because this operation is not automatically executed in the system rather, it is performed by visual examination of the three segments.
Evolutionary feature optimization for plant leaf disease detection by deep neural networks (Al-bayati et al., 2020) [38]	The researchers used a DNN for apple leaf disease identification, namely black rot, apple scab, and cedar rust, by applying the GOA and Robust Accelerated Feature SURF, where GOA was employed for feature optimization and SURF was applied for feature extraction. Prior to the implementation of DNN, many steps were performed, such as the image improvement in the pre-processing process and the ROI segmentation. Then, the features were extracted using the SURF descriptor, followed by the optimization by the GOA algorithm and, finally, the disease classification was carried out by the use of DNN.	- The experiments showed that the method based on DNN optimized by SURF provides a higher mean value of 98.28% in comparison with the other techniques; hence the accuracy of the model increases by 18.03%. Thus, the basic model has better transferability compared to the metric model.	- Only foliar diseases of apples were addressed in this paper.
Leaf Disease Detection using Image Processing (Karthikeyan et al., 2020) [65]	This research used IP techniques along with the SVM classifier to detect plant diseases. The identification of plant diseases requires the steps of transforming an RGB image to grayscale, then enhancing the image using the adaptive color histogram AHE, extracting 13 textural features using the GLCM and, finally, using SVM to classify the different types of diseases. Note that more than 500 images were taken for training and testing with intensity values ranging from 0 to 255.	- The system reveals the presence of disease in the leaves in a shorter time and at a lower cost than conventional systems.	- The accuracy rate is not given; also, the structure of the algorithm is complicated.
Plant disease detection using image processing techniques (Sawant et al., 2020) [70]	An IM and DL techniques-based approach was proposed for plant disease identification. First, pictures of healthy and unhealthy leaves are acquired and then stored in the database for preprocessing. Additionally, the images are pre-processed using different techniques such as histogram equalizer, Contrast Limited AHE, and image resizing. Then, once the RGB image is converted to CLAHE, it is resized to 70 × 70 for better resolution. Moreover, the leaf features are extracted and released to the CNN using the SoftMax function for plant disease classification. Note that the first layer has 1000 neurons.	- The CNN method enables accurate detection and classifies diverse plant diseases using IP techniques.	- Results were not provided in this paper.

Table 4. *Cont.*

Title	Methodology	Advantages	Disadvantages
Convolutional neural network for automatic identification of plant diseases with limited data (Afifi et al., 2020) [6]	Several approaches were developed in this study to identify plant diseases with little data. A DAML and triplet network approach was set up using three architectures of CNN (ResNet50, 34, and 18). Using a large dataset, the approaches were trained and then fitted from 5 to 50 images per disease for detecting new diseases.	- The model reached an accuracy of up to 99% when the change from the source domain to the targets was slight, but when the change was significant, the accuracy was up to 81%.	- The results showed a significant error rate for DAML methods of 22.2 per 50 shots and 42.6 per 5 shots compared to the other methods. - The basic model lags significantly behind the other methods.
Leaf disease detection using machine learning (Fulari et al., 2020) [71]	An efficient method for the identification of a healthy or infected leaf was presented using IP and ML techniques. The data were taken from the Kaggle website which contains more than 12,949 images. The method implies different steps such as image preprocessing, segmentation of the image, feature extraction (shape, color, and texture) with the use of GLCM, and classification with the use of SVM.	- The SVM offers a number of advantages over other classifiers, as it is efficient in high-dimensional spaces. - The SVM yielded an accuracy of up to 80% while the CNN provided an accuracy of up to 97.71%.	- The presented method, based on the CNN, provides good accuracy. Nevertheless, it is tedious and a lot of time is required to train the model.
Deep transfer learning models for tomato disease detection (Ouhami et al., 2020) [72]	The authors conducted a study to identify the most suitable DL model for identifying tomato diseases based on RGB leaf images, which were split into 6 different kinds of infections and parasitic attacks. Thus, two architectures of DL models, namely DensNet121 with 161 layers and VGG16, were used to perform the study.	- The results obtained were very promising with an accuracy of up to 95.65%, 94.93% and 90.5% respectively for the DensNet161, DensNet121, and VGG16 models, which shows that DensNet161 with 20 training periods surpassed the other two architectures. - DensNet models require significantly fewer parameters and calculations to achieve optimum performance.	- Adverse transfer can occur and may dramatically lower the model's accuracy.
A new segmentation method for plant disease diagnosis (Gurrala et al., 2019) [64]	IP and AI methods were used for the recognition of diseases such as scab, anthracnose, blight, and spots on plants. First, the RGB image was converted to grayscale, and the image was then segmented using the k-means and the modified CPDA algorithm. Thus, from the result obtained, a comparison was made between these two segmentation algorithms. The statistical parameters of the segmented image were calculated using the GLCM method; that is, the characteristics (entropy, mean, variance, type, RMS, contrast, correlation, energy, homogeneity, regularity, kurtosis, asymmetry, IDM moment of difference). Finally, the SVM classifier was trained with a dataset of about 100 images of leaves affected by the disease.	- The proposed modified CPDA detection algorithm yields more accurate results than the k-means clustering.	- Even if a disease is detected in a shorter time, the accuracy is limited.
IoT Enabled efficient Detection and Classification of Plant Diseases for Agricultural Applications (Devi et al., 2019) [61]	The authors proposed a simple and efficient IoT-based solution for the detection of bunch top and Sigatoka diseases in banana tree located on hills. First, 80 hill banana plant images are captured and then resized to 256 × 256 in the image preprocessing phase. Then, the preprocessed image is converted to a gray image. Thus, the histogram equalization technique is used to equalize the histogram of the resized gray image so that the intensities of the image are better distributed for better segmentation, which is performed by k-means clustering. From the segmented image, the GLCM features are extracted and uploaded to the cloud for further analysis. From these extracted characteristics, the hill banana diseases are classified using the RFC technique. Finally, the data is collected and analyzed by agricultural experts. In addition, this system allows remote monitoring of environmental parameters such as soil humidity and temperature to prevent diseases caused by climate change and pathogens as much as possible.	- The performance results showed an overall detection accuracy of 99.99% and demonstrate that RFC-GLCM-based leaf disease classification works best for the hill banana dataset. - Agricultural experts provide solutions to farmers in case of plant disease or massive changes in environmental parameters on the agricultural field.	- The accuracy of the system depends on the ambient conditions of the agricultural field, such as the angle of image capture and the lighting in the field.
A Preprocessing Approach for Accurate Identification of Plant Diseases in Leaves (Deepa et al., 2018) [73]	In this paper, the proposed method takes RGB images as the input and applies the preprocessing methods, such as image sharpness and median filters, to eliminate the noise from images, and for deblurring and edge detection. Then, these images were segmented using k-means clustering. It should be mentioned that they used the peak signal-to-noise ratio in order to measure the quality of the images.	- The results show that the method adopted allows better identification of plant leaf diseases.	- They studied only three diseases: Alternaria Alternata disease, Bacterial Blight, and Anthracnose disease. - They did not consider extracting characteristics or classifying diseases into different groups.
AI and IoT methods for plant disease detection in Myanmar (Win et al., 2018) [74]	The researchers developed two prototypes. The first is a mobile application that classifies diseases on rice plants. With this simple application, the farmer can easily identify the diseases or pests on rice plants, without using agronomists. The second is a system for monitoring temperature, atmospheric pressure, water level, and the sunlight level of rice fields. Using this system, the intensive work is reduced by remotely monitoring this environmental data from anywhere an Internet connection is available. For the development of the mobile app, they collected 6 kinds of rice pictures, of bacterial leaf blight, brown spot, rice blast, mice attack, insects, and healthy rice. In addition, they used the Arduino nano to turn off/on the solenoid valve, which was directly connected to the Raspberry Pi every 30 min, to ensure efficient power and a long system runtime. Diseases were classified using TL and DL models.	- They developed a simple Android app to monitor the temperature and battery level on a SensorTag. Thus, the farm sensor data reading can be easily viewed on smartphones or PCs.	- They installed 8 SensorTags in different areas. The distance between them and the Raspberry Pi was less than 50 m. To monitor a wide range of many rice fields, another type of communication system, rather than the Bluetooth technology of the SensorTags, must be considered. A problem was encountered with the SensorTag coil battery, Under normal conditions, the batteries last at least a year while keeping the SensorTag alive. However, firmware level changes were added to the SensorTag to advertise all the time so, the LEDs on the SensorTags blink and the batteries only last a few weeks. Hence, they needed to replace the button batteries many times.

Table 4. *Cont.*

Title	Methodology	Advantages	Disadvantages
Plant diseases recognition based on image processing technology (Sun et al., 2018) [75]	A multiple linear regression-based plant disease identification system is presented, along with a histogram-based segmentation method for accurate and automatic threshold calculation. The proposed system is based on IP techniques including four steps: preprocessing, image segmentation, feature extraction, and regression model. First, the spatial domain image denoising is used to filter the noise. Then, an improved histogram-based segmentation method is developed to distinguish lesions from normal foliage, which automatically determines the threshold and optimizes the segmentation process. Next, the regional growth method is used for multi-point selection to extract certain disease-affected areas. Then, lesion feature extraction is performed in terms of shape, texture, and color. Finally, a multiple linear regression model is implemented to determine the type of disease, and then the least-squares estimation algorithm is used to calculate the coefficients and confidence intervals to set up the disease recognition system.	- The obtained results proved that the proposed recognition system has high accuracy, reliability, and effective recognition ability of plant diseases. - The histogram segmentation method has great advantages, such as speed, efficiency, and accuracy.	- The error rate increases progressively as the disease state becomes more complex because, as the disease worsens, the characteristic parameters become more complex and so the results become unstable.
An IoT based smart solution for leaf disease detection (Thorat et al., 2017) [76]	This paper presents a solution based on a smart farming technique using WSN, a Raspberry PI module, and a camera to establish wireless communication. In addition, CV techniques were used, such as masking, segmentation, and feature extraction to identify leaf diseases. It should be noted that they used the Apache server to retrieve and send data.	- The proposed system allows remote monitoring of the farm. Thus, the recognition of different leaf diseases was carried out successfully.	- The power supply of the system is limited, so the whole process stops if the system fails, which is an inconvenience. Moreover, the images taken during the day can be affected by excessive sunlight or reflections, which means the leaf color cannot be identified by the camera or captured clearly at night.
Plant disease detection using hyperspectral imaging (Moghadam et al., 2017) [77]	The authors used hyperspectral imaging (SWIR and VNIR), ML techniques, and IP for detecting tomato wilt virus in capsicum plants. First, images were acquired from a hyperspectral imaging system consisting of two Headwall push-broom hyperspectral cameras, namely the SWIR hyperspectral camera that provides a spatial resolution of 384 pixels and a spectral range of 900 to 2500 with 168 spectral bands, and the VNIR hyperspectral camera, which provides a spatial resolution of 384 pixels. Then, these images were pre-processed using different pre-processing techniques such as the space-adaptive filtering approach for detection and grid removal. In addition, these images were segmented using an unsupervised k-means clustering algorithm. Further, discriminative feature extraction was performed using the full spectrum, VNIR, SWIR, and vegetation indices. Finally, these features were employed to train classifiers for discriminating leaves obtained from inoculated and healthy plants. They also used other techniques in the process of disease identification as follows: They used the KL divergence or relative entropy to estimate the distances between two distributions of the control group and of the inoculated group. Note that the high-pass filter used is a third-order Butterworth FIR filter, which is applied to flatten the power spectral density of the image in order to detect a known signal corrupted by additive white noise.	- The obtained results showed excellent discrimination based on the full spectrum. - The effectiveness of feature extraction techniques used for automatic disease classification in greenhouse experiments. - The cluster analysis was able to successfully classify the image spectra into two classes by using the significant difference in spectral profile between the vegetation and its surroundings.	- The researchers treated only one disease, TSWV. - They were not able to correlate the reduction in overall values of SWIR dissimilarity with a plant pathogenic biophysical interaction for DAI 7 and 10.
Deep Learning for Image-Based Cassava Disease Detection (Ramcharan et al., 2017) [41]	A new model for the identification of plant diseases is proposed based on TL to train a CNN using a dataset of 2756 images in order to identify two types of damage caused by pests and three diseases. This model was deployed on a mobile application.	- The CNN avoids the tedious and complex step of extracting features from images to train models on a mobile device. The results proved that the TL approach offers greater precision in cassava, and is also an affordable, fast, and easily deployable strategy for digital devices. The model accuracy was 96% for RMD and CMD, 95% for GMD, and 98% for brown spot and cassava.	- Due to several factors, such as the lighting in a complex environment, accurate identification of diseases is challenging.
Early detection and classification of tobacco leaves inoculated with tobacco mosaic virus based on hyperspectral imaging technique (Zhu et al., 2016) [44]	The authors proposed a procedure for the early detection of tobacco disease infected with the mosaic virus by different ML algorithms based on hyperspectral imaging techniques. Images of healthy leaves inoculated with TMV for a period of 7 days, i.e., after inoculation, were acquired by a hyperspectral imaging system every day with the VNIR wavelength region 380–1023 nm. In addition, the spectral reflectance of the predefined ROI was extracted from the hyperspectral images using the ENVI software. The different ML algorithms, namely, RF, SVM, BaBPNN, LS-SVM, PLS—DA, ELM, and LDA were used to quantitatively classify the stages of tobacco disease using EW that were selected using the SPA.	- BPNN and ELM models successfully detected healthy and diseased tobacco leaves (2 DPI, 4 DPI, 6 DPI); the detection rates were 98.33% and 96.67%, respectively. - The classification accuracy of the training set and the test set was 84.17% and 75% respectively.	- They used a single VNIR component from the electromagnetic spectrum. - Generation of average spectra from a single ROI rather than all pixels. - The performance of the SPA-PLS-DA model was relatively poor compared to other models and the accuracy was slightly lower 75%.
Plant disease detection using image processing (Khirade et al., 2015) [78]	This article discussed IP-based methods for the detection of plant diseases, in particular segmentation and feature extraction algorithms. Concerning the segmentation techniques, they exploited the k-mean clustering, boundary and spot detection algorithm, and the Otsu threshold algorithm. Regarding the feature extraction techniques, they studied various methods, such as the color co-occurrence method, and regarding the disease classification phase, they used various algorithms such as ANN, BBPN, and SVM to accurately classify various leaf diseases.	- A variety of IP and machine learning (ML) techniques are discussed.	- The authors did not suggest any system and no results were given.

Table 4. *Cont.*

Title	Methodology	Advantages	Disadvantages
Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features (Arivazhaga et al., 2013) [79]	A software solution is proposed for the detection of unhealthy regions and the automatic classification of diseases using the extracted texture characteristics. Thus, the scheme of the process consists of four steps. First, the acquired RGB leaf images were converted to HSV format. Then, the green pixels were masked and removed, followed by a segmentation process. Further, the texture characteristics were computed using the Color-Co-Occurrence Matrix and, finally, the classification was first performed using the minimum distance criterion, which yielded a gain of 86.7%. Results were then improved by the SVM classifier.	- The proposed method was tested on ten species of plants: beans, mango, lemon, jackfruit, banana, sapota, potato, and tomato. The results gave an accuracy of 94.74% using the SVM classifier. Therefore, the proposed approach can classify leaf diseases with little computational effort.	- Various reasons lead to an erroneous classification, namely, the identifying vectors of the taken features have to be optimized, and the symptoms of the diseased plant leaves vary from the early to the late stage.
Early detection of diseases on leaves by image processing (Han et al., 2013) [80]	The authors worked on images of vine leaves affected by mildew at different stages that were acquired by photographic sensors. They were able to detect diseases that are visible and barely visible to the naked eye using techniques based on the image representation in color space, and hybrids including information on both color and texture. Thus, for visual plant diseases, they were oriented towards changing the color space, in particular, to facilitate detection. To ensure better visualization, they were interested in the bands of each image (the V band of the YUV color space). Regarding diseases barely visible to the eyes, as a solution, they used the image analysis method that combines color and texture information. Otherwise, the most relevant challenge was the disease detection at an early stage. As an optimal solution, they worked on thermal imaging, which is very effective in detecting water stress. Finally, they calculated the Mahalanobis distance for image segmentation.	- As an advantage of the texture analysis (conversion of the image into hybrid space), they discerned 3 classes of textures and, for each class, 20 learning patches, to choose from. - The use of a thermal imaging device allows the premature detection of leaf diseases.	- They only treated mildew disease. - The downside of using the color space change method is that when the mildew stains are at the early stage, this method does not work well; hence, the need to couple this type of color treatment with one based on texture analysis. - They did not calculate the area infected with mildew disease.
Classification of cotton leaf spot disease using image processing edge detection techniques (Revathi et al., 2012) [81]	This article describes how to identify the part affected by leaf diseases using IP techniques. First, to segment the image, the authors used the Canny and Sobel edge detection technique and finally, they proposed an HPCCDD to analyze the images and classify the different diseases.	- The percentage was indicated to reduce leaf diseases. - The validation was undertaken via MATLAB.	- The authors only discussed cotton crops.
Color image segmentation using K-Medoids Clustering (Yerpude et al., 2012) [82]	A color image segmentation method was proposed using k-medoids clustering; the idea is to find groups of objects by finding the medoids for each group. The main objective of this paper is the critical analysis of different disease segmentation techniques.	- The obtained result shows the effectiveness of the k-medoids algorithm on different types of images, such as grayscale images. Moreover, the proposed method is not sensitive to noise.	- The segmented images are highly reliant on the centroids. However, they did not consider finding the optimal number of segments to obtain more accurate results. - The k-means algorithm is very sensitive to outliers, as the data distribution can be significantly distorted if an object has an extremely large value.
Color transform-based approach for disease spot detection on plant leaf (Chaudhary et al., 2012) [83]	A comparison was made between the effect of YCbCr, HSV, and CIELAB color spaces in the disease spot detection process since disease spots are different in color but not in intensity level. First, the different images of wheat, soybeans, rice, corn, cotton, mustard, apple, magnolia, and cherry leaf were taken. Then, these RGB images were converted to YCbCr color space using the color transformation formula and further to CIELAB and HSV color space. Then, for smoothing and enhancing the image, a median filter was applied. Finally, the segmentation of the image was carried out using the Otsu method on the components of the color space: Cr for filtered YCbCr, component H for the filtered HSV space, and component A for the LAB filtered color space.	- The experimental results show that the noise generated by the camera flash, background, and vein can be effectively removed using the CIELAB color model.	- Unfortunately, due to an imbalance in pigment formation and micronutrient deficiency, the color of the vein is different from the leaf spots. Thus, in some cases, these disease spots cannot be detected with precision using the CIELAB method.
Fast and accurate detection and classification of plant diseases (Al-Hiary et al., 2011) [84]	The acquisition of the leaf RGB images is undertaken to perform color space transformation. Once the segmentation of these images is performed using k-means clustering, the value of the green pixels is masked using the threshold obtained through Otsu's method. In addition, the affected clusters were converted to the hue saturation value. For texture analysis, the SGDM matrix is used for the formation of each image. Finally, the disease recognition process is performed by the ANN classifier.	- The color occurrence method is an advantageous strategy that gives exact precision.	- The methodology can be improved to increase the recognition rate of the classification process and to automatically estimate the severity of the disease detected.
Early detection of Fusarium infection in wheat using hyper-spectral imaging (Bauriegel et al., 2011) [85]	In this paper, Fusarium head disease was detected by spectral analysis in wheat, barley, oat, and rye plants. The PCA distinguishes affected from healthy tissue in the wavelength ranges 927–931 nm, 682–733 nm, 560–675 nm, and 500–533 nm. The SAM method is used to classify the degrees of infection. Finally, the best time to identify ear blight is the stage between 71–85 on the BBCH scale. They analyzed 292 spectra and 80 spectra for diseased and healthy tissue, respectively, in time series experiments. Smoothing was performed for these spectra with the "proc expand" function. Then, the first derivative was calculated and 104 individual spectra were subjected to PCA to assess the relevant wavelengths for discriminating between healthy and blight-infected tissue.	- The robustness and efficiency of the proposed algorithm are proved by experimental results of a database of about 500 plant leaves. - The SAM image analysis method correctly classifies the degree of disease at 87%, and the visual assessment error is 10%. - The SAM method yields accurate classification results; it is not practical for an online application because the analysis of 512 spectral bands involves a significant amount of computation.	- It is impossible to distinguish between different degrees of infection using only spectral analysis, due to the lack of symptoms. - The disadvantage of using the SAM method is that it is time consuming. Indeed, it involves the configuration of the reference spectra for the classification and, at the same time, the analysis of all the spectral bands.

- *The Difference between machine learning and deep learning*

The difference between machine and deep learning lies [86] first in the fact that machine learning algorithms deal with quantitative and structured data and, second, the operator is responsible for choosing the right algorithm to extract the features that will

influence the prediction. Deep learning algorithms deal with unstructured data and the algorithm is trained to extract the influential elements in the prediction as shown in Figure S1 in the Supplementary Materials. It should be noted that deep learning algorithms, compared to ML algorithms, demand a large amount of data and high computational power.

4. Discussion

In this paper, the authors reviewed many research articles and identified 129 studies eligible for systematic review using the PRISMA statement as presented in Figure S2 in the Supplementary Materials, these studies involve methodologies in image processing, machine learning, and deep learning particularly focused on the identification and classification of plant diseases. The study showed that the techniques most used in the literature, in general, are the support vector machine [22,59,61,65] (SVM), random forest [87] (RF), artificial neural network [84] (ANN) and convolutional neural network (CNN) [35,39,50].

Additionally, many scientific contributions have focused on the prediction of major diseases affecting wheat, rice, and potatoes, such as powdery mildew [88,89], late blight [90,91], and blast [92,93]. The challenging aspect of this work is the evaluation and investigation of the computational efficiency of each study compared to other studies, because each paper applies different metrics to a variety of diseases in different crops. In addition, many techniques and pretreatment approaches are used to predict disease presence or severity. Accordingly, it is nearly impossible to generalize and compare different articles because it is paramount to follow the same experimental conditions. Thus, the present comparison of the different approaches used was strictly constrained, for example, by considering the types of crops on which the work was undertaken, in addition to the kinds of diseases considered during the work. Therefore, based on these constraints, from the results obtained in related works, it is observed that deep learning-based models have outperformed the classical approaches such as random forest, support vector machine, and k-nearest neighbors classifiers, knowing that the performance of these algorithms has been proven and validated using metrics such as accuracy, sensitivity, specificity, and F1-score etc.

Table 4 shows that several researchers applied spectral analysis using thermal and optical remote sensing images, in addition to multispectral and hyperspectral images. As shown by Duarte-Carvajalino et al. [94], multispectral images were found to be relevant for the early-stage detection of disease, whereas hyperspectral images, which are the most widely used in the existing literature, can predict disease even before symptoms are visible to the naked eye. Note that this difference is due to the spectral resolution used by the two technologies. However, compared to hyperspectral imaging, multispectral imaging offers less data complexity [95]. However, hyperspectral image analysis has various limitations. Several authors have highlighted the high dimensionality of the data as one of the difficulties encountered. As pointed out by Mahlein et al. [95], the high degree of interband correlation leads to information redundancy, generating convergence instability in multivariate prediction models. It is observed that the dataset used by most of the researchers is taken from PlantVillage, and, in the image preprocessing process, most researchers used the histogram equalization to improve the contrast, and the median, Gaussian filter, and Gabor filter for denoising and image enhancement. Furthermore, for image segmentation, researchers have focused on the hue using the k-means and fuzzy c-means algorithm to segment the images; this procedure enables extraction of the region of interest from the given image. Using this, plant features such as texture, shape, and color have often been extracted using the gray-level co-occurrence matrix (GLCM), local binary patterns (LBPs), and histogram of oriented gradients (HOG). This is the most prominent step in the classification process. The researchers used different classification algorithms based on machine learning, and deep and transfer learning, for the classification phase, such as the decision tree classifier (TC), random forest (RF), naive Bayes (NB), support vector machine (SVM), artificial neural network (ANN), probabilistic neural network

(PNN), back-propagation neural network (BPNN), convolutional neural network (CNN), and InceptionV3.

The SVM and NN are mainly used in disease classification. The main advantage of NNs is that they can tolerate noise and are built from available data. The SVM, in turn, offers outstanding classification performance because it nonlinearly maps the input feature vector into a high dimensional space where it can be easily separated. Nevertheless, SVMs are not suitable when the data is very noisy. Hence, when many redundant variables form the input vector, it is possible to the use principal component analysis (PCA) dimensionality reduction method, as used by Kadir et al. [96]. In addition, the convolutional neural network (CNN), faster R-CNN, Vgg16, and ResNet50 models have been used to fully automate the classification process. Moreover, a new approach employed by Turkoglu et al. [97] is the extreme learning machine (ELM), which offers faster learning and better performance and generalization with lower computational cost. The advantages and disadvantages of the classifiers used in the literature are summarized in Table 5.

Table 5. Comparison of various classifiers.

Classifier	Advantages	Drawbacks
Artificial Neural Network ANN	Faster and more accurate than KNN and MMC	Strict because the data can only belong to 1 class
Random Forest	Can classify a large data set with excellent accuracy	Constraints on storage and processing time
Multiclass-Support Vector Machine	Helps to classify the data in several classes	Not suitable when the data is noisy
Least-Square SVM	Fast and not complicated	Pruning techniques must be applied to be sparse
K-Nearest Neighbours KNN	No time spent on training	More time spent on testing and it is expensive to test each instance as well sensitive to noise and yields
Extreme learning machine ELM	Faster training and better generalization	Overfitting (occurs when a complex model has several parameters)
Naïve Bayes	Less training data is required. It works better than its counterparts when the assumption of an independent variable is true	Conditional independence may reduce accuracy
Penalized Discriminant Analysis PDA	Beneficial when the problem has a large number of noisy features	High calculation cost
Bag of Words	Uncomplicated, robust, efficient	It supposes that all words are independent of each other
CNN/Deep learning	It removes the need for a feature extraction step and classification time is shortened	A large amount of data is required for training and it is expensive to compute. They require better hardware such as Graphical Processing Unit (GPU).
Transfer Learning	This helps to apply CNN to problems with a small amount of training data	The pretrained model may not have classes with the desired labels all the time

In this regard, a study was carried out by Ngugi et al. [98] to compare the performance of 10 deep learning models using the PlantVillage dataset, namely AlexNet, ResNet-101, GoogleNet, DenseNet201, Vgg16, Inceptionv3, InceptionResNetv2, SqueezeNet, ShuffleNet, and MobileNets. We present the results obtained by the different architectures for all the performance measures in Table 6 below.
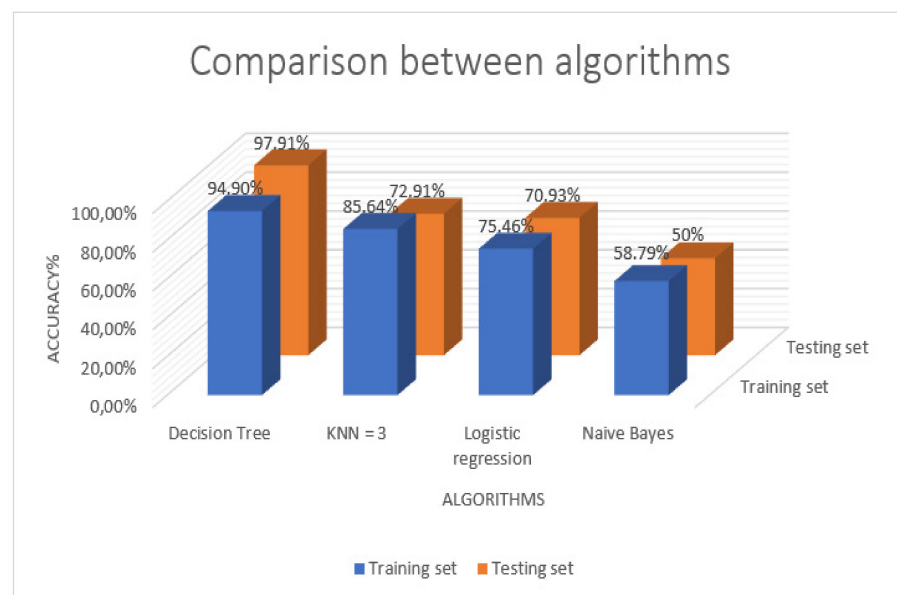
Table 6. The test set performance of 10 models considered in this comparative study.

Architecture	Recall	F1-score	Precision	Accuracy	Specificity
AlexNet	0.9843	0.9856	0.9871	0.9897	0.9997
InceptionV3	0.9906	0.9916	0.9926	0.9948	0.9999
GoogleNet	0.9874	0.9881	0.9891	0.9899	0.9997
SqueezeNet	0.9791	0.9787	0.9785	0.9837	0.9996
DenseNet201	0.9965	0.9961	0.9958	0.9973	0.9999
VGG16	0.9932	0.9930	0.9928	0.9951	0.9999
ResNet101	0.9936	0.9929	0.9924	0.9851	0.9999
ShuffleNet	0.9901	0.9897	0.9895	0.9929	0.9998
MobileNetv2	0.987	0.9862	0.9895	0.9905	0.9997
InceptionResNetv2	0.9887	0.9893	0.9901	0.9930	0.9998

According to Table 6, the DenseNet201 model is the most suitable because it requires less storage and has the advantage of having the best performance measures (accuracy = 0.9973, precision = 0.9958, recall = 0.9965, specificity = 0.9999, F1 score = 0.9961). However, it requires a longer learning time (82 h) compared to the InceptionV3 and ResNet-101 models; nonetheless, their accuracies are slightly lower than those of DenseNet201. Therefore, special care should be taken when choosing between these three architectures, as each model has certain advantages and limitations. By comparison, the small MobileNet, SqueezeNet, and ShuffleNet architectures are desirable in embedded and mobile applications where computing resources are limited, due to their short learning times and low storage requirements, while still achieving high accuracy.

In this regard, another comparative study of four machine learning algorithms—k-nearest neighbors, decision tree, naive Bayes, and logistic regression—was performed by Ahmed et al. [99] to detect three rice plant diseases where the images were taken from the same PlantVillage database.

As shown in Figure 3, the best accuracy (over 97% by applying it to the test dataset) was obtained by the decision tree algorithm after 10 cross-validations.

Figure 3. Comparison between machine learning algorithms.

To briefly summarize this section, it is inferred that multispectral and hyperspectral imagery represents a valuable source of useful information for developing autonomous non-

invasive systems to predict abiotic and biotic stresses in plants. Additionally, the integration of multiple data sources will strengthen and increase the stability and generalization capabilities of the algorithms. Furthermore, from the results obtained in the literature, it appears that the automatic extraction of leaf features performed by deep learning-based models is more relevant and efficient than the process of extracting these features using traditional approaches such as the grey level co-occurrence matrix (GLCM), area-based techniques (ABTs), and scale invariant feature transform (SIFT). However, it is noted that there is a lack of validation of the models used in real-world scenarios. Therefore, proper validation is necessary for the studies to have an accurate and general impact.

5. Unresolved Challenges in the Crop Disease Detection Field

The above section presents a wealth of promising research undertaken in the past few years in the area of crop foliar disease recognition and detection using a range of techniques. In the existing literature to date, there are numerous unresolved challenges that remain to be address and overcome to derive robust and feasible crop disease detection systems that can operate accurately under various field conditions. The most prominent of these highlighted challenges are:

5.1. Insufficient Data

The major problem in the use of deep learning models for plant disease detection is the insufficiency of datasets in terms of both diversity and size [100] because these models have extremely large data requirements. In the majority of cases, the identification of plant diseases has been performed under ideal and controlled conditions [43], such as the presence of a single disease with a homogeneous background. In addition, environmental conditions are not considered; hence, the accuracy rate obtained will be higher than that actually obtained in a practical application. Additionally, image labeling is a very laborious and tedious task. Due to these factors, the production of a reliable, efficient, and comprehensive dataset is extremely challenging. At present, there are six ways to deal with the lack of a dataset: data augmentation techniques, data sharing, citizen science, transfer learning, synthetic data, and few-shot learning.

5.2. Imbalanced Data

The most commonly used datasets for crop disease detection are cleaned or their unbalanced nature is ignored to fully concentrate on training algorithms and avoid being distracted by other problems. However, in real-world settings, the distribution across classes is skewed and unbalanced [101], ranging from mildly biased to severely unbalanced. This poses a challenge for predictive modeling and may require specialized techniques, such as re-sampling techniques, because the machine learning algorithms typically employed for classification are built around the assumption of an equal number of examples for each class. As a result, some of the models have poor predictive performance, especially for the minority class which is more susceptible to misclassification than the majority class.

5.3. Vanishing Gradient Problem

Hochreiter's work [102] showed an issue called the "vanishing gradient problem" that arises during the training phase when employing back-propagation learning techniques with neural networks. Specifically, each weight of the neural network is updated based on the current weight and is proportionally related to the partial derivative of the error function. However, this updating of the weights may not take place in some cases due to an extremely small gradient that approaches zero. As a result, the gradient descent does not converge to the optimum and the neural network stops completely [103].

5.4. Exploding Gradient Problem

The opposite problem to the vanishing problem is the gradient explosion problem [104]. Specifically, the gradients become increasingly larger as the back-propagation algorithm

advances. This will lead to extremely large updates of the network weights and causes the gradient descent to diverge, which means that the system becomes unstable [103]. Thus, the model will lose its ability to learn efficiently. In general, as we move up the network during back-propagation, the gradient grows exponentially by repeatedly multiplying the gradients. As a result, the weight values can become incredibly large and spill over to become a non-numerical value (NaN).

5.5. Overfitting and Underfitting Problem

Learning models have excessively high chances of overfitting and underfitting the data in the training stage due to the large number of parameters involved, which are correlated in complex ways. Such situations reduce the ability of the model to perform well on the tested data. Thus, it is considered that a learning algorithm is underfitting when it is unable to grasp the underlying trend in the data. Its occurrence simply means high bias, low variance, and that the model does not fit the data well enough. This usually occurs when fewer data are available to build an accurate model and also when a linear model is attempted to be built with non-linear data. Conversely, a model is said to be overfitted when it is trained with a large quantity of data; it then learns from noise and inaccurate data inputs in the dataset. Then, the model does not correctly categorize the data due to excessive detail and noise. Its occurrence simply means low bias and high variance. The overfitting occurs in nonlinear and nonparametric approaches, as these kinds of learning algorithms have greater leeway in setting up an unrealistic model. Ideally, both of these should not exist in models, but they are generally challenging to eliminate. This problem was noted by Ahmad et al. [105], whose model based on efficient convolutional neural networks tends to overfit during the training of the first epochs.

5.6. Image Acquisition: Conditions of Image Capture (Lighting, Spatial Location, Wind and Camera)

Ideally, images should be captured under similar conditions. However, in practice, this may only be feasible in the laboratory because it is extremely difficult to monitor the conditions of capture. Thus images may present unpredictable characteristics, making disease identification a daunting task. Moreover, the variable capture conditions have proven to be a challenging issue in measuring the severity of citrus leaf canker [106] and in identifying citrus diseases [107]. In light of this, several endeavors have been undertaken to develop methods of invariant illumination [108]. Nevertheless, their success to date is still relatively modest.

- **Lighting Issue**

Crops grow in natural environments that fluctuate greatly. Thus, images are impacted by numerous factors, such as wind, illumination, and other climatic conditions. Consequently, lighting issues are inevitable, and completely eliminating the variations is almost impossible. Nevertheless, certain endeavors have been made to mitigate them, e.g., Pourreza et al. [109] developed a system to detect real-time citrus Huanglongbing disease using a narrow-band imaging and polarizing filter set. Specular lighting, the simultaneous presence of light and shadow, is the most difficult problem to deal with. However, the presence of specular lighting can be lessened by changing either the angle at which the image is taken or the leaf position, although this likely causes some degree of reflection. Furthermore, specular reflections and shadows were noted by Zhou et al. [110] as the main source of error in monitoring Cercospora leaf spot on sugar beets, which occurred because of the automatic captures that complicate the prevention of lighting problems.

- **Camera**

The image resolution is one of the crucial factors that has a direct influence on the image features. A higher resolution enables the detection of small lesions and spores. Moreover, the device being used to capture the image also influences these features.

5.7. Image Preprocessing

During the preprocessing and storage of leaf images, more information is lost as the compression ratio is increased. This may not dramatically influence the analysis of large lesions, but may severely distort small symptoms. Therefore, compression should be kept to a minimum or even avoided, especially if the symptoms are tiny.

5.8. Image Segmentation and Symptom Discrimination

In general, symptoms do not have clear boundaries; they gradually disappear in normal tissue, making the distinction between healthy and diseased areas highly ambiguous. This clearly affects the accuracy of the threshold and extracted features. Although manual and visual representation cannot clearly determine the edges, any machine-based representation will be prone to many subjective issues. Notably, the issue of subjective delineation of affected regions was first addressed by Olmstead et al. [111] and later by Moya et al. [112], who emphasized that some sort of external reference needs to be established for proper validation of disease identification methods. However, without the use of a reference, Oberti et al. [113] observed for leaf powdery mildew that the number of false negatives or positives seen on the symptom discolored zones is too high. In summary, few solutions have been suggested for this problem because inconsistencies are intrinsic to the process. Furthermore, other difficulties are encountered when segmenting and locating regions of interest (ROIs):

- A leaf may overlap with another leaf or other parts of the plant, and they may even be tilted or covered with dew or dust.
- Images with complex backgrounds can render the segmentation of ROIs where symptoms appear challenging and intricate.

5.9. Feature Selection and Extraction

Although some plant species can be identified on the basis of leaf shape, other species have similar leaf shapes. Furthermore, symptoms do not necessarily arise in zones that are easily accessible; in practice, they can frequently be under the leaves or covered by other obstructions, or diseases can appear on the stems, fruits, or even flowers. Unfortunately, the latter problem has not attracted enough attention on the part of researchers. Furthermore, it is observed from the literature to date that researchers have mainly focused on the disease detection on the upper leaf surface. Nevertheless, Fuentes et al. [34] suggested using the faster network R-CNN for detecting a number of tomato plant diseases in several locations.

5.10. Disease Classification

In many of the cases listed below, the classifier used to identify plant diseases may not be able to distinguish between them; for example, if the symptoms presented by different diseases are visually very similar, as both Ahmad et al. [114] and Wiwartet al. [115] have stated. In addition, the difficulties stated below are highly relevant to measuring the disease severity:

- ***Differences in disease symptoms:*** According to the disease development stage, a specific disease can present very distinct characteristics in the symptoms' shape, color, and size, causing serious identification problems. It should be noted that many different diseases can occur at the same time, making it extremely complex to distinguish between combinations of symptoms and individual symptoms. This problem was noted by Camargo et al. [116] when handling symptoms produced by black streak disease on banana leaves, and Moya et al. [112] when evaluating powdery mildew severity on squash leaves.
- ***Diseases can occur simultaneously with many disorders, such as nutritional deficiencies, pests, and diseases:*** Typically, most techniques consider that there is only a single disease per image when, in reality, several other diseases can be present at the same time, in addition to other kinds of disorders, such as nutritional deficiencies and pests. These simultaneous symptoms can be either separate or physically combined, making

disease identification a significant challenge. In this regard, Bock et al. [106] observed the simultaneous presence of symptoms arising from different diseases and noted that this can lead to identification issues, and that more advances will be required to cope with this issue.

- *The symptoms' similarity between different disorder types:* Symptoms resulting from various disorders, such as diseases, phytotoxicity, presence of parasites, and nutritional deficiencies, can be visually similar. As a result, it can be extremely difficult to determine a symptom's source with certitude, particularly if only the visible spectrum is used in the identification process. This forces methods to rely on tiny differences to discriminate between the symptoms. Numerous researchers have stated that some disorders have close similarities, leading to major issues of discrimination. In this regard, Ahmad et al. [114] reported that symptoms resulting from Fusarium, Mosaic Potyvirus, Alternaria, and Phomopsis in soybean were very similar, and their classifier was unable to discern between them. This explains why the majority of studies conducted to date have chosen to tackle only diseases whose symptoms are quite dissimilar and, even then, their choices remain a significant challenge.

5.11. Other Challenges

Some other challenges facing automatic plant disease identification techniques cannot be categorized in the same way as those mentioned above. These challenges include reducing complexity, in addition to computational and memory demands [117], because low-cost computers and cameras have a very limited computational resource. At the same time, as image resolution is increasing, the computational resources are also growing. Another major concern is the lack of properly labeled [37] and sufficiently large datasets with high variability. This is notably the biggest hurdle when training recurrent neural network (RNN) models for plant disease detection, because collecting images in the field is not only a laborious task but also requires the guidance of agricultural experts for accurate annotation. Nevertheless, two free datasets exist [118]—PlantVillage and the Image Database of Plant Disease Symptoms dataset PDDB. Moreover, at present, no appropriate technology has been developed to automatically crop the leaf images around the affected area. A further issue is that hyperspectral data contain more than one hundred adjacent spectral bands and thus cannot be linearly trained [119]. Furthermore, these bands in different spectral regions are highly redundant [54,55] when extracting information to form an artificial neural network (ANN).

6. Future Work and Possible Solutions to Ongoing Limitations

In the previous section, gaps in the existing literature were highlighted to orient future research in this area. Thus, future work should first aim at acquiring diverse and large-size datasets to further promote research in this direction. Moreover, it is highly desirable to develop compact convolutional neural network CNN-based models that can achieve higher accuracy and promote the use of these technologies in the embedded platforms. Secondly, more emphasis should be placed in future research on the development of reliable methods and techniques to remove backgrounds and incorporate other forms of data, such as meteorological trends, disease occurrence history, and spatial location, to enhance the accuracy and reliability of disease identification systems. Additionally, disease recognition at different locations on plants and trees, such as the stems, blooms, and fruits, should receive greater attention from researchers due to its tremendous importance. One possible means to circumvent some of the limitations is to implement constraints to restrict variations in image capture conditions. However, even with very tight restrictions, many challenges will remain.

Some of the key challenges can be mitigated through the use of the most sophisticated approaches borrowed from the machine learning and computer vision fields. These include Markov random fields, mean shift, graph theory, and large margin nearest neighbor classification (LMNN), among other methods that have not yet been properly harnessed.

In this regard, the proposed solutions to remedy the challenges presented above can be summarized as follows.

6.1. Data Augmentation Techniques

If the aim is to avoid the overfitting problem and expand the size of the dataset without manually collecting new images, data augmentation techniques are a possible solution for any limited data problem [120,121]. Data augmentation incorporates a collection of methods that improve the attributes and size of training datasets. Thus, DL models can perform better when these techniques are exploited, such as rotation, canny edge detection, shear, image noise addition, shift, and flipping.

6.2. Tackling Overfitting Problem

Overfitting is one of the fundamental problems encountered when using learning models, and occurs due to the sensitivity to the scale of the cross-entropy loss and the continuous updating of the gradient. Three classes exist to avoid the overfitting problem. The first acts on both the model parameters and the model architecture. This includes the most familiar approaches, such as batch normalization [122], weight decay [123], and dropout [124]. Weight decay is the technique that is commonly used by default in all algorithms as a universal regularizer. The second class operates on model inputs such as data augmentation and corruption. One of the causes of the overfitting problem is the lack of training data; as a result, the learned distribution does not exactly reflect the real distribution. In contrast, the marginalized corrupted feature (MCF) exclusively improves the solution in data augmentation. MCF is a new approach to combat overfitting in supervised learning [125]. The main idea of the MCF is to allow the models to be regularized by training them on corrupted data copies, without raising the computational complexity. The final class works on the output of the model. A technique was recently proposed by Pereyra et al. [126] based on penalizing confident output distributions for model regularization. This method has demonstrated its high capacity to regularize CNN and RNN models. Hence, it will be judicious to explore these techniques in the field of crop disease detection.

6.3. Few-Shot Learning

In cases in which the dataset is extremely small, the techniques mentioned above may not be useful; that is, if there is a task in which the classification must be built with only one or two samples per class, and each sample is difficult to find. In such a case, innovative approaches are needed; one of these is few-shot learning (FSL) [127]. This is a relatively recent subfield of machine learning that needs more refinement and research. FSL allows the classification of new data when there are only a few training samples with supervised information. The approach of building an FSL classifier is suitable for solving the kind of problem related to rare plant pathologies, in which images are lacking for use in the training set. Typically, two major approaches are implicated in solving one-shot or few-shot machine learning issues, namely, the data-level approach and the parameter-level approach.

6.4. Transfer Learning

Recent research has revealed the extensive use of deep CNNs, which require a large quantity of data to perform effectively. The common challenge associated with the use of such models concerns the lack of training data. Specifically, collecting a large volume of data is an exhausting task, and no successful solution is available at this time. Therefore, in order to solve the fundamental dilemma of insufficient data, it is advisable to use TL models, which are highly effective in such cases [128]. In simple terms, transfer learning is the process by which the model trained for a specified task is reused as the starting point for training a new model. It attempts to transfer information from the original domain to the destination domain. This learning process is illustrated in Figure 4.
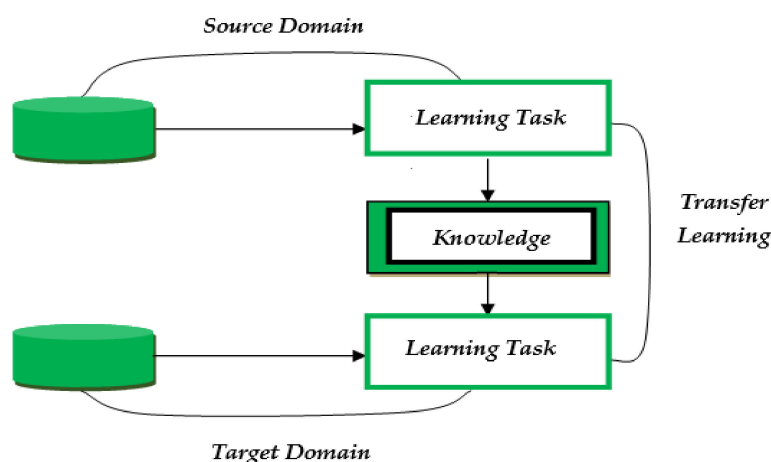
Figure 4. Transfer learning process.

Forthcoming research endeavors can be devoted to automatically estimate the detected disease severity, and expanded to attain the highest accuracy and speed via developing hybrid approaches such as genetic algorithms (GAs) and neural networks (NN) to increase the disease recognition rate, and combining particle swarm optimization (PSO) with other tools, such as gradient search techniques, to ensure a much higher speed. In addition, advanced and appropriate preprocessing techniques should be adopted to prevent noise interference in disease detection, in addition to partitioning the training and test data by employing more advanced techniques, such as stratified sampling, in order to create a well-balanced data partition, and avoid underfitting and overfitting. In addition, optimizing feature vectors should be contemplated to increase the disease recognition rate in these various stages, and recurrent neural network (RNN) models and the long-term memory function should be used to extract memory and temporal dimensions that can subsequently be harnessed for plant growth estimation. Finally, a web application can be designed with a range of features, such as displaying the identified diseases in the crops from leaf images taken by a smartphone camera. A discussion forum can also be developed for agronomists and farmers to talk about treatment and early preventive measures for the encountered diseases. Moreover, plant electrophysiology is a promising avenue for future research [129], i.e., the electrical signal response produced in plants can be used for real-time disease detection. This approach is based on the fact that plants perceive the environment, and this perception is translated by a generation of electrical signals that essentially represent changes in their underlying physiological processes. Under the influence of stress, the metabolic activities of plant tissues and cells are unstable, which is inevitably reflected in the plant's physiological electrical properties. As a result, the extraction of substantial characteristics from the generated electrical signals, such as impedance, varying capacitance, and conductivity, would be a highly interesting research direction for the classification of diseases in plants and crops.

7. Conclusions

Crop diseases are one of the main challenges in the farming sector. Thus, there is a need to identify crop diseases at the earliest stage to lessen disease severity and to curb disease propagation on farms. Accordingly, prominent and advanced research has been conducted in recent years on several kinds of disease identification techniques, as presented in this work. The main difference between other surveys and the present paper is the thorough technical analysis of the individual papers, and the approaches that have been applied to date. This provides a guideline and references to scientific communities. This paper also provides readers with insights into the automatic crop disease detection process and the key factors, namely the lack of sharp edges around the symptoms; fluctuating imaging conditions; variable symptoms presented by diseases; similar symptoms presented by different disorders; and the concomitant presence of symptoms arising from various

disorders. These issues have a relevant impact on the effectiveness of both the image processing methods and the analytical tools that have been introduced to date. From this survey, it is concluded that image preprocessing directly impacts the segmentation process. Moreover, the k-means clustering algorithm was found to be the most suitable technique for segmenting disease-affected leaves. In addition, convolutional neural network (CNN) models were revealed to be extremely powerful and proficient in locating visual patterns in images. Notably, the use of computer vision and artificial intelligence in crop diagnostics in the agricultural sector is still recent, which implies that their numerous alternatives and opportunities remain to be explored, which may help mitigate the above-mentioned challenges. Additionally, with the increase in available computing power, previously demanding strategies can now be easily executed. Thus, based on this in-depth study of the existing literature on crop foliar disease automatic detection, in upcoming work the researchers intend to develop an efficient, accurate, low-cost, and swift system capable of identifying crop diseases from foliar images. In addition, this identifying system will be implemented in a mobile application, allowing an alert to be sent to the farmer once the disease is detected to enable him to intervene as soon as possible.

Abbreviations

The following abbreviations are used throughout this manuscript.

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
TL	Transfer Learning
CV	Computer Vision
IoT	Internet of Things
IP	Image Processing
WSN	Wireless Sensor Network
NN	Neural Network
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DNN	Deep Neural Network
RNN	Recurrent Neural Network
PNN	Probabilistic Neural Network
BBPN	Back-Propagation Neural Network
GOA	Grasshopper Optimization Algorithm
DAML	Deep Adversarial Metric Learning
KNN	K-Nearest Neighbor
SVM	Support Vector Machine

M-SVM	Multiclass-Support Vector Machine
LS-SVM	Least Squares Supporting Vector Machine
RF	Random Forest
PDA	Penalized Discriminant Analysis
ELM	Extreme Learning machine
NB	Naïve Bayes
VGG	Visual Geometry Group
ResNet	Residual Neural Network
C-GAN	Conditional Generative Adversarial Network
GPU	Graphics Processing Unit
MEAN block	Mobile End Apple Net block
FSL	Few-Shot Learning
GOA	Grasshopper Optimization Algorithm
GA	Genetic Algorithms
PSO	Particle Swarm Optimization
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
PLS—DA	Partial Least Squares Discrimination Analysis
SPA	Successive Projection Algorithm.
SAM	Spectral Angle Mapper
BBCH	Biologische Bundesanstalt, Bundessortenamt and CHemical industry
RBM	Restricted Boltzmann Machine
AE	Auto-Encoder
EW	Effective Wavelengths
GLCM	Gray Level Cooccurrence Matrix
CPDA	Color Processing Detection Algorithm.
GUI	Graphical User Interface
LBP	Local Binary Patterns
HOG	Histogram-Oriented Gradient
HSV	Hue Saturation Value
RoI	Region of Interest
HPCCDD	Homogeneous Pixel Counting technique for Cotton Disease Detection
CMD	Cassava Mosaic Disease
RMD	Red Mite Damage
GMD	Green Mite Damage
TMV	Tobacco Mosaic Virus
VNIR	Visible and Near-Infrared
SWIR	Short Wavelength Infrared
ENVI	Environment for Visualizing Images
KL	Kullback Leibler
ABT	Area-Based Techniques
SIFT	Scale Invariant Feature Transform
MFC	Marginalized Corrupted Features
LMNN	Large Margin Nearest Neighbor
PRISMA	Preferred reporting items for systematic reviews and meta-analyses

References

1. Lucas, G.B.; Campbell, C.L.; Lucas, L.T. Causes of plant diseases. In *Introduction to Plant Diseases*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 9–14.
2. Shirahatti, J.; Patil, R.; Akulwar, P. A survey paper on plant disease identification using machine learning approach. In Proceedings of the 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 15–16 October 2018; pp. 1171–1174.
3. Liu, L.; Dong, Y.; Huang, W.; Du, X.; Ren, B.; Huang, L.; Zheng, Q.; Ma, H. A disease index for efficiently detecting wheat fusarium head blight using sentinel-2 multispectral imagery. *IEEE Access* **2020**, *8*, 52181–52191. [CrossRef]

4. Prasad, R.; Ranjan, K.R.; Sinha, A. AMRAPALIKA: An expert system for the diagnosis of pests, diseases, and disorders in Indian mango. *Knowl.-Based Syst.* **2006**, *19*, 9–21. [CrossRef]

5. Singh, V.; Misra, A.K. Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf. Process. Agric.* **2017**, *4*, 41–49. [CrossRef]

6. Afifi, A.; Alhumam, A.; Abdelwahab, A. Convolutional neural network for automatic identification of plant diseases with limited data. *Plants* **2021**, *10*, 28. [CrossRef] [PubMed]

7. Mugithe, P.K.; Mudunuri, R.V.; Rajasekar, B.; Karthikeyan, S. Image processing technique for automatic detection of plant diseases and alerting system in agricultural farms. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 28–30 July 2020; pp. 1603–1607.

8. Parikshith, H.; Rajath, S.N.; Kumar, S.P. Leaf disease detection using image processing and artificial intelligence—A survey. In Proceedings of the International Conference On Computational Vision and Bio Inspired Computing, Coimbatore, India, 25–26 September 2019; pp. 304–311.

9. Sethy, P.K.; Barpanda, N.K.; Rath, A.K.; Behera, S.K. Deep feature based rice leaf disease identification using support vector machine. *Comput. Electron. Agric.* **2020**, *175*, 105527. [CrossRef]

10. Mitra, D. Emerging plant diseases: Research status and challenges. *Emerg. Trends Plant Pathol.* **2021**, 1–17. [CrossRef]

11. Ichiki, T.U.; Shiba, T.; Matsukura, K.; Ueno, T.; Hirae, M.; Sasaya, T. Detection and diagnosis of rice-infecting viruses. *Front. Microbiol.* **2013**, *4*, 289.

12. Lacomme, C.; Holmes, R.; Evans, F. Molecular and serological methods for the diagnosis of viruses in potato tubers. In *Plant Pathology*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 161–176.

13. Balodi, R.; Bisht, S.; Ghatak, A.; Rao, K. Plant disease diagnosis: Technological advancements and challenges. *Indian Phytopathol.* **2017**, *70*, 275–281. [CrossRef]

14. Bachika, N.A.; Hashima, N.; Wayayoka, A.; Mana, H.C.; Alia, M.M. Optical imaging techniques for rice diseases detection: A review. *J. Agric. Food Eng.* **2020**, *2*. Available online: https://www.myjafe.com/wp-content/uploads/2020/04/MYJAFE2020-0001.pdf (accessed on 15 December 2021).

15. Galletti, P.A.; Carvalho, M.E.; Hirai, W.Y.; Brancaglioni, V.A.; Arthur, V.; da Silva, C.B. Integrating optical imaging tools for rapid and non-invasive characterization of seed quality: Tomato (*Solanum lycopersicum* L.) and carrot (*Daucus carota* L.) as study cases. *Front. Plant Sci.* **2020**, *11*, 577851. [CrossRef] [PubMed]

16. Bauriegel, E.; Herppich, W.B. Hyperspectral and chlorophyll fluorescence imaging for early detection of plant diseases, with special reference to Fusarium spec. infections on wheat. *Agriculture* **2014**, *4*, 32–57. [CrossRef]

17. Mishra, P.; Polder, G.; Vilfan, N. Close range spectral imaging for disease detection in plants using autonomous platforms: A review on recent studies. *Curr. Robot. Rep.* **2020**, *1*, 43–48. [CrossRef]

18. Neupane, K.; Baysal-Gurel, F. Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: A review. *Remote Sens.* **2021**, *13*, 3841. [CrossRef]

19. Abbas, A.; Jain, S.; Gour, M.; Vankudothu, S. Tomato plant disease detection using transfer learning with C-GAN synthetic images. *Comput. Electron. Agric.* **2021**, *187*, 106279. [CrossRef]

20. Hamdani, H.; Septiarini, A.; Sunyoto, A.; Suyanto, S.; Utaminingrum, F. Detection of oil palm leaf disease based on color histogram and supervised classifier. *Optik* **2021**, *245*, 167753. [CrossRef]

21. Sun, H.; Xu, H.; Liu, B.; He, D.; He, J.; Zhang, H.; Geng, N. MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks. *Comput. Electron. Agric.* **2021**, *189*, 106379. [CrossRef]

22. Pothen, M.E.; Pai, M.L. Detection of rice leaf diseases using image processing. In Proceedings of the 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 11–13 March 2020; pp. 424–430.

23. Kelman, A.; Pelczar Michael, J.; Shurtleff Malcolm, C.; Pelczar Rita, M. Plant Disease. Available online: https://www.britannica.com/science/plant-disease (accessed on 15 December 2021).

24. Cerda, R.; Avelino, J.; Gary, C.; Tixier, P.; Lechevallier, E.; Allinne, C. Primary and secondary yield losses caused by pests and diseases: Assessment and modeling in coffee. *PLoS ONE* **2017**, *12*, e0169133. [CrossRef] [PubMed]

25. Kaur, M.; Bhatia, R. Leaf disease detection and classification: A comprehensive survey. In Proceedings of the International Conference on IoT Inclusive Life (ICIIL 2019), Chandigarh, India, 19–20 December 2019; pp. 291–304.

26. Singh, V.; Misra, A. Detection of unhealthy region of plant leaves using image processing and genetic algorithm. In Proceedings of the 2015 International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India, 19–20 March 2015; pp. 1028–1032.

27. Karthika, J.; Santhose, M.; Sharan, T. Disease detection in cotton leaf spot using image processing. *J. Phys. Conf. Ser.* **2021**, *1916*, 012224. [CrossRef]

28. Devaraj, A.; Rathan, K.; Jaahnavi, S.; Indira, K. Identification of plant disease using image processing technique. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019; pp. 0749–0753.

29. Mojjada, R.K.; Kumar, K.K.; Yadav, A.; Prasad, B.S.V. Detection of plant leaf disease using digital image processing. *Mater. Today Proc.* **2020**. [CrossRef]

30. Iqbal, M.A.; Talukder, K.H. Detection of potato disease using image segmentation and machine learning. In Proceedings of the 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 4–6 August 2020; pp. 43–47.

31. Srivastava, A.R.; Venkatesan, M. Tea leaf disease prediction using texture-based image processing. In *Emerging Research in Data Engineering Systems and Computer Communications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 17–25.

32. Whitmire, C.D.; Vance, J.M.; Rasheed, H.K.; Missaoui, A.; Rasheed, K.M.; Maier, F.W. Using machine learning and feature selection for alfalfa yield prediction. *AI* **2021**, *2*, 6. [CrossRef]

33. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [CrossRef]

34. Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **2017**, *17*, 2022. [CrossRef]

35. Geetharamani, G.; Pandian, A. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Comput. Electr. Eng.* **2019**, *76*, 323–338.

36. Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y.A. Using deep transfer learning for image-based plant disease identification. *Comput. Electron. Agric.* **2020**, *173*, 105393. [CrossRef]

37. Singh, D.; Jain, N.; Jain, P.; Kayal, P.; Kumawat, S.; Batra, N. PlantDoc: A dataset for visual plant disease detection. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, Hyderabad, India, 5–7 January 2020; pp. 249–253.

38. Al-bayati, J.S.H.; Üstündağ, B.B. Evolutionary feature optimization for plant leaf disease detection by deep neural networks. *Int. J. Comput. Intell. Syst.* **2020**, *13*, 12–23. [CrossRef]

39. Arsenovic, M.; Karanovic, M.; Sladojevic, S.; Anderla, A.; Stefanovic, D. Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry* **2019**, *11*, 939. [CrossRef]

40. Costa, J.; Silva, C.; Ribeiro, B. Hierarchical deep learning approach for plant disease detection. In Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis, Madrid, Spain, 1–4 July 2019; pp. 383–393.

41. Ramcharan, A.; Baranowski, K.; McCloskey, P.; Ahmed, B.; Legg, J.; Hughes, D.P. Deep learning for image-based cassava disease detection. *Front. Plant Sci.* **2017**, *8*, 1852. [CrossRef] [PubMed]

42. Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. *Comput. Electron. Agric.* **2019**, *161*, 272–279. [CrossRef]

43. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef] [PubMed]

44. Zhu, H.; Cen, H.; Zhang, C.; He, Y. Early detection and classification of tobacco leaves inoculated with tobacco mosaic virus based on hyperspectral imaging technique. In Proceedings of the 2016 ASABE Annual International Meeting, Orlando, FL, USA, 17–20 July 2016; p. 1.

45. Zhu, H.; Chu, B.; Zhang, C.; Liu, F.; Jiang, L.; He, Y. Hyperspectral imaging for presymptomatic detection of tobacco disease with successive projections algorithm and machine-learning classifiers. *Sci. Rep.* **2017**, *7*, 4125. [CrossRef] [PubMed]

46. Cui, S.; Ling, P.; Zhu, H.; Keener, H.M. Plant pest detection using an artificial nose system: A review. *Sensors* **2018**, *18*, 378. [CrossRef] [PubMed]

47. Ma, J.; Du, K.; Zheng, F.; Zhang, L.; Gong, Z.; Sun, Z. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Comput. Electron. Agric.* **2018**, *154*, 18–24. [CrossRef]

48. Tran, T.-T.; Choi, J.-W.; Le, T.-T.H.; Kim, J.-W. A comparative study of deep CNN in forecasting and classifying the macronutrient deficiencies on development of tomato plant. *Appl. Sci.* **2019**, *9*, 1601. [CrossRef]

49. Tian, Y.; Yang, G.; Wang, Z.; Wang, H.; Li, E.; Liang, Z. Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Comput. Electron. Agric.* **2019**, *157*, 417–426. [CrossRef]

50. De Luna, R.G.; Dadios, E.P.; Bandala, A.A. Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. Proceedings of TENCON 2018-2018 IEEE Region 10 Conference, Jeju Island, Korea, 28−31 October 2018; pp. 1414–1419.

51. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–36. [CrossRef]

52. Mananze, S.; Pôças, I.; Cunha, M. Retrieval of maize leaf area index using hyperspectral and multispectral data. *Remote Sens.* **2018**, *10*, 1942. [CrossRef]

53. Rumpf, T.; Mahlein, A.-K.; Steiner, U.; Oerke, E.-C.; Dehne, H.-W.; Plümer, L. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agric.* **2010**, *74*, 91–99. [CrossRef]

54. Ng, W.; Minasny, B.; Malone, B.P.; Sarathjith, M.; Das, B.S. Optimizing wavelength selection by using informative vectors for parsimonious infrared spectra modelling. *Comput. Electron. Agric.* **2019**, *158*, 201–210. [CrossRef]

55. Wei, C.; Huang, J.; Wang, X.; Blackburn, G.A.; Zhang, Y.; Wang, S.; Mansaray, L.R. Hyperspectral characterization of freezing injury and its biochemical impacts in oilseed rape leaves. *Remote Sens. Environ.* **2017**, *195*, 56–66. [CrossRef]

56. Xie, C.; Shao, Y.; Li, X.; He, Y. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging. *Sci. Rep.* **2015**, *5*, 16564. [CrossRef]

57. Chen, W.-L.; Lin, Y.-B.; Ng, F.-L.; Liu, C.-Y.; Lin, Y.-W. RiceTalk: Rice blast detection using internet of things and artificial intelligence technologies. *IEEE Internet Things J.* **2019**, *7*, 1001–1010. [CrossRef]

58. Krishna, M.; Sulthana, S.F.; Sireesha, V.; Prasanna, Y.; Sucharitha, V. Plant disease detection and pesticide spraying using dip and IoT. *J. Emerg. Technol. Innov. Res.* **2019**, *6*, 54–58.

59. Mishra, M.; Choudhury, P.; Pati, B. Modified ride-NN optimizer for the IoT based plant disease detection. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 691–703. [CrossRef]

60. Truong, T.; Dinh, A.; Wahid, K. An IoT environmental data collection system for fungal detection in crop fields. In Proceedings of the 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, Canada, 30 April−3 May 2017; pp. 1–4.

61. Devi, R.D.; Nandhini, S.A.; Hemalatha, R.; Radha, S. IoT enabled efficient detection and classification of plant diseases for agricultural applications. In Proceedings of the 2019 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 21−23 March 2019; pp. 447–451.

62. Kumar, S.; Prasad, K.; Srilekha, A.; Suman, T.; Rao, B.D.; Krishna, J.N.V. Leaf disease detection and classification based on machine learning. In Proceedings of the 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, Karnataka, India, 10−11 July 2020; pp. 361–365.

63. Mallick, D.K.; Ray, R.; Dash, S.R. Detection and classification of crop diseases from its leaves using image processing. In *Smart Intelligent Computing and Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 215–228.

64. Gurrala, K.K.; Yemineni, L.; Rayana, K.S.R.; Vajja, L.K. A new segmentation method for plant disease diagnosis. In Proceedings of the 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 28−29 September 2019; pp. 137–141.

65. Karthikeyan, N.; Anjana, M.; Anusha, S.; Divya, R.; Vinod, A. Leaf disease detection using image processing. *Int. J. Innov. Sci. Eng. Technol.* **2020**, *7*. [CrossRef]

66. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]

67. Usman, A.; Bukht, T.F.N.; Ahmad, R.; Ahmad, J. Plant disease detection using internet of thing (IoT). *Plant Dis.* **2020**, *11*, 505–509.

68. Reddy, K.A.; Reddy, N.M.C.; Sujatha, S. Precision method for pest detection in plants using the clustering algorithm in image processing. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 28−30 July 2020; pp. 894–897.

69. Khan, M.A. Detection and classification of plant diseases using image processing and multiclass support vector machine. *Int. J. Comput. Trends Technol.* **2020**, *68*, 5–11. [CrossRef]

70. Sawant, C.; Shirgaonkar, M.; Khule, S.; Jadhav, P. Plant disease detection using image processing techniques. 2020. Available online: https://www.semanticscholar.org/paper/Plant-Disease-Detection-using-Image-Processing-Sawant-Shirgaonkar/d9a26b87f1879821fea0cd1944279cd51359d0c5 (accessed on 15 December 2021).

71. Fulari, U.N.; Shastri, R.K.; Fulari, A.N. Leaf disease detection using machine learning. *J. Seybold Rep.* **2020**, *1533*, 9211.

72. Ouhami, M.; Es-Saady, Y.; El Hajji, M.; Hafiane, A.; Canals, R.; El Yassa, M. Deep transfer learning models for tomato disease detection. In Proceedings of the International Conference on Image and Signal Processing, Marrakesh, Morocco, 4–6 June 2020; pp. 65–73.

73. Deepa. A pre processing approach for accurate identification of plant diseases in leaves. In Proceedings of the 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), Mysuru, India, 14–15 December 2018; pp. 249–252.

74. Win, T.T. AI and IoT Methods for Plant Disease Detection in Myanmar. Kobe Institute of Computing: Kobe, Tokyo, 2018.

75. Sun, G.; Jia, X.; Geng, T. Plant diseases recognition based on image processing technology. *J. Electr. Comput. Eng.* **2018**, *2018*, 6070129. [CrossRef]

76. Thorat, A.; Kumari, S.; Valakunde, N.D. An IoT based smart solution for leaf disease detection. In Proceedings of the 2017 International Conference on Big Data, IoT and Data Science (BID), Pune, India, 20–22 December 2017; pp. 193–198.

77. Moghadam, P.; Ward, D.; Goan, E.; Jayawardena, S.; Sikka, P.; Hernandez, E. Plant disease detection using hyperspectral imaging. In Proceedings of the 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia, 29 November–1 December 2017; pp. 1–8.

78. Khirade, S.D.; Patil, A. Plant disease detection using image processing. In Proceedings of the 2015 International Conference on Computing Communication Control and Automation, Pune, India, 26−27 February 2015; pp. 768–771.

79. Arivazhagan, S.; Shebiah, R.N.; Ananthi, S.; Varthini, S.V. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agric. Eng. Int. CIGR J.* **2013**, *15*, 211–217.

80. Han, S.; Cointault, F. Détection précoce de maladies sur feuilles par traitement d'images. HAL Open Sci. Available online: https://hal.archives-ouvertes.fr/hal-00829402/document (accessed on 15 December 2021).

81. Revathi, P.; Latha, M.H. *Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques*; IEEE: Piscataway, NJ, USA, 2012; pp. 169–173. [CrossRef]

82. Yerpude, A.; Dubey, S. Colour image segmentation using K-medoids clustering. *Int. J. Comput. Technol. Appl.* **2012**, *3*, 152–154.

83. Chaudhary, P.; Chaudhari, A.K.; Cheeran, A.; Godara, S. Color transform based approach for disease spot detection on plant leaf. *Int. J. Comput. Sci. Telecommun.* **2012**, *3*, 65–70.

84. Al-Hiary, H.; Bani-Ahmad, S.; Reyalat, M.; Braik, M.; Alrahamneh, Z. Fast and accurate detection and classification of plant diseases. *Int. J. Comput. Appl.* **2011**, *17*, 31–38. [CrossRef]

85. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W. Early detection of Fusarium infection in wheat using hyper-spectral imaging. *Comput. Electron. Agric.* **2011**, *75*, 304–312. [CrossRef]

86. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, 1–11. [CrossRef]

87. Panigrahi, K.P.; Das, H.; Sahoo, A.K.; Moharana, S.C. Maize leaf disease detection and classification using machine learning algorithms. In *Progress in Computing, Analytics and Networking*; Springer: Singapore, 2020; pp. 659–669.

88. Zhao, J.; Xu, C.; Xu, J.; Huang, L.; Zhang, D.; Liang, D. Forecasting the wheat powdery mildew (Blumeria graminis f. Sp. tritici) using a remote sensing-based decision-tree classification at a provincial scale. *Australas. Plant Pathol.* **2018**, *47*, 53–61. [CrossRef]

89. Zhang, J.; Pu, R.; Yuan, L.; Huang, W.; Nie, C.; Yang, G. Integrating remotely sensed and meteorological observations to forecast wheat powdery mildew at a regional scale. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4328–4339. [CrossRef]

90. Fenu, G.; Malloci, F.M. Artificial intelligence technique in crop disease forecasting: A case study on potato late blight prediction. In Proceedings of the International Conference on Intelligent Decision Technologies, Split, Croatia, 17–19 June 2020; pp. 79–89.

91. Fenu, G.; Malloci, F.M. An application of machine learning technique in forecasting crop disease. In Proceedings of the 2019 3rd International Conference on Big Data Research, Paris, France, 20–22 November 2019; pp. 76–82.

92. Hsieh, J.-Y.; Huang, W.; Yang, H.-T.; Lin, C.-C.; Fan, Y.-C.; Chen, H. *Building the Rice Blast Disease Prediction Model Based on Machine Learning and Neural Networks*; EasyChair: Manchester, UK, 2019.

93. Kim, Y.; Roh, J.-H.; Kim, H.Y. Early forecasting of rice blast disease using long short-term memory recurrent neural networks. *Sustainability* **2018**, *10*, 34. [CrossRef]

94. Duarte-Carvajalino, J.M.; Alzate, D.F.; Ramirez, A.A.; Santa-Sepulveda, J.D.; Fajardo-Rojas, A.E.; Soto-Suárez, M. Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sens.* **2018**, *10*, 1513. [CrossRef]

95. Mahlein, A.-K.; Kuska, M.T.; Behmann, J.; Polder, G.; Walter, A. Hyperspectral sensors and imaging technologies in phytopathology: State of the art. *Annu. Rev. Phytopathol.* **2018**, *56*, 535–558. [CrossRef]

96. Kadir, A.; Nugroho, L.E.; Susanto, A.; Santosa, P.I. Performance improvement of leaf identification system using principal component analysis. *Int. J. Adv. Sci. Technol.* **2012**, *44*, 113–124.

97. Turkoglu, M.; Hanbay, D. Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts. *Appl. Math. Comput.* **2019**, *352*, 1–14. [CrossRef]

98. Ngugi, L.C.; Abelwahab, M.; Abo-Zahhad, M. Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. *Inf. Process. Agric.* **2021**, *8*, 27–51. [CrossRef]

99. Ahmed, K.; Shahidi, T.R.; Alam, S.M.I.; Momen, S. Rice leaf disease detection using machine learning techniques. In Proceedings of the 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 24–25 December 2019; pp. 1–5.

100. Barbedo, J.G. Factors influencing the use of deep learning for plant disease recognition. *Biosyst. Eng.* **2018**, *172*, 84–91. [CrossRef]

101. Sambasivam, G.; Opiyo, G.D. A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. *Egypt. Inform. J.* **2021**, *22*, 27–34. [CrossRef]

102. Hochreiter, S. Untersuchungen zu dynamischen neuronalen Netzen. Master Thesis, Technische Universität München, München, Germany, 1991.

103. Or, B. The Exploding and Vanishing Gradients Problem in Time Series. Available online: https://towardsdatascience.com/the-exploding-and-vanishing-gradients-problem-in-time-series-6b87d558d22 (accessed on 15 December 2021).

104. Hasan, R.I.; Yusuf, S.M.; Alzubaidi, L. Review of the state of the art of deep learning for plant diseases: A broad analysis and discussion. *Plants* **2020**, *9*, 1302. [CrossRef]

105. Ahmad, M.; Abdullah, M.; Moon, H.; Han, D. Plant disease detection in imbalanced datasets using efficient convolutional neural networks with stepwise transfer learning. *IEEE Access* **2021**, *9*, 140565–140580. [CrossRef]

106. Bock, C.; Cook, A.; Parker, P.; Gottwald, T. Automated image analysis of the severity of foliar citrus canker symptoms. *Plant Dis.* **2009**, *93*, 660–665. [CrossRef] [PubMed]

107. Pydipati, R.; Burks, T.; Lee, W. Identification of citrus disease using color texture features and discriminant analysis. *Comput. Electron. Agric.* **2006**, *52*, 49–59. [CrossRef]

108. Guo, W.; Rage, U.K.; Ninomiya, S. Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Comput. Electron. Agric.* **2013**, *96*, 58–66. [CrossRef]

109. Pourreza, A.; Lee, W.S.; Raveh, E.; Ehsani, R.; Etxeberria, E. Citrus Huanglongbing detection using narrow-band imaging and polarized illumination. *Trans. ASABE* **2014**, *57*, 259–272.

110. Zhou, R.; Kaneko, S.I.; Tanaka, F.; Kayamori, M.; Shimizu, M. Image-based field monitoring of Cercospora leaf spot in sugar beet by robust template matching and pattern recognition. *Comput. Electron. Agric.* **2015**, *116*, 65–79. [CrossRef]

111. Olmstead, J.W.; Lang, G.A.; Grove, G.G. Assessment of severity of powdery mildew infection of sweet cherry leaves by digital image analysis. *HortScience* **2001**, *36*, 107–111. [CrossRef]

112. Moya, E.; Barrales, L.; Apablaza, G. Assessment of the disease severity of squash powdery mildew through visual analysis, digital image analysis and validation of these methodologies. *Crop. Prot.* **2005**, *24*, 785–789. [CrossRef]

113. Oberti, R.; Marchi, M.; Tirelli, P.; Calcante, A.; Iriti, M.; Borghese, A.N. Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity. *Comput. Electron. Agric.* **2014**, *104*, 1–8. [CrossRef]

114. Ahmad, I.S.; Reid, J.F.; Paulsen, M.R.; Sinclair, J.B. Color classifier for symptomatic soybean seeds using image processing. *Plant Dis.* **1999**, *83*, 320–327. [CrossRef] [PubMed]

115. Wiwart, M.; Fordoński, G.; Żuk-Gołaszewska, K.; Suchowilska, E. Early diagnostics of macronutrient deficiencies in three legume species by color image analysis. *Comput. Electron. Agric.* **2009**, *65*, 125–132. [CrossRef]

116. Camargo, A.; Smith, J. An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosyst. Eng.* **2009**, *102*, 9–21. [CrossRef]

117. Liu, J.; Wang, X. Plant diseases and pests detection based on deep learning: A review. *Plant Methods* **2021**, *17*, 1–18. [CrossRef]

118. Hughes, D.; Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv* **2015**, arXiv:1511.08060.

119. Cucci, C.; Casini, A. Hyperspectral imaging for artworks investigation. In *Data Handling in Science and Technology*; Elsevier: Amsterdam, The Netherlands, 2020; Volume 32, pp. 583–604.

120. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 1–74.

121. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 1–48. [CrossRef]

122. Laurent, C.; Pereyra, G.; Brakel, P.; Zhang, Y.; Bengio, Y. Batch normalized recurrent neural networks. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2657–2661.

123. Zhang, G.; Wang, C.; Xu, B.; Grosse, R. Three mechanisms of weight decay regularization. *arXiv* **2018**, arXiv:1810.12281.

124. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.

125. Maaten, L.; Chen, M.; Tyree, S.; Weinberger, K. Learning with marginalized corrupted features. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 410–418.

126. Pereyra, G.; Tucker, G.; Chorowski, J.; Kaiser, Ł.; Hinton, G. Regularizing neural networks by penalizing confident output distributions. *arXiv* **2017**, arXiv:1701.06548.

127. Kadam, S.; Vaidya, V. Review and analysis of zero, one and few shot learning approaches. In Proceedings of the International Conference on Intelligent Systems Design and Applications, Vellore, India, 6–8 December 2018; pp. 100–112.

128. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [CrossRef]

129. Chatterjee, S.K.; Malik, O.; Gupta, S. Chemical sensing employing plant electrical signal response-classification of stimuli using curve fitting coefficients as features. *Biosensors* **2018**, *8*, 83. [CrossRef]