Smart Agricultural Technology 3 (2023) 100083

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools

Aanis Ahmad [a,*], Dharmendra Saraswat [b], Aly El Gamal [a]

[a] Elmore Family School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana, USA
[b] Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana, USA

ARTICLE INFO

Keywords:
Plant diseases
Deep learning
Precision agriculture
Generalization
Review
Survey

ABSTRACT

Several factors associated with disease diagnosis in plants using deep learning techniques must be considered to develop a robust system for accurate disease management. A considerable number of studies have investigated the potential of deep learning techniques for precision agriculture in the last decade. However, despite the range of applications, several gaps within plant disease research are yet to be addressed to support disease management on farms. Thus, there is a need to establish a knowledge base of existing applications and identify the challenges and opportunities to help advance the development of tools that address farmers' needs. This study presents a comprehensive overview of 70 studies on deep learning applications and the trends associated with their use for disease diagnosis and management in agriculture. The studies were sourced from four indexing services, namely Scopus, IEEE Xplore, Science Direct, and Google Scholar, and 11 main keywords used were Plant Diseases, Precision Agriculture, Unmanned Aerial System (UAS), Imagery Datasets, Image Processing, Machine Learning, Deep Learning, Transfer Learning, Image Classification, Object Detection, and Semantic Segmentation. The review is focused on providing a detailed assessment and considerations for developing deep learning-based tools for plant disease diagnosis in the form of seven key questions pertaining to (i) dataset requirements, availability, and usability, (ii) imaging sensors and data collection platforms, (iii) deep learning techniques, (iv) generalization of deep learning models, (v) disease severity estimation, (vi) deep learning and human accuracy comparison, and (vii) open research topics. These questions can help address existing research gaps by guiding further development and application of tools to support plant disease diagnosis and provide disease management support to farmers.

1. Introduction

Plant diseases are responsible for yield losses that directly impact national and global food production systems, resulting in economic losses. According to the Food and Agriculture Organization (FAO), plant diseases and pests are responsible for 20% to 40% loss in global food production (Food and Agriculture Organization of the United Nations International Plant Protection Convention, 2017). Plant diseases are responsible for an estimated 13% of global crop yield loss [77]. These statistics highlight the importance of identifying plant diseases to mitigate yield losses. However, first, it is crucial to understand the factors responsible for plant diseases.

Three factors aid disease formation in plants: the host, a favorable environment, and the pathogen. These factors create the plant disease triangle shown in Fig. 1. In most cases, diseases begin to show symptoms and affect the plant from the bottom up. Many plant diseases spread throughout the crop after infection. Therefore, crops need to be monitored regularly as early management of the disease will help prevent their spread [26]. In several cases, plant diseases also appear later in the season after pollination. Diseases in plants are of different types and affect different plant organs. Plant diseases that show symptoms on leaves, i.e., foliar diseases, show the most distinguishable features that plant pathologists can identify via visual inspection. In particular, fungal diseases are responsible for up to 50% of yield losses [82]. Therefore, most modern studies use images of plant leaves to identify diseases using computer vision, machine learning, and deep learning techniques. Effective plant disease diagnosis will comprise early-season plant disease identification [48,105], identification of multiple diseases in different crops and multiple simultaneous diseases [12], estimating the severity of the disease [116], estimating the appropriate volume of

* Corresponding author at: Purdue University, West Lafayette, IN, 47906, USA.
E-mail address: ahmad31@purdue.edu (A. Ahmad).

https://doi.org/10.1016/j.atech.2022.100083
Received 10 May 2022; Received in revised form 9 June 2022; Accepted 14 June 2022
Available online 16 June 2022

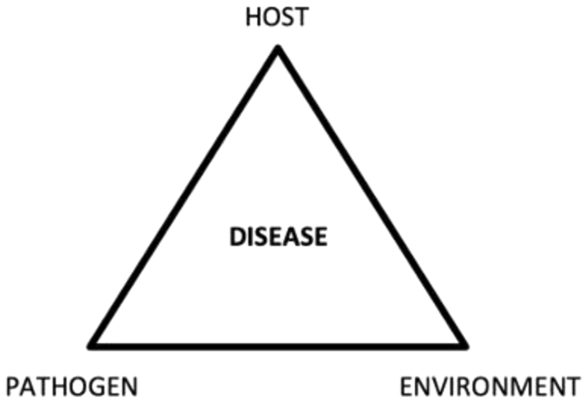

Fig. 1. Plant Disease Triangle.

pesticide to apply [110,114], and useful steps to take for managing disease to limit its spread [105].

Plant disease identification is necessary for precision agriculture and plant phenotyping. Both these fields are data, information, and technology-intensive. Conventional plant disease diagnosis and monitoring methods are expensive, time-consuming, rely on experts, and therefore ineffective for precision agriculture as they involve a manual visual inspection. Moreover, these approaches are likely to be affected by human bias and fatigue, resulting in reduced accuracy [15,21,52]. To overcome these unreliable methods for disease identification, studies have investigated the application of image processing techniques using plant images. An automated disease assessment device using black and white images and videos of leaves belonging to potted tomato and blackened fern plants was one of the earliest studies that utilized traditional image processing for plant diseases in 1983 [11,58]. Image processing techniques were also used to quantify streak disease in corn, and it was reported that computer-based methods were more accurate than traditional visual analysis [68]. Reliance on traditional image processing techniques has become popular over the past 30 years as they can help objectively diagnose diseases [69]. However, these techniques require manual feature extraction, which is time-consuming and subject to personal bias as different studies may consider different features more important.

Machine learning started to gain interest for plant disease identification approximately 20 years ago when its applications were discussed and studies were reviewed for agriculture and plant diseases [118]. Traditional machine learning techniques were common in the research community for identifying plant diseases. Some traditional machine learning techniques that were used for plant disease identification include support vector machines (SVM) for tomato disease identification [73], random forests for tomato disease identification [37], and K-nearest neighbors (KNN) for soybean disease identification [95]. Machine learning techniques were also used to identify the disease and estimate severity [45]. In addition, SVM, KNN, and Naïve Bayes were used for tomato powdery mildew disease identification [19]. A tomato powdery mildew disease forecasting technique was also proposed [18]. However, traditional machine learning techniques still relied on extracting important features for training models [96], which were time-consuming. Additionally, traditional image processing and machine learning were only successful under specific conditions [76]. Therefore, studies within the last decade started to incorporate deep learning techniques that inherently carry out automatic feature extraction and achieved higher accuracies than traditional approaches [1,50].

Deep learning, a subset of machine learning, has become a preferred approach for disease identification due to increased computational power, storage capacities, and availability of large datasets. Many researchers across different disciplines adopted deep learning after the ILSVR ImageNet competition in 2012 [53]. Common deep learning techniques include using Convolutional Neural Networks (CNN) for image classification, object detection, and semantic segmentation. However, deep learning is data-hungry and relies on large datasets consisting of hundreds and thousands of images [67]. For plant disease identification, deep learning using CNN has become a hot spot research area [59] after the creation of the PlantVillage dataset in 2015 [42]. PlantVillage has become one of the most popular datasets that was used for disease identification [72], severity estimation [56], and development of different management systems [100]. Recently, multiple additional plant disease datasets were made publicly available for training deep learning models, including the Digipathos dataset [16], the PlantDoc dataset [99], the Northern Leaf Blight (NLB) dataset [117], the RoCoLe coffee disease dataset [79], the rice disease dataset [83], and the cassava disease dataset [78]. Many studies have used publicly available datasets for training deep learning models for identifying diseases in different crops to tackle the problem of yield reduction. However, some studies have also used custom datasets. The use of deep learning techniques for plant disease diagnosis provides multiple advantages, including separating disease symptoms, identifying multiple diseases and instances, estimating disease severity, and developing cost-effective solutions [52]. A timeline for deep learning architectures, datasets, and important studies is shown in Fig. 2.

The correct sensors and platforms must first be identified to acquire datasets with a good representation of plant diseases. Most recent studies for plant disease identification have relied on RGB imagery; however, deep learning techniques have also been used on images acquired from multispectral and hyperspectral sensors to identify regions of a field where diseases were present. Thermal sensors and fluorescent sensors have also been used. Researchers have also relied on different data collection platforms, including handheld sensors, fixed platforms, and unmanned aerial systems (UAS).

In this review, we have evaluated 70 studies in which plant disease datasets were created, plant disease identification and disease severity estimation models were developed, different imaging sensors on various data collection platforms were used, and different gaps in the research area were tackled. The studies were sourced using four indexing services, namely Scopus, IEEE Xplore, Science Direct, and Google Scholar with 11 main keywords, namely Plant Diseases, Precision Agriculture, UAS, Imagery Datasets, Image Processing, Machine Learning, Deep Learning, Transfer Learning, Image Classification, Object Detection, and Semantic Segmentation. In particular, this review provides the advancements in plant disease identification over time, along with the development of different datasets, the usage of different imaging sensors, and data collection platforms. Limitations of existing research are also discussed to help identify a research direction for creating a robust deep learning-based real-time disease management system for farmers.

Although multiple reviews and surveys have been conducted for plant disease identification [12,22,34,54,61,63,66,98], to the best of our knowledge, a comprehensive study that is focused on evaluating disease datasets for deep learning, generalization of deep learning models, answering research-specific questions, and identifying open research gaps, have not been conducted.

In this study, each of the following sections was organized to answer important research questions for plant disease identification:

- Which publicly available datasets are commonly used for plant diseases?
- What are the different imaging sensors and data collection platforms that have been used for disease image acquisition?
- What deep learning techniques have been used for plant disease identification?
- How well have deep learning models generalized to unseen testing data, different datasets, and field images?
- How was disease severity estimated?
- Can deep learning outperform humans for plant disease identification?

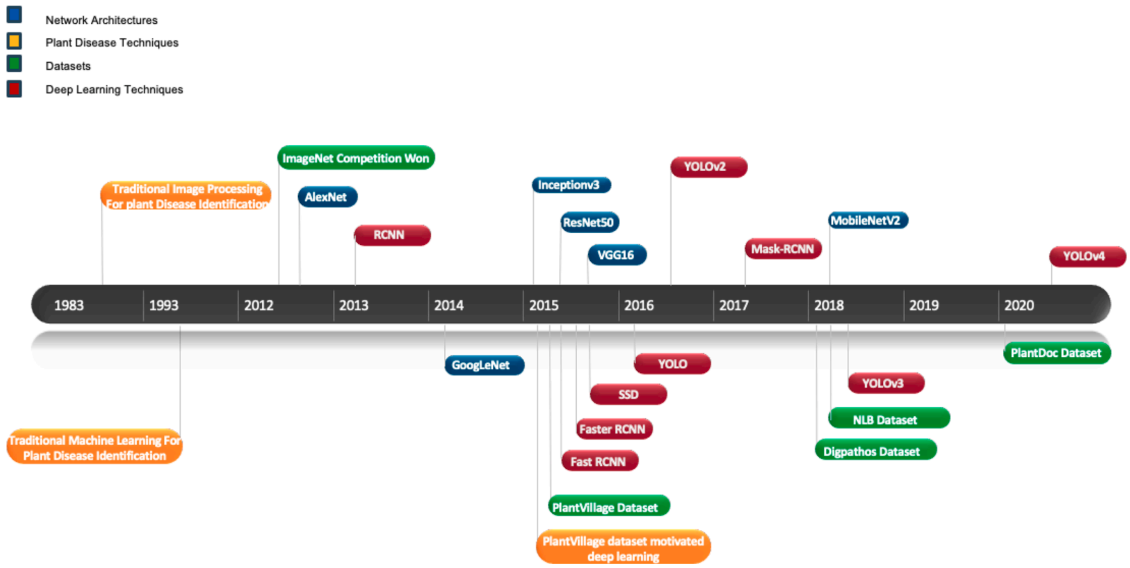

Fig. 2. Plant Disease Identification Research Timeline.

- What are some open research topics?

2. What publicly available datasets are commonly used for plant diseases?

To develop algorithms and train computers to identify plant diseases from acquired images, representative datasets are required to identify plant diseases. Large datasets with hundreds or thousands of images are required for deep learning-based solutions. In Section 2, multiple publicly available datasets were summarized in the reviewed studies. The distribution of studies that have used plant disease images from different datasets is shown in Fig. 3. It was observed that the most popular dataset that has been used is the PlantVillage dataset. However, most researchers from the reviewed studies have opted to acquire custom datasets not made available to the public.

2.1. PlantVillage

Since the introduction of the PlantVillage dataset [42], it has become the most commonly used dataset for training and developing deep learning-based plant disease identification and severity estimation models [22]. The PlantVillage dataset consists of a total of 54,309 images. The images are divided across 38 different diseases affecting 14 crops. Most of the images were acquired under controlled lab conditions with uniform backgrounds, as shown in Fig. 4. As images within the PlantVillage dataset are not representative of real-field conditions, deep learning models trained to identify diseases using these images could not generalize with higher accuracy to images acquired under field conditions [30]. Furthermore, the dataset is unbalanced, and a summary is shown in Table 1. After visually inspecting the images within the PlantVillage dataset, it was also observed that some of the diseases from different classes overlap. For example, some of the Gray Leaf Spot (GLS) disease images (Fig. 4) also contain NLB lesions, resulting in an erroneous performance for deep learning models.

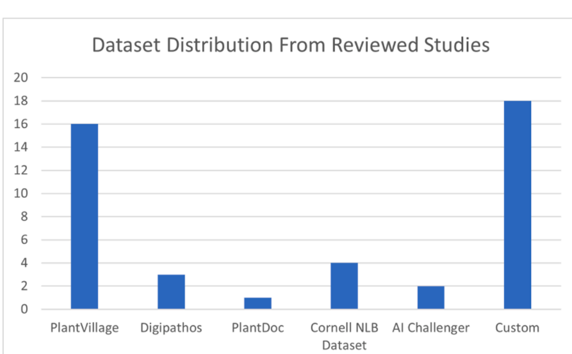

Fig. 3. Number of studies conducted for identifying diseases in different crops using publicly available and custom datasets.

Fig. 4. Gray Leaf Spot Disease of Corn from the PlantVillage Dataset.

2.2. Digipathos

Another large plant disease dataset called Digipathos was recently introduced, and it consists of 46,513 images across 171 diseases affecting 21 different crops [16]. Only 2326 of the images in this dataset represent diseases of leaves acquired under controlled lab conditions with uniform backgrounds, as shown in Fig. 5 (a), or under field conditions with complex backgrounds, as shown in Fig. 5 (b). The remaining 44,187 images consist of cropped disease lesions, as shown in Fig. 5 (c).

Table 1
PlantVillage Dataset Summary.

Crop	Disease	Images
Apple	Healthy	1,645
	Black Rot	621
	Cedar Apple Rust	275
	Apple Scab	630
Blueberry	Healthy	1,502
Cherry	Healthy	854
	Powdery Mildew	1,052
Corn	Healthy	1,162
	Grey Leaf Spot	513
	Common Rust	1,192
	Northern Leaf Blight	985
Grape	Healthy	423
	Black Rot	1,180
	Black Measles	1,383
	Isariopsis Leaf Spot	1,076
Orange	Citrus Greening	5,507
Peach	Healthy	360
	Bacterial Spot	2,297
Bell Pepper	Healthy	1,478
	Bacterial Spot	997
Potato	Healthy	152
	Early Blight	1,000
	Late Blight	1,000
Raspberry	Healthy	371
Soybean	Healthy	5,090
Squash	Powdery Mildew	1,835
Strawberry	Healthy	456
	Leaf Scorch	1,109
Tomato	Healthy	1,592
	Bacterial Spot	2,127
	Early Blight	1,000
	Late Blight	1,909
	Leaf Mold	952
	Septoria Leaf Spot	1,771
	Spider Mites	1,676
	Target Spot	1,404
	Yellow Leaf Curl	5,357
	Mosaic Virus	373

Therefore, it is clear that more than 95% of the images in this dataset do not represent images of plant diseases under real-field conditions. Nevertheless, the Digipathos dataset has been used to identify diseases in multiple studies, and the dataset can be found at the following link: https://www.digipathos-rep.cnptia.embrapa.br.

2.3. PlantDoc

A smaller plant disease dataset called PlantDoc was recently acquired, and it consists of 2598 images across 17 diseases affecting 13 different crops [99]. Most images were acquired under field conditions, as shown in Fig. 6 (a); however, some images also consist of uniform backgrounds, as in Fig. 6 (b). Variation in image acquisition conditions can lead to training robust deep learning-based disease identification models. However, some images within this dataset consist of multiple diseased leaves or entire crops (Fig. 6(c)), limiting the ability of deep learning models to learn important disease features. Additionally, the PlantDoc dataset is unbalanced and consists of a limited number of images per class (i.e., less than 200 images). A summary of the PlantDoc dataset is shown in Table 2. As this dataset is far from perfect, it is difficult to implement accurate deep learning models for disease identification. Therefore, its use in the research community is limited, as it was used by only two of the reviewed studies in this survey for plant disease identification [6,24].

2.4. NLB Dataset

A dataset consisting of 18,222 field images of Northern Leaf Blight (NLB) infected corn, a common foliar disease of corn, was acquired using UAS-based aerial imagery, handheld imagery, and mounting a camera on a boom [117]. The NLB dataset consists of real-field images with 105, 735 lesion annotations, as shown in Table 3 below. This dataset only consists of corn leaf images with a single disease, it cannot be used to identify multiple diseases. Therefore, the best application for this dataset lies in training deep learning models to discriminate between healthy and diseased corn plants or for object detection to identify and locate NLB lesions. Furthermore, this dataset can provide testing images to test the generalization capability of deep learning models to identify diseases across datasets. Examples of UAS-based, handheld, and boom images can be seen in Fig. 7 below.

2.5. RoCoLe: coffee disease dataset

Another crop-specific annotated dataset for coffee diseases with severity information was acquired [79]. The dataset consists of a total of 1560 images acquired using a 5 MP smartphone camera in a field occupied by 390 coffee plants in Ecuador. Four images were acquired from each coffee plant, and each image was annotated with ground truth using an open-source tool. A summary of the dataset is shown in Table 4.

2.6. Rice disease dataset

A rice disease dataset consisting of 3355 images was also acquired [83]. A total of four disease classes are present, namely healthy, brown spot, hispa, and leaf blast (Table 5). The dataset, however, consists of images acquired under controlled lab conditions with uniform white backgrounds, as shown in Fig. 8. Therefore, such a dataset will make it difficult to train robust deep learning models capable of identifying rice diseases under field conditions.

2.7. Cassava disease dataset

Another dataset consisting of cassava disease images was recently

Fig. 5. (a) NLB of corn in Digipathos Dataset with Controlled Background (b) NLB of corn in Digipathos Dataset under real-field conditions (c) NLB Lesion of Corn from Digipathos Dataset.

Fig. 6. (a) NLB of corn in PlantDoc dataset with uniform background (b) NLB of corn in PlantDoc dataset with field conditions (c) Entire corn plants infected with NLB in PlantDoc dataset.

Table 2
PlantDoc Dataset Summary.

Crop	Disease	Images
Apple	Healthy	91
	Scab	93
	Rust	89
Bell Pepper	Healthy	61
	Leaf Spot	71
Blueberry	Healthy	117
Cherry	Healthy	57
Corn	Leaf Blight	192
	Grey Leaf Spot	68
	Rust	116
Grape	Healthy	69
	Black Rot	64
Peach	Healthy	112
Potato	Early Blight	117
	Late Blight	105
Raspberry	Healthy	119
Soybean	Healthy	65
Squash	Powdery Mildew	130
Strawberry	Healthy	96
Tomato	Healthy	63
	Bacterial Spot	110
	Early Blight	88
	Late Blight	111
	Leaf Mold	91
	Septoria Leaf Spot	151
	Mosaic Virus	54
	Yellow Virus	76
	Spider Mite	2

Table 3
NLB Disease Dataset Summary.

Platform	Images	Annotations
Drone	7,669	42,117
Handheld	1,787	7,669
Boom	8,766	55,919

acquired under field conditions [78]. The dataset consists of 5656 images for five different disease classes: healthy, cassava bacterial blight, cassava brown streak disease, cassava green mite, and cassava mosaic disease (Table 6). The dataset was acquired under field conditions and consisted of complex backgrounds, as shown in Fig. 9. Therefore, this dataset may be used to train models capable of identifying cassava diseases in field fields.

2.8. CD&S (corn disease & severity) dataset

The Corn Disease and Severity (CD&S) dataset that we acquired under field conditions was also made available to the public [5]. The dataset consists of 4455 images for three common corn diseases, namely NLB, GLS, and Northern Leaf Spot (NLS), for training deep learning models capable of identifying different corn disease types under field conditions. For the images within the training set, three additional versions of the original dataset were created by removing the background, replacing the field background with a black background, and

Table 4
RoCoLe: Coffee Disease Dataset Summary.

Class	Images
Healthy	791
Red Spider Mite	167
Rust Level 1	344
Rust Level 2	166
Rust Level 3	62
Rust Level 4	30

Table 5
Rice Disease Dataset Summary.

Class	Images
Healthy	791
Brown Spot	167
Hispa	344
Leaf Blast	166

Fig. 7. (a) Handheld image (b) Boom acquired image (c) UAS acquired image.

Fig. 8. Rice disease sample with white uniform backgrounds.

Table 6
Cassava Disease Dataset Summary.

Class	Images
Healthy	316
Bacterial Blight	466
Brown Streak	1,443
Green Mite	773
Mosaic Disease	2,658

Fig. 9. Cassava disease sample with complex field backgrounds.

replacing the field background with a white background (Fig. 10). In addition, images for five different severity levels for the NLS disease are available for training severity estimation models. The CD&S dataset was recently used [4]. A summary is shown in Table 7.

3. What are the different imaging sensors and data collection platforms that have been used for disease image acquisition?

3.1. Imaging sensors

The most commonly used imaging sensors are red, green, and blue (RGB) sensors due to their low costs, ease of use, and accessibility. Therefore, RGB sensors can acquire important features to help identify diseased plants [33]. Most studies reviewed for the interest of this study used RGB sensors to acquire images to train deep learning models for implementing solutions capable of identifying plant diseases with high accuracy. As discussed in Section 2, most publicly available datasets also consist of RGB imagery.

Multispectral sensors are imaging sensors capable of extracting specific bands such as red, green, blue, near-infrared, and red-edge. The images acquired are then stacked. High-stress levels in plants can easily be picked up by multispectral sensors and can be identified easily as stress increases visible reflectance. Popular bands for stress detection include red, green, near-infrared, and red-edge [33]. These sensors were used for detecting plant diseases, weeds, water stress levels, nutrient deficiencies, etc. Plant disease identification or discrimination was studied using multispectral sensors ([121]b; [115]). Although multispectral sensors could discriminate diseased regions from healthy regions in crops, a limiting factor of using these sensors is attributed to the limited number of bands compared to hyperspectral sensors. It poses a challenge when identifying different types of diseases in multiple crops. Therefore, multispectral sensors may be most useful for providing information regarding the regions of a field where crop health is affected. However, a sensor that can pick up more bands may be more useful for distinguishing between multiple diseases.

Hyperspectral sensors are expensive sensors capable of capturing hundreds of different bands and picking up different signatures from multiple diseases. These sensors were used to acquire data for deep learning-based analysis ([121]b; [115]). In addition, hyperspectral sensors help acquire rich spectral information about plants that may aid in detecting diseases before the appearance of visual symptoms [115]. However, studies using hyperspectral images for deep learning are limited due to high costs and memory requirements.

Fluorescence and thermal sensors can also provide useful information regarding plant diseases. These sensors can provide information about the photosynthetic ability of different plants and help identify diseased regions with varying heat signatures. Healthier plants will have the greater photosynthetic ability, while diseased and stressed plants will be limited in their ability to conduct photosynthesis. Fluorescence sensors can pick up these signatures.

For plant disease identification, a recent study discussed RGB, multispectral, hyperspectral, fluorescence, and thermal sensors [66]. As distance plays an important role in spatial resolution accuracy, extra care is required when mounting spectral sensors on UAS systems. In addition, it was noted that spectral properties could help identify different plant disease patterns, which could ultimately help in plant disease identification. Thermal sensors can also provide useful information regarding plant diseases as thermal sensors help identify the temperature of plants with water status and transpiration. However, thermal sensors are subject to environmental factors. It was reported that fluorescence sensors might be useful for discriminating and identifying fungal diseases. However, the use of fluorescence sensors requires controlled conditions. The main conclusion was that thermal and fluorescence sensors are useful for identifying diseases early on before visual symptoms appear. However, they cannot identify certain diseases. Therefore, hyperspectral or RGB sensors will be preferable in such cases. Different imaging sensors were also reviewed for plant stress identification. In particular, RGB, multispectral, hyperspectral, fluorescence, and LIDAR sensors were reviewed and discussed [33]. As RGB sensors are more accessible and most publicly available datasets consist of RGB imagery, deep learning-based models trained using RGB imagery for plant disease management systems are currently preferable.

3.2. Data collection platform

Once an image acquisition sensor has been chosen for the task, it is important to decide on a data collection platform. Multiple studies

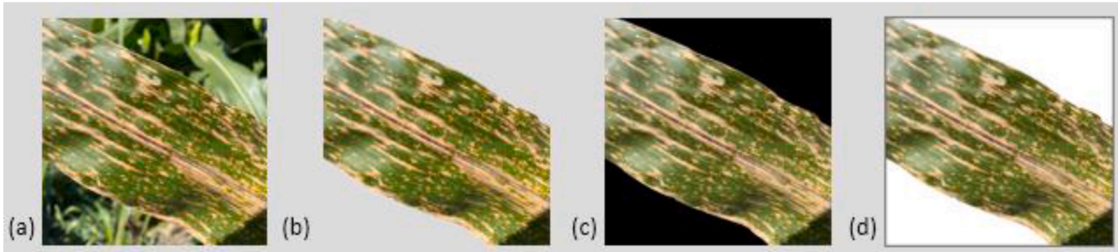

Fig. 10. (a) Original GLS Disease of Corn acquired under field conditions. (b) Same GLS image with the removed background. (c) Same image with a black background. (d) Same image with a white background.

Table 7
CD&S Dataset Summary.

Class	Images
Northern Leaf Blight	1,237
Gray Leaf Spot	1,313
Northern Leaf Spot	1,905

within this review have used different methods to acquire datasets for deep learning-based plant disease identification.

One common approach that researchers have utilized is the use of online resources to compile datasets. The PlantDoc dataset discussed in Section 2 consists of carefully selected plant disease images from Google and Ecosia [99]. Diseases were also identified using corn disease images via Google and corn disease images from the PlantVillage dataset [122]. Images from publicly available datasets, Google, and personally acquired datasets were then used for disease identification in corn and peach [94]. Although images from online sources can be a quick dataset acquisition solution for researchers with limited access to crops, such datasets introduce unwanted noise due to climate, location, image size, sensor quality, etc. Therefore, it is more common for researchers to acquire data over a given field in specified locations using different platforms.

A common approach for acquiring datasets is by using a handheld sensor. A handheld imaging platform was used to acquire a small dataset of 300 images for six diseases. An 8 MP iPad camera was used to acquire images of plant diseases from 30 cm above the plant. The images were acquired in the shade with no flash to reduce the effect of reflection [81]. Multiple smartphone cameras, namely ASUS Zenfone 2, Xiaomi Redmi 5A, Xiaomi S2, Galaxy S8, and iPhone 6s, were used for acquiring handheld images of coffee leaf diseases [28]. Images were acquired from the underside of the leaves in Espirito Santo, Brazil. The conditions for acquiring the images were somewhat controlled with white backgrounds. Some additional studies also used handheld imaging platforms [59]. Handheld sensors ensure that each image is ideal for training deep learning models; however, it is time-consuming to manually scout a field and acquire individual leaf images. Furthermore, handheld imagery comes at the cost of controlling the height and angle of the sensor, which can result in highly variable data. To tackle this limitation presented by handheld images, cameras and sensors are mounted on booms to control factors such as height and angle.

Mounting a sensor on a boom or fixed autonomous system can help acquire a more reliable dataset. A controlled imaging platform with a boom was used to acquire 1070 hyperspectral images of charcoal rot disease using the Pika XC Hyperspectral sensor under lab conditions [74]. An autonomous imaging platform was also created using the Motorman 5L Robotic Manipulator to pan a V10E Specim ImSpector Hyperspectral sensor for acquiring images [115]. Similarly, a PTZ Automated System was created to acquire 584 images of onion diseases. The system was placed 10 meters from the cultivation area, and images were acquired every 30 minutes between 7:00 and 19:00 from April 20 to May 20, 2017. The soil temperature of the onion fields was between 25 and 28 degrees Celsius [52]. However, in most conditions, the images are acquired under lab conditions, or an observer is still required to scout the field and manually capture the images. To overcome these issues, modern-day researchers are adapting to UAS imagery.

UAS imagery helps farmers implement efficient disease management practices [75,106]. A UAS imaging platform consisting of a DJI Phantom Pro 4 mounted with a Canon EOS 80 DSLR at 10 meters altitude to acquire 120 images [41]. The UHD 185 Firefly hyperspectral sensor was also mounted on a UAS to acquire 15,000 hyperspectral images ([121] b). As UAS images can be acquired and analyzed quickly, they can help reduce costs in the long run. However, UAS's need to fly low to acquire images with high spatial resolution. Modern UAS's suffer from limited flight times due to battery limitations. Furthermore, attaching multiple sensors increases the load on the UAS, resulting in further reduction of flight times [38], which motivates high-altitude flights at the expense of spatial resolution. Therefore, to acquire details of each lesion, close-up imagery such as that offered by handheld imagery and mounting sensors on booms can be more useful. The NLB dataset discussed in Section 2 was acquired using each of the three platforms: handheld, boom, and UAS [117]. A Sony A600 and Canon EOS Rebel sensors were used. When the sensor was mounted on a boom, the height was set to five meters. The UAS was flown at an altitude of six meters, a speed of 1.5 m/s, and images were acquired every two seconds. This resulted in a dataset consisting of 18,222 images of the NLB disease in corn. UAS imagery datasets can be very useful as they are easy to use and are capable of acquiring data at field and plant scales. Furthermore, a mosaic of UAS images acquired over a field can be created to help analyze the entire field at once [38]. However, creating a mosaic may cause loss of disease-specific information due to small disease lesion characteristics. Therefore, it is more useful to analyze specific diseases using single images and mosaics to analyze disease spread.

Although UAS are becoming increasingly popular, they are prone to limitations [75]. Nevertheless, as robust deep learning models continue to be developed, UAS technology continues to improve, and regulations are eased [88], disease management systems will rely heavily on the integration of UAS.

4. What deep learning techniques have been used for plant disease identification?

4.1. Artificial neural networks

Deep learning is a subset of machine learning, a subset of artificial intelligence. Deep learning relies on deep artificial neural networks to learn patterns and draw inferences from provided data. There are three different deep learning and more generally machine learning approaches based on the availability of labeled data, namely supervised learning, where all the training data is labeled, semi-supervised learning, where part of the training data is labeled, and unsupervised learning, where only unlabeled data is provided.

Using deep learning techniques, plant disease identification has proven to be a promising approach as convolutional layers have been successfully used to automatically identify important plant features, including the colors and textures of lesions. Furthermore, similar performance was attainable while removing 75% of the parameters, further emphasizing the usefulness and practicality of deep learning [9]. CNNs can also identify the features within images and perform classification concurrently. However, this approach is limited and time-intensive due to the need for large datasets. More researchers are starting to adopt these techniques as more authors are beginning to provide their datasets for public use and push research further [9,46,107].

Deep learning is being promoted in precision agriculture and phenotyping as it holds a lot of promise for developing the research area [33]. However, one limitation of deep learning-based solutions for plant disease identification is the interpretability of the results [80]. Therefore, a thorough review of deep learning-based plant disease identification studies is crucial for understanding the techniques, the ability to derive useful information from the results, and the importance of different evaluation metrics.

A typical convolutional neural network consists of an input layer, hidden layers, convolutional layers, pooling layers, fully connected layers, and an output layer cost function used during training. In addition, different activation functions are used and, in some cases, dropout layers as a regularization technique.

4.2. Image classification

Image classification via deep learning is a popular technique for agricultural applications [50], such as plant disease identification, as observed from the reviewed studies. Deep learning-based image classification is typically carried out via supervised training, where labeled image datasets are used to identify the class that best represents the object in images using the softmax activation function in the final output layer. Over the years, image classification has become increasingly popular and adapted by many researchers across disciplines as different CNN architectures have been developed and used for different applications. The advancements in image classification CNN architectures, namely LeNet, AlexNet, ZFNet, VGGNet, GoogLeNet, ResNet, DenseNet, CapsNet, and SENet, were discussed [103].

Image classification models are generally trained on a network from scratch or by using transfer learning on pre-trained models. Training a model accurately from scratch requires access to large datasets such as ImageNet [27] and MS COCO [57], consisting of 3.2 million and 328, 000 images. Transfer learning is a method in which pre-trained weights from a model trained to identify similar objects are loaded prior to training. Transfer learning using pre-trained weights from the ImageNet dataset is a method that researchers commonly use to help improve the performance of deep learning models when access to large training datasets is limited [50]. In addition, transfer learning was reported to yield higher accuracies for plant disease identification [25,65]. Out of the 70 studies, 43 studies used image classification. Among the reviewed studies, image classification was the most commonly used deep learning technique for plant disease identification and transfer learning was used in most studies.

Image classification models are commonly evaluated based on training, validation, and testing accuracies. However, testing accuracies are used to report the ability of a deep learning model to identify objects across different datasets and conditions. Therefore, it is crucial to maximizing testing accuracy on unseen data from different datasets. Additional metrics commonly used to evaluate the performance of deep learning models include precision, recall, and F1-score. The F1-score metric is particularly important in studies where datasets contain class imbalance. For example, F1-score was used to evaluate the performance of image classification models and testing accuracies for weed identification, as there was a class imbalance in the dataset used for training [3]. Image classification models have successfully identified plant diseases across different crops and in different conditions with validation and testing accuracies of up to 99.8% for single crop identification [2] and 99.75% for the entire PlantVillage dataset [108]. However, studies have been unable to generalize their models and report high testing accuracies across datasets and image acquisition conditions, as discussed in the next section.

Although image classification has been the most commonly used deep learning technique within the research community, it is difficult to perceive how it would efficiently provide information regarding the location of disease lesions within images and identify multiple diseases at once.

4.3. Object detection

Object detection refers to locating and identifying multiple instances of objects within images and videos. Here, we are interested in deep learning-based object detection, where supervised training is used via images provided with bounding box annotations. Popular deep learning-based object detection algorithms include R-CNN [36], Fast R-CNN [35], Faster R-CNN [87], YOLO [86], and Single Shot MultiBox Detector (SSD) [60]. To train object detection models, bounding box annotations must provide ground truth information to the deep learning model. Popular tools for creating the bounding box annotations are MATLAB, YOLO-Mark, and LabelImg.

Object detection has also been used to identify plant diseases within the research community. However, some studies have focused on identifying entire diseased leaves, while others have identified specific disease lesions. Eight studies used object detection to identify and locate plant diseases among the reviewed studies.

Object detection models are commonly evaluated using average precision (AP) and mean average precision (mAP). It is important to report these metrics because reporting the mAP may not be the best metric to evaluate models when the AP varies significantly across different classes. The highest mAP of 84.1% for plant diseases was reported for detecting tomato disease lesions [24]. Furthermore, studies have also reported precision, recall, and F1-scores for deep learning-based object detection models.

Deep learning-based object detection is promising for plant disease identification as it can identify multiple different plant diseases and instances of diseases simultaneously within images and videos. Therefore, such techniques can be very useful in real-time systems. However, the objects are detected within bounding boxes, containing more than necessary information. For example, if the disease lesion in the frame is diagonal, the bounding box may contain an extra leaf area that does not consist of the disease.

4.4. Segmentation

Semantic segmentation via deep learning is also a popular technique that can be used for plant disease identification. Models are trained by segmenting entire objects within images and assigning a class to each pixel from the training labels. Some popular segmentation algorithms include Mask R-CNN and U-Net. Ground truth information is also required for supervised training of segmentation models. The annotations for segmentation are more accurate as they localize the exact shape of the object rather than a bounding box. However, the annotations time required is greater, increasing the labor cost. Some common annotation tools used for image segmentation include MATLAB and LabelMe.

Three studies used semantic segmentation to identify corn, wheat, and alfalfa disease. Accuracies of up to 94.74% were reported for alfalfa disease identification [84].

5. How well have deep learning models generalized to unseen testing data, different datasets, and field images?

5.1. What is generalization?

Generalization refers to how accurately a trained model can identify unseen testing images. Generalization can be further categorized into same dataset generalization and cross dataset generalization. Same dataset generalization applies when providing unseen testing data from the same dataset as the training data. Although the testing data is kept hidden during model training, the testing accuracy is expected to be high as the images used for both testing and training data were acquired under the same conditions. However, the added complexity reduces the testing accuracy when images acquired from different datasets, backgrounds, and acquisition conditions are provided as testing data to the same models. This is referred to as cross dataset generalization in this study. Due to the complexity of potentially varying circumstances affecting agricultural data acquired under field conditions, it is important to maximize the testing accuracy for cross dataset generalization by training robust deep learning models capable of distinguishing disease features accurately.

5.2. Same dataset generalization

Recently, there have been many success stories with designing deep learning models that generalize well to unseen testing images from the same dataset used to train the models. For example, a testing accuracy of 99.48% was reported for the VGG model trained to identify diseases from the PlantVillage dataset when 20% of the dataset was used as testing data [30]. A similar result was observed when the PlantVillage dataset was used for training the InceptionV3 pre-trained model. Testing accuracy of 99.01% was observed when unseen testing images from the PlantVillage dataset were provided [55]. Generalization within the PlantVillage dataset was also observed when the following testing accuracies were reported: 99.185% was reported for tomato disease identification [23], 98.9% for corn disease identification with the GoogLeNet pre-trained model [122], 99.75% for the entire PlantVillage dataset with the DenseNet121 pre-trained model [108], 91.7% for identifying apple, cherry, and corn diseases [39], 96.5% when identifying 27 diseases across six crops [70], 96.7% for identifying corn diseases [71], and 98% for potato disease identification [44]. The Digipathos dataset was also used to report the same dataset generalization with a testing accuracy of 80% for identifying all diseases within the dataset [10] and 96.96% for identifying soybean diseases [51]. Many studies where custom datasets were used for plant disease identification have also reported testing accuracies of greater than 90%, which supports the hypothesis that deep learning models are, in general, capable of generalizing well to unseen testing data within the same dataset for plant disease identification [64].

5.3. Cross dataset generalization

Although researchers have reported high accuracy for same dataset generalization, it is also important to test the generalization accuracy across datasets and conditions, particularly for agricultural applications, where trained models are anticipated to be tested on data corresponding to upcoming seasons. Like a well-trained plant pathologist, a well-trained deep learning model should be capable of identifying plant diseases from crops worldwide under real-field and lab conditions. However, the accuracies reported for cross dataset generalization are poor compared to the same dataset generalization accuracies. After high testing accuracies were obtained when testing data from the same dataset was provided, two additional experiments were performed to test the ability of the trained models to identify diseases from images acquired under real-field or lab conditions for cross-dataset generalization [30]. First, deep learning models were trained using field images and tested on lab images, which resulted in the highest testing accuracy of 65.69%. The process was reversed for the next experiment, where the models were trained on lab images and tested on field images resulting in the highest testing accuracy of 33.27%. A similar approach was also followed when field images from IPM and Bing were used as testing images when the model was trained on the PlantVillage dataset, resulting in the highest testing accuracy of 45.95% and 33.97% for IPM and Bing acquired images, respectively [55]. Generalization capabilities of trained deep learning models were also tested by removing complex backgrounds from field acquired images for testing data, which resulted in 89% testing accuracy [13].

Although one instance of high generalization accuracy was observed [13], the accuracy reported for cross dataset generalization for plant disease identification was lower than 50% when the models were trained on lab images with uniform backgrounds but tested on field images. However, accuracy improved when the models were trained on field-acquired images and tested on images acquired under lab conditions. As it is easier and faster to acquire images under lab conditions and farmers will be using the models in the field, it is important to develop models that can be trained on lab acquired images with uniform backgrounds capable of accurately identifying diseases under field acquired images.

6. How was disease severity estimated?

The severity of diseases refers to how a particular disease spreads throughout the plant and affects growth and yield. It is most often expressed as a percentage or proportion. Accurately estimating the degree/level of disease severity will help predict crop yield and recommend effective management practices. Traditionally, raters and plant pathologists visually analyze plants and report disease severity estimates. However, such methods for analyzing severity are inefficient as they incur high costs and personal bias [116]. There is a great promise to alleviate these issues by using a deep learning-based system for analyzing disease severity from images.

The problem of estimating plant disease severity has been tackled by many researchers, however, different studies have used different definitions of severity to estimate disease severity. Most studies have defined disease severity as the area of disease lesions or symptoms with respect to the area of the leaf [29,81]. Therefore, researchers have used deep learning-based image classification to solve the problem based on the area of leaves covered by lesions. A total of six studies in this review used images for estimating the severity of different plant diseases.

Before deep learning-based severity estimation models, traditional image processing was used to identify pomegranate disease severity [17]. Fuzzy logic was also used for corn disease severity estimation [97]. Deep learning-based solutions are becoming accepted widely within the research community to learn important features for automatically estimating the disease severity. Four different disease severity levels/classes for the black rot disease of apple, namely healthy, early-stage, middle-stage, and end-stage, were identified using deep learning. Images of leaves with no spots were labeled as healthy, leaves with small circular spot lesions with diameters of less than 5 mm were labeled as early-stage, and leaves with greater than three spots and a frog-eye spot were labeled as middle-stage. Finally, the end-stage was used to label leaves just before falling. After training the models, the highest accuracy of 90.4% was achieved using the VGG16 model [116]. A new network architecture, PD2SE, was then created for disease identification and severity estimation. Accuracies of up to 91% were reported for disease severity estimation after training the PD2SE model to identify if the disease was mild or severe from the PlantVillage dataset [56]. Finally, five levels of coffee disease severity, namely healthy, very low, low, high, and very high, were identified using deep learning. Leaves with less than 0.1% lesion coverage were labeled as healthy, from 0.1% to 5% as very low, from 5.1% to 10% as low, from 10.1% to 15% as high, and greater than 15% as very high. An accuracy of 86.51% was reported for severity estimation using the VGG16 model [28]. A multi-label binary relevance CNN (BR-CNN) identified crops, disease type, and severity estimation. The diseases and severity were identified with the highest accuracies of 98.45% with DenseNet121 and 92.93% with ResNet50, respectively. However, only three severity classes, namely general, normal, and serious, were classified [47]. Deep learning models based on the squeeze and excitation networks, CapsNet, AlexNet, and ResNet were recently used to identify four severity classes for the late blight disease in tomatoes: namely healthy, early, middle, and end, with testing accuracies of up to 93.75% [112].

The use of deep learning-based disease severity estimation represents a gap in the literature as most researchers have custom definitions for different levels of severity. Another challenge in this area was also discussed, where each different plant disease has a different severity estimation criterion. For example, the tan spot disease in lima bean had a minimum severity of 0.29% and a maximum of 35.42%, while the powdery mildew disease in lilac had a minimum severity of 14.21% and a maximum of 93.19% [81]. This further complicates the severity estimation problem as hundreds of different crops with multiple diseases exist. Therefore, a standard method must be developed to estimate disease severity accurately. A diagrammatic scale was developed to improve the precision and accuracy of estimating the severity level of NLB-infected corn leaves [113].

7. Can deep learning outperform humans for plant disease identification?

It is crucial for the success of deep learning-based solutions to outperform currently deployed human manual identification. Humans can identify objects with an accuracy of 94% [40]. Therefore, it is reasonable to hypothesize that a deep learning model capable of identifying plant diseases with an accuracy greater than 94% can outperform humans.

Deep learning-based image classification was used to identify all the plant diseases in the entire PlantVillage dataset with accuracies of 99.48% [30], 98.65% [49], 99.01% [55], 91.4% [7], 99.75% [108], and 99.81% [92], given that the training and testing images were from the same dataset. On the other hand, all the diseases for each crop present in the Digipathos dataset were identified using image classification with average accuracies of only 87% [13] and 82% [14].

Some researchers only focused on identifying diseases for some crops within datasets using image classification from the reviewed studies. Apple, cherry, and corn diseases were identified with an accuracy of 98.8% [109]. Corn, grape, tomato, apple, and sugarcane diseases were identified with an accuracy of 96.5% [70]. Corn and peach diseases were identified with an accuracy of 99.28% [94]. Grape and tomato diseases were identified with an accuracy of 100% [100]. Custom acquired peach, apple, pear, and grapevine diseases were identified with an accuracy of 96.3% [101].

Many researchers focused on identifying diseases specific to a single crop. Image classification was used to identify tomato diseases from the PlantVillage dataset with accuracies of 98.6% [93], 99.185% [23], 91.15% ([120]a), 99.65% [20], and 97.53% [8]. Tomato diseases were again identified via image classification using custom datasets with accuracies of 95.75% [64] and 96.25% [115]. The powdery mildew disease in tomatoes was identified using the three different deep learning models: namely Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), and Gated Recurrent Unit (GRU), with testing accuracy of up to 93.88% [111]. Corn diseases from the PlantVillage dataset were also identified using image classification with accuracies of 98.9% [122], 96.7% [71], and 82.8% [104]. Image classification was also used for identifying corn diseases with an accuracy of 87% [10] and for identifying soybean diseases with an accuracy of 96.96% [51] using the Digipathos dataset for training models. Charcoal rot in soybean was identified with an accuracy of 95.37% [74]. Bean disease classes: namely angular leaf spot, bean rust, and healthy, were also identified using deep learning with a testing accuracy of up to 95.31% using a balanced dataset acquired under field conditions [89]. Diseases in cassava were identified using a custom acquired dataset under field conditions with an accuracy of 96.75% [78]. Image classification was also used for potato disease identification with an accuracy of 99.8% [1] and banana disease identification with an accuracy of 95.31% [90]. A new L2MXception CNN model was recently created by using L2 regularization for identifying peach diseases under field conditions with 93.85% accuracy [119]. Lastly, wheat diseases were identified using image classification with accuracies of 87% [82] and 85% ([121]b).

In most cases, deep learning-based image classification has been employed to achieve accuracy greater than 94% for plant disease identification. However, lower accuracies were reported for a non-negligible number of cases. Most cases in which classification accuracies of under 94% were reported were cases in which the learning task requires identifying diseases from multiple different crops. In such studies, the dataset classes were unbalanced. According to the review of studies conducted on deep learning-based plant disease identification [12,22,54,61,63,66,100], balanced datasets are very important in ensuring high classification accuracies. Additionally, deep learning-based object detection and segmentation have also been used. Tomato diseases were detected with an accuracy of 91.6% [64] and mAP of 0.83 at an IoU of 0.5 [32]. Wheat diseases were also identified using object detection with an accuracy of 97% [62]. Object detection was also used for identifying and locating downy mildew in onions with an accuracy of 87.9% and mAP of 0.8 at an IoU of 0.6 [52]. In addition, the SSD object detection model was also trained using the entire PlantVillage dataset to identify and locate multiple diseases with a mAP of 73.07% [91]. However, the PlantVillage dataset only consists of single leaf images with uniform backgrounds. Therefore, the recent study [85] is more applicable to real-world applications as real field data was used to train the SSD object detection model to identify cassava diseases from images and videos with a mAP of 73% and accuracy of up to 80%, respectively. Segmentation was used to identify NLB lesions in corn with a precision of 96% at an IoU of 0.5 [102]. A new segmentation approach was also used to identify disease lesions with an accuracy of 94% by providing segmented lesion images to the pre-trained AlexNet model [65].

Although traditional machine learning solutions could identify diseases with an accuracy of greater than 94% [31,43,44,84], their use is limited due to feature extraction and definition requirements. Therefore, deep learning can help improve generalization as CNN's allow for automatic feature extraction.

It may appear, at first glance, that deep learning solutions for disease identification are superior to human-based identification. However, as discussed in Section 6 of this study, a major limitation to current solutions is the generalization capability of models. For example, this was shown when diseases were first identified with an accuracy of 99.01%, given that the training and testing images were from the same dataset. However, when testing images from a different dataset were provided to the model, the accuracy dropped down to 45.95% [55].

Based on the literature review, deep learning-based solutions for plant disease management have shown promising results. Therefore, deep learning-based solutions are quickly becoming a standard for disease identification and severity estimation. However, due to crucial open research areas and limitations with existing solutions, it is still difficult to achieve an end-to-end system for disease management purposes.

8. What are some open research topics?

Plant disease diagnosis and monitoring are complex tasks that rely on accurate disease identification and severity estimation. Since the introduction of the PlantVillage dataset in 2015, multiple studies have used different image processing, machine learning, and deep learning techniques to identify different diseases in different crops that are essential to improving global food production and food security. As discussed in this study, many researchers have successfully created different algorithms and models to identify plant diseases. However, limitations and gaps within the research are responsible for the lack of end-to-end plant disease management systems.

A change in the approach is required as current deep learning-based plant disease identification methods require every user to train a deep learning model in their field under controlled or pre-defined image acquisition conditions. Although diseases were identified accurately by most studies reviewed so far, the trained models did not generalize well to images from different datasets and conditions. Therefore, current methods will be as ineffective as scouting of fields by farmers. Hence, to create a robust model for identifying diseases accurately, it is important to ensure that a trained model is capable of global disease identification from all datasets and conditions.

Current deep learning-based solutions are also limited in identifying multiple disease types. Most studies conducted on plant disease identification focused on training deep learning models for single crop disease identification or identifying a small number of diseases. Although these solutions may result in accurate disease identification, diseases outside the training set will not be identified. This problem must be addressed to implement an effective disease identification system for precision agriculture.

Severity estimation definitions have varied across multiple studies. However, a standard definition much be used for training deep learning models capable of accurately estimating the severity of diseases. Since image classification cannot accurately localize plant disease lesions and symptoms, its use for severity estimation will result in inaccurate analysis. Object detection overcomes the limitations of image classification and helps localize the disease lesions. Nevertheless, object detection can only create rectangular bounding boxes, resulting in the localization of additional areas with no disease lesions and symptoms. Therefore, the segmentation approach will result in the most accurate calculation of infected leaf regions with respect to the entire leaf.

Finally, disease identification timing within the season is also an important factor that must be considered to ensure that correct management techniques are implemented to maximize crop yield.

9. Conclusion

This study reviewed 70 studies consisting of review papers, image processing, machine learning, and deep learning studies for plant disease diagnosis. A set of seven critical research questions were identified and addressed. Multiple publicly available datasets and descriptions of popular plant diseases were reviewed. Different imaging sensors and data collection platforms were reviewed for plant disease identification. Finally, we discussed the generalization performance of deep learning models and their ability to outperform humans. We conclude that useful plant disease analysis will identify multiple crops, their respective diseases early in the season, and accurate disease severity estimation to develop an automated end-to-end plant disease management system.

CRediT authorship contribution statement

Aanis Ahmad: Investigation, Software, Writing – original draft. **Dharmendra Saraswat:** Conceptualization, Supervision, Funding acquisition. **Aly El Gamal:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was made possible by the funding provided by Wabash Heartland Innovation Network (WHIN) grant number 18024589 and the USDA National Institute of Food and Agriculture (NIFA) Hatch project 1012501.

References

[1] M. Agarwal, S.K. Gupta, K. Biswas, Development of efficient cnn model for tomato crop disease identification, Sustain. Comput. 28 (2020), 100407.
[2] M. Agarwal, A. Sinha, S.K. Gupta, D. Mishra, R. Mishra, Potato crop disease classification using convolutional neural network. Smart Systems and IoT: Innovations in Computing, Springer, 2020, pp. 391–400.
[3] A. Ahmad, D. Saraswat, V. Aggarwal, A. Etienne, B. Hancock, Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems, Comput. Electron. Agric. 184 (2021), 106081.
[4] A. Ahmad, D. Saraswat, A. El Gamal, G.S. Johal, Comparison of deep learning models for corn disease identification, tracking, and severity estimation using images acquired from UAV-mounted and handheld sensors. 2021 ASABE Annual International Virtual Meeting, American Society of Agricultural and Biological Engineers, 2021, p. 1.
[5] Ahmad, A., Saraswat, D., Gamal, A.E., & Johal, G. (2021c). Cd&s dataset: Handheld imagery dataset acquired under field conditions for corn disease identification and severity estimation. arXiv:2110.12084.
[6] Anjanadevi, B., Charmila, I., Akhil, N., & Anusha, R. (2020). An improved deep815learning model for plant disease detection.
[7] D. Arg¨ueso, A. Picon, U. Irusta, A. Medela, M.G. San-Emeterio, A. Bereciartua, A. Alvarez-Gila, Few-shot learning approach for plant disease classification using images taken in the field, Comput. Electron. Agric. 175 (2020), 105542.
[8] H.A. Atabay, Deep residual learning for tomato plant leaf disease identification, J. Theoret. Appl. Inf. Technol. (2017) 95.
[9] M.A.F. Azlah, L.S. Chua, F.R. Rahmad, F.I. Abdullah, S.R. Wan Alwi, Review on techniques for plant leaf classification and recognition, Computers 8 (2019) 77.
[10] J.G. Barbedo, Factors influencing the use of deep learning for plant disease recognition, Biosystems Eng. 172 (2018) 84–91.
[11] J.G.A. Barbedo, Digital image processing techniques for detecting, quantifying and classifying plant diseases, SpringerPlus 2 (2013) 660.
[12] J.G.A. Barbedo, A review on the main challenges in automatic plant disease identification based on visible range images, Biosystems Eng. 144 (2016) 52–60.
[13] J.G.A. Barbedo, Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification, Comput. Electron. Agric. 153 (2018) 46–53.
[14] J.G.A. Barbedo, Plant disease identification from individual lesions and836spots using deep learning, Biosystems Eng. 180 (2019) 96–107.
[15] J.G.A. Barbedo, A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses, Drones 3 (2019) 40.
[16] J.G.A. Barbedo, L.V. Koenigkan, B.A. Halfeld-Vieira, R.V. Costa, K.L. Nechet, C. V. Godoy, M.L. Junior, F.R.A. Patricio, V. Talamini, L.G. Chitarra, et al., Annotated plant pathology databases for image-based detection and recognition of diseases, IEEE Lat. Am. Trans. 16 (2018) 8431749–8431757.
[17] M. Bhange, H. Hingoliwala, Smart farming: pomegranate disease detection using image processing, Procedia Comput. Sci. 58 (2015) 280–288.
[18] A. Bhatia, A. Chug, A.P. Singh, R.P. Singh, D. Singh, A forecasting technique for powdery mildew disease prediction in tomato plants. Proceedings of Second Doctoral Symposium on Computational Intelligence, Springer, 2022, pp. 509–520, 861.
[19] A. Bhatia, A. Chug, A.P. Singh, R.P. Singh, D. Singh, A machine learning-based spray prediction model for tomato powdery mildew disease, Indian Phytopathol. 75 (2022) 225–230.
[20] P. Bhatt, S. Sarangi, S. Pappula, Comparison of cnn models for application in crop health assessment with participatory sensing. 2017 IEEE Global Humanitarian Technology Conference (GHTC), IEEE, 2017, pp. 1–7.
[21] C. Bock, G. Poole, P. Parker, T. Gottwald, Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging, Crit. Rev. Plant Sci. 29 (2010) 59–107.
[22] J. Boulent, S. Foucher, J. Th´eau, P.-L. St-Charles, Convolutional neural networks for the automatic identification of plant diseases, Front. Plant Sci. (2019) 10.
[23] M. Brahimi, K. Boukhalfa, A. Moussaoui, Deep learning for tomato diseases: classification and symptoms visualization, Appl. Artif. Intell. 31 (2017) 299–315.
[24] Chandra, M., Patil, P.S., Roy, S., & Redkar, S.S. (2020). Classification of various859plant diseases using deep siamese network.
[25] J. Chen, J. Chen, D. Zhang, Y. Sun, Y.A. Nanehkaran, Using deep transfer learning for image-based plant disease identification, Comput. Electron. Agric. 173 (2020), 105393.
[26] C. DeChant, T. Wiesner-Hanks, S. Chen, E.L. Stewart, J. Yosinski, M.A. Gore, R. J. Nelson, H. Lipson, Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning, Phytopathology 107 (2017) 1426–1432.
[27] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition, Ieee, 2009, pp. 248–255.
[28] J.G. Esgario, R.A. Krohling, J.A. Ventura, Deep learning for classification and severity estimation of coffee leaf biotic stress, Comput. Electron. Agric. 169 (2020), 105162.
[29] G. Fenu, F.M. Malloci, Forecasting plant and crop disease: an explorative study on current algorithms, Big Data Cogn. Comput. 5 (2021) 2.
[30] K.P. Ferentinos, Deep learning models for plant disease detection and diagnosis, Comput. Electron. Agric. 145 (2018) 311–318.
[31] N. Fu, C. Wang, X. Ji, Study on visual detection device of plant leaf disease. 2019 IEEE International Conference on Mechatronics and Automation (ICMA), IEEE, 2019, pp. 86–90.

[32] A. Fuentes, S. Yoon, S.C. Kim, D.S. Park, A robust deep-learning- based detector for real-time tomato plant diseases and pests recognition, Sensors 17 (2017) 2022.
[33] Z. Gao, Z. Luo, W. Zhang, Z. Lv, Y. Xu, Deep learning application in plant stress imaging: a review, AgriEngineering 2 (2020) 430–446.
[34] K.R. Gavhale, U. Gawande, An overview of the research on plant leaves disease detection using image processing techniques, IOSR J. Comput. Eng. (IOSR-JCE) 16 (2014) 10–16.
[35] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1440–1448.
[36] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587.
[37] M. Govardhan, M. Veena, Diagnosis of tomato plant diseases using random forest. 2019 Global Conference for Advancement in Technology (GCAT), IEEE, 2019, pp. 1–5.
[38] W. Guo, M.E. Carroll, A. Singh, T.L. Swetnam, N. Merchant, S. Sarkar, A.K. Singh, B. Ganapathysubramanian, Uas-based plant phenotyping for research and breeding applications, Plant Phenomics (2021) 2021.
[39] J. Hang, D. Zhang, P. Chen, J. Zhang, B. Wang, Classification of plant leaf diseases based on improved convolutional neural network, Sensors 19 (2019) 4161.
[40] J. Hearty, Advanced Machine Learning with Python, Packt Publishing Ltd, 2016.
[41] G. Hu, H. Wu, Y. Zhang, M. Wan, A low shot learning method for tea leaf's disease identification, Comput. Electron. Agric. 163 (2019), 104852.
[42] Hughes, D., & Salath ́e, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv 2015. *arXiv preprint* arXiv:1511.08060.
[43] M.A. Iqbal, K.H. Talukder, Detection of potato disease using image segmentation and machine learning. 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), IEEE, 2020, pp. 43–47.
[44] M. Islam, A. Dinh, K. Wahid, P. Bhowmik, Detection of potato diseases using image segmentation and multiclass support vector machine. 2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE), IEEE, 2017, pp. 1–4.
[45] S.B. Jadhav, V.R. Udup, S.B. Patil, Soybean leaf disease detection and severity measurement using multiclass svm and knn classifier, Int. J. Electr. Comput. Eng. 9 (2019) 4092.
[46] W.-S. Jeon, S.-Y. Rhee, Plant leaf recognition using a convolution neural network, Int. J. Fuzzy Logic Intell. Syst. 17 (2017) 26–34.
[47] M. Ji, K. Zhang, Q. Wu, Z. Deng, Multi-label learning for crop leaf diseases recognition and severity estimation based on convolutional neural networks, Soft Comput. 24 (2020) 15327–15340.
[48] A. Johannes, A. Picon, A. Alvarez-Gila, J. Echazarra, S. Rodriguez-Vaamonde, A. D. Navajas, A. Ortiz-Barredo, Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case, Comput. Electron. Agric. 138 (2017) 200–209.
[49] K. Kamal, Z. Yin, M. Wu, Z. Wu, Depthwise separable convolution architectures for plant disease classification, Comput. Electron. Agric. 165 (2019), 104948.
[50] A. Kamilaris, F.X. Prenafeta-Bold ́u, Deep learning in agriculture: a survey, Comput. Electron. Agric. 147 (2018) 70–90.
[51] A. Karlekar, A. Seal, Soynet: soybean leaf diseases classification, Comput. Electron. Agric. 172 (2020), 105342.
[52] W.-S. Kim, D.-H. Lee, Y.-J. Kim, Machine vision-based automatic disease symptom detection of onion downy mildew, Comput. Electron. Agric. 168 (2020), 105099.
[53] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012, pp. 1097–1105.
[54] S.H. Lee, C.S. Chan, S.J. Mayo, P. Remagnino, How deep learning extracts and learns leaf features for plant classification, Pattern Recognit. 71 (2017) 1–13.
[55] S.H. Lee, H. Go ̈eau, P. Bonnet, A. Joly, New perspectives on plant disease characterization based on deep learning, Comput. Electron. Agric. 170 (2020), 105220.
[56] Q. Liang, S. Xiang, Y. Hu, G. Coppola, D. Zhang, W. Sun, Pd2se-net: computer-assisted plant disease diagnosis and severity estimation network, Comput. Electron. Agric. 157 (2019) 518–529.
[57] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll ́ar, C. L. Zitnick, Microsoft coco: common objects in context. European conference on computer vision, Springer, 2014, pp. 740–755.
[58] S. Lindow, R. Webb, Quantification of foliar plant disease symptoms by microcomputer-digitized video image analysis, Phytopathology 73 (1983) 520–524.
[59] B. Liu, Y. Zhang, D. He, Y. Li, Identification of apple leaf diseases based on deep convolutional neural networks, Symmetry 10 (2018) 11.
[60] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: Single shot multibox detector. European conference on computer vision, Springer, 2016, pp. 21–37.
[61] M. Loey, A. ElSawy, M. Afify, Deep learning in plant diseases detection for agricultural crops: a survey, Int. J. Serv. Sci. Manag. Eng. Technol. (IJSSMET) 11 (2020) 41–58.
[62] J. Lu, J. Hu, G. Zhao, F. Mei, C. Zhang, An in-field automatic wheat disease diagnosis system, Comput. Electron. Agric. 142 (2017) 369–379.
[63] Y. Lu, S. Young, A survey of public datasets for computer vision tasks in precision agriculture, Comput. Electron. Agric. 178 (2020), 105760.
[64] R.G. de Luna, E.P. Dadios, A.A. Bandala, Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. TENCON 2018-2018 IEEE Region 10 Conference, IEEE, 2018, pp. 1414–1419.
[65] J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, Z. Sun, A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network, Comput. Electron. Agric. 154 (2018) 18–24.
[66] A.-K. Mahlein, Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping, Plant Dis. 100 (2016) 241–251.
[67] Marcus, G. (2018). Deep learning: A critical appraisal. ***arXiv preprint*** arXiv:1801.00631.
[68] D.P. Martin, E.P. Rybicki, Microcomputer-based quantification of maize streak virus symptoms in zea mays, Phytopathology 88 (1998) 422–427.
[69] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L.R. Goulart, et al., Advanced methods of plant disease detection. A review, Agron. Sustain. Dev. 35 (2015) 1–25.
[70] S.V. Militante, B.D. Gerardo, N.V. Dionisio, Plant leaf detection and disease recognition using deep learning, 2019 IEEE Eurasia Conf. IOT, Commun. Eng. (ECICE) (2019) 579–582, https://doi.org/10.1109/ECICE47484.2019.8942686.
[71] S. Mishra, R. Sachan, D. Rajpal, Deep convolutional neural network based detection system for real-time corn plant disease recognition, Procedia Comput. Sci. 167 (2020) 2003–2010.
[72] S.P. Mohanty, D.P. Hughes, M. Salath ́e, Using deep learning for image-based plant disease detection, Front. Plant Sci. 7 (2016) 1419.
[73] U. Mokhtar, M.A. Ali, A.E. Hassenian, H. Hefny, Tomato leaves diseases detection approach based on support vector machines. 2015 11th International computer engineering conference (ICENCO), IEEE, 2015, pp. 246–250.
[74] K. Nagasubramanian, S. Jones, A.K. Singh, S. Sarkar, A. Singh, B. Ganapathysubramanian, Plant disease identification using explainable 3d deep learning on hyperspectral images, Plant Methods 15 (2019) 98.
[75] K. Neupane, F. Baysal-Gurel, Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: a review, Remote Sens. 13 (2021) 3841.
[76] S. Nigam, R. Jain, Plant disease identification using deep learning: a review, Indian J. Agric. Sci. 90 (2020) 249–257.
[77] E.-C. Oerke, H.-W. Dehne, Safeguarding production—losses in major crops and the role of crop protection, Crop Prot. 23 (2004) 275–285.
[78] D.O. Oyewola, E.G. Dada, S. Misra, R. Damaˇseviˇcius, Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing, PeerJ Comput. Sci. 7 (2021) e352.
[79] J. Parraga-Alava, K. Cusme, A. Loor, E. Santander, Rocole: a robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition, Data Brief 25 (2019), 104414.
[80] M. P ́erez-Enciso, L.M. Zingaretti, A guide on deep learning for complex trait genomic prediction, Genes 10 (2019) 553.
[81] S.J. Pethybridge, S.C. Nelson, Leaf doctor: a new portable application for quantifying plant disease severity, Plant Dis. 99 (2015) 1310–1316.
[82] A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, A. Johannes, Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild, Comput. Electron. Agric. 161 (2019) 280–290.
[83] H.B. Prajapati, J.P. Shah, V.K. Dabhi, Detection and classification of rice plant diseases, Intell. Decis. Technol. 11 (2017) 357–373.
[84] F. Qin, D. Liu, B. Sun, L. Ruan, Z. Ma, H. Wang, Identification of alfalfa leaf diseases using image recognition technology, PLoS One 11 (2016), e0168274.
[85] A. Ramcharan, P. McCloskey, K. Baranowski, N. Mbilinyi, L. Mrisho, M. Ndalahwa, J. Legg, D.P. Hughes, A mobile-based deep learning model for cassava disease diagnosis, Front. Plant Sci. 10 (2019) 272.
[86] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788.
[87] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real- time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. (2015) 91–99.
[88] R. Rodriguez, et al., Perspective: agricultural aerial application with unmanned aircraft systems: current regulatory framework and analysis of operators in the united states, Trans. ASABE 64 (2021) 1475–1481.
[89] P. Sahu, A. Chug, A.P. Singh, D. Singh, R.P. Singh, Deep learning models for beans crop diseases: classification and visualization techniques, Int. J. Modern Agric. 10 (2021) 796–812, 1062.
[90] P. Sahu, A. Chug, A.P. Singh, D. Singh, R.P. Singh, Classification and activation map visualization of banana diseases using deep learning models. International Conference on Innovative Computing and Communications, Springer, 2022, pp. 751–767.
[91] M.H. Saleem, S. Khanchi, J. Potgieter, K.M. Arif, Image-based plant disease identification by deep learning meta-architectures, Plants 9 (2020) 1451.
[92] M.H. Saleem, J. Potgieter, K.M. Arif, Plant disease classification: a comparative evaluation of convolutional neural networks and deep learning optimizers, Plants 9 (2020) 1319.
[93] P. Sharma, Y.P.S. Berwal, W. Ghai, Performance analysis of deep learning cnn models for disease detection in plants using image segmentation, Inf. Process. Agric. (2019).
[94] M.H. Sheikh, T.T. Mim, M.S. Reza, A.S.A. Rabby, S.A. Hossain, Detection of maize and peach leaf diseases using image processing. 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2019, pp. 1–7.
[95] S. Shrivastava, S.K. Singh, D.S. Hooda, Soybean plant foliar disease detection using image retrieval approaches, Multim. Tools Appl. 76 (2017) 26647–26674.

[96] V.K. Shrivastava, M.K. Pradhan, Rice plant disease classification using color features: a machine learning paradigm, J. Plant Pathol. 103 (2021) 17–26.
[97] M. Sibiya, M. Sumbwanyambe, An algorithm for severity estimation of plant leaf diseases by the use of colour threshold image segmentation and fuzzy logic inference: a proposed algorithm to update a "leaf doctor" application, AgriEngineering 1 (2019) 205–219.
[98] A.K. Singh, B. Ganapathysubramanian, S. Sarkar, A. Singh, Deep learning for plant stress phenotyping: trends and future perspectives, Trends Plant Sci. 23 (2018) 883–898.
[99] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, N. Batra, Plantdoc: a dataset for visual plant disease detection, in: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, 2020, pp. 249–253.
[100] K.K. Singh, An artificial intelligence and cloud based collaborative platform for plant disease identification, tracking and forecasting for farmers. 2018 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), IEEE, 2018, pp. 49–56.
[101] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, D. Stefanovic, Deep neural networks based recognition of plant diseases by leaf image classification, Comput. Intell. Neurosci. (2016) 2016.
[102] E.L. Stewart, T. Wiesner-Hanks, N. Kaczmar, C. DeChant, H. Wu, H. Lipson, R. J. Nelson, M.A. Gore, Quantitative phenotyping of northern leaf blight in uav images using deep learning, Remote Sens. 11 (2019) 2209.
[103] F. Sultana, A. Sufian, P. Dutta, Advancements in image classifica-1053tion using convolutional neural network, in: 2018 *Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2018, https://doi.org/10.1109/icrcicn.2018.8718718.
[104] X. Sun, J. Wei, Identification of maize disease based on transfer learning. Journal of Physics: Conference Series, IOP Publishing, 2020, 012080 volume 1437.
[105] S.F. Syed-Ab-Rahman, M.H. Hesamian, M. Prasad, Citrus disease1 detection and classification using end-to-end anchor-based deep learning model, Appl. Intell. 52 (2022) 927–938.
[106] E.C. Tetila, B.B. Machado, G.K. Menezes, A.D.S. Oliveira, M. Alvarez, W. P. Amorim, N.A.D.S. Belete, G.G. Da Silva, H. Pistori, Automatic recognition of soybean leaf diseases using uav images and deep convolutional neural networks, IEEE Geosci. Remote Sens. Lett. 17 (2019) 903–907.
[107] Y. Toda, F. Okura, et al., How convolutional neural networks diagnose plant disease, Plant Phenomics 2019 (2019), 9237136.
[108] E.C. Too, L. Yujian, S. Njuki, L. Yingchun, A comparative study of1068fine-tuning deep learning models for plant disease identification, Comput. Electron. Agric. 161 (2019) 272–279.
[109] K. Trang, L. TonThat, N.G.M. Thao, Plant leaf disease identification by deep convolutional autoencoder as a feature extraction approach. *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (ECTI-CON), IEEE, 2020, pp. 522–526.
[110] S.K. Upadhyay, A. Kumar, Early-stage brown spot disease recognition in paddy using image processing and deep learning techniques, Traitement du Signal 38 (2021) 1755–1766.
[111] T. Varshney, A. Chug, A.P. Singh, Deep learning models for prediction of tomato powdery mildew disease. 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), IEEE, 2021, pp. 1036–1041.
[112] S. Verma, S. Jahangir, A. Chug, R.P. Singh, A.P. Singh, D. Singh, Se-capsnet: automated evaluation of plant disease severity based on feature extraction through squeeze and excitation (se) networks and capsule networks, Kuwait J. Sci. (2022) 49.
[113] R.A. Vieira, R.M. Mesquini, C.N. Silva, F.T. Hata, D.J. Tessmann, C.A. Scapim, A new diagrammatic scale for the assessment of northern corn leaf blight, Crop Prot. 56 (2014) 55–57.
[114] C. Wang, P. Du, H. Wu, J. Li, C. Zhao, H. Zhu, A cucumber leaf disease severity classification method based on the fusion of deeplabv3+ and u-net, Comput. Electron. Agric. 189 (2021), 106373.
[115] D. Wang, R. Vinson, M. Holmes, G. Seibel, A. Bechar, S. Nof, Y. Tao, Early detection of tomato spotted wilt virus by hyperspectral imaging and outlier removal auxiliary classifier generative adversarial nets (or-ac-gan), Sci. Rep. 9 (2019) 1–14.
[116] G. Wang, Y. Sun, J. Wang, Automatic image-based plant disease severity estimation using deep learning, Comput. Intell. Neurosci. (2017) 2017.
[117] T. Wiesner-Hanks, E.L. Stewart, N. Kaczmar, C. DeChant, H. Wu, R.J. Nelson, H. Lipson, M.A. Gore, Image set for deep learning: field images of maize annotated with disease symptoms, BMC Res. Notes 11 (2018) 440.
[118] Witten, I.H., Holmes, G., McQueen, R.J., Smith, L.A., & Cunningham, S.J. (1993). Practical machine learning and its application to problems in agriculture.
[119] N. Yao, F. Ni, Z. Wang, J. Luo, W.-K. Sung, C. Luo, G. Li, L2mxception: an improved xception network for classification of peach diseases, Plant Methods 17 (2021) 1–13.
[120] S. Zhang, W. Huang, C. Zhang, Three-channel convolutional neural networks for vegetable leaf disease recognition, Cogn. Syst. Res. 53 (2019) 31–41.
[121] X. Zhang, L. Han, Y. Dong, Y. Shi, W. Huang, L. Han, P. Gonź alez-Moreno, H. Ma, H. Ye, T. Sobeih, A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral uav images, Remote Sens. 11 (2019) 1554.
[122] X. Zhang, Y. Qiao, F. Meng, C. Fan, M. Zhang, Identification of maize leaf diseases using improved deep convolutional neural networks, IEEE Access 6 (2018) 30370–30377.