

## Review Article

# Image-Based Detection of Plant Diseases: From Classical Machine Learning to Deep Learning Journey

Rehan Ullah Khan ,<sup>1</sup> Khalil Khan ,<sup>2</sup> Waleed Albattah ,<sup>1</sup> and Ali Mustafa Qamar ,<sup>3</sup>

<sup>1</sup>Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

<sup>2</sup>Department of Information Technology and Computer Science, Pak-Austria Fachhochschule,  
Institute of Applied Sciences and Technology, Pakistan

<sup>3</sup>Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia

Correspondence should be addressed to Ali Mustafa Qamar; al.khan@qu.edu.sa

Received 22 February 2021; Accepted 24 May 2021; Published 3 June 2021

Academic Editor: Farman Ullah

Copyright © 2021 Rehan Ullah Khan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Plant disease automation in agriculture science is the primary concern for every country, as the food demand is increasing at a fast rate due to an increase in population. Moreover, the increased use of technology today has increased the efficacy and accuracy of detecting diseases in plants and animals. The detection process marks the beginning of a series of activities to fight the diseases and reduce their spread. Some diseases are also transmitted between animals and human beings, making it hard to fight them. For many years, scientists have researched how to deal with the common diseases that affect humans and plants. However, there are still many parts of the detection and discovery process that have not been completed. The technology used in medical procedures has not been adequate to detect all diseases on time, and that is why some diseases turn out to become pandemics because they are hard to detect on time. Our focus is to clarify the details about the diseases and how to detect them promptly with artificial intelligence. We discuss the use of machine learning and deep learning to detect diseases in plants automatically. Our study also focuses on how machine learning methods have been moved from conventional machine learning to deep learning in the last five years. Furthermore, different data sets related to plant diseases are discussed in detail. The challenges and problems associated with the existing systems are also presented.

## 1. Introduction

The use of technology in the detection and analysis process increases the accuracy and reliability of these processes. For example, the people who use the latest technology to analyze the diseases that arise unexpectedly are at a higher chance of controlling them than those that do not. In the recent occurrence of coronavirus, the world relied on the latest technology to develop preventive measures that have helped reduce the rate at which the disease is transmitted. Crop diseases are a significant threat to human existence because they are likely to lead to droughts and famines. They also cause substantial losses in cases where farming is done for commercial purposes. The use of computer vision (CV) and machine learning (ML) could improve the detection and fighting of diseases. Computer vision is a form of artificial intelligence

(AI) that involves using computers to understand and identify objects. It is primarily applied in testing drivers, parking, and driving of self-driven vehicles and now in medical processes to detect and analyze objects [1]. Computer vision helps increase the accuracy of disease protection in plants, making it easy to have food security.

One of the areas that CV has helped most is the detection of the severity of the diseases. Deep learning (DL), a part of the CV, is useful and promising in determining the severity of diseases in plants and animals [2]. It is also used to classify diseases and avoid the late detection of diseases [1]. Plant diseases are slightly different from those that affect human beings. Many factors make diseases similar as well. However, the diseases that can be transmitted from humans to plants and vice versa are rare. The analysis of the data related to this field helps identify how the use of the latest technology can be

improved. The images of leaves and other parts of the plants can be used to detect diseases in plants [3]. The technology could be applied in analyzing images in human beings that also prove the presence of diseases and determine the extent of their destruction. This research study is aimed at analyzing the way image-based technology can be used in detecting diseases in both plants and animals.

## 2. Background

ML is the technology that allows machines to communicate with human beings and understand their needs. It also makes machines act like human beings and make the decision on behalf of humans. It is one of the areas that have grown fast over the past few years. ML helps in classifying plant diseases. The use of this technology is seen as a significant beginning and achievement in dealing with plant diseases. It has also increased productivity in the field of cultivation. Visualization techniques have also been included in this technology, and it has been improved over the last three years to the current improved levels [4]. The challenges that face the world today, related to the diseases affecting plants and humans, can be reduced if the diseases are identified before they spread to vast areas. The use of ML is widespread in the world today. Diverse methods used in ML and DL help the experts to analyze the plant diseases and know their source in time [4]. The detection of these diseases is affected mainly by several challenges that affect the effectiveness and accuracy of this technology.

The first challenge is the time complexity associated with the use of ML and DL, whereby some of the technologies used in the detection of these diseases are outdated or based on some information from the past. The other challenge is segmentation sensitivity [5]. It means that the region of interest (RoI) requires a high level of accuracy and sensitivity to acquire the required usage and accuracy. The other challenge is that there is a language barrier that affects the way the technology is applied. Another challenge is the inadequate resources that are required to support the application of this technology. Most of the ML and DL activities need many resources to use and implement. Private and government entities usually fund the institutes that use this technology to detect diseases in humans and plants, which could affect the success of the research and implementation of the technology.

The importance of plants in the world has increased over time. The discoveries about the critical roles that plants could play in medicine, energy production, and the recent concerns about the reduction of global warming have for long been a significant part of science and technology [6]. A reduction in the plant cover in the world increases the risk of higher global warming and an increase in the related challenges. The need to build a state-of-the-art convolutional system that supports the image detection technology and classification of plant diseases has led to many research programs to provide the scientists with the required knowledge [7]. Image detection could be applied when necessary to differentiate healthy leaves from those that are not healthy. The convolutional neural networks (CNNs) provide the differences among plant

images that help determine the abnormalities that could exist in the plants in the natural environment [6]. The background study shows that the scanning of the images that show the healthy and unhealthy plants forms a basis for comparison by the scientists in this field.

DL can be used to detect abnormalities in both humans and plants. The pixel-wise operations are used to analyze the leaves collected from sick plants, and this is used to classify the diseases according to their impact on the plants. The visible patterns in these leaves are used to decide the diseases that affect the plants and how they can be dealt with to prevent them from spreading. Research shows that the use of DL technology is up to 98.59% accurate [8]. The field of plant pathology has contributed immensely to the control of diseases and reduced global warming. One of the essential background knowledge that guides the use of image detection technology is that the leaves of the infected plants are different from the healthy ones. The leaves are likely to have dark parts, and some may be dry along the edges. The dried parts are also likely to fold up, and this is easy to detect even with a bare eye. The use of ML is to detect these differences without human intervention.

The ML methods used to make decisions on the detection of diseases include Artificial Neural Networks (ANN), decision trees, support vector machines (SVM), and K-means, among others [9]. The computers may not work directly with the images taken in the fields. The images are converted into data that the machines and computers can interpret. It means that the technology requires the coding of the images into data that can be fed into the computer systems. Therefore, the basic knowledge in computer codes and programming forms an integral part of the background knowledge required in this field.

Apart from the aforementioned application of the knowledge in plants and detecting the diseases on time, the image-based detection and recognition of diseases are applied in humans to detect the diseases that affect different body parts. Some human diseases have significant impacts on the tissues and organs that they affect. Adenocarcinoma of the prostate is one of the most common forms of cancer. It is detected using the image-based detection methods whereby the body is scanned for the abnormalities and the images obtained are used to determine if the patient has the disease. It is the second most diagnosed type of cancer in men, with about one in every nine men diagnosed with it in their lifetime [10]. The use of subjective tissue examination has been the primary way of detecting cancer in men suspected of it. The examination of these body tissues is largely dependent on the Gleason system [11]. The field of AI has offered many solutions to the challenge of inaccuracy in the Gleason scale and how it is used to provide solutions to people who have prostate cancer. The AI solutions are applied in analyzing the prostatectomy specimens to determine the impact of cancer on them. Before and after they have been affected by a disease like cancer, the appearance of the body organs is used to design the image-based detection process that is more accurate than the human eyes. In other words, computer systems can detect the difference in the organs more efficiently than humans. It helps to detect the disease long before they get to a fatal stage.

As for some plants like Cassava, it is possible to provide some protection to them from common diseases through image-based detection technology. Agriculture depends on some drought-resistant crops like Cassava to regulate the food supply and ensure an adequate supply of nutrients [12]. However, diseases threaten some of these crops' survival, which makes it hard for the agriculture departments to meet their targets. The background knowledge about the diseases could help avoid such instances. The use of CNN generates a reliable platform where the diseases are analyzed in detail [13]. The accuracy of this method is high and reliable. Further research by experts on the diseases that affect plants like Cassava has classified the diseases in terms of their impact on the leaves and other parts of the plants [14]. For example, many diseases affect the part of the leaves, some affect the entire leaf, and others attack the edges or the stalk. The images of the leaves can be analyzed using the image-based detection system to determine the classification that suits the disease.

In places like sub-Saharan Africa, Cassava is one of the essential foods because it provides people with many carbohydrates. However, due to its vulnerability to viral diseases, it has not effectively sustained the nutritional value it could achieve in this region. In 2014, about 145 million tons of Cassava was harvested in Africa [13]. The food control methods applied in the world today mainly focus on increasing production. The application of the latest technology to detect and control the disease that affects production is a reliable way of reducing the challenges of food inadequacy. Most of the plants are also used as raw materials in many industries. If such plants are of low quality, it means that they will lead to the production of low-standard products.

### 3. Literature Survey on Types of Plant Diseases

Fungi usually cause diseases that affect the plants, and they typically attack the leaves. Viral and bacterial pathogens cause many others. Precision in agriculture has improved with the increased use of ML and its related features [15]. The reduced production quantity in agriculture hurts many people and animals, which requires modern technology to solve. The extraction and detection of diseases are easier when the image-based detection system is used because of its high accuracy and reduced complications and duplication of data. In some plants like tomatoes, the use of the images to determine the diseases that affect them and the extent of the damages cannot be achieved unless there is a high accuracy rate [16]. The survey on plant diseases shows that many diverse factors determine how technology-based image detection is applied. In other words, the diseases that cause visible dents and changes on the plants are the ones that can be detected using this technology as opposed to the ones that cause damages that cannot be detected from the plants' images [17]. The analysis in this research shows that plant diseases are usually detected when they start showing an impact on the physical appearance of the plants.

The main challenge affecting the field of agriculture is the reduction in production and poor-quality production in plants. The challenge is a result of the poor detection and

management of the diseases that affect the plants. The challenge is also extended to affect human beings in several ways. The reduced plant cover due to plant diseases means that global warming, famine, and reduced air purification ensue. Hyperspectral imaging has become a reliable way of detecting crop diseases on time [18]. It is hard to determine the factors that lead to the diseases unless they are detected on time. In other words, if a disease is detected on time, it is easy to relate it to the possible factors that lead to its occurrence. For example, scientists could determine if there was a change in weather or climate that could have led to the occurrence of the disease.

Further research by [6] shows an inadequate database that could be used to provide background knowledge for comparing the images taken. The other challenge is that the symptoms and characteristics of the diseases are diverse and could be similar to a certain degree [19]. For example, many diseases could lead to the wilting of leaves. The challenge is yet to be resolved because more and new images are uploaded progressively by experts.

The other challenge is the lack of suitable instruments for use in the work of image detection. Most of the experts in the field do not have the equipment they require to analyze the images they get from the field, and this makes it hard for them to acquire accurate data and identify the diseases [20]. The other one is that there is a low rate of implementation in some areas due to the regulations put in place to ensure the credibility and reliability of the data from these analyses. For example, after the 4th and 6th International Conference on Machine Learning and Soft Computing, there have been many regulations that may derail the use of ML in some parts [21]. The rules discourage some of the results from the ML functions from being applied in practice because they do not meet the required parameters.

The technology has been in existence for several years now. However, there are still many issues that have not been clarified about its application. The other challenge is related to this fact. Some of the important images that could help determine if disease exists have not been captured. The other one is that the future perspectives of the research are not clear, and this is because of the increased diversity in the diseases that affect both humans and animals [19]. The application of image-based detection is also affected by the increased diversity in the way the diseases appear. Some of the diseases that used to affect the plants a few years ago have evolved into new forms, and they have different impacts and outcomes. It is difficult for the images to be used alone to conclude the diseases and choose a solution. Some of the solutions used in the past have also become ineffective, reducing the effectiveness of the technology.

The aforementioned challenges show that there are many possible ways in which the image-based detection could be applied, but the challenges reduce its usability. The first solution is to provide adequate data that can be used to identify the diseases accurately without confusing the ones that are closely related. The changes in weather, global warming, and other impacts have led to many diseases that have not been documented. The solution is to increase scientists' coverage and promote a better way of collecting information

[22]. The other solution would be to offer training to the scientists in this field to ensure that they are equipped to collect the data. Yet another solution is to create better ways of capturing the data collected about the diseases. The challenge of inadequate information about the diseases can be solved if there is an improved data-captioning process that involves fine details of the images taken and the differences that define them [6]. The images should be analyzed keenly to determine the ones that are affected or infected.

Another solution would be to focus on using the latest technology that is reliable and valid. The confusion that comes with the inadequate database for use in detecting the diseases results from inferior technology and low storage abilities of the existing systems. Most of the images are not stored correctly, which affects the accessibility of the information. It could be solved by the use of modern methods of storing information. For example, the use of cloud computing could help increase the accuracy of storage and accessibility. The other solution is training the people in charge of the research and analysis of the information. A trained DL algorithm increases the accuracy of the technology [23]. The other solution would be to understand the phenotypes used in detecting diseases [24]. The phenotype used in the detection of diseases is usually a result of climate and weather [23, 25]. The other solution would be to update the systems to ensure the data captured is up to date. The high level of uncertainty in the detection of diseases affects the way the technology is implemented. For example, the use of Bayesian DL is associated with several uncertainties [26]. It means that this method is unreliable if used alone.

A combination of several methods could reduce the inefficiencies of the processes. For example, the application of DL meta-architectures provides a solution to some of the errors that are experienced when using other methods of identifying diseases [27]. The other solution would be the application of deep convolutional generative adversarial networks that help in the identification and analysis of the images [28]. The contribution of the adversarial networks increases the accuracy of the detection process. The use of CNN methods could also be effective in dealing with the inaccuracy and slow identification processes of the diseases [29]. The methods have been used to detect the diseases that affect rice and have many benefits [28]. Image-based detection requires many resources, and the authorities should ensure they are available to ensure the activities are smooth.

#### 4. Plant Disease Image Data Sets

The data sets used in the research include the descriptions of the leaves before and after the diseases affect them. The data comprises tables and images of the leaves that are taken in the fields. The data is analyzed and classified in a way that is easy for the readers to understand. For example, Walleigh et al. [9] show the leaves used to determine the soybean plants affected by the diseases. The data set shows healthy leaves and the ones that had dents due to the attack by septorial leaf blight, others by frogeye leaf spot, and those affected by downy mildew as shown in Figure 1.

The images in Figure 1 show that there were visible differences between the leaves affected by the disease and those that had not. The data set was clear and easy to understand. Another form of data was the table that showed the number of leaves that were classified under each disease. The data set is clear and indicates the total number of leaves analyzed and classified into four categories. The other set of data is by [4], which shows the graphical representation of the captured information in the tables.

The data set used in the research can be shown in tables, texts, graphs, and other forms. However, an essential aspect of all of them is the ease of analysis and ease of understanding. Some of the data are also grouped as per the required levels. For example, data can be grouped in terms of the diseases they represent, the time they were collected, or the analysis method [30]. The data captured in some of the research outcomes also show the use of technology and its effectiveness. For instance, the data sets captured using technology allow for a controlled environment; the data sets show the type of control used and its effectiveness. For example, computer-aided diagnosis (CAD) systems were captured in the data analyzed by [31]. The data sets are provided to help understand the usage of this technology and its impact on the quality of the research. The same data shows the classification techniques used and gives reasons for the choice.

The other feature is that the data sets for leaves' analysis are based on the primary data collected in the fields. The reliability of the data is high because it is based on the observable features of the leaves. The data sets are also divided into sections that are easy to understand. For example, the work of [32] shows the divisions of the work in terms of the diseases like Rice Blast (RB), Bacterial leaf Blight (BLB), and Sheath Blight (SB). The use of a PlantVillage data set was also applied in the research by [32]. The data set consists of 54,306 images of 14 different crops representing 26 plant diseases. The images that were included in the data set included leaves having different colors. Figure 2 shows some samples of the PlantVillage data set. The colors indicate the parts of the leaves affected by the diseases that were under investigation. The authors also used the augmented data set proposed by Geetharamani and Pandian [33]. The ImageNet data set was also used in the research, and this led to high-quality research outcomes because of the synergy of combining various methods [32].

The data sets used in the research studies are dependent on the type of information contained. For example, the research by [14] focused on the effects of uncontrolled pests in China and the impacts on the total food produced. The research shows that poor control of pests in China leads to a loss of about 30% of the total foods produced [35]. The data sets in the research are large and show the different production levels and how the pests affect them. The use of a public data set also made it possible to understand the different ways in which the research was done and verified using data that are available in the public domain [36]. The PlantVillage data set was also used in [37] where 14 different types of leaves of cucumber plant are analyzed for seven different diseases. The data sets were mainly combined to provide a good presentation of the data that was collected. The use of the PlantVillage

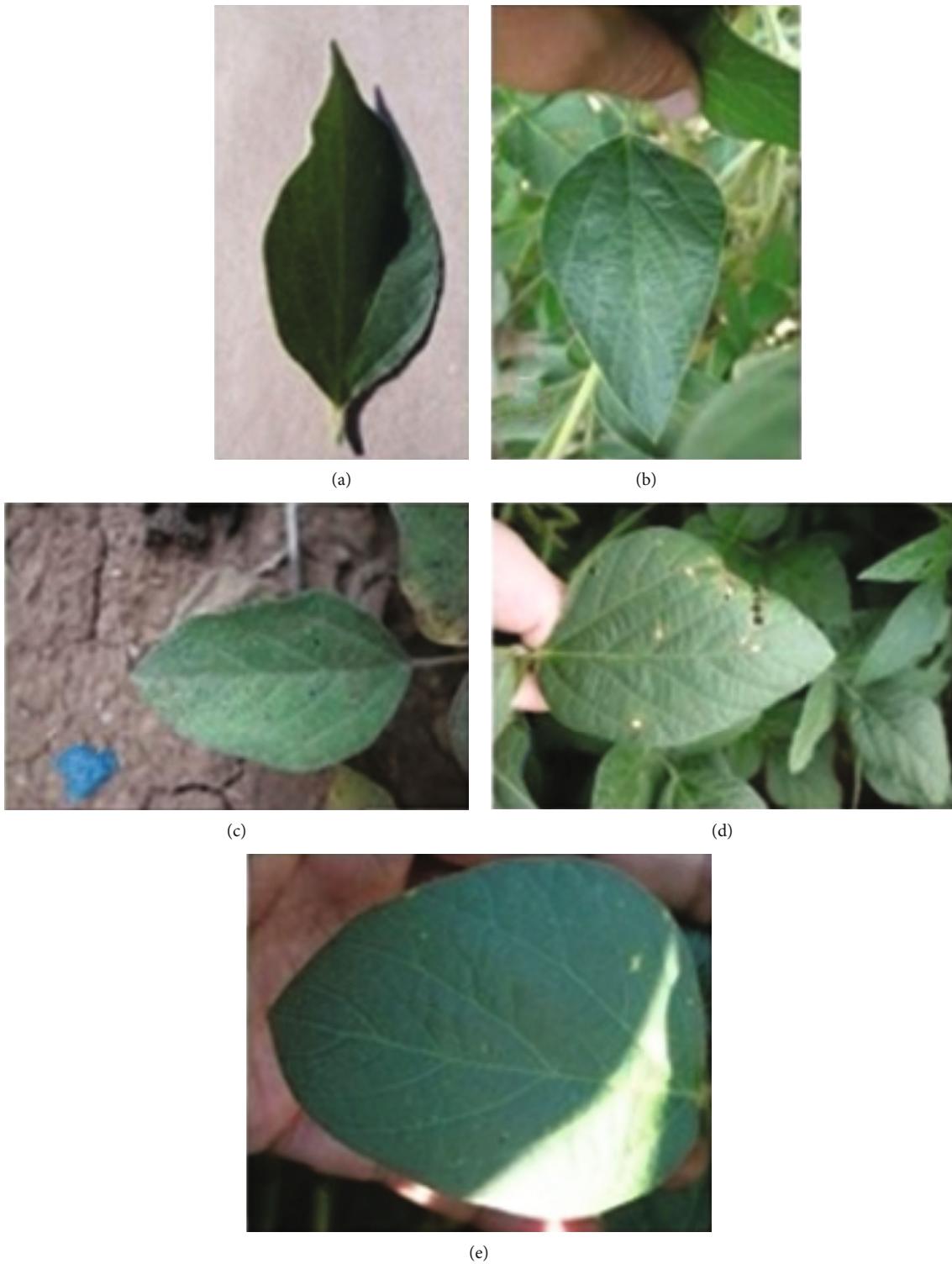


FIGURE 1: Data set used for classification by Wallelign et al. [9].

data set by [38] was used to show how the collected information was helpful in understanding plant diseases.

The data sets used in most references were from the data collected by different researchers and combined in one set. The reliability of the research could be compromised if there is no control over what is to be included in the data sets [39].

The nature of the research also determines how the data sets are used. For example, the collection of leaves and combining them in a table usually involves using the PlantVillage data set. The use of coffee leaf data sets in the research by [40] is backed by the need to show the diseases that affect the coffee leaves and tomatoes and how they can be detected using

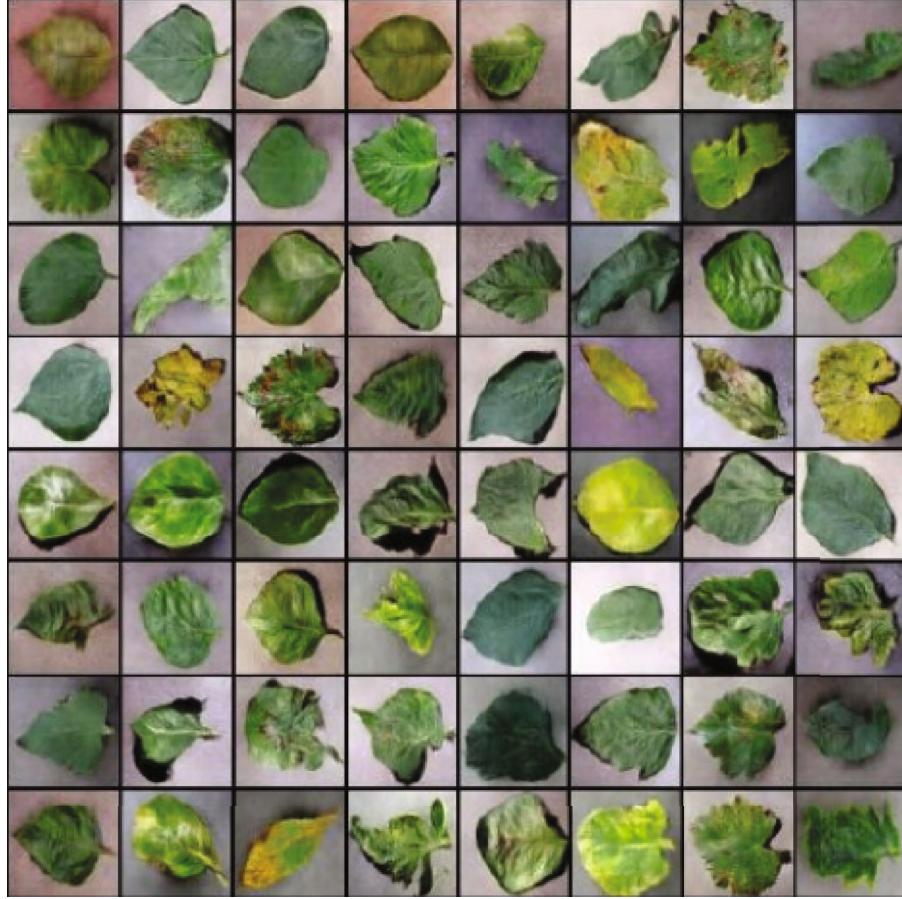


FIGURE 2: PlantVillage data set—examples of different phenotypes of various plants [34].

image-based detection methods [36]. In the research by [41], the authors used an information-rich color data set and large numerical data sets to display the data collected. The machine learning process also employs the training data set to predict and analyze unseen data. The data sets also include the expected data, making it easy to rely on the data in the research and determine if the data is valid. The use of data sets related to the nature of the research helps achieve the goals of the data collection and analysis [42]. The aforementioned data sets are affected by some challenges and limitations that reduce their applicability.

Another well-known data set used by the research community is Northern Leaf Blight (NLB), which contains infected maize leaves [1]. Some sample images from the NLB data set are shown in Figure 3. NLB consists of 1787 images having 7669 lesions. The images were obtained from maize plants in the field while using a handheld camera. The images in NLB were captured in uncontrolled conditions as opposed to the PlantVillage data set.

**4.1. Challenges and Problems with the Data Sets.** The explored data sets have several challenges and problems that affect their application. One of the challenges is how the data sets are organized to show the data collected from different fields. For example, when using the PlantVillage data set, the data collected from different fields could be hard to pres-

ent in the same way. Most of the data may be diverse. The missing uniformity could lead to a challenge in how the data is shown in the tables and other exhibits [43]. The other problem is that the data sets may be hard to convert into graphical representations. The information in the data sets could make more sense if the information is shown in graphical formats [44]. The data could also make more sense to the readers when presented in graphs. However, some of the data sets cannot be transformed into graphs directly and may require redoing the graphs entirely. For the data sets that include tables, it is possible to convert the data into graphs without having to draw manually. The other challenge is that some of the data sets do not summarize the data contained [45]. For example, when data is presented in the PlantVillage, a summary of the total value is given at the end of the table. However, in some data sets, this information is missing.

The other challenge with the data sets is that they may not capture all the information about the factors that lead to the values shown in them. For example, the data presented in work by [27] does not show the factors that affect the classification and detection process. Such missing details could affect the way the data is applied in making decisions about the diseases [46]. The data sets could also be complex for people to use if they do not have some prior training and education related to the statistical analysis of the data. The data sets also contain some details that may be inconsistent with the

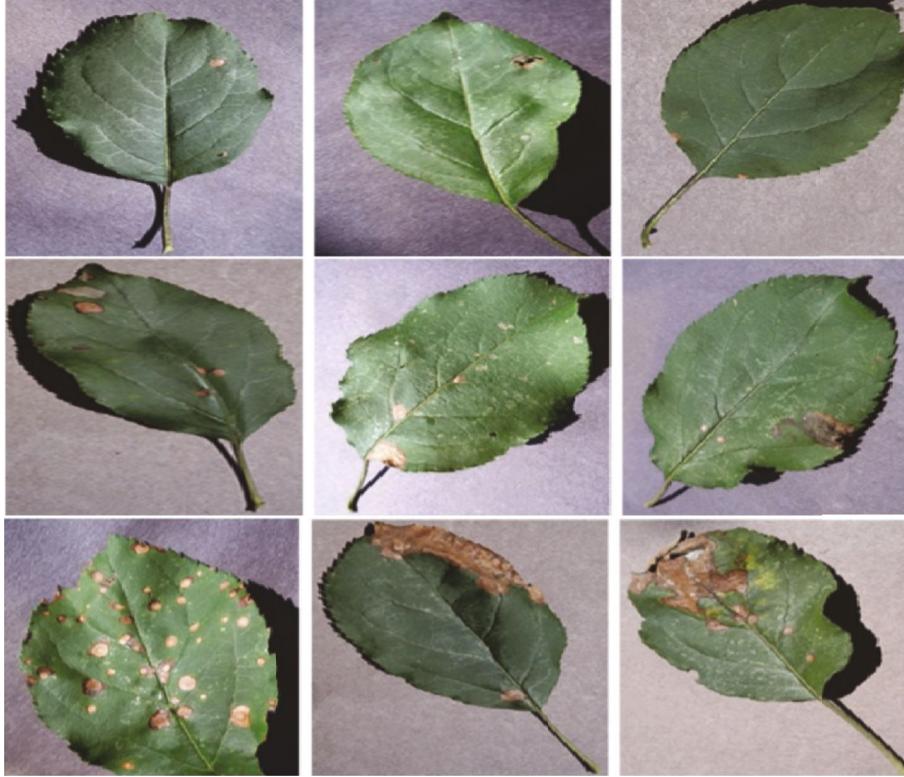


FIGURE 3: Sample leaf images from the Northern Leaf Blight data set [1].

research. For example, the data sets could contain information about the classification of the diseases and fail to capture the prevention measures that could be used to reduce the chances of having the diseases affecting the plants [47]. In some instances, the data sets are unclear, which could confuse the people who interpret them. Some of the data could show information about the diseases that affect coffee or rice and fail to contain the exact impact on production [20]. The exact impact on the quantity and quality of the product should be captured in such research. The other challenge is that some of the data sets may not allow the use of mobile gadgets [13]. Their complications make it hard to analyze their contents using mobile devices because of a higher screen resolution requirement.

## 5. Feature Extraction for Disease Identification

The images of the plants have three key features, namely, color, shape, and texture. Compared to color and texture, the shape feature cannot help find the plant's diseases [48]. Hlaing and Zaw [48] classified tomato plant disease using a combination of texture and color features. They used the Scale Invariant Feature Transform (SIFT) to find the texture information, containing details about the shape, location, and scale. Similarly, they gathered the color details from the RGB channel.

Dandawate and Kokare [49] developed an approach for the automatic detection of diseases in soybean plants. They converted the image from RGB to HSV (Hue Saturation Value) color space. Color and cluster-based methods were

employed for segmentation. The SIFT method was used to detect the type of plant, based on the shape of the leaf.

Pydipati et al. [50] identified the citrus disease using color texture features along with discriminant analysis. They also employed the color cooccurrence method (CCM) to determine if hue, saturation, and intensity (HSI) color features and the statistical classification algorithms could help identify the diseased leaves. Their method achieved an accuracy of more than 0.95.

Al-bayati and Üstündağ [51] extracted only the area of the leaf affected by the disease. Furthermore, they used feature fusion which helped in feature reduction.

Image-based detection requires many resources, and the authorities should ensure they are available so that the activities are smooth. DL in general and CNN in particular have been developed to analyze multidimensional data such as images. The underlying model is based on the multilayer ANN. Nevertheless, a convolutional layer performs kernel operations over various areas of the provided image. The obtained representation is independent of the operations such as translation or rotation. These kinds of features have been proved to work better as compared to the traditional features earlier used in the detection of plant diseases [2].

## 6. Comparison of Performance and Result Discussion

Our current research work covers state-of-the-art plant disease recognition using AI in the last five years. We summarize a series of observations that emerge from this work in the following paragraphs:

- (i) *Available Databases and Size Issue.* It is difficult to obtain leaf images for specific plant infections. Due to this fact, the sizes of the available plant data sets are very small. Only limited works have reported thousands of images for research purposes [22, 52–56]. Due to the small database size problem, a large portion of the data set is used for the training phase in most of the deep learning methods. However, very few exceptions are there, for example [57–59]. Furthermore, the available database images are collected in very constrained environmental conditions. We believe that images must be gathered in real-world conditions to make the algorithms more practical. Efficient image acquisition of leaf images is the need of the hour. If these images are captured in real-time scenarios, such databases would be warmly welcomed in the research community. In most of the recently reported works, the images captured with smart mobile devices are gaining popularity. Some single-click image systems are also introduced, but much more is supposed to be done by the researchers to automate plant disease identification algorithms. The transition of image-capturing systems to smart devices may help overcome serious issues related to database size.
- (ii) *Issues with Available Feature Extraction Methods.* Performing the tasks of preprocessing, feature extraction, and segmentation plays a key role in developing a machine learning-based algorithm. Selecting the most suitable method for preprocessing and segmentation further depends on the nature of the data set. Among many techniques, one that is most suitable for a specific acquisition usually serves the purpose. We observe variability span in the reported algorithms so far under different modules. We observe somehow similar observations for various feature extraction techniques. In a nutshell, the standardization of the reported methods is yet to be fixed and achieved.
- (iii) *Difficulties in Classification Module.* Plant disease automation and detection is an active area of research for a long time. Considering very few images for training and testing, highly acceptable results are reported by researchers. Many classifiers are explored by researchers in this domain. This study concluded that backpropagation neural network, SVM, and discriminant analysis (particularly linear) perform much better than others. These are then followed by Naïve Bayes, random forest, k-nearest neighbor, and multilayer perceptron. However, state-of-the-art results are much improved with recently introduced optimized deep neural networks. More proper utilization of deep convolutional neural networks can help in improving the results for large data sets.
- (iv) *Limitations of Available Systems.* We argue that image analysis methods are comparatively better

than the techniques that visually rate the severity of a particular disease. However, systems which are designed using these imaging techniques are not perfect. The performance of a system highly depends on the quality of the training data. In plant disease automation, it is the training images and certain extracted features, which significantly affect the performance of a system. A system trained with good quality data is trained well. However, most of the existing systems have a specific set of requirements needed to be fulfilled for a system to perform accurately. If some of these constraints are not fulfilled, the system may give inaccurate results, ultimately leading to wrong disease detection. For example, most of the DL-based methods particularly and conventional machine learning methods generally are faced with the problem of overfitting. Researchers must think of adaptive systems which are designed with more flexible requirements. Additionally, some generalized methods should be adapted which work in heterogeneous environments. For improving efficiency, in-depth knowledge of the methods and proper usage of the tools are also necessary.

- (v) *Evaluation Measures.* Many measures are available to compare different models for disease classification. These measures are based on four statuses: true-positive (TP), indicating the number of infected samples correctly identified; true-negative (TN), describing the correctly identified healthy images. Similarly, false-positives (FPs) showing the number of healthy samples that have been incorrectly classified as infected ones. Lastly, false-negative (FN) represents the infected samples wrongly categorized as healthy ones. Accuracy is the ratio of the correct classifications ( $TP + FP$ ) to the total number of classifications ( $TP + FP + TN + FN$ ). Precision represents the ratio of the correctly identified samples as infected (TP) to the total samples identified as infected (sum of TP and FP). Similarly, recall is the ratio of TP to the actual number of infected samples (sum of TP and FN). Lastly, F-measure represents the harmonic mean of precision and recall.
- (vi) *Comparison of Results.* State-of-the-art results for plants disease detection are compared and summarized for various data sets and methods in Table 1. We present a summary of the observations in the following paragraphs:
- (a) Dandawate and Kokare [49] used SVM, a linear supervised learning classifier, for the classification of soybean plant diseases. Their proposed system got an average accuracy of 0.938. Al-bayati and Üstündag [51] used SVM and ANN for the detection of plant diseases. They tested their approaches on the PlantVillage data set. Among various fruits and vegetables, strawberry got the highest precision of

TABLE 1: Image-based detection and recognition of diseases (year wise from 2016 to 2021).

Paper	Year	Database used	Images used	Accuracy
Atila et al. [32]	2021	PlantVillage	61,486	0.984
Reddy et al. [61]	2021	PlantVillage	54,305	0.900
Sembiring et al. [62]	2021	PlantVillage	17,641	0.972
Adit et al. [12]	2020	PlantVillage	76,000	0.980
Guo et al. [46]	2020	PlantVillage	61,486	0.992
Khan et al. [60]	2020	PlantVillage	7,733	0.978
Saleem et al. [27]	2020	PlantVillage	61,486	0.730
Ashqar and Abu-Naser [20]	2019	PlantVillage	9,000	0.994
Goncharov et al. [38]	2019	PlantVillage	54,306	0.870
Lin et al. [5]	2019	Powdery mildew data set	20	0.961
Nagasubramanian et al. [18]	2019	Hyperspectral image data set	111	0.957
Ozguven et al. [39]	2019	Sugar beet data set	155	0.955
Ramcharan et al. [7]	2019	Cassava disease data set	720	0.790
Shrivastava et al. [31]	2019	Oryza sativa database	619	0.914
Stewart et al. [25]	2019	Aerial image data set	3,000	0.780
Tian et al. [45]	2019	Apple disease image data set	640	0.994
Turkoglu et al. [15]	2019	Turkey plant data set	1,965	0.956
Verma et al. [24]	2019	IPPN phenotyping data set	345	0.904
Barbedo [6]	2018	PlantVillage augmented	50,000	0.960
Barbedo [19]	2018	PlantVillage augmented	50,000	0.980
Cap et al. [37]	2018	Saitama research center data set	60,000	0.780
Dhakal and Shakya [8]	2018	PlantVillage	54,000	0.985
Ferentinos [3]	2018	Open plant database	87,848	0.995
Rangarajan et al. [23]	2018	PlantVillage	13,262	0.950
Ubbens et al. [14]	2018	PlantVillage	18,160	0.962
Walleigh et al. [9]	2018	Soybean image data set	12,673	0.993
Wiesner-Hanks et al. [10]	2018	Cornell University data set	18,222	0.850
Amara et al. [35]	2017	PlantVillage	3700	0.997
Brahimi et al. [36]	2017	PlantVillage	14,828	0.992
Cruz et al. [30]	2017	PlantVillage	299	0.986
Durmus et al. [21]	2017	Korean tomato disease data set	5,000	0.972
Fuentes et al. [16]	2017	Tomato disease image data set	5,000	0.870
Ramcharan et al. [13]	2017	Cassava disease image data set	11,670	0.930
Wang et al. [1]	2017	PlantVillage	2,265	0.904
Ale et al. [43]	2016	Real-world plant disease data set	21,917	0.980
Sladojevic et al. [22]	2016	Images collected from Internet	33,469	0.982

0.991. Moreover, the results improved with the fusion of classifiers. In terms of recall, the best results were also obtained for strawberry (0.959). It was observed that the fusion improved the results for recall, just like the case for precision. As expected, the best F-measure was observed for strawberry (0.975). Furthermore, the F-measure results were better while using the fusion of classifiers

- (b) Saleem et al. [27] developed three DL architectures to detect various plant diseases, namely, Single Shot Multibox Detector (SSD), Faster Region-based CNN (Faster-RCNN), and Region-based Fully Convolutional Networks (RFCN). SSD performed all

computations in one network and used smaller convolution filters, such as  $4 \times 4$  and  $8 \times 8$ . Similarly, in Faster-RCNN, object detection is carried out at two different stages in contrast to SSD. In the first step, also known as the region proposal network stage, the images are processed to come up with the region proposals using feature extractors. Consequently, the features help in determining the class-specific proposals for each of the intermediate convolutional layers. Later, during the second step, the characteristics of the same intermediate layer of an image are detected. RFCN is very similar to Faster-RCNN but does not include fully convolutional layers after the region of interest (ROI) pooling

- (c) Hernandez and Lopez [26] used Stochastic Gradient Descent (SGD) to train a softmax layer. The convergence was achieved after 60 epochs. SGD got an accuracy of 0.96 and an F1-score of 0.96 as well, while fine-tuning VGG16, a state-of-the-art image classifier, using Bayesian learning algorithms. Monte Carlo (MC) dropout was able to get both accuracy and an F1-score of 0.94. Lastly, Stochastic gradient Langevin dynamics (SGLD) achieved an accuracy of 0.89, whereas an F1-score of 0.88. SGD achieved an accuracy of 0.49 for out-of-sample classification. On the other hand, Monte Carlo (MC) dropout was able to achieve an accuracy of 0.55. It is pertinent to note that MC dropout can be used both during training as well as testing. SGLD got an accuracy of just 0.54, whereas an F1-score of 0.15 for out-of-sample classification. Kamaris and Prenafeta-Boldu [2] report that many papers related to plant and leave disease detection and employing DL techniques report excellent results, i.e., an accuracy greater than 0.95 or F1-score greater than 0.92. The reason is probably the peculiar characteristics of (sick) leaves/plants and fruits in the images
- (d) Brahim et al. [36] used AlexNet and GoogLeNet on the PlantVillage data set and obtained an accuracy of 0.99. Similarly, Dhakal and Shakya [8] also used DL for the PlantVillage data set. They considered four class labels: Bacterial Spot, Yellow Leaf Curl Virus, Late Blight, and Healthy Leaf. They were able to get a test accuracy of 0.956. Khan et al. [60] proposed a CNN-based model and worked on five different diseases found in the PlantVillage data set. The method got an accuracy of 0.978. Walleigh et al. [9] used CNN to identify the diseases found in the soybean plant. They used 12,673 images from the PlantVillage data set. The data set was unbalanced and contained 6234 instances of healthy leaf, 3565 images of septorial leaf blight, 2023 images of frogeye leaf spot, whereas just 851 images of downy mildew. They used three convolutional layers, each followed by a max-pooling layer. ReLU activation function was employed in each of the convolutional layers and the fully connected (FC) layer. Their results showed that it is much better to use color images than grayscale or segmented ones. Their approach was able to get a test accuracy of 0.993 after dropout. Nevertheless, the best model got an F1-score of 0.99
- (e) Reddy et al. [61] used CNN to identify the plant species while using color images of leaves. They used five data sets, including PlantVillage, Leaf Snap, UCI leaf, Flavia, and Swedish leaf. Their model uses four convolutional layers, two FC layers, and a softmax layer. The proposed method gets perfect accuracy for three data sets, Flavia, Swedish leaf, and UCI leaf. Similarly, an accuracy of 0.980 is obtained for Leaf Snap, whereas 0.900 is obtained for the PlantVillage data

set. Sembiring et al. [62] developed a concise CNN for detecting the diseases in the tomato plant. They also used the PlantVillage data set and got an accuracy of 0.972. The accuracy is slightly less than the one achieved by VGG16 (0.983). Nevertheless, the proposed method required less than one-fourth of the time needed by VGG16. Table 1 provides a comparison of different approaches

## 7. Conclusion and Future Works

In this paper, we have discussed how ML in general and DL in particular have helped identify the diseases in plants. If the diseases are not correctly identified, they affect the crop yield and ultimately result in long-term issues, such as global warming and even famine. The proposed work summarizes multiple studies regarding plant disease automation and identification through different ML methods. The proposed manuscript also shows well-acceptance of a range of CV methods in this domain, making it a wide area of research to be explored in the near future. Some points are summarized in the following paragraphs, which may help to improve and enhance the current state-of-the-art and give researchers some potential ideas to further explore the field in the future.

- (i) *Disease Stage Identification.* Disease stage identification is one of the main areas to be explored regarding plant disease identification. Each disease has several stages. Most of the researchers only focused their work on the type of disease identification, but none of these works target a particular disease stage identification. Additionally, such systems must possess the capability to suggest some measures depending particularly on specific disease stages. Identification of disease forecasting will help agriculturists to take proper actions and precautions to reduce damage percentage.
- (ii) *Quantification of a Disease.* Another exciting area to be explored is the quantification of a particular disease. Although there has been much work done in this field, very few researchers have identified the extent of the damage caused by the disease. They can help a lot since remedial actions could be taken according to the severity of the disease. Such kind of quantification will detect the infected proportion of a particular culture with some disease. This research perspective is vital as the number of pesticides can be controlled. Normally, the farmers apply chemicals to cure diseases without any prior analysis or quantification. Such kind of practice is extremely harmful to human health. Developing an efficient image processing application will help to determine if specific chemicals are needed or not.
- (iii) *Mobile and Online Applications.* Literature reported several solutions regarding applications of disease identification. However, few of the portals and mobile applications are publicly available and

accessible online. Some of these applications are Assess Software and Leaf Doctor, which are publicly available for use. However, these applications work on leave images with a flat and only black background. Therefore, such kinds of online systems and applications are highly needed for plant diseases' recognition. The availability of these types of software will help farmers to identify a particular kind of disease. Such kinds of software may be used to get analysis reports which can be sent to a disease expert for getting some suggestions.

- (iv) *Exploring Transfer Learning to Increase Data Size.* Similarly, noting the current trends of the developments in CV, which are moving very fast towards DL methods, for plant disease detection is not very satisfactory. Given the difficulty of the data, particularly the training stage, the best option to be explored is transfer learning. For investigating knowledge transfer, a heterogeneous domain strategy may be adapted. Considering automatic plant disease identification, the keywords which may be explored are LSTMs, optical flow frames, temporal pooling, and 3D convolution. The last point that must be remembered is that better and carefully engineered methods are needed for further exploring this area. For example, the case of data augmentation may be further investigated.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgments

The authors gratefully acknowledge Qassim University, represented by the Deanship of Scientific Research, on the financial support for this research under the number (10269-coc-2020-1-3-I) during the academic year 1441 AH/2020 AD.

## References

- [1] G. Wang, Y. Sun, and J. Wang, "Automatic image-based plant disease severity estimation using deep learning," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 2917536, 8 pages, 2017.
- [2] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: a survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [3] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- [4] R. I. Hasan, S. M. Yusuf, and L. Alzubaidi, "Review of the state of the art of deep learning for plant diseases: a broad analysis and discussion," *Plants*, vol. 9, no. 10, article 1302, 2020.
- [5] K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, "Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network," *Frontiers in Plant Science*, vol. 10, p. 155, 2019.
- [6] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.
- [7] A. Ramcharan, P. McCloskey, K. Baranowski et al., "A mobile-based deep learning model for cassava disease diagnosis," *Frontiers in Plant Science*, vol. 10, p. 272, 2019.
- [8] A. Dhakal and S. Shakya, "Image-based plant disease detection with deep learning," *International Journal of Computer Trends and Technology*, vol. 61, no. 1, pp. 26–29, 2018.
- [9] S. Wallelign, M. Polceanu, and C. Buche, "Soybean plant disease identification using convolutional neural network," in *Proc. Thirty-First International Florida Artificial Intelligence Research Society Conference (FLAIRS-31)*, pp. 146–151, Melbourne, FL, USA, 2018.
- [10] T. Wiesner-Hanks, E. L. Stewart, N. Kaczmar et al., "Image set for deep learning: field images of maize annotated with disease symptoms," *BMC Research Notes*, vol. 11, no. 1, p. 440, 2018.
- [11] K. Nagpal, D. Foote, Y. Liu et al., "Publisher correction: development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer," *NPJ Digital Medicine*, vol. 2, no. 1, p. 113, 2019.
- [12] V. V. Adit, C. V. Rubesh, S. S. Bharathi, G. Santhiya, and R. Anuradha, "A Comparison of Deep Learning Algorithms for Plant Disease Classification," in *Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies, Lecture Notes in Electrical Engineering*, vol. 643, pp. 153–161, Springer, Singapore, 2020.
- [13] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Frontiers in Plant Science*, vol. 8, article 1852, 2017.
- [14] S. Verma, A. Chug, A. P. Singh, S. Sharma, and P. Rajvanshi, "Deep Learning-Based Mobile Application for Plant Disease Diagnosis: A Proof of Concept With a Case Study on Tomato Plant," in *Applications of Image Processing and Soft Computing Systems in Agriculture*, pp. 242–271, IGI Global, 2019.
- [15] M. Türkoğlu and D. Hanbay, "Plant disease and pest detection using deep learning-based features," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 3, pp. 1636–1651, 2019.
- [16] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, article 2022, 2017.
- [17] M. Loey, A. ElSawy, and M. Afify, "Deep learning in plant diseases detection for agricultural crops," *International Journal of Service Science, Management, Engineering, and Technology*, vol. 11, no. 2, pp. 41–58, 2020.
- [18] K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease identification using explainable 3D deep learning on hyperspectral images," *Plant Methods*, vol. 15, no. 1, p. 98, 2019.
- [19] J. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Computers and Electronics in Agriculture*, vol. 153, pp. 46–53, 2018.
- [20] B. Ashqar and S. Abu-Naser, "Image-based tomato leaves diseases detection using deep learning," *International Journal of Engineering Research*, vol. 2, no. 12, pp. 10–16, 2019.
- [21] H. Durmus, E. O. Gunes, and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," in *Proc.*

- 6th International Conference on Agro-Geoinformatics, pp. 1–5, Fairfax, VA, 2017.
- [22] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep neural networks based recognition of plant diseases by leaf image classification,” *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 11 pages, 2016.
  - [23] A. Rangarajan, R. Purushothaman, and A. Ramesh, “Tomato crop disease classification using pre-trained deep learning algorithm,” *Procedia Computer Science*, vol. 133, pp. 1040–1047, 2018.
  - [24] J. Ubbens and I. Stavness, “Corrigendum: deep plant phenomics: a deep learning platform for complex plant phenotyping tasks,” *Frontiers in Plant Science*, vol. 8, article 2245, 2018.
  - [25] E. L. Stewart, T. Wiesner-Hanks, N. Kaczmar et al., “Quantitative phenotyping of Northern Leaf Blight in UAV images using deep learning,” *Remote Sensing*, vol. 11, no. 19, article 2209, 2019.
  - [26] S. Hernández and J. L. López, “Uncertainty quantification for plant disease detection using Bayesian deep learning,” *Applied Soft Computing*, vol. 96, article 106597, 2020.
  - [27] M. H. Saleem, S. Khanchi, J. Potgieter, and K. M. Arif, “Image-based plant disease identification by deep learning meta-architectures,” *Plants*, vol. 9, no. 11, article 1451, 2020.
  - [28] J. Li, J. Jia, and D. Xu, “Unsupervised representation learning of image-based plant disease with deep convolutional generative adversarial networks,” in *Proc. 37th Chinese Control Conference (CCC)*, pp. 9159–9163, Wuhan, China, 2018.
  - [29] A. Singh, B. Ganapathysubramanian, S. Sarkar, and A. Singh, “Deep learning for plant stress phenotyping: trends and future perspectives,” *Trends in Plant Science*, vol. 23, no. 10, pp. 883–898, 2018.
  - [30] A. Cruz, A. Luvisi, L. De Bellis, and Y. Ampatzidis, “Vision-based plant disease detection system using transfer and deep learning,” in *Proc. 2017 ASABE Annual International Meeting*, Spokane, WA, USA, 2017.
  - [31] V. Shrivastava, M. Pradhan, S. Minz, and M. Thakur, “Rice Plant disease classification using transfer learning of deep convolution neural network,” *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-3/W6, pp. 631–635, 2019.
  - [32] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, “Plant leaf disease classification using EfficientNet deep learning model,” *Ecological Informatics*, vol. 61, article 101182, 2021.
  - [33] G. Geetharamani and J. A. Pandian, “Identification of plant leaf diseases using a nine-layer deep convolutional neural network,” *Computers & Electrical Engineering*, vol. 76, pp. 323–338, 2019.
  - [34] D. P. Hughes and M. Salathe, “An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing,” 2015, <https://arxiv.org/abs/1511.08060>.
  - [35] J. Amara, B. Bouaziz, and A. Albergawy, “A deep learning-based approach for banana leaf diseases classification,” in *Datenbanksysteme für Business, Technologie und Web*, Stuttgart, 2017.
  - [36] M. Brahimi, K. Boukhalfa, and A. Moussaoui, “Deep learning for tomato diseases: classification and symptoms visualization,” *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, 2017.
  - [37] H. Q. Cap, K. Suwa, E. Fujita, S. Kagiwada, H. Uga, and H. Iyatomi, “A deep learning approach for on-site plant leaf detection,” in *Proc. IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)*, pp. 118–122, Batu Ferringhi, 2018.
  - [38] P. Goncharov, G. Ososkov, A. Nechaevskiy, and I. Nestsiarenko, “Disease detection on the plant leaves by deep learning,” in *Selected Papers from the XX International Conference on Neuroinformatics, in Advances in Neural Computation, Machine Learning, and Cognitive Research II*, pp. 151–159, Moscow, Russia, 2019.
  - [39] M. M. Ozguven and K. Adem, “Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms,” *Physica A: Statistical Mechanics and its Applications*, vol. 535, article 122537, 2019.
  - [40] J. G. M. Esgario, R. A. Krohling, and J. A. Ventura, “Deep learning for classification and severity estimation of coffee leaf biotic stress,” *Computers and Electronics in Agriculture*, vol. 169, article 105162, 2020.
  - [41] A. Lowe, N. Harrison, and A. P. French, “Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress,” *Plant Methods*, vol. 13, no. 1, p. 80, 2017.
  - [42] A. Mahlein, “Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping,” *Plant Disease*, vol. 100, no. 2, pp. 241–251, 2016.
  - [43] L. Ale, A. Sheta, L. Li, Y. Wang, and N. Zhang, “Deep learning based plant disease detection for smart agriculture,” in *Proc. 2019 IEEE Globecom Workshops*, pp. 1–6, Waikoloa, HI, USA, 2019.
  - [44] R. Chalapathy and S. Chawla, “Deep learning for anomaly detection: a survey,” 2019, 2020, <https://arxiv.org/abs/1901.03407>.
  - [45] Y. Tian, G. Yang, Z. Wang, E. Li, and Z. Liang, “Detection of apple lesions in orchards based on deep learning methods of cyclegan and yolov3-dense,” *Journal of Sensors*, vol. 2019, Article ID 7630926, 13 pages, 2019.
  - [46] Y. Guo, J. Zhang, C. Yin et al., “Plant disease identification based on deep learning algorithm in smart farming,” *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 2479172, 11 pages, 2020.
  - [47] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, “Solving current limitations of deep learning based approaches for plant disease detection,” *Symmetry*, vol. 11, no. 7, p. 939, 2019.
  - [48] C. S. Hlaing and S. M. M. Zaw, “Tomato plant diseases classification using statistical texture feature and color feature,” in *Proc. IEEE/ACIS 17th International Conference on Computer and Information Science*, Singapore, 2018.
  - [49] Y. Dandawate and R. Kokare, “An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective,” in *Proc. IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Kochi, India, 2015.
  - [50] R. Pydipati, T. F. Burks, and W. S. Lee, “Identification of citrus disease using color texture features and discriminant analysis,” *Computers and Electronics in Agriculture*, vol. 52, no. 1–2, pp. 49–59, 2006.
  - [51] J. S. H. Al-bayati and B. B. Üstündağ, “Evolutionary feature optimization for plant leaf disease detection by deep neural networks,” *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 12–23, 2020.

- [52] A. Meunkaewjinda, P. Kumsawat, K. Attakitmongkol, and A. Srikaew, "Grape leaf disease detection from color imagery using hybrid intelligent system," in *Proc. IEEE 5th International Conference on Electrical engineering/electronics, Computer, Telecommunications and Information technology ECTI-CON*, pp. 513–516, Krabi, 2008.
- [53] R. D. L. Pires, D. N. Gonçalves, J. P. M. Oruê et al., "Local descriptors for soybean disease recognition," *Computers and Electronics in Agriculture*, vol. 125, pp. 48–55, 2016.
- [54] S. Shrivastava, S. K. Singh, and D. S. Hooda, "Color sensing and image processing-based automatic soybean plant foliar disease severity detection and estimation," *Multimedia Tools and Applications*, vol. 74, no. 24, pp. 11467–11484, 2015.
- [55] M. Schikora, B. Neupane, S. Madhogaria et al., "An image classification approach to analyze the suppression of plant immunity by the human pathogen *Salmonella Typhimurium*," *BMC Bioinformatics*, vol. 13, no. 1, p. 171, 2012.
- [56] J. G. A. Barbedo, L. V. Koenigkan, and T. T. Santos, "Identifying multiple plant diseases using digital image processing," *Biosystems Engineering*, vol. 147, pp. 104–116, 2016.
- [57] D. Majumdar, D. K. Kole, A. Chakraborty, and D. D. Majumder, "An integrated digital image analysis system for detection, recognition and diagnosis of disease in wheat leaves," in *Proc. 3rd international ACM Symposium on Women in Computing and Informatics*, pp. 400–405, New York, USA, 2015.
- [58] S. Kai, L. Zhikun, S. Hang, and G. Chunhong, "A research of maize disease image recognition of corn based on BP networks," in *Proc. of 3rd IEEE International conference on measuring technology and mechatronics automation (ICMTMA)*, pp. 246–249, Shanghai, China, 2011.
- [59] S. N. Ghaiwat and P. Arora, "Cotton Leaf Disease Detection by Feature Extraction," in *Research Advances in the Integration of Big Data and Smart Computing*, pp. 89–104, IGI Global, 2016.
- [60] M. A. Khan, T. Akram, M. Sharif, and T. Saba, "Fruits diseases classification: exploiting a hierarchical framework for deep features fusion and selection," *Multimedia Tools and Applications*, vol. 79, no. 35-36, pp. 25763–25783, 2020.
- [61] S. R. G. Reddy, G. P. S. Varma, and R. L. Davuluri, "Optimized convolutional neural network model for plant species identification from leaf images using computer vision," *International Journal of Speech Technology*, 2021.
- [62] A. Sembiring, Y. Away, F. Arnia, and R. Muharar, "Development of concise convolutional neural network for tomato plant disease classification based on leaf images," *Journal of Physics: Conference Series*, vol. 1845, article 012009, 2021.