A Review of CNN-Based Techniques for Accurate Plant Disease Detection

Bahaa Samy Hamed
Computer Science Department
(Faculty of Computers and Information Menoufia University)
Shebin Elkom, 32511, Egypt
samybahaa642@gmail.com

Mahmoud Hussein
Computer Science Department
(Faculty of Computers and Information Menoufia University)
Shebin Elkom, 32511, Egypt
mahmoud.hussein@ci.menofia.edu.eg

Afaf Mousa
Computer Science Department
(Faculty of Computers and Information Menoufia University)
Shebin Elkom, 32511, Egypt
afafmousa11@gmail.com

***Abstract*— Various techniques have revolutionized the field of plant disease detection, offering accurate approaches for timely detection and recognition of crop diseases. This comprehensive review explores the current utilization of diverse techniques for plant disease detection and classification. It analyzes recent publications, considering aspects such as disease detection methods and dataset characteristics. These techniques have significantly advanced object detection and recognition in agriculture, facilitating efficient crop management and higher yields. However, the complexity of identifying and detecting plant diseases from images necessitates species-specific detection for customized control strategies. This study discusses the challenges and proposed solutions associated with the use of different techniques in early disease detection concentrated on deep learning methods. Overall, the review demonstrates the considerable potential of these techniques in disease detection and emphasizes the ongoing need for research and development to address current challenges and optimize their benefits in agriculture. and also underscores the importance of incorporating emerging technologies and data-driven approaches to further enhance the precision and scalability of plant disease detection systems.**

Keywords—Deep learning; CNN; plant disease datasets; pre-trained models.

I. INTRODUCTION

Plant diseases have a profound impact on food production, resulting in considerable reductions in yield and economic consequences. The Food and Agriculture Organization (FAO) reports that plant diseases and pests contribute to a loss of 20% to 40% in global food production. Additionally, approximately 13% of the global crop yield loss can be attributed to plant diseases [1-2]. These alarming statistics emphasize the critical importance of identifying and effectively managing plant diseases.

In the majority of cases, plant diseases initially appear at the lower sections of plants and subsequently extend throughout the entire crop. It is essential to monitor crops regularly to detect diseases early and prevent their further proliferation. Certain diseases may also emerge later in the season, specifically after the pollination stage [3]. Plant diseases can affect different organs, with foliar diseases being the most visibly recognizable through visual examination as they exhibit symptoms on the leaves. Fungal diseases, in particular, have a substantial impact on crop yield, leading to significant losses. In recent studies, computer vision, machine learning, and deep learning techniques have been utilized to identify plant diseases using images of plant leaves. Effective plant disease diagnosis involves early detection of diseases during the growing season, detection of multiple diseases across different crops, estimation of disease severity, determination of optimal pesticide volumes for application, and implementation of appropriate disease management strategies to limit its spread. The integration of advanced technologies and techniques, such as computer vision, image processing, and deep learning, presents promising opportunities for precise and timely detection and management of plant diseases [4]. Researchers are leveraging these tools to improve early detection capabilities, optimize treatment strategies, and mitigate the impact of plant diseases on crop yields [5]. Traditional methods of plant disease detection often require manual feature extraction, which is time-consuming and inconsistent. Machine learning approaches, while explored, still rely on feature extraction and may have limited effectiveness. However, deep learning techniques have emerged as a promising solution. Deep learning models can automatically extract relevant features, eliminating the need for manual extraction. In the field of plant disease detection, classical machine learning algorithms combined with image processing techniques have shown excellent results. Image segmentation, involving thresholding, clustering, edge detection, and region extraction, is a crucial step in disease detection systems. The K-means clustering algorithm and SVM are widely used for segmentation and classification, respectively. For study[49] achieved 93% accuracy in detecting and classifying five plant diseases using a K-means and neural network-based model. Another study obtained a 90% overall accuracy in classifying tomato diseases using pre-processing, segmentation with k-means, feature extraction, and SVM classification. Further improvements can be made by increasing the number of training images.

The emergence of deep learning, a branch of machine learning, has brought significant advancements in the field of disease detection. These advancements can be attributed to improvements in computational power, storage capabilities, and the availability of large datasets. The ILSVR ImageNet competition in 2012 played a crucial role in popularizing and adopting deep learning techniques across various domains [8]. Convolutional Neural Networks (CNNs) have become widely used in tasks such as image classification, object detection, and semantic segmentation. However, it is important to highlight that the effectiveness of deep learning heavily depends on the availability of substantial amounts of data, necessitating large datasets with numerous images for successful model training and optimal performance [5]. Deep learning, specifically utilizing Convolutional Neural Networks (CNNs), has gained significant attention in the field of disease detection, particularly with the introduction of the PlantVillage dataset in 2015. This dataset has become a valuable resource for detecting diseases, estimating severity, and developing management systems. In addition to PlantVillage, several other publicly available plant disease datasets, such as Plant Pathology, Plant Pathology Challenge Dataset, Plant Disease Detection Dataset, PlantDoc dataset, Rice disease dataset, and Northern Leaf Blight (NLB) dataset, have facilitated the training of deep learning models for disease detection in various crops, addressing the challenges of yield reduction [31]. While these publicly available datasets are widely utilized, some studies also employ custom datasets tailored to their specific research objectives. As the field advances, researchers continue to explore and improve deep learning architectures, accompanied by the creation of new datasets to further enhance the capabilities of plant disease detection.

In this comprehensive review, we have evaluated several studies that encompassed the creation of plant disease datasets, the development of disease detection and severity estimation models and the exploration of researches. This review provides an analysis of the advancements in plant disease detection over time, including the progression of dataset development and the utilization of different data collection platforms. While previous reviews and surveys [6,7,8] have been conducted there has been a lack of studies in the field of plant disease detection that cover all aspects and identify open research gaps. This review explores the current techniques revolutionizing plant disease detection, enabling accurate and timely detection and classification of crop diseases. The analysis includes recent publications, focusing on disease detection methods, dataset characteristics, and advancements in object recognition in agriculture. While these techniques have improved crop management and yields, the study also discusses challenges and proposed solutions for customized control strategies and early disease detection. This study addressed the following research questions:

- Which datasets are currently available and can be effectively utilized for conducting research on plant diseases?
- What are the techniques that have been implemented for disease detection in the field of plant pathology?

By addressing these research questions, this study aims to provide a comprehensive understanding of the utilization of datasets, techniques of classification, data generality, disease severity estimation, and the potential of deep learning in comparison to human expertise. Furthermore, it highlights the existing research gaps and identifies promising areas for future investigations.

II. Utilizing Available Datasets for Plant Disease Research

In order to facilitate the development of models to accurate plant disease detection using image-based approaches, the availability of representative datasets is of utmost importance. These datasets should encompass a substantial number of images, typically in the thousands, to effectively support deep learning solutions. Among these datasets, the PlantVillage dataset stands out as the most commonly utilized choice by researchers. However, it is important to note that many researchers, as observed in the reviewed studies, opt to create their own datasets, which are not available to be accessible [9]. This trend signifies the growing preference for developing specialized datasets tailored to specific research requirements in the domain of plant disease detection.

1. plantVillage Datase

The PlantVillage dataset, created in 2015, is a widely used and influential collection of high-quality images representing various plant diseases. It consists of tens of thousands of images, including both healthy plants and plants affected by different diseases, covering a wide range of crops and regions. Exactly contain 54,309 images representing in 38 different class in 14 crops. The dataset has been extensively utilized for training and evaluating machine learning and deep learning models in the field of plant disease detection and management [10]. It has played a significant role in the development and assessment of algorithms for plant disease detection, classification, and severity estimation. However, it is important to note that the dataset may not fully represent real-field conditions and suffers from class imbalance and overlap between disease classes, which can impact the generalization and accuracy of deep learning models trained solely on this dataset. Example of plant village: Crop: Tomato, Disease: Late Blight, Description: The image displays a close-up view of a tomato leaf with distinct symptoms of Late Blight disease. The affected leaf shows irregularly shaped, dark lesions surrounded by a yellowish halo. The disease progression can be observed, with the lesions expanding and causing wilting and decay of the leaf tissue.

2. Plant Pathology 2020 Dataset

One widely used plant disease dataset is the "Plant Pathology 2020 Dataset". It focuses on plant disease classification and contains a variety of images showcasing different diseases affecting crops like apples, grapes, peaches, and more. The dataset encompasses diseases such as rust, scab, and various types of blight, with each image labeled accordingly. Researchers and data scientists can leverage this dataset to develop and evaluate machine learning models for plant disease detection and classification [11]. With its large size and diverse visual characteristics, including variations in disease severity and leaf appearance, the dataset allows for robust model training and simulates real-world scenarios. By utilizing the Plant Pathology 2020 dataset, researchers can contribute to the creation of accurate and efficient disease diagnosis systems, benefiting farmers, agronomists, and plant pathologists in early detection and effective management of crop diseases, ultimately leading to improved yields and reduced economic losses.

3. PlantDoc Dataset

PlantDoc is a dataset specifically designed for plant disease detection, consisting of 2,598 images that represent 17 different diseases across 13 diverse crops. There are Apple had 273 samples, Bell Pepper had 132 samples, Blueberry had 117 samples, Cherry had 57 samples, Corn had 376 samples, Grape had 133 samples, Peach had 122 samples, Potato had 222 samples, Raspberry had 119 samples, Soybean had 65 samples, Squash had 130 samples, Strawberry had 96 samples, and Tomato had 746 samples [12]. Each image is labeled with the corresponding disease type. The dataset aims to capture real-field conditions, providing a realistic representation of the challenges involved in identifying plant diseases. It covers a variety of crop species and showcases different disease symptoms. However, one limitation of this dataset is its class imbalance, with a little image in the classes, usually less than 200. Despite this limitation, researchers can still utilize the dataset to develop algorithms and models for accurate disease detection and classification, contributing to improved crop health and productivity.

4. Rice disease dataset

The rice disease dataset is a widely utilized and essential resource in the field of rice disease research. Specifically curated to facilitate the development of machine learning models for rice disease detection and classification. The dataset comprises 3,355 images that are classified into four distinct categories: 1488 healthy samples, 523 samples of Brown Spot, 565 samples of Hispa, and 779 samples of Leaf Blast [13]. With comprehensive labels for each image, researchers and practitioners can train and evaluate machine learning algorithms to automate the detection of rice diseases [14]. The dataset offers a wide array of visual characteristics, capturing different disease severities and leaf appearances, enabling the creation of robust models capable of handling real-world scenarios. Overall, the Rice Disease Image Dataset plays a vital role in advancing the field of rice pathology and contributing to effective disease management strategies.

5. Coffee disease dataset

In a separate study, a crop-specific annotated dataset was created specifically for coffee diseases, including severity information. The dataset consists of 1,560 images taken in a coffee field located in Ecuador. There are 791 healthy samples, 167 samples of Red Mite, 344 samples of Rust Level 1, 166 samples of Rust Level 2, 62 samples of Rust Level 3, and 30 samples of Rust Level 4. The field contained a total of 390 coffee plants [15]. To ensure accuracy, each image was meticulously annotated using an open-source tool, providing ground truth annotations for disease detection and severity estimation.

III. ADVANCEMENTS IN DISEASE DETECTION TECHNIQUES IN PLANT PATHOLOGY

1. Harnessing Deep Learning for Agriculture

Deep learning, a branch of machine learning and artificial intelligence, utilizes to process and analyze data in order to extract meaningful patterns and make predictions. It encompasses a range of techniques that depend on the availability of labeled data, including supervised learning, semi-supervised learning, and unsupervised learning. In plant disease detection, deep learning techniques have shown promise, with convolutional layers effectively capturing important features such as lesion colors and textures. Convolutional neural networks have demonstrated the ability to extract features and perform classification simultaneously, leading to increased adoption in plant disease detection as more datasets become publicly available [16].

Deep learning has emerged as a valuable tool in precision agriculture and phenotyping research. However, one of the challenges faced in utilizing deep learning for plant disease detection is the interpretability of the results. To address this, conducting a comprehensive review of deep learning-based studies in this field becomes crucial. Such a review helps in understanding the techniques employed, extracting meaningful insights, and evaluating different performance metrics [17]. A typical convolutional neural network (CNN) architecture consists of various layers, including input, hidden, convolutional, pooling, fully connected, and output layers. During training, a cost function is used to improve the network's performance. Activation functions and regularization techniques like dropout layers may also be applied to enhance model performance and generalization.

2. Plant Classification

Deep learning, especially in the agricultural field of plant disease detection, has revolutionized image classification. Through supervised training, deep learning utilizes labeled image datasets to classify objects in images. The SoftMax activation function assigns probabilities to each category, indicating the likelihood of the object belonging to a specific category. By optimizing the model's parameters using backpropagation and gradient descent, the network progressively improves its prediction accuracy. Supervised training with labeled datasets is a crucial aspect of deep learning, enabling the development of highly accurate classification models. Convolutional Neural Networks (CNNs) serve as the foundation for these models, leveraging their hierarchical architecture to effectively classify images in various computer vision tasks.

The (Table 1) illustrates key contributions made by various models in different years. The development of CNNs has seen several significant milestones. In 1998, the introduction of LeNet-5 demonstrated the effectiveness of convolutional layers in recognizing handwritten digits, laying the foundation for future advancements. In 2012, AlexNet achieved a breakthrough by winning the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), showcasing the power of deep architectures with multiple convolutional layers and the importance of large-scale labeled datasets for training CNNs [7].

Subsequent models further pushed the boundaries of accuracy and efficiency. VGGNet, proposed in 2014, demonstrated the benefits of deeper networks with stacked convolutional layers, although at the expense of increased computational complexity [18]. GoogLeNet (Inception) in 2015 introduced inception modules, reducing parameters and computational complexity while maintaining competitive performance. The ResNet model, also introduced in 2015, addressed the challenge of training extremely deep networks through the use of residual connections, leading to improved accuracy.

Other notable CNN models include DenseNet, MobileNet, EfficientNet, and Vision Transformer (ViT). DenseNet, developed in 2016, introduced dense connections between layers to facilitate better gradient flow, parameter efficiency, and feature reuse. MobileNet, introduced in 2017, focused on efficient models for mobile and embedded vision applications using depth-wise separable convolutions. EfficientNet, proposed in 2019, aimed to achieve state-of-the-art performance by scaling the depth, width, and resolution of the CNN architecture. Vision Transformer (ViT), introduced in 2020, applied the transformer architecture from natural language processing to computer vision tasks, achieving competitive performance.

Evaluation of image classification models involves assessing training, validation, and testing accuracies, with additional metrics like precision, recall, and F1-score for handling class imbalance [19]. High testing accuracies have been reported in various studies, although challenges remain in generalizing models across different datasets and image acquisition conditions. While image classification is widely used, it may not provide detailed lesion location information or enable the detection of multiple diseases simultaneously.

TABLE I. CNN MODELS EVOLVE

Year	Model	Key Contributions
1998	LeNet-5	Demonstrated effectiveness of convolutional layers
2012	AlexNet	Deep architecture, large-scale labeled datasets
2014	VGGNet	Deeper networks, fixed structure
2015	GoogLeNet	Inception modules, reduced parameters
2015	ResNet	Residual connections, training deep networks
2016	DenseNet	Dense connections, better gradient flow
2017	MobileNet	Efficient models for mobile and embedded vision
2019	EfficientNet	Optimal model scaling, compound scaling method
2020	Vision Transformer (ViT)	Transformer architecture for computer vision

3. Detecting of Diseases

Deep learning-based object detection is a powerful technique employed to locate and identify multiple objects within images and videos. This approach involves supervised training, where labeled images with boundary annotations are used to train deep neural networks. By learning from the labeled data, these networks can accurately detect and classify objects in new, unseen images. Various deep learning models have been developed for object detection, including SSD (Single Shot MultiBox Detector), YOLO (You Only Look Once), Faster R-CNN (Region-based Convolutional Neural Network), Fast R-CNN, and R-CNN (Region-based Convolutional Neural Network). These models differ in their architectures and methodologies, but they all aim to achieve efficient and accurate object detection in different applications.

In the context of plant disease detection, object detection has been applied to locate and identify plant diseases. Some studies focus on detecting entire diseased leaves, while others aim to detect specific disease pest. Among the reviewed studies, eight utilized object detection for plant disease detection.

Evaluation of object detection models for plant disease detection are evaluated using metrics such as average precision (AP), mean average precision (mAP), precision, recall, and F1-score [20]. AP and mAP measure the overall performance and accuracy of the model across different classes, while precision, recall, and F1-score provide insights into the model's ability to correctly detect and classify objects. These evaluation metrics play a vital role in comparing and selecting the most effective models for accurate plant disease detection.

Deep learning-based object detection techniques have shown promise in plant disease detection by enabling the detection of multiple disease instances simultaneously in images and videos, making them suitable for real-time systems. However, a limitation of these techniques is the localization of objects within bounding boxes. This can lead to the inclusion of healthy areas alongside disease lesions within the bounding boxes, introducing noise and potentially impacting the accuracy of disease detection [21]. The presence of healthy areas within the bounding boxes poses a challenge to the analysis and interpretation of the results. Therefore, further research is necessary to address this challenge and enhance the accuracy of object detection in plant disease detection tasks.

The performance of the models can be influenced by various factors, including training data, model architecture, and the nature of the plant disease. While deep learning has shown impressive results, human expertise remains crucial in plant disease detection. Human experts consider factors beyond visual symptoms, like environmental conditions and historical context. Combining deep learning models with human expertise leads to accurate and reliable results. While surpassing human accuracy (approximately 94%) is a goal for deep learning, the combination of deep learning and human involvement is often the most effective approach for plant disease detection [32].

Deep learning-based image classification has demonstrated impressive accuracies in identifying plant diseases within specific datasets. In the comprehensive PlantVillage dataset, accuracies ranging from 91.4% to 99.81% were achieved when both training and testing images were sourced from the same dataset [33,10 ,34].

Several studies have concentrated on the detection of diseases in specific crops using image classification techniques, demonstrating impressive accuracies. For instance, accuracies as high as 98.8% were achieved for identifying diseases in apples, cherries, and corn [36]. In another study, diseases in corn, grapes, tomatoes, apples, and sugarcane were identified with an accuracy of 96.5% [35]. Additionally, custom acquired diseases for peaches, apples, pears, and grapevines were successfully identified with an accuracy of 96.3% [37]. These results emphasize the preference of deep learning-based image classification in accurately detecting plant diseases in specific crops.

Several studies have specifically targeted the detection of diseases in individual crops using image classification techniques. For instance, in the case of tomato diseases from the PlantVillage dataset, accuracies ranging from 91.15% to 99.65% were achieved [10, 38, 39]. Custom datasets were also utilized, resulting in accuracies of 95.75% and 96.25% for tomato disease detection [40, 41]. Similarly, image classification techniques were employed to identify corn diseases, leading to an accuracy of 98.9% [42]. Furthermore, accuracies of 87% and 96.96% were attained for corn disease detection using the Digipathos dataset for training models [43, 44]. These results demonstrate the effectiveness of image classification in accurately identifying diseases in specific crops, particularly tomatoes and corn.

In the field of bean diseases, including angular leaf spot, bean rust, and healthy classes, deep learning techniques were employed to successfully identify these diseases. With a balanced dataset captured under field conditions, testing accuracies of up to 95.31% were achieved [20]. Similarly, for cassava diseases, a custom acquired dataset collected under field conditions led to an accuracy of 96.75% in disease detection [45]. These studies highlight the effectiveness of image classification and deep learning approaches in accurately detecting and classifying diseases within specific crops, such as beans and cassava.

Classification models has generally achieved high accuracies exceeding 94% for plant disease detection. However, there are cases where lower accuracies have been reported, particularly when the task involves identifying diseases across multiple crops. Traditional machine learning solutions could achieve accuracies higher than 94% for disease detection, but their application is limited by the need for manual feature extraction and definition [47,48]. Deep learning, on the other hand, enables automatic feature extraction through convolutional neural networks (CNNs), enhancing generalization capabilities. While deep learning solutions for disease detection may initially appear superior to human-based detection, a significant limitation lies in the generalization capability of models. For instance, in one study, diseases were initially detecting with an accuracy of 99.01% when testing and training images were from the same dataset. despite of, when testing images from a different dataset were introduced, the accuracy decreased to 45.95% [46]. This highlights the challenge of generalizing deep learning models to new datasets and emphasizes the need for robustness in real-world applications.

TABLE II. A COMPARATIVE ANALYSIS FOR CROP DISEASE DETECTION ACCURACY AND DATASET LIMITATIONS ACROSS VARIOUS CROPS.

References	Dataset	Result	Limitations
[33, 10, 34,36]	PlantVillage dataset	Accuracies ranging from 91.4% to 99.81%	Limited to the same dataset for train and test and some model more Complex deep learning architectures and the challenging to interpret and explain model decisions, which crucial in certain contexts.
[37]	Dataset of 13 crops types	Accuracy of 96.3%	Models struggle with real-world challenges such as varying lighting and background interference, affecting their performance in practical agricultural settings.
[10, 38, 39]	Tomato diseases in PlantVillage dataset	Accuracies ranging from 91.15% to 99.65%	Only use the single plant tomato diseases and that the same dataset for train and test
[40, 41]	Tomato (custom dataset)	Accuracies of 95.75% and 96.25%	Models not perform well on new datasets due to their reliance on specific datasets for training and testing.
[42]	Corn diseases	Accuracy of 98.9%	Datasets used not fully capture the range of variations and disease manifestations seen in real-world scenarios, potentially leading to biased or incomplete training.
[43, 44]	Corn diseases	Accuracies of 87% and 96.96%	Specific to corn diseases
[20]	Bean diseases	Testing accuracy of up to 95.31%	The scarcity of comprehensive and diverse datasets training and evaluation, leading to suboptimal performance and limited generalization
[45]	Cassava diseases	Accuracy of 96.75%	Specific to cassava diseases
[47, 48]	VillagePlant	Accuracies exceeding 94%	Manual feature extraction techniques can be time-consuming and subjective, limiting the models' ability to capture complex disease patterns effectively.

4. Planet disease estimation

The measurement of plant disease severity plays a crucial role in assessing disease spread and its impact on plant growth and yield. Traditionally, this has relied on costly and subjective visual analysis by experts. However, deep learning-based systems offer an automated and improved solution for disease severity estimation from images. By leveraging deep learning algorithms, these systems can overcome the limitations of manual analysis and deliver accurate results [22].

Researchers have adopted various definitions and metrics for estimating disease severity. One common approach involves quantifying severity as the proportion of leaf area affected by disease lesions or symptoms [23]. Deep learning-based image classification techniques have been developed to estimate severity based on the extent of leaf area impacted by these lesions. Numerous studies have utilized deep learning and image classification to estimate the severity of different plant diseases [24]. These approaches aim to provide precise quantification of disease severity and valuable insights for effective disease management in agriculture.

In the past, traditional image processing techniques were commonly used for pomegranate disease severity detection, while fuzzy logic was utilized for estimating corn disease severity [25]. However, the emergence of deep learning-based models has brought about a significant shift in the research community. These models have gained acceptance due to their capability to automatically learn essential features for disease severity estimation. The use of deep learning has providing more accurate and efficient methods for assessing the severity of plant diseases.

In a study related to coffee diseases, a deep learning model was trained to assess severity levels as healthy, very low, low, high, and very high. The VGG16 model achieved an accuracy of 86.51% in estimating severity based on the percentage of lesion coverage on coffee leaves [26]. In another specific study focused on black rot disease in apples, a deep learning model was trained to identify four severity levels: healthy, early-stage, middle-stage, and end-stage. By analyzing spot lesions in the images, the VGG16 model achieved an accuracy of 90.4% in severity classification [22]. Another model called PD2SE was then developed, achieving up to 91% accuracy in disease severity estimation using the PlantVillage dataset [27].

This study investigates the application of deep convolutional neural networks, namely VGGNet and Inception module, for the detection of plant leaf diseases. Transfer learning is utilized by initializing the weights with pre-trained models from ImageNet. The approach achieves a validation accuracy of at least 91.83% and an average accuracy of 92.00% for rice plant images [30].

In the case of late blight disease in tomatoes, deep learning models based on squeeze and excitation networks, CapsNet, AlexNet, and ResNet were employed. These models successfully identified four severity classes: healthy, early, middle, and end, achieving testing accuracies of up to 93.75% [28]. For multi-label severity estimation, a binary relevance CNN (BR-CNN) model was utilized. It accurately identified crops, disease types, and severity levels. In disease detection, DenseNet121 achieved a 98.45% accuracy, while ResNet50 achieved a 92.93% accuracy in severity estimation. However, the severity was classified into three classes: general, normal, and serious [29].

Standardizing severity estimation methods would facilitate comparative studies, enhance understanding of disease progression, and support the development of effective management strategies. By establishing uniform criteria, researchers can improve the reliability and reproducibility of severity estimation across different plant diseases and crops.

IV. DISCUSSION

Plant disease diagnosis and monitoring are critical tasks in plant pathology, and recent studies have employed image processing, machine learning, and deep learning techniques, utilizing datasets such as PlantVillage, to identify and classify plant diseases. Although advancements have been made in algorithm development and disease detection, there are still limitations and gaps that need to be addressed for the development of comprehensive plant disease management systems. These include challenges in generalization to new datasets, interpretability of deep learning models, and the need for robust and scalable solutions that can be applied in real-world settings. Further research is required to bridge these gaps and create effective and practical solutions for plant disease diagnosis and monitoring.

One limitation is the focus on individual crop diseases or limited disease coverage within studies. While these approaches can accurately detect specific targeted diseases, they may struggle to identify diseases outside the training set, posing a challenge for comprehensive agriculture disease detection. Standardizing severity estimation definitions is another challenge. Deep learning models require consistent and well-defined criteria for accurate severity estimation. Image classification alone is insufficient, as it cannot precisely localize disease lesions. Object detection improves localization but may include non-infected areas. Segmentation approaches can provide more accurate calculations of infected leaf regions. Considering the season of disease detection is crucial, enabling timely and effective management techniques for maximum crop yield. Incorporating this aspect into plant disease management systems ensures efficient disease control and prevention strategies. Addressing these limitations and gaps in research will be instrumental in developing comprehensive end-to-end plant disease management systems. By incorporating multiple disease types, standardized severity definitions, and considering disease detection timing, future solutions can enhance food production and improve global food security.

V. CONCLUSION

This comprehensive study examined several research papers covering a wide range of topics related to plant disease diagnosis, including review papers, computer vision and processing, machine learning and deep learning studies. The study addressed two critical research questions and explored an available datasets and descriptions of diseases that affecting plants as in section 2. It also discussed the evolves of deep learning models and their potential to surpass human capabilities as in section 3. The study concluded that effective plant disease analysis should involve the detection of multiple plants and their respective diseases at an early stage, as well as accurate estimation of disease severity. These findings are crucial for the development of an automated end-to-end plant disease detection system.

VI. References

[1] E.-C. Oerke, H.-W. Dehne Safeguarding production—losses in major crops and the role of crop protection Crop Prot., 23 (2004), pp. 275-285.

[2] M. Loey, A. ElSawy, M. Afify Deep learning in plant diseases detection for agricultural crops: a surveyInt. J. Serv. Sci. Manag. Eng. Technol. (IJSSMET), 11 (2020), pp. 41-58.

[3] 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), pp. 280-290.

[4] F. Martinelli, R. Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, Boschetti, L.R. Goulart, et al. Advanced methods of plant disease detection. A review Agron. Sustain. Dev., 35 (2015), pp. 1-25.

[5] Marcus, G. (2018). Deep learning: A critical appraisal. *arXiv preprint* arXiv:1801.00631.

[6] J.G.A. Barbedo A review on the main challenges in automatic plant disease identification based on visible range images Biosystems Eng., 144 (2016), pp. 52-60.

[7] deep learning in plant diseases detection for agricultural crops: a survey Int. J. Serv. Sci. Manag. Eng. Technol. (IJSSMET), 11 (2020), pp. 41-58.

[8] A.K. Singh, B. Ganapathysubramanian, S. Sarkar, A. Singh Deep learning for plant stress phenotyping: trends and future perspectives Trends Plant Sci., 23 (2018), pp. 883-898.

[9] Ahmad, A., Saraswat, D., & Gamal, A. E. (2022). A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agricultural Technology*, *3*, 100083. j. atech.2022.100083.

[10] K.P. Ferentinos Deep learning models for plant disease detection and diagnosis Comput. Electron. Agric., 145 (2018), pp. 311-318.

[11] Thapa, R., Zhang, K., Snavely, N., Belongie, S., & Khan, A. (2020). The Plant Pathology Challenge 2020 data set to classify foliar disease of apples. Applications in Plant Sciences, 8(9). 10.1002/aps3.11390.

[12] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, N. Batra Plantdoc: a dataset for visual plant disease detection Proceedings of the 7th ACM IKDD CoDS and 25th COMAD (2020), pp. 249-253.

[13] Rallapalli, S., & Durai, M. S. (2021). A contemporary approach for disease identification in rice leaf. International Journal of Systems Assurance Engineering and Management. s13198-021-01159-y.

[14] H.B. Prajapati, J.P. Shah, V.K. Dabhi Detection and classification of rice plant diseases Intell. Decis. Technol., 11 (2017), pp. 357-373.

[15] 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), Article 104414.

[16] 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), p. 77.

[17] M. P´erez-Enciso, L.M. Zingaretti A guide on deep learning for complex trait genomic prediction Genes, 10 (2019), p. 553.

[18] 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.

[19] 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.

[20] Chandra, M., Patil, P.S., Roy, S., & Redkar, S.S. (2020). Classification of various859plant diseases using deep siamese network.

[21] 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), Article e0168274.

[22] G. Wang, Y. Sun, J. Wang Automatic image-based plant disease severity estimation using deep learning Comput. Intell. Neurosci. (2017), p. 2017.

[23] G. Fenu, F.M. Malloci Forecasting plant and crop disease: an explorative study on current algorithms Big Data Cogn. Comput., 5 (2021), p. 2.

[24] M. Bhange, H. Hingoliwala Smart farming: pomegranate disease detection using image processing Procedia Comput. Sci., 58 (2015), pp. 280-288.

[25] 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), pp. 205-219.

[26] 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), Article 105162.

[27] 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), pp. 518-529.

[28] 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), p. 49.

[29] 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), pp. 15327-15340.

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

[31] 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), pp. 55-57.

[32] J. Hearty Advanced Machine Learning with Python Packt Publishing Ltd (2016).

[33] 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), Article 105542.

[34] 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), p. 1319.

[35] 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), pp. 579-582, 10.1109/ECICE47484.2019.8942686.

[36] 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.

[37] 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), p. 2016.

[38] 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).

[39] S. Zhang, W. Huang, C. Zhang Three-channel convolutional neural networks for vegetable leaf disease recognition Cogn. Syst. Res., 53 (2019), pp. 31-41.

[40] 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.

[41] 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), pp. 1-14.

[42] 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), pp. 30370-30377.

[43] J.G. Barbedo Factors influencing the use of deep learning for plant disease recognition Biosystems Eng., 172 (2018), pp. 84-91.

[44] A. Karlekar, A. Seal Soynet: soybean leaf diseases classification Comput. Electron. Agric., 172 (2020), Article 105342.

[45] 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), p. e352.

[46] 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), Article 105220.

[47] 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.

[48] 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.

[49] Bashish, D.A., Braik, M., Bani-Ahmad, S.: Detection and classification of leaf diseases using K-means-based segmentation and neural-networks-based classification. Inf. Technol. J. **10**, 267–275 (2011).