Exploring the Deep Learning Techniques in Plant Disease Detection: A Review of Recent Advances

Saurabh Singh and Rahul Katarya

Abstract In agriculture, protecting crop yield is one of the most critical aspects of avoiding crop waste and ensuring food security around the world. One of the most critical aspects of preserving yield is protecting it from pests and plant diseases. With the advancement in the field of Artificial Intelligence (AI), it has been applied to different domains, and one such field is agriculture, where we can incorporate AI. Deep learning (DL), which is a subset of Artificial Intelligence, has gained lots of attention toward plant disease detection in the present day because of its better accuracy and performance in comparison with other techniques like machine learning (ML), etc. In this paper, we provide a comprehensive review of the current research work by utilizing deep learning for plant disease detection. We study the different models and architectures proposed by different authors and try to identify the pros and cons of the proposed methodology. We also discuss the various datasets that have been used in research work for detecting plant diseases. Finally, we describe the possible challenges in implementing deep learning models and discuss the future roadmap that can be followed by trying to identify the research gaps.

Keywords Agriculture · Plant diseases · Artificial Intelligence (AI) · Deep learning (DL) · Machine learning · Plant disease detection

1 Introduction

The agricultural sector holds significant importance in the economy, serving as the primary means of sustenance for a significant portion of the population and playing a vital role in livelihoods worldwide. Here's a brief overview of agriculture in India and the world.

S. Singh · R. Katarya (✉)
Delhi Technological University, New Delhi, India
e-mail: rahuldtu@gmail.com

S. Singh
e-mail: saurabh_2k22afi20@dtu.ac.in

S. Das et al. (eds.), *Advances in Data-Driven Computing and Intelligent Systems*,
Lecture Notes in Networks and Systems 892,
https://doi.org/10.1007/978-981-99-9521-9_21

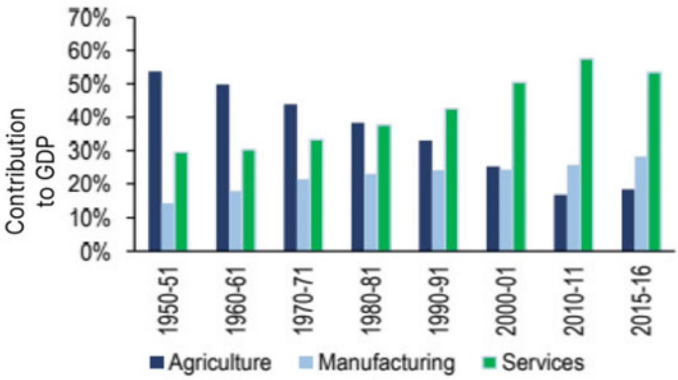

Fig. 1 Sector-wise contribution to GDP of India [2]

Agriculture in the World: Over 26% of the world's workforce is employed in the agriculture sector, which is an important one. The agricultural sector only made up 4.00% of the global GDP, while in low-income nations it accounts for an average of 30.00% of the GDP. International trade in agriculture contributes to meeting the various demands of nations and the availability of food [1].

Agriculture in India: In India, agriculture employs about 50% of the population and generates close to 17.5% of the GDP. Figure 1 shows the contribution of different sectors to the GDP of India over the years. It plays a crucial role in the country's development, alleviation of poverty, and food security. India is a major producer of several goods, including rice, wheat, pulses, oil seeds, fruits, and vegetables [2].

We can see agriculture holds great significance on a global scale and any disruption in agriculture can have significant impacts on various aspects, including food security, economy, environment, and social well-being. Agriculture faces various challenges, including fragmented landholdings, dependence on monsoons, crop damage due to plant diseases, water scarcity, inadequate infrastructure, post-harvest losses, farmer indebtedness, and market volatility. One such biggest challenge faced in agriculture is plant diseases. Plant diseases can lead to significant economic losses by reducing crop yields, quality, and market value. They can affect both food crops and cash crops, impacting farmers' livelihoods and global food supplies.

The five main crops grown across the world include wheat, rice, maize, soybeans, and potatoes—contributing roughly 18.3, 18.9, 5.4, 3.3, and 2.2 percent of the calories consumed worldwide, respectively. Each of these crops has an estimated 21.5, 30.0, 22.6, 21.4%, and 17.2% loss worldwide as a result of illnesses and pests infecting these plants [3]. Figure 2 shows the percentage of crop loss caused by different plant pathogens.

At the current time deep learning [5] has spread across different domains and areas with advancements in hardware and technologies and the same can be employed in the field of agriculture. In this paper, we try to review different plant disease detection

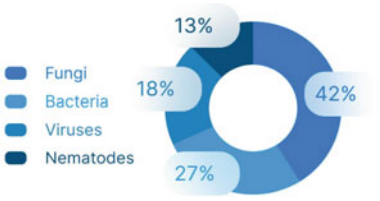

Fig. 2 Crop loss by plants pathogens [4]

techniques based on deep learning. This study's key contribution can be summed up as follows:

- To provide a comprehensive review of the current research work by utilizing deep learning for plant disease detection.
- Study and discuss the pros and cons of different models and architectures proposed for plant disease detection.
- Identify the challenges and limitations for future research direction.

The paper is organized with Sect. 2 giving background about plant diseases and deep learning and its components. Section 3 provides information on different plant disease datasets available and the latest research work done utilizing deep learning for plant disease detection. Section 4 describes the various limitations and research gaps that need to be addressed in the future. Section 5 gives the conclusion of this study work.

2 Background

2.1 Plant Diseases

Plant diseases refer to the abnormal conditions or disorders that affect plants, leading to a decline in their health, growth, and productivity. These diseases can be caused by various factors, including pathogens, parasitic plants, and abiotic factors (such as nutrient deficiencies, extreme temperatures, and pollution). Here, we discuss different types of plant diseases [6]:

- **Fungal Diseases**: Fungi can impact various parts of plants including leaves, stems, fruits, and roots. Some common types of plant fungal diseases include powdery mildew, rust, downy mildew, smuts, leaf spots, and root rots.
- **Bacterial Diseases**: Bacteria can infect plants and cause various diseases like bacterial blight, bacterial canker, crown gall, and bacterial wilt. Wilting, leaf spots, cankers, or galls are some examples of the symptoms of bacterial infections.

- **Viral Diseases**: Viruses, such as mosaic viruses, leaf curl viruses, and necrotic ring spot viruses, are contagious agents that can harm plants. Symptoms of viral infections frequently include stunted or deformed development, yellowing, mosaic patterns, and leaf mottling.
- **Nematode Diseases**: Microscopic worms that can parasitize plant roots are the cause of nematode infections, which include cyst nematodes and root-knot nematodes. Plants that are impacted could exhibit signs of growth stunting, root galling, or nutritional shortages.
- **Parasitic Plant Diseases**: Diseases caused by parasitic plants include dodder and witchweed, which physically attach to their host plants and siphon off nutrients and water.
- **Abiotic Diseases**: Various abiotic factors like extreme temperatures, nutrient imbalances or deficiencies, water stress, chemical toxicity, or air pollution can cause plant illnesses. These illnesses can cause tissue death (necrosis), chlorosis (yellowing), or an overall decline in plant health.

It's important to note that the specific types and names of plant diseases can vary across different plant species, regions, and environmental conditions. Figure 3 shows sample images of some of the plant diseases. Proper diagnosis and identification of plant diseases are crucial for implementing effective disease management strategies.

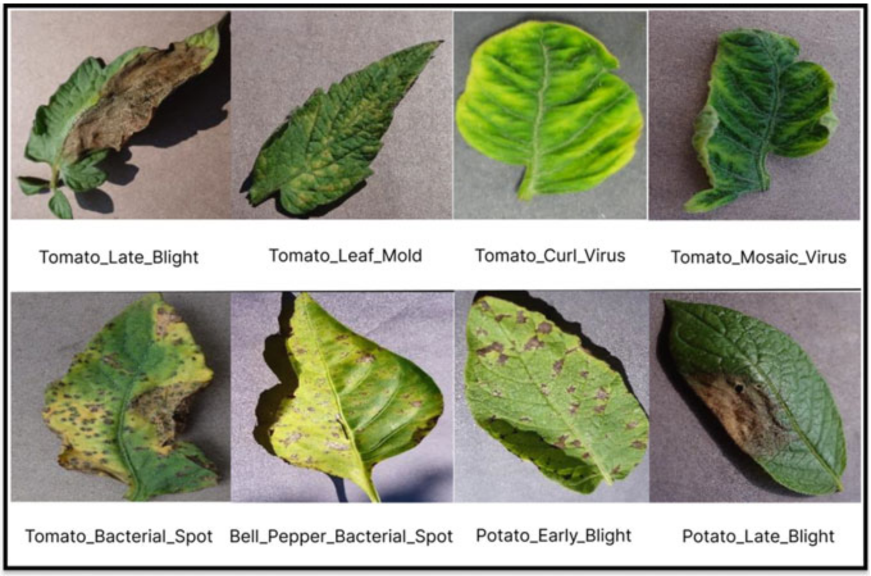

Fig. 3 Sample images of plant diseases

2.2 Deep Learning

Deep learning is a subfield of machine learning that focuses on building artificial neural networks that can learn from complicated data and make predictions based on that data. These networks are modeled after the structure and operation of the neural network in the human brain. The key components of a deep learning architecture are as follows:

- **Input Layer**: It is the first layer in the model that receives the input data or features that are fed into the deep learning model. A feature or property of the input data is represented by each neuron in the input layer. The input layer is connected to hidden layers.
- **Hidden Layers**: Deep learning models have one or more hidden layers associated with them. The hidden layer is composed of multiple neurons that are connected to one another. The weighted sum of the inputs received from the last layer is performed for each neuron in the hidden layer, which is followed by applying the activation function to finally produce an output. The hidden layer helps the network learn complex information and extract hierarchical features from the input data.
- **Activation Functions**: Activation functions are used to add nonlinearity to the network. Some of the commonly used activation functions are sigmoid, hyperbolic tanh, Rectified Linear Unit, etc. This nonlinearity helps the network learn the complex patterns and relationships from the data.
- **Weight Parameters**: The connections between neurons in different layers are associated with weight parameters. These weights are adjusted to minimize the loss function during the training of the model, using techniques like gradient descent and backpropagation.
- **Output Layer**: It is the final layer of the model which produces the final output or predictions based upon learned knowledge from the previous layers. The activation function and neurons in the final layer depend upon the nature of the task being performed.
- **Loss Function**: The difference between the actual output and the expected output is measured by the loss function. During model training, the weights are adjusted using the gradients of the loss function to enhance the model's ability to predict outcomes.
- **Optimization Algorithm**: Optimization algorithms are used to adjust the weights and bias associated with the neurons using the gradients of the loss function. Some commonly used optimization algorithms are stochastic gradient descent, Adam, AdaGrad, etc. These algorithms decide how much weight should be updated and in what direction to iteratively boost the model's performance.

Deep learning architectures can vary depending on the specific task or application. Well-known architectures encompass convolutional neural networks (CNNs) [7] that excel in image and video processing, recurrent neural networks (RNNs) [7] which are ideal for handling sequential data, and transformer models that are widely used in

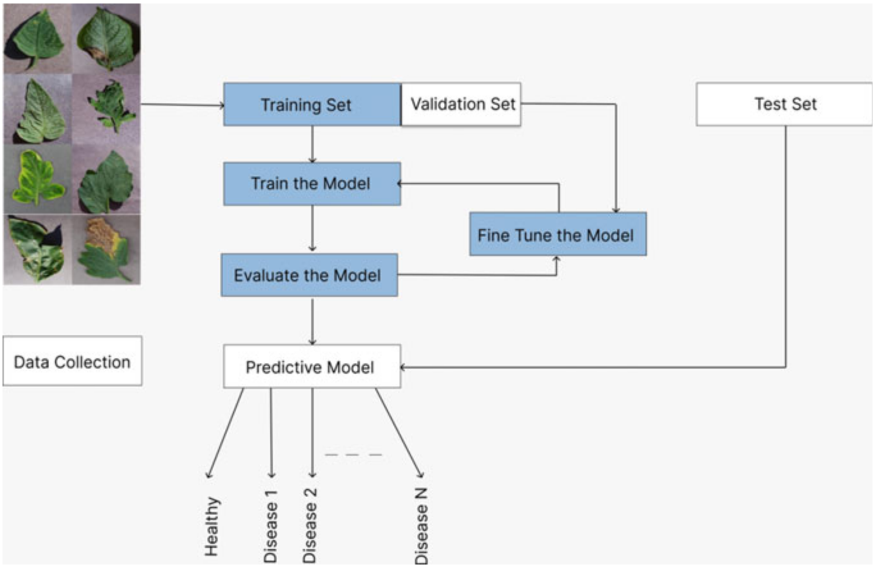

Fig. 4 General deep learning framework for plant disease detection

natural language processing tasks. These architectures incorporate specialized layers and components to handle the unique characteristics and challenges of different data types and tasks. Figure 4 shows a general deep learning framework for plant disease detection.

3 Related Work

3.1 *Literature Survey*

In recent years, extensive work has been done using deep learning toward plant disease detection. This literature review seeks to give an overview of the major advancements and developments in deep learning-based plant disease detection. The literature survey highlights the rapid advancements and significant contributions of deep learning techniques in plant disease detection. Deep learning-based models tend to show better accuracy and performance in comparison with other methods provided quality data is available to train and evaluate the model. Some of the recent works are discussed below:

Gehlot et al. [8] proposed an architecture named EffiNet-TS, which consists of two classifiers and one decoder. The proposed model is based on state-of-the-art Teacher/Student architecture built around EfficientNetV2. The suggested model highlights the important feature for classifying plant disease, which improves classification and offers a clearer visual representation of specific plant disease symptoms. Many authors have used state-of-the-art object detection algorithms with some internal modifications to improve the performance. Li et al. [9] proposed an integrated model that combines single-stage and two-stage target detection networks. The single-stage network is built upon the YOLO with internal structure optimization. The two-stage network, on the other hand, is based on the Faster R-CNN (Region Convolutional Neural Network). Initially, the target frames are clustered using clustering techniques followed by the integration of two models to perform the disease detection task. Another author Mahum et al. [10] proposed a model based on DenseNet-201 for classifying potato leaves into five different categories.

Some of the authors have proposed a novel deep learning model for plant disease detection like Yu et al. [11] and Ramamoorthy et al. [12] both the authors proposed a novel deep learning-based model with the former proposing a model based on inception convolution and vision transformer and the latter based on MobileNet V1 architecture. Wang et al. [13] proposed a lightweight model which is based on state-of-the-start YOLOV5 architecture for plant disease detection which has better accuracy and performance in comparison with other state-of-the-art techniques. Another author Elaraby et al. [14] proposed a model based on AlexNet for classifying the diseases in plants and use of Particle Swarm Optimization (PSO) for feature selection which helped in optimizing the performance of the overall model.

Saleem et al. [15] proposed a model named region-based fully convolution network (RFCN). The author also studied the use of different data augmentation techniques and the effect of hyperparameter tuning on the performance of the model. Shah et al. [16] proposed a model named ResTS (Residual Teacher/ Student) which is based on CNN architecture. It consists of two classifiers and a decoder. The proposed model is capable of finer visualization for disease detection. Panchal et al. [17] performed a comparative study on four models, namely Inception-v3, ResNet50, VGG16, and VGG19. The author also performed parameter tuning on these models to obtain better results and gave insights for each of these models.

The studies discussed demonstrate the effectiveness of deep learning models, such as CNNs, in accurately identifying and classifying plant diseases across various crops and imaging modalities. In Table 1, we highlight the key findings of our literature review and perform a comparative study of the different models proposed for plant disease detection. The various deep learning approaches that have been studied can prove to be highly beneficial to the farmers in monitoring and identifying the different types of diseases and taking appropriate action.

3.2 Datasets

Datasets play a crucial role in deep learning-based models, and they tend to show better accuracy and performance in comparison with other methods provided quality data is available to train and evaluate the model. Some of the commonly used datasets are discussed in brief and mentioned in Table 2.

Table 1 Literature survey of recent deep learning-based plant disease detection

Year, References	Model used	Performance	Dataset	Pros	Cons
2023, [8]	EfficientNetV2	F1 score: 0.989, Accuracy: 0.990	PlantVillage dataset	Better performance in comparison with the state-of-the-art architecture of ResTS	The dataset does not represent real-world crop scenarios Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.
2023, [9]	YOLO + RCNN	Accuracy 85.2%	A self-made dataset consisting of 7199 images belonging to 6 species namely peach, pepper, potato, squash, tomato, and strawberry	Computationally fast and efficient	The model accuracy is improved but at the cost of operation speed The quality and quantity of the dataset is questionable
2023, [10]	Efficient DenseNet	Accuracy 97.2%	PlantVillage dataset + manually gathered data	Computationally fast and efficient	Focuses only on potato leaf disease detection The quality of the dataset is questionable
2023, [11]	CNN	Accuracy (PlantVillage): 99.94, Accuracy (ibean): 99.22, Accuracy(AI2018):86.89, Accuracy (PlantDoc):77.54	PlantVillage dataset + ibean leaf image dataset + AI2018 + PlantDoc dataset	Better performance in comparison with the state-of-the-art models	Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.

(continued)

Table 1 (continued)

Year, References	Model used	Performance	Dataset	Pros	Cons
2023, [12]	CNN	Accuracy of 95%	PlantVillage dataset	Simple and efficient	Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.
2022, [13]	Optimized YOLOv5	Precision 93.73, Recall 92.94, Accuracy (PlantDoc): 90.26% Accuracy (Peanut Rust): 92.57%	PlantVillage dataset, PlantDoc, and self-made Peanut Rust dataset	The proposed model, i.e., optimized YOLOv5 is efficient in terms of memory and operation time in comparison with other networks	The dataset does not represent real-world crop scenarios Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.
2022, [14]	AlexNet + PSO	Accuracy 98.83	The dataset consists of nearly 13,000 images of 5 crop species	Significant improvement in accuracy of AlexNet from 95.6 to 98.83 with PSO	Room for improvement in the quality of the dataset for plant disease detection
2022, [15]	Region-based fully convolutional network (RFCN)	mAP 93.8%	NZDLPlantDisease-v1	Focuses on other organs of plants like fruits, stems, etc., instead of just leaves	Disease detection for multiple organs is not considered in all plant species The dataset is limited to fields from New Zealand

(continued)

Table 1 (continued)

Year, References	Model used	Performance	Dataset	Pros	Cons
2021, [16]	CNN	F1 score of 0.991	PlantVillage dataset	Better performance in comparison with the state-of-the-art Teacher/Student architecture	Requires more memory and operation time in comparison with benchmark Teacher/Student architecture Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.
2021, [17]	CNN	Accuracy of 93.5%	PlantVillage dataset	Simple and efficient model	More training data could have helped further improve the accuracy of the model Focuses only on the leaf organ of plants. Diseases are associated with other organs like fruits, stems, etc.

Table 2 Datasets used in plant disease detection

Dataset name	No. of images	Year released
Plant pathology 2021—FGVC8 [18]	23,249	2021
DiaMOS plant dataset [19]	3505	2021
PlantDoc [20]	2598	2019
RoCoLe dataset [21]	1560	2019
BRACOL dataset [22]	1747	2019
PlantVillage dataset [23]	54,303	2015

The Plant Pathology 2021—FGVC8 [18] is a dataset for apple foliar disease detection in apple plants released in the year 2021. Another dataset released in the year 2021 is the DiaMOS [19] plant dataset belonging to pear fruit which consists of 3505 images of leaves belonging to four different classes of disease.

The PlantDoc [20] dataset consists of 2598 images of 13 plant species belonging to 17 different classes of disease released in the year 2019. RoCoLe [21] dataset was also released in the year 2019 focusing on the Robusta Coffee Leaf image divided into 6 classes. BRACOL [22] dataset was also released in the year 2019 consisting of 1747 images focusing on Arabica coffee leaves affected by different diseases belonging to five different classes.

PlantVillage [23] is one of the most extensively used datasets in different studies. The PlantVillage dataset consists of high-quality images of diseased and healthy plant leaves captured under controlled conditions. It consists of a total of 54,303 images spanning across 14 crop species. The dataset contains images depicting a wide range of plant diseases caused by fungi, bacteria, viruses, and other pathogens.

4 Research Gap

From the above literature review in Table 1, we identify various limitations that can be addressed in the future. Following are the research gaps we could identify from this study:

- **Generalization**: Most of the research work has been performed using the PlantVillage dataset but it may not represent a practical real-world scenario as it was developed under a controlled environment. The generality of the algorithms is impacted by this issue, making them unsuitable for real-world deployment.
- **Limited scope**: Most of the study focus only on plant leaf disease detection but diseases may also be associated with other organs of plant like stem, fruits, etc. Hence there is a need for comprehensive plant disease detection beyond the leaf organ of plants.

- **Recognition efficiency at the cost of inference efficiency**: The vast majority of current studies concentrate on recognition efficiency while ignoring inference efficiency, which restricts their practical real-world application.
- **Small datasets**: Dataset is very crucial for deep learning-based models for better performance and accuracy. There are limited and small datasets for plant disease detection which needs to be addressed.
- **Multiple disease detection**: Furthermore, it hasn't been done to simultaneously detect several illnesses in a single plant organ. Also, the effectiveness of the same optimized/modified model has not been studied in complicated horticulture settings consisting of different crops.

5 Conclusion

This paper aims to provide a comprehensive review of the latest research work by utilizing deep learning for plant disease detection. Large datasets and advances in deep learning architectures have allowed researchers to detect plant diseases with high accuracy rates, outperforming earlier techniques and enabling quick and accurate diagnosis. Deep learning is advantageous for detecting plant diseases because it has the potential for real-time monitoring and is scalable for use in large-scale agricultural applications. Deep learning algorithms can be used to develop intelligent systems that can identify diseases early on, assisting with timely disease management decisions. The various deep learning approaches that have been studied can prove to be highly beneficial to the farmers in monitoring and identifying the different types of diseases and taking appropriate action, which would ultimately prevent crop wastage and financial loss to the farmer. After a thorough analysis and study, we could identify some limitations and research gaps that could be addressed in the future work. Future research directions might also involve combining deep learning with cutting-edge technologies like drones and the Internet of Things (IoT).

In conclusion, deep learning-based plant disease detection has enormous potential to transform crop protection and disease management techniques. Deep learning techniques will continue to evolve, along with the incorporation of complementing technology, opening the door for more precise, effective, and sustainable agricultural practices that will ultimately improve crop health and contribute to global food security.

References

1. Anik R, Asif, SR, Sarker JR (2020) Five decades of productivity and efficiency changes in world agriculture (1969–2013). Agriculture 10(6):200
2. Deshpande T (2017) State of agriculture in India. PRS Legislative Res 53(8):6–7
3. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A (2019) The global burden of pathogens and pests on major food crops. Nat Ecol Evolut 3(3):430–439

4. Khan MR, Sharma RK (2020) Fusarium-nematode wilt disease complexes, etiology and mechanism of development. Ind Phytopathol 73(4):615–628
5. Hao X, Zhang G, Ma S (2016) Deep learning. Int J Semant Comput 10(03):417–439
6. Singh RS (2018) Plant diseases. Oxford and IBH Publishing
7. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8:1–74
8. Gehlot M, Gandhi GC (2023) EffiNet-TS: a deep interpretable architecture using EfficientNet for plant disease detection and visualization. J Plant Diseases Protect 130(2):413–430
9. Li M, Cheng S, Cui J, Li C, Li Z, Zhou C, Lv C (2023) High-performance plant pest and disease detection based on model ensemble with inception module and cluster algorithm. Plants 12(1):200
10. Mahum R, Munir H, Mughal Z-U-N, Awais M, Khan FS, Saqlain M, Mahamad S, Tlili I (2023) A novel framework for potato leaf disease detection using an efficient deep learning model. Human and Ecological Risk Assessment: Int J 29(2):303–326
11. Yu S, Xie Li, Huang Q (2023) Inception convolutional vision transformers for plant disease identification. Internet of Things 21:100650
12. Ramamoorthy R, Saravana Kumar E, Naidu RCA, Shruthi K (2023) Reliable and accurate plant leaf disease detection with treatment suggestions using enhanced deep learning techniques. SN Comput Sci 4(2):158
13. Wang H, Shang S, Wang D, He X, Feng K, Zhu H (2022) Plant disease detection and classification method based on the optimized lightweight YOLOv5 model. Agriculture 12(7):931
14. Elaraby A, Hamdy W, Alruwaili M (2022) Optimization of deep learning model for plant disease detection using particle swarm optimizer. Comput Mater Cont 71(2)
15. Saleem MH, Potgieter J, Mahmood Arif K (2022) A performance-optimized deep learning-based plant disease detection approach for horticultural crops of New Zealand. IEEE Access 10:89798–89822
16. Shah D, Trivedi V, Sheth V, Shah A, Chauhan U (2022) ResTS: residual deep interpretable architecture for plant disease detection. Inf Proc Agricul 9(2):212–223
17. Panchal AV, Patel SC, Bagyalakshmi K, Kumar P, Khan IR, Soni M (2023) Image-based plant diseases detection using deep learning. Mater Today: Proc 80:3500–3506
18. Thapa R, Zhang K, Snavely N, Belongie S, Khan A (2020) The plant pathology challenge 2020 data set to classify foliar disease of apples. Appl Plant Sci 8(9):e11390
19. Fenu G, Malloci FM (2021) DiaMOS plant: a dataset for diagnosis and monitoring plant disease. Agronomy 11(11):2107
20. Singh D, Jain N, Jain P, Kayal P, Kumawat S, Batra N (2020) PlantDoc: a dataset for visual plant disease detection. In: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, pp 249–253
21. Parraga-Alava J, Cusme K, Loor A, Santander E (2019) RoCoLe: a robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. Data Brief 25:104414
22. Krohling RA, Esgario J, Ventura JA (2019) BRACOL–a Brazilian Arabica coffee leaf images dataset to identification and quantification of coffee diseases and pests. Mendeley Data 1
23. Hughes D, Salathe M (2015) An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing