A-6371
[1-11]

REVIEW ARTICLE Indian Journal of Agricultural Research

AI-powered Innovations in Agriculture: A Systematic Review on Plant Disease Detection and Classification

Anita Rani Mehta[1], Pardeep Kumar[1], Guru Prem[2], Shalini Aggarwal[3], Rakesh Kumar[4]

10.18805/IJARe.A-6371

ABSTRACT

One of the primary needs of humans is food, which can be obtained through farming. Not only does agriculture meet the necessities of humankind, but it is also a primary source of employment. Agriculture is the main driver of employment and economic growth for a growing country such as India. For a thriving agricultural and economic sector, plant disease identification is a crucial concern to enhance productivity requires attention. Conventional approaches for detecting plant diseases are gruelling, time-consuming and demand a great deal of experience. Plant disease can be detected in a timely manner as it emerges on plant leaves, with the utilization of numerous machine learning and deep learning approaches. These high precision and time-efficient techniques are the way forward to generate qualitative farm produce. The current article examined a few methods currently used for analysing data sources, feature extraction, data augmentation and classification of plant diseases. PRISMA guidelines have been utilized to select the articles, using various keywords from peer-reviewed articles published in several databases between 2016 and 2025. Following the removal of studies based on the abstract, title, full text and conclusion, 75 publications were found and examined for their immediate relevance to plant disease recognition and categorization. Of these, 45 publications have been selected for this systematic review.

Key words: Deep learning techniques, Machine learning techniques, Plant diseases detection and classification.

India is the seventh-largest nation geographically, occupying 328 million hectares (Mha) area with 2.4% land area of the world (Bhattacharyya *et al.,* 2015). To feed the ever-increasing population, India has increased 6.2 times the food grain production of 314 million tons (Mt) from a meagre of 51 Mt in 1950-51. Moreover, the production of fruits and vegetables increased 13.3 times, *i.e.* from 25 to 333 Mt during the same period (Pathak *et al.,* 2022). Providing safe and healthy food to 600 million Indians expected to reside in urban areas by 2030 is the biggest challenge as land, water and air, the important natural resources, are continuously deteriorating in nature (Hoegh-Guldberg *et al.,* 2018). Diseases inception in plants may seriously affect the growth of the crop. Various infectious agents, *i.e.* pathogens like fungi, viruses, bacteria, nematodes, protozoa, prokaryotes and eukaryotes are causative agents for disease in plants (Abdulkhair and Alghuthaymi, 2016). Pathogen-generated diseases can migrate from one plant to another and destroy plant cell tissues, roots, leaves, stems and fruit (Savary *et al.,* 2012). The diseases can cause significant loss of total agricultural production (Goel and Nagpal, 2023).

Farmers rely on traditional ways of determining agricultural disease by visually inspecting the crops (Pham *et al.,* 2020). Sometimes, agricultural specialists are also unable to recognize the specific disease in any plant. The reasons were the high degree of complexity, variety of cultivated plants and conventional techniques involved in symptomatic identification of the plant leave ailment, which led to the selection of wrong protection treatments for plants (Ferentinos, 2018). Conventional scientific methods consist of gathering samples from the field and chemically analysing them, which are labour-intensive, time-consuming and have a limited scope of disease detection in plants (Kang *et al.,* 2023). The earliest detection of illnesses in the plants is the way forward to prevent crop damage and increase the quality and productivity of the crops. For effective control of the ailment, it is crucial to utilize accurate and reliable ways to assess the severity of diseases in plants. Artificial intelligence (AI) and computer vision have been extensively employed to recognize plant leaf diseases, classify plant species and find the severity of plant diseases. Continuous development in electronic device hardware performance and computer imaging technology in recent years enables the scientific community to develop modern tools for disease diagnoses in

[1]Department of Computer Science and Applications, Kurukshetra University, Kurukshetra-136 119, Haryana, India.
[2]School of Core Engineering, Shoolini University, Solan-173 229, Himachal Pradesh, India.
[3]Shaheed Udham Singh Government College, Matak Majri, Indri, Karnal-132 001, Haryana, India.
[4]Natural Resource Management Division, Indian Agricultural Research Institute, New Delhi-110 012, India.

Corresponding Author: Anita Rani Mehta, Department of Computer Science and Applications, Kurukshetra University, Kurukshetra-136 119, Haryana, India. Email: anita_mehta.dcsa@kuk.ac.in
ORCID: 0009-0005-7940-5653.

How to cite this article: Mehta, A.R., Kumar. P., Prem. G., Aggarwal. S. and Kumar. R. (2025). AI-powered Innovations in Agriculture: A Systematic Review on Plant Disease Detection and Classification. Indian Journal of Agricultural Research. 1-11. doi: 10.18805/IJARe.A-6371.

Submitted: 31-01-2025 **Accepted:** 21-05-2025 **Online:** 04-06-2025

agriculture (Liang *et al.,* 2019). Machine learning (ML) and deep learning (DL) approaches are progressively employed to recognize crop ailment from imaging data, incorporating feature extraction with classification. The illness classification and recognition strategies aim to assist non-expert users, such as those who are neither pathologists nor botanists. These techniques have been extensively utilized to detect diseases like early and late blight, scrab, black rot, powdery mildew, bacterial blight, brown leaf spot and leaf rust in fruit, vegetable and cereal crops using ML and DL models (Joseph *et al.,* 2024; Patil and More, 2025; Shafik *et al.* 2024; Mehta *et al.,* 2025)

These technological developments in computer vision have the potential to decrease labour costs, minimize time wastage and upgrade the quality of crops and yield overall. Thus, this systematic review aims to analyze the existing methods to detect and classify plant ailments by applying numerous ML and DL architectures. It also examined different optimization and classification techniques for datasets, accuracy, difficulties and outcomes. It can help farmers plan appropriate and proper treatments and researchers can create more dependable prediction systems. Additionally, this review will assist academics in determining how learning algorithms might be tailored to fit more complex models, making disease prediction more accurate for use in agriculture. In light of this, recommendations and suggestions for the future are also offered, which may help researchers working in this context to make progress in forecasting techniques.

Contribution and organization of the paper

This study provides an in-depth comparison of current research works applying AI-based methodologies. The majority of the methods in the various papers have significant results and are based on the implementation of ML and DL frameworks. It includes the Analysis of various crop illnesses, distinct augmentation techniques utilized to increase dataset, various features extracted and feature extraction techniques, crop disease recognition and classification utilizing DL and ML architectures. It also includes challenges corresponding to recognize and categorising plant illnesses in AI based methods. Following the introduction, the paper is arranged as follows. The research methodology and approaches used to select the literature are discussed in section 2. Section 3 discussed the challenges faced by researchers in conducting research. The conclusion of this paper and potential future directions are covered in Section 4.

The PRISMA guidelines are followed in this systematic literature review utilizing keywords such as plant disease detection, crop disease recognition and image processing using ML and DL. Fig 1 shows the many screening phases of the PRISMA flowchart, which guides the process of choosing from the collected papers. This review evaluates the relevance of 45 selected papers on crop disease recognition and categorization using deep learning and machine learning architectures. The papers are empirically investigated, with results tabulated and graphed to answer study questions.

Literature review

A comprehensive evaluation process was carried out by keeping an eye on the current models to address the following research questions formulated to write a systematic review.

RQ1: What are the most prevalent diseases which impact various plants?

RQ2: Which methods are employed for data augmentation?

RQ3: Which feature extraction techniques are employed in plant disease detection?

RQ4: Which ML and DL techniques are utilized to identify crop diseases?

RQ5: Which datasets are frequently employed to identify and categorize plant ailments?

RQ6: Which challenges are faced by the researchers in the classification process?

Analysis of various diseases detected in plants

Timely, fast, accurate and automatic recognition of plant diseases is critically required to reduce the horizontal spread of crop infections and the associated yield and economic losses. Utilizing modern image analysis tools to identify plant diseases automatically will also lower the labour costs of closely monitoring crops for potential infections. Using leaves images, three different kinds of diseases were identified in lady finger, *i.e.* yellow mosaic vein, powdery mildew and leaf spot (Sahithya *et al.,* 2019). Different leaf and fruit diseases such as whitefly, rust, algal leaf spot, fruit canker, fruit rot and anthracnose in guava were identified by (Howlader *et al.,* 2019; Farhan Al Haque *et al.,* 2019). Images of mango leaves were utilized to recognize three distinct diseases, *i.e.* anthracnose, gall midge and powdery mildew (Pham *et al.,* 2020). (Abbas *et al.,* 2021) utilized DenseNet121, a DL tool for diagnosing nine kinds of diseases yellow leaf curl virus, mosaic virus, two-spotted spider mite, septoria leaf spot, target spot, late blight, mosaic virus, leaf mould and early blight from leaf photos of tomato. (Kaur *et al.,* 2022) utilized CNN to recognise black measles, mosaic virus and leaf blight in grape plants using transfer learning and EfficientNet B7. A model was proposed by (Datta and Gupta, 2023) to recognize gray blight, brown blight, algal spot, red spot and helopeltis in tea leaves. Using images (Singh *et al.,* 2024) performed a comparative analysis of various DL approaches to detect late blight and early blight in potato. (Joseph *et al.,* 2024) provided a methodology for diagnosing four fungal illnesses of wheat, four fungal illnesses of maize, two fungal and bacterial illnesses of rice. Research on diseases in crops like wheat, grapes, potato, tomato, guava, mango, maize, rice and tea highlights advancements in agricultural disease detection.

Analysis of various features and feature extraction techniques

Features are essential to recognise patterns since they assist in item description. To create feature vectors, features recommended by specialists are extracted from the picture. Texture, shape and colour are the three different feature categories that (Es-saady *et al.,* 2016) identified. The colour histogram, colour moments (mean, skewness and standard de*via*tion) and colour structure descriptor were used to extract color features, whereas texture features were obtained by applying Grey Level Co-occurence Matrix. Shape features were complexity, circularity, area and perimeter. (Kumari *et al.,* 2019) utilized K-means clustering to extract different features, whereas (Pham *et al.,* 2020) used CLAHE and Wrapper techniques to find Area, Eccentricity, Convex area, Mean, Standard de*via*tion, Kurtosis, Skewness, Contrast and homogeneity. Scale-invariant feature transform was used by (Chouhan *et al.,* 2021) to extract features. Texture and colour features are the two primary categories of information retrieved from the pictures (Jain and Dharavath, 2023).

The colour moment equations mean, skewness, standard de*via*tion and kurtosis were applied to determine colour characteristics and texture features were retrieved using GLCM. More recent studies (Aboelenin *et al.,* 2025) applied DL model Inception-V3, DenseNet20 and VGG16 for feature extraction in apple and corn crops. (Chavan *et al.,* 2025) extensively used the important features extraction based on color, shape and texture using LGXP (Local Gabor XOR pattern) technique for detailed classification of crops and identification of diseases in crops. The summary of various feature extraction techniques is given in Table 1.

Accurate feature extraction is crucial for early illness recognition, reducing manual inspections and promoting sustainable farming practices. Techniques like color histograms, color moments, GLCM, LBP, HOG and SIFT help identify disease-related structural abnormalities in plant images. Deep learning techniques have revolutionized feature extraction by automatically learning and retrieving features from unprocessed picture data, reducing crop loss and optimizing disease management. These methods help identify patterns of discolouration, texture features and disease-related structural abnormalities.

Analysis of various augmentation techniques

Data augmentation is an approach to enlarging the dataset by utilising distinct approaches like rotation, scaling, shearing, flipping, cropping and zooming to reduce overfitting and improve effectiveness. Data augmentation helps to expand the quantity of positive and negative instances for contrastive learning, which improves the effect of contrastive learning. This also makes a limited amount of data generate value similar to greater data without increasing the data (Shorten and Khoshgoftaar, 2019). In order to reduce the problem of overfitting, (Farhan Al Haque *et al.,* 2019) used several augmentation methods, including horizontal flipping, nearest fill, width and height shifting, zooming, shearing and rotation. To enhance the quantity of photos, (Zhang *et al.,* 2019) employed geometric and intensity adjustments. Five methods were employed for the intensity changes: brightness enhancement, PCA jittering, colour jittering, blur (radial) and contrast enhancement. Pictures underwent horizontal and vertical geometric modifications, including enlargement, cropping, rotation and flipping. To enhance the number of leaf images of the lady finger crop, (Selvam and Kavitha, 2020) carried out distinct augmentation operations: zooming, rotation, shearing, horizontal flipping and shifting of width and height. (Akshai and Anitha, 2021) applied various augmentation methods, including zooming, rotation and shifting, to enlarge the dataset and minimize the issue of overfitting.

The SMOTE (synthetic minority oversampling approach) was employed by (Divakar *et al.,* 2021) to increase the quantity of images, which is a statistical way to enlarge the dataset in a balanced way. To expand the dataset size and decrease overfitting, (Vallabhajosyula *et al.,* 2022) used four distinct augmentation techniques, *i.e.* scaling (resizing), rotation, translation and picture enhancement. The size of dataset was increased from 55,448 pictures to 234,008 pictures (Pandian *et al.,* 2022) by applying position and colour augmentation, principal component analysis (PCA), deep convolutional generative adversarial network (DCGAN) and neural style transfer. (Joseph *et al.,* 2024) applied different augmentation approaches like vertical flipping, horizontal flipping, rotation of images by 90 degrees, Shearing and Zooming by 0.2, width and height Shifting by the range of 0.2 and Brightness enhancement (0.2-0.8) for expanding the dataset from 1500 images to 25000 images. (Ashurov *et al.,* 2025) employed luminance adjustment, flipping and rotation techniques to enhance the dataset size. The summary of different augmentation techniques is given in Table 2. The datasets diversity is enhanced through augmentation techniques like random changes in object orientation, aspect ratio adjustments, luminance variations, scaling, rotation, cropping, zooming and flipping are used to improve the robustness and representativeness of crop disease data. These techniques solve overfitting problems, improving accuracy and reliability in identifying plant illnesses. By providing a wider range of datasets, models can perform better in real-world problems and generalize more effectively.

Analysis of various classification techniques used in plant disease detection

In image processing and computer vision, disease classification is the most prominent phase in detecting plant diseases. The effectiveness of this phase, which is crucial for disease identification, also depends on preliminary techniques such as acquisition, preprocessing and feature extraction. Numerous classification approaches have been investigated and applied to detect and categorize plant diseases utilizing leaf images. Support Vector Machine was utilized to recognise various diseases by (Hossain *et al.,* 2018; Hou *et al.,* 2021; Mukhopadhyay *et al.,* 2021; Singh and Kaur, 2021), achieving an accuracy of 93%, 97.40%, 83% and 95.99% respectively. (Kumar and Vani, 2019) applied VGG16 to identify multiple diseases in tomato having accuracy of 99.11%. (Tiwari *et al.,* 2020) utilized VGG19 architecture to recognize late blight and early blight in leaf images of potato and attained 97.80% accuracy. (Jadhav *et al.,* 2021) utilized GoogleNet and AlexNet to detect brown spot, bacterial blight and frogeye spot in soyabean leaves and attained 96.25% and 98.75% accuracy, respectively. Moreover (Ajra *et al.,* 2020) employed AlexNet and ResNet50 to classify unhealthy and healthy leaves with an accuracy of 96.5%, 97% in potato, 96% and 95.3% in tomato (Iqbal and Talukder, 2020) applied random forest for classification of healthy, late blight and early blight in potato leaves with 97% accuracy.

(Hong *et al.,* 2020) used DenseNet and (Kibriya *et al.,* 2021) employed VGG16 and GoogleNet to classify multiple diseases on tomato, achieving an accuracy of 97.10%, 98% and 99.23%. (Pham *et al.,* 2020) employed various deep learning models to find diseases such as powdery mildew, anthracnose and gall midge on mango with an accuracy of 89.41%, 78.64%, 79.92% and 84.88%, respectively. (Sambasivam and Opiyo, 2021; Khalifa *et al.,*

Table 1: Feature extraction techniques to extract distinct features.

Feature extraction techniques applied	Features extracted	Reference and year
GLCM, color moments, structure descriptor and histogram	Color, shape and texture features	(Es-saady *et al.,* 2016)
K-means clustering	Energy, Contrast, standard deviation, Correlation, mean, variance and homogeneity	(Kumari *et al.,* 2019)
CLAHE and Wrapper based	Texture and shape features	(Pham *et al.,* 2020)
SIFT	Shape features	(Chouhan *et al.,* 2021)
GLCM	Texture and color features	(Jainand Dharavath,2023)
Local Gabor XOR pattern	Color, shape and texture features	(Chavan *et al.,* 2025)
Inception-V3, DenseNet20 and VGG16	Global features	(Aboelenin *et al.,* 2025)

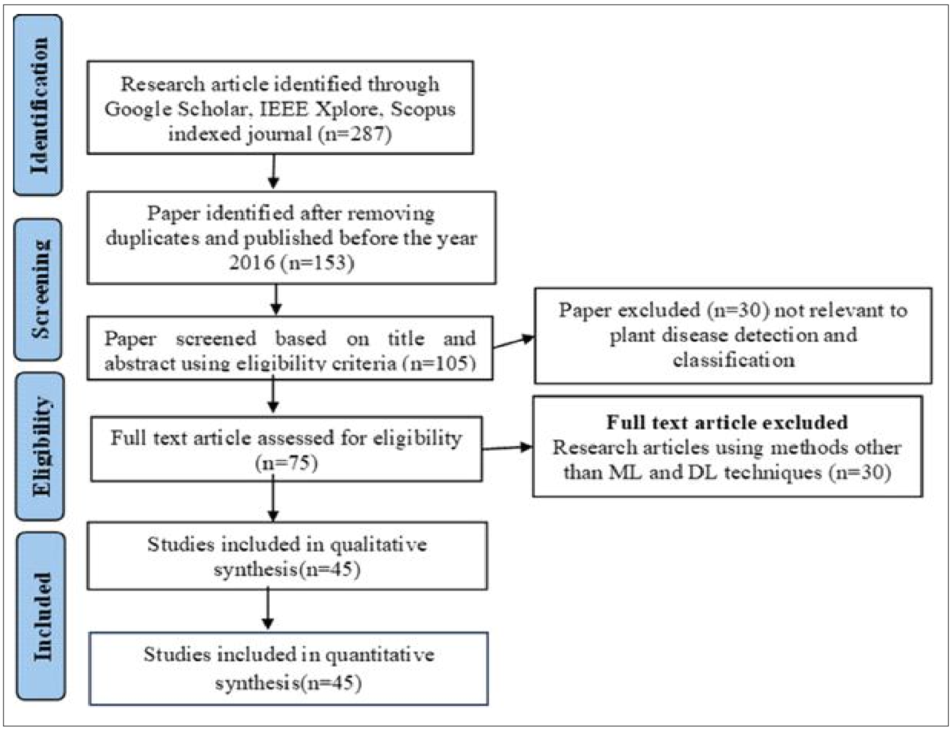

Fig 1: PRISMA guidelines of the search process.

2021; Akbar *et al.,* 2022; Datta and Gupta, 2023) applied CNN to detect different diseases on potato, peach and tea, attaining an accuracy of 98%, 99% and 96.56%, respectively. (Kaur *et al.,* 2022) developed a model by applying Efficient NetB7 for feature reduction and logistic regression for categorization of black rot, leaf blight and black measles on grapes leaves and achieved 98.7% accuracy. (Attallah, 2023) proposed a novel technique with transfer learning to retrieve features and KNN and SVM to classify distinct diseases on tomato, attaining 99.92% and 99.90% accuracy. (Singh and Yogi, 2023) did a comparison of Resnet with other deep learning models to classify healthy leaves, late blight and early blight of potato and 99.62% accuracy was achieved.

(Joseph *et al.,* 2024) utilized eight fine-tuned deep learning architectures and demonstrated that MobileNet, Xception and CNN performed the best in identifying illnesses of maize leaves, with testing accuracy of 94.64%,95.80% and 97.04%respectively. Similarly, the architectures MobileNetV2, MobileNet and CNN outperformed in identifying the leaf illnesses of wheat, with testing accuracy of 96.32%, 96.28% and 98.08%, respectively. In terms of identifying leaf illnesses of rice, the Inception V3, Xception and CNN models outperformed, with testing accuracy of 96.20%, 97.28% and 97.06%, respectively. (Sutiaji *et al.,* 2024) suggested a novel weighted deep learning ensemble technique to enhance the performance of plant disease detection. By ensembling the architecture with a combination of two and three pretrained CNN models, applied transfer learning to individual CNN architectures by using weight updation on the final few layers to prioritise high-dimensional features. Grid search was used to find each model's optimal weights for ensembling the models. Evaluation metrics showed that the three-model ensemble outperformed the two-model ensemble. The MobileNetV2-Xception-DenseNet121 and MobileNetV2-DenseNet121 ensemble models have the best accuracy values, which are 99.49% and 99.56%, respectively. For effective plant disease diagnosis and classification, (Shafik *et al.,* 2024) presented two Plant Disease Detection (PDDNet) architectures: the lead voting ensemble (LVE) and early fusion (AE), which are merged with some pre-trained convolutional neural networks and fine-tuning is done by applying deep feature extraction. Prominent pre-trained networks such as ResNet50, ResNet101, DenseNet201, AlexNet, EfficientNetB7, GoogleNet, ResNet18, Conv Next Small and NAS Net Mobile, were used for hyper parameter fine-tuning. Finally, logistic regression is applied to assess the effectiveness of distinct CNN architectures. Additionally, the suggested model, deep learning classifiers and related recent studies were compared. The evaluations showed that PDDNet-LVE and PDDNet-AE outperformed existing CNNs in terms of robustness and generalisation, achieving 97.79% and 96.74%, respectively, on tests of multiple plant diseases. Kanade *et al.* (2025) utilized 4,480 real-time leaf photographs and found CNN-based YOLOv8 model exhibited superior precision, recall and F1-score of 97.2%, 78.6% and 86.9%, respectively. Table 3 shows classification approaches used in plant disease detection.

Discussion on classification techniques

Numerous classification approaches have demonstrated high accuracy rates in various studies. ML algorithms, such as Decision Trees, Support Vector Machines (SVM) and Random Forests, have been widely applied because of their potential to manage huge and complicated datasets and provide interpretable results. Depending on the usage of feature selection, preprocessing methods and the size of dataset, these algorithms often attain high accuracy levels, ranging from 80% to 99%. These are particularly effective when trained on well-curated datasets with distinct features. In contrast, DL models, particularly Convolutional Neural Networks (CNNs), have revolutionized the recognition and categorization of crop ailments by instinctively learning hierarchical features from raw image data. CNNs excel in capturing complex patterns and peculiarities in plant diseases, achieving accuracies upwards of 95% in numerous studies. Their ability to adapt and generalize across diverse datasets and environmental conditions underscores their superiority in complex image classification

Table 2: Augmentation techniques used in plant disease detection and classification.

Applied augmentation techniques	Reference and year
Rotation, zooming, nearest fill, horizontal flipping, height and breadth shifting and Shearing	(Farhan Al Haque *et al.,* 2019)
Geometric transformations and Intensity transformations	(Zhang *et al.,* 2019)
Shearing, horizontal flipping, zooming, height and width shifting and rotation	(Selvam and Kavitha, 2020)
Shifting, image Zooming and rotation	(Akshai and Anitha, 2021)
Synthetic minority oversampling technique (SMOTE)	(Divakar *et al.,* 2021)
Translation, enhancement, rotation and scaling	(Vallabhajosyula *et al.,* 2022)
Augmentation of color and position, Principal component analysis, neural style transfer and deep convolutional GAN	(Pandian *et al.,* 2022)
Horizontal flipping, vertical flipping, shear, zoom, height shift, width shift, rotation and brightness	(Joseph *et al.,* 2024)
Luminance adjustment, flipping and rotation	(Ashurov *et al.,* 2025)

Table 3: Various classification approaches in plant disease detection.

Techniques used	Crop	Diseases detected	Dataset	Accuracy	Reference and year
SVM	Tea	Brown blight, algal	Self-captured	93%	(Hossain *et al.,* 2018)
VGG16	Tomato	Multiple diseases	Plant village	99.11%	(Kumar and Vani, 2019)
VGG19	Potato	Early blight, late blight, healthy	Plant village	97.80%	(Tiwari *et al.,* 2020)
AlexNet and ResNet50	(a)Potato, (b)Tomato	Healthy and classification unhealthy	Kaggle dataset (b) 96 and 95.3%	(a) 97 and 96.5%	(Ajra *et al.,* 2020)
Random Forest	Potato	Early blight, late blight and healthy	Plant village	97.00%	(Iqbal and Talukder, 2020)
DenseNet	Tomato	Multiple diseases	Plant village	97.10%	(Hong *et al.,* 2020)
ANN, AlexNet, VGG16, ResNet50	Mango	Anthracnose, powdery mildew and gall midge	Self-captured (450 images)	89.41, 78.64, 79.92 and 84.88%	(Pham *et al.,* 2020)
GoogleNet, VGG16	Tomato	Multiple diseases	Plant village	99.23% and 98%	(Kibriya *et al.,*
AlexNet and GoogleNet	Soybean	Bacterial blight, Frogeye leaf spot and Brown spot	Self-captured images (1199)	98.75 and 96.25%	2021)
SVM	Potato	Early blight, late blight and healthy	Plant village	95.99%	(Singh and Kaur, 2021)
SVM	Tea	Red rust, thrips, heloptis, red spider, sunlight	Own	83.00%	(Mukhopadhyay *al., et* 2021)
			scorching		
SVM	Potato	Early blight, late blight	AI Challenger Global AI Contest dataset	97.40%	(Hou *et al.,* 2021)
CNN	Potato	Late blight, Early blight and healthy	Kaggle compitition (1722 images)	98.00%	(Khalifa *et al.,* 2021)
LWNet (Light Weight CNN)	Peach	Bacterial leaf spot	Self-captured images 4500 bacteriosis and 5500 healthy	99.00%	(Akbar *et al.,* 2022)
Logistic regression	Grapes	Black rot, leaf blight and black measles	Plant village	98.70%	(Kaur *et al.,* 2022)
CNN	Tea	Gray blight, algal spot, red spot, brown blight and helopeltis	Self-created (2800 images)	96.56%	(Datta and Gupta, 2023)
KNN, SVM	Tomato	Multiple diseases	Plant village	99.92 and 99.90%	(Attallah, 2023)
ResNet, VGG16, InceptionV3, MobileNetV2, NasNetV2, VGG19	Potato	Early blight, late blight and healthy	Plant village	99.62, 98.88, 98.0, 94.56, 96.0 and 97.40%	(Singh and Yogi, 2023)
Xception, Mobilenet, CNN	Rice, Wheat, Maize	Multiple diseases	ImageNet Dataset and Self-Developed	(Rice-97.28, 96.20, 97.06%), (Wheat-96.32%,	(Joseph *et al.,* 2024)

Table 3: Continue...

Table 3: Continue...

			Dataset	96.28%, 98.08%), (Maize-95.80%, 94.64%,97.04%)	
Mobile netV2 -Dense net121 and Mobile net V2-Xception -Dense net	Multiple Crops	Multiple diseases	Plant Village Dataset	99.49% and 99.56%	(Sutiaji *et al.*, 2024)
121 PDD net- Early fusion and PDDNet- Lead voting ensemble	Multiple Crops	15 Multiple diseases	Plant Village Dataset	96.74% and 97.79%	(Shafik *et al.*, 2024)

Table 4: Most common type of datasets for detection of plant diseases.

Name of dataset	Dataset description	Link
Plant pathology 2020- FGVC7	This dataset contains 3642 leaf images of Apple having three categories Scab, Rust and Healthy.	https://www.kaggle.com/c/plant-pathology-2020-fgvc7/data?select=images
Plant village dataset	This dataset contains 54303 unhealthy and healthy leave images distributed into 38 classes by disease and species.	https://github.com/spMohanty/PlantVillage-Dataset
Plant doc dataset	It has 13 types of plants and 17 types of diseases.	https://github.com/pratikkayal/PlantDoc-dataset
New plant disease	This dataset contain pictures of infected and healthy leaves of several plants, *i.e.* Tomato, corn, apple, squash, orange, grape, peach, strawberry, pepper bell, blueberry, potato, cherry, soybean, and raspberry.	https://www.kaggle.com/vipoooool/new-plant-diseases-dataset
UCI	This dataset has 19 distinct classes. It has 307 leaf images of soybean and 120 leaf images of rice.	https://archive.ics.uci.edu/ml/datasets/
OSF	This dataset has 123,957 images of maize leaves (18,222 real images and 105,735 images anno tated by human experts)	https://osf.io/p67rz/
Rice Diseases Image Dataset	This dataset contains 513 pictures of Brown spot diseases, 779 pictures of leafblast, 565 pictures of Hispa and 1488 healthy leaves images of rice	https://www.kaggle.com/minhhuy2810/
Disease prediction dataset of cotton	The National College of Ireland publishes this dataset containing 1935 images of cotton leaves.	https://universe.roboflow.com/national-college-of-ireland/cotton-plant-disease-prediction-igthk

tasks. These developments show significant promise for real-world agricultural applications by highlighting the ability of complex classification techniques to improve the precision and robustness of crop disease diagnosis.

The accuracy rate of various ML and DL architectures is demonstrated in Fig 2. With 99.90% accuracy, SVM performs the best classification tasks. Convolutional Neural Networks (CNN), with 99% accuracy, demonstrate their efficacy in intricate data patterns. The resilience of various deep learning architectures such as ResNet50, Google Net and VGG16 in image processing tasks is demonstrated by the accuracy of 99.62%, 99.23% and 99.11%, respectively. The remarkable accuracy of 99.56% achieved by ensemble methods such as the MobileNetV2-Xception-DenseNet21 combination suggests that incorporating different models can improve performance. Overall, deep learning models high accuracies highlight their superiority in managing challenging tasks. In contrast, the effectiveness of conventional approaches highlights the necessity of choosing the right models depending on particular application requirements.

Datasets used in plant disease detection

Gathering data from plant leaf imaging is the first step in identifying and classifying leaf diseases. Using a camera device, one can obtain image data from open-source repositories or personally snap pictures of plant leaves.

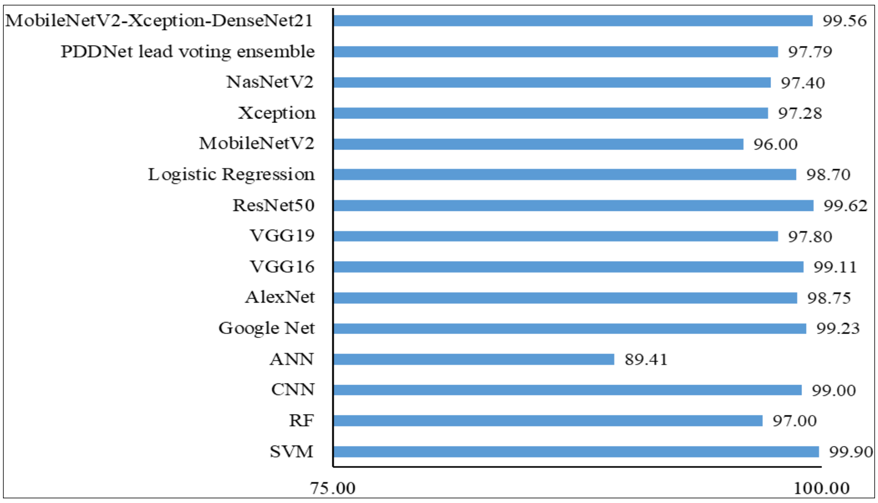

Fig 2: ML and DL based classification techniques with their accuracy (%).

This section discusses various resources many researchers have used to gather image data for their studies. Table 4 shows a detailed description of various datasets and their corresponding links. In conclusion, collecting image data is essential for determining the kind and severity of plant leaf diseases. Researchers have employed open-source repositories and customized datasets like Plant Pathology Dataset, Plant Village, New Plant Diseases Dataset and PlantDoc to collect image data for analysing and detecting plant leaf diseases, enhancing efficiency.

Critical issues and challenges

i. The shortage of sufficient data to train the models data is a common problem for researchers. Smaller data sets are more difficult for researchers since they might lead to over-fitting issues, which provide new challenges.

ii. Neural networks have high computation time and high complexity because they have many layers. Another frequent problem faced by researchers is the increased dimensionality of data. It is because of data having a large number of features referred as dimensionality.

iii. Despite the numerous studies to apply machine learning approaches in farming, farmers may only use a few mobile-based applications. Mobile platforms must be coupled with the models created by different researchers in order for to use and apply them more easily by the farmers.

iv. In certain studies, CNN performs better when certain hyperparameters are adjusted, including weight, max epoch, bias learning rate and minibatch size. Tuning specific parameters in pre-established techniques, such as voting classifier, AdaBoost, XGBoost, Random Forest and Decision Tree and combining them with nature-inspired optimization techniques to improve performances.

v. The researchers have mainly used lab-created datasets with high illumination, contrast, regions and other characteristics. It is possible that some requirements would not be satisfied in real life when preparing the dataset. Therefore, real-world scenarios may not yield optimal performance from models trained on lab-created datasets.

CONCLUSION

This systematic literature review summarised 45 publications on plant illness identification and classification utilizing distinct ML and DL approaches. The review presented numerous approaches to optimize the models and classify plant diseases. Furthermore, the review included the datasets that were utilized in different studies. The review also emphasized how accurate the current technologies are. It was discovered that although deep learning and machine learning techniques are widely used for the recognition and categorization of crop diseases, there is still uncertainty regarding the models interpretability. The practical implementation of crop disease recognition techniques cannot find a promising scope in agriculture unless these challenges are resolved. Therefore, by taking into account the limitations of the current approaches, researchers in the future can contribute by suggesting way to increase the interpretability, generalisability and accuracy for detection and categorization. In addition to the previously analysed ML and DL techniques, emerging AI techniques in detecting plant diseases will be analysed for future scope and their potential outcomes will be examined to help researchers and developers propose new approaches for identifying and categorising accurately and early crop ailments.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

REFERENCES

Abbas, A., Jain, S., Gour, M. and Vankudothu, S. (2021). Tomato plant disease detection using transfer learning with C-GAN synthetic images. Computers and Electronics in Agriculture. 187: 106279. https://doi.org/10.1016/j. compag.2021.106279.

Abdulkhair, W.M. and Alghuthaymi, M.A. (2016). Plant Pathogens. In Plant Growth. InTech. https://doi.org/10.5772/65325

Aboelenin, S., Elbasheer, F.A., Eltoukhy, M.M., El-hady, W.M. and Hosny, K.M. (2025). A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. Complex and Intelligent Systems. 11(2): 142. https://doi.org/10.1007/s40747-024-01764-x

Ajra, H., Nahar, Mst. K., Sarkar, L. and Islam, Md. S. (2020). Disease detection of plant leaf using image processing and CNN with preventive measures. 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE). 1-6. https://doi.org/10.1109/ETCCE51779.2020. 9350890.

Akbar, M., Ullah, M., Shah, B., Khan, R.U., Hussain, T., Ali, F., Alenezi, F., Syed, I. and Kwak, K.S. (2022). An effective deep learning approach for the classification of Bacteriosis in peach leave. Frontiers in Plant Science. 13: https://doi.org/10.3389/fpls.2022.1064854

Akshai, K.P., and Anitha, J. (2021). Plant disease classification using deep learning. 2021 3rd International Conference on Signal Processing and Communication (ICPSC). 407-411. https://doi.org/10.1109/ICSPC51351.2021.9451696

Ashurov, A.Y., Al-Gaashani, M.S.A.M., Samee, N.A., Alkanhel, R., Atteia, G., Abdallah, H.A. and Saleh Ali Muthanna, M. (2025). Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections. Frontiers in Plant Science. 15: https://doi.org/10.3389/fpls.2024.1505857.

Attallah, O. (2023). Tomato leaf disease classification *via* compact convolutional neural networks with transfer learning and feature selection. Horticulturae. 9(2): https://doi.org/10.3390/horticulturae9020149

Bhattacharyya, R., Ghosh, B., Mishra, P., Mandal, B., Rao, C., Sarkar, D., Das, K., Anil, K., Lalitha, M., Hati, K. and Franzluebbers, A. (2015). Soil Degradation in India: Challenges and Potential Solutions. Sustainability. 7(4): 3528-3570. https://doi.org/10.3390/su7043528

Chavan, P., Chavan, P.P. and Chavan, A. (2025). Hybrid architecture for crop detection and leaf disease detection with improved U-net segmentation model and image processing. Crop Protection: 107117. https://doi: org/10.1016/j.cropro.2025.107117

Chouhan, S.S., Singh, U.P. and Jain, S. (2021). Automated Plant Leaf Disease Detection and Classification Using Fuzzy Based Function Network. Wireless Personal Communications. 121(3): 1757-1779. https://doi: org/10.1007/s11277-021-08734-3.

Datta, S. and Gupta, N. (2023). A novel approach for the detection of tea leaf disease using deep neural network. Procedia Computer Science. 218: 2273-2286. https://doi: org/10.1016/j.procs.2023.01.203.

Divakar, S., Bhattacharjee, A. and Priyadarshini, R. (2021). Smote-DL: A Deep Learning Based Plant Disease Detection Method. 2021 6th International Conference for Convergence in Technology (I2CT). 1-6. https://doi: org/10.1109/I2CT51068.2021.9417920

Es-saady, Y., El Massi, I., El Yassa, M., Mammass, D. and Benazoun, A. (2016). Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers. 2016 International Conference on Electrical and Information Technologies (ICEIT). 561-566. https://doi: org/10.1109/EITech.2016.7519661

Farhan Al Haque, A.S.M., Hafiz, R., Hakim, Md.A. and Rasiqul Islam, G.M. (2019). A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach. 2019 22nd International Conference on Computer and Information Technology (ICCIT). 1-6. https://doi: org/10.1109/ICCIT48885.2019.9038598.

Ferentinos, K.P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture. 145: 311-318. https://doi: org/10.1016/j. compag.2018.01.009.

Goel, L. and Nagpal, J. (2023). A Systematic Review of Recent Machine Learning Techniques for Plant Disease Identification and Classification. IETE Technical Review. 40(3): 423-439. https://doi: org/10.1080/02564602.2022.2121772

Hoegh-guldberg, O., Jacob, D., Taylor, M., Bindi, M., Abdul Halim, S., Achlatis Australia, M., Alexander, L.V, Allen, M.R., Berry, P., Boyer, C., Brilli, L., Buckeridge, M., Byers Austria, E., Antonio Marengo Brazil, J., Pereira, J., Sherstyukov, B., Jacob, D., Taylor, M., Bindi, M., … Waterfield, T. (2018). Hubertus Fischer (Switzerland), Klaus Fraedrich (Germany). Sabine Fuss (Germany).

Hong, H., Lin, J. and Huang, F. (2020). Tomato disease detection and classification by deep learning. 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). 25-29. https://doi: org/10.1109/ICBAIE49996.2020.00012.

Hossain, S., Mou, R.M., Hasan, M.M., Chakraborty, S. and Razzak, M.A. (2018). Recognition and detection of tea leaf's diseases using support vector machine. 2018 IEEE 14th International Colloquium on Signal Processing and Its Applications (CSPA). 150-154. https://doi: org/10.1109/CSPA.2018.8368703

Hou, C., Zhuang, J., Tang, Y., He, Y., Miao, A., Huang, H. and Luo, S. (2021). Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation. Journal of Agriculture and Food Research. 5: 100154. https://doi: org/10.1016/j.jafr.2021.100154

Howlader, Md.R., Habiba, U., Faisal, R.H. and Rahman, Md.M. (2019). Automatic Recognition of Guava Leaf Diseases using Deep Convolution Neural Network. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). 1-5. https://doi: org/10.1109/ECACE.2019.8679421

Iqbal, Md.A. and Talukder, K.H. (2020). Detection of potato disease using image segmentation and machine learning. 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). 43-47. https://doi: org/10.1109/WiSPNET48689.2020.9198563

Jadhav, S.B., Udupi, V.R. and Patil, S.B. (2021). Identification of plant diseases using convolutional neural networks. International Journal of Information Technology. 13(6): 2461-2470. https://doi: org/10.1007/s41870-020-00437-5

Jain, S. and Dharavath, R. (2023). Memetic salp swarm optimization algorithm based feature selection approach for crop disease detection system. Journal of Ambient Intelligence and Humanized Computing. 14(3): 1817-1835. https://doi: org/10.1007/s12652-021-03406-3

Joseph, D.S., Pawar, P.M. and Chakradeo, K. (2024). Real-time plant disease dataset development and detection of plant disease using deep learning. IEEE Access. 12: 16310-16333. https://doi: org/10.1109/ACCESS.2024.3358333.

Kanade, A.K., Potdar, M.P., Balol, G. and Rathod, N. (2025) Artificial intelligence based precise disease detection in soybean using real time object detectors. Legume Research-An International Journal. 1: 6. https://doi: org/10.18805/LR-5431

Kang, F., Li, J., Wang, C. and Wang, F. (2023). A light weight neural Network-based method for identifying Early-blight and Late-blight leaves of potato. Applied Sciences (Switzerland). 13(3): https://doi: org/10.3390/app13031487.

Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A. and Alabdali, A. M. (2022). Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction. Sensors. 22(2): https://doi: org/10.3390/s22020575.

Khalifa, N.E.M., Taha, M.H.N., Abou El-maged, L.M. and Hassanien, A.E. (2021). Artificial intelligence in potato leaf disease classification: A Deep Learning Approach. In Studies in Big Data (Vol. 77, pp. 63-79). Springer Science and Business Media Deutschland GmbH. https://doi: org/10.1007/978-3-030-59338-4-4.

Kibriya, H., Rafique, R., Ahmad, W. and Adnan, S. M. (2021). Tomato leaf disease detection using convolution neural network. 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST). 346-351. https://doi: org/10.1109/IBCAST51254.2021.9393311

Kumar, A. and Vani, M. (2019). Image based tomato leaf disease detection. 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT): 1-6. https://doi.org/10.1109/ICCCNT45670.2019.8944692

Kumari, Ch.U., Jeevan Prasad, S. and Mounika, G. (2019). Leaf disease detection: Feature Extraction with K-means clustering and classification with ANN. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC).1095-1098.https://doi: org/10.1109/ICCMC. 2019.8819750.

Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D. and Sun, W. (2019). PD2SE-Net: Computer-assisted plant disease diagnosis and severity estimation network. Computers and Electronics in Agriculture. 157: 518-529. https://doi: org/10.1016/j.compag.2019.01.034.

Mehta, A.R., Kumar, P., Prem, G., Aggarwal, S. and Kumar, R. (2025). Leveraging Artificial Intelligence for Disease Diagnosis in Agricultural Crops: A Review. Indian Journal of Agricultural Research. 1-10. https://doi: org/10.18805/IJARe.A-6363

Mokhtar, U., El Bendary, N., Hassenian, A.E., Emary, E., Mahmoud, M.A., Hefny, H. and Tolba, M.F. (2015). SVM-based Detection of Tomato Leaves Diseases (pp. 641-652). https://doi: org/10.1007/978-3-319-11310-4-55.

Mukhopadhyay, S., Paul, M., Pal, R. and De, D. (2021). Tea leaf disease detection using multi-objective image segmentation. Multimedia Tools and Applications. 80(1): 753-771. https://doi: org/10.1007/s11042-020-09567-1

Padol, P.B. and Yadav, A.A. (2016). SVM classifier based grape leaf disease detection. 2016 Conference on Advances in Signal Processing (CASP). 175-179. https://doi: org/10.1109/CASP.2016.7746160.

Pandian, J. A., Kanchanadevi, K., Kumar, V. D., Jasiñska, E., Goòo, R., Leonowicz, Z. and Jasiñski, M. (2022). A five convo lutional layer deep convolutional neural network for plant leaf disease detection. Electronics (Switzerland). 11(8). https://doi.org/10.3390/electronics11081266.

Pathak, H., Mishra, J.P. and Mohapatra, T. (n.d.). Editors Indian Agriculture after Independence. www.icar.org.in.

Patil, R.G. and More, A. (2025). A comparative study and optimization of deep learning models for grape leaf disease identification. Indian Journal of Agricultural Research. 59(4): 654-663. https://doi: org/10.18805/IJARe.A-6242.

Pham, T.N., Tran, L. Van and Dao, S.V.T. (2020). Early disease classification of mango leaves using Feed-forward neural network and hybrid metaheuristic feature selection. IEEE Access. 8: 189960-189973. https://doi: org/10.1109/ACCESS.2020.3031914.

Sahithya, V., Saivihari, B., Vamsi, V.K., Reddy, P.S. and Balamurugan, K. (2019). GUI based detection of unhealthy leaves using image processing techniques. 2019 International Conference on Communication and Signal Processing (ICCSP). 0818-0822. https://doi: org/10.1109/ICCSP.2019.8697946.

Sambasivam, G. and Opiyo, G.D. (2021). A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. Egyptian Informatics Journal. *22*(1): 27-34. https://doi: org/10.1016/j.eij.2020.02.007.

Savary, S., Ficke, A., Aubertot, J.N. and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. Food Security. 4(4): 519-537. https://doi: org/10.1007/s12571-012-0200-5.

Selvam, L. and Kavitha, P. (2020). Classification of ladies finger plant leaf using deep learning. Journal of Ambient Intelligence and Humanized Computing. https://doi: org/10.1007/s12652-020-02671-y.

Shafik, W., Tufail, A., De Silva Liyanage, C. and Apong, R.A.A.H.M. (2024). Using transfer learning-based plant disease classification and detection for sustainable agriculture. BMC Plant Biology. 24(1): 136. https://doi: org/10.1186/s12870-024-04825-y.

Shorten, C. and Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data. *6*(1): 60. https://doi: org/10.1186/s40537-019-0197-0.

Singh, A. and Kaur, H. (2021). Potato Plant Leaves Disease Detection and Classification using Machine Learning Methodologies. IOP Conference Series: Materials Science and Engineering. 1022(1): 012121. https://doi: org/10.1088/1757-899X/1022/1/012121.

Singh, G. and Yogi, K. K. (2023). Comparison of RSNET model with existing models for potato leaf disease detection. Biocatalysis and Agricultural Biotechnology, 50. https://doi.org/10.1016/j.bcab.2023.102726.

Singh, I., Jaiswal, A. and Sachdeva, N. (2024). Comparative analysis of deep learning models for potato leaf disease detection. 2024 14th International Conference on Cloud Computing, Data Science andamp; Engineering (Confluence). 421-425.https://doi: org/10.1109/Confluence 60223.2024.10463314.

Sutiaji, D., Yildiz, O., Rosyid, H. and Chotijah, U. (2024). WECNN-PDP: Weighted Ensemble Convolutional Neural Networks Models to Improve the Plant Disease Prediction. E3S Web of Conferences. 482: 03003. https://doi: org/10.1051/e3sconf/202448203003.

Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S. and Bhardwaj, S. (2020). Potato Leaf Diseases Detection Using Deep Learning. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). 461-466. https://doi: org/10.1109/ICICCS48265.2020.9121067.

Vallabhajosyula, S., Sistla, V. and Kolli, V.K.K. (2022). Transfer learning-based deep ensemble neural network for plant leaf disease detection. Journal of Plant Diseases and Protection. 129(3): 545-558. https://doi: org/10.1007/s41348-021-00465-8.

Zhang, S., Zhang, S., Zhang, C., Wang, X. and Shi, Y. (2019). Cucumber leaf disease identification with global pooling dilated convolutional neural network. Computers and Electronics in Agriculture. 162: 422-430. https://doi: org/10.1016/j.compag.2019.03.012.