Majlesi J. Electr. Eng. 2026; 20(2) : 229-249

Majlesi Journal of Electrical Engineering (MJEE)

https://doi.org/10.57647/j.mjee.2025.17383

Review

A Review on Machine Learning Approaches for Plant Disease Diagnosis

Vijayan Subramanian, Chiranji Lal Chowdhary*

School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India

*Corresponding author: chiranji.lal@vit.ac.in

Article History

Received:
30 November 2024

Revised:
18 February 2025

Accepted:
7 March 2025

Published in Issue:
30 June 2026

Abstract:
In recent years, the identification and classification of plant and crop diseases have significantly increased due to the use of machine learning (ML) algorithms. This rise is attributed to the remarkable precision and potential of ML methods in automating plant disease identification. This comprehensive analysis examines various ML approaches and image processing techniques used to identify and classify diseases such as false smut, bacterial leaf blight, brown leaf spot, leaf scald, rice blast, rice tungro, sheath blight, and stem rot also we embed the rice crop major disease, the disease caused by, symptoms of diseases, factors that trigger the disease and in what stage the disease affect the crops. Here we register the common issues in generalization of models, crop disease diagnosis process stages. We explore advanced ML techniques as well as conventional ML methods like random forest (RF), naive bayes (NB), decision tree (DT), k-nearest neighbors (KNN) and support vector machine (SVM). A tabular summary of the results of these methods is provided in this paper. This article also discusses about the challenges and limitations associated with applying ML methods to detect plant disease, offering insights into the constraints practitioners face. Our goal is to provide a thorough overview of the latest developments in machine learning algorithms for crop disease diagnostics by integrating existing research. Ultimately, this study enhances understanding of the diseases, methodologies, challenges, and future directions for using ML in sustainable agriculture.

Keywords: Crop; Disease; Deep learning; Transfer learning; Machine learning; Plant

Cite this article: Subramanian V, Chowdhary CL. A Review on Machine Learning Approaches for Plant Disease Diagnosis. Majlesi J. Electr. Eng. 2026;20(2): 229-249. https://doi.org/10.57647/j.mjee.2025.17383

1. Introduction

Agriculture is essential to our survival and one of the most vital sectors of the world. However, crops are often affected with diseases and damage by animals [1]. These diseases can be influenced by factors such as climate, weeds, lack of soil minerals, and the planting of the wrong crops in the season. Diseases can appear in the leaves, stems, flowers, vegetables and fruits of plants, with leaves usually being the first to show signs [2]. Generally, identifying plant diseases involves examining the upper side of the leaves because symptoms can differ on each side of the plant. Recently, machine learning has become a key tool for identifying and classifying plant diseases. Machine learning algorithms can analyze large sets of plant images to accurately recognize and categorize diseases [3]. The field of artificial intelligence (AI) attempts to replicate intelligent human behavior using machines. A machine learning system creates prediction models based on historical data and predicts outcomes when new data are introduced. In addition machine learning and deep learning enhance the accuracy and speed of result recognition [4]. This review explores the current use of machine learning and pre-trained models in identifying and classifying plant diseases, discard discusses benefits, drawbacks and future possibilities.

1.1 Various crop diseases

Numerous diseases can affect rice crops, including symptoms, yield effects, and favourable environmental conditions [5]. Comprehending these diseases is essential for efficient handling and enhanced agricultural yield. The diseases, types, and factors affecting crops are depicted in Fig. 1. The two types of these causes are biotic and abiotic. Living things such as fungi, bacteria, and viruses are considered biotic factors. Air, light, soil and water are examples of substances that are not living and are considered abiotic.

Table 1 provides the essential information such as disease caused by, symptoms, yield loss, favorable environment and the crop growth stage of rice crop diseases [6, 7, 8].

1.2 Associated challenges in crop disease diagnosis

Various significant challenges are associated with the effective implementation of crop disease diagnostics:

- Limited awareness and training among farmers hinder advanced technique adoption, impacting disease management [9].
- The high cost of advanced diagnostic technologies creates financial barriers for small-scale farmers and communities [10].
- In remote agricultural regions uncommon diagnostic facilities delay disease identification [11].
- Uneven regulations for crop disease diagnostics lead to inconsistencies, hindering unified approaches [12].

1.3 Diagnosis of crop disease

The benefits of using machine learning algorithms for plant disease identification and classification are manifold. These algorithms have the capability to quickly and exactly process huge datasets. They can distinguish subtle patterns in images that might avoid human perception. Furthermore, these algorithms can be trained to differentiate between diverse disease types, thereby facilitating more precise and comprehensive categorization. By connecting the competences of machine learning algorithms, researchers can accelerate and enhance disease identification, leading to more efficient disease management and control [13]. In future, machine learning algorithms may enable real-time disease detection, there by aiding in more proactive disease management strategies. Crop diseases pose a serious threat to agricultural output and global food security [14]. The increasing global populations, imperative for accurate and efficient crop disease diagnosis have never been more pronounced. Left unchecked, these diseases can lead to substantial yield losses, impacting not only farmers, livelihoods but also the stability of worldwide food supplies [15]. In this context, the integration of machine learning techniques for crop disease diagnosis has immense potential to revolutionize the landscape of disease detection, monitoring, and mitigation [16]. This survey study aims to comprehensively outline techniques for crop disease diagnosis using machine learning. By integrating existing research, our goal is to provide an overview of the advancements in this domain and their potential implications for sustainable agriculture. Through this survey, readers will gain insights into the numerous approaches, methodologies, and challenges researchers and practitioners handle when applying machine learning to crop disease diagnosis. Figure 2 depicts the different stages of crop or plant disease identification and classification.

These stages encompass pre-processing, segmentation, feature extraction, model selection, and the subsequent training and testing of the model. Image pre-processing

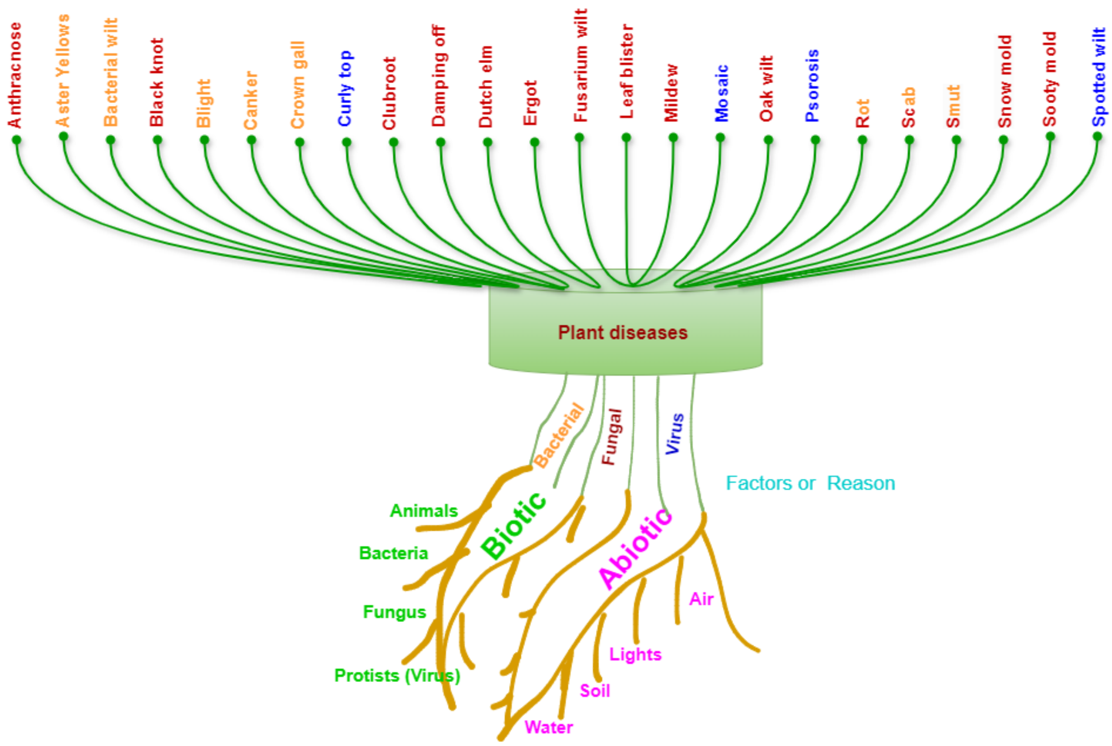

Figure 1. Different crop diseases and classification.

https://doi.org/10.57647/j.mjee.2025.17383

Table 1. Various rice crop diseases.

	Rice Tungro	
Caused by	Rice Tungro bacilliform virus and Rice Tungro spherical	
Symptoms	Stunted growth and yellow orange leaves, Reduced tillering, Delayed flowering, Incomplete grain filling	
Yield loss	<= 70%	
Factors	Green leafhopper, warm and humidity conditions, High density of infected plants nearby	
Stage	Seedling, Vegetative, Reproductive	
	Rice blast	
Caused by	Magnapothe oryzae	
Symptoms	Diamond shaped lesions on leaves, Black discoloration an nodes, Lesions on panicle neck	
Yield loss	<= 50%	
Factors	High humidity (> 90%), Frequent rainfall, Dense planting, nitrogen over - fertilization.	
Stage	Seedling, vegetative, reproductive stages	
	Sheath blight	
Caused by	Rhizoctonia solani	
Symptoms	Irregular, greenish-gray.	
Yield loss	5-50%.	
Factors	High humidity, high temperature (25-30°C), Dense canopy, High nitrogen levels, Flooded conations.	
Stage	Tillering, Reproductive.	
	Bacterial leaf blight	
Caused by	Xanthomonas oryzae pv.oryzae.	
Symptoms	Water soaked to yellowish stripes, Leaves turn grayish green, roll up and wilt	
Yield loss	20-30%.	
Factors	Warm temperatures (25-30°C) High humidity, Rainfall or heavy dew, Mechanical damage.	
Stage	Seedling, Vegetative, Reproductive	

https://doi.org/10.57647/j.mjee.2025.17383

Continued of Table 1.

	Brown spot	
Caused by	Cochliobolus miyabeanus	
Symptoms	Small circular brown lesions on leaves, Lesions have dark brown margin and gray light brown center.	
Yield loss	10-50%.	
Factors	Poorly drained soils, Nutrient deficiencies, and Prolonged wet conditions.	
Stage	Vegetative, Reproductive	
	Rice False smut	
Caused by	Ustilaginoidea virens	
Symptoms	Greenish spore balls replacing grains, Balls turn yellowish orange and then greenish black,	
Yield loss	1%-3% and upto 30% based on severity	
Factors	High humidity, high nitrogen application flowering stage of rice	
Stage	Reproductive stage	
	Stem rot	
Caused by	Sclerotium oryzae	
Symptoms	Black lesions at waterline on stems, Internal rotting of stem base, Lodging of plants	
Yield loss	20-30%	
Factors	High humidity, Warm temperature (20-30°C), Poorly drained soils.	
Stage	Tillering to maturity	
	Grain discoloration	
Caused by	Bipolaris oryzae, Curvularia spp., Fusarium spp.	
Symptoms	Discolored grains on husk, Reduced grain quality	
Yield loss	up-and-down	
Factors	High humidity during grain filling and maturation, warm conditions	
Stage	Reproductive stage	

https://doi.org/10.57647/j.mjee.2025.17383

Continued of Table 1.

Leaf scald	
Caused by	Microdochium oryzae
Symptoms	Light brown to grayish lesions on leaves, Lesions can coalesce causing leaf blight.
Yield loss	05-15%
Factors	High humidity, Frequent rains or heavy dew, Dense planting.
Stage	Tillering to booting stages

Narrow Brown leaf spot	
Caused by	Cercospora janseana fungus
Symptoms	Narrow elongated brown lesions on leaves, Premature leaf senescence
Yield loss	05-20%
Factors	High humidity, Wet conditions, Poor crop rotation practice.
Stage	Vegetative to reproductive stages

Bakanae disease	
Caused by	Fusarium fujikuroi
Symptoms	Abnormal elongation of seedlings, Taller pale green and weak plants, plant death
Yield loss	<= 20%
Factors	Infected seeds, warm temperature (25-30°C)
Stage	Seedling stage

entails several steps to enhance an image, rendering it suitable for computer vision tasks. These techniques not only enhance the accuracy but also moderate computational time and costs. Common pre-processing techniques encompass resizing, contrast enhancement, brightness adjustment, noise removal, and color correction [17]. The segmentation process partitions images into smaller units. The objective of this process is to extract pertinent features from segmented images, thereby reducing the execution time. Feature extraction is pivotal for isolating key characteristics for distinguishing healthy from unhealthy images [18]. The model selection phase must select suitable models from various machine learning techniques for plant disease identification and classification [19]. Crop diseases have long been a formidable adversary to agricultural productivity, posing substantial challenges to global food security. These diseases caused by a variety of factors, including fungi, bacteria, viruses, and environmental stressors, result in reduced crop yield, diminished quality, and increased susceptibility to other stresses [20]. The algorithms and models that learn from data and machine learning offer the potential to revolutionize crop disease diagnosis [21]. A machine learning model for crop disease identification and classification is mentioned in Fig. 3.

The economic implications of unchecked crop diseases are vast, affecting not only farmers' livelihoods but also market prices and food availability worldwide. The traditional methods of crop disease diagnosis, which rely on visual inspection and manual assessment by experts, are time consuming, labor-intensive and often prone to subjectivity. In addition, these methods may fail to

https://doi.org/10.57647/j.mjee.2025.17383

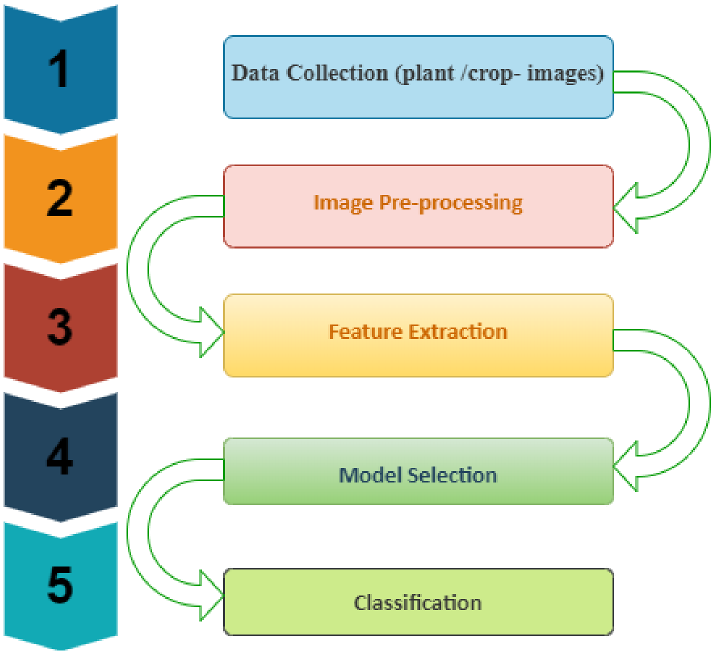

Figure 2. Common steps involved in plant disease identification and classification.

accurately detect diseases in their early stages, leading to delayed intervention and exacerbated losses. As climate change and globalization continue to impact agriculture, the need for efficient, accurate, and scalable methods of disease detection and management becomes more pressing. Machine learning has emerged as a powerful tool in addressing the challenges posed by crop diseases, with its capacity to process large volumes of data and identify complex patterns.

1.4 Research contributions

The main contributions of this paper are as follows:

- The research surveys covers rice crop diseases and describe their details.
- Discusses machine learning techniques for crop disease identification and classification, along with their respective advantages and disadvantages.
- The work explains the segmentation, feature extraction and selection, classification techniques.

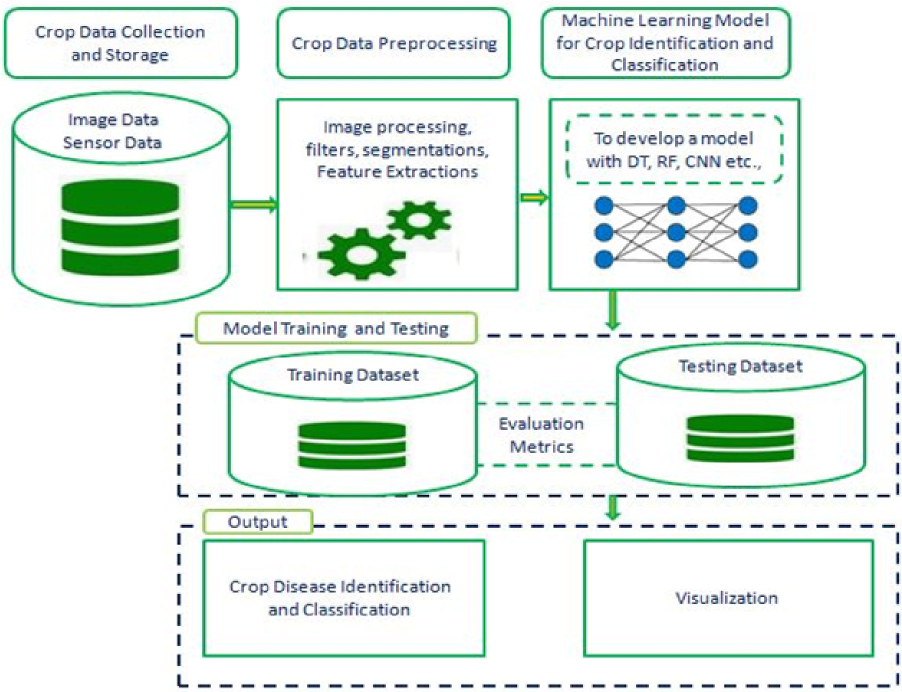

Figure 3. ML model for crop disease identification and classification.

https://doi.org/10.57647/j.mjee.2025.17383

- Presents different approaches for crop disease diagnosis using machine learning present era.
- The proposed research survey also outlines challenges, insights and future research directions for crop disease diagnosis using machine learning.

1.5 Organization of the work

The organization of this paper is as follows: The introduction is presented in section 1 with the motivation and contribution. The research methodology of the survey paper along with the subsection background, literature survey, the dataset, crop diseases overview and comparative analysis of ML techniques are presented in sections 2. The results discussion, challenges and case studies are presented in section 3. Finally conclusions presented in section 4.

2. Research methodology

The information presented in this survey paper, which pertains to the integration of ML into the realm of plant and crop diseases, was acquired by conducting searches on notable platforms such as Science Direct, Springer, IEEE, MDPI, Wiley and Research Gate. During the process of researching pertinent papers, keywords like "machine learning", "crop diseases" or "plant diseases", "deep learning", and "transfer learning", were employed in various combinations. We focused on the articles published from the period of 2015 to 2024. Figure 4 exhibits the plant disease-related manuscript published in region-based statics, the contents extracted from Clarivate's official website formerly the Web of Science. Figure 5 represent the number of articles cited for this article. Figure 6 shows the number of articles published about crop disease. Figure 7 Describes the article's concern with crop disease and machine learning published by the various publishers.

2.1 Background

Agriculture is a major source of raw food and the production of products. The production of agricultural raw products is shortened by disease, which leads to economic crises for the farming country. The farmers generally spot the issues at the end of the diseases covering the cultivated crop or plant area that reduces the profit of the farmer, the early detection will improve the crop yield. Human disease identification is difficult when the cultivation area is large and the person is not skilled, to solve the problem. Image processing and machine learning techniques are integrated into detecting and classifying plant/crop diseases [22].

2.2 Literature review

Machine Learning Techniques

This review paper explores numerous machine learning methods for detecting and classifying plant or crop diseases, including upgraded versions of machine learning models that achieve improved results with less complexity. Author in [23] experimented with a prototype for the detection and classification of rice plants or crops,

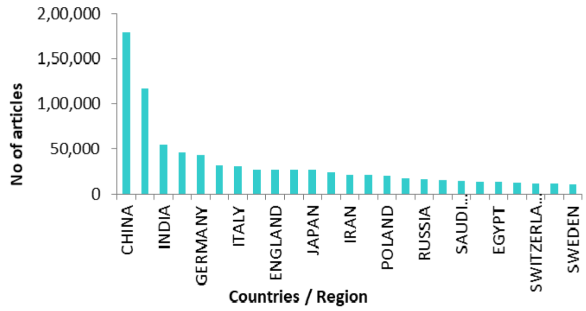

Figure 4. Number of articles published based on countries/region (Courtesy Clarivate).

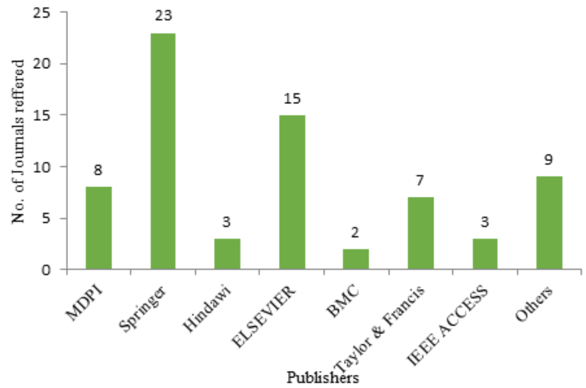

Figure 5. Number of articles referred concern with plant disease.

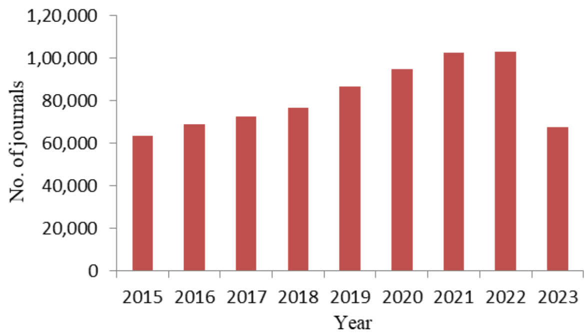

Figure 6. Number of articles published year-wise related with crops disease (Courtesy Clarivate).

and they concentrated on bacterial leaf blight, brown spot, and leaf smut diseases, the images acquired directly from the field. The k-means clustering method was applied to extract the accurate features by green pixels removed at the disease section. Author in [24] developed a novel algorithm for features extracted under the color, shape, and texture categories. The segmentation task was applied in the affected area of the apple leaf image. Before segmentation, the leaf images are enhanced by brightness preserving the dynamic fuzzy histogram equalization method. Author in [25] conducted a comparison of diverse feature extraction techniques using various classifiers and an extensive review of prevalent classification techniques, encompassing support vector machine, k-nearest neighbors, naive bayes, and decision tree. Their analysis indicated that support vector machine and naïve bayes require a larger sample size

https://doi.org/10.57647/j.mjee.2025.17383

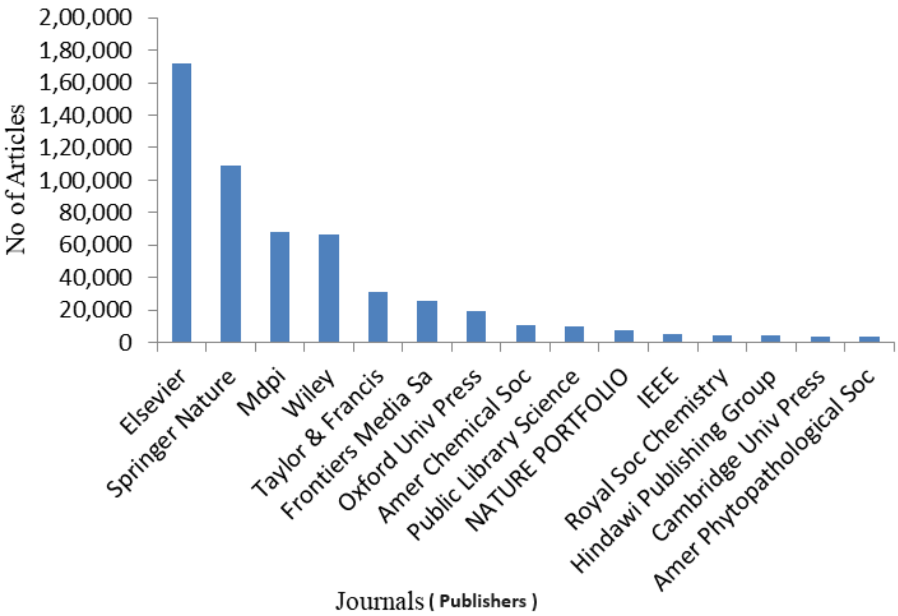

Figure 7. Number of articles published based on publishers (Journals)(Courtesy Clarivate).

to achieve maximum predictive accuracy. The support vector machine classifier achieved an accuracy of 89.8% with a smaller dataset (242 instances), to identify affected crops via support sensor development required additional work, the crop investigation in the field is controlled by spectral data sample. Author in [26] categorized and identified diseases in citrus fruit leaf images. Gaussian function enhanced the images, weighted segmentation was used for segmentation, and a saliency map was employed for feature extraction using skewness, Principal Component Analysis (PCA), and Entropy techniques. The Fig. 8 shows the road map of machine learning and basic concepts along with founder of the respective algorithms.

Author in [27] analyzed various machine learning algorithms and compared their performances with online datasets; RF stands out as the best performer, followed by support vector machine and decision tree. Author in [40] discussed various segmentation techniques, and feature extraction techniques with pros and cons of the same. Author in [41] provided insights into machine learning algorithms like random forest, support vector machine, decision tree, k-nearest neighbors, and naive bayes for plant illness classification. Random forest exhibited the highest accuracy when compared to other algorithms. They gathered real-time tomato leaf images, addressing noise through RGB to grayscale conversion and image size reduction. Image segmentation was executed using label edge detection and features included color, texture, and shape attributes. The classifier was built on these features to identify tomato plant diseases. Table 2 shows the various segmentation techniques and performance

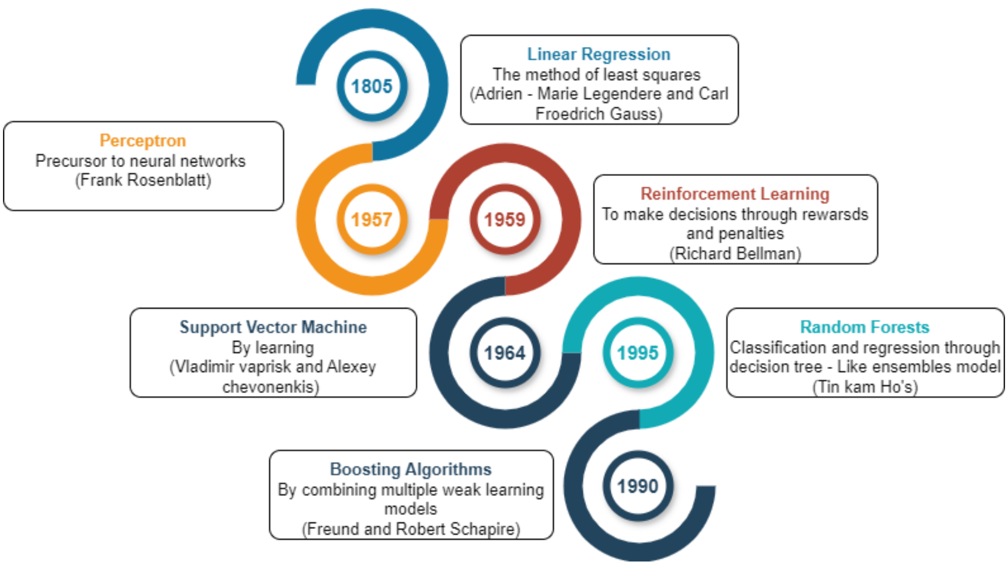

Figure 8. History of machine learning algorithms.

https://doi.org/10.57647/j.mjee.2025.17383

Table 2. Details of segmentation technique with ML models used by various researchers.

Reference	Segmentation	Dataset	ML model	Accuracy in %
[28]	k-means clustering	IRRI-International Rice Research Institute	KNN	97.5
			ANN	94.4
			NB	85
			SVM	98.63
[29],[30], [31]	GLCM TRGS Extended ROI	Real time Bahawalpur Agriculture Research Lab Pakistan Plant village	ELM	90
			RF	92.78
			BN	94.33
			LB	94.44
			MLP	94.67
			KNN	96
			NB	89
			SVM	99
[32]	Image Mask	Real time	LR	87.33
			LDA	80.7
			NB	86.74
			KNN	85.46
			DT	78.81
			SVM	89.73
			RF	93.38
[33]	C-Means clustering	Real time, Pakistan	RFC	99.8
			DT	99.2
			KNN	99.0
			NB	97.7
			SVM	94.4
			LR	89.6
[34]	Laplacian filter	Real time	KNN	97
[35]	k-means	Real time	RBF-SVM	92
			SVM	90
			RF	89
			ID3	82
[36]	k-means	CCDDB, PDDB	GA-SVM	91.3
[37]	Color based-thresholding, Edge detection	Real time	KNN	93
			SVM	87.40
			RF	90
			DT	90
[38]	Background subtracts.	Real time	IKNN	97.01
[39]	Superpixel fuzzy C–mean clustering	Plant village	HRF-MCSVM	98.9

accuracy of the respective classification algorithms.

Author in [42] identified diseases in cashew leaves using machine learning techniques. Their approach included image collection, pre-processing, image segmentation, feature extraction, and classification. Images were sourced from benchmark datasets like cashew crop disease database (CDDB) and plant disease database (PDDB), with pre-processing operations including rotation and flipping for reduced execution time. The segmentation was performed using k-means, and various machine learning algorithms were used for classification. Among them, the random forest classifier with k-means segmentation showed exceptional performance with an accuracy of 99.7%.

Table 3 shows the various feature extraction techniques and performance accuracy of the respective datasets and classifications algorithm.

Author in [45] developed a model for detecting and identifying crop diseases using a hybrid technique that combines the Firefly Algorithm with Support Vector Machines (SVM). The model utilized shape and color attributes for feature extraction and employed k-means clustering for image segmentation. This approach achieved a notable accuracy.

We reviewed approximately 75 research and survey manuscripts from this literature study, and we found that machine learning models perform well on limited or small dataset based on the literature study. The findings and challenges were explained in respective sections.

2.3 Datasets

The literature survey exhibits the various datasets such as their own dataset and benchmark dataset for their research work. A few datasets are listed below:

https://doi.org/10.57647/j.mjee.2025.17383

- International Rice Research Institute (IRRI) contains a rice research data repository which includes various diseased rice datasets for rice research [8].

- Rice diseases image dataset contains bacterial leaf blight, brown spot, and leaf blast affected image data sets of rice plants, Plant village dataset covers maximum rice crop disease images [46].

- UCI Machine learning repository has brow spot, leaf smut, and bacterial leaf blight affected image dataset [47].

- Mendeley has leaf smut, bacterial leaf blight, and brown spot rice leaf diseased dataset [48].

2.4 Crop disease diagnosis process

The crop diagnosis processes consist of image collection, pre-processing, segmentation, feature extraction, dataset splitting (training (70% or 80%), validation (10% or 15%) and testing (10% or 15%), model selection, classification and performance evaluation of the respective model, each process explained with few lines [49].

2.5 Data collection or acquisition

The first step in plant disease identification and classification using machine learning is collecting the data. These data can be collected from various sources such as agricultural research centers, online databases, and captured images from the field. The data should include information about the plant species, the symptoms of the disease, and the environmental conditions of the affected

Table 3. Details of feature extraction technique with ML models used by various researchers.

Reference	Feature extraction techniques	Dataset	ML model	Accuracy in %
[43]	GLCM	Own dataset-Real	SVM	98.97
[28]	KNN BPNN NB SVM&Hybrid DWT, SIFT GLCM	IRRI Indian Rice Research Institute	KNN ANN NB SVM	97.5 94.4 85 98.63
[29]	GLCM	Real-time	ELM	90
[30]	CVIP-GLCM	Bahawalpur Agriculture Research Lab Pakistan	RF BN LB MLP	92.78 94.33 94.44 94.67
[31]	Color coherence vector	Plant Village	KNN NB SVM	96 89 99
[32]	Harlick descriptors	Real-time dataset Collected From filed	LR LDA NB KNN DT SVM RF	87.33 80.7 86.74 85.46 78.81 89.73 93.38
[33]	Haralick Texture (HT) Hue moments (HM), Color Histogram (CH)	Real-time Pakistan	RFC DT KNN NB SVM LR	99.8 99.2 99.0 97.7 94.4 89.6
[34]	Histogram, Texture, RST, Binary Spectral	Real-time	KNN	97
[44]	Feature Color, Shape, Text.	Plant village	RF	89
[35]	Shape and Color	Own & Plant village	RF	99.7
[36]	Color, Shape	CCDDB, PDDB	GA-SVM	91.3
[37]	Texture based Color based Shape Local binary & wavelet	Real time	KNN SVM RF DT	93 87.40 90 90
[38]	GCR [GIDOR + GPOR]	Real time	IKNN	97.01

https://doi.org/10.57647/j.mjee.2025.17383

plants.

Image Pre-processing

Once the data are collected, the next step is to preprocess the data. This involves cleaning the data, removing any outliers, Resizing and Cropping, Color Spaces, Normalization, Image augmentation, noise removal, gray scaling, histogram equalization, smoothing, sharpening, and color correction at the end of that section. Feature engineering is performed to extract the most relevant features from the data. This was performed to ensure that the data were suitable for machine learning and that the processing time and cost were reduced. Author in [50] states that the crop disease diagnosis process; has different stages for the successful identification and classification of diseases, such as data collection or data acquisition, pre-processing, segmentation, feature extraction, and finally identifying and classification of the diseases. Authors in [51], [52] discussed different image pre-processing techniques such as auto-orientation, object isolation, resizing, gray scale static crop, contrast adjustment, and tiling. The model selection is summarized using a simple step-by-step procedure with flow chart Fig. 9.

Step 1: Model selection Start Step
Steo 2: Check datasets quantity
  If (Samples > 10, 000) then
Step 3: Determine the resource availability
    If (high end resources like GPU) then
Step 4 : Are you expect more accuracy
      If(Accuracy need high)
        Select DL/TL
      else
        Select ML (RD,NB,DT,KNN)
    else
      Select ML (RD,NB,DT,KNN)/TL
  else
    Select ML (RD,NB,DT,KNN)/TL
Step 5: Complete model selection

Image segmentation

Segmentation techniques are used to partition an image into multiple segments, each containing different types

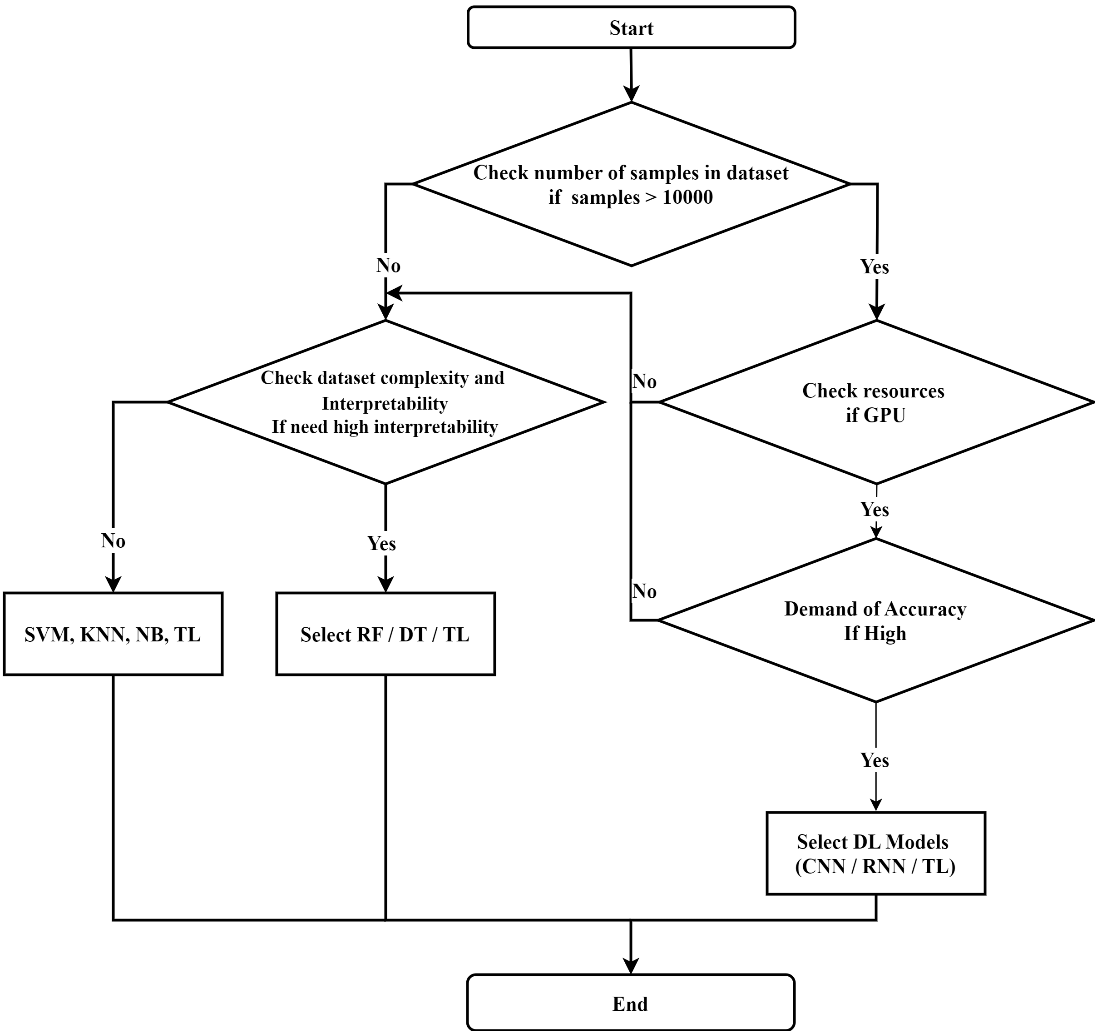

Figure 9. History of machine learning algorithms.

https://doi.org/10.57647/j.mjee.2025.17383

of pixels or regions. Common segmentation techniques used in image segmentation include support vector machine, decision tree, k-means clustering, threshold-based, and region growing, color clustering, edge detection, and graph-based methods. Each technique has its advantages and drawbacks and can be used in different applications. Author in [53] explained the (1) for threshold-based imaged segmentation. Figure 10 shows the crop or plant disease diagnosis model with various stages and techniques involved.

$$segimg(x, y) = \begin{cases} obji1 & \text{if } thobj1 < f(x, y) < thobj2 \\ obji2 & \text{if } f(x, y) > thobj2 \\ bimgi & \text{if } f(x, y) \le thobj1. \end{cases} \tag{1}$$

where:
$f(x, y)$: Input image
$segimg(x, y)$: Segmented image
$obji1$: Intensity of object 1
$obji2$: Intensity of object 2
$bimgi$: Intensity of background
$thobjx$: Threshold values of respective objects
Feature selection
Feature selection is an important step in plant disease identification and classification using machine learning. The proposed method can improve the accuracy of the model and reduce its complexity of the model. Different best methods are there to extract the features such as Histogram of Oriented Gradients (HoG), Scale-Invariant Feature Transform (SIFT), Local Binary Patterns (LBP), Speeded-Up Robust Features (SURF), Bag of Words (BoW), Color Histograms, Hu Moments, Wavelet transform, Gabor filters. Reducing the number of resources needed to accurately represent a big collection of data is the goal of feature extraction. Features like edges, textures, and forms are frequently extracted in image processing.

Authors in [54, 55] applied the (2) to facilitate the identification of object edges in a picture, which helps determine the limits of the affected region.

$$Gmx = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * f(x, y),$$

$$Gmy = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * f(x, y). \tag{2}$$

where:
Gmx: Gradient in the x direction
Gmy: Gradient in the y direction
$f(x, y)$ is the input image and '$*$' represent convolution operation
The computation of the gradient magnitude of the input image is mentioned in (3).

$$Gm = \sqrt{(Gmx)^2 + (Gmy)^2} \tag{3}$$

The simple color feature extraction from an image is demonstrated in (4).

$$F_v = [f_{p1}, f_{p2}, f_{p3}, \ldots, f_{pn}] \tag{4}$$

Fv: Feature vector
f_{px}: Pixel value frequency in a given input
To categorize or forecast crop diseases, machine learning algorithms use these feature vectors as input. Upon feature extraction, machine learning models use the features as vectors for additional analysis.
Dataset splitting -Training and testing
The machine learning model classifications are contingent on how well they simplify from the training and testing data (unseen), so the machine learning models need a structured training and testing processes to show their reliability in real world applications. Testing data evaluates how effectively a machine learning model has learnt, training data instructs the model on how to behave.

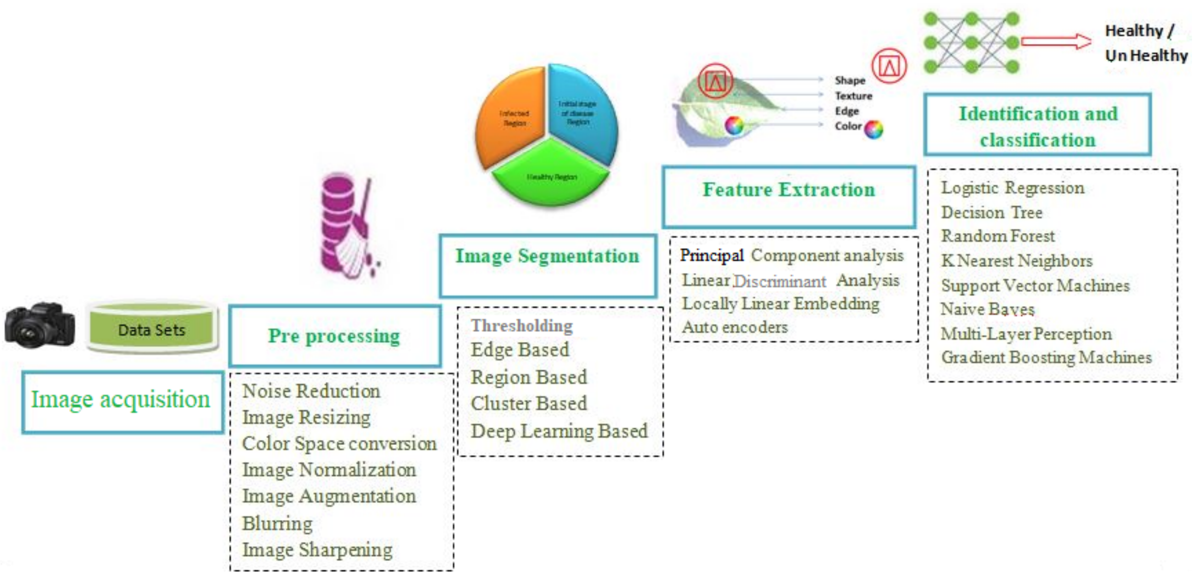

Figure 10. Crop disease diagnose process various stages and respective techniques.

Machine learning divides the dataset into three categories such as training, testing and validation, validation data is an optional one the same represented by Fig. 11. The training data is bigger than testing data. The dataset is divided as follows: 20% are used for validation, 60% are used for training, and 20% are used for holdout (Training set (70-80%), Validation (10-15%), Testing (10-15%)). To learn the model parameters, utilize the training subset. Accuracy improvement is monitored using the validation subset following each learning round. Once training is complete, the holdout subgroup is utilized to verify the outcomes. The amount of training samples is decreased by this division, but overfitting on the training images is prevented, greatly increasing the reliability of the results [56]. The dataset to be ensures before training that each disease class is proportionally represented in all subsets to preventing class imbalance issues.

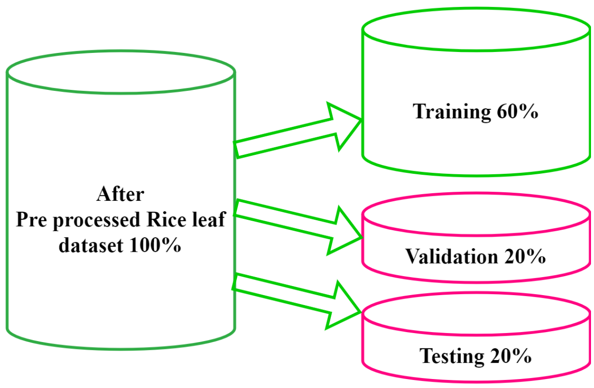

Figure 11. Crop disease diagnose process various stages and respective techniques.

A machine learning algorithm builds a decision-making model by identifying patterns in the data it receives from records.

Algorithms enable a business to make decisions based on its prior experience. The model evaluates all prior cases and their outcomes, and then uses this information to build models that assess and forecast the results of on-going cases.

Overfitting

A model performs very well on training data but poorly on the testing data says that overfitting. Here the model has leaned the training data too closely, including noise and cannot generalize to new data. The overfitting we may reduce by applying the following techniques.

Data augmentation: The images were increased by rotation, flipping, brightness adjustment and gaussian noise addition.

Regularisation: Reduce excessive reliance on specific features applying through the techniques dropout and L2 regularization.

Transfer learning: Use a pretrained model (VGG16,ResNet), when dataset is small.

Cross validation Cross validation to minimize (mitigate) overfitting and obtain a more robust evaluation a techniques called cross-validation is often used, here the data is split into k equal parts that is k folds and the model is trained and tested on each fold, with the results averaged across all folds.

Classification

Classification refers to classifying crop or plant diseases using machine learning algorithms. The following are some famous machine learning classification algorithms such as support vector machine [57], random forest [58], k-nearest Neighbors [59], decision tree [60], Linear Regression [61] and naive bayes [62].

Model performance evaluation

After completing machine learning, is completed the model is evaluated on the unseen test dataset. The machine learning model performance measured by applying the equations (5) and (6).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

where:

TP - True positives

TN - True Negatives

FP - False Positives

FN - False Negatives

F1-score – To assess the model performance while using imbalanced dataset

$$\text{F1-score} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

Confusion matrix – Shows detailed classification and identify misclassification.

3. Results and discussion

Although plant disease detection can be automated, there are still several challenges to overcome at each step or process. For instance, obtaining samples for specific diseases can be quite difficult. Moreover, the performance of the system can be greatly affected by the type of dataset used, such as laboratory-controlled or real-time datasets. Real-time data obtained in an uncontrolled environment can increase system complexity, but it is crucial for agricultural development and more suitable for current research. During the feature extraction process, some difficulties have been observed, such as the similarity in the infected area leading to the extraction of inaccurate features and false classification based on irrelevant feature matching. Therefore, it is necessary to select a set of features since each feature has different levels of importance [63]. Various classification techniques, including support vector machine, artificial neural network, naive bayes, back propagation neural networks, decision tree, and k-nearest neighbors, have been used for plant disease detection [64]. However, deep neural networks, particularly convolutional neural network, have shown more accurate results for large samples. Nevertheless, over-fitting of convolutional neural network is a significant issue in plant disease detection systems [65]. Several approaches have been proposed to enhance the efficiency of convolutional neural network based classification systems for multiple cultures. The

https://doi.org/10.57647/j.mjee.2025.17383

literature suggests that the proposed system must meet specific specifications, or it may result in inaccurate disease detection. Therefore, a flexible system with an adjustable set of specifications must be designed by researchers instead of a fixed one. Over-fitting can also affect the actual use of machine learning, so a highly generalized and adaptable system is needed for plant disease detection in various cultures. Machine learning is a popular domain due to the availability of techniques and tools, but accuracy should not be compromised in the process [66, 67, 68]. Figure 12, is the graphical representation of the distinct machine learning algorithm performance analysis in crop disease identification and classification tasks used in various articles.

Figure 13 presents a graphical representation of performance analysis of distinct segmentation techniques in crop disease identification and classification tasks applied by various researchers. Table 4 displays the performance analysis of different machine learning techniques, datasets, and accuracy concerning the various literature papers reviewed. Figure 14 shows a graphical representation of the performance analysis of distinct feature extraction techniques in crop disease identification and classification tasks applied by various researchers.

3.1 Challenges

The challenge in plant disease detection is the scarcity of experts who can accurately annotate the data. This issue arises when experts are unable to distinguish between diseased and dead plants. This process requires skilled and knowledgeable professionals to detect rare or new diseases, and it can be time-consuming and expensive. Shallow architectures are more suitable for small datasets. The latest disease detection models provide a different perspective for building a plant disease detection model. Machine learning and deep learning are believed to enhance disease detection models. Several factors and issues that may influence plant disease identification and classification are listed below. The effectiveness of the plant disease detection system is mainly determined by the feature extraction and classification methods employed. The majority of the studies reviewed utilized the Plant Village Dataset, which comprises laboratory images rather than real-time images. The performance of the classifier relies heavily on the dataset used for testing and training. Real field images often have intricate backgrounds, making it challenging to segment affected areas and affect the overall performance of the system. Plants may experience nutrient deficiencies and contamination at an early stage of growth [70]. The esti-

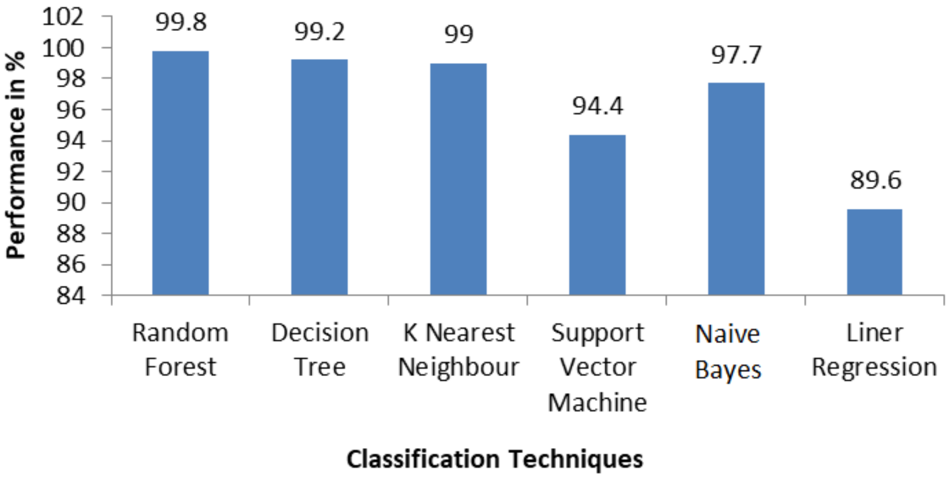

Figure 12. Machine learning algorithm performance comparison used by various researchers.

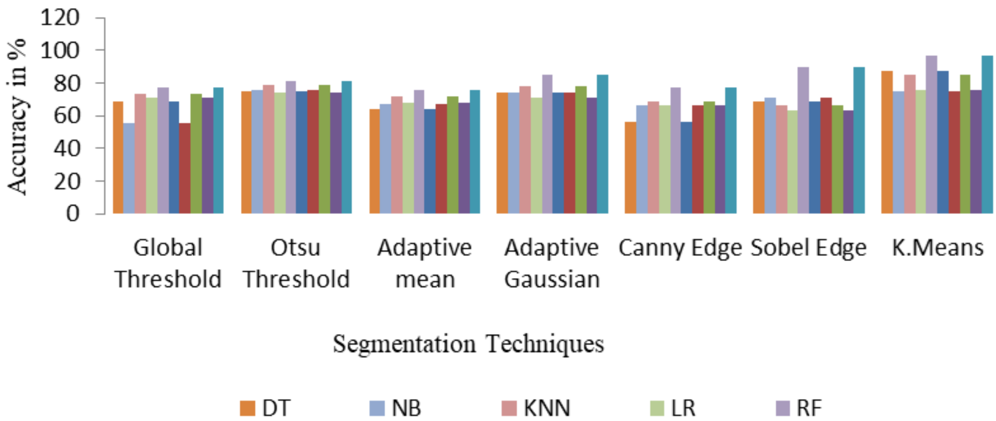

Figure 13. Segmentation techniques performance comparison used by various researchers.

https://doi.org/10.57647/j.mjee.2025.17383

Table 4. Details of ML models and datasets used by various researchers.

Reference	Dataset	Crop	No. of images	ML Techniques	Accuracy in %
[28]	IRRI (Indian Rice Research Institute)	Rice	970	KNN ANN NB SVM	97.5 94.4 85 98.63
[29]	Real time	Rice	73	ELM	90
[30]	Bahawalpur Agriculture Research Lab Pakistan	Corn seeds	6000	RF BN NB MLP	92.78 94.33 94.44 94.67
[31]	Plant village	Potato	1499	KNN NB SVM	96 89 99
[32]	Real time Collected from filed	Apple Black gram	3121 760	LR LDA NB KNN DT SVM RF	87.33 80.7 86.74 85.46 78.81 89.73 93.38
[34]	Real time	Rice	5000	KNN SVM RF DT	93 87.40 90 90
[44]	Real time Pakistan	Wheat	3150	RFC DT KNN NB SVM LR	99.8 99.2 99.0 97.7 94.4 89.6
[44]	Plant village	Common	54303	HRF-MCSVM	98.9
[35]	Real time	Tomato	1090	RF	89
[36]	Real time, rlant village	Grape	3885	SVM	98.97
[37]	CCDDB, PDDB *	Cashew	2326	RF	99.7
[38]	Plant village	Grape	4062	IKNN	98.07
[69]	RoCole (Robusta coffee leaf images dataset)	Coffee	1560	SVM	96.23

mation of infected areas and severity management with disease detection can help control pesticide use. The real-time efficiency of disease detection on constrained devices can also be considered. Simultaneous occurrence of various diseases may result in distinct symptoms, making it difficult to identify and combine them. The performance of the disease identification system can be significantly impacted by hyperparameters tuning and selection. Additionally, disease detection systems face significant challenges due to the uniformity of diseases and the selection of appropriate attributes. In addition, we listed some additional information presented here.

Data collection

Accurate identification and classification of plant diseases using machine learning hinge on the availability of vast datasets. Gather from diverse sources such as research papers, online databases, and a field experiment is essential. These datasets must undergo precise labeling and annotation to ensure the efficacy of machine learning models.

Feature extraction

The pivotal stage in machine learning for plant disease diagnosis is involves feature extraction. Extracting relevant features from the data is crucial for pattern identification and disease classification. However, this process can prove to be intricate and time-consuming.

Model selection

The array of M: Models accessible for plant disease identification and classification, which presents a challenge in model selection. The selection of the most suitable model necessitates meticulous consideration of both the data characteristics and the specific problem at hand.

https://doi.org/10.57647/j.mjee.2025.17383

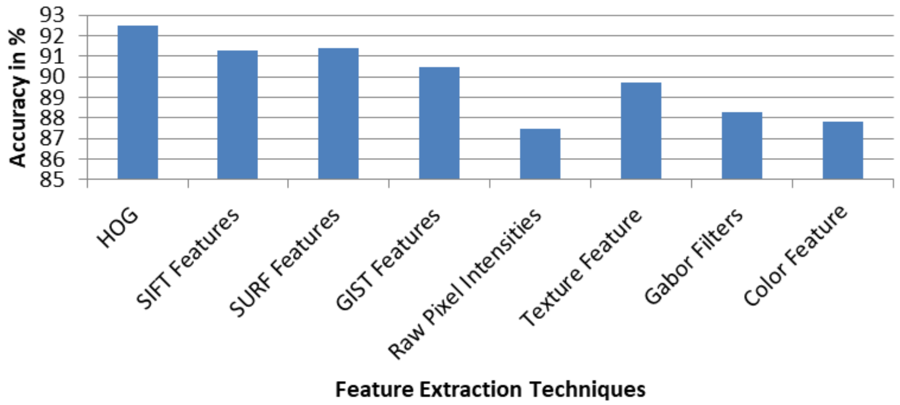

Figure 14. Feature extraction techniques performance comparison used by various researchers.

Figure 15 presents the challenges in crop or plant disease identification and classification using machine learning.

Overfitting

Overfitting remains a pervasive obstacle in machine learning. It occurs when a model becomes excessively intricate, closely aligning with the training data while faltering to generatively generalize new and unseen data. This phenomenon yields inaccuracies and subpar performance.

Generalized model

Authors in [71], [72] elaborate the hindrances in creating generalized model. Higher temperatures had no discernible effect on rice yield, but they had a negative effect on long an output and a positive effect on maize. Crop yields were unaffected by the trend in rainfall. Individual crop types had no impact on farm net income, despite diminishing trends in the production of various cultivated crops. Geographical location and cumulative productivity were key factors in farm revenue.

par Crop yields are still being accelerated hampered by elements including global warming, recurrent droughts, shifting atmospheric carbon dioxide (CO_2) concentrations, weather disturbances, and other climate-related factors, such as temperature, water availability, and weather patterns, are the main determinants of agricultural activity and crop yield. Given how heavily agriculture depends on climate variables, modifications or disturbances to these variables could have a significant impact on agricultural productivity and yield.

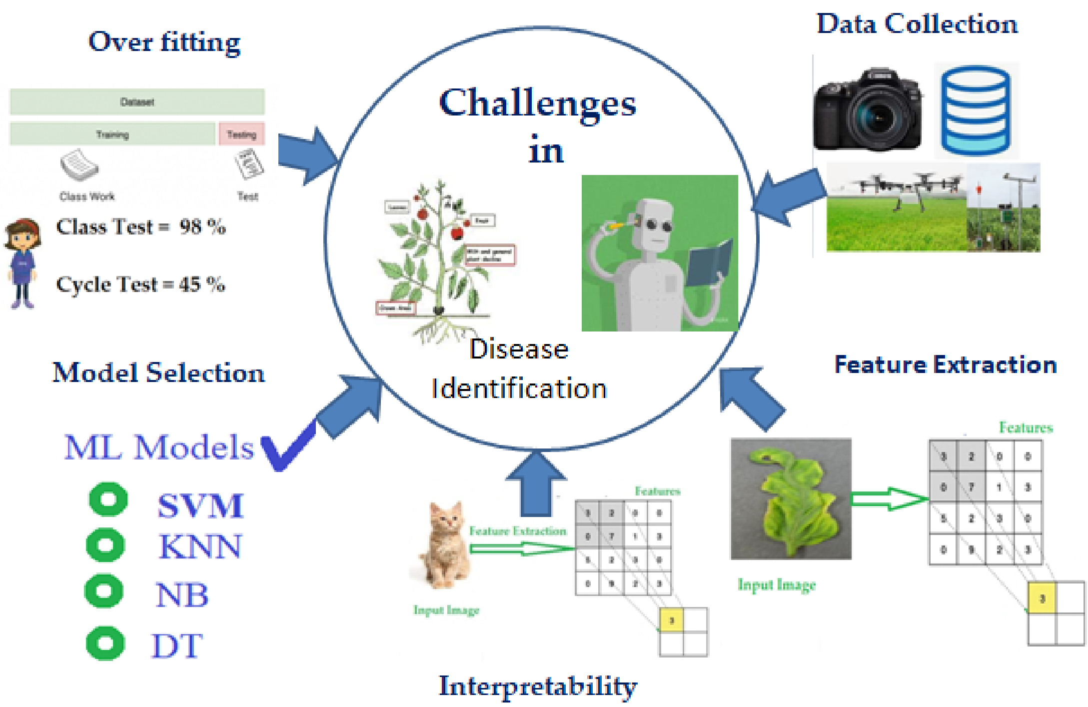

Figure 15. Crop disease identification and classification challenges.

https://doi.org/10.57647/j.mjee.2025.17383

Rice disorders show the importance of cyclical and geographical changes due to changes in climatic conditions, agricultural practices and soil health. Factors such as temperature, humidity, rainfall patterns and pest infestations influence the prevalence and severity of diseases, making it challenging to develop a single generalized model applicable to all regions.
Issues

- Season and climatic variations: The occurrence of rice disease fluctuates across different seasons. For example brown spots are more prevalent in warm and humid conditions, whereas Leaf blast inclines to occur during periods of high moisture and moderate temperatures. This seasonal dependency impact disease manifestation, making it difficult for a fixed machine learning or deep learning models to perform consistency across all seasons.

- Geographic variability and pathogen evolution: Rice-growing regions across different countries or even within the same country exhibit significant variations in soil properties, irrigation methods and pest interactions. Additionally pathogens responsible for diseases can evolve leading to variations in symptoms and severity. For instance research has shown that the same disease strain may exhibit different visual characteristics in tropical versus temperate climates, affecting model generalizability.

- Dataset limitations and bias: Most machine learning models are trained on specific datasets, which may bias toward particular geographical locations or seasonal conditions. A model trained on images from one region may not perform well in another region due to differences disease symptoms and environments conditions. Domain adaption techniques or transfer learning can help address these limitations by fine tuning models with local datasets.

Suggestions

- Region specific model training: Multiple models for different climatic zones rather than relying on a single global model.

- Integrating diverse datasets: Using a dataset that includes images collected from various geographical locations and different seasons can enhance model stoutness.

- Adaptive learning models: Meta learning of few-shot learning approaches is implemented to allow models to quickly adapt to new environments with minimal retraining.

- Hybrid approaches: Combining image based deep learning with sensor based environmental data (temperature, humidity) to improve prediction accuracy based on contextual conditions.

Interpretability
Interpreting machine learning models is often difficult thus, they are rendered less transparent and comprehensible. This lack of interpretability hinders the explanation of results and the detection of potential issues or avenues for enhancement.

3.2 Case studies

Machine learning plays a significant role in agriculture, which is often referred to as smart farming. This section elaborates on two case studies focusing on the application of machine learning in rice crop disease identification and classification.
Case study 1
Author in [73] discusses how plant or crop diseases are a major cause of crop losses, significantly impacting agricultural yields and a country's economic development. These diseases can be triggered by various factors, including environmental conditions, pests, and pathogens. To effectively manage and control plant diseases at an early stage and reduce agricultural losses, accurate identification and classification are crucial. However, manual methods are tedious. Machine learning offers a way to automate the process of identifying and classifying plant diseases. This case study shows an machine learning model that can accurately predict and classify rice plant diseases. A dataset of 1,000 images of rice plants, including both healthy and diseased samples, was collected from various sectors of Kashmore, Pakistan. The process began with image acquisition, followed by various image pre-processing techniques like cropping, enhancement, and resizing. Data augmentation was then applied to avoid overfitting. The dataset was segmented into augmented and un-augmented forms before being fed into the transfer learning section. The convolutional neural network model begins with convolution and pooling, where images are extracted into features and autonomously verified. The output is then forwarded to a fully connected neural network, which performs the classification. This case study demonstrates that machine learning can be effectively used to identify and classify plant diseases. The model's performance was evaluated using metrics such as accuracy, precision, recall, and F1 score. The performance analysis revealed that the validation accuracy of the trained model was 64% with the un-augmented dataset and 71.28% with the augmented dataset. After applying principal component analysis with 10-fold cross-validation, the random forest model achieved 73.12% accuracy on the testing dataset, which, while improved, is still modest compared to transfer learning. The model will also be tested on unseen data to ensure its accuracy in identifying and classifying plant diseases. Figure 16 shows the transfer learning model used to identify fungal blast disease in rice plants.
Case study 2
Authors in [74], [75] noted that only 21% of farmers in agriculture have sufficient knowledge to identify crop diseases, while the majority are unaware of how to identify diseases and pests, particularly biotic stress in crops.

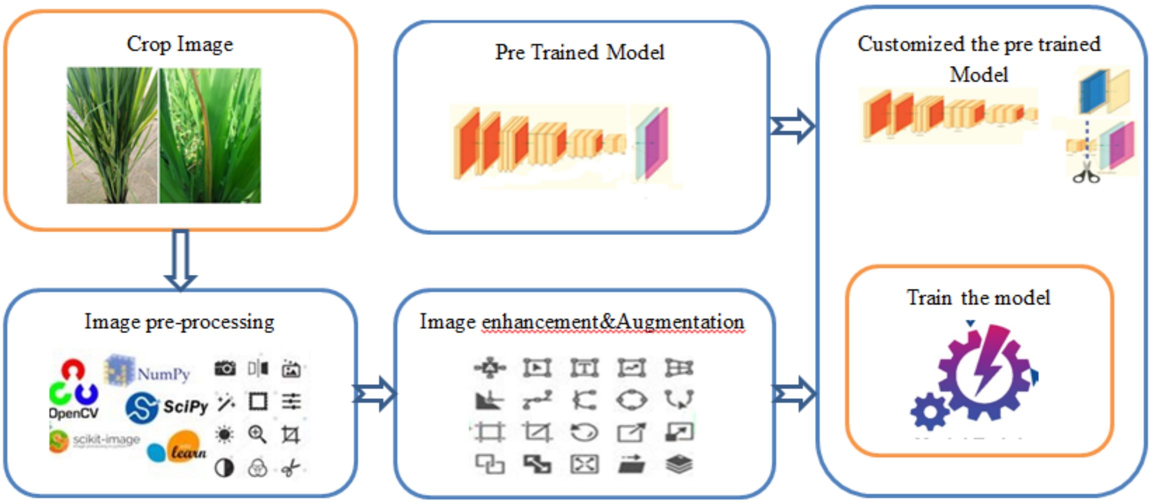

Figure 16. Transfer Learning model for crop disease classification.

Manual crop disease identification on large farms is time consuming and costly. The dataset for this study was supplied by Krishibhavan-Tarur and Krishibhavan-Thenkurussy in Kerala, India, during fieldwork and was downloaded from the online crop pest surveillance system managed by the Government of Kerala. The dataset comprises 2,000 images of both infected and healthy rice leaves. To predict early stage rice crop diseases such as brown leaf spot, leaf blast, false smut, and bacterial blight, the authors designed a model integrating machine learning algorithms with a convolutional neural network (CNN) [76]. The processes include image collection, data augmentation, labeling, image pre-processing, segmentation, feature extraction, and classification. Data was collected using web scraping tools, cameras, and sensor-equipped devices, followed by data augmentation, which increased the number and variety of images through techniques like scaling, rotating, and skewing. This step helps prevent overfitting and improves model performance. Image pre-processing removed noise, while segmentation was carried out using an image segmentation algorithm. Important features such as color, shape, and texture were extracted, followed by feature selection using Principal Component Analysis (PCA) to improve classification accuracy. The final step involved classification, where either a Support Vector Machine (SVM) or convolutional neural network could be used. The study found that convolutional neural network outperformed when compare to support vector machine in classification accuracy. Otsu segmentation proved superior to k-means, which was more costly, and Sobel edge detection was less effective due to noise. The proposed method achieved a high accuracy of 91.45% and 88.93% for the training and testing datasets, respectively, using a combination of Softmax, ReLU, and Parametric ReLU (P-ReLU) activation functions. The study suggests that model efficiency heavily depends on the selection of pre-processing techniques and machine learning algorithms.

4. Conclusion

In agriculture, plant diseases pose a major threat to crop success. Machine learning has become a crucial tool in addressing these challenges, enabling quick and accurate diagnosis and classification of crop diseases. This survey highlights the current state of machine learning in plant disease identification and classification, focusing on their benefits, limitations, and future potential. Image processing is key to accurately locating diseases. By discussing the strengths and weaknesses of various methods and providing valuable insights for researchers and industry professionals, this study advances crop disease detection and protection. The review demonstrates that machine learning algorithms are effective in identification and classification of plant diseases, suitable for moderate datasets, and both simple and cost-effective. The performance of machine learning depends largely on the quality of the dataset and pre-processing techniques. This article explores advancements in machine learning algorithms for plant disease identification and classification. However, it also points out areas for improvement, such as the focus on leaf images of disease symptoms from the top side only, neglecting the back side of the leaf, missing to focus the visible symptoms of disease apart from leaf, such as lower stems, root, limited discussion about image lighting and background impact on accuracy of crop images. Additionally, there is a need to focus on deep learning based automatic early stage plant disease identification, developing more robust datasets with divers environmental conditions to improve generalizability, integrating domain adaptation techniques to enhance transfer learning models for varying geographic and climatic conditions, exploring explainable artificial methods to increase transparency and trust in machine learning based models, Combining machine learning with internet of things for real time disease detection and monitoring smart agriculture.

Authors contributions
All authors contributed equally to the conception, design, execution, and writing of this work. All authors read and approved the final manuscript.

Availability of data and materials
The authors declare that the data supporting the findings of this study are available within the paper.

Conflict of interests
The authors assert that they do not have any identifiable conflicting financial interests or personal relationships that might be perceived to influence the work presented in this paper.

References

1. Ushadevi G et al. "**A survey on plant disease prediction using machine learning and deep learning techniques.**" Inteligencia Artificial 2020; 23:136–154. DOI: 10.4114/intartif.vol23iss65pp136-154
2. Kumar S, Sharma B, Sharma V, Sharma H, and Bansal J. "**Plant leaf disease identification using exponential spider monkey optimization.**" Sustain. Comput. 2020; 28:100283. DOI: 10.1016/j.suscom.2018.10.004
3. Dawod RG and Dobre C. "**Upper and lower leaf side detection with machine learning methods.**" Sensors 2022; 22:2696. DOI: 10.3390/s22072696
4. OpenAI. 2024; ChatGPT. Available from: https://www.openai.com/chatgpt
5. Azizi F, Malek M, and Lau HY. "**Advanced diagnostic approaches developed for the global menace of rice diseases: A review.**" Can. J. Plant Pathol. 2022; 44:627–651. DOI: 10.1080/07060661.2022.2053588
6. Archana K et al. "**A novel method to improve computational and classification performance of rice plant disease identification.** " J. Supercomput. 2022; 78:1–21. DOI: 10.1007/s11227-021-04245-x
7. Upadhyay S and Kumar A. "**A novel approach for rice plant diseases classification with deep convolutional neural network.**" Int. J. Inf. Technol. 2022; 14:185–199. DOI: 10.1007/s41870-021-00817-5
8. International Rice Research Institute. "**Rice Knowledge Bank.**" Available from: http://www.knowledgebank.irri.org
9. Adams A and Jumpah E. "**Agricultural technologies adoption and smallholder farmers' welfare.**" Cogent Econ. Finance 2021; 9:2006905. DOI: 10.1080/23322039.2021.2006905
10. Buja I. "**Advances in plant disease detection and monitoring.** " Sensors 2021; 21:2129. DOI: 10.3390/s21062129
11. Panth M, Hassler S, and Baysal-Gurel F. "**Methods for management of soilborne diseases.**" Agriculture 2020; 10:16. DOI: 10.3390/agriculture10010016
12. Jacquet F et al. "**Pesticide-free agriculture as a new paradigm.**" Agron. Sustain. Dev. 2022; 42:8. DOI: 10.1007/s13593-021-00742-8
13. Meenakshi T. "**Automatic detection of diseases in medicinal plant leaves.**" Wirel. Pers. Commun. 2023; 131:2573–2597. DOI: 10.1007/s11277-023-10555-5
14. Basaligheh P and Dhabliya R. "**Precision agriculture through deep learning.**" Res. J. Comput. Syst. Eng. 2023; 4:31–45. DOI: 10.52710/rjcse.72
15. Srinivas L et al. "**An optimized machine learning framework for crop disease detection.**" Multimed. Tools Appl. 2024; 83:1539–1558. DOI: 10.1007/s11042-023-15446-2.
16. Das S, Pattanayak S, and Behera P. "**Application of machine learning in plant disease detection.**" J. Plant Prot. Res. 2022 :122–135. DOI: 10.24425/jppr.2022.141360
17. Kurmi Y et al. "**Leaf image classification for crop disease detection.**" Multimed. Tools Appl. 2022; 81:8155–8178. DOI: 10.1007/s11042-022-11910-7
18. Barman U et al. "**Performance analysis of support vector machine for early identification of citrus diseases.**" J. Appl. Nat. Sci. 2023; 15:852–859. DOI: 10.31018/jans.v15i2.4630
19. Varshney D et al. "**Machine learning techniques for plant disease detection.**" 5th Int. Conf. Trends Electron. Informat. (ICOEI) 2021 :1574–1581. DOI: 10.1109/ICOEI51242.2021.9453053
20. Pandit MA et al. "**Major biological control strategies for plant pathogens.**" Pathogens 2022; 11:273. DOI: 10.3390/pathogens11020273
21. Anderson R, Bayer PE, and Edwards D. "**Climate change and the need for agricultural adaptation.** " Curr. Opin. Plant Biol. 2020; 56:197–202. DOI: 10.1016/j.pbi.2019.12.006
22. Brevik E et al. "**Soil and human health: Current status and future needs.**" Air Soil Water Res. 2020; 13:1178622120934441. DOI: 10.1177/1178622120934441
23. Prajapati H, Shah J, and Dabhi V. "**Detection and classification of rice plant diseases.**" Intell. Decis. Technol. 2017; 11:357–373. DOI: 10.3233/IDT-170301
24. Singh S et al. "**Extraction of multiple apple leaf diseases using machine learning.**" Int. J. Image Graph. 2022; 22:2140009. DOI: 10.1142/S021946782140009X
25. Soofi A and Awan A. "**Classification techniques in machine learning.**" J. Basic Appl. Sci. 2017; 13:459–465. DOI: 10.6000/1927-5129.2017.13.76
26. Rauf HT et al. "**Citrus fruit and leaf dataset for disease detection.**" Data Brief 2019; 26:104340. DOI: 10.1016/j.dib.2019.104340

27. Uddin S et al. “**Comparison of supervised machine learning algorithms for disease prediction.**” BMC Med. Inform. Decis. Mak. 2019; 19:281. DOI: 10.1186/s12911-019-1004-8

28. Devi TG and Neelamegam P. “**Image processing based rice plant leaf disease identification.**” Cluster Comput. 2019; 22:13415–13428. DOI: 10.1007/s10586-018-1949-x

29. Aqel D et al. “**Extreme learning machine for plant disease classification.**” Cluster Comput. 2022; 25:1–14. DOI: 10.1007/s10586-021-03397-y

30. Ali A et al. “**Machine learning approach for classification of corn seed using hybrid features.**” Int. J. Food Prop. 2020; 23:1110–1124. DOI: 10.1080/10942912.2020.1778724

31. Abdu M, Mokji M, and Sheikh U. “**Automatic vegetable disease identification using lesion features.**” Comput. Electron. Agric. 2020; 176:105660. DOI: 10.1016/j.compag.2020.105660

32. Vishnoi V, Kumar K, and Kumar B. “**Feature extraction techniques for plant leaf disease detection: A review.**” Multimed. Tools Appl. 2022; 81:1–53. DOI: 10.1007/s11042-021-11375-0

33. Khan H et al. “**Automated wheat disease classification using machine learning.** ” Agriculture

34. Aslam T et al. “**Machine learning based classification of Mangifera indica leaves.**” Int. J. Food Prop. 2022; 25:1987–1999. DOI: 10.1080/10942912.2022.2117822

35. Sudha P and Kumaran P. “**Early detection of anthracnose disease in cashew leaves.**” Signal Image Video Process. 2023; 17:1–8. DOI: 10.1007/s11760-023-02552-9

36. Kethineni K and Pradeepini G. “**Leaf disease identification using machine learning.**” J. Adv. Inf. Technol. 2023; 14:122–129. DOI: 10.12720/jait.14.1.122-129

37. Sharma M, Kumar C, and Deka A. “**Early diagnosis of rice plant disease using machine learning.**” Arch. Phytopathol. Pflanzenschutz 2022; 55:259–283. DOI: 10.1080/03235408.2021.2015866

38. Shantkumari M and Uma S. “**Machine learning for grape leaf disease detection.** ” Multimed. Tools Appl. 2023; 82:1–23. DOI: 10.1007/s11042-023-14441-x

39. Shahoveisi F et al. “**Image processing and transfer learning for rust disease detection.**” Sci. Rep. 2023; 13:5133. DOI: 10.1038/s41598-023-31942-9

40. Chowdhary CL and Acharjya DP. “**Segmentation and feature extraction: A review.**” Procedia Comput. Sci. 2020; 167:26–36. DOI: 10.1016/j.procs.2020.03.179

41. Mokhtar U et al. “**SVM-based detection of tomato leaf diseases.**” 7th IEEE Int. Conf. Intell. Syst. 2014 :641–652. DOI: 10.1007/978-3-319-11310-4_55

42. Timothy M et al. “**CNN-based detection of cashew plant diseases.**” 5th Int. Conf. Future Netw. Distrib. Syst. 2021 :225–232. DOI: 10.1145/3508072.3508107

43. Javidan S et al. “**Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning.** ” Smart Agric. Technol. 2023; 3:100081. DOI: 10.1016/j.atech.2022.100081

44. Neelakantan P. “**Analyzing machine learning algorithms for plant disease classification.**” Mater. Today Proc.

45. AlMaki W and Makrufi A. “**SVM with firefly optimization for apple disease classification.**” MATRIK 2022; 22:177–188. DOI: 10.30812/matrik.v22i1.2365

46. Huy. “**Rice diseases image dataset.**” Kaggle 2019. DOI: 10.34740/KAGGLE/DSV/796736

47. Shah H, Prajapati J, and Dabhi V. “**Rice leaf diseases dataset.**” UCI Mach. Learn. Repos. 2017. DOI: 10.24432/C5R013

48. Mensah et al. “**Crop pest and disease detection dataset.**” Data Brief 2023. DOI: 10.1016/j.dib.2023.109306

49. Kahloot M and Ekler P. “**Algorithmic splitting for dataset preparation.**” IEEE 2021; 9:125229–125237. DOI: 10.1109/ACCESS.2021.3110745

50. Bedi P and Gole P. “**Plant disease detection using convolutional autoencoders.** ” Artif. Intell. Agric. 2021; 5:90–101. DOI: 10.1016/j.aiia.2021.05.002

51. Ullah K, Jan M, and Sayyed A. “**Automatic maize disease detection using CNN.**” Int. J. Adv. Comput. Sci. Appl. 2021; 10. DOI: 10.30534/ijatcse/2021/301022021

52. Karantoumanis V et al. “**Comparison of image preprocessing techniques.**” SEEDA-CECNSM 2022 :1–5. DOI: 10.1109/SEEDA-CECNSM57760.2022.9932972

53. Gonzalez RC and Woods RE. 2009; Digital Image Processing, 3rd ed:India: Pearson

54. Fan P et al. “**Quantum image edge extraction using Sobel operator.**” Quantum Inf. Process. 2019; 18:24. DOI: 10.1007/s11128-018-2131-3

55. Ma Y, Ma H, and Chu P. “**Enhanced Sobel operator for quantum image processing.**” IEEE 2020; 8:210277–210285. DOI: 10.1109/ACCESS.2020.3038891

56. Nguyen QH et al. “**Influence of data splitting on ML performance.** ” Math. Probl. Eng. 2021; 2021:4832864. DOI: 10.1155/2021/4832864

57. Demilie WB. “**Plant disease detection techniques: A comparative study.**” J. Big Data 2024; 11:5. DOI: 10.1186/s40537-023-00863-9

58. Wójtowicz A et al. "**Random forest model for wheat and rye leaf rust.**" J. Photochem. Photobiol. B 2021; 223:112278. DOI: 10.1016/j.jphotobiol.2021.112278
59. Nalini T and Rama A. "**Impact of temperature on crop disease analysis.**" Meas. Sensors 2022; 24:100408. DOI: 10.1016/j.measen.2022.100408
60. Basavaiah J and Anthony AA. "**Tomato leaf disease classification using feature extraction.** " Wirel. Pers. Commun. 2020; 115:633–651. DOI: 10.1007/s11277-020-07590-x
61. Fenu G and Malloci FM. "**Forecasting plant disease: An explorative study.** " Big Data Cogn. Comput. 2021; 5:2. DOI: 10.3390/bdcc5010002
62. Yogeshwari and Thailambal G. "**Plant leaf disease detection using GLCM and CNN.** " Mater. Today Proc. 2023; 81:530–536. DOI: 10.1016/j.matpr.2021.03.700
63. Medjahed S. "**Feature extraction methods in image classification.** " Int. J. Image Graph. Signal Process. 2015; 7:16–23. DOI: 10.5815/ijigsp.2015.03.03
64. Kumar S et al. "**Machine learning algorithms for cotton disease detection.** " Math. Probl. Eng. 2021. DOI: 10.1155/2021/1790171
65. Veerendra G et al. "**Digital image processing techniques for plant disease detection.** " Mater. Today Proc. 2022; 51:837–841. DOI: 10.1016/j.matpr.2021.06.271
66. Rath A and Meher J. "**Computational detection of plant leaf disease.** " Arch. Phytopathol. Pflanzenschutz 2019; 19–20:1348–1358. DOI: 10.1080/03235408.2019.1708546
67. Shrivastava V and Pradhan M. "**Rice plant disease classification using color features.** " J. Plant Pathol. 2021; 103:17–26. DOI: 10.1007/s42161-020-00683-3
68. Sengar MD and Travieso C. "**Powdery mildew detection in cherry leaves.** " Computing 2018; 100:1189–1201. DOI: 10.1007/s00607-018-0638-1
69. Parraga-Alava J et al. "**RoCoLe coffee leaf dataset.** " Data Brief 2019; 25:104414. DOI: 10.1016/j.dib.2019.104414
70. Griffel L et al. "**SVM classification of potato virus infection.** " Comput. Electron. Agric. 2018; 153:318–324. DOI: 10.1016/j.compag.2018.08.027
71. Kyaw Y et al. "**Effect of climate variability on crop yield.** " Climate 2023; 11:204. DOI: 10.3390/cli11100204
72. Subedi B et al. "**Impact of climate change on insect pests.** " J. Agric. Food Res. 2023; 14:100733. DOI: 10.1016/j.jafr.2023.100733
73. Kumar R et al. "**Fungal blast disease detection in rice seed.** " Int. J. Adv. Comput. Sci. Appl. 2021; 12. DOI: 10.14569/IJACSA.2021.0120232
74. Haridasan A, Thomas J, and Raj ED. "**Deep learning system for paddy disease detection.** " Environ. Monit. Assess. 2022; 195:120. DOI: 10.1007/s10661-022-10656-x
75. Simhadri CG and Kondaveeti HK. "**Rice leaf disease recognition using transfer learning.** " Agronomy 2023; 13:961. DOI: 10.3390/agronomy13040961
76. Vijayan S and Chowdhary CL. "**Hybrid feature optimized CNN for rice crop disease prediction.** " Sci. Rep. 2025; 15:7904