Indonesian Journal of Electrical Engineering and Computer Science
Vol. 38, No. 2, May 2025, pp. 1308~1323
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v38.i2.pp1308-1323

Plant leaf disease detection and classification using artificial intelligence techniques: a review

Kusuma R, R. Rajkumar
[1]Department of CS&E, RNS Institute of Technology, Affiliated to Visvesvaraya Technological University, Belagavi, India
[2]Department of CS&E (Cyber Security), RNS Institute of Technology, Affiliated to Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jul 23, 2024
Revised Nov 8, 2024
Accepted Nov 24, 2024

Keywords:

Agriculture
Classification and detection
Computer vision
Image processing
Plant leaf disease

ABSTRACT

Agriculture is a cornerstone of human civilization, providing both food and economic stability. While not necessarily fatal, leaf diseases are a crucial threat to plant health. Accurate detection and classification of diseases in early stages are essential to minimize damage. Manual identification can be challenging, and delays in detection can lead to crop devastation. Fortunately, computer-aided image processing offers a solution. Researchers have explored several techniques for disease detection and classification by usage of affected leaf images, making significant progress over time. However, there's always room for improvement. Machine learning (ML), deep learning (DL) techniques have shown hopeful results. ML, DL approaches act as black-box; eXplainable AI (XAI) provides clear explanations on decisions made by these black-boxes. This study aims to present a comprehensive review on plant leaf disease detection and classification by means of ML, DL and XAI methods with an overview of the outcomes of existing techniques, summarizes their performance, evaluation metrics, and analyses the challenges in existing systems, and offers the study's inferences.

Corresponding Author:

Kusuma R
Department of CS&E, RNS Institute of Technology, Affiliated to Visvesvaraya Technological University
Belagavi, Karnataka, India
Email: 1rn22pcs04.kusuma@rnsit.ac.in

1. INTRODUCTION

In countries like India, agriculture contributes a significantly higher share approximately 16% to the gross domestic product (GDP) in comparison with the 4% of global average [1]. India is an agriculturally based country and a leading contributor in agricultural producers of the entire world. Agriculture is the foremost income source and cornerstone for Indian economy. As the population growth is increasing rapidly the demand for food is raising. To fulfil this need, increasing crop productivity and protection of crops plays a prime role.

However, crop productivity is significantly affected by various parameters like change in environmental factors and spread of plant diseases. Unpredictable weather patterns significantly impact agriculture and cannot be controlled by humans. Diseases of plants are conditions caused by biotic factors like pathogens: bacteria, fungi, viruses, micro-organisms, pests: mites, slugs, mammals, rodents, weeds: monocots, dicots and abiotic factors like temperature variation, rainfall and humidity, or nutrient deficiencies that affect the productivity [2]. Figure 1 shows the factors causing plant diseases. A major challenge in agriculture is plant disease, leading to substantial losses in crop yield, quality, and ultimately, economic productivity. If the diseases are not identified on-time, then the disease outbreaks it has harmful impact on the food security. Thus, early detection of plant diseases acts as the major area of concern in agriculture and it is crucial for effectively preventing and controlling their spread [3], [4].

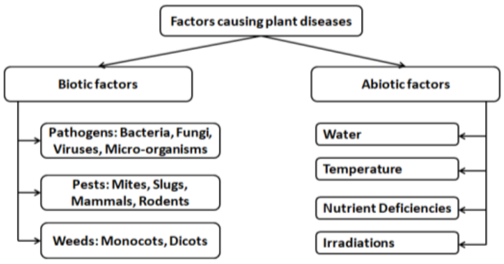

Figure 1. Factors causing plant diseases

Typically, plants afflicted by disease show signs of visible marks or lesion on various parts of it like leaf, flower, stem, fruit. Every disease tends to show a distinct visible pattern, aiding in its unique diagnosis. Generally, the leaves serve as the foremost indicators of plant diseases, with symptoms often first appearing on them. Presently, manual and visual inspection of diseased plants is performed by plant pathologist with his experience and diagnoses the disease. Manual inspection is time-consuming practice and also depends on proficiency of pathologist in recognizing disease. The failure to notice diseases early and take the measures to prevent out spread can be a devastating result as plant diseases are highly contagious, which in turn burdens agriculture personnel financially [5].

At the forefront of agricultural advancements is automatic model can use for early disease identification. Machine learning (ML), deep learning (DL), eXplainable AI (XAI) techniques are employed in identification of plant diseases in earliest stage and reduction of loss with respect to farmers. Agriculture field is witnessing a flow of ML, DL, and XAI algorithm usage as a significant step towards obtaining precision in early detection. This review paper lists crop-based common plant leaf diseases, discusses general process of disease detection system, and surveys the landscape of various ML, DL, and XAI approaches use for early detection of plant diseases. The available datasets are analyzed and explored the performance evaluation metrics employed to assess these techniques. The different crop plants will come across different diseases. The list of plants crops along with their leaves are given in Table 1.

Table 1. Crop based common diseases

Crop/fruit	Possible diseases
Apple [6]	Marsonina Leaf Blotch, Apple Necrosis Leaves, Apple Black Rot Apple Frogeye Leaf Spot, Powdery Mildew, Apple Scab, Cedar Apple Rust, Apple Mosaic, Little Brown Dots
Corn/Maize [7], [8]	Anthracnose Leaf, Common Rust, Southern Rust Eyespot, Gray Leaf Spot, Goss's Wilt, Corn Leaf Blight, Tar Spot
Cherry [9], [10]	Leaf Spot, Silver Leaf Canker, Coryneum Blight, Powdery Mildew, Prunus Necrotic, Ringspot Virus
Banana [11]	Panama Wilt, Leaf Spot- Yellow Sigatoka, Black Sigatoka
Grape [12]	Black Measles, Black Rot, Isariopsis Leaf Spot
Guava [13]	Algal leaf Spot, Rust, Whitefly
Mango [14]	Mango Anthracnose, Gray Blight, Powdery Mildew, Phoma Blight, Scab, Red Rust, Leaf Blight, Bacterial Canker
Cirtus [15]	Canker, Black Spot, Huanglongbing
Potato [16]	Early Blight, Late Blight
Rice [17]	Bacterial Leaf Blight, Brown Spot, Rice Blast
Strawberry [18]	Leaf Scorch
Tomato [19]	Bacterial Spot, Early Blight, Late Blight, Leaf Mold, Septoria Leaf Spot, Spider Mites, Target Spot, Tomato Mosaic Virus Yellow Leaf Curl Virus

A multi-step process used to identify diseases accurately is shown below. Figure 2 represents the general architecture for plant disease recognition system.

- Image acquisition: performs dataset collection for detection of diseases. Datasets can be obtained by self-collection of images directly from the fields or use the publicly available datasets.
- Image pre-processing: pre-processing of images does noise removal, and using standard techniques feature enhancement is also done. Region of interest is obtained by image segmentation.

– Feature extraction: features related to area, centroids, correlation, morphology, color, shape, texture, wavelets, size and textural descriptors are extracted.
– Classification: identification and classification of the diseases is performed through different ML and DL based classification algorithms [20].

To do research the essential part is dataset [20]. Dataset is required for the training of ML, DL, XAI algorithms for the automated disease detection. Both the magnitude and excellence of a dataset significantly impact the result of disease analyzing and classification models. The datasets can be generated through personal collection or by publicly available datasets [21]. The self-collection of datasets involves using the photography using grounded cameras; drone based aerial photography, sensor-driven data collection through an IoT network and capturing of images. The various available datasets for detection models for plants diseases [22], [23] are shown in Table 2.

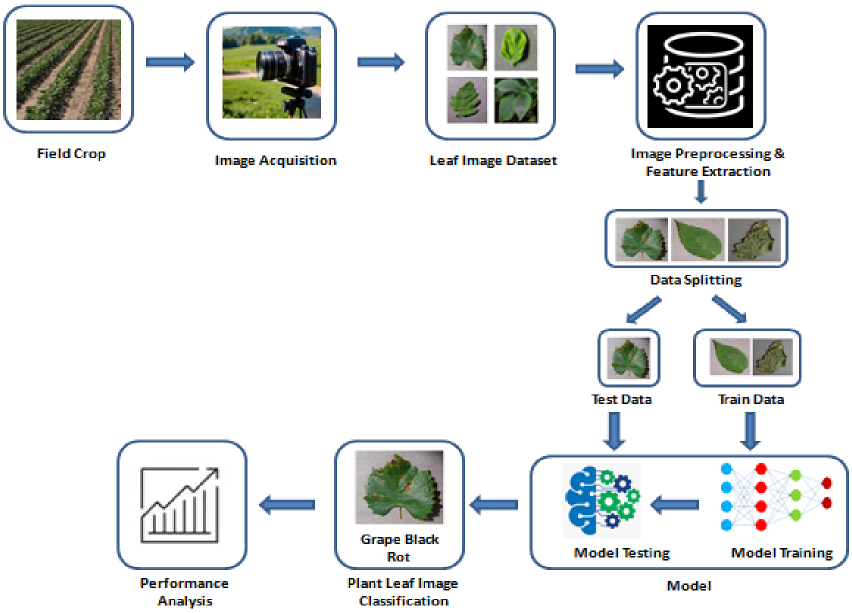

Figure 2. Architecture of plant disease recognition system

Table 2. Publicly available datasets

Reference	Dataset with description	Source
Mohanty *et al.* [24] Ali [25]	PlantVillage: Dataset contains 54303 images of 14 crops classified plant diseases divided into 38 categories.	Kaggle
Kour *et al.* [26]	PlantaeK: Dataset has 2157 images of eight different plants divided in 16 subclasses.	Mendeley
Singh *et al.* [27] Khan *et al.* [28]	PlantDoc: PlantDoc has 2,598 images of 13 plant variety with 17 unique disease categories.	Github
Ranjita *et al.* [29]	Plant Pathology 2021 - FGVC8: Dataset has 3,651 RGB images with 6 foliar diseases of apples. Dataset is updated roughly with 23,000 apple images with annotation.	Kaggle
Prajapati *et al.* [30] Chikati *et al.* [31]	UCI Rice Leaf Diseases Dataset: Dataset includes 40 images from all 3 rice leaf disease categories.	UC Irvine Machine Learning Repository
JG et al. [32] Uzhinskiy *et al.* [33]	Image Database of Plant Disease Symptoms (PDDB): Dataset has 2326 images capturing the diversity of 171 diseases across 21 plant varieties.	Repositorio Digipathos
Bhattarai, [34]	New Plant Diseases Dataset (Augmented): 87000 RGB images with normal and diseased leaves are grouped across 38 distinct groups/classes.	Kaggle
Wiesner-Hanks *et al.* [35] Sun *et al.* [36]	Maize Leaf (NLB) Dataset: Dataset contains 18,222 images, captured in field, and 105,735 annotations by human experts categorized into 2 diseases.	Open software foundation
Rauf *et al.* [37] Huang *et al.* [38]	Citrus Dataset: Dataset is made of Citrus - 150 fruits images, 609 leaves images, 5 categories in both fruits and leaves. Every image was annotated.	Mendeley
Minh Do [39]	Rice Diseases Image Dataset: Dataset comprise of four categories: 523 - Brown Spot, 1488 - Healthy, 565 - Hispa and 779 - Leaf Blast images.	Kaggle
Jepkoech *et al.* [40]	JMuBEN Datasets: The dataset contains 22591 images leaf images from Arabica coffee, and it shows three sets of unhealthy leaf images caused by Coffee Rust, Cescospora and Phoma.	Mendeley
Owomugisha, *et al.* [41] IYER [42]	Cassava Disease Dataset: Dataset include 9430 images with 1 healthy and 4 categories of disease - Brown Streak, Mosaic, Bacterial Blight and Green Mite. Images were captured by 200 farmers by phones. Experts annotated the labels.	Kaggle

2. METHOD

This segment provides a comprehensive review on recent researches done by ML, DL and XAI algorithms in recognition along with classification of plant leaf diseases.

2.1. Machine learning approaches

This segment provides review of the majority of regularly used ML algorithms like naive bayes (NB), K-nearest neighbors (KNN), support vector machine (SVM), decision tree (DT), random forest (RF), ensemble model, and logistic regression (LR) in plant leaf disease diagnosis and categorization of diseases.

Savitha *et al*. [43], a model to identify five distinct diseases that affect plants, namely: early blight, mosaic virus, down mildew, white fly, and leaf miner is developed. Model performs image pre-processing by converting all images into grayscale, then followed by image segmentation using K-means technique to remove noises, and feature extraction is carried out by gray level co-occurrence matrix (GLCM) algorithm then KNN classifier categorizes plant leaves into distinguished category. KNN classifier uses train and test ratio 80:20 of dataset. The prediction accuracy of model using KNN is 98.56% whereas existing system SVM is 97.6%. The limitation found is only 5 types of diseases are categorized and datasets used is of small size.

As per [44], an AI based system which does disease detection was developed. It makes use of 20,000 images with 19 different classes. Gaussian blur technique is used to remove noise in images and red, green, blue (RGB) images are translated into hue, saturation, value (HSV) images. Segmentation is done through K-means clustering. Dataset comprise both normal and diseased leaves with diseases like rust, rot, bacterial spot, scorch, blight leaf, mosaic virus and target spot from apple, grape, strawberry, corn, potato, and tomato. The accuracy obtained by different classifiers is logistic regression-66.4%, KNN-54.5%, SVM-53.4%, and CNN-98.0% respectively. The model can be enhanced for diseases detection on aerial photos of orchards as well vineyards captured by drones.

Kaur and Kaur [45], various fusion techniques combine different ML algorithms in the projected system which in turn improves the accuracy of the disease detection system. Images are pre-processed by HSV translation, along with segmentation, and then feature extraction is performed, continued by dimensionality reduction by principal component analysis (PCA). Hybrid fusion model employ XGBoost and Decision Tree to perform classification. This hybrid model obtained a remarkable accuracy of 100% against existing CNN model provides accuracy of 99.06%. This model innovates to build robust accurate solution.

As discussed in [46], a model is developed based on computer vision (CV) and ML techniques. Images are pre-processed by performing the grayscale conversion, smoothing, binarizing and morphological transfer. Foreground detection along with bitwise AND on binarized image plus original image results in RGB image of segmented leaves. RGB image is transformed to HSV color space. Image characteristics like color, shape and texture are extracted using GLCM algorithm. A mixed method of classifier random decision forest is used for classification. Model is tested 5-fold for more robustness with respect to tomato, potato, grapes, apple and corn with split of 60-40 train and test data. An average of 93.7% accuracy was achieved. Demerit is related diseases for only limited crops were investigated.

Harakannanavar [47], authors have presented an approach to identify tomato plant diseases using CV and ML techniques. This model performs pre-processing of images by grayscale conversion, histogram equalization, k-means clustering and tracing of contour. The feature extraction taken place by PCA, GLCM and discrete wavelet transform (DWT). Tomato dataset with 600 sample images were tested in this model using different classifiers. Classification is performed using different classification algorithms like SVM, CNN and K-NN. Model accuracy achieved using SVM, KNN and CNN is 88%, 97% and 99.6% respectively. This model was limited to tomato-related diseases.

As per [48], a model is presented which uses the images and performs global extraction of features and labels the images in H5 file format. RGB-HSV conversion and segmentation is performed for color extraction. Model uses hu moments, histogram of oriented gradients (HOG), Haralick texture feature descriptors. Usage of descriptors resulted in feature vector representation of image. Once all the features are extracted labels are assigned. Label Encoder function encodes the labels and MinMaxScaler function performs normalization of images. Then the model saves the features and labels in an H5 file format using the h5py library. This output is fed into Naive Bayes to carry out classification of plant leaf disease. Output Accuracy is not discussed.

Panigrahi *et al*. [49], the projected model performs classification of maize leaf diseases using various supervised ML algorithms. Maize leaf images are converted from RGB to Gray scale in pre-processing, image segmentation does Label edge detection method, RGB feature extraction extracts color, shape and texture features. Maize dataset contains 3823 images with four classes are split into 90:10 ratio for train-test in the model. The accuracy obtained by these classifiers is SVM-77.56%, NB-77.46%, KNN-76.16%, DT-74.35%, and RF-79.23%. The RF classifier outperforms among others. This model was limited to maize-related diseases.

Hatuwal *et al.* [50], authors have proposed the work which classifies the PlantVillage dataset using different classifiers and compares the results. Contrast, correlation, entropy, RGB color standard deviation and inverse difference moment's features extraction is done by Haralick texture algorithm using GLCM. The train and test data are split into ratio of 80:20. Classification is performed by ML classifiers like SVM, RF, KNN and CNN. CNN has obtained highest accuracy of 97.89% then RF-87.436%, SVM -78.61% and KNN-76.96% respectively. Limitation of algorithms like grid search is not incorporated to get best value of hyper parameters.

According to Selvakumar and Balasundaram [51], a model to classify mango leave diseases is presented. The model uses mango leaves images taken from PlantVillage dataset with 4335 images. Images are pre-processed and segmented. Image dataset is spilt up into 70:30 ratios for train and test propose. DT, SVM, neural network, and linear regression are used to categorize mango leaf images. The performance metrics such as classification accuracy, area of curvature (AOC), precision, F1-score and recall are found. Neural network outperforms among all with AOC-0.988, accuracy-0.943, F1-score-0.952, precision-0.970 and recall-0.937. This model uses an ensemble technique (stacking) of all the four algorithms, which resulted in improved accuracy of model by 98.6% and further Adam optimizer is used in model which improved accuracy to 99.2%.

As discussed in [52], a model for rice disease prediction using Hadoop platform is presented. Map Reduce framework is used for the disease prediction. Feature selections by Map phase and classification module by Reduce Phase are performed. Set theory-based feature selection algorithm performs the feature selection and different ensemble classifier made up of Adaboost, random forest and bagging is used as the base classifier and majority voting principle be employed to get the ultimate output of classification. Model used rice leaf of dataset of 500 images with four diseases and achieved accuracy of about 88.19% in rice leaf disease detection. Generalized prediction model was built from presented work which does the classification of human disease too and accuracy achieved was good. Table 3 (see appendix) shows the summary of discussed ML models.

2.2. Deep learning approaches

A review of the majority repeatedly used deep learning approaches in plant disease diagnosis and classification such as CNN, LeNet, long short-term memory networks (LSTMs), AlexNet, GoogleNet, ResNet50 and InceptionV3.

As per [53], authors have anticipated a model for real-time recognition of apple leaf diseases by means of enhanced CNN. Basically, captured apple foliar pictures are data augmentation and annotated to create a large foliar disease dataset. In this model pre-processing is performed to make all images to a particular standard size. Model is prepared to use hierarchical data format version 5 (HDF5) file format. LeNet CNN architecture performs the sub-sampling and reduced feature maps, and then classification is performed by model which was trained on MNIST database with 60000 training and 10000 testing images. Features are extracted using techniques like PCA, independent component analysis (ICA), plus linear discriminant analysis (LDA) followed by Maxpooling. Classification says whether the leaf is healthy or unhealthy. Model has achieved accuracy 0.95 and precision 0.90 for the given dataset.

Sengupta *et al.* [54], optimal trained DL based LSTM model is presented for detecting diseases. It uses 50,000 images of various crops containing both normal and infected leaves from PlantVillage dataset. Histogram equalization performs data pre-processing. Statistical features, LGP and color features are extracted. Self-improved blue monkey optimization (SI-BMO) algorithm optimizes the weight in the LSTM classifier. Optimized LSTM performs the detection and classification. The performance of adopted LSTM + SI-BMO was analyses over other classifiers CNN, SVM, deep belief networks (DBN) and gated recurrent units (GRU) and it outperformed over them. LSTM with optimization has achieved accuracy 0.944262 whereas LSTM with no optimization has achieved an accuracy of 0.739268.

Deputy *et al.* [55], an automated image recognition-based disease diagnosis system based on deep learning techniques is developed. Feature extraction done by two ways. In first technique image processing methodologies performs features extraction of color, shape and texture, HOG, speeded-up robust features (SURF), are extracted. In second method AlexNet's pre-trained DL model extract's feature. Gathering of features is done by back propagation in neural network (BPNN) algorithm. Model uses PlantVillage dataset. Training from scratch and transfer learning using DL architectures like AlexNet, ResNet50, InceptionV3 and GoogleNet was performed. Leaves with color, grayscale and segmented were used for training. All the models were able to handle the approach of training from scratch, whereas ResNet50 and InceptionV3 were able to handle the approach of transfer learning. Model uses different dataset ratios like 90:10, 80:20, 70:30 and 60:40 for train and testing. With 90:10 train-tests ratio highest accuracy achieved by GoogLeNet architecture was 0.999 and 0.996 for color and segmented images respectively, and 0.994 for grayscale images by InceptionV3. All the architectures achieved maximum F1-score 1.0 for color, segmented images. GoogleNet and InceptionV3 achieved 0.999 F1-score for grayscale images.

Padshetty and Ambika [56], authors propose leaky rectilinear residual network (LRRN) model. Leaves for training-testing are obtained from PlantVillage dataset. Resizing and scaling, data augmentation and noise reduction are performed in data pre-processing phase. Color, structure, texture and size are the features that are extracted. ResNet architecture was integrated with Leaky ReLU activation function to classify diseases. The outcome of LRRN model is compared with hybrid random forest multiclass support vector machine (HRF-MCSVM), multi-feature fusion faster region-based convolutional neural network (MF3R-CNN), optimal mobile network-based convolutional neural network (OMN-CNN), and convolutional neural network-based visual geometry group 19 (CNN-VGG19). The LRRN model achieves accuracy-94.56%, precision-93.48%, recall-93.12%, F1-score-93.82%, and specificity-92.58% that outperforms when compared with other models.

As per [57], a convolutional neural network-based squeeze and excitation network (SENet-CNN) model does classification of potato leaf diseases. Initially data is collected in data acquisition phase followed by data annotation and augmentation. Image resizing and noise reduction is performed in image pre-processing phase. Improved median filter is used to noise removal. Classification is done using visual geometry group-VGG16 convolutional layers uses ReLU activation function followed my Maxpooling along with inception and SE block, then average polling and SoftMax. Experiments are carried out using Dataset 38, Dataset 15, and PlantVillage dataset. Performance of SENet-CNN classification model is compared with SVM, VGG-16 CNN, deep convolutional neural networks (DCNN), mask region-based CNN (Mask-R CNN), 1-dimensional CNN (1D-CNN), Shallow-CNN. The proposed SENet-CNN model has superior flexibility with changing images and achieved highest accuracy of 99.3%.

Balaji *et al*. [58], a model with transfer learning technique for multimodal diseases with enhanced CNN is proposed. Pre-processing is done using gaussian filter. It reduces the noise as well as converts the images into grayscale images. Genetic algorithm (GA) performs the feature extraction. GA is based on fitness function of given system. It uses three functions internally like operation in selection, crossover and mutation function. Transfer learning is applied in this model. ResNet-50 is used to extract the features along with SoftMax functions during classification. Enhanced multiple CNN (ECNN) is used for finding and categorization of diseases in rice plant. Transfer learning (TL), CNN+TL, artificial neural networks (ANN) algorithms performance compared with ECNN+GA. Enhanced accuracy of 95% is obtained in this model.

As discussed in [59], authors have presented a robust DL technique ResNet-34 based faster region-convolutional neural network (Faster-RCNN) model. From PlantVillage dataset only tomato plant images with 10 specified classes are used here. The annotation of specified images is done to obtain RoI. ResNet-34 along with convolutional block attention module (CBAM) extracts the features like size, colour, and texture. Faster-RCNN performs the classification which is performed through various steps like region proposal networks (RPN), pooling, find and categorize a number of plant diseases. Class based tomato plant evaluation is done with various metrics accuracy, precision, recall, F1-Score and error rate. GoogleNet, VGG-19, ResNet-101, Xception and SE-ResNet50 are used to examine the robustness of Faster-RCNN. Faster-RCNN approach with ResNet-34 obtained the results with average precision 99.48%, recall 99.32%, F1-score 99.42%, and accuracy 99.97% values highest over other DL approaches. It’s a low-cost solution framework.

Patani *et al*. [60], authors have presented GoogLeNet based model for recognition of diseases. PlantVillage dataset is used with division ratio of 80:20 for train and test. Transfer learning method does training. Image pre-processing is performed using filters to remove noises and normalize the intensity of images followed by image annotation and image augmentation. Features are extracted and then classification is done using GoogLeNet with 22 layers’ deep architecture and 9 inception modules. Model achieves 99.35% accuracy.

As per [61], AlexNet modification architecture-based CNN model has been presented. A dataset has collection of tomato images from PlantVillage with 8,345 images for training and 4,585 images for testing in this predictive model using exploratory data analysis function. The CNN model has different layers as convolutional, sub sampling, fully connected layers. To get more accurate classification, an AlexNet based modified architecture with CNN is used. It has 3 convolutional layers, 3 fully connected layers, and 1 output layer. Convolutional layer uses ReLU, output layer uses the SoftMax activation function. To get better efficiency Adam optimizer is included in modified architecture. A mobile-based android platform is developed using this architecture. Performance measure has resulted in average accuracy 96%, precision 98%, recall 95%, and F1-Score 97%.

As discussed in [62], CNN-based pre-trained model is proposed. Transfer learning based pre-trained models DenseNet-121, VGG-1, ResNet-50, and InceptionV4 are used in this model and majorly concentrated on fine tuning the hyper parameters. The model works on PlantVillage dataset containing 54,305 image samples of 14 crop species belonging to 38 classes. Dataset is pre-processed along with augmentation, then feature extraction and categorization. Metrics used are accuracy, F1 score, sensitivity and specificity. A comparative examination of entire models was performed and results showed DenseNet-121 attained higher accuracy-99.81% which was superior of all other models. Summary of the discussed DL approaches is represented in Table 4.

Table 4. Summary of DL approaches

Reference	Feature extraction	Techniques	Crops	Size	Performance metric
[53]	PCA, ICA, and LDA followed by Maxpooling	LeNet CNN	Apple	60000 training and 10000 testing images	Accuracy 0.95 Precision 0.90
[54]	Histogram equalization, LGP	Self-Improved Blue Monkey Optimization (SI-BMO)	Multiple crops	50,000 images	LSTM with optimization accuracy 0.945 LSTM with no optimization accuracy 0.739
[55]	HOG, SURF, BPNN algorithm	AlexNet, GoogleNet, ResNet50, InceptionV3	Multiple crops	Dataset ratio like 90:10	GoogLeNet gets Highest accuracy - 0.999, F1-score-1.0 for colour images
[56]	Leaky Rectilinear Residual Network	ResNet50	Potato pepper corn grape	24,000 images with 80:20 ratio	Accuracy 94.56% F1-score 93.82% Precision 93.48% Recall 93.12% Specificity 92.58%
[57]	improved median filter	SENet-CNN	Potato	Train-validation-test data ratio 70-20-10	Accuracy 99.3%
[58]	Gaussian filter Genetic algorithm ResNet-50	Enhanced multiple CNN with transfer learning technique	Rice	Not Available	Accuracy 95%
[59]	ResNet-34 with CBAM	Faster-RCNN	Tomato	18,160 Original images + annotated images	Accuracy 99.97% Average Precision 99.48% F1-score 99.42% Recall 99.32%
[60]	Not Available	GoogLeNet	Multiple crops	Train - test data ratio 80:20	Accuracy 99.35%
[61]	Not Available	AlexNet modification architecture-based CNN	Tomato	18,345 images for training and 4,585 images for testing	Average Accuracy 96% Precision 98% Recall 95% F1-Measure 97%
[62]	Filters	DenseNet-121 VGG-16 ResNet-50 Inception V4	Multiple crops	54,305 images	DenseNet-121 attained higher accuracy-99.81%

2.3. eXplainable AI approaches

Artificial intelligence (AI) builds the models to imitate human intelligence and try to find the solutions to real-world problems. ML and DL are AI techniques use algorithms which envisage outcome more correctly without human interference [63], [64]. The output of ML and DL models act as black box and it is not easy to understand the process of getting those results. DL and ML approaches cannot give explanation to users on how the model operates and critical decision-making is done. For users this gap results in trust issues over technology, eventually stop them from final application usage.

AI has potential to revolutionize fields like healthcare, agriculture and finance is undeniable; it's the concept of responsible AI that unlocks its true power. This approach emphasizes building trustworthy AI systems that prioritize:

– Minimizing bias: ensuring AI algorithms are fair and unbiased, preventing discriminatory outcomes.

– Protecting privacy: safeguard individual data privacy, ensuring responsible data collection and usage.
– Supporting security: building robust AI systems that are resistant to cyberattacks and manipulation.
– Enhancing transparency and accountability: making AI decision-making processes understandable and allowing for human oversight and control [65].

Transparency is a cornerstone of responsible AI, and (XAI) acts as the key that unlocks it. XAI plays vital roles in building the AI models which are easy to interpret and understood by humans. XAI bridges the gap by presenting clear explanations on AI decisions, allowing humans to understand the reasoning behind them. This transparency is vital for: building trust, identifying bias: ensuring accountability [66].

Review is carried out on frequently used XAI approaches: Gradient-weighted class activation mapping (GradCAM), locally interpretable model agnostic explanations (LIME), SmoothGrad, XAI saliency method and shapley additive explanations (SHAP) in explanation of plant disease diagnosis process and categorization done by various ML, DL algorithms. Figure 3 shows the representation of XAI architecture for disease detection.

As per [67], authors have presented a model using transfer learning and XAI approaches. The usage of Kaggle PlantVillage dataset is done to detect the plant leaf diseases. Data is pre-processed and then split into train, validate and test data. Three pretrained CNN models such as are used EfficientNetV2L, MobileNetV2, and ResNet152V2 along with Adam optimizer performs the detection of various plant diseases. The accuracy achieved by EfficientNetV2L, MobileNetV2, and ResNet152V2 are 99.63%, 98.86% and 98.44% respectively. EfficientNetV2L has achieved highest accuracy among the three pretrained models. LIME interprets the decisions obtained by EfficientNetV2L during classifications.

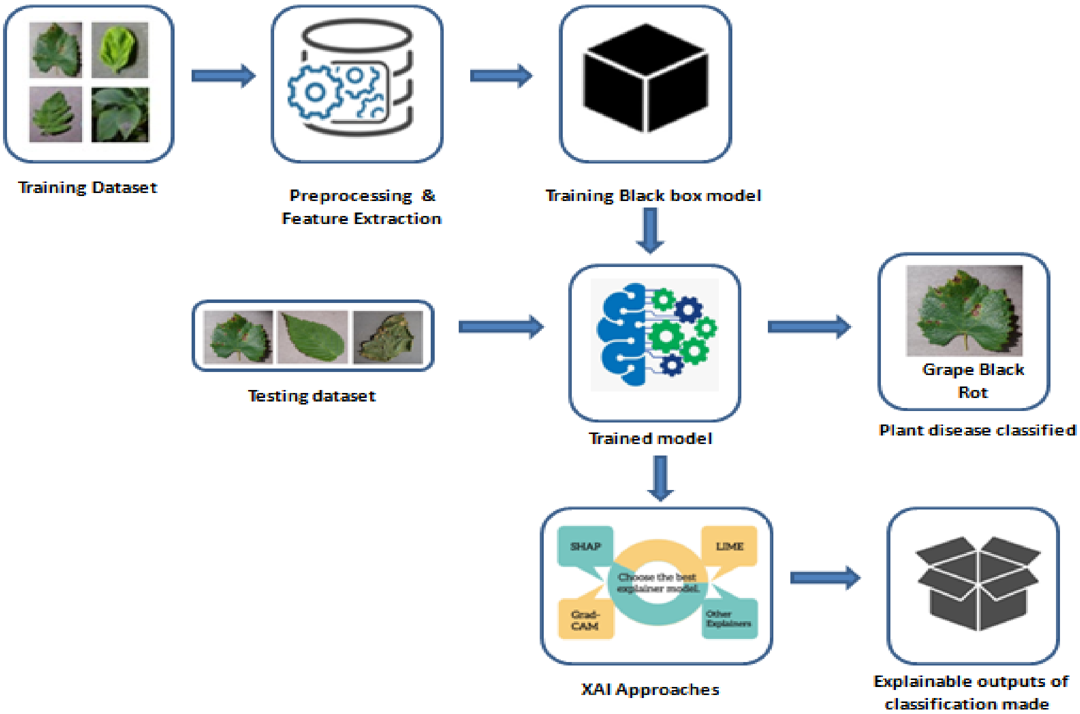

Figure 3. Representation of XAI architecture for leaf-based plant disease detection

Bhandari *et al.* [68], an explanation driven DL approach for leaf disease recognition and classification on tomato is proposed. Tomato leaf dataset is obtained from PlantVillage dataset and is split into 90:10 ratios. Model uses DL algorithm for leaf disease recognition, classification and XAI algorithms for explanation of the classification. EfficientNetB5 does the disease identification along with classification. EfficientNetB5 DL model performance is compared with other models and it has obtained highest test accuracy of 99.0% whereas MobileNet-94.00%, Xception-95.32%, VGG16-93.35%, ResNet50-(96.03%, and DenseNet121-96.30%. GradCAM and LIME were used to explain results of classification by identifying the specific regions that are accountable for classification. LIME performed better than GradCAM in identifying the regions of images.

As discussed in [69], authors have presented leaf disease classification model using explainable deep learning algorithms. The Dataset contains only different kinds of fruit leaf images from PlantVillage dataset. Data is augmented and texture, shape and disease spot features were extracted. ResNet was associated with attention model CBAM to advance the ability to do key features extraction and improve power of generalization. 99.11%, 99.4% and 99.89% accuracy is achieved from the experimental results. Three XAI visualization methods such as GradCAM, LIME, and SmoothGrad were used. Three model's interpretabilities of the classification of different kinds on fruit leaf diseases were compared and GradCAM was the most suitable model among the three.

The authors have proposed a model which uses XAI for DL disease detection system [70]. Several DL architectures are available for disease detection but they act as black box systems and they don't provide any classification explanations. Aim of this model is to explain why the model has classified for the given image into this category of disease/healthy. PlantVillage dataset was used. Mask R-CNN performs detection and classification. GradCAM++ explains the model. Model identifies the location and regions which contributes to disease and aids in the classification with an accuracy of 99.67%.

Arvind *et al*. [71], a computer aided distinct system with deep learning-based disease classification with XAI pipeline is discussed. Tomato plant images are chosen from PlantVillage dataset and different augmentation methods are applied to make the dataset of 43,212 images. Model is trained by transfer learning technique. VGG16, VGG19, ResNet, Inception V3, MobileNet, EfficientNet architectures were used for the categorization of tomato leaf images. Among all of them EfficientNet has attain the highest accuracy of 0.92 and 0.91 for original as well as augmented images. To explain the model working LIME and GradCAM are used along with validation of you only look once (YOLO) V4. LIME uses linear regression with salient features and GradCAM uses gradient information for the explanation of classification of images.

As per [72], a TL-based model for mango leaf disease detection is presented. Dataset contains only mango plant leaves obtained from PlantVillage dataset. Data is pre-processed and augmented. Dataset includes ratio of 70:20:10 with respect to train, validate and test purpose. Pre trained transfer learning approach based DenseNet169 does mango leaf disease recognition and classification. Accuracy, F1 score, Precision, and AUC are evaluated. Model has obtained the results with 97.41% accuracy rate. XAI algorithm LIME is used to explain the obtained results.

As discussed in [73], authors have presented the DL model with XAI for disease detection. Model uses only potato leaf images with augmentation from PlantDoc dataset. Pre-trained ResNet50 architecture with Faster-RCNN model on ImageNet dataset detects diseases on potato leaves. Model has three phases feature extraction network, region proposal network and detection generator. Quantitative and qualitative experiments were performed. XAI saliency method was used to explain the potato disease detection based on intermediate object detection results. Detector randomized input sampling for explanation (D-RISE) method and XAI saliency method were analyzed with respect to the explanation of classification. XAI saliency method explanation was better when compared to D-RISE.

Nahiduzzaman *et al.* [74], an explainable DL model for mulberry leaf disease detection is presented. Dataset is collected from Bangladesh region with leaf diseases: rust, spot and normal leaves followed by annotation of the leaf images by experts. Dataset was pre-processed and augmentation techniques were applied. Dataset is split and used to train, validate and test the model. A Lightweight Parallel Depth-Wise Separable CNN (PDS-CNN) model is provided to progress the performance of classification. SHAP explains the outcomes of PDS-CNN. The PDS-CNN performs preeminent in comparison with other six models like ResNet152, Xception, MobileNetV2, VGG19, MobileNet and DenseNet121. PDS-CNN has attained an optimistic accuracy - for three class classifications as 95.05 ± 2.86% and for binary classifications as 96.06 ± 3.01%.

Bandi *et al*. [75], authors have developed explainable AI based leaf disease severity classification model using transformer networks. Data acquisition of only apple leaf images is done from PlantDoc and PlantVillage datasets. Data is pre-processed and augmented. DL model YOLOv5 performs disease detection. Background removal was made through U2-Net architecture which aids in accurate classification. Depending on severity of infection, disease can be classified into multiple stages like low, moderate and high. Model's output explanation is given by Grad-CAM. Model obtained an F1 score 91% on original as well enhanced datasets. The Vision transformer (ViT) performed the classification and results achieved were more accurate with ViT networks. Table 5 shows summary of XAI models.

Figure 4 represents the plants investigated in the overall survey. Represents the papers surveyed in reference to each fruit – Apple, grape, peach, tomato, mango; strawberry and multiple fruits survey. represents the surveyed in reference to each crop- Corn, pepper, potato, maize, multiple crops, rice and mulberry.

Table 5. Summary of XAI approaches

Reference	Techniques	eXplainable-AI Technique	Crops	Size	Performance metric
[67]	EfficientNetV2L MobileNetV2 ResNet152V2	LIME	Multiple crops	43,429 Train, 5,417 validation, and 5,459 test images	EfficientNetV2L, MobileNetV2, and ResNet152V2 are 99.63%, 98.86% and 98.44%
[68]	EfficientNetB5	GradCAM LIME	Tomato	Train–test images ratio 90:10	EfficientNetB5 highest accuracy of 99.0% MobileNet-94.00%, Xception-95.32%, VGG16-93.35%, ResNet50-96.03% DenseNet121-96.30%
[69]	VGG GoogLeNet ResNet with CBAM	SmoothGrad LIME GradCAM	Multiple fruits	34,000 images	VGG, GoogLeNet and ResNet with attention model CBAM 99.11%,99.4% and 99.89%
[70]	Mask R-CNN	GradCAM++	Multiple crops	70,000 images	Accuracy 99.67%
[71]	VGG16 VGG19 ResNet Inception V3 MobileNet EfficientNet	LIME GradCAM YOLOV4	Tomato	Augmented 43,212 images	EfficientNet got an accuracy of 0.92, 0.91 for original and augmented images respectively
[72]	DenseNet169	LIME	Mango	Train-test-validation images ratio 70:20:10	97.41% accuracy rate.
[73]	Faster RCNN based ResNet50	D-RISE XAI saliency method	Potato	2,568 images over 13 plants and 27 classes	D-RISE achieves mAP of 46.22% whereas Faster RCNN achieves mAP of 39.37%.
[74]	PDS-CNN	SHAP	Mulberry	Original 1,091 images used to generate 6000 synthetic images with 4000 train and 2000 test images	PDS-CNN has achieved an optimistic accuracy - for three class classifications as 95.05 ± 2.86% and for binary classifications as 96.06 ± 3.01%.
[75]	Deep learning modelYOLOv5, U2-Net architecture	Grad-CAM	Apple	2572 images	Apple leaf disease multi stage classification by ViT with and without background obtained F1-Score 0.758 and 0.91respetively.

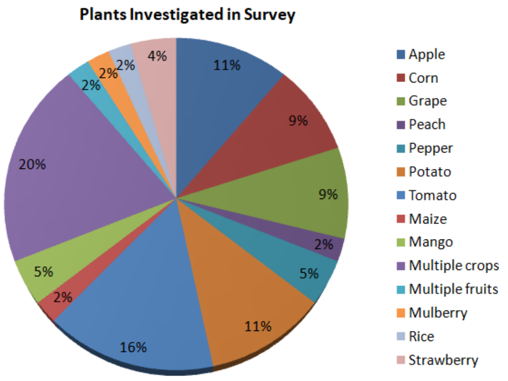

Figure 4. Percentage of studies per crop and/or fruit species

3. RESULTS AND DISCUSSION

In the results and discussion section, inference of study and open doors for future works based on the challenges of the present system [76], [77] are discussed:

3.1. Key findings of study

The practical inference of this study includes:
- Revolutionizing agriculture with ML and DL as disease detectives and XAI as interpreter: ML and DL cutting-edge technologies are poised to revolutionize the way we identify and manage plant diseases. By analysing vast amounts of data, including images and sensor readings, ML and DL models can outsmart the human eye, act as early warning systems and scale up for large-scale agriculture. Automate disease detection. XAI helps understand the model in better way. The benefits for farmers and plant specialists are immense: boosted crop yields and cost savings.
- Development of generalizable models: by saving time and effort, these models open door for researchers and practitioners, to recognize and categorize diseases in various situation facilitating faster response.
- Open-source datasets in training and evaluation: increased number of public datasets helps researchers to design extra precise robust models with improved performance.
- Reduction in possible expenses: usage of ML, DL methods offer cost-effective and time-saving solution as manual labour required by farmers and agricultural workers is reduce and eliminates the need for costly equipment or pathology expert.

Knowledge transfer to other domain: research has ability to put forward transferable framework for disease identification and classification with the usage of ML, DL, XAI techniques. Framework can be adapted to other domain like analysis of medical images, and remote sensing.

3.2. Challenges faced and scope for future works

The various challenges faced in present plant disease detection systems and these challenges can be used as the open doors of future works in this domain are:

3.2.1. Limited dataset

Although DL has revolutionized computer vision tasks, agricultural applications like disease detection in plant has a unique hurdle: limited data availability. Unlike publicly available datasets, collecting high-quality agricultural data can be labor-intensive and expensive. This scarcity, particularly for low-incidence diseases, restricts the potential of deep learning in this crucial field. Researchers are tackling this challenge with three innovative solutions: data amplification, synthesis, and generation.

All techniques used to deal with the problem of small datasets, particularly in disease detection with respect to plants. Here's a breakdown of each:
- Data amplification: focuses on stretching existing data using techniques like rotations, flips, cropping of images to create variations. Adding noise or slight blur effects to simulate real-world conditions.
- Data synthesis: aims to artificially create new data points using generative models that resemble your real data.
- Data generation: a broader term encompassing both amplification and synthesis. It refers to any method that increases the size and diversity of your dataset. Apart from above solutions, researchers leveraged generative adversarial networks (GANs) [78] and automated encoders [79] to deal with the challenge of small datasets. These powerful tools generate a flood of fresh, diverse samples, significantly enriching the data available for training.

3.2.2. Usage of transfer learning model

Traditionally training AI for plant disease detection need a massive quantity of data which is labelled can be scarce. Transfer learning leverages pre-trained models, already good at recognizing general features from tons of images, and tailor them to identify specific plant diseases with less data. By fine-tuning these pre-trained models, researchers can achieve high accuracy while saving time and resources.

For example, VGG network was used to identify infected potatoes images of various size, colors, and form under natural light images [80]. Shows by application of fine tuning and contrast parameters improve accuracy of Dense Net with increased iterations [81]. Shows using TL along with fine-tuning, images of rice diseases under complex background conditions were successfully identified, with an average accuracy of 92.00% [82]. This result demonstrates that TL performs best in comparison with training from scratch.

3.2.3. Sensible network structure design

Proper sensible network design results in reduction of model requirements [83]. Vegetable disease identification is relied on three-channel convolutional neural networks (TCCNN) to extract color information using high-level discriminant features, as a replacement for complex pre-processing, segmentation of lesion and manual feature extraction.

Grape leaf diseases identification was performed by improved CNN [84]. Deep separable convolution network usage improves over fitting and reduces parameters. The multi-scale feature extraction was performed through inception structures. Dense connectivity facilitates feature reuse and propagation, leading to 97.22% recognition accuracy, exceeding the performance of other transfer learning methods.

3.2.4. Early detection of tiny lesions

The attention mechanism prioritizes relevant areas in plant images, effectively filtering out background noise [85]. By analyzing image features through a weighted sum approach, it isolates the key elements. The SoftMax function permits noise reduction. This function extracts a prominent image, separates it from the background, and fuses it with the original image, creating a clearer representation. This focus on crucial data enables attention mechanism to allocate resources efficiently, leading to more accurate identification of subtle lesions, especially in the premature stages of diseases.

While attention mechanisms like network residual attention [86] have proven highly accurate, future research should focus on further enhancing precision for tiny lesion detection. This could involve developing robust pre-processing algorithms that effectively reduce background noise and improve image resolution through techniques like image enhancement, denoising, and super-resolution.

3.2.5. Impact of lighting on detection accuracy

Lighting problems: While indoor light boxes simplify image processing by eliminating external light influences [87], they significantly differ from real-world scenarios with natural light. Natural light's dynamic nature and camera limitations often lead to color distortions beyond the camera's acceptable range. Furthermore, varying view angles and distances during image capture significantly alter the appearance of plant diseases and pests, posing challenges for visual recognition algorithms.

Occlusion problem: current research in plant disease identification often overlooks the complexities of real-world environments. Focusing solely on single-background images and ignoring occlusion issues leads to significantly lower accuracy and limited practical application. Occlusion, a common occurrence in natural settings due to factors like leaf posture, branches, lighting, and combinations thereof, poses significant challenges. Be short of visible features and noise due to occluded areas hinder accurate recognition, leading to potential misidentification or missed detection entirely.

Recent advancements in DL have challenged the limitations of prior work like Liu and Wang, demonstrating significant progress in identifying plant diseases under occluded conditions [88], [89]. This opens doors for real-world applications. However, the inherent randomness and complexity of occlusion pose challenges: training basic frameworks is difficult, and hardware dependence remains. To address these issues, research should focus on: Lightweight network architectures, GAN exploration, network performance quantification. By tackling these areas, researchers can create more robust and practical solutions for plant disease identification in real-world scenarios with occlusion.

Detection speed problem, DL offers greater performance compared to conventional methods for plant disease detection, its high computational complexity presents a significant challenge. Achieving high accuracy often demands extensive image feature learning, leading to increased computational load and slower detection speeds, hindering real-time applications. To overcome this, finding an efficient algorithm that balances accuracy and speed is crucial. This might involve model optimization, training optimization and hardware optimization. By striking a balance between accuracy and speed through these approaches, researchers can develop practical deep learning solutions for real-time scenarios.

In real-world agricultural applications, utilizing DL for plant disease and pest detection, model inference speed takes center stage. While high accuracy remains crucial, traditional methods often prioritize it over efficiency. KC *et al*. [90] addresses this by introducing a deep convolution model for disease detection. This model, named Reduced MobileNet, achieve accuracy of 98.34% in comparison with VGG and MobileNet (29x and 6x less, respectively). This demonstrates a valuable trade-off between accuracy and efficiency, making reduced MobileNet suitable for real-time disease diagnosis of crop on limited resource mobile devices.

4. CONCLUSION

Early recognition of plant diseases using automatic plant disease detection systems plays critical role for preventing their spread and managing their impact. Automatic plant leaf disease detection systems are built using ML and DL black box approaches and XAI approaches are used for more transparent to trust and understand decision process of systems by users. Challenges faced by present disease detection systems such as availability of limited datasets, difficult to implementation transfer learning models, early detection of tiny lesions and impact of lighting on detection accuracy provides as an insight for researchers the scope for further work and to improve the strength and generality of disease detection models, it’s crucial to collect images from diverse plant development stages, various time period, and different geographical area. Future research possibilities in improving and enhancing the existing state-of-the-art in plant disease identification can be disease stage identification, multiple disease infection, and quantification of a disease.

APPENDIX

Table 3. Summary of ML approaches

Reference	Feature extraction	Techniques	Crops	Size	Diseases detected	Accuracy
[43]	K-means GLCM algorithm	• SVM • KNN	Multiple crops	75 images	Early Blight Mosaic Virus Down Mildew White Fly Leaf Miner	SVM 97.6% KNN 98.56%
[44]	K-means	• Logistic Regression • KNN • SVM • CNN	Apple Grape Strawberry Corn Potato Tomato	20,000 images	Black Rot Bacterial Spot Early Blight Late Blight Leaf Scorch Mosaic Virus Rust Target Spot	LR 66.4% KNN 54.5% SVM 53.4% CNN 98.0%
[45]	PCA	• CNN • XGBoost and Decision Tree	Multiple crops	Not Available	Black Mold Bacterial Spot Gray Spot Late Mold Powdery Mildew	CNN 99.06% XGBoost and DT 100%
[46]	GCLM	• Random Decision Forest	Apple Corn Grapes Potato Tomato	87,000 images	Multiple disease classification of Diseased Leaves	93.7%
[47]	DWT PCA GLCM	• SVM • KNN • CNN	Tomato	Not Available	Late Blight Septorial spot Tomato Mosaic Yellow Curved Bacterial Spot	SVM 88% KNN 97% CNN 99.09%
[48]	Hu Moments, Haralick Texture, and Histogram of Oriented Gradients	• NB	Multiple crops	Not Available	Diseased Leaves	Not Available
[49]	Label edge detection method, RGB feature extraction	• NB • DT • KNN • SVM • RF	Maize	3,823 images	Gray Leaf Spot Common Rust Northern Leaf Blight	RF 79.23%. SVM 77.56% NB 77.46% KNN 76.16% DT 74.35%
[50]	Haralick texture algorithm using GLCM	• SVM • RF • KNN • CNN	Apple Corn Grape Peach Peeper Strawberry	37,058 images	Diseased Leaves	CNN 97.89% RF 87.436% SVM 78.61% KNN 76.96%
[51]	Not Available	• Ensemble Technique (Stacking) • Adam optimizer	Mango	4,335 images	Diseased Leaves	ensemble technique 98.6% Adam optimizer 99.2%
[52]	Set theory based selection of features	• Ensemble Classifier	Rice	500 images	Bacterial Blight Brown Spot Rice Blast Sheath Rot	88.19%

REFERENCES

[1] World Bank Group, 2024. [Online] Available: https://www.worldbank.org/en/topic/agriculture/overview.
[2] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, vol. 11, no. 1, p. 5, Jan. 2024, doi: 10.1186/s40537-023-00863-9.

[3] D. Tejaswi, T. Sri Vaishnavi, B. Nandini, P. Nuka Raju, and D. Jayanth Babu, "Plant disease detection using deep learning," *International Journal of Science and Research Archive*, vol. 12, no. 1, pp. 2476–2488, Jun. 2024, doi: 10.30574/ijsra.2024.12.1.1043.

[4] P. Dey, T. Mahmud, S. R. Nahar, M. S. Hossain, and K. Andersson, "Plant Disease Detection in Precision Agriculture: Deep Learning Approaches," in *2nd International Conference on Intelligent Data Communication Technologies and Internet of Things, IDCIoT 2024*, Jan. 2024, pp. 661–667, doi: 10.1109/IDCIoT59759.2024.10467525.

[5] S. Dasari and S. Kathula, "A Study of Plant Disease Detection and Classification by Deep Learning Approaches," *Proceedings of International Conference on Agriculture*, vol. 7, no. 2, pp. 31-40, May 2023, doi: 10.17501/26827018.2023.7202.

[6] A. Bonkra *et al.*, "Apple Leave Disease Detection Using Collaborative ML/DL and Artificial Intelligence Methods: Scientometric Analysis," *International Journal of Environmental Research and Public Health*, vol. 20, no. 4, p. 3222, Feb. 2023, doi: 10.3390/ijerph20043222.

[7] P. Shenbagam and N. Sanjana, "Corn Leaf Disease Detection; A Survey," in *5th International Conference on Inventive Computation Technologies, ICICT 2022 - Proceedings*, Jul. 2022, pp. 1287–1294, doi: 10.1109/ICICT54344.2022.9850647.

[8] "Corn Leaf Diseases," *Crop Science United States*, 2023. [Online] Available:https://www.cropscience.bayer.us/articles/bayer/general-leaf-diseases-of-corn.

[9] N. Sengar, M. K. Dutta, and C. M. Travieso, "Computer vision based technique for identification and quantification of powdery mildew disease in cherry leaves," *Computing*, vol. 100, no. 11, pp. 1189–1201, Nov. 2018, doi: 10.1007/s00607-018-0638-1.

[10] "Cherry Diseases," 2020. [Online] Available: https://vikaspedia.in/agriculture/crop-production/integrated-pest-managment/ipm-for-fruit-crops/ipm-strategies-for-cherry/cherry-diseases.

[11] J. G. Esguera, M. A. Balendres, and D. P. Paguntalan, "Overview of the *Sigatoka* leaf spot complex in banana and its current management," *Tropical Plants*, vol. 3, no. 1, pp. 0–0, 2024, doi: 10.48130/tp-0024-0001.

[12] M. Fraiwan, E. Faouri, and N. Khasawneh, "Multiclass Classification of Grape Diseases Using Deep Artificial Intelligence," *Agriculture (Switzerland)*, vol. 12, no. 10, p. 1542, Sep. 2022, doi: 10.3390/agriculture12101542.

[13] M. Susai, N. S. A. M. Taujuddin, and S. Sari, "Guava Leaf Disease Detection Using Colour Region Segmentation and Circularity Value Techniques," *Evolution in Electrical and Electronic Engineering*, vol. 4, no. 1, pp. 458–467, 2023, doi: 10.30880/eeee.2023.04.01.054.

[14] L. B. Koppal, T. M. Rajesh, and K. B. Vedamurthy, "A Novel Model for Disease Identification in Mango Plant Leaves Using Multimodal Conventional and Technological Approach," 2021, pp. 131–140, doi: 10.1007/978-981-16-0171-2_13.

[15] S. F. Syed-Ab-Rahman, M. H. Hesamian, and M. Prasad, "Citrus disease detection and classification using end-to-end anchor-based deep learning model," *Applied Intelligence*, vol. 52, no. 1, pp. 927–938, Jan. 2022, doi: 10.1007/s10489-021-02452-w.

[16] S. Durai, T. Sujithra, and M. M. Iqbal, "Image Classification for Potato Plant Leaf Disease Detection using Deep Learning," in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pp. 154–158, 2023, doi: 10.1109/ICSCSS57650.2023.10169446.

[17] K. S. Archana, S. Srinivasan, S. P. Bharathi, R. Balamurugan, T. N. Prabakar, and A. S. F. Britto, "A novel method to improve computational and classification performance of rice plant disease identification," *Journal of Supercomputing*, vol. 78, no. 6, pp. 8925–8945, Apr. 2022, doi: 10.1007/s11227-021-04245-x.

[18] I. Abbas, J. Liu, M. Amin, A. Tariq, and M. H. Tunio, "Strawberry fungal leaf scorch disease identification in real-time strawberry field using deep learning architectures," *Plants*, vol. 10, no. 12, p. 2643, Dec. 2021, doi: 10.3390/plants10122643.

[19] O. Attallah, "Tomato Leaf Disease Classification via Compact Convolutional Neural Networks with Transfer Learning and Feature Selection," *Horticulturae*, vol. 9, no. 2, p. 149, Jan. 2023, doi: 10.3390/horticulturae9020149.

[20] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.

[21] M. Shoaib *et al.*, "An advanced deep learning models-based plant disease detection: A review of recent research," *Frontiers in Plant Science*, vol. 14, Mar. 2023, doi: 10.3389/fpls.2023.1158933.

[22] J. Yao, S. N. Tran, S. Sawyer, and S. Garg, "Machine learning for leaf disease classification: data, techniques and applications," *Artificial Intelligence Review*, vol. 56, no. S3, pp. 3571–3616, Dec. 2023, doi: 10.1007/s10462-023-10610-4.

[23] L. Li, S. Zhang, and B. Wang, "Plant Disease Detection and Classification by Deep Learning - A Review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.

[24] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, Sep. 2016, doi: 10.3389/fpls.2016.01419.

[25] A. Ali, "PlantVillage Dataset," 2019. [Online] Available: https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset.

[26] V. P. Kour and S. Arora, "PlantaeK: A Leaf Database of Native Plants of Jammu and Kashmir," *Lecture Notes in Electrical Engineering*, vol. 832, pp. 359–368, 2022, doi: 10.1007/978-981-16-8248-3_29.

[27] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A dataset for visual plant disease detection," in *ACM International Conference Proceeding Series*, Jan. 2020, pp. 249–253, doi: 10.1145/3371158.3371196.

[28] A. Khan and Dr. Atul Srivastava, "PlantDoc-Plant Disease Detection using AI", *Journal of Informatics Electrical and Electronics Engineering.*, vol. 4, no. 1, pp. 1–10, Apr. 2023, doi: 10.54060/jieee.v4i1.86.

[29] T. Ranjita, K. Zhang, N. Snavely, S. Belongie, and A. Khan, "Plant Pathology 2021 - FGVC8," *Plant Pathology*, 2021. [Online] Available: https://kaggle.com/competitions/plant-pathology-2021-fgvc8.

[30] H. B. Prajapati, J. Shah, and V. Dabhi, "Detection and classification of rice plant diseases," *Intelligent Decision Technologies*, vol. 11, no. 3, pp. 357–373, 2017, doi: 10.3233/IDT-170301.

[31] Divya Chikati, Rajasekhar Reddy Manyam, Ramkumar Kancharla, "An Effective Machine Learning Technique for Rice Leaf Disease Detection", *International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE*, 2023, doi: 10.1109/ICCCIS60361.2023.10425299.

[32] B. JG *et al.*, "Annotated plant pathology databases for image-based detection and recognition of diseases," *IEEE Latin America Transactions*, vol. 16, no. 6, pp. 1749–1757, 2018, doi: 10.1109/TLA.2018.8444395.

[33] Uzhinskiy. A, Ososkov. G, Goncharov. P, Nechaevskiy. A., "Multifunctional platform and mobile application for plant disease detection". *In Proceedings of the 27th International Symposium Nuclear Electronics and Computing (NEC'2019)*, Budva, Montenegro, 30 September–4 October 2019.

[34] S. Bhattarai, "New Plant Diseases Dataset," 2018. [Online] Available: https://www.kaggle.com/vipoooool/new-plant-diseases-dataset/.

[35] T. Wiesner-Hanks *et al.*, "Image set for deep learning: Field images of maize annotated with disease symptoms," *BMC Research Notes*, vol. 11, no. 1, 2018, doi: 10.1186/s13104-018-3548-6.

[36] J. Sun, Y. Yang, X. He, and X. Wu, "Northern Maize Leaf Blight Detection Under Complex Field Environment Based on Deep Learning," *IEEE Access*, vol. 8, pp. 33679–33688, 2020, doi: 10.1109/access.2020.2973658.

[37] H. T. Rauf, B. A. Saleem, M. I. U. Lali, M. A. Khan, M. Sharif, and S. A. C. Bukhari, "A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning," *Data in Brief*, vol. 26, 2019, doi: 10.1016/j.dib.2019.104340.
[38] M.-Ling Huang and Y.-An Chen, "Citrus dataset for image classification," *Data in Brief*, vol. 52, p. 109628, 2023, doi: 10.1016/j.dib.2023.109628.
[39] H. Minh Do, "Rice Diseases Image Dataset," *Kaggle*, 2019. [Online] Available: https://www.kaggle.com/datasets/minhhuy2810/rice-diseases-image-dataset.
[40] J. Jepkoech, D. M. Mugo, B. K. Kenduiywo, E. C. Too, "Arabica coffee leaf images dataset for coffee leaf disease detection and classification," *Data in Brief*, vol. 36, 2021, doi: 10.1016/j.dib.2021.107142.
[41] G. Owomugisha, J. Nakatumba-Nabende, J. J. Dhikusooka, E. Taravera, E. Nuwamanya, and E. Mwebaze, "A labeled spectral dataset with cassava disease occurrences using virus titre determination protocol," *Data in Brief*, vol. 49, 2023, doi: 10.1016/j.dib.2023.109387.
[42] V. IYER, "Cassava Image Dataset," *Kaggle*, 2022. [Online] Available: https://www.kaggle.com/datasets/visalakshiiyer/cassava-image-dataset.
[43] J. Savitha, M. N, S. R. D, and V. R, "Utilizing Machine Learning to find Plant Leaf Disease," *International Journal of Innovative Research in Technology*, vol. 10, no. 3, pp. 1–7, 2023.
[44] P. V. Shetkar, A. Jagdale, and S. Potkule, "Leaf Disease Detection in Plants Using Aiml and Image Preprocessing Techniques," *International Research Journal of Modernization in Engineering Technology and Science*, no. 05, pp. 5–8, 2023.
[45] P. Kaur and B. Kaur, "A Hybrid Approach to Identify Plant Leaf Disease using Machine Learning and Image Processing," *International Journal of Computer Applications*, vol. 185, no. 27, pp. 25–30, 2023, doi: 10.5120/ijca2023923020.
[46] K. Joshi, R. Awale, S. Ahmad, S. Patil, and V. Pisal, "Plant Leaf Disease Detection Using Computer Vision Techniques and Machine Learning," *ITM Web of Conferences*, vol. 44, p. 03002, 2022, doi: 10.1051/itmconf/20224403002.
[47] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, 2022, doi: 10.1016/j.gltp.2022.03.016.
[48] V. Vineeth, P. Rithwik, M. Neha, R. Vaishnai, Smt. N. Madhuri, "Plant Leaf Disease Detection Using Naive Bayes ML Algorithm," *International Journal of Emerging Technologies and Innovative Research (www.jetir.org)*, vol.10, no. 4, 2023.
[49] K. P. Panigrahi, H. Das, A. K. Sahoo, and S. C. Moharana, "Maize Leaf Disease Detection and Classification Using Machine Learning Algorithms," in *Advances in Intelligent Systems and Computing*, vol. 1119, 2020, pp. 659–669, doi: 10.1007/978-981-15-2414-1_66.
[50] B. Hatuwal, B. Joshi, B. K. Hatuwal, and A. Shakya, "Plant Leaf Disease Recognition Using Random Forest, KNN, SVM and CNN," *Polibits*, vol. 62, pp. 13–19, 2021. [Online]. Available: https://doi.org/10.17562/PB-62-2.
[51] A. Selvakumar and A. Balasundaram, "Machine Learning based Classification of Diseased Mango Leaves," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 10, no. 7, pp. 38–44, Jul. 2022, doi: 10.17762/ijritcc.v10i7.5563.
[52] S. Sengupta, A. Dutta, S. A. M. Abdelmohsen, H. A. Alyousef, and M. Rahimi-Gorji, "Development of a Rice Plant Disease Classification Model in Big Data Environment," *Bioengineering*, vol. 9, no. 12, 2022, doi: 10.3390/bioengineering9120758.
[53] K. Ashritha, K. Sandhya, Y. U. Kiran, and V. N. L. N. Murthy, "Plant-Leaf Disease Prediction Using Deep Learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, no. 3, pp. 121–128, Mar. 2023, doi: 10.22214/ijraset.2023.49338.
[54] B. P. Sarita Das, Bhabani Shankar Prasad Mishra, Abhinandan Kumar Tiwar, "Optimal Trained Deep LSTM Model for Detecting Plant Leave Diseases," *International Journal of Engineering Research and Technology (IJERT)*, vol. 12, no. 12, 2023, doi: 10.17577/IJERTV12IS120028.
[55] K. V. Deputy, K. Passi, and C. K. Jain, "Crop Disease Detection Using Deep Learning Techniques on Images," *Journal of Computer Science*, vol. 19, no. 12, pp. 1438–1449, 2023, doi: 10.3844/jcssp.2023.1438.1449.
[56] S. Padshetty and Ambika, "Leaky ReLU-ResNet for Plant Leaf Disease Detection: A Deep Learning Approach †," in *Engineering Proceedings*, Dec. 2023, vol. 59, no. 1, p. 39, doi: 10.3390/engproc2023059039.
[57] R. S. Krishnan and E. G. Julie, "Computer aided detection of leaf disease in agriculture using convolution neural network based squeeze and excitation network," *Automatika*, vol. 64, no. 4, pp. 1038–1053, Oct. 2023, doi: 10.1080/00051144.2023.2241792.
[58] V. Balaji *et al.*, "Deep Transfer Learning Technique for Multimodal Disease Classification in Plant Images," *Contrast Media & Molecular Imaging*, vol. 2023, pp. 1–8, May 2023, doi: 10.1155/2023/5644727.
[59] M. Nawaz *et al.*, "A robust deep learning approach for tomato plant leaf disease localization and classification," *Scientific Reports*, vol. 12, no. 1, p. 18568, Nov. 2022, doi: 10.1038/s41598-022-21498-5.
[60] A. Patani, I. Pardikar, P. Doshi, S. Rodge, and S. Khachane, "Plant Leaf Recognition and Disease Detection Using GoogLeNet," *Journal of Emerging Technologies and Innovative Research*, vol. 8, no. 5, pp. 595–602, 2021.
[61] H. C. Chen *et al.*, "AlexNet Convolutional Neural Network for Disease Detection and Classification of Tomato Leaf," *Electronics (Switzerland)*, vol. 11, no. 6, p. 951, Mar. 2022, doi: 10.3390/electronics11060951.
[62] J. Andrew, J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, "Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications," *Agronomy*, vol. 12, no. 10, p. 2395, Oct. 2022, doi: 10.3390/agronomy12102395.
[63] K. Letrache and M. Ramdani, "Explainable Artificial Intelligence: A Review and Case Study on Model-Agnostic Methods," in *Proceedings - SITA 2023: 2023 14th International Conference on Intelligent Systems: Theories and Applications*, Nov. 2023, pp. 1–8, doi: 10.1109/SITA60746.2023.10373722.
[64] R. Rajkumar and M. V Sudhamani, "Content based Image Retrieval System using Combination of Color and Shape Features, and Siamese Neural Network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 2S, pp. 71–77, Dec. 2019, doi: 10.35940/ijitee.b1053.1292s19.
[65] A. Saranya and R. Subhashini, "A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends," *Decision Analytics Journal*, vol. 7, p. 100230, Jun. 2023, doi: 10.1016/j.dajour.2023.100230.
[66] S. Ali *et al.*, "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Information Fusion*, vol. 99, p. 101805, Nov. 2023, doi: 10.1016/j.inffus.2023.101805.
[67] M. H. K. Mehedi *et al.*, "Plant Leaf Disease Detection using Transfer Learning and Explainable AI," *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2022*, pp. 166–170, 2022, doi: 10.1109/IEMCON56893.2022.9946513.
[68] M. Bhandari, T. B. Shahi, A. Neupane, and K. B. Walsh, "BotanicX-AI: Identification of Tomato Leaf Diseases Using an Explanation-Driven Deep-Learning Model," *Journal of Imaging*, vol. 9, no. 2, p. 53, Feb. 2023, doi: 10.3390/jimaging9020053.
[69] K. Wei *et al.*, "Explainable Deep Learning Study for Leaf Disease Classification," *Agronomy*, vol. 12, no. 5, 2022, doi: 10.3390/agronomy12051035.

[70] S. Kinger and V. Kulkarni, "Explainable AI for deep learning based disease detection," in *ACM International Conference Proceeding Series*, Aug. 2021, pp. 209–216, doi: 10.1145/3474124.3474154.
[71] C. S. Arvind *et al.*, "Deep Learning Based Plant Disease Classification with Explainable AI and Mitigation Recommendation," in *2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings*, Dec. 2021, pp. 01–08, doi: 10.1109/SSCI50451.2021.9659869.
[72] V. Mistry, "Explainable AI(XAI) using LIME and Disease Detection in Mango Leaf by Transfer Learning Approach," *International Research Journal of Engineering and Technology*, vol. 11, no. 02, pp. 72–76, 2024.
[73] S. B. Bidouk, E. Zagrouba, and A. Bigand, "Explainable AI for deep learning-based potato leaf disease detection," *IEEE International Conference on Fuzzy Systems*, 2023.
[74] M. Nahiduzzaman *et al.*, "Explainable deep learning model for automatic mulberry leaf disease classification," *Frontiers in Plant Science*, vol. 14, Sep. 2023, doi: 10.3389/fpls.2023.1175515.
[75] R. Bandi, S. Swamy, and C. S. Arvind, "Leaf disease severity classification with explainable artificial intelligence using transformer networks," *International Journal of Advanced Technology and Engineering Exploration*, vol. 10, no. 100, pp. 278–302, Mar. 2023, doi: 10.19101/IJATEE.2022.10100136.
[76] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, vol. 11, no. 1, p. 5, Jan. 2024, doi: 10.1186/s40537-023-00863-9.
[77] A. Bajaj, "Performance Metrics in Machine Learning," *Neptune Blog*, 2023. [Online] Available: https://neptune.ai/blog/performance-metrics-in-machine-learning-complete-guide
[78] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*, 2023, doi: 10.1109/ICCCNT56998.2023.10306417.
[79] Y. Pu *et al.*, "Variational autoencoder for deep learning of images, labels and captions," *Advances in Neural Information Processing Systems*, pp. 2360–2368, 2016, doi: 10.48550/arXiv.1609.08976.
[80] D. Oppenheim, G. Shani, O. Erlich, and L. Tsror, "Using deep learning for image-based potato tuber disease detection," *Phytopathology*, vol. 109, no. 6, pp. 1083–1087, Jun. 2019, doi: 10.1094/PHYTO-08-18-0288-R.
[81] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, Jun. 2019, doi: 10.1016/j.compag.2018.03.032.
[82] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Computers and Electronics in Agriculture*, vol. 173, p. 105393, Jun. 2020, doi: 10.1016/j.compag.2020.105393.
[83] S. Zhang, W. Huang, and C. Zhang, "Three-channel convolutional neural networks for vegetable leaf disease recognition," *Cognitive Systems Research*, vol. 53, pp. 31–41, Jan. 2019, doi: 10.1016/j.cogsys.2018.04.006.
[84] B. Liu, Z. Ding, L. Tian, D. He, S. Li, and H. Wang, "Grape Leaf Disease Identification Using Improved Deep Convolutional Neural Networks," *Frontiers in Plant Science*, vol. 11, Jul. 2020, doi: 10.3389/fpls.2020.01082.
[85] R. Thangaraj, S. Anandamurugan, and V. K. Kaliappan, "Automated tomato leaf disease classification using transfer learning-based deep convolution neural network," *Journal of Plant Diseases and Protection*, vol. 128, no. 1, pp. 73–86, Feb. 2021, doi: 10.1007/s41348-020-00403-0.
[86] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," *Applied Soft Computing Journal*, vol. 86, p. 105933, Jan. 2020, doi: 10.1016/j.asoc.2019.105933.
[87] F. Martinelli *et al.*, "Advanced methods of plant disease detection. A review," *Agronomy for Sustainable Development*, vol. 35, no. 1, pp. 1–25, Jan. 2015, doi: 10.1007/s13593-014-0246-1.
[88] J. Liu and X. Wang, "Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model," *Plant Methods*, vol. 16, no. 1, p. 83, Dec. 2020, doi: 10.1186/s13007-020-00624-2.
[89] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.
[90] K. KC, Z. Yin, M. Wu, and Z. Wu, "Depthwise separable convolution architectures for plant disease classification," *Computers and Electronics in Agriculture*, vol. 165, p. 104948, Oct. 2019, doi: 10.1016/j.compag.2019.104948.

BIOGRAPHIES OF AUTHORS

Kusuma R received B.E. in Computer Science and Engineering and M.Tech. in Computer Science and Engineering from VTU, Belagavi during 2008 and 2011 respectively. She is currently full time Research Scholar in RNSIT Bengaluru, Karnataka, India. She has 10 Years of academic experience. Her research area includes Machine Learning, Deep Learning, Computer Vision, and explainable Artificial Intelligence. She has been awarded as "Wipro Certified Faculty for project-based learning in Java J2EE". Her publication includes 1 International Conference paper and 4 National Conference paper. One patent has been filled by her. She was Resource person for three FDPs. She can be contacted at email: 1rn22pcs04.kusuma@rnsit.ac.in.

R. Rajkumar is having 21+ years of experience in the field of academics and Associate Professor in CSE-CY in RNSIT. He has Ph.D. degree from VTU for his research on Image Processing & Deep Learning Techniques. His area of interest includes Cyber security, Machine learning, Deep Learning, Computer Vision and Artificial intelligence. He has presented and published more than 15 papers in International Conferences and Journals. One patent has been filled by him. He can be contacted at email: rajkumar.r@rnsit.ac.in.