An Hybrid Feature Extraction and Classification using Xception-RF for Multiclass Disease Classification in Plant Leaves

Abisha A & Bharathi N

Published online: 22 Feb 2023.

Submit your article to this journal

Article views: 822

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

An Hybrid Feature Extraction and Classification using Xception-RF for Multiclass Disease Classification in Plant Leaves

Abisha A 🆔 and Bharathi N 🆔

Department of Computer Science and Engineering, SRM Institute of Science and Technology, College of Engineering and Technology, Chennai, India

ABSTRACT

In today's world, agriculture is one of the most important assets as it affects the economy of the country. Biotic stress caused by fungi, bacteria, and viruses is common in plants; hence plant disease recognition is essential as it could identify diseases at the early stages of infection. Most of the previous plant disease works involve PlantVillage dataset, which is taken in a controlled environment. In this paper, the Banana leaf dataset is used which is combined from two datasets. CNNs are becoming more popular in many forms, such as pre-trained models, and in this research, we employed a pre-trained model in conjunction with an ML classifier to predict plant disease. Pre-trained models such as VGG 16, RESNET, MobileNet, InceptionResNetV2, Inception, and Xception models were deployed for feature extraction. Based on the above models, the proposed Xception-RF model achieved 100% accuracy with no misclassified classes. Using an Xception model with depth-wise convolutions reduces the number of parameters and the cost of implementation. The proposed work merges machine learning and deep learning by extracting features using CNN and then using ML-based RF for classification, which has resulted in an increase in performance in terms of accuracy.

Introduction

In a world with more than seven billion people and more than 90% lacking access to tools or features that would diagnose and address the issue, crop diseases are currently a major threat to the world's food supplies. To comprehend India's economy in terms of agricultural productivity, plant disease detection is essential. Plant diseases must be identified and classified as soon as possible since they can harm a species' ability to grow and thrive. Since trained CNN models can distinguish between certain items even before being deployed in this plant disease application, they perform more accurately and with higher quality.

CONTACT Abisha A ✉ aa7111@srmist.edu.in Department of Computer Science and Engineering, SRM Institute of Science and Technology, College of Engineering and Technology, Vadapalani Campus, Chennai, India

Although many algorithms that have been evolving, there is no proper algorithm to justify the best technique. In previous works, the authors have worked on datasets taken from controlled environments that are not suitable for real-time image processing. Hence, here we have combined two datasets that consist of banana leaf images taken under a controlled environment as well as real-time images. Hence, here some combinations of pre-trained models have been introduced to boost the existing research methods and have proven that Xception-RF has performed well compared to the other models. The process used here is that the layers of Xception model are used for the feature extraction process and RF is used as a classifier.

Related work

Combining techniques from transfer learning and deep feature extraction, Mohameth, Bingcai, and Amath Sada (2020) make use of CNN architectures. SVM and KNN are used to categorize all the collected features. Free and available to anyone, the PlantVillage Dataset is put to good use here. Based on the data, it was determined that SVM is the most effective classifier for identifying illnesses in leaves. Abisha and Bharathi (2021) studied and highlighted the detection of numerous defects, disorders, and methods employing ML, DL, and ANN methodologies. Additionally, a thorough analysis is emphasized to assist in identifying the gaps and allow us to come up with an improved key for the improvement of plant health. This study provides an overview of plant health and the stresses as well as a critique of previous research on the use of deep learning, machine learning, and artificial neural networks (ANN) to identify, categorize, and forecast various plant health issues.

This research (Lili, Zhang, and Wang 2021) discusses the state of the art in the detection of plant leaf disease utilizing deep learning and state-of-the-art imaging techniques, as well as the opportunities and threats that lie ahead. This work gives a useful insight into deep learning techniques. The most recent CNN networks with relevance to plant leaf disease classification are discussed in the study by Lu, Tan, and Jiang (2021). CNNs have shown great results in machine vision, therefore Hassan et al. (2021) use CNNs to recognize and diagnose plant illnesses from their leaves. Scarpa et al. (2018) suggest using data fusion and deep learning to estimate missing visual characteristics. The use of optical sequences, SAR sequences, and digital elevation models is considered to take advantage of cross-sensor and temporal interdependence. Different fusion systems, both causal and non-causal, single-sensor or joint-sensor, are taken into consideration. According to all performance measures, experimental results are quite encouraging, demonstrating a considerable improvement compared with baseline techniques.

The authors of the aforementioned (Lamba, Gigras, and Dhull 2021) research proposed a model for plant disease detection that uses Auto-Color Correlogram as an image filter and DL as classifiers with various activation functions. This suggested model is used to address binary and multiclass subclasses of plant diseases across four datasets. The proposed model has been shown to achieve better results than the state-of-the-art methods LibSVM, SMO (sequential minimal optimization), and DL (deep learning) with activation functions softmax and soft sign in terms of the F-measure, recall, MCC (Matthews correlation coefficient), specificity, and sensitivity. In the study by Toda and Okura (2019), the CNN was used with an openly accessible dataset of images depicting plant diseases and was used to apply a range of neuron-wise and layer-wise visualization techniques. It demonstrated how neural networks, when used for diagnosis, can take the hues and textures of the diseased area particular to a given disease, mimicking human judgment. The parameters were reduced to 75% without compromising classification accuracy by identifying several layers inside the network that were not contributing to inference by evaluating the attention maps created.

An efficient loss-fused convolutional neural network model is put forth in (Gokulnath and Usha Devi 2021) to identify the plants suffering from a certain sort of sickness. This approach improves prediction by combining the benefits of two separate loss functions. In the last layer of the model, the diseases were categorized using the attributes that were collected from the plant leaves. When separating the damaged samples from the unaffected ones, this technique achieved an accuracy of 98.93%. (Kumar, Chaudhary, and Khaitan Chandra 2021) concentrates on identifying plant diseases and minimizing financial damages. A deep learning-based technique for image recognition is suggested in the study. The system suggested in the research can handle complex situations and is effective at detecting various illness types. The validation results demonstrate the Convolution Neural Network's viability and suggest a route for an AI-based Deep Learning solution to this complex problem with an accuracy of 94.6%.

Using pre-trained convolutional neural networks (CNNs) like AlexNet, GoogleNet, VGG16, ResNet101, and DensNet201, an effective method for identifying soybean diseases has been described in (Jadhav, Udupi, and Patil 2019). To recognize three soybean illnesses from healthy leaves, the CNNs were trained by 1200 PlantVillage dataset images with infected and healthy leaves. The CNN, which has already been trained, is employed to make system implementation quick and simple in practice. In this study, feature extractors and classifiers were created using a pre-trained CNN. (Zhao and Liu 2020) have suggested a system that uses CNN for FE using the MNIST dataset. Mathematical and algebraic models were deployed for feature selection to select feature sets obtained from FE. Classifier fusion was carried out, which gave an accuracy above 98%. According to the experimental findings,

a classifier's fusion can produce a classification accuracy of about 98%. In a study by Swinney and Woods (2021), the impact of the concurrent factual digital world with Bluetooth and Wi-Fi signal intrusion on the feature extraction and ML categorization of UASs is assessed. The paper evaluated several UASs that utilize various communication technologies. The study also demonstrates how to distinguish between UASs that are using the same communication technology.

The likelihood of a fusion classification technique dubbed CNN and RF, which syndicates the reflex FE capabilities of CNNs with the improved discernment capabilities, is examined in the context of early crop mapping by (Kwak et al. 2021). Two experiments on incremental crop categorization using images from unmanned aerial vehicles have been led to evaluate the algorithm combined with CNN and RF. The experimental results demonstrate that CNN-RF, even with a limited number of input images and training samples, is a useful classifier for early crop mapping.

Vijayalakshmi and Joseph Peter (2021) suggest a 5-layered CNN for banana recognition that is made up of a convolution layer, a pooling layer, and a fully connected layer. Using a CNN system built on deep learning, numerous features were retrieved from apples, strawberries, oranges, mangoes, and bananas. Experiments were carried out on 5887 fruit image data in our database. The accuracy rate for the deep feature-random forest classification combination procedure has been achieved at 96.98%, which has performed better compared with the other deep feature algorithms with KNN and RF algorithms.

(Singh et al. 2020) suggests using neural networks and visualization technology to identify Android malware (SARVOTAM). To retrieve valuable data, the technology transforms the malware's confusing traits into fingerprint photos. By automatically extracting rich features from displayed malware, a tailored Convolutional Neural Network (CNN) reduces the cost of feature engineering and domain expertise. The CNN-SVM model outpaced the basic CNN model and the other hybrid models.

In a study by Xie et al. (2022), the fault detection effectiveness and efficiency of four hybrid models built on CNNs have been surveyed. The comparison analysis revealed that the SVM and RF could completely utilize the CNN FE capacity. Using DL and ML techniques, Gayathri, Gopi, and Palanisamy (2021) suggest a computerized DR classifying approach in which features from the fundus images can be retrieved and categorized according to severity. Here, multipath CNN(MCNN) is implemented. The input is then categorized based on severity using a machine learning classifier. The results of the studies demonstrate that the M-CNN network with the J48 classifier generates the best response. The classifiers are assessed using pre-trained network characteristics as well as currently used DR grading techniques and the suggested work has a 99.62% average.

To classify crops in multi-temporal optical remote sensing photos, Yang et al. (2020) used the fusion CNN-RF technique with optimum feature selection (OFSM). To recognize crop types in the northeast Chinese area of Jilin, researchers evaluated 234 variables from three scenes of Sentinel-2 pictures, including spectral, segmentation, color, and texture features. The OFSM method has been compared with traditional methods such as feature selection with RF and other methods and the OFSM method has been proven to be effective.

(Li et al. 2019) proposed a deep cube CNN model called DCNR, using deep CNN, a random forest classifier, and a cube neighbor HSI pixel method. In the DCNR model, each target pixel and its neighbors are combined to create cubic samples that contain spectral-spatial information. Following that, the cube CNN model is used to extract features with high representational ability. The DCNR model has outperformed other models tested.

DL-based FE and classification are used to propose a new KOA detection at an initial stage in (Mahum et al. 2021). After preprocessing the raw X-ray pictures, segmentation is used to extract the Region of Interest (ROI). There are four grades in the Kellgren-Lawrence system ranging from grade I to grade IV. Different combinations of the suggested framework are tested experimentally. The experimental results demonstrate that the HOD feature descriptor gives the highest accuracy.

Abisha and Bharathi (2022) divide the data into six cases and consider two classifications, healthy and infected. Hu moments, Haralick texture, and color histogram are the feature descriptors to extract features from plant photos. Following feature extraction, the images are categorized as healthy or contaminated using feature descriptors. The retrieved data are validated after training using different classifier algorithms such as LR, CART, KNN, LDA, and RF. Compared with the other models, RF, CART, and KNN have each demonstrated superiorly well.

The above study describes various works that work in standalone mode as well as in hybrid. A summary of related works is detailed in Table 1. Although there are various methods proposed to enhance the performance, the novel method proposed in this paper has achieved 100% accuracy, which is not seen in any other works described above for plant health applications. The integration of Xception and RF methods is also a novel method, which has not been proposed in any other works so far. This paper emphasizes the following:

(1) A Novel method Xception-RF that integrates ML and DL.
(2) Feature extraction using the Xception model that can enhance the feature extraction process with 100% accuracy.

Table 1. A comparison of studies pertaining to plant classification and algorithms.

S.No	Reference	Models Used	Database	Performance and Limitation
1	Gayathri, Gopi, and Palanisamy (2021)	Multipath Convolutional Neural Network (M-CNN) with e J48 classifier	(1) IDRID (2) Kaggle (3) MESSIDOR	This work achieved a high level of accuracy of 99.62%. Some limitations include taking up more time and overfitting. There are no detailed dataset name links or descriptions available.
2	(Lamba, Gigras, and Dhull 2021)	Auto-Color Correlogram as image filter, DL as classifiers	Rice dataset and the Kaggle PlantVillage dataset.	The accuracy for binary class is 99.4%, while the accuracy for multiclass is 99.2%. Training DL classifiers takes time, depending on the number of epochs.
3	(Toda and Okura 2019)	Convolutional neural networks (CNNs)	PlantVillage dataset	99% have been achieved. Because the PlantVillage dataset was obtained in a controlled environment, the model must be validated using alternative datasets or real-time data.
4	(Abdu, Mohd Mokji, and Ullah Sheikh 2020)	Color coherence vector (CCV) and local binary pattern (LPB) for FE with Machine learning classifiers	PlantVillage	This work has given an accuracy of 99%. As PlantVillage dataset has been collected in a controlled environment, the model has to be tested with other datasets or real-time data.
5	(Gokulnath and Usha Devi 2021)	LF-CNN	PlantVillage Database	This work provides an accuracy of 98.93%. PlantVillage dataset was taken in a controlled environment and it is a commonly used dataset. The main limitation is that the model isn't been tested under different conditions such as illumination, occlusion, pose variation, lightning factors etc.
6	(Zhao and Liu 2020)	CNN	MINST dataset	The experimental results show that the classifiers fusion can achieve the classification accuracy of ≥98%. CNN classifiers slow down the training time with respect to number of epochs. Hence ML hybrid methods could be considered.
7	(Hassan et al. 2021)	Deep convolutional-neural-network (CNN)	PlantVillage dataset	The performance accuracy was found to be 98% and above. The convolutional neural network classifier doesn't give higher performance.
8	(Mohameth, Bingcai, and Amath Sada 2020)	CNN-SVM	PlantVillage	An accuracy of 97.82% have been attained. Collecting the own dataset and comparing it to the PlantVillage dataset will provide best efficiency of the model.
9	(Mahum et al. 2021)	CNN through Local Binary Patterns (LBP) and Histogram of oriented gradient (HOG) with SVM, RF, and KNN	Mendeley dataset IV	Achieved an accuracy of 97%. LBP and HOG is ineffective at capturing discriminating characteristics which slow down the recognition speed especially on large-scale face database.
10	(Vijayalakshmi and Joseph Peter 2021)	Convolution neural network (CNN)	Fruits 360 dataset	Have achieved an accuracy of 96.98%. Could infuse ML classifiers as it is fast and increases the performance.

(*Continued*)

Table 1. (Continued).

S.No	Reference	Models Used	Database	Performance and Limitation
11	(Li et al. 2019)	DCNR: deep cube CNN with random forest	Images from KSC (Kennedy Space Center), PU (Pavia University) and IP (Indian Pines)	96.78%, 96.08% and 94.85% accuracy was obtained on KSC, IP and SA datasets. As DCNR is trained with random sampling instead of active learning may lead to the loss of certain key pixels. Moreover, the time cost in DCNR training is much higher than CCS model.
12	(Jadhav, Udupi, and Patil 2019)	AlexNet, GoogleNet, VGG16, ResNet101, and DensNet201	PlantVillage database	95%, 96.4%, 96.4%, 92.1%, 93.6%. Attempt to develop our own CNN model in order to increase the performance rate of the model by varying the minibatch size, bias learning rate, and weight.
13	(Kumar, Chaudhary, and Khaitan Chandra 2021)	Neural Network: Faster Region-based CNN (Faster R-CNN), Region-based Fully CNN(R-CNN) and Single shot Multibook Detector (SSD)	PlantVillage Dataset	Accuracy of 94.6% have been attained but the network may take a long time to attain convergence because all samples from a single image may be correlated (i.e. their features are similar).
14	(Yang et al. 2020)	CNN-RF	Sentinel-2 images	94.27% have been obtained. Performance could be improved. Because clouds can disrupt Sentinel-2 data, future work could focus on establishing a multi-source remote sensing imagery fusion strategy for crop categorization.
15	(Li, Nie, and Chao 2020)	SCNN-KSVM (Shallow CNN with Kernel SVM) and SCNN-RF (Shallow CNN with RF)	Maize dataset, Apple dataset and Grape dataset from PlantVillage datset	This work obtaines an accuracy of 94%. As PlantVillage is commonly used dataset, and taken under controlled environments, the accuracy could have been improved.
16	(Singh et al. 2020)	Convolutional Neural Network (CNN)- Support Vector Machine (SVM)	DREBIN dataset	92.59% accuracy has been achieved. The scope of this experiment was limited to evaluating the performance of the proposed model using malware images. It could be tested with another dataset to know the model's actual performance.
17	(Scarpa et al. 2018)	Convolutional neural network (CNN)	Sentinel-1, Sentinel-2, SAR image	90% obtained from the average of five months. Accuracy could be increased by changing the classifier to ML.
18	(Barbedo 2019)	CNN	Manually collected images(OWN).	Above 75% in most of the cases. Requires more time and performsnce could be improved.
19	(Kwak et al. 2021)	CNN-random forest (CNN-RF)	UAV images using eBee unmanned aerial system	Overall accuracy increased by maximum 6.7%p and 4.6%p in two study areas. As no proper accuracy value have been mentioned, the exact performance is unknown.
20	(Simonyan and Zisserman 2014)	Convolutional networks (ConvNets)	ILSVRC-2012 dataset	Have provided performance results as top-1 validation error of 23.7, top-5 validation error of 6.8 and top-5 test error of 6.8. Could be tried with other datasets and could be explicit in displaying the accurate accuracy/performance of the model.
21	(Xie et al. 2022)	Convolutional neural network (CNN), SVM, RF	Case Western Reserve University (CWRU) bearing fault datasets	Increase of about 1–2% accuracy compared to traditional methods. Up to 99% in some cases. Deep learning models with high interpretability could be considered.

Methodology

Data and image pre-processing

Generally, input data could be taken as images, CSV files, etc. In this paper, images are taken as input. As the PlantVillage dataset is one of the most commonly used datasets taken under controlled environments. Here, we have used banana leaf images as datasets collected and combined from open-source platforms that consist of real-time images and ground truth images. It consists of 1600 images from the dataset Medhi and Deb (2022) and combined with images from Al Mahmud (2021). All the images are being resized to the size of 299×299. Sample images of the dataset are shown in Figure 1.

Feature extraction

Feature extraction is an important step in any image-processing application. CNN provides better pattern recognition results. Hence, feature extraction is done here using CNN-pre-trained models. VGG 16, RESNET, MobileNet, InceptionResNetV2, Inception, and Xception have been deployed and the Xception model outperformed other models. The classification was done using the random forest algorithm. The proposed work is a combination of Xception and RF algorithms. Agricultural production and quality are analyzed by Li, Nie, and Chao (2020). CNNs have been the main focus of this study, and they have effectively identified the deep CNN. The results provided that CNN has been more efficient for image processing. Hence, this work has proposed two hybrid CNNs, namely SCNN-KSVM (Shallow CNN with Kernel SVM) and SCNN-RF (Shallow CNN with Random Forest). Similarly, we have also proposed the hybrid Xception-RF that combines pre-trained Xception with the RF algorithm. The Xception uses depthwise separability, which helps to effectively identify the features, and RF is itself an ensemble algorithm, which makes it the most accurate ML algorithm.

Vgg 16

A CNN with 16 layers is termed VGG-16. Millions of images were pre-trained and formed as a database called as ImageNet database. These types of pre-trained models can distinguish many types of images like different objects, such as cats, dogs, balls, and bats. The vgg16 command will return an ImageNet-trained VGG-16 network on successful completion. Simonyan and Zisserman (2014) presented the VGG hypothesis. The deep CNN network VGG-16 was presented for publication in the year 2015. The Visual Geometry Group (VGG) has uncovered an effect that includes the depth of the CNN in terms of the accuracy of the results. They have implemented an architecture with small convolution filters of size 3×3, and it demonstrates some visible
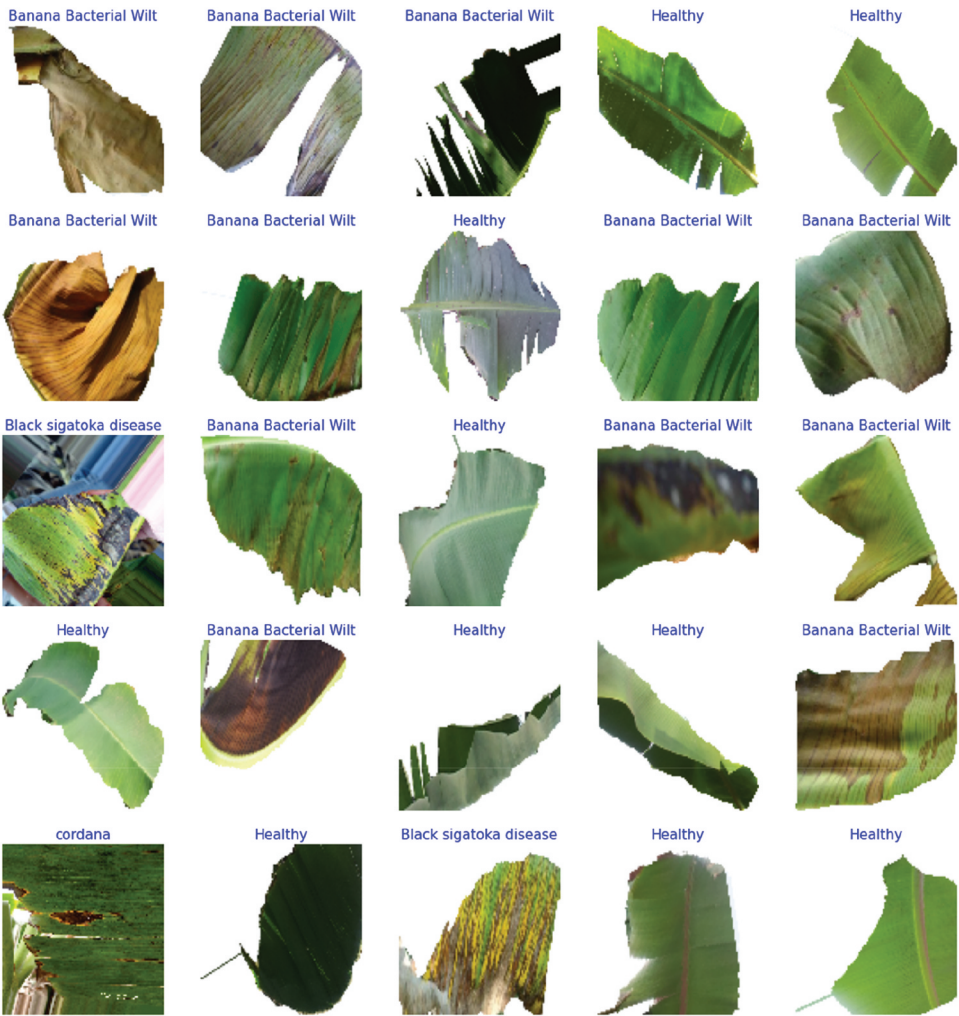
Figure 1. Sample images from the dataset.

results in the state-of-the-art settings. This configuration was entered into the ImageNet Contest in 2014 and was recognized for its excellence.

Reset 50

In 2015, specialists from Microsoft Research came up with the idea of ResNet, which resulted in the creation of an innovative architectural concept known as Residual Network. The problem of the vanishing or exploding gradients inspired this design to come up with the concept of the residual blocks, which was then implemented. To connect the layers to the activations, this model makes use of a method that is known as skip connections. This method skips some levels in the model. This results in an extra block being left over. The resnets are formed by stacking the remaining blocks in this case.

Theckedath and Sedamkar (2020) describe the short-form ResNet50 for residual networks that include 50 layers. It is quite similar to VGG-16, with the exception that Resnet50 possesses an additional capacity for identity mapping. ResNet can accurately forecast the delta that must be achieved to arrive at the final prediction when moving from one layer to the next [20]. By enabling this alternative, more direct route for the gradient to flow through, ResNet can mitigate the problem of vanishing gradients. Because of the way identity mapping works in ResNet, the model can skip over a CNN weight layer if it determines that the layer in question is not required. This contributes to the problem of overfitting the training set not be avoided as a result. There are 50 layers in ResNet50.

MobileNet

The MobileNet structure is built with depthwise separable convolutions, but the first layer is skipped over in the process. This is done in the same way as it is explained in (Howard et al. 2017). It can conduct an examination into the topologies of networks in a short amount of time to discover a good network because the network is defined using such basic words. The final fully connected layer, which does not contain any nonlinearities and instead feeds into a softmax layer for classification, is the only exception to this rule. Every other layer is followed by a batch norm and a ReLU nonlinearity.

Inception and inception resnetv2

In the past, Inception models were trained in a partitioned manner such that the complete model could be retained in memory. This allowed the memory to hold the model in its entirety. During the training process, this meant that each replica was partitioned into several separate, more manageable networks. Residual connections, as opposed to filter concatenation, are utilized in networks of the type Inception-ResNet-v2. According to the criteria outlined, both of these approaches are successful (Christian et al. 2017).

Proposed Xception-RF hybrid method

Image pre-processing, feature extraction, and classification are the modules that are used in the work, and it has been explained in detail below. The use of Xception net, which is a pre-trained model, is the advantage and helps to improve the model performance by automatically giving out the features needed for classification. It uses pointwise convolution followed by depthwise convolution, which is the boon of the feature extraction process. The overall workflow diagram of the paper is shown in Figure 2, and it shows the overall models and methods used in this paper. Figure 3 demonstrates the actual proposed work of the paper.
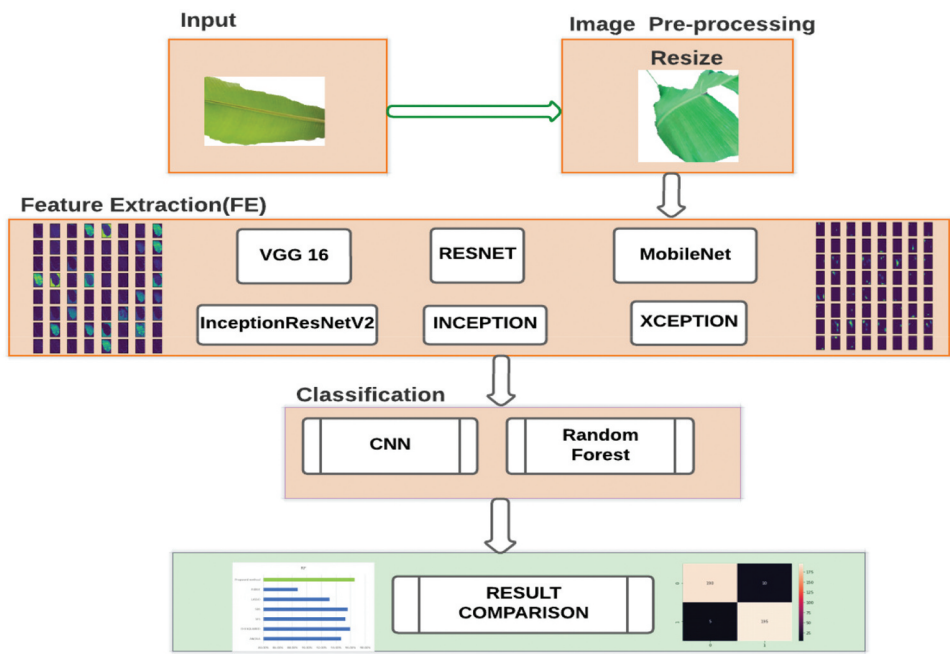
Figure 2. Overall workflow diagram.

Xception

The Xception model rests entirely on the depthwise separable convolution layers as its primary underlying structure as shown in Figure 4.

Inception modules, which are found in convolutional neural networks, are said to represent a transitional stage between the depthwise separable convolution operation and the ordinary convolution, as stated in (Chollet 2017). When viewed in this light, a depthwise separable convolution can be interpreted as an Inception module consisting of the greatest possible number of towers. Based on this observation, a ground-breaking new architecture for deep convolutional neural networks that involve the replacement of Inception modules with depthwise separable convolutions is found. Howard et al. (2017) show that this architecture, which they call Xception, performs significantly better than Inception V3 on a larger image classification dataset that includes 350 million images and 17,000 classes. In our work, we have used an Xception network that has a pointwise convolution followed by a depthwise convolution. It was built with three phases, namely, entry, middle, and exit flow. Usually, an Xception model has 36 Conv layers, but we use the feature extraction part and vomit the classification layers since the classification is done by RF.
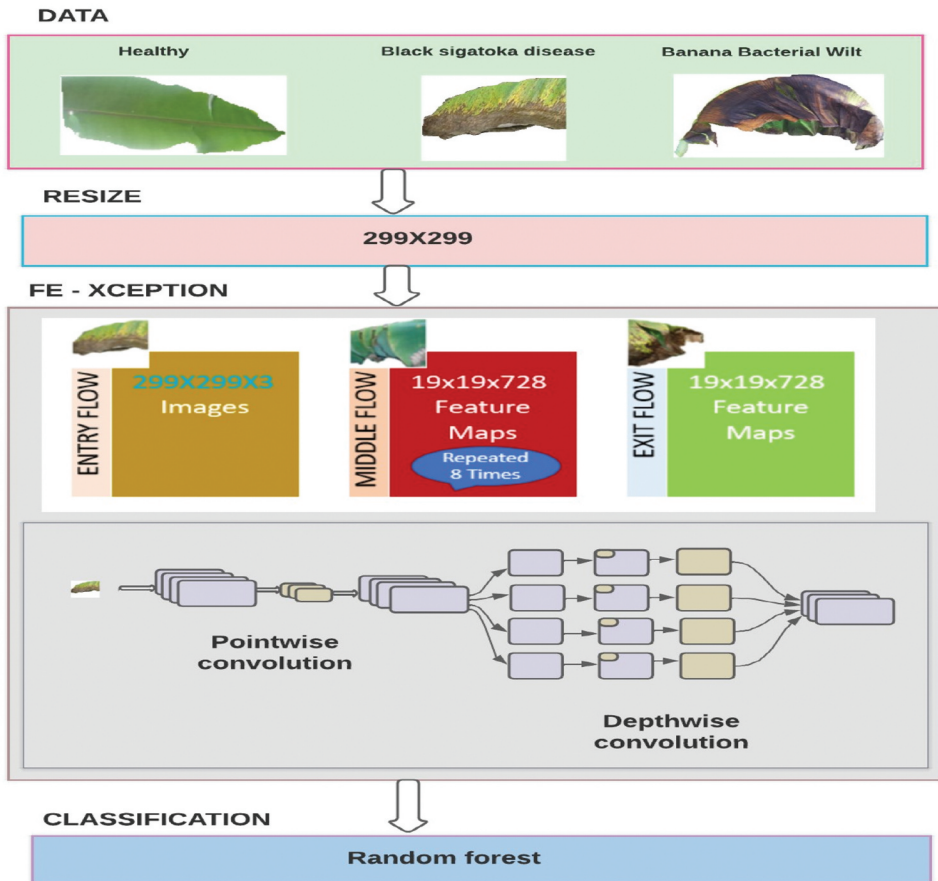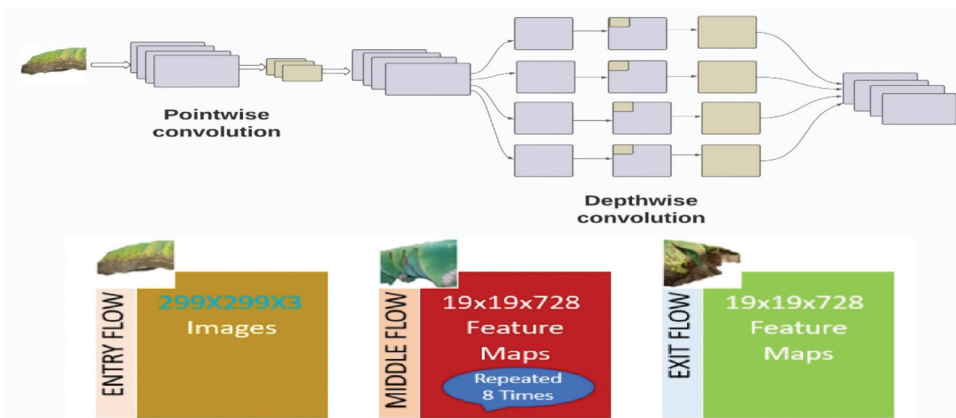
Figure 3. Workflow of proposed method.

Figure 4. Working of Xception model.

Classification

Classification is done using the random forest algorithm. The extracted feature values from the CNN are sent random forest algorithm for classification. The RF technique was first developed by L. Breiman in 2001 (Biau and Scornet 2016), and since then, it has gained popularity as a potent tool for general-purpose classification and regression analysis. The method has been demonstrated to be useful in circumstances in which there are more variables than observations. This is because it syndicates the predictions from numerous randomized decision trees and takes an average of them. In addition, it may be built up to cope with complex difficulties, adapted to meet a wide range of impromptu learning initiatives, and create metrics of different significances. These are only a few of its many advantages. Here in our work, the random forest takes input from the Xception feature extraction and then splits the data obtained from feature extraction. After splitting the data, it trains the data using the RF algorithm, as shown in Figure 5

Evaluation Metrics

It is essential to analyze a specific classified method to ascertain the level of excellence of the proposed approach. (Dalianis 2018) investigates a variety of evaluation criteria and assists in analyzing the results with a more transparent perspective. The accuracy was determined with the help of a confusion matrix, which makes a comparison between the values that were predicted and the values that were obtained using a confusion matrix that is two by two. True positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) are represented in each of the four quadrants (TN). The proportion of correctly classified results with actual values relative to the total number of other values can be used to quantify the predictive accuracy of a model, as shown in the equation below (1). In a similar vein, precision is defined as the number of correct instances retrieved divided by the total number of instances
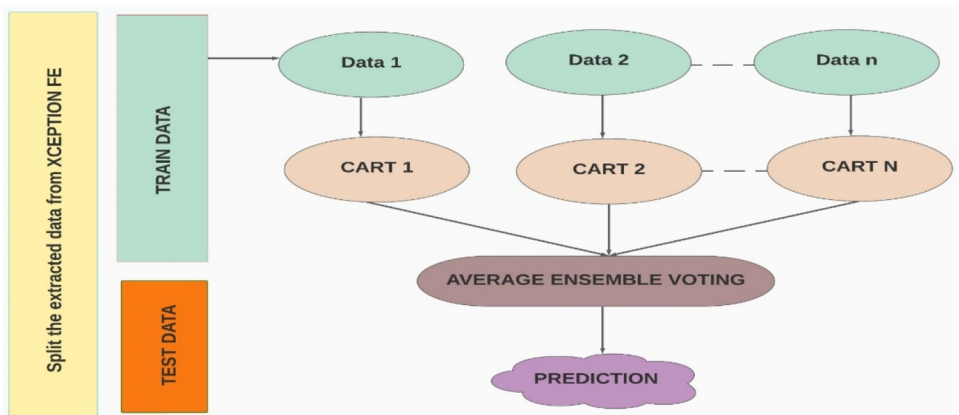

Figure 5. RF Architecture.

retrieved as in Equation (2). As indicated in Equation (3), recall is defined as the number of accurate instances retrieved divided by the total number of correct instances, and the F1-score is defined as the weighted average of both precision and recall dependent on the weight function given in detail in Dalianis (2018).

$$\text{Accuracy} = (TP + TN)/(TP + FP + FN + TN) \tag{1}$$

$$Precision(P) = \frac{TP}{TP + FP} \tag{2}$$

$$Recall(R) = \frac{TP}{TP + FN} \tag{3}$$

$$F1Score = 2 * \frac{P * R}{P + R} \tag{4}$$

Experimental results and discussion

Standard CNN models include several parameters and are computationally expensive. Standard convolution was replaced by depth-wise separable convolution, reducing parameter number and computation cost. Hence, CNNs are rising in different forms, such as pre-trained models, and we have used a pre-trained model combined with RF classifier to predict plant disease. Table 2 describes the findings of existing and proposed methods. It is evident from Table 2 that the proposed method could outperform the existing done methodologies in terms of various factors. Some existing methods have performed with data taken from controlled environments, which is not suitable for real-time processing. Hence, we have tested with the datasets available in open platforms combined with the real-time data. The findings that were obtained from the various pretrained models are presented in Table 3. CNNs are efficient at extracting features; however, the CNN classifiers are not suited for use in our work and hence cannot be used for classification due to low performance shown in Table 3. As a result, the RF algorithm was implemented to advance the performance of the network, and the result shown a significant increase in

Table 2. Comparison with existing hybrid works of DL+ML.

Models	Accuracy
Multipath-CNN(M-CNN) + J48 classifier	99.62%
CNN+SVM	97.82%
DCNR: deep cube CNN + RF	96.78%
CNN-RF	94.27%
SCNN-KSVM (Shallow CNN with Kernel SVM)	94%
CNN+SVM	92.59%
Proposed work	**100%**

Table 3. Accuracy percentage comparison of classifiers.

Pre-trained CNN	CNN CLASSIFIER	RF
VGG 16	50.2857%	98.85%
RESNET	46.857%	97.14%
MOBILENET	58.28%	98.85%
INCEPTIONRESNETV2	57.142%	99.42%
INCEPTION	51.999%	98.28%
XCEPTION	58.714%	**100%**

Table 4. F1 score, precision, and recall of pretrained models with RF.

Hybrid combinations	F1 SCORE	PRECISION	RECALL
VGG 16+RF	0.989	0.987	0.989
RESNET+RF	0.971	0.973	0.975
MOBILENET+RF	0.989	0.990	0.990
INCEPTIONRESNETV2+RF	0.994	0.995	0.995
INCEPTION+RF	0.983	0.983	0.983
XCEPTION+RF	**1.000**	**1.000**	**1.000**

the network's performance. As it is shown in Table 3, the Xception model combined with RF has performed better than other models. The F1 score, precision, and recall were all computed, and the results are presented in Table 4. These three scores have scores that are comparable to the accuracy score presented in Table 3. The confusion matrix of correctly categorized and incorrectly classed class images can be shown in Figures 6 and 7, respectively. Also in addition, as shown in Figure 8, the accuracy comparisons of all of the algorithms are compared, and the suggested Xception-RF algorithm has performed well in comparison to the other algorithms. The F1 score, precision, and recall were also calculated as shown in Figure 9, and the figure exhibits that the proposed Xception-RF has performed well compared to the other models.
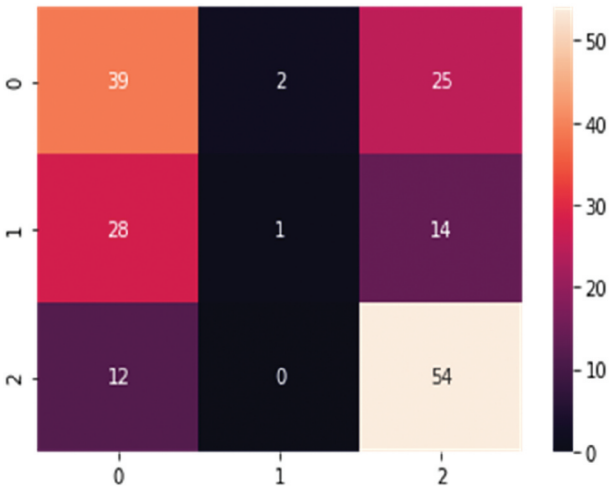

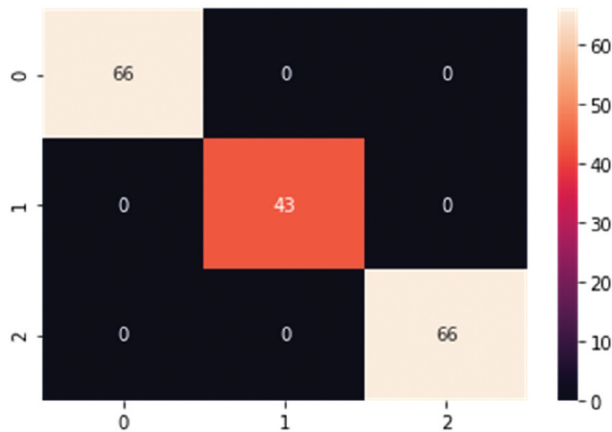**Figure 6.** Confusion matrix with CNN classifier.
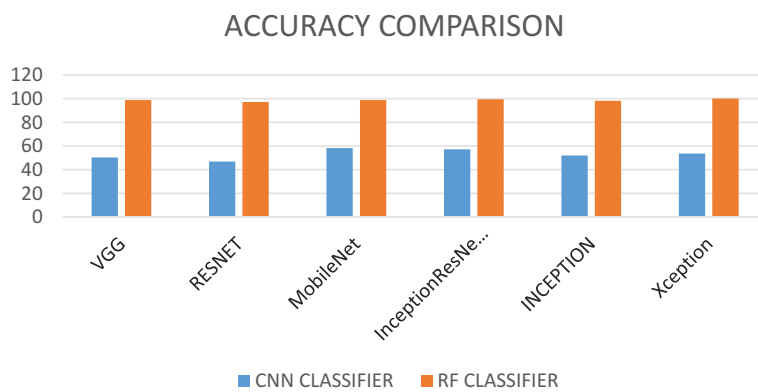
Figure 7. Confusion matrix with RF classifier.

Figure 8. Accuracy comparison of the models.
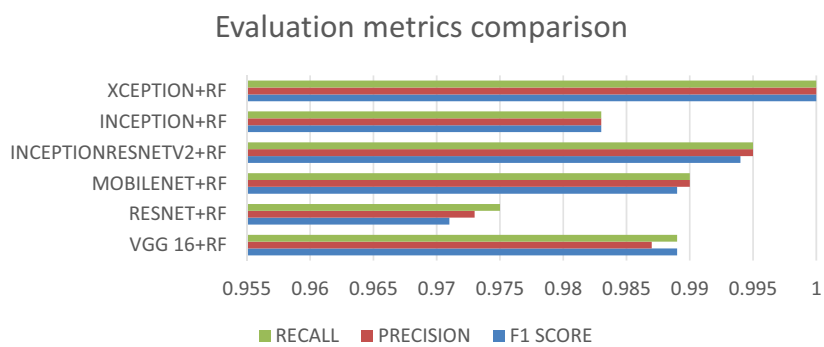

Figure 9. F1 Score, Precision and Recall.

Conclusion

This work, therefore, presents the results of experiments in the analysis of numerous pre-trained models employing machine learning techniques. The proposed Xception-RF model provides perfect accuracy so far by surpassing other competing methods. This work contributes very much to the agricultural countries and enhances the quality of farming by predicting diseases. Deep convolution layers for feature extraction and the RF method both appear to have played major roles in the model's effectiveness. We have evaluated the model using accuracy, F1 score, precision, and recall, which have evaluated the model with similar results and have displayed the algorithm's effectiveness. Using the Xception model in tandem with the RF classifier to categorize plant diseases is the contribution of this work. The inclusion of abiotic stress with real-time cloud processing will further enhance the work in the future.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Abisha A ⓘ http://orcid.org/0000-0002-0464-2073
Bharathi N ⓘ http://orcid.org/0000-0003-4026-6754

References

Abdu, A. M., M. Mohd Mokji, and U. Ullah Sheikh. 2020. Automatic vegetable disease identification approach using individual lesion features. *Computers and Electronics in Agriculture* 176:105660. doi:10.1016/j.compag.2020.105660.

Abisha, A., and N. Bharathi. 2021. "Review on Plant health and Stress with various AI techniques and Big data." In 2021 International Conference on System, Computation, Automation and Networking (ICSCAN) Puducherry, India, pp. 1–6 doi:10.1109/ICSCAN53069.2021.9526370. IEEE.

Abisha, A., and N. Bharathi. 2022. Feature extraction from plant leaves and classification of plant health using machine learning Gupta, Deepak, Sambyo, Koj, Prasad, Mukesh, Agarwal, Sonali. In *Advanced Machine Intelligence and Signal Processing*, 867–76. Singapore: Springer.

Al Mahmud, K. 2021, June. Banana leaf dataset, version 1. Retrieved January 20, 2022 from https://www.kaggle.com/datasets/kaiesalmahmud/banana-leaf-dataset.

Barbedo, J. G. A. 2019. Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering* 180:96–107. doi:10.1016/j.biosystemseng.2019.02.002.

Biau, G., and E. Scornet. 2016. A random forest guided tour. *Test* 25 (2):197–227. doi:10.1007/s11749-016-0481-7.

Chollet, F. 2017. "Xception: Deep learning with depthwise separable convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition Honolulu, HI, USA, pp. 1251–58.

Christian, S., S. Ioffe, V. Vanhoucke, and A. A. Alemi. 2017. "Inception-v4, inception-resnet and the impact of residual connections on learning." In Thirty-first AAAI conference on artificial intelligence San Francisco, California, USA.

Dalianis, H. 2018. Evaluation metrics and evaluation. In *Clinical text mining*, 45–53. Cham: Springer.

Gayathri, S., V. P. Gopi, and P. Palanisamy. 2021. Diabetic retinopathy classification based on multipath CNN and machine learning classifiers. *Physical and Engineering Sciences in Medicine* 44 (3):639–53. doi:10.1007/s13246-021-01012-3.

Gokulnath, B. V., and G. Usha Devi. 2021. Identifying and classifying plant disease using resilient LF-CNN. *Ecological Informatics* 63:101283. doi:10.1016/j.ecoinf.2021.101283.

Hassan, S. M., A. Kumar Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska. 2021. Identification of plant-leaf diseases using CNN and transfer-learning approach. *Electronics* 10 (12):1388. doi:10.3390/electronics10121388.

Howard, A. G., M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv: 170404861* 1 1–9 .

Jadhav, S. B., V. R. Udupi, and S. B. Patil. 2019. Convolutional neural networks for leaf image-based plant disease classification. *IAES International Journal of Artificial Intelligence* 8 (4):328. doi:10.11591/ijai.v8.i4.pp328-341.

Kumar, S., V. Chaudhary, and S. Khaitan Chandra. 2021. Plant disease detection using CNN. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12 (12):2106–12.

Kwak, G.H., C.W. Park, K.D. Lee, S. -I. Na, H.Y. Ahn, and N.W. Park. 2021. Potential of hybrid CNN-RF model for early crop mapping with limited input data. *Remote Sensing* 13 (9):1629. doi:10.3390/rs13091629.

Lamba, M., Y. Gigras, and A. Dhull. 2021. Classification of plant diseases using machine and deep learning. *Open Computer Science* 11 (1):491–508. doi:10.1515/comp-2020-0122.

Li, T., J. Leng, L. Kong, S. Guo, G. Bai, and K. Wang. 2019. DCNR: Deep cube CNN with random forest for hyperspectral image classification. *Multimedia Tools and Applications* 78 (3):3411–33. doi:10.1007/s11042-018-5986-5.

Lili, L., S. Zhang, and B. Wang. 2021. Plant disease detection and classification by deep learning —a review. *IEEE Access* 9:56683–98. doi:10.1109/ACCESS.2021.3069646.

Li, Y., J. Nie, and X. Chao. 2020. Do we really need deep CNN for plant diseases identification? *Computers and Electronics in Agriculture* 178:105803. doi:10.1016/j.compag.2020.105803.

Lu, J., L. Tan, and H. Jiang. 2021. Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture* 11 (8):707. doi:10.3390/agriculture11080707.

Mahum, R., S. Ur Rehman, T. Meraj, H. Tayyab Rauf, A. Irtaza, A. M. El-Sherbeeny, and M. A. El-Meligy. 2021. A novel hybrid approach based on deep cnn features to detect knee osteoarthritis. *Sensors* 21 (18):6189. doi:10.3390/s21186189.

Medhi, E., and N. Deb. 2022. PSFD-Musa: A dataset of banana plant, stem, fruit, leaf, and disease. *Data in Brief* 43:108427. doi:10.1016/j.dib.2022.108427.

Mohameth, F., C. Bingcai, and K. Amath Sada. 2020. Plant disease detection with deep learning and feature extraction using plant village. *Journal of Computer and Communications* 8 (6):10–22. doi:10.4236/jcc.2020.86002.

Scarpa, G., M. Gargiulo, A. Mazza, and R. Gaetano. 2018. A CNN-based fusion method for feature extraction from sentinel data. *Remote Sensing* 10 (2):236. doi:10.3390/rs10020236.

Simonyan, K., and A. Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv: 14091556* 1 1–14 .

Singh, J., D. Thakur, F. Ali, T. Gera, and K. Sup Kwak. 2020. Deep feature extraction and classification of android malware images. *Sensors* 20 (24):7013. doi:10.3390/s20247013.

Swinney, C. J., and J. C. Woods. 2021. The Effect of Real-World Interference on CNN Feature Extraction and Machine Learning Classification of Unmanned Aerial Systems. *Aerospace* 8 (7):179. doi:10.3390/aerospace8070179.

Theckedath, D., and R. R. Sedamkar. 2020. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Computer Science* 1 (2):1–7. doi:10.1007/s42979-020-0114-9.

Toda, Y., and F. Okura. 2019. How convolutional neural networks diagnose plant disease. *Plant Phenomics* 2019. doi:10.34133/2019/9237136.

Vijayalakshmi, M., and V. Joseph Peter. 2021. CNN based approach for identifying banana species from fruits. *International Journal of Information Technology* 13 (1):27–32. doi:10.1007/s41870-020-00554-1.

Xie, W., Z. Li, Y. Xu, P. Gardoni, and W. Li. 2022. Evaluation of different bearing fault classifiers in utilizing CNN feature extraction ability. *Sensors* 22 (9):3314. doi:10.3390/s22093314.

Yang, S., L. Gu, X. Li, T. Jiang, and R. Ren. 2020. Crop classification method based on optimal feature selection and hybrid CNN-RF networks for multi-temporal remote sensing imagery. *Remote Sensing* 12 (19):3119. doi:10.3390/rs12193119.

Zhao, H.H., and H. Liu. 2020. Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition. *Granular Computing* 5 (3):411–18. doi:10.1007/s41066-019-00158-6.