INTELIGENCIA ARTIFICIAL

A Survey on Plant Disease Prediction using Machine Learning and Deep Learning Techniques

Gokulnath BV, Usha Devi G*

School of Information Technology and Engineering
Vellore Institute of Technology, Vellore, Tamil Nadu, INDIA.
gokul19.gn@gmail.com

School of Information Technology and Engineering
Vellore Institute of Technology, Vellore, Tamil Nadu, INDIA.
ushadevi.g@vit.ac.in (*Corresponding Author)

Abstract The major agricultural products in India are rice, wheat, pulses, and spices. As our population is increasing rapidly the demand for agriculture products also increasing alarmingly. A huge amount of data are incremented from various field of agriculture. Analysis of this data helps in predicting the crop yield, analyzing soil quality, predicting disease in a plant, and how meteorological factor affects crop productivity. Crop protection plays a vital role in maintaining agriculture product. Pathogen, pest, weed, and animals are responsible for the productivity loss in agriculture product. Machine learning techniques like Random Forest, Bayesian Network, Decision Tree, Support Vector Machine etc. help in automatic detection of plant disease from visual symptoms in the plant. A survey of different existing machine learning techniques used for plant disease prediction was presented in this paper. Automatic detection of disease in plant helps in early diagnosis and prevention of disease which leads to an increase in agriculture productivity.

Keywords:Plant disease prediction, Crop productivity, Support vector machine, Deep learning, Meteorological factor, Visual symptoms, Random forest

1.   Introduction

Productivity, food security, sustainability, and environmental impact are handled using smart farming techniques. As the population level is increasing day by day food production plays a major role in it Turner *et al.* [2018]. The quality of the food products is essential for good healthy human beings. For providing this healthy environment analysis of various factor that causes a productive loss in agriculture has to be done. Identification of diseased part, Image classification, Detection of an anomaly in plant leaves open wide area research named image processing in agriculture Dhingra *et al.* [2019].

Cuadro 1: List of abbreviations	
PH	Potential of hydrogen
SVI	Spectral Vegetation Indices
CART	Classification and Regression Tree
FCN	Fully Convolution Network
BPNN	Back Propagation Neural network
R2	R Square
PCA	Principal Component Analysis
MAE	Mean Absolute Error
KNN	K-Nearest Neighbors
RMSE	Root Mean Square Error
SVM	Support Vector Machine

1.1. Challenges in Agriculture

As the population increases, day by day the challenges in agriculture also increases exponentially. Land limitation and water scarcity play a vital role in agriculture. Modern agriculture integrates many technologies to resolve these problems in real time.

- **Choosing the right seed:** One of the important aspects of agriculture is choosing the right seed to cultivate in the right area Kouadio *et al.* [2018]. Classification analysis of the field soil, water, and PH level help in cultivating the right crop in the right area, so the result the yield of the crop is comparatively high.

- **Irrigation Techniques:** Maintain proper irrigation in the field helps in increasing the crop yield. Machine learning technique helps in improving the crop yield by maintaining the water level in the root region, average vegetable yield ratio is an increase, and plant need is measured using soil moisture sensor and low labor input Elavarasan *et al.* [2018].

- **Predictive Analysis:** It helps in making the sowing decision, finding healthy crops, and tells when fertilizer has to be added

- **Soil Defect Diagnosis:** anomaly analysis helps in identifying the strength of the soil, so it helps the farmer during the crop cultivation process. Crop yield loss can be reduced by identifying the defects.

- **Forecasting Weather Condition:** Climate plays a vital role in crop growth. One of the main reason for the productivity loss in the crop is due to the climate factor. For providing the stable growth weather forecasting is the key factor and regression analysis helps in forecasting weather.

- **Detection of weeds in the field:** Nearly 50% loss in the crop yield is mainly due to the disease, pest and harvest loss. Image processing technique mainly helps in analyzing the disease Krishnakumar and Narayanan [2018]. Classification of plant among weeds helps in weed detection which provides a larger growth.

- **Water Treatment:** minerals are essential for plant growth. Unsupervised analysis and anomaly detection help in finding the right amount of minerals needed for the plant.

1.2. Environmental Condition

The environmental condition also plays a key role in plant disease. Pathogen grows at different environmental condition. Fungal infection is more common where the moisture is a high and specific temperature. Soil quality, temperature, wind, nutrition, sunlight influence plant growth. Spreading of disease is also influenced by environmental factors.

■ **Temperature:** Pathogen grows at a specific temperature condition. There are two types of disease that come under temperature and they are a warm-weather and cool-weather disease. Soil temperature like cool and wet soil is mainly responsible for fungal disease in the root Bugingo [2018]. Temperature mainly increases sustainability. Wind and Sun ? a sufficient amount of sunlight is needed for growth of the plant. Spores will spread the infected plant tissues.

■ **Soil Type and its Fertility:** Soil may help in the development of pathogen. If organic matter is low in soil then it helps in the growth of nematodes. Soil PH is mainly responsible for the development of pathogen. Scabs on potato surface are developed when there is high PH. Similarly, the excessive amount of nitrogen in the soil also promotes the growth of the pathogen.

■ **Moisture:** Rain, dew, humidity, and water can be included in moisture. Fire blight, black spot, and apple scab can be spread rapidly when the environmental condition is moisture. Wet foliage condition promotes more disease-causing pathogens.

2.   Factors Causing Plant Disease

Pathogens are the main reason for causing diseases in the plant. There is a department named after it called plant pathology it mainly deals with the study of the pathogen. There are two main factors which cause disease in plants and they are pathogens and environmental condition.

2.1.   Viral

It is a living organism which has living cells in it and it affects the plant. The region affected by the virus in plant leaf and fruits are seen as yellow streaking, yellow spots, deformed leaves, and stunted growth Tennant *et al.* [2018]. In cucumber viral infection is mainly caused by cucumber mosaic virus and these viral diseases is a communicable disease which can be spread from one plant to another plant either by insect or touch. The best way to prevent viral disease is the disposal of the viral affected region.

2.2.   Fungi

Fungi is also one main reason for the productivity loss in the plant. Ascomycetes and basidiomycetes are two main fungi mainly responsible for disease in the plant. Fungicides are widely used for controlling fungal infection in a plant Collinge *et al.* [2019]. Magnaporthe grisea which is commonly known as rice blast disease. Sclerotinia sclerotiorum is responsible for cotton rot. Oomycetes and Phytomyxea are the fungal-like organisms which contain destructive pathogen in the plant.

2.3.   Bacteria

When a plant is infected by microscopic living organisms then it is a bacterial infection. Bacteria is a one-celled organism. The bacterially infected regions in plants can be seen as wilts, spots, and scabs. The spots which are caused by the blights can spread rapidly in plants. Tropical plants and vegetables are affected mainly by wilts. The absorption water by the plant is blocked by the bacterial infection. Some of the bacterial plant pathogens are Burkholderia and Proteobacteria.
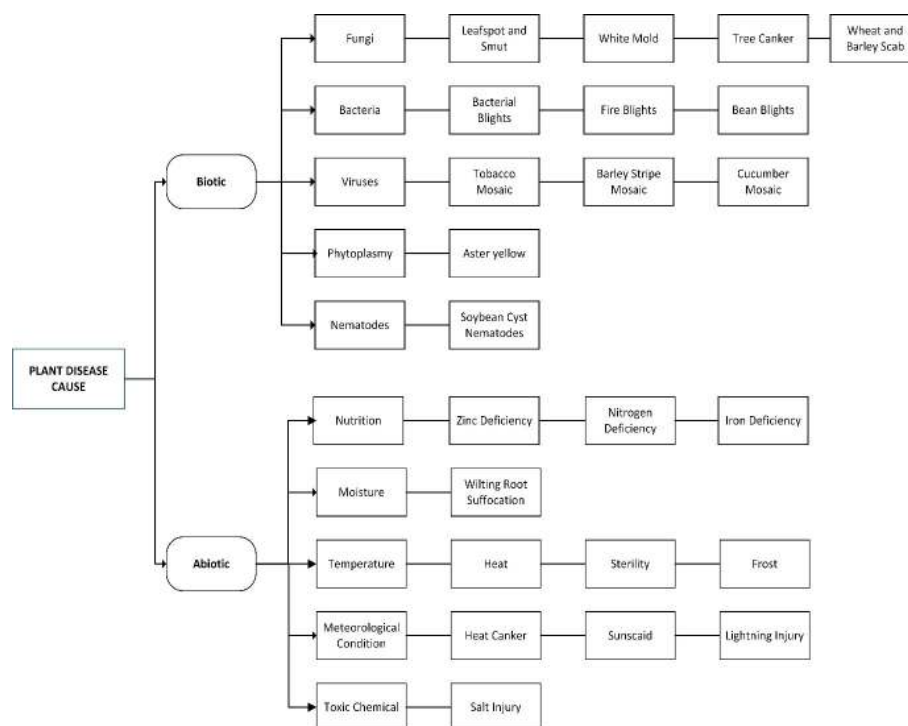
Figura 1: Factors causing diseases in plant

The first section presents an introduction about challenges in agriculture. The second section deals with the factor causing plant disease. The third section deals with the techniques used in plant disease prediction. The fourth section brief about image analysis for disease detection in the plant. The fifth section details the machine learning and deep learning technique used in plant disease prediction. The sixth section concludes the survey.

3.   Techniques used in Plant Disease Prediction

Machine learning is the successor of the statistical approach and it comes under the broader segment of artificial intelligence. Machine learning is broadly classified into four types as follows

- Supervised learning method

- Unsupervised learning method

- Semi-supervised learning method

- Reinforcement learning method

It has many efforts, new chances, new procedures, evolution technique in the agriculture field. This technique can be applied to water preservation, supplement utilization which provides a better healthy environment Oliver *et al.* [2018]. Crop management leads to several areas such as crop yield forecasting, plant weed detection, plant disease detection, agro-climate prediction, and pest control analysis. Classification problem and regression problem both can be easily solved using a supervised algorithm. It mainly consists of training data?s with labeled data are in it, so it easily compares the new data with it and predicts the output.

In unsupervised learning, the neural networks help to find the structure and find the features in it and then analyze it. Clustering, anomaly detection, autoencoders, and association can be organized by unsupervised learning. Whereas semi-supervised learning consist of both labeled and unlabeled data and

this learning mainly helps when the data are difficult to understand. In reinforcement learning, it works like video games like previous feedback of game helps in the improvement of the next game. It mainly helps in finding the best way to predict the outcome.

Applications like crop selection, crop yield prediction, weather forecasting, smart irrigation system, crop disease prediction and minimum support price system are implemented by the government organisation to help the farmers and increase the crop productivity and crop health.
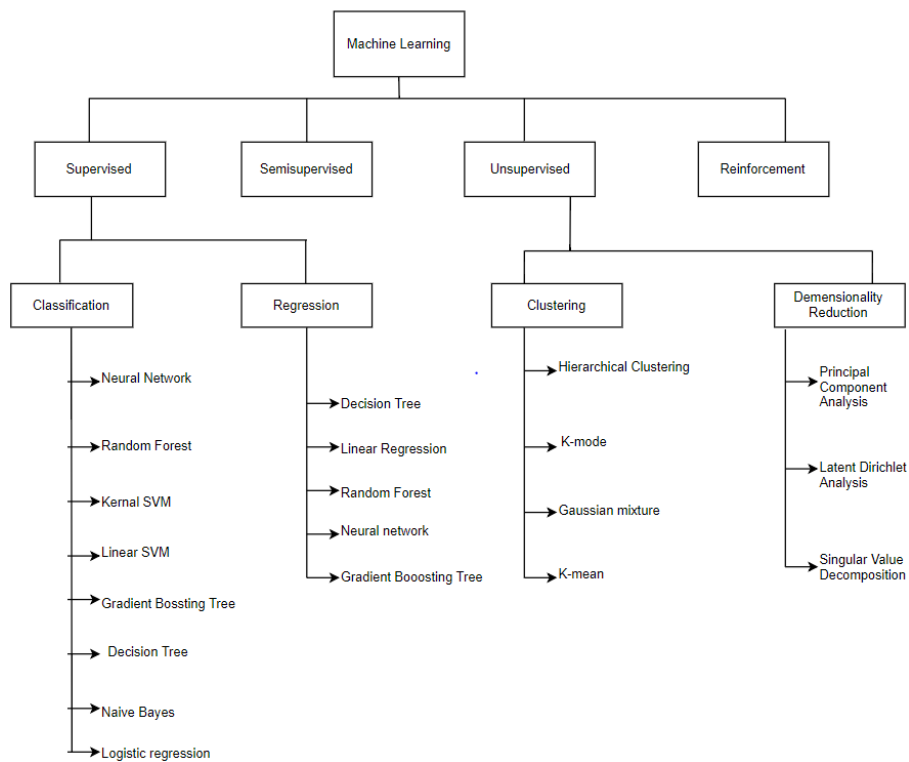

Figura 2: Types of machine learning algorithms

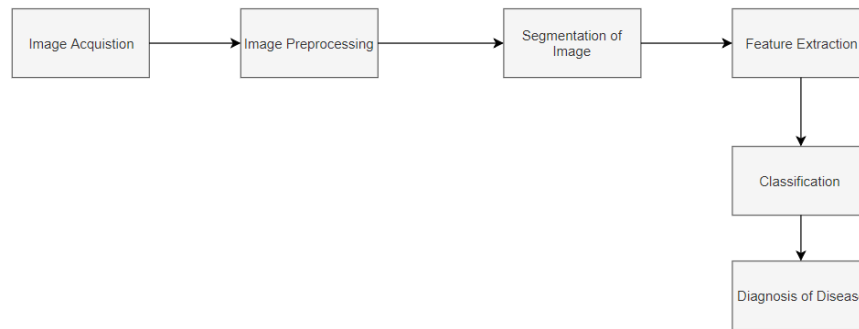# 4.   Image Analysis for Disease Detection in Plant

Figura 3: General workflow of plant disease prediction model

The image acquisition process is the first stage in the process of image analysis. It is also known as digital image acquisition. It can be defined as the visual character of an object can be represented as a digitally encoded one. In simple terms, it can be defined as an image captured using a camera. Nowadays digital image is extended to a mobile phone it makes the process of image acquisition a user-friendly. Photographs, printed paper, and photographic film are the media used for it. It mainly captures visual moments.

In the pre-processing of the image, there are two types in it namely digital image processing and analog image processing. Removal of unwanted features in the image is the main process involved in it. More algorithms are used for the removal of unwanted features in the image. The major steps involved in the image pre-processing are 1. Image Acquisition 2. Image Normalization 3. Image Enhancement 4. Segmentation 5. Morphology.

Separating image into pixel and their similar attributes is image segmentation. It mainly helps in the image interpretation process Oliver *et al.* [2018]. It transfers the low-level image into a high-level image. While analyzing an image it success mainly depends on reliability in the segmentation process. Segmentation process involves both contextual and non-contextual. Several algorithms are used for the segmentation process.

A copy of the original feature is kept during the feature selection process. In the feature extraction process, new sets of features are created and these two processes mainly deal with removing unwanted noise in the image and choosing the needed features only for the image analysis process. Transformation of the attribute process takes place during the feature extraction process. The speed and effectiveness of the process are enhanced during this process.

During the classification process, the data are categorized into a number of classes. If new observation came into the process then identifying whether the new observation belongs to which class Ferentinos [2018a]. There are several classification algorithms available for the process of classification which gives accurate classification result as well.

Cuadro 2: List of publicly available plant leaf dataset

Dataset Name	URL
Plant Village	https://www.plantvillage.org/
Plant Image Analysis	https://www.plant-image-analysis.org/dataset
DATA.GOV	https://www.catalog.data.gov/dataset
UCI Machine Learning Repository	https://www.archies.ics.uci.edu/ml/datasets/plants
Flavia	https://www.flavia.sourceforge.net
Image Database of Plant Disease Symptoms (PDDB)	https://www.digipathos-rep.anptia.embrop.br/
Swedish Leaf Dataset	https://www.cvl.isy.liu.se
Plant Phenotyping dataset	https://plant-phenotyping.org/dataset

Several IoT devices are used for smart farming and it helps in improving the agriculture system. Many sensor are used for the purpose of smart farming as a result it helps in saving time and increase productiviy for farmers. Some of the key application of it are precision farming, Agriculture drones, livestock monitoring, and smart green house.

5. Related Works

5.1. Machine Learning Techniques for Plant Disease Prediction

Azadbakht *et al.* [2019] analyzed the performance for identifying leaf rust of wheat at canopy scale and at a different level of leaf area index (LAI) - high, medium and low. Four methods had been used for analyzing the performance in identifying the rust detection namely Gaussian process regression, random forest regression, v-support vector regression and boosted regression tree. The severity level of disease is also predicted in this. For analyzing the performance the experiment was conducted at North West of Iran in 7000 hectares of wheat cultivation. Hyperspectral reflectance data recorded at different environmental condition is used for it. Based on the results from the experiment hyperspectral images can be used for identifying the wheat rust. Comparison of spectral vegetation indices (SVIs) and ML shows that ML performs better than SVIs.

Ip *et al.* [2018] presented a review about crop protection using big data and how it helps in controlling the weed. Discussed mainly about invasive species detection, predicting and modeling of resistance in herbicide, support system used for the purpose of crop protection and robotic weed control system. The machine learning approaches used for the above-mentioned problem are discussed clearly.

Markov random field uses spatial component for analyzing is explored in it. Probabilistic graphical model, traditional neural networks and support vector machine are the generative approach used in it. For the purpose of modeling and prediction of resistance in herbicide rule-based approach and CART is used. For detecting species and weeds random forest classifier is used. RIM is used to control the ryegrass in the field. AgBot is robotics used as autonomous weed control. Markov Random Field helps in content-based fresh fruit grading and it has the ability to handle spatial data with irregular spacing. For modeling, the crop disease in corn Bayesian network is used and SVM helps in the semi-autonomous estimation of vegetables.

Chlingaryan *et al.* [2018] nitrogen estimation is an important factor in precision agriculture. Remote sensing system there is a problem in processing huge amount of data from a different platform. The machine learning techniques help in handling the non-linear task. The mainly hybrid system of integrating machine learning techniques and remote sensing technique is used in nitrogen estimation. For estimating the chemical concentration in dry leaf Continual Removal (CR) method is used. For identifying more informative feature (BPNNs) is used. For calculating the vegetation indices back propagation neural networks are used. Feature extraction can be done by combining a convolution neural network and Gaussian process. The Gaussian process also helps in identifying different plant leaves characteristic. For the purpose multi-class crop prediction M5 prime regression tree is used. And for nitrogen estimation, least square support vector machine is used. Fuzzy Cognitive Map is used in the field of crop management.

Elavarasan *et al.* [2018] presented the models that are used for analyzing several factors like crop disease analysis, yield prediction, crop improvement. And also provides the comparison of these models with the error measures. In this paper comparison of the existing survey objective and current survey, the objective is provided. And a detailed description of what are the algorithms used in machine learning for agriculture problems. Some of the main parameters that are used in the agriculture sector are also listed as well. And the performance metric of both supervised and unsupervised machine learning algorithm is compared with the error measures MAE, RRMSE, RMSE, and R2. A tabulated description of ideology, application, advantage and parameter considered are listed in it.

Ashourloo *et al.* [2016] disease severity detection can be analyzed using spectral vegetation indices (SVIs) and the author had analyzed the method of leaf rust detection like partial least square regression (PLSR), Gaussian process regression (GPR) and v-Support Vector Regression (v-SVR) and also analyzed how training data size has its impact on the result and how the disease symptoms has its effects on the prediction accuracy. Based on the analysis made GPR performs better and SVIs is sensible to different disease symptoms. The accuracy is high when there is a small sample data set are used for the GPR method. PLSR, v-SVR and GPR can be used to test various variety in a plant using a different sensor.

Lu *et al.* [2017a] presented a system based on the deep learning framework named in field automatic wheat diagnosis system (DNIL- WDDS). In wild condition, the image-level annotation is done for the wheat disease. VGG-FCN-S and VGG-FCN-VD16 are the two architecture used in it. The proposed system achieves the accuracy of 97.95 % and 95.12 % when compared to CNN architecture achieving the accuracy of 93.27 % and 73.00 %. It has an added advantage of integrating the system with the mobile application. Nearly four types of deep learning method are used for analyzing the disease prediction accuracy in wheat. The DMIL-WDDS model shows improvement in accuracy when compared to other models.

Kukar *et al.* [2019] proposed a decision support system for the farm management system. The decision support system is an integration of cloud-based decision support toolbox. This decision support system helps the farmer to upload their own field data and then the support system do the analysis portion and then predict the outcome to the farmers. This system includes structural change detection, time series clustering and decomposition in it. It is mainly used in pest population dynamics. It uses data mining principle for extracting useful information from a large volume of the dataset. The system has an easy user interface which makes it more useful for all kind of users. It makes the system accessible for a non-technical person. This system follows (CRISP-DM). For predictive modeling purpose, random forest algorithm was used and the performance is assessed using out-of-bag (OOB). Data mining is implemented using a wild fly application server and the core analyzing and processing is done using R software.

Barbedo [2016] presented various problem in plant disease identification. It uses visible range image for disease identification purpose. Both intrinsic and extrinsic factor is involved in this. While considering the extrinsic factor that influence plant disease identification is image background. If the background is white or blue then it can be processed easily. Image capture condition has to be maintained while segmenting image. If the segmentation process is done in a laboratory condition is ideal for analysis purpose. An aspect of illumination is also considered while analyzing. While considering the intrinsic factor symptom segmentation is one of the important aspects that affect plant disease identification. Boundaries and edges have to be defined clearly in disease analysis. Symptom variation is also an important aspect of considering the intrinsic factor. Different stage of the disease has different growth of tissue so identifying them and analyzing them is an important aspect of symptom variation. Multiple simultaneous disorder and different disorders with similar symptoms are some of the intrinsic aspects that influence plant disease identification.

Kamilaris *et al.* [2017] presented a survey on how big data techniques used in the field of agriculture using three V as a parameter that is volume, velocity, variety and also analyzed what are the techniques and solution provided for it. Problems like weather and climate change, land, animal research, crops, soil, weeds, food availability and security, biodiversity, remote sensing, farmer insurance, and finance are analyzing indicators like Low (L), Medium (M), and High (H) in three V. Some of its application in agriculture is described briefly. The source of data for the agriculture application is also listed briefly in this work. Techniques like image processing, statistical analysis, cloud computing, machine learning, NDVI are most commonly used for the problems related to the agriculture application. List of software tools used in each platform for the purpose of agriculture application is also listed in it.

Johannes *et al.* [2017] prediction of symptoms helps in early diagnosis of disease. Image-based automatic identification of plant disease mostly fails under real field condition so an image processing algorithm was proposed based on hot spot identification. Rust, septoria and tan disease are considered for identification using this algorithm in wheat plant and the images used for this analysis are taken from the mobile devices. The proposed algorithm shows 0.80 higher accuracies than another algorithm. This algorithm can be applied to another crop as well. Many algorithms fail under natural condition but this algorithm provides a better result in the natural condition. This algorithm provides versatility and has been tested for three different wheat disease. Colour constancy algorithm helps in improving accuracy.

Park *et al.* [2018] PCA helps in reducing the dimensionality but does not provide which spectral band is essential so a feature selection technique named minimum redundancy and maximum relevance (mRMR) was proposed in this. This technique enables to choose a raw band from the spectral image. A deep neural network was proposed along with convolution neural network and a fully connected network. FCN is used to classify the apple leaf according to its condition namely normal, young, early, late, and mal nutrient. This method reduces the problem that occurs in the hyperspectral image which leads to better disease prediction. Through the proposed technique the hyperspectral image of 519 bands is reduced to 5 bands using automatic feature selection (mRMR) and deep neural network (VGG net + FCN) is used for the leaf state classification.

Khanal *et al.* [2018] proposed a technique integrating machine learning and remotely sensed data. Performance of different machine learning algorithm is analyzed. Using remote sensing data spatial property of soil and corn yield are measured. Soil and vegetation indices are derived at m resolution. Mainly linear regression is used for prediction of soil property and corn yield. Performance of machine learning algorithms like a neural network, random forest, gradient boosting model, support vector machine and cubist are also analyzed. R2 value and RMSE value are calculated for accessing the performance of machine learning algorithm. For the purpose of statistical analysis, R software is used. Six statistical models are used for predicting the property of soil. The parameter for measuring the property of soil and the corn yield is caution Exchange capacity (CEC), Soil Organic Matter (SOM), Magnesium (Mg), Potassium (Mg), pH, Yield (Approach 1), Yield (Approach 2). The effectiveness of the model is calculated using R2 and Root mean squared error.

Phadikar *et al.* [2013] proposed an autonomous system for classifying disease affected in rice crop. It mainly extracts an infected region from the rice plant leaf. For the purpose extracting an infected region from the rice plant leaf Fermi energy based segmentation method is used. For the purpose of symptom classification, a novel algorithm is developed which classify the symptom based on the color, shape, and position. Information loss reduction and selection of an important feature are done through Roush Set Theory (RST). The classification portion is done using rule-based classifier and it classifies whether it is leaf brown spot, rice blast, sheath rot, and bacterial blight. The feature extraction includes shape feature, center detection, shape detection, position detection and based on that rule is generated as well. The computational complexity is less in the proposed system. To compare the performance of the proposed algorithm with other algorithms 10 fold cross validation is performed.

Chapman *et al.* [2018] for predicting the yield in agriculture Bayesian network are used it helps in predicting the fruit yield, average bunch number per hectare. When comparing the prediction accuracy of Bayesian with another algorithm it shows high accuracy of 0.6 and 0.9 r2 values. When the Bayesian is compared with the artificial neural network it produces a similar result as well. The yield production of major crops mainly depends on rainfall and fertilizer. Analysis of the soil texture, fertility status, drainage, and flooding and moisture status is done. R software is used to compare the performance of Bayesian with ANN. Backpropagation algorithm is used in the artificial neural network. PALMSIM helps in making an evidence-based decision in agriculture.

Karadağ *et al.* [2019] proposed a technique that helps in identifying pepper fusarium disease. Light reflection from plant helps in knowing information about plant through specter radiometer. Four types of leaves are used for it namely mycorrhizal fungus, fusarium diseased, healthy and mycorrhizal leaves. The wavelength should be between 350nm and 2500nm. The processing takes place in two levels 1. Feature vector 2. Feature vector classification. For the purpose of classification three methods are used they are ANN, KNN, and NB. The average success rate is calculated for it KNN achieving a 100 % success rate whereas 97.5 % for ANN and 90 % for Naïve Bayes. Large data dimension reduces the classification performance. Wavelet Transformation (WT) helps in reducing the dimension. Sym5, db2, and Haar are

used for this purpose. Backpropagation algorithm is trained in ANN for the purpose of classification. Its performance is tested using 10 field cross-validation. Mat lab R2015b is used for the classification algorithm implementation purpose. For KNN when K=2 the success rate for sym5, Haar and db2 is 100 %, 100 %, and 98 %. When K=3 the average success rate is 100 %, 100 %, and 98 %. When K=5 the average success rate is 92.5 %, 91.5 %, and 85 %. Compared with other algorithm KNN provides better results. Wavelet decomposition reduces the features from 2150 to 75.

Liang *et al.* [2019] proposed a technique based on the robust image-based plant disease diagnosis and severity estimation network (PD2SE-Net). It mainly focuses on practical diagnosis. This proposed technique also helps in measuring the severity of the disease. For the purpose of improving the accuracy convolution, a neural network is used. ResNet50 architecture is used as the base model. The accuracy achieved through this proposed technique over existing approaches is 0.91, 0.99 and 0.98. The implementation code is from PyTorch framework. Overfitting occurs due to the partition of classes. The computational cost of the proposed technique is low. The plant disease diagnosis accuracy and plant disease severity are plotted against the class ID.

Ebrahimi *et al.* [2017] proposed a technique helps in detecting thrips (Thysanoptera). From the crops, canopy image thrips can be identified using Support Vector Machine classification. This technique helps in identifying whether it is found in the strawberry plant through an image processing technique. Different type of kernel function is used in the SVM parasite classification and detection of Thysanoptera. MAE, RMSE, MPE, and MSE is also calculated. The error value is less than 2.25 % when the color index and region index is used for the classification. For the purpose of removing the background from an image, MATLAB R2010a was used in it.

Iqbal *et al.* [2018] mainly focus on the citrus plant disease and the classification of the different disease that occurs in the citrus plant. It also gives a detailed description of the different technique used for the segmentation process, feature extraction, feature selection, image processing, and classification method. It also discusses the automated tools used for the detection and classification purpose. Canker, black spot, citrus scab, melanose, gearing are the disease that occurs in the citrus plant. The techniques used for the different stage of analysis is compared with the existing survey for the purpose of disease extraction K-mean algorithm is used. Back Propagation Neural Network (BPNN) and Grey Level Co-Occurrences Matrix (GLCM) are used for color feature computation and classification. The techniques used for pre-processing, color based transformation, image enhancement, noise reduction and resize and segmentation are discussed. Different feature extraction technique based on texture, color, and shape. The summary of different classifier technique along with its application is given in it. From the analysis, it is clear that segmentation accuracy is improved by the pre-processing technique.

Pantazi *et al.* [2019] proposed a technique for identifying the disease from the wine leaf (healthy, downy mildew, powdery mildew, and black rot). For the purpose of feature extraction, local binary pattern technique is used. From the wine leaf image data set the disease are classified with the help of one class classification method. This algorithm is trained for the various crop as well and the accuracy results are high. It achieves an accuracy of 95 % when tested in 46 plant condition. Grab Cut algorithm is used for the image segmentation process. Local binary pattern sets the pixel label. One class classifier is trained using LBP histograms. One class SVM is used for classifying the infected and the diseased leaf portion. One class support vector is used for classifying each class and conflict the resolution. Nearest support vector machine helps in the labeling process based on the proximity. Its novelty is achieved by testing through different plant species.

Sharif *et al.* [2018] The main source of nutrition are present in the citrus plant. Previously some image-based technique is used for identifying the citrus plant disease. The author proposed a hybrid technique for disease identification and classification in a citrus plant. It mainly focuses on identifying the lesion spot in citrus leaf and classification of the disease occurs in a citrus plant leaf. For identifying the lesion spot optimized weighing segmentation method is used. The hybrid technique helps in the process of classification. Entropy, PCA score, skewness are the factors considered during the process of selecting the feature. The final classification is done by Multi-Class Support Vector Machine. Anthracnose, black spot, melanize, scab, canker, and greening are the disease that can be detected from the citrus plant. The results show the accuracy of 97 % in classifying the disease. Image enhancement, infected region identification and segmentation, feature extraction and feature extraction are the steps involved in it. 10 fold cross validation is done during the evaluation process. The proposed technique is compared with

Ensemble Boosted Tree, Weighted K- Nearest Neighbour, Linear Discriminant Analysis and Decision Tree. The performance of the classification algorithm is accessed by TPR, FPR, FNR, PPV, and FDR.

dos Santos Ferreira *et al.* [2017] The loss in the production of potato is mainly due to the (Potyviridae, PVY) virus. Griffel proposed a technique based on the integration of machine learning and remote sensing technology to detect and differentiate the disease that occurs in the potato plant. Unmanned aerial system and unmanned ground vehicle provide high spatial devices to examine agriculture production. A number of spectral reflectances are found in the PVY affected plants. SVM classifier helps in classifying the infected PVY virus plant from the non-infected plant. The accuracy achieved during this process is 89.8 %. There is a drop in accuracy of 46.95 when RGB wavelength is used. SVM checks mainly whether it classify correctly infected from the non-infected plant. False Negative (FN) rate is calculated for analyzing its performance. SVM classification is done at different wavelength like 500-900 nm, 720-900nm, 720-1300 nm, and 900-1300nm. Highest accuracy is achieved when the wavelength is 500-900 nm and 720-1300nm.

5.2. Deep Learning Techniques for Plant Disease Prediction

dos Santos Ferreira *et al.* [2017] proposed a method to identify unwanted weeds in the soybean field. Unwanted weed includes unwanted grasses and broadleaf. Convolution neural network technique is applied in the process of identifying the weeds in the soybean field. For the purpose of capturing the image, drones were used in it. The database used for analyzing purpose includes fifteen thousand pictures weeds, soil, soybean, grass weed, and broadleaf. SafeNet architecture is used for training the neural network. The cafe software includes Alex Net in it. Pynovisao algorithm is used to build a robust image database. The results are compared with Support Vector Machine, Ada Boost, and Random Forest. The accuracy of 99 % is achieved using the convolution neural network. The steps involved in this identification process are as follows 1. Image capture using UAVs 2. Image segmentation using superpixel algorithm 3. Feature extraction based on shape, color and texture 4. Classifier training 5. Segmentation and classification. Superpixel algorithm (Simple Linear Iterative Clustering (SLIC) Superpixel) mainly focus on object localization and segmentation of the image.

Carranza-Rojas *et al.* [2017] proposed a technique for herbarium species identification using deep learning technique. It mainly focuses on how convolution neural networks help in automatic identification of plant species. Image-Net classification performs very well in convolution neural network process. Transfer learning is also used for domain related training. Results show a greater accuracy when it is trained and tested for a different set of species. It has been shown in it that by using herbarium dataset transfer learning is possible to another region even when the species don?t match. Handwritten tags and noise can be removed by the pre-processing technique. The transfer learning from herbarium to non-dried plants are clearly listed in the table.

Lu *et al.* [2017b] proposed a technique for identifying the pathogen in the vegetable. Deep convolution neural network technique is used for the identification of the rice disease. Training and testing the model consist of 500 images of rice leaves and stem with 10 types of rice disease in it. 10 fold cross validation method is used for identification of rice disease. The proposed novel model provides an accuracy of 95.48 %. The structure of 10 cross field deep convolution network consist of input (3@512 * 512), convolution (362@244 * 224), stochastic pooling (32@112 * 1112), convolution (16@56 * 56), stochastic pooling (16@28 * 28), convolution (16@28 * 28), stochastic pooling (16@14 *14), and two fully connected one. In the pre-processing stage scale normalization and mean normalization is done for color image and grey image and then PCA and whitening method is applied. Finally trained and tested feature map is plotted. Recognition accuracy for mean, max and stochastic pooling is as follows 92.11, 93.24,95.48 and the recognition accuracy for different filter(5*5, 9*9, 16*16, 32*32) are 93.15, 92.56, 93.29, 92.48. The proposed method is compared with BP, SVM, and PSO.

Barbedo [2019] proposed a technique based on deep learning for the purpose of image classification. Data augmentation technique helps in the lack of a database for plant image. This paper mainly focuses on identifying the individual lesion and spot instead of considering the whole leaf for identification. While using only lesion and spot the accuracy is 12 % higher than using the entire leaf. The complete details about the recent architecture used for identifying the plant disease and where the data are collected for identifying the plant disease and its accuracy after identification is also mentioned clearly. The list of disorder found in the plant specimen is also listed out clearly in it. Google Net CNN was used in the

experimental setup. In the experiment, three different types of images were used and they are 1.Image with-out any modification 2. Image with background removed 3. Expanded dataset. Accuracy for both original and expanded images are calculated.

Barbedo [2018] the problems faced in the machine learning technique has been overcoming by the deep learning concepts such as Convolution Neural Networks (CNN). Large data sets are needed for processing this technique. This paper mainly focuses on how the size of data and its variety affects the performance of the deep learning concepts. 12 plant species with different samples, different disease, and different character are taken into consideration. This analysis describes the different CNN network used for disease classification along with where this large amount of data are collected for classification. Accuracy is also calculated for each deep learning concepts. The number of correctly classified sample divided by the total number of samples provides the accuracy value. List of different plant species and its disease are listed in it. Removing background from image improves the prediction accuracy. This analysis was performed mainly using dataset obtained from different sources.

Tavakoli and Gebbers [2019] presented an analysis of winter wheat nitrogen and assessment of water in the field by using a camera. This experiment was conducted during a period of three years (2012, 2013, and 2014). Nitrogen fertilization and different level of water are applied in the field for the purpose of the experiment. Two machine learning algorithm was developed for the purpose of analysis namely Random Forest (RF) and Partial Least Square Regression (PLSR). Specter radiometer was used for radial measurement. Separately Vegetation Index (VI) is also calculated. For analyzing the nitrogen content R2 (RMSE) model is used and it is calculated separately for both data type. Random forest algorithm performs better in combined-date data. Nitrogen estimation calculation performs better while using the digital camera. It can also be integrated with the smartphone. It has a limitation of accessing only three spectral bands so that the analysis of plant status is also limited.

Kaya *et al.* [2019] some major issues are faced while classifying the data manually and they are expensive, time-consuming and required experts for this process. Deep Neural Network is proposed for solving this problem during the classification process. The author analyzed four types of transfer learning method applied in the deep neural network for the purpose of plant classification. The performance of plant classification model is improved. Comparative analysis of different method and its best results are listed in it like (DF-VGG16/LDA = 99.00, DF-Alex net/LDA = 96.20, CNN-RNN = 98.80, CNN = 99.60, (CNN, SVM = 97.47)). The proposed architecture of CNN consists of input, 3*3 conv, ReLu, pool, 3*3 conv, ReLu, pool ReLu, FC-class size, Softmax. The classification accuracy for each model for the training dataset and pre-trained model is listed in it.

Grinblat *et al.* [2016] proposed a method used for the identification of plant using leaf vein pattern. The classification of white bean, red bean and soybean are also done in this. Referred pipeline accuracy is also improved in it. The vein pattern is obtained by analyzing the visualization technique with the obtained results. The image processing is done in four different stages namely vein segmentation, central patch extraction, vein measure, and classification. Random forest, support vector machine, and penalized discriminant analysis algorithm are used for classification purpose. Central patch extraction and vein measure are replaced by the convolution neural network technique where it learns from the data set and solve this problem. In the proposed system of CNN, the depth of the model is increased from 2 layers to 6 layers. While analyzing the results it shows that the accuracy gets improved when we go deep into the layer at the 5th layer an accuracy of 92.6 is achieved.

Ferentinos [2018b] proposed a technique on convolution neural network. For the purpose of training, the model 87,848 images of healthy and diseased plant leaves are taken which includes 25 plant variety. These plants are tested under two different condition namely laboratory and field condition. Alex Net, AlexNetOWTBn, Google Net, over feat and VGG architecture are used for identification of plant disease from the leaves. Its implementation is done using Torch7. It is a machine learning framework. Its training portion is implemented in the Linux environment. 80 % of training data and 20 % testing data for CNN. 99.49 % success rate is achieved when using AlexNetOWTBn and 99.53 % of success rate is achieved using VGG model. The success rate for both the original image and the pre-processed image is analyzed as well for all the five models. The success rate is more when the model is first trained for field condition and then laboratory condition. The success rate is low when it is tested under laboratory condition and then field condition. It can be integrated with the mobile device due to low computational power.

Cuadro 3: Analysis of related works

| AUTHOR	PLANT SPECIES	ALGORITHM USED	ACCURACY
Siedliska *et al.* [2014]	Apple	NB	Cultivar- 86.7 %, Bruise- 83.3 %
Mursalin and Mesbah [2014]	Capsicum	NB	10 fold cross validation 98.9 %
Bandi *et al.* [2013]	Citrus	NB	95 %
Romualdo *et al.* [2014]	Corn	NB	V4- 82.5 %, V7- 87.5 %
Cen *et al.* [2016]	Cucumber	NB	Two class- (91.6-97.6 %), Three class- (81-85.7 %)
Asadollahi *et al.* [2009]	Tomatoes	NB	80.10 %
Piron *et al.* [2009]	Carrot	DA	72 %
Chen *et al.* [2010]	Corn	DA	94 ? 100 %
Gowen *et al.* [2009]	Mushroom	DA	Damaged- 97.9 % Un damaged- 100 %
Okamoto *et al.* [2007]	Sugar beet	DA	81.30 %
Zhang *et al.* [2017a]	Wheat	DA	77.90 %
Sun *et al.* [2017]	Peaches	DA	92.96 % - 97.28 %
Omrani *et al.* [2014]	Apple	SVM	96.30 %
Karimi *et al.* [2006]	Corn	SVM	Nitrogen- 81 %, Weed- 86 %
Camargo and Smith [2009]	Cotton	SVM	90 %
Meunkaewjinda *et al.* [2008]	Grapes	SVM	97.80 %
Islam *et al.* [2017]	Potato	SVM	95 %
Ahmed *et al.* [2012]	Chilli	SVM	Classification of weed 97 %
Miller *et al.* [1998]	Apple	k-NN	91.9-95.7 %
Li *et al.* [2014]	Blueberry	k-NN	85 ? 98 %
Sankaran and Ehsani [2013]	Citrus	k-NN	99 % - infrared and visible light
Rivera *et al.* [2014]	Mango	k-NN	94.87 - 98.08 %
Åstrand and Baerveldt [2002]	Sugar beet	k-NN	97 %
Guru *et al.* [2012]	Tobacco	k-NN	LBN - 72.75 %, LBP - 75.93 %, GLTP ? 80.91 %
Dubey and Jalal [2012]	Apple	k-means	Disease Detection k-means 93.17 %
Hu *et al.* [2014]	Banana	k-means	Visual categorization
Zhang *et al.* [2017b]	Cucumber	SVM	Disease Detection 85.70 %
Makky and Soni [2013]	Oil palm	k-means	93.53 %
Costa *et al.* [2009]	Orange	k-means	74.50 %
Olgun *et al.* [2016]	Wheat	k-means	88.33 %
Boydell and McBratney [2002]	Cotton	FCM	Not mentioned
Romeo *et al.* [2012]	Maize	FCM	97.32 %
Majumdar and Jayas [2000]	Wheat	FCM	56 %
Neto *et al.* [2006]	Soya bean	FCM	46 %
Meyer *et al.* [2004]	Sunflower and velvet	FCM	69 %
Guijarro *et al.* [2011]	Barley	FCM	91.60 %
De Rainville *et al.* [2014]	Soybean and corn	GGMM	Soybean- 85 %, Corn- 94 %
Dey *et al.* [2012]	Grapevines	GMM	98 %
Bauer *et al.* [2011]	Sugar beet	GMM	86 %- Fungus, 91 %- Leaf spot, 94 %- Healthy
Bai *et al.* [2013a]	Rice	GMM	85.80 %
Bai *et al.* [2013b]	Soybean	GMM	99 %
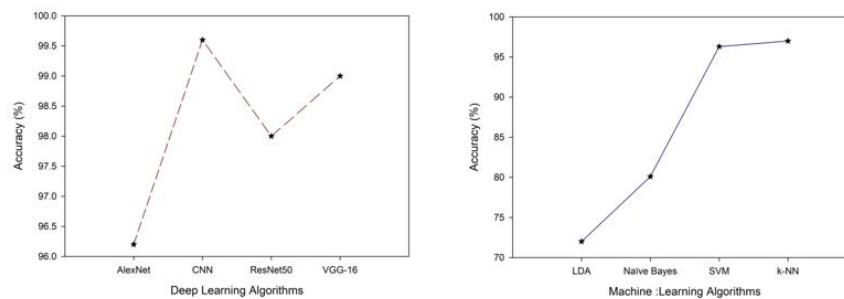
Figura 4: Performance of Machine learning and Deep learning algorithms on plant disease classification

6. Discussions

Plant disease diagnosis is an important process to determine the standard of the crop through which many factors such as yield capacity, richness of the grains, nutrition retention are evaluated. This paper aims to make a comprehensive study on various computational methods embedded in the plant disease identification and classification system. Many intelligent algorithms played a vital role in achieving the desired task. In Fig. 4, the best performing deep learning and machine learning algorithms are plotted to portray the performance of each algorithm. Also, many other fusion models were built to improve the predictability of the computational model and are discussed in detail. The outcome of this study reveals the significance of automated tools for assisting the end users to find the plant disease without the human intervention. In future, prescriptive models are to be developed, which is in great demand near in future.

7. Conclusion

The key objective of this work is to analyze different machine learning techniques widely used in the prediction of plant diseases and how advancement can be made in the future in this technique to achieve higher accuracy, robustness, cost-efficient disease prediction system. The steps involved in image processing techniques like pre-processing, segmentation, extracting feature and classification based on symptoms in the plant are discussed in this survey. Machine learning techniques play a key role in the machine vision system. In the future, deep learning framework can be used for disease prediction system. Integrating image processing techniques and deep learning techniques proved to be more potential in disease prediction system. Still, more investigations have to be made in these techniques for achieving better prediction system.

Acknowledgements

Referencias

Faisal Ahmed, Hawlader Abdullah Al-Mamun, ASM Hossain Bari, Emam Hossain, and Paul Kwan. Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection*, 40:98–104, 2012.

Hassan Asadollahi, Morteza Sabery Kamarposhty, and Mir Majid Teymoori. Classification and evaluation of tomato images using several classifier. In *2009 International Association of Computer Science and Information Technology-Spring Conference*, pages 471–474. IEEE, 2009.

Davoud Ashourloo, Hossein Aghighi, Ali Akbar Matkan, Mohammad Reza Mobasheri, and Amir Moeini Rad. An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9):4344–4351, 2016.

Björn Åstrand and Albert-Jan Baerveldt. An agricultural mobile robot with vision-based perception for mechanical weed control. *Autonomous robots*, 13(1):21–35, 2002.

Mohsen Azadbakht, Davoud Ashourloo, Hossein Aghighi, Soheil Radiom, and Abbas Alimohammadi. Wheat leaf rust detection at canopy scale under different lai levels using machine learning techniques. *Computers and Electronics in Agriculture*, 156:119–128, 2019.

XD Bai, Zhi-Guo Cao, Y Wang, ZH Yu, XF Zhang, and CN Li. Crop segmentation from images by morphology modeling in the cie l* a* b* color space. *Computers and electronics in agriculture*, 99:21–34, 2013.

XD Bai, Zhi-Guo Cao, Y Wang, ZH Yu, XF Zhang, and CN Li. Crop segmentation from images by morphology modeling in the cie l* a* b* color space. *Computers and electronics in agriculture*, 99:21–34, 2013.

Sudheer Reddy Bandi, A Varadharajan, and A Chinnasamy. Performance evaluation of various statistical classifiers in detecting the diseased citrus leaves. *International Journal of Engineering Science and Technology*, 5(2):298–307, 2013.

Jayme Garcia Arnal Barbedo. A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems engineering*, 144:52–60, 2016.

Jayme Garcia Arnal Barbedo. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and electronics in agriculture*, 153:46–53, 2018.

Jayme Garcia Arnal Barbedo. Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, 180:96–107, 2019.

Sabine D Bauer, Filip Korč, and Wolfgang Förstner. The potential of automatic methods of classification to identify leaf diseases from multispectral images. *Precision Agriculture*, 12(3):361–377, 2011.

B Boydell and AB McBratney. Identifying potential within-field management zones from cotton-yield estimates. *Precision agriculture*, 3(1):9–23, 2002.

Collins Bugingo. Efficacy of seed treatments, microbial and biochemical pesticides for managing early tan spot and stripe rust diseases of wheat. 2018.

A Camargo and JS Smith. Image pattern classification for the identification of disease causing agents in plants. *Computers and Electronics in Agriculture*, 66(2):121–125, 2009.

Jose Carranza-Rojas, Herve Goeau, Pierre Bonnet, Erick Mata-Montero, and Alexis Joly. Going deeper in the automated identification of herbarium specimens. *BMC Evolutionary Biology*, 17(1):1–14, 2017.

Haiyan Cen, Renfu Lu, Qibing Zhu, and Fernando Mendoza. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification. *Postharvest Biology and Technology*, 111:352–361, 2016.

Ross Chapman, Simon Cook, Christopher Donough, Ya Li Lim, Philip Vun Vui Ho, Koon Wai Lo, and Thomas Oberthür. Using bayesian networks to predict future yield functions with data from commercial oil palm plantations: A proof of concept analysis. *Computers and Electronics in Agriculture*, 151:338–348, 2018.

Xiao Chen, Yi Xun, Wei Li, and Junxiong Zhang. Combining discriminant analysis and neural networks for corn variety identification. *Computers and electronics in agriculture*, 71:S48–S53, 2010.

Anna Chlingaryan, Salah Sukkarieh, and Brett Whelan. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and electronics in agriculture*, 151:61–69, 2018.

David B Collinge, Hans JL Jørgensen, Meike AC Latz, Andrea Manzotti, Fani Ntana, Edward Camilo Rojas Tayo, and Birgit Jensen. Searching for novel fungal biological control agents for plant disease control among endophytes. *Endophytes for a Growing World*, pages 25–51, 2019.

Corrado Costa, Paolo Menesatti, Graziella Paglia, Federico Pallottino, Jacopo Aguzzi, Valentina Rimatori, Giuseppe Russo, Santo Recupero, and Giuseppe Reforgiato Recupero. Quantitative evaluation of tarocco sweet orange fruit shape using optoelectronic elliptic fourier based analysis. *Postharvest Biology and Technology*, 54(1):38–47, 2009.

François-Michel De Rainville, Audrey Durand, Félix-Antoine Fortin, Kevin Tanguy, Xavier Maldague, Bernard Panneton, and Marie-Josée Simard. Bayesian classification and unsupervised learning for isolating weeds in row crops. *Pattern Analysis and Applications*, 17(2):401–414, 2014.

Debadeepta Dey, Lily Mummert, and Rahul Sukthankar. Classification of plant structures from uncalibrated image sequences. In *2012 IEEE workshop on the applications of computer vision (WACV)*, pages 329–336. IEEE, 2012.

Gittaly Dhingra, Vinay Kumar, and Hem Dutt Joshi. A novel computer vision based neutrosophic approach for leaf disease identification and classification. *Measurement*, 135:782–794, 2019.

Alessandro dos Santos Ferreira, Daniel Matte Freitas, Gercina Gonçalves da Silva, Hemerson Pistori, and Marcelo Theophilo Folhes. Weed detection in soybean crops using convnets. *Computers and Electronics in Agriculture*, 143:314–324, 2017.

Shiv Ram Dubey and Anand Singh Jalal. Adapted approach for fruit disease identification using images. *International journal of computer vision and image processing (IJCVIP)*, 2(3):44–58, 2012.

MA Ebrahimi, MH Khoshtaghaza, Saeid Minaei, and B Jamshidi. Vision-based pest detection based on svm classification method. *Computers and Electronics in Agriculture*, 137:52–58, 2017.

Dhivya Elavarasan, Durai Raj Vincent, Vishal Sharma, Albert Y Zomaya, and Kathiravan Srinivasan. Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Computers and Electronics in Agriculture*, 155:257–282, 2018.

Konstantinos P Ferentinos. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145:311–318, 2018.

Konstantinos P Ferentinos. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145:311–318, 2018.

Aoife A Gowen, Masoud Taghizadeh, and Colm P OâDonnell. Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *Journal of Food Engineering*, 93(1):7–12, 2009.

Guillermo L Grinblat, Lucas C Uzal, Mónica G Larese, and Pablo M Granitto. Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture*, 127:418–424, 2016.

Marıa Guijarro, Gonzalo Pajares, Isabel Riomoros, PJ Herrera, XP Burgos-Artizzu, and Angela Ribeiro. Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1):75–83, 2011.

DS Guru, PB Mallikarjuna, S Manjunath, and MM Shenoi. Machine vision based classification of tobacco leaves for automatic harvesting. *Intelligent Automation & Soft Computing*, 18(5):581–590, 2012.

Meng-han Hu, Qing-li Dong, Bao-lin Liu, and Pradeep K Malakar. The potential of double k-means clustering for banana image segmentation. *Journal of Food Process Engineering*, 37(1):10–18, 2014.

Ryan HL Ip, Li-Minn Ang, Kah Phooi Seng, JC Broster, and JE Pratley. Big data and machine learning for crop protection. *Computers and Electronics in Agriculture*, 151:376–383, 2018.

Zahid Iqbal, Muhammad Attique Khan, Muhammad Sharif, Jamal Hussain Shah, Muhammad Habib ur Rehman, and Kashif Javed. An automated detection and classification of citrus plant diseases using image processing techniques: A review. *Computers and electronics in agriculture*, 153:12–32, 2018.

Monzurul Islam, Anh Dinh, Khan Wahid, and Pankaj Bhowmik. Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th Canadian conference on electrical and computer engineering (CCECE)*, pages 1–4. IEEE, 2017.

Alexander Johannes, Artzai Picon, Aitor Alvarez-Gila, Jone Echazarra, Sergio Rodriguez-Vaamonde, Ana Díez Navajas, and Amaia Ortiz-Barredo. Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Computers and electronics in agriculture*, 138:200–209, 2017.

Andreas Kamilaris, Andreas Kartakoullis, and Francesc X Prenafeta-Boldú. A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143:23–37, 2017.

Kerim Karadağ, Mehmet Emin Tenekeci, Ramazan Taşaltın, and Ayşin Bilgili. Detection of pepper fusarium disease using machine learning algorithms based on spectral reflectance. *Sustainable Computing: Informatics and Systems*, 2019.

Y Karimi, SO Prasher, RM Patel, and SH Kim. Application of support vector machine technology for weed and nitrogen stress detection in corn. *Computers and electronics in agriculture*, 51(1-2):99–109, 2006.

Aydin Kaya, Ali Seydi Keceli, Cagatay Catal, Hamdi Yalin Yalic, Huseyin Temucin, and Bedir Tekinerdogan. Analysis of transfer learning for deep neural network based plant classification models. *Computers and electronics in agriculture*, 158:20–29, 2019.

Sami Khanal, John Fulton, Andrew Klopfenstein, Nathan Douridas, and Scott Shearer. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and electronics in agriculture*, 153:213–225, 2018.

Louis Kouadio, Ravinesh C Deo, Vivekananda Byrareddy, Jan F Adamowski, Shahbaz Mushtaq, et al. Artificial intelligence approach for the prediction of robusta coffee yield using soil fertility properties. *Computers and Electronics in Agriculture*, 155:324–338, 2018.

Anakha Krishnakumar and Athi Narayanan. A system for plant disease classification and severity estimation using machine learning techniques. In *International Conference on ISMAC in Computational Vision and Bio-Engineering*, pages 447–457. Springer, 2018.

Matjaž Kukar, Petar Vračar, Domen Košir, Darko Pevec, Zoran Bosnić, et al. Agrodss: A decision support system for agriculture and farming. *Computers and Electronics in Agriculture*, 161:260–271, 2019.

Han Li, Won Suk Lee, and Ku Wang. Identifying blueberry fruit of different growth stages using natural outdoor color images. *Computers and electronics in agriculture*, 106:91–101, 2014.

Qiaokang Liang, Shao Xiang, Yucheng Hu, Gianmarc Coppola, Dan Zhang, and Wei Sun. Pd2se-net: Computer-assisted plant disease diagnosis and severity estimation network. *Computers and electronics in agriculture*, 157:518–529, 2019.

Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, and Changshui Zhang. An in-field automatic wheat disease diagnosis system. *Computers and electronics in agriculture*, 142:369–379, 2017.

Yang Lu, Shujuan Yi, Nianyin Zeng, Yurong Liu, and Yong Zhang. Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267:378–384, 2017.

S Majumdar and DS Jayas. Classification of cereal grains using machine vision: Iii. texture models. *Transactions of the ASAE*, 43(6):1681, 2000.

Muhammad Makky and Peeyush Soni. Development of an automatic grading machine for oil palm fresh fruits bunches (ffbs) based on machine vision. *Computers and electronics in agriculture*, 93:129–139, 2013.

A Meunkaewjinda, P Kumsawat, K Attakitmongcol, and A Srikaew. Grape leaf disease detection from color imagery using hybrid intelligent system. In *2008 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology*, volume 1, pages 513–516. IEEE, 2008.

George E Meyer, Joao Camargo Neto, David D Jones, and Timothy W Hindman. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and electronics in agriculture*, 42(3):161–180, 2004.

WM Miller, JA Throop, and BL Upchurch. Pattern recognition models for spectral reflectance evaluation of apple blemishes. *Postharvest Biology and Technology*, 14(1):11–20, 1998.

Md Mursalin and Md Mesbah. Towards classification of weeds through digital image. In *2014 Fourth International Conference on Advanced Computing & Communication Technologies*, pages 1–4. IEEE, 2014.

Joao Camargo Neto, George E Meyer, and David D Jones. Individual leaf extractions from young canopy images using gustafson–kessel clustering and a genetic algorithm. *Computers and electronics in agriculture*, 51(1-2):66–85, 2006.

Hiroshi Okamoto, Tetsuro Murata, Takashi Kataoka, and SHUN-ICHI HATA. Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed Biology and Management*, 7(1):31–37, 2007.

Murat Olgun, Ahmet Okan Onarcan, Kemal Özkan, Şahin Işik, Okan Sezer, Kurtuluş Özgişi, Nazife Gözde Ayter, Zekiye Budak Başçiftçi, Murat Ardiç, and Onur Koyuncu. Wheat grain classification by using dense sift features with svm classifier. *Computers and Electronics in Agriculture*, 122:185–190, 2016.

Avital Oliver, Augustus Odena, Colin A Raffel, Ekin Dogus Cubuk, and Ian Goodfellow. Realistic evaluation of deep semi-supervised learning algorithms. In *Advances in neural information processing systems*, pages 3235–3246, 2018.

Elham Omrani, Benyamin Khoshnevisan, Shahaboddin Shamshirband, Hadi Saboohi, Nor Badrul Anuar, and Mohd Hairul Nizam Md Nasir. Potential of radial basis function-based support vector regression for apple disease detection. *Measurement*, 55:512–519, 2014.

Xanthoula Eirini Pantazi, Dimitrios Moshou, and Alexandra A Tamouridou. Automated leaf disease detection in different crop species through image features analysis and one class classifiers. *Computers and electronics in agriculture*, 156:96–104, 2019.

Keunho Park, Young ki Hong, Gook hwan Kim, and Joonwhoan Lee. Classification of apple leaf conditions in hyper-spectral images for diagnosis of marssonina blotch using mrmr and deep neural network. *Computers and Electronics in Agriculture*, 148:179–187, 2018.

Santanu Phadikar, Jaya Sil, and Asit Kumar Das. Rice diseases classification using feature selection and rule generation techniques. *Computers and electronics in agriculture*, 90:76–85, 2013.

Alexis Piron, Vincent Leemans, Frédéric Lebeau, and M-F Destain. Improving in-row weed detection in multispectral stereoscopic images. *Computers and electronics in agriculture*, 69(1):73–79, 2009.

Nayeli Vélez Rivera, Juan Gómez-Sanchis, Jorge Chanona-Pérez, Juan José Carrasco, Mónica Millán-Giraldo, Delia Lorente, Sergio Cubero, and José Blasco. Early detection of mechanical damage in mango using nir hyperspectral images and machine learning. *Biosystems Engineering*, 122:91–98, 2014.

J Romeo, G Pajares, M Montalvo, JM Guerrero, M Guijarro, and A Ribeiro. Crop row detection in maize fields inspired on the human visual perception. *The Scientific World Journal*, 2012, 2012.

LM Romualdo, PHC Luz, FFS Devechio, MA Marin, AMG Zúñiga, OM Bruno, and VR Herling. Use of artificial vision techniques for diagnostic of nitrogen nutritional status in maize plants. *Computers and electronics in agriculture*, 104:63–70, 2014.

Sindhuja Sankaran and Reza Ehsani. Comparison of visible-near infrared and mid-infrared spectroscopy for classification of huanglongbing and citrus canker infected leaves. *Agricultural Engineering International: CIGR Journal*, 15(3):75–79, 2013.

Muhammad Sharif, Muhammad Attique Khan, Zahid Iqbal, Muhammad Faisal Azam, M Ikram Ullah Lali, and Muhammad Younus Javed. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and electronics in agriculture*, 150:220–234, 2018.

Anna Siedliska, Piotr Baranowski, and Wojciech Mazurek. Classification models of bruise and cultivar detection on the basis of hyperspectral imaging data. *Computers and Electronics in Agriculture*, 106:66–74, 2014.

Ye Sun, Xinzhe Gu, Ke Sun, Haijiang Hu, Miao Xu, Zhengjie Wang, Kang Tu, and Leiqing Pan. Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches. *Lwt*, 75:557–564, 2017.

Hamed Tavakoli and Robin Gebbers. Assessing nitrogen and water status of winter wheat using a digital camera. *Computers and Electronics in Agriculture*, 157:558–567, 2019.

Paula Tennant, Gustavo Fermin, and Jerome E Foster. *Viruses: Molecular Biology, Host Interactions, and Applications to Biotechnology*. Academic Press, 2018.

Graham M Turner, Kirsten A Larsen, Seona Candy, Sue Ogilvy, Jaithri Ananthapavan, Marj Moodie, Sarah W James, Sharon Friel, Chris J Ryan, and Mark A Lawrence. Squandering australiaâs food securityâthe environmental and economic costs of our unhealthy diet and the policy path weâre on. *Journal of cleaner production*, 195:1581–1599, 2018.

Jingcheng Zhang, Ning Wang, Lin Yuan, Fengnong Chen, and Kaihua Wu. Discrimination of winter wheat disease and insect stresses using continuous wavelet features extracted from foliar spectral measurements. *Biosystems Engineering*, 162:20–29, 2017.

Shanwen Zhang, Xiaowei Wu, Zhuhong You, and Liqing Zhang. Leaf image based cucumber disease recognition using sparse representation classification. *Computers and electronics in agriculture*, 134:135–141, 2017.