

Review on Emerging Trends in Detection of Plant Diseases using Image Processing with Machine Learning

Punitha Kartikeyan
Ph.D. Research Scholar
School of Information Technology
MATS University, Raipur
Chhattisgarh, India

Gyanesh Shrivastava
Head of the Department
School of Information Technology
MATS University, Raipur
Chhattisgarh, India

ABSTRACT

In India, about 70% of the population is involved in agriculture and farming. Today plant diseases are significant concern as it reduces the production and quality of agriculture produce. Most plant diseases are caused by bacteria, fungi and virus. Manual detection and identification of leaf disease involve more man power and expensive in large farm. Detection of disease and healthy monitoring of plant are major challenges for sustainable agriculture. Hence there is need to detect plant disease automatically using image processing technique at an early stage with more accuracy. It involves image acquisition, image pre-processing, image segmentation, feature extraction and classification of disease. To increase the quality of produce and yield, a regular monitoring technique for detection of diseases in plants becomes essential. The digital image processing system is one such powerful technique to diagnose the difficult symptoms much earlier than the human eye could recognize. It enables the farmers to take appropriate actions timely in order to safeguard the crop and get the desired quality and yield of agriculture produce. Different techniques used for the classification of plant disease using various classifiers such as Support Vector Machine, Artificial Neural Network, K-Nearest Neighbors and other classifier methods have been discussed. The purpose of this paper is to give an overview of established methods for plant disease detection, classification systems.

Keywords

Image Processing, Disease Detection, Segmentation, Feature Extraction, Classification

1. INTRODUCTION

India is an agriculture country and majority of its population depends on agriculture as this is the source of income as well as job employment. Agriculture crops heavily suffer from losses due to diseases that affect leaves, crops, stem etc. The conventional way of leaf disease detection and identification involve naked eye observation by experts. This requires skill and time and is not feasible for monitoring large farms. Hence, automatic detection of leaf disease is essential as it helps to overcome the drawbacks of manual detection. Further, earlier detection of leaf diseases is also possible which prevents huge losses.

2. METHODOLOGY

The detection of plant disease involves five major steps viz., Image acquisition, Image Pre-processing, Image segmentation, Feature extraction and Classification. In image processing, acquisition of images is done through digital camera or scanner, image pre-processing involves image enhancement, image segmentation where the affected and healthy areas are

segmented, feature extraction defines the area of infection and classification helps to detect the type of diseases. Table.1 shows the framework of the system.

Table 1. Framework for plant disease detection

1. Image Acquisition
2. Image Pre-processing
3. Image Segmentation
4. Feature Extraction
5. Classification

2.1. Image Acquisition

The initial process is to collect the image data which are used as input for further processing. The input of image data should be in .bmp, .jpg, .png, .gif format.

2.2. Image Pre-processing

Pre-processing techniques are applied to improve the image data such as image cropping technique to change the image size and shape, image smoothing for filtering of noise and image enhancement to enhance contrast and colour conversion of images.

2.3. Image Segmentation

Segmentation of leaf image involves partitioning of image into different parts of same features. The segmentation can be done using various methods like Otsu's, k-means clustering, thresholding, region and edge based methods etc.

2.4. Feature Extraction

Feature extraction plays an important role in the classification of diseases. In many applications, feature extraction is carried out by the color, texture, and morphology in plant disease classification. Texture means how the color is distributed in the image, the roughness, the hardness of the image. Morphological feature extraction is better than color and texture feature for detecting leaf diseases.

2.5. Disease Classification

Classification is the challenging task in image processing technique. The main purpose of the classification is to correctly predict the value of a designated discrete class variable, given a vector of predictors or attributes. In plant disease detection, classification is done based on whether the image is infected or not.

3. LITERATURE REVIEW

In the past, many researchers have done their work to detect the disease automatically, quickly with more accuracy using

different classification methods.

3.1 Support Vector Machine

Support vector machine (SVM) is a supervised learning algorithm also used for classification, regression and outlier detection of plant diseases. In this algorithm [1], each data item as a point in n-dimensional space where n is number of features with value of each feature being the value of a particular coordinate. By using hyper-plane as shown in Fig 1, to maximize the classification boundaries for identifying two classes separately. The adjacent points of hyper-plane were recognized as support vectors. To make a separation between these vectors, a fixed maximum margin is used, through which the SVM can be trained effectively.

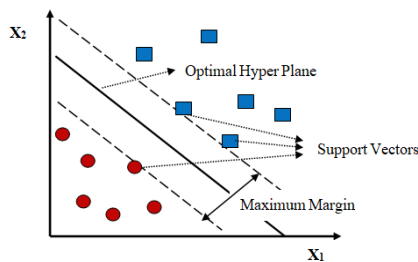


Fig 1. Classification using SVM

The SVM technique widely used by many researchers for the classification of plant diseases is discussed below:

Youwen et al. [2] identified Powdery Mildew and Downey Mildew diseases in Cucumber leaves. They applied median filter to remove the noise and used statistic pattern and mathematics morphology to segment the leaf images. The features of cucumber leaf image disease were extracted and these features were fed to SVM classification for identification of disease. Experimental results indicated that the classification performance by SVM was better than that of neural networks. Shape and texture features in SVM technique was better than of using only shape feature.

Meunkaewjinda et al. [3] proposed a diagnosis system for detecting the Grape leaf diseases. The system comprised of three main parts viz., first part involved grape leaf color extraction from complex background, secondly grape leaf disease color extraction and finally grape leaf disease classification. For analysing, Back-Propagation Neural Network with a Self-Organizing Feature Map together was used to recognize colors of grape leaf. Further, MSOFM and GA (Genetic Algorithm) were used for grape leaf disease segmentation and SVM for its classification. Finally, filtration of resulting segmented image was done by Gabor Wavelet and then SVM was again applied to classify the types of grape leaf diseases. This system could classify the grape leaf diseases into three classes: Scab disease, Rust disease and no disease.

Two stage SVM based MCS (multiple classifier system) for detection of disease in Wheat leaves was proposed by Tian et al. [4]. The colour, texture and shape feature were utilized for training the classifiers. GLCM (Gray Level Cooccurrence Matrix) was used to extract texture features, whereas, invariant moments were used as shape features.

Zhang and Zhang [5] used SVM to recognize Cucumber diseases Downey Mildew, Brown and Leaf Spots. They separated the disease spots from the background and threshold concept was applied in segmentation and three features color,

shape and texture were retrieved and these features were submitted to classifier that are SVM based on Kernel functions with Radial Basis Function (RBF), Polynomial Function (PF) and Sigmoid Function (SF). The authors stated that the RBF and SF kernel function reported better results as compared to PF and SF kernel functions.

Prasad et al. [6] attempted to distinguish different crop diseases using image processing algorithms. They reported that imaging framework comprised of ailment spot recognition using histogram based segmentation, extraction of Gabor Wavelet Transform (GWT) features and classification by SVM classifiers. This approach yielded 89% accuracy.

Pawar et al. [7] developed an automatic grading and sorting system for Pomegranate fruits. Color, texture feature analysis was used for detection of surface defects on pomegranates fruits. Important features were used as an input to SVM classifier and tests were performed to identify best classification model.

Arivazhagan et al. [8] developed an image analysis technology in which images of various leaves were acquired using a digital camera to identify the diseases from the symptoms of the leaves. The digital RGB images of leaves were converted into HSI color image in that the green pixels were masked and removed using specific threshold value followed by the segmentation process and the statistics of texture were computed. The SVM was used to classify the disease using texture features. They reported that an accuracy of 94.74% was possible in the identification of disease infection in plant leaves.

Kiran et al. [9] attempted to detect Citrus leaf disease such as Anthracnose, Citrus Canker. They pre-processed the image that involved color space conversion by applying YCbCr color system and L*a*b* color space. Color image enhancement by applying discrete cosine transform. Used K-mean clustering for disease part segmentation. In feature extraction, GLCM was used to see various statistics such as energy, contrast, homogeneity and entropy. Finally, the citrus leaf disease classification was done by SVM RBF and SVM POLY that resulted 96% and 95% accuracy respectively.

Singh and Mishra [10] detected plant leaf diseases of Banana, Beans, Lemon and Rose using image segmentation and soft computing techniques. They used genetic algorithm for segmentation that worked effectively. For classifying the disease, minimum distance criteria with Kmean, minimum distance criteria with GA and SVM with GA algorithms were used and these methods resulted 86.54%, 93.63% and 95.71% accuracy respectively.

Jayamoorthy and Palanivel [11] proposed Spatial Fuzzy C Mean (SFCM) to detect plant disease and also compared with other clustering techniques K-means, Fuzzy C-Means (FCM) and Kernel based FCM (KFCM). Then the features such as color, texture were extracted from diseased leaf image. The neural network method was used to classify the diseases. The proposed method gave better performance and suggested required pesticide for the diseases.

Akhtar et al. [12] tried to detect and classify the Black Spot and Anthracnose disease in Rose leaf using the SVM approach. They adopted the threshold method for segmentation and Otsu's algorithm was used to define the threshold values. In this approach, features of Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT) and eleven haralick texture

features were extracted and further used with SVM approach. This method showed accuracy value of 94.45% in leaf disease detection.

The concept of detection and classification of Apple fruit diseases namely Scab, Apple Rot and Apple Blotch was explored by Dubey and Jalal [13]. The segmentation was carried out using K-means clustering technique and the features were extracted from the segmented images. They used Multiclass SVM for classification to achieve an accuracy of 93%.

Mokhtar et al. [14] detected the Powdery Mildew and Early Blight Tomato leaves diseases. They carried out image pre-processing that involved various techniques such as smoothness, removal of noise, image resizing, image isolation and background removing for image enhancement. They applied Gabor Wavelet Transformation in feature extraction for feature vectors and also in classification. Cauchy Kernel, Laplacian Kernel and Invmult Kernel were also applied in SVM for output decision and training for disease identification. The author reported that this method could yield 99.5% accuracy.

Yao et al. [15] employed SVM method to classify Rice Bacterial Leaf Blight, Rice Sheath Blight, and Rice Blast. They proposed segmentation of rice disease spots, and extracting the shape and texture features from these segments. The results showed that SVM could effectively detect and classify these disease spots to an accuracy of 97.2%.

Support Vector Regression (SVR) based on radial basis functions to identify and classify diseases of the Apple tree was proposed by Omrani et al. [16]. This process involved captured images of the leaves had to be changed into a device independent color space such as CIELAB, from a device depended format such as Red-Green-Blue (RGB) color space, then the image was segmented to extract the infected area from the overall leaf image. The segmentation technique employed was a region-based one using K-means clustering, wavelet, and grey-level co-occurrence matrix. The features extracted include the color, shape, and texture. Finally, the segmented image was classified using SVR-RBF, SVR-POLY with ANN. They concluded that SVR model gave better results than ANN.

Table 1: Detail of SVM technique for Detection and Classification of Plant Diseases

Author & Year	Classification Methods	Plant	Disease	Result
Youwen et al [2] (2008)	SVM	Cucumber	Powdery Mildew, Downey Mildew	Better accuracy than ANN method
Menukaewjinda et al [3] (2008)	SVM	Grape	Scab and Rust Disease	Better accuracy
Tian et al [4] (2010)	SVM	Wheat	Different Diseases	Good accuracy
Zhang Jian and Zhang Wei [5] (2010)	SVM with different kernel functions	Cucumber	Downey Mildew, Brown and Leaf spots	Radial Basis Kernel function gives good result than other kernel functions.
Prasad et al [6] (2012)	SVM	Crop	Diseases	Around 89%
Meenakshi Pawar et al [7] (2012)	SVM	Pomegranate	Sorting and Grading of Pomegranate Fruit	Better accuracy
Arivazhagan et al [8] (2013)	SVM	Banana, Beans, Jackfruit, Lemon, Mango, Potato, Tomato	Late Scorch, Bacterial Spot, Fungal Spot, Sun Burn, Sooty Mold, Early Blight, Late Blight, Scorch, Ashen Mold, Leaf Lesion	94.74%
Kiran Gavhale et al [9] (2014)	SVM RBF, SVM POLY	Citrus	Citrus Canker, Anthracnose	96% and 95%
Singh and Mishra [10] (2016)	Minimum Distance Criteria with Kmean and GA, SVM with GA	Banana, Beans, Lemon, Rose	Bacterial Leaf Spot, Frog Eye Leaf Spot, Sun Burn, Early Scorch	86.54, 93.63% and 95.71%
Jayamoorthy and Palanivel [11] (2017)	Spatial Fuzzy C Mean (SFCM), SVM	Plants	Bacterial Blight, Foot Rot	Better accuracy
Akhtar et al [12] (2013)	DCT+DWT+SVM	Rose	Black Spot, Anthracnose	94.45%
Dubey and Jalal [13] (2012)	MSVM	Apple	Scab, Apple Rot, Apple Blotch	93%
Mokhtar et al [14] (2015)	SVM with Cauchy kernel,	Tomato	Powdery Mildew, Early Blight	99.5%.

	InvmultKernel and Laplacian Kernel			
Yao et al [15] (2009)	SVM	Rice	Rice Bacterial Leaf Blight, Rice Sheath Blight, Rice Blast	97.2%.
Omrani et al [16] (2014)	SVR RBF, SVR POLY, ANN	Apple	Alternaria, Apple Black Spot, Apple Leaf Miner	SVR is better than ANN

3.2 Artificial Neural Network

Artificial Neural Network (ANN) is basically an engineering approach of biological neuron, the number of nodes in ANN is called as neurons. ANN is organized layer by layer. They are represented in the form of input, hidden and output layers. In network, each neuron in hidden layer receives signals from other neurons of input layer[1]. The strength of each signal and biases are represented by weights and constants that are calculated by training phase. After the inputs are weighted and added, the result is transformed by transfer function to the output. Sigmoid, tanh etc are used as activation

functions. The Fig. 2 represents the architecture of ANN.

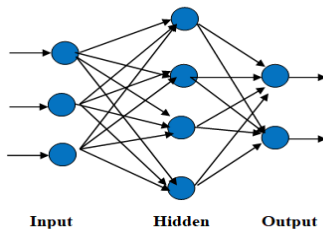


Fig 2. Classification using ANN

The work done on the classification of diseases in plants using ANN is discussed below:

Phadikar and Sil [17] proposed a method to identify the Rice crop diseases such as Blast and Brown Spots. The image was converted to the HSI color space, followed by entropy-based thresholding and an edge detector was applied to the segment the image for identifying the infected parts of the leaf. The features were fed into Self Organizing Map (SOM) neural network to classify the diseases. This technique resulted in an accuracy of 92% in detecting the disease.

Al Bashish et al. [18] developed a technique to identify five diseases such as Early and Late Scorch, Cottony Mold, Ashen Mold, Tiny Whiteness diseases of the plant. Images were divided into clusters by using a k-means clustering algorithm after image pre-processing. Feature extraction process used color co-occurrence method for extracting color and texture features. These features were fed to a Multi Layer Perceptron (MLP) Neural Network with ten hidden layers, which performed the final classification. Precision of around 93% was achieved in detecting the disease through this method.

Al-Hiary et al. [19] used Otsu segmentation and K-means clustering procedure to identify the infected regions of plant and stem diseases. They used color co-occurrence methodology for the features extraction of infected parts. The ANN was used for detecting and classifying the diseases like Early Scorch, Cottony Mold, Ashen Mold, Late Scorch, Tiny whiteness and White Mold. The proposed technique was robust and yielded a precision value between 83% and 94%.

Song et al. [20] studied Corn diseases detection through image

processing technique. They used YCbCr color space for image segmentation of disease spots and spatial gray level matrix to extract features. They extracted energy, entropy, moment of inertia, related and local stationary texture features and fed to Back Propagation neural network to recognize and classify the Corn Leaf Blight, Sheath Blight and Southern Leaf Blight diseases. The experimental results showed that the algorithm could effectively identify the disease to an accuracy of as high as 98%.

Revathi and Hemalatha [21] applied homogeneity based edge-detector technique for detecting Leaf Spot disease in Cotton. The symptoms of the disease were captured and neural network was used to classify the disease. The homogeneity operator calculates the difference between the core pixel and the neighbouring pixels and divides the output of any edge detector by the average output area. They identified disease like Fusarium Wilt, Verticillium Wilt, Root Rot, Boll Rot, Grey Mildew, Leaf Blight, Bacterial Blight, Leaf Curl in Cotton plant. Classification of diseases was easier and excellent accuracy was achieved through this technique.

Asfarian et al. [22] used texture analysis with fractal descriptors based on Fourier Spectrum method to identify Paddy diseases like Leaf Blast, Brown Spot, Bacterial Leaf Blight, and Tungro. The disease affected images were manually extracted, later converted into HIS color space from which the saturation components were finally extracted. They performed histogram equalization and used Laplacian filter for reducing the lighting effect and to sharpen images respectively. At the end, fractal descriptors were extracted from each lesion and fed to Probabilistic Neural Networks (PNN) for classify the disease. They concluded that this fractal descriptor method could be employed to identify four common paddy diseases with 83% accuracy.

The two-layer feed-forward network with color, shape, texture, boundary and entropy features were used by Bindu and Toran [23] to identify the Apple diseases. The neural network with sigmoid function yielded accurate result for healthy and disease leaf of apple.

To detect diseases at an early stage with more accuracy in plants, Hong and Huijie [24] used a combination of image processing technique and Back Propagation (BP) neural network. Median filter technique was used for removing the noise of an image. They extracted seven HU invariant moment eigen values, three shape eigen values and eight texture eigen values features. They proposed BP network, which was optimized by L-M (Local Minima) algorithm to identify the classes of the plant diseases such as Cercospora Leaf Spot, Canker, Rust and Mosaic Virus. This method gave an accuracy 99%.

Sumathi and SenthilKumar [25] detected and classified plant leaves diseases. In this study, MLP batch propagation was proposed. MLP with incremental back propagation and Levenberg–Marquardt based learning were used in existing

method and was compared with the proposed method for estimating the accuracy values. Among the methods studied, batch propagation method yielded good accuracy than back propagation and L-M based learning method. The authors reported that an accuracy 93% could be achieved by this technique.

Kiran and Gavhale [26] , in this paper comprehensive study of plant leaves disease detection system was done for improving the productivity of crops. They compared the benefits and limitations of these potential methods such as KNN, BPNN, SVM and PNN.

Cotton leaf diseases such as Bacterial Blight, Myrothecium Leaf Spot and Alternaria were identified and classified using a pattern recognition system by Rothe and Kshirsagar [27]. Active contour model was used for image segmentation and Hu's moments were extracted as features for the training of adaptive neuro fuzzy inference system. This method resulted in 85% accuracy.

Leaf disease detection and grading were performed by Rastogi et al. [28] through an image processing and machine vision based technology. Image pre-processing of leaf was done and segmentation was determined by K-means clustering and Euclidean distance technique. The GLCM matrix feature such as contrast, energy, homogeneity, and correlation were extracted. The classification of disease was identified by ANN. For disease grading, percentage infection was calculated by using total leaf area (AT) and diseased area (AD). Percentage infection grading was done by fuzzy logic.

Yadav and Verma [29] used hybrid approach of Neural Network and Genetic Algorithm to recognize Black Mold disease in Tomato. After the image pre-processing, region growing algorithm was applied for disease part segmentation. Then color and texture features were extracted and genetic algorithm was applied for optimization of features. These features were forwarded to the Backpropagation Neural Network (BPN) for detecting the disease. Thus this classifier was used to obtain better accuracy.

Khirade and Patil [30] recorded various image processing and machine learning techniques applied in plant diseases detection. In this processing, pre-processing methods such as clipping, image smoothing, image enhancement, and histogram equalization were presented. Segmentation techniques such as boundary detection and spot detection, K-means clustering, and Otsu's threshold were also briefly described. Features extraction such as color, texture, morphology, and edges were also dealt. Classification based on ANN with the working principle was discussed.

Wang et al. [31] presented Grape and Wheat plant disease. Total 185 images including 85 Grape leaves and 100 Wheat

leaves were considered. They compressed the images using nearest neighbour interpolation method without changing the image resolution. To de-noise the images, they applied median filter and K-means clustering was used for segmentation. For image segmentation, RGB image was converted into CIE XYZ color space, followed by conversion of XYZ color space into l^*a^*b color space. The color, shape, and texture features were extracted. Principal Component Analysis (PCA) was applied to reduce the dimensions of the feature data extracted from images. This technique helped to reduce the number of neurons and thereby increased the speed of Backpropagation neural network to achieve an accuracy of 97%

The concept of neural network for the classification of cotton leaf disease like Red Spot, White Spot, Yellow Spot, Alternaria and Cercospora was studied by Bhog et al [32]. For segmentation and experimentation K-means clustering and MATLAB toolbox were used respectively. The recognition accuracy for K-mean clustering method using euclidean distance was 89.56% and the execution time for K-means clustering method using euclidean distance was 436.95 second.

Kutty et al. [33] proposed a method to classify Anthracnose and Downey Mildew disease in Watermelon leaf. The infected leaves were identified from normal one based on RGB color component. In segmentation, median filter was used to reduce noise while for classification, neural network pattern recognition toolbox was used. By this method, 75.9% accuracy could be achieved .

Sannaki et al. [34] aimed to diagnose the disease using image processing and artificial intelligence techniques on images of Grape plant leaf. They classified Downy Mildew and Powdery Mildew of grape leaf disease. Masking was used to remove background to improve accuracy. Anisotropic Diffusion was used for preserving information of affected portion of leaf. Segmentation was carried out by K-means clustering method and feature extraction was carried out by calculating GLCM. Finally classification was done using Feed Forward Back Propagation Network classifier. They have used only Hue feature which gives more accurate result.

Principal Component Analysis (PCA) and neural networks for disease identification and recognition in Wheat crop and Grape plants were recommended by Wang et al. [35]. They used PCA to extract twenty one colors, four shapes, and twenty five texture features from the plants. After the feature extraction process, various types of neural networks were used for identification and classification of the diseases including Back Propagation (BP-NN), Radial Basis Function (RBF-NN), Generalized Regression (GR-NN), and Probabilistic Neural Networks (PNN). The results of different neural networks were compared.

Table 2: Detail of ANN technique for Detection and Classification of Plant Diseases

Author & Year	Classification Methods	Plant	Disease	Result
Phadikar and Sil [17](2008)	SOM with Neural Network	Rice	Rice Blast, Brown Spot	92%
Al Bashish et al [18] (2010)	MLP Neural Network	Plants	Early and Late Scorch, Cottony Mold, Ashen Mold, Tiny Whiteness	Precision is around 93%.
Al-Hiary et al [19] (2011)	ANN	Plant and	Early Scorch, Cottony Mold, Ashen Mold, Late Scorch,	Precision between 83%

		Stem	Tiny Whiteness , White Mold.	and 94%
Song et al [20] (2011)	BPNN	Maize	Corn Leaf Blight, Sheath Blight , Southern Leaf Blight	Accuracy 98 % more
Revathi and Hemalatha [21](2012)	NN	Cotton	Fusarium and Verticillium Wilt ,Root and Boll rot , Grey Mildew ,Leaf and Bacterial Blight, Leaf Curl	Excellent accuracy
Asfarian et al [22] (2013)	PNN	Paddy	Leaf Blast, Brown Spot, Bacterial Leaf Blight	83.00%
Bindu and Toran [23] (2013)	ANN	Apple	Infected or Normal	Good accuracy
Hong and Huijie [24] (2014)	BPNN with LM optimization	Plant	Cercospora Leaf Spot, Canker, Rust , Mosaic Virus	99%
Sumathi and SenthilKumar [25] (2014)	MLP with Batch Propagation	Plant	Common Diseases	93%
Kiran and Gavhale [26] (2014)	BPNN,PNN, SVM, KNN	Plant	Bacterial Leaf Spot , Mosaic Virus, Late Blight, Early Blight, Downy Mildew	Good
Rothe and Ksirsagar [27](2015)	Neuro-Fuzzy Inference System	Cotton	Bacterial Blight, Leaf Spot, Alternaria	85%
Rastogi et al. [28](2015)	ANN, Fuzzy logic	Plant	Leaf Spot	Good
Yadav and Verma [29] (2016)	BPNN and GA	Tomato	Black Mold	Better accuracy
Khirade and Patil [30] (2015)	ANN	Plant	Diseases	Good
Wang et al. [31] (2012)	BPNN	Grape Wheat	Powdery Mildew , Downey Mildew, Wheat Stripe Rust , Wheat Leaf Rust	Above 97%
Bhog and Pawar [32] (2016)	Neural Network	Cotton	Red Spot, White Spot, Yellow Spot, Alternaria, Cercospora	89.56%
Kutty et al.[33] (2013)	Neural Network	Watermelon	Anthracoise, Downey Mildew	75.9%
Sannaki et al. [34](2013)	Feed Forward BPNN	Grape	Powdery Mildew ,Downey Mildew	Better accuracy
Wang et al. [35] (2012)	PCA, BPNN, RBF-NN, GR-NN , PNN	Wheat Grape	Wheat Stripe Rust and Leaf Rust, Grape Downy Mildew and Powdery Mildew	Better accuracy

3.3 K Nearest Neighbour

KNN algorithm is used to classify by finding the K nearest matches in training data followed by using the label of closest matches to predict. Traditionally, distance such as euclidean is used to find the closest match. It is also known as a lazy learner because it simply stores all the training tuples given to it as inputs in its learning phase without performing any calculations[1]. This feature prevents it from being used in areas where dynamic classification is needed for large databases. The star test data point is bounded through the circles, and squares that differentiate between classes as shown in Fig. 3. The three essential aspects of K-NN classification which includes easy interpretation of output results, computational time is short and prediction rates are high.

Many researchers have used the K-NN approach for detecting plant diseases that are discussed below:

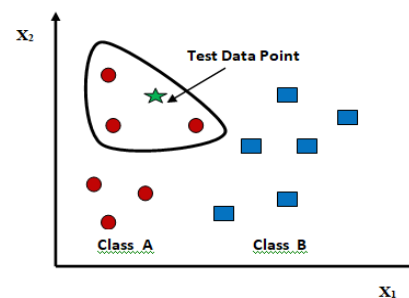


Fig 3. Classification using KNN

Xu et al. [36] demonstrated a technique to identify Nitrogen and Potassium deficiencies in Tomato plant. Fourier transform, and Wavelet form for texture feature extraction were used, while the color feature extraction was done using LAB color space. The selected features were extracted by a genetic algorithm for optimization and fuzzy K-nearest neighbour classifier were applied to identify nitrogen and potassium deficiencies. They reported that the system diagnosed to an accuracy above 82.5% .

Paddy plant diseases like Blast and Brown Spot were identified by Suresha et al. [37] using KNN classification method. About 300 images were used as input, Otsu with global threshold technique was

used for segmenting the diseased images. After segmentation, feature extraction with geometrical features like area, major axis, minor axis and perimeter of diseased part of paddy plant was extracted. Finally, KNN technique was used for classification of diseases that resulted in 76.59% accuracy.

Diseases Alternaria Alternata, Anthracnose, Bacterial Blight, Leaf Spot, and Canker of various plants were detected by Hossain et al. [38] using KNN classification technique. About 237 leaf images were acquired from the Arkansas plant disease

database for input process. After segmentation, GLCM matrix was used for feature extraction of diseased plant images and to prevent overfitting, the 5-fold cross validation was applied on the training dataset. This method provided 96.76% accuracy.

Abdulridha et al. [39] identified Avocado plants diseases such as Laurel Wilt, Phytophthora Root Rot as well as Nutrient deficiencies like Iron and Nitrogen. After image processing, overall canopy region of interest (OVROI) and polygon region of interest (PROI) were used for segmentation of plants. Then MLP and K-NN classification techniques were used to identify the disease type and deficiency of nutrient in plants. MLP approach achieved better classification results than the KNN.

Table 3: Detail of KNN technique for Detection and Classification of Plant Diseases

Author & Year	Classification Methods	Plant	Disease	Result
Xu et al. [36] (2011)	Fuzzy KNN	Tomato	Nitrogen , Potassium Nutrient Deficiencies	Accuracy was above 82.5% .
Suresha et al.[37](2017)	KNN	Plant	Diseases	76.59%
Hossain et al[38] (2019)	KNN	Plant	Alternaria Alternata, Anthracnose, Bacterial Blight, Leaf Spot, Canker	96.76%
Abdulridha et al [39] (2019)	KNN, MLP	Avacoda	Laurel Wilt, Phytophthora Root Rot, Iron and Nitrogen Nutrient Deficiencies	MLP achieved better accuracy than KNN

3.4. Other Classifiers

Yan Cheng Zhang et al. [40] tried to identify Cotton disease and developed the Fuzzy Feature selection approach, Fuzzy Curves (FC) and Surfaces (FS) to select features of diseased leaves image. They showed that the effectiveness of features selected by the FC and FS method was much better than that selected by human randomly or other methods due to FC could quickly isolate important and significant subset features, while, FS could eliminate those features which were dependent on other important features.

Pang et al. [41] identified diseases of Maize crops such as Leaf Blight, Cercospora Leaf Spot, Leaf Shealths and Brown Spots. The algorithm identified all pixels for which the level of the red channels (R) gray level was higher than the level of the green channel (G) gray levels. The red pixels contributed 98% of the diseased region; the connected regions were identified and labelled appropriately, the connected regions as seeds and applied a region growing technique to detect diseased regions more accurately and efficiently to an extent of 96%.

An algorithm to detect and classify Rice Plant-Hopper (RPH) infestation was proposed by Zhou et al. [42]. They acquired the infected Rice stem images and then applied smoothing, denoising, color space transformation, followed by box-counting dimension algorithm for extracting Eigen values. Linear Regression Model (LDM) and Fuzzy C-means Clustering (FCM) techniques were applied for final identification and classification. They concluded that use of fractal Eigen values resulted in more accuracy in differentiating RPH infected images.

Nitin and Sanjay [43] used color transformation structure to explain texture statistics for detecting the plant leaf disease. RGB was converted into HSV. They used pre-computed threshold level, masking and removing of green pixels of leaves

was done, segmentation was done to identify the healthy and un-healthy regions of the leaf. These segments were used for texture analysis by color co-occurrence matrix. SGDM texture parameters of infected leaves were compared with texture parameters of normal leaf. Finally the occurrence of diseases on the plant leaf was evaluated.

Qinghai et al. [44] used the histogram and histogram equalization techniques for extracting the disease infected Cotton leaves. They designed a color and image enhancement analyzes to identify the pests and disease affected cotton leaves. They used high resolution digital camera for capturing the the images of uninfected and infected cotton leaves and processed by digital image processing techniques. The RGB color images were converted into gray scale, HIS and YCbCr images and were plotted for infected and uninfected gray scale cotton images The comparisons of results proved very good accuracy in extracting the disease infected cotton leaves and YCbCr color space was considered as the best color model for extracting the damaged image.

In an attempt to detect Cotton Mite disease, Zhihua and Huan [45] used modern histogram technique. The RGB color images were converted into gray scale. Firstly, the disease spots were extracted from green plants followed by histograms were plotted for all healthy and cotton mite disease leaves. Image segmentation achieved by color features and area thresholding. The diseased areas were calculated and compared. This technique resulted in 94.79% accuracy in terms of separation of the healthy region and cotton mite disease infected region.

A system for disease identification and grading in Pomegranate leaf and fruit was suggested by Deshpande et al. [46]. For segmentation, K-means clustering method was used. They

determined infection percentage grade of the disease is dividing total disease area by total leaf or fruit area was calculated. They reported that, this system provided accurate and acceptable results.

Wang et al. [47] aimed to recognize the Stripe Rust and Powdery Mildew diseases in Wheat. They designed a new version of Rotational Kernel Transformation directional (IRKT) feature for disease recognition in the spatial domain. The images were converted from RGB space to HSV and YCbCr color spaces. The Otsu's method was performed for segmentation process. The kernel based IRKT directional feature extracted from segmented areas, were used to detect edges. They reported that this method resulted 97.5 % recognition accuracy.

Narvear and Patil [48] performed novel algorithms based image processing in Grape leaf diseases detection. The RGB image was converted into HSV format in the image pre-processing followed by feature extraction method was carried out using SGDM method. Five features were calculated like energy, homogeneity, contrast, cluster prominence and cluster shade. Since segmentation technique was not applied, they could not classify diseases such as Powdery Mildew and Downy Mildew properly.

Bhavini and Sheshang [49] used Random Forest classifier to identify Apple fruit diseases like Scab, Rot and Blotch. For image segmentation, K mean clustering was applied to identify infected regions and features were extracted with the help of fusion technique. The accuracy of the diseases classification was improved by feature level fusion.

Lu et al. [50] described an in-field automatic Wheat disease diagnosis system. This involved weekly supervised deep learning framework to achieve an integration of identification of Wheat diseases and localization of disease areas using image-level annotation for training images in wild conditions. Wheat Disease Database 2017 (WDD2017) was collected to verify the effectiveness of the system. They used two different architectures, i.e. VGG-FCNVD16 and VGG-FCN-S that resulted an accuracy of 97.95% and 95.12% respectively over 5-fold cross validation on WDD2017, exceeding the results of 93.27% and 73.00% by two conventional CNN frameworks, i.e. VGG-CNN-VD16 and VGG-CNN-S. They concluded the proposed system was better than conventional CNN architectures.

Mrinal Kumar et al. [51] detected plant leaf diseases using image processing technique with K-mean clustering, GLCM feature extraction, PNN classification methods. After that classification method, which indicated the disease name, the affected region percentage and accuracy was calculated.

4. COMPARATIVE ANALYSIS

The models are evaluated based on their average accuracy of results and the performance. This work presented different approaches using machine learning for plant disease detection and classification. From the tables (1, 2 and 3) and fig 4 it can be seen that out of several models compared, the SVM method gave the highest accuracy followed by ANN in detecting and classification of plant disease.

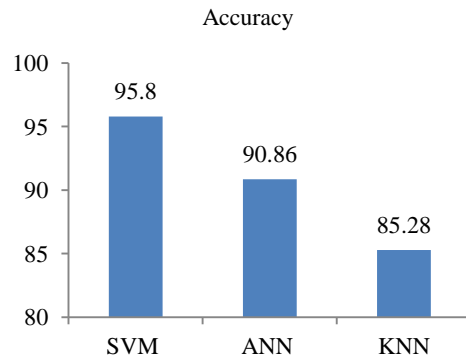


Fig 4. Comparative Analysis

5. CONCLUSION AND FUTURE WORK

Plant disease detection is very momentous and important research area. Further, the accurate disease detection and classification of the plant leaf image is very important for the successful cultivation and reduce the loss of crops. During the review, it was observed that classification techniques were widely used for the identification and detection of diseases in plant leaves. Among the classification techniques, many researchers have reported SVM and ANN methods gave the highest accuracy and recommended for plant disease detection.

In future work, the development in hybrid algorithms by using genetic algorithms, ant colony, cuckoo optimization, and particle swarm optimization with SVM, ANN, and KNN would increase the efficiency in plant disease detection. Mobile application can be developed with inbuilt remedial solution that can be used by farmers to detect any kind of leaf, stem, fruit, flower disease including nutrient deficiency easily. Deep learning is most powerful technique that has promising features for accurate plant disease detection and classification than machine learning technique.

6. REFERENCES

- [1] V. Sharma, A. Verma and N. Goel, "Classification Techniques for Plant Disease Detection", International Journal of Recent Technology and Engineering (IJRTE), 2020, ISSN: 2277-3878, vol-8, Issue-6
- [2] T. Youwen, L. Tianlai and N. Yan, "The recognition of cucumber disease based on image processing and support vector machine", In proceeding of 2008 Congress on image and signal processing, IEEE, Sanya, 262-267, DOI: 10.1109/CISP.2008.29.
- [3] A. Meunkaewjinda, P. Kumsawat and K. Attakitmongcol, Sri kaew, "Grape leaf disease detection from color imagery using hybrid intelligent system", proceedings of ECTI-CON 2008, IEEE, Krabi, pp 513-516, DOI:10.1109/ecticon.2008.4600483
- [4] Y. Tian, C. Zhao, S. Lu and X. Guo, "SVM- based Multiple Classifier System for Recognition of Wheat Leaf Diseases", Proceedings of 2010 Conference on Dependable Computing (CDC'2010), 2010, Yichang, China.
- [5] J. Zhang and W. Zhang, "Support Vector Machine for Recognition of Cucumber Leaf Diseases", In Proceeding of 2nd International Conference on Advanced Computer Control (ICACC), IEEE, Shenyang, 2010, 590-594. DOI: 10.1109/ICACC.2010.5487242.
- [6] S. Prasad, P. Kumar, R. Hazra and A. Kumar, "Plant Leaf Disease Detection Using Gabor Wavelet Transform", Proceedings of 3rd International Conference

- on Swarm, Evolutionary and Memetic Computing, Springer-Verlag, Berlin, Heidelberg, 2012.
- [7] M. M. Pawar, S. Bhusari and A. Gungewar, "Identification of infected pomegranates using color texture feature analysis", *International journal of computer application*, (0975-8887), 2012.
- [8] S. Arivazhagan, S. Newlin, S. Ananthi and S.V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features", *Agricultural Engineering International: CIGR Journal*, 2013, vol.15, no.1, pp.211-217.
- [9] R. Kiran Gavhale, U. Gawande, Kamal and O. Hajari, "Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques", *IEEE International Conference for Convergence of Technology*, 2014, pp. 978-1-4799-3759-2/14.
- [10] V. Singh and A.K. Mishra, "Detection of plant leaf diseases using image segmentation and soft computing techniques", *Information Processing in Agriculture 4 (1)*, 2016
- [11] S. Jayamoorthy and N. Palanivel, "Identification of Leaf Disease Using Fuzzy C-MEAN and Kernel Fuzzy C-MEAN and Suggesting the Pesticides", *International Journal of Advanced Research in Science, Engineering and Technology*, May 2017, Vol. 4, Issue 5.
- [12] Akhtar, Asma, A. Khanum, Shoab A. Khan and A. Shaukat, "Automated Plant Disease Analysis (APDA): Performance Comparison of Machine Learning Techniques", *IEEE International Conference on Frontiers of Information Technology (FIT)*, pp. 60-65, 2013.
- [13] S. R. Dubey and A.S. Jalal, "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns", *IEEE Third International Conference on Computer and Communication Technology*, pp. 978-0-7695-4872, 2012.
- [14] U. Mokhtar, A.S. Mona, Alit, A. E. Hassenian and H. Hefny, "Tomato leaves diseases detection approach based on support vector machines", *IEEE pp. 978-1-5090-0275-7/15*, 2015.
- [15] Q. Yao, Z. Guan, Y. Zhou, J. Tang; Y. Hu and B. Yang, "Application of support vector machine for detecting rice diseases using shape and color texture features", *International Conference on Engineering Computation*, 2009, pp.79-83.
- [16] E. Omrani, B. Khoshnevisan, S. Shamshirb and, H. Saboohi, N. B. Anuar and M. H. N. Nasir "Potential of radial basis function-based support vector regression for apple disease detection", *Measurement*, vol.55, pp.512–519, Sept-2014
- [17] S. Phadikar and J. Sil, "Rice Disease Identification using Pattern Recognition Techniques", *Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008)*, 25-27 December, 2008, IEEE, Khulna, Bangladesh, pp. 420-423.
- [18] D. Al Bashish, M. Braik and S. Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases", *In proceeding of 2010 International conference on signal and image processing. IEEE, Chennai*, 113-118.
- [19] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. AlRahamneh, "Fast and accurate detection and classification of plant diseases", *International journal of computer applications (0975-8887)*, 2011, vol.17- no.1, pp-31-38.
- [20] K. Song, Z. Liu, H. Su and C. Guo, "A Research of maize disease image recognition of Corn Based on BP Networks", *Third international conference on Measuring Technology and Mechatronics Automation, IEEE*, 2011:246-249, DOI:10.1109/ICMTMA.2011.66
- [21] P. Revathi and M. Hemlatha, "Homogeneous Segmentation Based Edge Detection Techniques for Proficient Identification of the Cotton Leaf Spot Diseases", *International Journal of Computer Applications*, 2012, vol 47, no 2, pp 0975-888.
- [22] A. Asfarian, Y. Herdiyeni, A. Rauf and KH. Mutaqin, "Paddy Diseases Identification with Texture Analysis using Fractal Descriptors Based on Fourier Spectrum", *In proceeding of International Conference on Computer Control Informatics and Its Applications, IEEE, Jakarta*, 77-81, 2013,
- [23] T. Bindu and V.Toran, "Identification and classification of normal and infected apples using neural network", *International Journal of Science and Research*, 2013, vol.2, no.6, pp.160-163.
- [24] F. Hong and L. Huijie, "Plant leaves recognition and classification model based on image features and neural network", *IJCSI International Journal of Computer Science*, 2014, vol.11, no.1, pp.100-104.
- [25] C.S. Sumathi and A.V. Senthil Kumar, "Enhancing accuracy of plant leaf classification techniques", *International Journal of Engineering Research and Applications*, 2014, vol.4, no.3, pp.40-46.
- [26] R Kiran Gavhale and U. Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques", *IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727*, vol 16, Issue 1, 2014, pp 10-16.
- [27] P.R. Rothe and R. V. Kshirsagar, "Cotton Leaf Disease Identification using Pattern Recognition Techniques", *International Conference on Pervasive Computing (ICPC)*, IEEE, 2015.
- [28] A. Rastogi, R. Arora and S. Sharma, "Leaf Disease Detection and Grading using Computer Vision Technology and Fuzzy Logic", *IEEE 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, Noida, 2015, pp 500-505, 2015.
- [29] M. Yadav and T. Verma, "Hybrid Approach of Neural Network and Genetic Algorithm to Recognize Black Mold Disease in Tomato", *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*. 2016, Vol 4, Issue 5.
- [30] S. D. Khirade and A.B. Patil, "Plant Disease Detection Using Image Processing", *In International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2015 (pp. 768- 771). IEEE.
- [31] H. Wang, G. Li, Z. Ma, and X. Li, "Image recognition of plant diseases based on backpropagation networks", *In 5th International Congress on Image and Signal Processing (CISP)*, 2012 pp. 894-900, IEEE.
- [32] Bhong, B. V. Pawar and S. Vijay, "Study and Analysis of Cotton Leaf Disease Detection Using Image Processing." *International Journal of Advanced Research in Science, Engineering and Technology*, 2016, vol 3, Issue 2
- [33] S.B Kutty, N.E Abdullah, H. Hashim, A. Afhzan Ab Rahim, A. S. Kusim, T.N. Tuan Yaakub, Puteri N. A. M. Yunus and M. F. A. Rahman, "Classification of Watermelon Leaf Diseases Using Neural Network Analysis", *IEEE, Business Engineering and Industrial Applications Colloquium (BEIAC)*, Langkawi, pp 459 – 464, 2013.

- [34] S. Sanjeev Sannaki, S. Vijay, Rajpurohit, V and B. N. P. Kulkarni "Diagnosis and Classification of Grape Leaf Diseases using Neural Network", IEEE, Tiruchengode, pp1 – 5, 2013.
- [35] H. Wang, G. Li, Z. Ma and X. Li, "Image Recognition of Plant Diseases Based on Principal Component Analysis and Neural Networks", 8th International Conference on Natural Computation, 2012, pp.246-251
- [36] G. Xu, F. Zhang, S.G. Shah, Y. Ye and H. Mao, "Use of leaf color images to identify nitrogen and potassium deficient tomatoes", 2011, Pattern Recognition Lett vol.32, 1584-1590.
- [37] M. Suresha, K. N. Shreekanth and B. V. Thirumalesh, "Recognition of diseases in paddy leaves using knn classifier," 2nd International Conference for Convergence in Technology (I2CT), Mumbai, 2017, pp. 663-666. DOI: 10.1109/I2CT.2017.8226213.
- [38] E. Hossain, M. F. Hossain and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier," International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 2019, pp. 1-6.
- [39] J. Abdulridha, R. Ehsani, A. Abd-Elrahman and Y. Ampatzidis, "A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses", Computers and electronics in agriculture, 2019, 156, 549-557
- [40] Y.C Zhang, H.P Mao, B. Hu and M.X Li, "Features selection of cotton disease leaves image based on fuzzy feature selection techniques", IEEE Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, China, 2-4 Nov 2007.
- [41] J. Pang, Bai Zy, Lai Jc and Li Sk, "Automatic segmentation of crop leaf spot disease images by integrating local threshold and seeded region growing", International conference on image analysis and signal processing, IEEE, Hubei, 2010, pp 590–594.
- [42] Z. Zhuo, Y. Zang, Y. Li, Y. Zhang, P. Wang and X. Luo, "Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means", Mathematical and Computer Modelling 58(2011), 701–709.
- [43] P.K. Nitin and B.D Sanjay, "Crop Disease Detection Using CBIR", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS), 2013.
- [44] H. Qinghai, M. Benxue, Q. Duanyang, Z. Qiang, H. Xinmin and Z. Jing, "Cotton pests and diseases detection based on image processing", Telkomnika Indonesian Journal of Electrical Engineering, 2013, vol.11, no.6, pp.3445-3450.
- [45] D. Zhihua and W. Huan, "Image segmentation method for cotton mite disease based on color features and area thresholding", Journal of Theoretical and Applied Information Technology, 2013, vol. 48, no.1, pp.527-533.
- [46] T. Deshpande, S. Sengupta and K.S. Raghuvanshi, "Grading & Identification of Disease in Pomegranate Leaf and Fruit", International Journal of Computer Science and Information Technologies, IJCSIT, 2014, vol. 5 (3), pp 4638-4645.
- [47] L. Wang, F. Dong, Q. Guo, C. Nie and S. Sun, "Improved Rotational Kernel Transformation Directional Feature for Recognition of Wheat Stripe Rust and Powdery Mildew", In proceeding of Seventh International Conference on Image and Signal Processing, IEEE, Dalian, 286-291, DOI: 10.1109/CISP.2014.7003793
- [48] P. Narvekar and S. N. Patil, "Novel Algorithm for Grape Leaf Diseases Detection," International Journal of Engineering Research and General Science, 2015, vol.3, pp 1240-1244.
- [49] J. Bhavini Samajpati and D. Sheshang Degadwala, "Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier", IEEE International Conference on Communication and Signal Processing, 2016, pp. 978-5090-0396, DOI:10.1109/ICCSP.2016.7754302.
- [50] J. Lu, J. Hu, G. Zhao, F. Mei and C. Zhang, "An in-field automatic wheat disease diagnosis system", Computers and Electronics in Agriculture 142, 2017, 369–379.
- [51] M. Kumar, T. Hazra and S. S. Tripathy, "Wheat Leaf Disease Detection Using Image Processing", International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS), Apr 2017, vol 6, Issue 4.