RESEARCH ARTICLE OPEN ACCESS

A Survey on Plant Disease Identification

Vipinadas. M. J [1], Thamizharasi. A [2]
M.Tech Student [1], Assistant Professor [2]
Department of Computer Science and Engineering
Mohandas College of Engineering & Technology
Nedumangad, Anad, Thiruvananthapuram
Kerala – India

ABSTRACT

Plant disease is a wreckage of the normal state of a plant. It interrupts or modifies its vital function. Plant disease causes, loss of crops that may result in hunger and starvation especially in less developed countries. Automatic detection of plant disease helps in monitoring the crop and automatically detects the symptoms of the disease as soon as they appear on plant leaves. Most plant diseases are caused by fungi, bacteria and viruses. Image processing techniques are used to detect various plant diseases. This includes several steps viz. input images, image preprocessing, and extraction of features and classify them on the different basis. It uses many classification techniques such as K Nearest Neighbor classifier, k-means Classifier, Neural Networks, Support Vector Machine, Artificial neural network, and Fuzzy logic. Selecting a best classification method is cumbersome because the efficiency of result can vary based on input data. This paper focuses on different image processing techniques along with different classification techniques used for plant leaf disease classification.

Keywords:- Automatic leaf disease detection, NN, BPNN, SVM, Image processing.

I. INTRODUCTION

Agriculture is a base of the world economy and for India, the role of agriculture in the economy is much more than other Product. World Agricultural production is affected by the annual loss of about 20% to 30% on an average due to plant diseases. From the survey of NABARD (National Bank for agriculture and rural development) in 2013, India loses about 30% of its crops due to diseases each year. The loss due to these is estimated to be Rs.60, 000 cores annually. So, if agriculture affected by any disease it will directly or indirectly affect our economy and the population which is dependent on agriculture. It needs to analyze plant diseases very accurately within the specific time.

The Early system uses bare eye based plant disease identification. But that is not an effective method so image processing techniques can make it task easy to process all kind of disease images very accurately. It also gives a system with automatically detect the diseases without having an expert on the field. Photographic images of plant disease symptoms are used for the plant disease identification and in research, teaching and diagnostics etc.

Plant pathologists can arrange these digital images using digital image tools for identification of plant diseases. Images often do not contain sufficient details to assist in diagnosis; resulting in wastage of time also, it leads to the imprecise diagnosis. Farmers experience great pressure and also in changing from one disease control policy to another policy. Computer processing Systems are developed for agricultural applications, such as detection of leaf diseases, fruits diseases etc. In all these techniques, digital images are collected using a digital camera and image processing techniques are applied on these images to extract useful information that are necessary for further analysis. The main aim of this paper is to concentrate on the plant leaf. Disease detection based on the texture, color and shape of the Leaf Shows several advantages over flowers and fruits.

Rest of the paper is listed as follows. Section II presents plant disease symptoms. Section III includes the brief review on various image processing techniques. Then section IV concludes this paper along with possible future directions.

II. PLANT DISEASE SYMPTOMS

Following are describes some common symptoms of fungal, bacterial and viral plant leaf diseases.

A. Bacterial disease symptoms:

The disease is mainly referred to as "bacterial leaf spot" Symptoms begin as small, yellowgreen lesions on young leaves which usually seen as deformed and twisted, or as dark, water-soaked, greasy-appearing lesions on older foliage as shown in Fig.1.

Fig.1.Bacterial disease on leaf

B. Viral disease symptoms:

All virus disease presents some degree of reduction in production and the length of life of virus infected plants is usually short. The most accessible symptoms of virus-infected plants are usually those appearing on the leaves, but some viruses may cause strike on the leaves, fruits and, roots. The Viral disease is very difficult to diagnose. In Fig.2. Leaves are seen as wrinkled, curled and growth may be stunted due to the virus.

Fig. 2. Viral disease on leaf

C. Fungal disease symptoms:

It is a type of plant pathogen and is responsible for the serious plant diseases. Most diseases in vegetable are caused by fungi. They damage plants by killing cells. The main Source of fungal disease is the infected seed, soil, crop, and weeds. It is spread by the wind and water and through the movement of contaminated soil, animals, workers, machinery, tools. In, initial stage it appears on lower or older leaves as water-soaked, gray-green spots. Later, these spots darken and then white fungal growth spread on the undersides. In downy mildew yellow to white streak on the upper surfaces of older leaves occurs. These areas are surrounded with white to Gray fungal growth on the undersides as shown in Fig. 3. In leaf-late blight, water-soaked lesions are present. Initially, it seen as a small brown or black spot on the under leaves later it expanded over the whole region shown in figure Fig.4. Early blight is a fungal disease caused by the fungus Alternaria solani shown in Fig. 5. In the early time, it shown on the lower side, older leaves like small brown spots with concentric rings that form a bull's eye pattern. When disease rate increases, it spreads outward on the leaf surface causing it to turn yellow.

Fig.3. Fungal disease on leaf-downy mildew

Fig. 4. Fungal disease on leaf-late blight

Fig.5. Fungal disease on leaf-early blight

III. LITERATURE SURVEY

P. R. Rothe and R. V. Kshirsagar [1] proposed an Active Contour model (Snake segmentation) technique for segmenting the diseased region from the cotton leaf. Hu's moments [26] are used as the features for the classification. For training and classification, it uses a set of seven moments and Back Propagation Neural network has been used for classification with an accuracy of 85.52%. Back-propagation neural networks are highly efficient for solving Multiple Class problems. Its weight is updated using Levenberg Marquardt Optimization. The proposed methods can be applied to other crops like orange, citrus, wheat, corn and maize etc. Aakanksha Rastogi, Ritika Arora and Shanu Sharma [2] suggested a Fuzzy system for leaf disease detection and grading. K-means clustering technique has been used for segmentation. which groups similar pixels of an image. RGB color space is converted to L*a*b space, where L is the luminosity and a*b are the color space. The reason for this conversion is that luminosity factor is not important for the color image. GLCM [27] matrix including contrast, correlation, energy and homogeneity has been measured for disease grading. Artificial Neural Networks as been used for training the data. Fuzzy logic is used for grading the disease.

SmitaNakwadi and NiketAmoda[3] recommended a k-means clustering technique for segmentation. RGB has been converted to HIS, where H is the hue, I indicate the intensity and S indicate the saturation value. Color Co-occurrence method or CCM method has been used for color feature extraction. Plant disease is detected using Histogram matching. The Threshold value for the pixel is computed using Otsu's method.

S. S. Sannakki and V. S. Rajpurohit [4] suggested a Back-propagation Neural Network based classifier (BPNN) for detecting the disease in Pomegranate leaf. Features have been selected as color and texture. BPNN detects and classifies the diseases with a precision of around 97.30 %.

Dr. K. Thangadurai and K. Padmavathi [5] recommended computer vision image enhancement for leaf disease identification. It includes color conversion and Histogram equalization. Histogram equalization increases the image clarity. RGB to Grayscale conversion is used to retain the luminance information rather than Hue and Saturation information. For encoding of linear intensity values, Gamma expansions are used. Cumulative Gaussian distribution function distributes the intensity value of the image. Histogram Equalization provides the better quality image in Grayscale.

YuanTian, ChunjiangZhao, ShenglianLu and XinyuGuo[6] proposed an SVM-based Multiple Classifier System (MCS)[25] for wheat leaf diseases. It uses a stacked generalization structure to join the classification decisions obtained from three kinds of support vector machines (SVM) based classifiers. The features like color, texture and shape features are used as training sets for classifiers. Firstly, features are classified using a classifier in low-level of MCS to corresponding mid-level categories, which can partially detect the symptom of crop diseases according to the knowledge of plant pathology. Then the mid-level

features are generated from these mid-categories generated from low-level classifiers. Finally, high-level SVM has been trained and correct errors made by the color, texture and shape SVM [24] to improve the performance of detection. Compared with other classifiers, it can provide better success rate of detection. The classifiers like SVM Artificial Neural Network classifier, k-nearest neighbor (k-NN) classifier's, the MCS can obtain better recognition accuracy than others classifiers. Color, texture and shape SVMs [24] to improve the performance of detection. Compared with other classifiers, it can provide better success rate of detection. The classifiers like SVM, Artificial Neural Network classifier, k-nearest neighbor (k-NN) classifier's, the MCS can obtain better recognition accuracy than others classifiers.

Neetu Chahal and Anuradha[7] proposed a C Means Clustering Approach for the identification leaf disease. Neural Networks are used for the classification.

Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa [9] recommended a Machine learning system which converts RGB to HSV, RGB to L*a*b*. The Shape was also taken as Connectivity opening [11] which are used to calculate all the components in each threshold image. Shape features selected as Area of minimum enclosing rectangle, elongation, small compactness, small perimeter and Moment of Inertia. It uses seven classifiers such as Nearest Neighbors [12], Decision tree [13], [14], Random forest [15], [16], Extremely Randomized Trees [17], Naive Bayes [18] and support vector classifier (Linear SVM and RBF SVM) [19], [20], [21], [22].

TITLE	METHOD	CLASSIFIER	ADVANTAGES	DISADVANTAG-ES
Cotton Leaf Disease Identification using Pattern Recognition Techniques [1].	Uses Snake segmentation, Hu's moments are taken as feature.	Back Propagation Neural Network	Active contour model used to minimize the energy inside the disease spot, BPNN solves the multiple class problems, average classification is found to be 85.52%.	Snake segmentation is a very slow process.
Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic [2].	K-means clustering used to segment the defected area; GLCM is used for the extraction of Texture features, Disease grading using fuzzy logic.	Artificial Neural Network	Severity of the disease is checked, Fast and highly efficient.	Low-level segmentation.

Classification of Pomegranate Diseases Based on Back Propagation Neural Network [4].	K-means clustering used to segment the defected area, color and texture are used as the features,	Back Propagation Neural Network	RGB image is converted to L*a*b to extract chromaticity layers of image, classification is found to be 97.30%.	Only applicable for limited crops.
Automated VisionBased Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease [8].	Color histograms are extracted And transformed from RGB to HSV, RGB to L*a*b*. Peak components are used to create max tree, Five shape attributes are used, seven classifiers are used, and area under the curve (AUC) analysis is used for classification.	Nearest Neighbors ,Decision tree , Random forest , Extremely Randomized Trees , Naive Bayes and support vector classifier (Linear SVM and RBF SVM)	In seven classifiers, Extremely Randomized Trees Yield a very high score, provide realtime information, Provide flexibility to the application since farmers are able to move with their phones to the fields	It is not useful when the system is off-line, needs essing more proc time.
Feature Extraction for Identification of Sugarcane Rust Disease [9].	Texture, shape and color features are extracted, Gray Level Cooccurrence Matrix (GLCM) used to extracts the second order statistical texture features,	Support vector machine classifier(SVM)	combination of texture and color feature extraction with polynomial kernel results in 98.5 %	Only efficient when considering the narrow leaves.
SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases [6].	Color features are represented in RGB to HIS, Gray Level Cooccurrence Matrix (GLCM) is used, seven invariant moments are taken as shape parameter,	Support vector machine classifier(SVM)	Off-line two stages SVM based MCS for the detection of wheat leaf diseases.	Segmentation is difficult.
Advances In Image Processing For Detection Of Plant Diseases [3].	K-means clustering used to segment the defected area, masking green pixels, SGDM matrix generation, otsu's segmentation, Histogram matching	Neural Network	otsu's segmentation is efficient, combination color and texture feature provide efficient disease detection.	Histogram is not accurate.

Table 1.Leaf Disease Identification Techniques

The splitting of dataset into training and testing was done by the k-fold cross-validation or called rotation estimation method. The dataset was randomly split into mutually exclusive subsets (folds) of equal size of 10 [23].Comparing the seven classifiers, Extremely Randomized Trees yield a very high score.

Ratih Kartika Dewi and R. V. Hari Ginardi[9] recommended an image pattern classification techniques to identify the rust disease in sugarcane leaf. Features have been selected as color and Texture. Shape features include solidity, extent, minor axis length and eccentricity. Texture features are extracted from contrast, correlation, energy and homogeneity. It converts RGB to LAB because LAB color channel is consistent in terms of human perception.SVM classifiers are used for the classification of Training sample in support vector machine is separable by a hyperplane.

Eccentricity is used to determine whether the disease shape is a circle or line segment. It is the ratio of the distance between major axis length and the foci. An ellipse whose eccentricity is 0 can represent as a circle; however, an ellipse whose eccentricity is 1 can represent a line segment. Minor axis length is used to determine the length of the axis of the diseased region. Extent is the area of the diseased region that is divided by the area of the bounding box. The Extent can be computed as the area divided by the area of the bounding box.

Solidity is used to determine the area of the diseased region

Divided by the pixels in the convex hull. From Gray-Level Co-occurrence Matrix (GLCM) the texture features [10] can be extracted. Texture features are contrast, correlation, energy, and homogeneity. The Otsu-threshold is used to extract the shape feature. Mean, skewness and kurtosis are used to represent color as features. For this, it transforms RGB to LAB. SVM is the binary classifier because it can classify accurately [6]. There are different types of kernel function in SVM classifier. Table I shows different types of Leaf Disease Identification Techniques.

Table II shows different types of SVM. They are linear, quadratic, polynomial, and radial basis function. A polynomial kernel with order=3 has the highest accuracy with 98.5 % classification accuracy.

Table II.Different types of svm

linear	$K(x_i, x_j) = x_i^T x_j$
quadratic	$K(x_i, x_j) = ((x_i, x_j^1)+1))^2$
polynomial	$K(x_i, x_j) = ((x_i, x_j^1)+1))^3$
Radial basis function	$K(x_i, x_j) = \exp\left(-\frac{\|xi-xj^T\|^2}{2\sigma^2}\right)$

IV. CONCLUSION

This paper focused on image processing techniques that have been used for detection of plant diseases. Bare eye based plant disease detection is not efficient so image processing techniques are used. The major techniques for detection of plant diseases are; BPNN, NN, Histogram, SVM, K-means clustering, and SGDM, Fuzzy logic. These techniques are used to analyse the diseased plants leaves in the early stage. Some of the problem in this technique's viz. effect of background in the resulting image, segmentation of diseased spot in plant leaf and automation of the technique for continuous monitoring of plant leaf diseases. The review proposed that this disease detection technique shows a good result with an ability to detect plant leaf diseases.

REFERENCES

[1] P. R. Rothe and R. V. Kshirsagar," Cotton Leaf Disease Identification using Pattern Recognition Techniques", International Conference on Pervasive Computing (ICPC),2015.

[2] Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic" 2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.

[3] Smita Naikwadi, Niket Amoda," Advances In Image Processing For Detection Of Plant Diseases," International Journal of Application or Innovation in Engineering & Management (IJAIEM), Vol2, Issue 11, November 2013.

[4] S. S. Sannakki and V. S. Rajpurohit,"Classification of Pomegranate Diseases Based on Back Propagation Neural Network," International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015.

[5] Dr.K.Thangaduraiand K.Padmavathi," Computer Visionimage Enhancement For Plant Leaves Disease Detection," World Congress on Computing and Communication Technologies2014.

[6] Yuan Tian,Chunjiang Zhao, Shenglian Lu and Xinyu Guo," SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases," Proceedings of 2010 Conference on Dependable Computing (CDC'2010),November 20-22, 2010.

[7] Neetu Chahal and Anuradha," A Clustering Adaptive Neural Network Approach for Leaf Disease Identification," International Journal of Computer Applications (0975 – 8887)Vol 120 – No.11, June 2015.

[8] Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa," Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease ",Proceding of the 1'st interational conference on the use of mobile ICT in Africa ,2014.

[9] Ratih Kartika Dewi and R. V. Hari Ginardi," Feature Extraction for Identification of

Sugarcane Rust Disease", International Conference on Information, Communication Technology and System, 2014.

[10] Bernardes, J. G. Rogeri, N. Marranghello, and A. S Pereira,"Identification of foliar diseases in cotton crop," Topics in Medical Image Processing and Computational Vision, vol. 8, pp. pp 67–85, 2013.

[11] Ronse, "Set-theoretical algebraic approaches to connectivity in continuous or digital spaces," Journal of Mathematical Imaging and Vision, 1998.

[12] Z. Ma and K. Ata, "K-nearest-neighbours with a novel similarity measure for intrusion detection," UKCI'13, pp. 266–271, 2013.

[13] O. Maimon and L. Rokach, "Data mining and knowledge discovery handbook, second edition," April 2010.

[14] L. Ruey-Hsia and B. Geneva G, "Instability of decision tree classification algorithms," Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 570–575, 2002.

[15] B. S. Leo and L. Breiman, "Random forests," Machine Learning, pp. 5– 32, 2001.

[16] L. Andy and W. Matthew, "Classification and regression by randomforest," R News, vol. 2, no. 3, pp. 18–22, 2002.

[17] H. Zhang, "The optimality of naive bayes," Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2004), 2004.

[18] S. R. Gunn, "Support vector machines for classification and regression," University of Southampton, Technical Report, 1998.

[19] S. Keerthi, O. Chapelle, and D. DeCoste, "Building support vector machines with reduced classifier complexity," Journal of Machine Learning Research, vol. 7, pp. 1493–1515, 2006.

[20] wei Hsu, C. chung Chang, and C. jen Lin, "A practical guide to support vector classification," National Taiwan University, Taipei 106, Taiwan, 2010.

[21] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines: And other kernel-based learning methods," Cambridge University Press, 2000.

[22] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2, pp. 1137–1143, 1995. [Online]. Available:

[23] http://dl.acm.org/citation.cfm?id=1643031.1643047.

[24] Tian Y W, Wang B, Tang X M. Recognition of maize disease based on texture feature and support vector machine. Journal of Shenyang Agricultural University, 2005, 36(6): 730-732. (in Chinese)

[25] Roli, F. Andgiacinto,G. 2002. Design of multiple classifier systems. In Hybrid Methods in Pattern Recognition, H. Bunke, and A. Kandel, eds. World Scientific, River Edge, NJ.

[26] Dr. K.Sri Rama Krishna ,B.Vanajakshi "Classification Of Boundary And Region Shapes Using Hu-Moment Invariants ",B.Vanajakshi et al / Indian Journal of Computer Science and Engineering (IJCSE),Vol. 3 No. 2 Apr-May 2012.

[27] Miroslav Benco, Robert Hudec, Patrik Kamencay, Martina Zachariasova and Slavomir Matuska,"An Advanced Approach to Extraction of Colour Texture Features Based on GLCM",International Journal of Advanced Robotic Systems, 14 May 2014