Review on Plant Disease Identification Using Image Processing Techniques

Priya Soni
Student of Master Degree in Communication
Shri Sant Gadge Baba College of Engineering & Technology, Bhusawal
psoni070392@gmail.com

Abstract: This paper presents a survey on different studies done in the field of identification of plant diseases, necessity of identification of such diseases in early stage, review of researches done, highly recommended method of plant disease identification. This includes several steps viz. input images, image pre-processing, and extraction of features and classify them on different basis. Early information on crop health and disease detection can facilitate the control of diseases through proper management strategies. Plant leaf disease classifications have wide applications in various fields such as in biological research, in Agriculture etc. This paper provides an overview of different classification techniques used for plant leaf disease classification.

Keywords: Artificial neural networks (ANNs), Support Vector Machine (SVM), Discrete Wavelength Transform (DWT)

I. INTRODUCTION

Agriculture plays an important role in world economy and for India the role of agriculture in economy is much more as most of the Indian population depends on agriculture. So, if agriculture affected by any reason it will directly or indirectly affect our economy and the population which is dependent on agriculture. One of such reason which is adversely affecting the crop is diseases in plants or crop.

To avoid this, farmers require continuous monitoring of experts which might be prohibitively expensive and time consuming. Therefore, looking for fast, less expensive and accurate method to automatically detect the diseases from the symptoms that appear on the plant leaf is of great realistic significance. Author is leaving in a banana belt. Production of banana is more here so it is necessary to monitor the diseases in banana crop time to time so that production does not affect the productivity and quality of crop. But, instead of monitoring only single plant leave we have to move ahead in direction to identify diseases in almost all the plants.

II. METHODOLOGY

According to the research done in this field mainly chosen method of plant disease identification can be consists of mainly 4 steps –

1. Input Images
2. Pre – Processing of Images
3. Feature Extraction
4. Recognition & Classification

- **Input Images**

The very first step of this methodology starts with capturing images using cameras or scanners. The images is captured from the digital camera and then stored in the database. To start identification of plant diseases we have to first gather all the defected plant images so that we can work on it and find the exact solution to cure the diseases in plants. These all images store in the database of the system so that classification of this can be easier.

- **Pre - Processing of Images**

Image processing is the second and foremost step in this direction due to this extraction of images has been done so that the quality of image has been improved and the undesired distortion of the image has been suppressed and enhancement can be done on that part of the images where features need to be enhanced for further processing & analysis tasks.

It is used in considerable redundancy of the images. Neighboring pixels corresponding to one real object have the same or similar brightness value. If a distorted pixel can be picked out from the image, it can be restored as an average value of neighboring pixels .In the proposed approach image pre-processing methods are applied to the captured image which is stored in image database.

This includes – color space conversion, image enhancement, and image segmentation. The RGB images of leaves are converted into color space representation. The purpose of the color space is to facilitate the specification of colors in some standard accepted way. RGB images converted into Hue Saturation Value (HSV) color space representation. RGB is for color generation and his for color descriptor.

HSV model is an ideal tool for color perception. Hue is a color attribute that describes pure color as perceived by an observer. Saturation termed as relative purity or the amount of white light added to hue and value means amplitude of light.

Ycbcr color system is a common color space, which is applied by the most widely used jpeg image. Y, cb and cr, indicates a luminance component and two color component signals respectively. Different from other color space, Ycbcr color space is orthogonal, which fully takes important factors of composition of RGB from other colors into account. Ycbcr color space model is often used in image compression.

Image segmentation is process used to simplify the representation of an image into something that is more meaningful and easier to analyze. As the premise of feature extraction and pattern recognition, image segmentation is also

.

the fundamental approaches of digital image processing. There are various techniques for image segmentation includes – Region based, Edge based, Threshold based, feature based clustering and Model based etc.

- **Feature Extraction**

After segmentation the area of interest is extraction of the diseased on infected part. In this significant features are extracted and those features. Actually, image features usually includes color, shape and texture features. Currently most of the researchers targeting plant leaf texture as the most important feature in classifying plants. With the help of texture features, plant diseases are classified into different types. The various methods for feature extraction are as follows:

1) Texture Analysis
2) Texture Feature Extraction

1. Texture Analysis

Textures are a pattern of non-uniform spatial distribution of different image intensities, which focus generally on the individual pixels that make up an image. Texture is termed as quantifying the spatial relationship between materials in an image. Some properties are playing an important role in recitation of texture i.e. uniformity, regularity, density, linearity, directionality, roughness, coarseness, phase and frequency. Texture Analysis characterizes texture of an image into four major categories such as statistical, structural, fractals, and signal processing.

- Statistical: Statistical type includes grey-level histogram, grey-level co-occurrence matrix, auto-correlation features, and run length matrices for texture extraction.
- Structural: The structural models of texture presume that textures are combinations of texture primitives. Conceptually, structural texture analysis carried out into two major steps i.e. extraction of the texture elements, and inference of the placement rule. Two different structural methods are considered i.e. two dimensional wavelet transform and Gabor transform.
- Fractals: Many natural surfaces possess a statistical quality of roughness and self-similarity at different scales. Fractals have become very useful and popular in modeling these properties in the image processing field.
- Signal processing: Texture is especially suited for this type of analysis because of its properties. Includes spatial domain filters, Fourier domain, two-dimensional Gabor function.

2. Texture Feature Extraction

The extraction methods are used for extracting interesting and important features from the input images. Texture feature extraction method is the one which is used for the extraction of texture feature from images. The extraction techniques in texture field includes –

i) Color Co-occurrence Method

In statistical texture analysis, the texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. Grey Level Co-occurrence Matrices (GLCM) is a statistical method. It is an old and used feature extraction method for texture classification. It has been an important feature extraction method in the domain of texture classification that computes the relationship between pixel pairs in the image. The textural features can be calculated from the generated GLCMs, e.g. contrast, correlation, energy, entropy and homogeneity. However, in recent years, instead of using the GLCM alone, it is combined with other methods. Here are few other implementations of the GLCM, other than the conventional implementation e.g. one-dimensional GLCM, second-order statistical GLCM. It can be also applied on different color space for color co-occurrence matrix. Spatial Gray-level Dependence Matrices (SGDM) method is a way of extracting statistical texture features. Spatial Gray-level Dependence Matrices (SGDM) includes properties viz. contrast, energy, local homogeneity, and correlation.

ii) Gabor Filters

Gabor filters also popular as the Gabor wavelets. It is a widely used signal processing method. The Gabor filters consists of parameters such as the radial centre frequency, standard deviation and orientation. It can be can be used by defining a set of radial centre frequencies and orientations. As signal processing method produces large feature size, Gabor filters requires to be downsized for the prevention of the dimensionality issues. Principal Component Analysis (PCA) can be a good choice to downsize the feature space Though Gabor filters are popular in texture classification it sometimes combined with other methods too.

iii) Wavelets Transform

Another popularly used signal processing method in image processing and pattern recognition is wavelet transforms. At present, it became one of an important texture feature extraction method to be used in texture classification. Several wavelet transforms are used popularly nowadays such as Discrete Wavelet Transforms (DWT), Haar wavelet and Daubechies wavelets. Among these DWT is most widely used wavelet transform. The information on the frequency domain is usually more stable as compared to the spatial domain. Therefore, despite being more complex and slower, wavelet transforms usually produces better features with a higher accuracy.

iv) Principal Component Analysis

PCA is a statistical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of original variables is greater than or equal to the number of principal components. It is sensitive to the relative scaling of the original variables. PCA is the simplest of the true eigenvector-based multivariate analyses; PCA is closely related to factor analysis.

- **Recognition & Classification**

Now, it's time to recognize & classify the information gathered by applying image processing techniques often allows not only detecting the disease, but also estimating its severity, there are two main classifier are –

.

1. Linear Classifiers

A software routine was written in MATLAB. In which training and testing performed via several neural network classifier which are linear classifier. Texture Feature Classification Methods include Neural Networks – Neural Networks or Artificial Neural Networks (ANNs) are popular machine learning algorithms that are in a widely used nowadays. Multilayer Perception (MLP) is the basic form of ANN that updates the weights through back propagation during the training. There are other variations in neural networks, which are recently, became popular in texture classification.

Probabilistic Neural Network (PNN): It is derived from Radial Basis Function (RBF) network and it has parallel distributed processor that has a natural tendency for storing data. PNN is an implementation of a statistical algorithm called kernel discriminate analysis in which the operations are organized into a multilayered feed forward network having four layers i.e. input layer, pattern layer, summation layer, and output layer.

Convolutional Neural Network: It is a neural network that has convolution input layers acts as a self learning feature extractor directly from input images. Hence, it can perform both feature extraction and classification under the same architecture. Back propagation network: A typical BP network consists of three parts: input layer, hidden layer and output layer. Three parts in turn connect through the collection weight value between nodes. The largest characteristic of BP network is that network weight value reach expectations through the sum of error squares between the network output and the sample output, and then it continuously adjusted network structure's weight value. It is popular and extensively used for training feed forward networks. Also it has no inherent novelty detection, so it must be trained on known outcomes for training feed forward networks.

2. Non – Linear Classifier

Support vector machine (SVM) is a non-linear classifier, and is a newer trend in machine learning algorithm. SVM is popularly used in many pattern recognition problems including texture classification. Given a set of examples, this classifier marked each as belonging to one of two categories; an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

III. REFERENCES

Many Researches has been done in the field of "Identification of Plant Diseases".

A methodology for detecting plant diseases early and accurately using diverse image processing techniques has been proposed by Anand H.Kulkarni et al. [1], where Gabor filter has been used for feature extraction and ANN based classifier has been used for classification with recognition rate up to 91%. F. Argenti, et al. [2] proposed a fast algorithm for calculating parameters of co-occurrence matrix by supervised learning and maximum likelihood method for fast classification. Homogenize techniques like sobel and canny filter has been used to identify the edges by P.Revathi et al. [3]. These extracted edge features have been used in classification to identify the disease spots.

Brendon J. Woodford, Nikola K. Kasabov and C. Howard Wearing in paper titled "Fruit Image Analysis using Wavelets"[4] proposed wavelet based image processing technique and neural network to develop a method of on line identification of pest damage in pip fruit in orchards. Spatial gray-level dependence matrices (SGDM) method has been used for extracting statistical texture features by Sanjay B. Dhaygude et al. [5]. Diseases are impairment to the normal state of the plant that modifies or interrupts its vital functions such as photosynthesis, transpiration, pollination, fertilization, germination etc. These diseases are caused by pathogens viz., fungi, bacteria and viruses, and due to adverse environmental conditions.

Tian et al. developed a "Knowledge Based Machine Vision System for Outdoor Plant Identification" in a commercial agricultural environment in 1995 [6]. Images acquired in agricultural tomato fields under natural illumination were studied extensively and an environmentally adaptive segmentation algorithm, which could adapt to changes in natural light illumination, was developed. The method used four semantic shape features to distinguish tomato cotyledons from weed leaves and a whole plant syntactic algorithm was used to predict stem location of whole plant. Using these techniques, accuracies of 65% for detection of tomato plants were reported.

Yang et al. (1998) developed an artificial neural networks (ANNs) to distinguish between images of corn plants and seven different weeds species commonly found in experimental fields [7]. The performance of the neural networks was compared and the success rate for the identification of corn was observed to be as high as 80 to 100%, while the success rate for weed classification was as high as 60 to 80%.

To avoid larger damage in crops the early stage diagnosis of plant disease is an important task. Many researchers had already initiated the process in same direction considering the crop or plant available in their nearby place. As I am leaving in a banana belt I am considering Banana Plant as my main focus of this research along with the other plants

IV. RESEARCH GAP

According to my studies on all of the researches, mostly plant diseases identification is done on 1 or 2 types of plants. In my research, I am focusing on around 9 types of plants i.e. Banana, Cotton, Lemon, Rose, Papaya, Mango, Okra, Coconut, and Hibiscus. Along with that, to classify these images of plant leaves I am using three types of classifier i.e.

SVM (Support Vector Machine), KNN (Known Nearest Neighbor), and NN (Neural Network). For feature extraction DWT has been used

V. CONCLUSION

Early stage identification of plant disease is necessary to avoid damage in crop. This paper focused on identification of plant diseases which can be identified in initial stages so that pest control tools can be used to solve pest problems while minimizing risks to people and the environment. The paper mainly summarizes image processing techniques for several plant species that have been used for recognizing plant diseases. The major techniques for detection of plant diseases are: BPNN, SVM, and KNN. These techniques are used to analyses the healthy and diseased plants leaves. Some of the challenges in these techniques viz. effect of background data in the resulting image, optimization of the technique for a specific plant leaf diseases, and automation of the technique for continuous automated monitoring of plant leaf diseases under real world field conditions.

ACKNOWLEDGMENT

The successful completion on any task would not be complete without expression of the gratitude of all those who helped in doing that task. I hereby express my heartfelt gratitude towards the people whose help provided useful in completing my paper on “Review on Plant Diseases Identification Using Image Processing Techniques”. I would like to thank our honorable H.O.D for providing necessary facilities to carry out this seminar work and also my guide Prof. G.A.Kulkarni to give his valuable and sufficient time toward seminar and gave me proper guidance in the direction of completing this work. It would be untrue if I forget the kind and polite help of the staff member’s of Electronics & Communication department. Finally, I would like to thank all my friends and well wishers for having a very important role in nurturing and strengthening my career.

REFERENCES

[1] Anand H. Kulkarni, Ashwin Patil R. K., Applying image processing technique to detect plant diseases, International Journal of Modern Engineering Research, vol.2, Issue.5, pp: 3661-3664, 2012.

[2] F. Argenti,L. Alparone,G. Benelli ,” Fast algorithms for texture analysis using co-occurrence matrices” Radar and Signal Processing, IEEE Proceedings , vol. 137, Issue 6, pp:443-448 , No. 6, December 1990.

[3] P. Revathi, M. Hemalatha, Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques, IEEE International Conference on Emerging Trends in Science, Engineering and Technology, pp-169-173, Tiruchirappalli, Tamilnadu, India, 2012.

[4] Brendon J. Woodford , Nikola K. Kasabov and C.Howard Wearing[1999] Fruit Image Analysis using Wavelets, Proceedings of the ICONIP/ANZIIS/ANNES.

[5] Ananthi, S. Vishnu Varthini, and Detection & Classification of Plant Leaf Diseases International Journal of Research in Engineering & Applied Sciences, Volume 2, Issue 2 (February 2012) ISSN: 2249-3905.

[6] Tian, L. 1995. Knowledge Based Machine Vision System for Outdoor Plant Identification. Ph.D. Dissertation. University of California, Davis, CA.

[7] Yang, C.-C., S.O. Prasher, J.-A. Landry and A. DiTommaso. 2000a. Application of artificial neural networks in image recognition and classification of crop and weeds. Canadian Agricultural Engineering 42(3): 147-152.

[8] Hetzroni A, Miles GE, Engel BA, Hammer PA, Latin RX (1994) Machine vision monitoring of plant health. Adv Space Res. 14(11):203-212

[9] Zhangjian, Zhangwei, Inst. Of Intel. Machines, Chinese Acad. Of Sci., Hefei, China, Support Vector Machine For Recognition Of Cucumber Leaf Diseases, Advanced, and Issue Date: 27-29 March 2010 PN: 264 – 266.

[10] Pranjali Vinayak Keskar, Shubhangi Nimba Masare, Manjusha Suresh Kadam and Prof. Mrs. Seema U. Deoghare.2013 “Leaf Diseases Diagnosis & Detection” International Journal of Emerging Trends in Electrical and Electronics (IJETEE) Vol. 2, Issue. 2, April-2013.

About Authors:

First Author Priya Soni was born on 07.03.1992. In 2013 she completed her Graduation (B.E) with distinction at the department of Electronics & Communication of the Shri Sant Gadge Baba College of Engineering & Technology, Bhusawal. She is pursuing her Master degree in Communication. With learning on fly ability she achieved best results in whatever role she is assigned to. She passionate about playing Badminton and know how to enjoy her life well.

.