International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

Plant Leaf Disease Detection Recognition using Machine Learning

Shrutika Ingale
ME Student,
Dept. of Electronics and Telecommunication,
Sinhgad College of Engineering,
Vadgaon/Bk, Pune, India,

Prof. V. B. Baru
Professor,
Dept. of Electronics and Telecommunication,
Sinhgad College of Engineering,
Vadgaon/Bk, Pune, India,

***Abstract*— Indian economy generally depends on agricultural productivity. It is natural to have a disease on plants so that detection of disease plays an important role in agriculture. The main motto is to implement image analysis and classification techniques for detection of leaf diseases and classification. This method consists of four parts. They are Image preprocessing, Segmentation of the leaf, feature extraction and classification of diseases.**

Keywords*— Plant diseases, Image Segmentation, masking, disease detection, feature extraction

I. INTRODUCTION

In India, agriculture is known as backbone of the economy. According to the survey, 50% of the overall population is involved in farming directly or indirectly. The main goal of agriculture is not only to feed ever-increasing population but it is an important source of energy and it is a solution to solve the problem related to global warming [1].

The main diseases of plants are viral, fungus and bacterial disease like Alternaria, Anthracnose, bacterial spot, canker, etc. The viral disease is due to environmental changes, fungus disease is due to the presence of fungus in the leaf and bacterial disease is due to the presence of germs in leaf or plants. Segmentation procedure is done on different appearance found in a picture such as color orientation, texture, boundaries etc. A procedure of separating a picture into a different partition is called Image segmentation. In this paper, segmentation of leaves is done using a Gradient Boosting Algorithm. If the image processing is applied for the automatic disease detection then a low amount of effort will be needed, it will have low cost an on the other side it will be less time consuming and will be more accurate. This paper discusses the importance of image processing techniques in detection and identification of plant diseases in the earlier stages and hence the quality of the product could be increased.

II. LITERATURE SURVEY

There are various methods of image processing technique used for the segmentation process of plant leaf disease detection. But they are unable to differentiate different types of diseases on the various crop. Generally, color images are in three color red, green and blue. It gives an error when implementing the application using RGB model as their range is varies from 0 to 255. For this, they convert this RGB images into the grey images and for the result histogram is evaluated.

Monica Jhuria uses image processing for the detection of disease and the fruit grading in [3]. Artificial neural network for detection of disease is used by them. Firstly, they created two separate databases, one for the training which is of already stored images and others for the implementation of the query images. The adjustment of weight of training databases was done by back propagation. When results are calculated they identified that the morphological features give a better result than the other two features.

Zulkifli Bin Husin [4] captured the chili plant leaf images and processed it to determine the health status of the chili plants. Their technique is ensuring that the chemicals should apply to the diseased chili plant only. In this digital camera is used for image capturing and LABVIEW software tool is used to build the GUI (Graphic User Interface).

The RGB (Red,Green, Blue) image is converted into the HSV (Hue, Saturation, Value) translation. For the texture calculation, the SDGM matrix is generated and using GLCM function the features are calculated. It has two data compress and transmission method to meet user's different need and uses multi-channel wireless communication to lower the whole system cost. Shantanu Phadikar and Jaya Sil use pattern recognition technique for the identification of rice diseases. The DSP and FPGA based system is developed by Chunxia Zhang, Xiuquing Wang for monitoring and control of plant diseases. The FPGA is used to get the field plant image or video data for monitoring and control plant diseases.

III. BASIC STEPS FOR DISEASE DETECTION

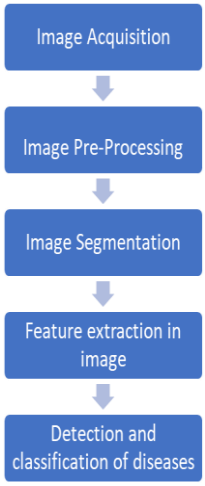

Fig. 1) Steps for plant disease detection and classification

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

A. Image Acquisition:

The images of plant leaf first captured by the camera. The captured image is in RGB form that is in Red, Green, Blue color. This is the first step. This step is then taken as an input.

B. Image Pre-processing:

The main moto of this step is an improvement of the captured image data that destroy the unwanted distortions which results in the enhancement of the features.

C. Image Segmentation:

Segmentation means partitioning of one image into various parts. The parts having same features or having somewhat similarity. The segmentation can be done in various methods like k-means clustering, converting RGB into HIS model, Otsu method etc.

1] Segmentation using Boundary and spot detection algorithm: The RGB image is converted into the HIS model for segmentation. This method helps to find the infected part of the leaf. For boundary detection, the 8 connectivity of pixel is checked and the boundary algorithm is applied.

2] Masking of green pixels: Masking means the pixel value in an image set to zero or some other value. Since the green region represents the healthy region of the leaf. When the intensity of a green pixel is greater than the previous one then all those values set to zero. After masking, pixels with zero values are discarded. In masking the disease portion of the leaf is identified by H and S plane value and value of '1' which is allocated to a particular portion. Rest of the regions '0' value is given. As a result, a binary image contains only zeros and ones. Thus, the diseased area of the leaf can be extracted.

3] Threshold based Segmentation: Similarly based segmentation is adapted to segment input images on the basis of the similarity in intensity or grey levels in an image. This method was usually used for light objects in a dark background or dark objects in a light background. The threshold-based algorithm will choose a proper threshold value T to divide image's pixels into several classes and separate objects from the background. Binary image contains only values of ones and zeros. Then this binary image is multiplied with the original RGB color image. In this way, the infected portion get extracted.

D. Feature Extraction:

The three fundamental patterns used in human interpretation of images are spectral, textural and contextual features. In this work, shape, color and texture features are extracted from every leaf image of the dataset.

1) Gray-level co-occurrence matrix (GLCM):

It is the statistical method of identifying texture which considers some spatial relation with the pixel. The statistical features are then extracted from the matrix. The features like color, texture, morphology, homogeneity, contrast, correlation, energy are extracted from a given image. This extracted feature then give some value to respected image. It can also be used for detection of infected plant areas.

2) Shape Features:

There are many features based on the geometry of the leaves. The length of the leaf is found by taking the Euclidean Distance between the two points on either side of the long axis whereas breadth is corresponds to the length of the minor axis as shown in fig.2

$$\text{Aspect Ratio} = \frac{\text{Length of the leaf}}{\text{Breadth of the leaf}}$$

By using length and breadth of leaf aspect ratio is found.

- The area is calculated by initially finding the area of one pixel.

$$\text{Area} = \text{Area of one pixel} * \text{Total no. of pixels present in the leaf}$$

- The perimeter of the leaf is given by the count of pixels having the leaf margin.
- Rectangularity shows the similarity between a leaf and a rectangle.

$$\text{Rectangularity} = \frac{L*W}{A}$$

Where L is the length, W is the width and A is the area of leaf.

3) Venation:

The leaves can be better identified by its veins. These veins are unique for each plant species. These can be extracted by applying morphological operations. So, the information occupied by the background part is subtracted and only the vein pattern appears.

E. Classification:

1) Using ANN:

The database is classified in accordance with the neural network. These feature vectors are considered as neuron in ANN while classification is done. The output obtained is a function of the weighted sum of the inputs given. The various types of classification methods are used such as SVM, Back Propagation algorithm can be used.

- **Advantage:** Good potential with ability to detect plant leaf disease.
- **Disadvantage:** Require more time

2) K-means clustering:

Partitioning is done by k means clustering. Basically K- means clustering is used for classification of the object based on the set of the features into K number of classes. This classification is minimized by the squares of the distances between the object and the corresponding clusters.

- **Advantage:** It is easy to implement and gives quick result.
- **Disadvantage:** Slow learner and fully not robust to the noise data in large training.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

3) *Otsu Threshold Algorithm:*

Binary images from grey level image are created using thresholding by setting all pixels below some threshold to zero and all the pixels above threshold to one.

- **Advantages:**
  - Regardless of uniformity and shape measures
  - It works on real world images
- **Disadvantage:** takes more time

4) *Support Vector Machine (SVM)):*

Support Vector Machine is a kernel based supervised learning algorithm used as classification tool. In this, the training algorithm maximize the margin between training data and class boundary. The result depends on training data is called as support vectors as shown in Fig 2. It is effective in high dimensional space where a number of dimension are greater than the number of training data.

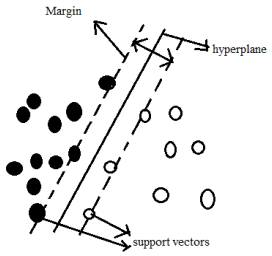

Fig. 4) Support Vector Machine Classifier

- **Advantage:** Prediction accuracy is high and having robust working when training example have error in them.
- **Disadvantages:**
  - Require long training time
  - Difficult to understand learned function
  - Large nos. of support vectors are used for training in classification task

5) *Back Propogation:*

BPNN algorithm is used in a recurrent network. Once trained, the neural weights then fixed and can be used to calculate output values for the new query images which are not present in the learning database.

- **Advantages**:
  - It is fast, simple and easy to program.
  - It has no parameters to tune (except for the number of input).
  - It is adaptable to any situation
  - It is powerful and it can adopt complex functions as well.

- **Disadvantages:**
  - Not exact
  - Large complexity of network structure

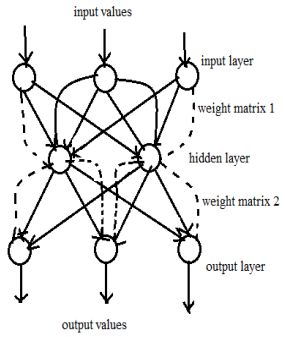

Fig. 5) Back Propagation Network

IV. CONCLUSION

This paper represents a review of the technical implementation in the research area of plant disease detection using image processing technique. From the literature, it is proof that color, texture and morphological features are most suitable to identify and classify the diseases in plants. ANN and SVM are most commonly used classification techniques used for classification of diseases in plant leaf. This detection of diseases in plant nowadays helps the farmer to improve the crop quality which helps in improvement of Indian gross domestic product (GDP).

	Technique name	Merit	Demerit
1	Otsu's method	Regardless of uniformity & shape measures, it works on real world images.	Takes more time for image processing.
2	K-means	Minimize sum of square distance between object and centroid.	Difficult to predict value of k with fixed number of clusters.
3	Grey-level Thresholding	It provides contrast for disease region and background	Every time need to select proper threshold value for getting better result in segmentation.
4	Fuzzy c - means	Use partial membership, therefore more useful for real problems.	Sensitive to initialization condition of cluster no. and cluster center.

V. REFERENCES

[1] Wenjiang Haung, Quigsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Linsheng Huang, and Dongyan Zhang, "New Optimized Spectral Indices for identifying and Monitoring Winter Wheat Diseases", IEEE Journal of selected topics in applied earth observation and remote sensing, vol. 7, No. 6, June 2014.

[2] R. Meena Prakash, G. P. Sarasathy, G. Ramalakshmi, K. H. Mangleshwari, T. Kaviya, "Detection of leaf Diseases and Classification using Digital Image Processing", 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS).

[3] Monica Jhuria, Ashwani Kumar, and Rushikesh Borse, "Image Processing and Fruit Grading", Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).

[4] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook, "Feasibility study on plants chili diseases using image processing Techniques", 2012

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

Third International Conference on Intelligent Systems Modelling and Simulation.

[5] Saradhambal G1, Dhivya.R2, Latha.S3, R. Rajesh4, "Plant Diseases Detection and its Solution using Image Classification", International Journal of Pure and Applied Mathematics Volume 119 No. 14 2018, 879-884.

[6] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. A L Rahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Application (0975-8887) Volume 17-No. 1, March 2011.

[7] Chunxia Zhang, Xiuquing Wang, Xudong Li, "Design of Monitoring and control Plant Disease System Based on DSP & FPGA", 2010 Second International Conference on Network Security.

[8] D Venkatraman, Mangayarkarasi N, "Computer Vision Based Feature Extraction of leaves for Identification of Medicinal Values of Plants", 2016 IEEE International Conference on Computational Intelligence and Computing Research.

[9] Jitesh P. Shah, Harshadkumar B. Prajapati, Vipul K. Dabhi, "A Survey on Detection and Classification of Rice Plant Diseases", 2016 IEEE.