IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DETECTION OF PLANT DISEASE USING IMAGE PROCESSING TECHNIQUES

[1]Sona Vijayan, [2]Prameeja Vimal

[1]Msc Scholar, [2]Assistant Professor

[1,2]Department of Computer Science,

[1,2]St. Joseph's College (Autonomous), Irinjalakuda, Thrissur, India

Abstract: Diseases in plants can reduce the production and quality of food, fibre and biofuel crops. Plant disease detection system has been a necessity for cultivators or farmers as the detection of a correct disease can lead to a proper yield and effective growth of the plant. Here the proposed plant disease detection methodology aims at identifying tomato plant diseases from images of affected plant leaves. It is very difficult to detect the correct plant disease manually as it take lot of time and energy. This proposed technique will help farmers or any cultivators to detect plant diseases digitally. In the proposed project, various image processing steps which are performed on the leaf images. We use the classification methods such as K-Nearest Neighbor Algorithm and Naïve Bayes Algorithm to distinguish the different plant diseases.

Key words- ***Image acquisition, Image segmentation, Image post-processing, Feature extraction and Classification.***

I. INTRODUCTION

Agriculture is the backbone of our country. According to survey nearly 70% of India's population is depends on agriculture. The increase in human population requires corresponding increase in the food production. The quantity of food consumed by the United Kingdom is almost equal to the quantity of food wasted in India. As reported by the United Nations up to 40% of food produced in India is wasted. Food wastage and shortage has become a major issue now a days.

The primary types of Food losses are,

1. Production stage: Losses due to the disease, spillage and mechanical error.
2. Handling and storage: Losses due to the storage and transportation during farms and distribution networks.
3. Processing: These are the losses during the domestic and industrial processing like juice production, canning, sorting for unsuitability.
4. Consumption: These are the wastes produced by consuming food at a domestic level.

This project focuses on a potential solution to reduce the waste caused during the production stage due to the plant diseases. It is very difficult to find the plant diseases manually. One of the existing methods, a simple observation is done by the experienced farmers or experts to detect a plant disease. For large plantation, a team of experts is required to examine the plant. This procedure can consume large amount of time and cost. Here the proposed system can predict the plant disease using image processing and other techniques which makes the disease detection process much faster and easier. It also assists new farmers in the disease detection process.

In the production of food, controlling plant diseases is very necessary since it involves the use of various resources like land and water for cultivation of the crops, fuel and other energy resources used for machinery and equipment. All the plants have their own intrinsic disease resistance, yet there are various examples of plant diseases which have caused devastating impacts for instance the Great Famine of Ireland, soybean cyst nematode, ice blast and citrus canker. Detecting plant disease may be a key to stop agricultural losses.

II. EXISTING SYSTEM

Plant leaf disease is a common problem in most of the plantations and farms. This system can help cultivators and farmers in detecting the plant disease with the help of an image of the affected leaf. Health monitoring and identification of disease is very difficult in the traditional methods. The traditional methods are slow, time-consuming and might also lead to incorrect detection. So it is very essential to implement a modern method that would help in such problem with accuracy and speed.

All the plant leaves are differ from each other based on their different physical features such as shape, color etc. We cannot predict some of the plant diseases by looking at the external features and guessing the disease. The laboratory predications of some diseases might be 100% accurate but it is a lengthy process. The plants which are very sensitive require much faster detection and faster measures to cure them. So farmers cannot opt for laboratory detection technique since it takes more time. It may also lead to wrong recognition, so they need an efficient and fool proof system to aid in their work. The aim of this system is to help farmers in detecting plant diseases without any delay and inaccuracies.

III. PROPOSED SYTEM

The objective of this system is to develop a plant disease detection system using the domains of image processing. This focuses on detecting plant diseases up to a certain extent. For implementing the system to detect plant diseases, we extract the data from the plant leaf image using image processing. Then the image is classified as either a healthy leaf or an infectious leaf by using a classifier.

Architecture of the proposed system

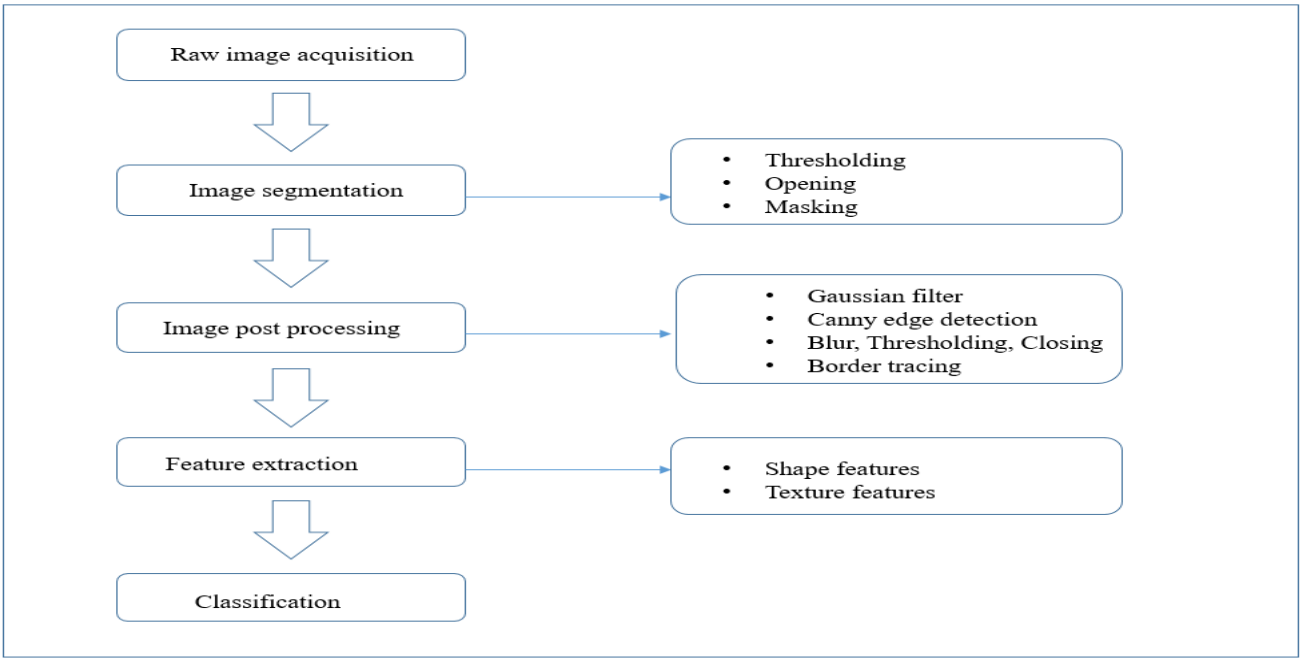

Fig.1 Block diagram

i. IMAGE ACQUISITION

The first step in disease recognition is image acquisition. In image acquisition, acquiring an image of the target leaf using a digital camera. This image is stored for further processing. All the possible optical details of the target leaf are captured. In this project all the images used are of tomato leaves. The dataset is categorized into 6 categories. The primary dataset consists of healthy tomato leaves. The rest of datasets consist of leaf images with diseases namely bacterial spots, late blight, leaf blight, leaf curl, mosaic virus and septoria leaf spot. This categorization will be used as an input for training and testing the classifiers in the classification phase later.

Fig.2 Image acquired

ii. IMAGE SEGMENTATION

Image segmentation means fragmenting the digital image into multiple parts. We use image segmentation to separate background of the leaf from the foreground on which the further analysis is done. To extract the numerical features, we require the leaf part which is shades of green and yellow colors in majority and the shades of black and brown for the diseases. The rest of the pixel data is dumped to avoid the outliers. To accomplish this, we use color image thresholding and other procedures. The RGB image is transformed into HSV to perform thresholding. This is done due to the following reasons:

- RGB is the composite color model.
- HSV is very good for visualization.
- The algorithm will make is less prone to lighting variations when using the Hue component only.
- Inaccurate and wrong results may be provided when the channel-wise arithmetical operation is performed in the RCB space.

Masking is the next step. Here a mask is created of the original RCB image using the HSV image obtained in the previous step. Then in order to obtain the segmented image, a bitwise AND operation is performed between the actual image and the masked image.

Fig.3 After image segmentation

iii. IMAGE POST PROCESSING

Among other things pre-processing include filtering techniques. This application need the use of leaf edges for the functioning which may be softened on such smoothing. For this, first segment the image to detect the outer edge accurately. Then use the post-processing to smoothen the image which removes various noises.

1) Gaussian filter

Gaussian filter is a linear filter. Gaussian filtering operator is a 2-D convolution operator and is used to remove off the details and noise by blurring off images much like mean filtering but with a different kernel which has the shape of a Gaussian hump.

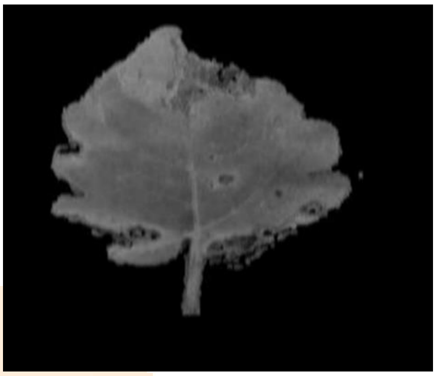

Fig.4 after Gaussian filter

2) Canny edge detection

The canny edge detection algorithm is a multi-stage algorithm which detects the edges in the images. Based on the Gaussian derivative it uses a filter to compute the intensity of the gradients. The Gaussian can reduce the effect of noise present in an image. Then it uses Non-Maximal Suppression which thins down the potential edges to near 1-pixel curves. After this Hysteresis Thresholding is done which decides whether the edge pixels are kept or removed. The output of the canny can be adjusted using the three parameters: width or the kernel of the Gaussian, the lower threshold and the higher threshold of the hysteresis.

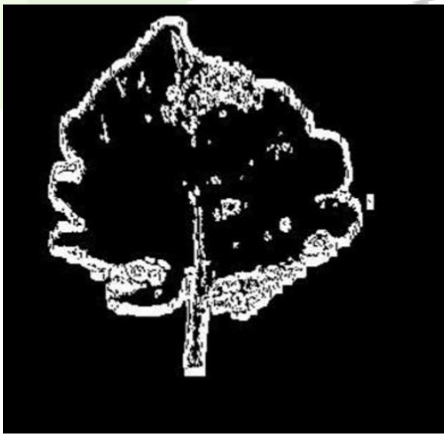

Fig.5 after Canny edge detection

3) Blur-Threshold-Close

It involves three steps. In the first step blurring the image with a Gaussian filter. The Gaussian filter spreading the intensity of the pixel blobs in its neighboring pixels. In the second step performing thresholding on these pixel spots and removing the outer pixel belt using a certain threshold value. The third step is a closing operation which removes the reduced pixel spots in the image. This process is done several times which eliminates the pixel spots in the background of the leaf.

Fig.6 After the blur-threshold-close

4) Border Tracing

Canny edge detection leads to a result which is an image with a thin multi-pixel border. This method of Border Tracing yields a border which have 1-pixel thickness. It starts with an initial border point of the object in the binary image and it is followed by an iterative search in its 8-neighborhood periphery. To check for a border pixel it hypothetically draws a line segment in all of its neighborhood. So the algorithm is called Radical Sweep.

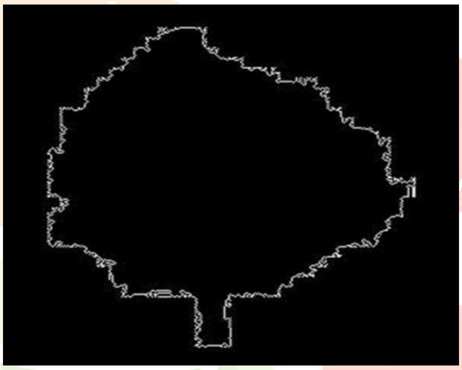

Fig.7 after Border Tracing

iv. FEATURE EXTRACTION

A feature in an image can be parameters such as a piece of spatial information, properties such as shape, curves or edges, borders, they can also be features such as blobs and ridges. In a successful extraction of a feature, the result; called the feature descriptor or feature vector is then stored for further processing. Then this complete set of features will be given to the classifier for training and testing.

The features are extracted from the leaf in two ways. They are shape features and texture features.

1) Shape Features:

- Length – It refers to the length of the two terminals of the central vein.
- Width – The physiological is the highest distance between the points of the orthogonal intersection of the length and the border.
- Area – It is the number of set pixels on the segmented image.
- Leaf Perimeter – It refers to the number of border pixels in the leaf margin.
- Aspect Ratio – It refers to the ratio of physiological length to the physiological width of a leaf.
- Form Factor – It is used to specify the difference between a leaf and a circle. Form Factor is defined as $4\Pi A/P^2$, where P is the perimeter of the leaf margin and A is the leaf area.
- Rectangularity – It refers to the likeness between a rectangle and a leaf. It is defined as (L*W)/A, where L is the length, W is the width and A is the leaf area.

2) Texture Features:

- Angular Second Moment (ASM) – It measures orderliness, how regular or orderly the pixel values are in the window.
- Energy – It obtains the sum of squares of every element in the GLCM.
- Entropy – Entropy measures uncertainty of the image (variations).
- Contrast – It calculates the local variations of the gray-level co-occurrence matrix.
- Homogeneity – Homogeneity refers to the adjacency of the elemental distribution in the GLCM and GLCM diagonal.

Table 1 after Feature extraction

Feature	Value
Area	26376
Form factor	0.19402943676497836
Entropy mean	1.6719866350768626
Entropy range	4.204318021092474
Homogeneity mean	0.682312377753786
Homogeneity range	0.08485436691438598
Perimeter	1307.0
Aspect ratio	0.9004080511633751
Contrast mean	267.73086219520883
Contrast range	1268.4025336131135
Sum of squares mean	3202.456126228885
Sum of square range	9838.438988702219
Rectangularity	1.9748485653549155
Angular second moment mean	0.4357475156795096
Angular second moment range	6.465582331463026E-4

v. CLASSIFICATION

The main goal of classification is to predict the target class for each case in the data accurately. In the proposed system to classification of the diseases is done based on the numerical dataset derived from the feature extraction. Here we are using the classification algorithms like K-Nearest Neighbor and Naïve Bayes algorithms.

1) K-Nearest Neighbor Algorithm

KNN algorithm is used to solve classification and regression problems. Here we use KNN for the classification purpose. KNN is a supervised machine learning algorithm which relies on input data whose characteristics or parameters are known as labelled data and gives as output a result based on its learning patterns. In the proposed system, KNN takes as input the image of the leaf and all its characteristics and features that are extracted in the previous phases and it creates a data model. Then this data model is given as an input a test data that is leaf image whose characteristics are not known and the algorithm predicts the characteristics, in this case the disease of the leaf.

2) Naïve Bayes Algorithm

It is used to differentiate between the different data objects based on some characteristics. This follows a probabilistic machine learning model which is used for learning task. The core of the classifier is based on the Bayes theorem.

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(A)}$$

The Naïve Bayes classifier presumes that a feature value that is independent of other feature values of a class variable. Simply, Naïve Bayes model follow conditional probability for the classification.

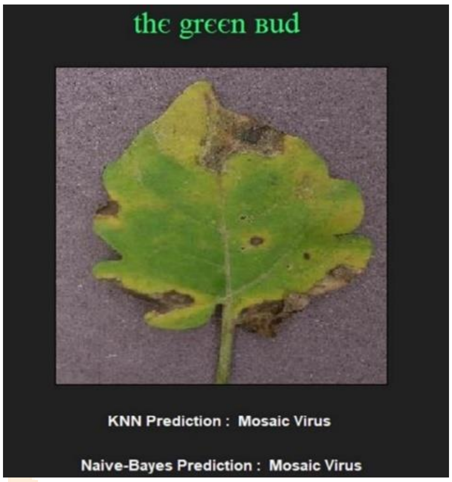

Fig.9 after the K-Nearest Neighbor Algorithm and Naïve Bayes Algorithm

IV. CONCLUSION

Identifying diseases in plant leave is challenging for farmers and cultivators. The major motive of this project was to recognize diseases which are still difficult for farmers to detect without any assistance. During this project, we researched about the diseases of tomato leaves and its characteristics, various image processing techniques and classification algorithms. We used the Naïve Bayes Algorithm and then the K-Nearest Neighbor Algorithm for the classification purpose. The Naïve Bayes classifier has an accuracy of 65% and the K-Nearest Neighbor has an average accuracy of 80%. So KNN was found to be good as it has higher accuracy. This paper provides an efficient and accurate approach for the detection of plant diseases.

REFERENCES

[1] Shubham Malvankar, Bilwa Prabhu Moye, Riya Shirodkar, Vipul Soman, Nora Naik, Saurabh Naik, “Plant Disease Detection using Image Processing”, IJRASET, ISSN: 2321-9653, Volume: 8, Issue: V, May 2020.

[2] Sharath D M, Akhilesh, S Arun Kumar, Rohan M G, Prathap C, “Image Based Plant Disease Detection In Pomegranate Plant For Bacterial Blight”, International Conference On Communication and Signal Processing, April 4-6, 2019.

[3] Saradhambal.G, Dhivya.R, Latha.S, R.Rajesh, “PLANT DISEASE DETECTION AND IT’S SOLUTION USING IMAGE CLASSIFICATION”, International Journal of Pure and Applied Mathematics, Volume 119 No.14 2018, 879-884, ISSN:1314-3395(on-line version).

[4] Ashwini C, Anusha B, Divyashree B R, Impana V, Nisarga S P, “Plant Disease Detection using Image Processing”, IJERT, 2278-0181, 07-08-2020.

[5] V Suresh, D Gopinath, M Hemavarthini, K Jayanthan, Mohana Krishnan, “Plant Disease Detection using Image Processing”, International Journal of Engineering Research & Technology(IJERT), ISSN: 2278-0181, Vol.9 Issue 03, March-2020.

[6] Karthikeyan N, Anjana M S, Anusha S, Divya R M and Asha Vinod, “Leaf Disease Detection using Image Processing”,IJISET, ISSN(online) 2348-7968, Vol. 7 Issue 5, May 2020.

[7] Abhang Kalyani, Tayade Payal, Tilekar Nilesh, Wadekar Satyam, Prof.Borkar B S, “DETECTION OF LEAF DISEASE AND CLASSIFICATION USING IMAGE PROCESSING”, ISSN(online) 2456-3293, Volume 5, Issue 5, May 2020.

[8] Simranjeet Kaur, Geetanjali Babbar, Navneet Sandhu, Dr. Gagan Jindal, “VARIOUS PLANT DISEASES DETECTION USING IMAGE PROCESSING METHODS”, IJSDR, ISSN: 2455-2631, Volume 4, Issue 6, June 2019.