ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 6, June 2017

# Fruits and Leafs Disease Detection using Naive Bayes Algorithm

[1] Pramod D Huddar, [2] S Sujatha
[1] DEC. M-tech. Bangalore Institute of Technology Bangalore
[2] Dept. of ECE. Bangalore Institute of Technology Bangalore

***Abstract*: India is an agricultural country wherein most of the population depends on agriculture. Research in agriculture is aimed towards increase of productivity and food quality at reduced expenditure, with increased profit. Here in this paper using naive bayes algorithm we will be storing the data and with help of probability we will be checking for different leaves and fruits weather they are effected with disease or not, if yes characteristics of the disease will also be know. This project is implemented in matlab with help of image processing technique.**

***Keywords*- naïvebayesalgorithm, glcm, pre-processing.**

## I. INTRODUCTION

Agricultural production system is an outcome of a complex interaction of soil, seed, and agro chemicals. Vegetables and fruits are the most important agricultural products. In order to obtain more valuable products, a product quality control is basically mandatory.

Many studies show that quality of agricultural products may be reduced due to plant diseases. 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. Therefore, the early stage diagnosis of plant disease is an important task. 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. This enables machine vision that is to provide image based automatic inspection, process control and robot guidance. The objective of this paper is to concentrate on the plant leaf disease detection based on the texture of the leaf. Leaf presents several advantages over flowers and fruits at all seasons worldwide.

## II. EXISTING METHODOLOGY

There are five main steps used for the detection of plant leaf diseases as shown in fig.1. The processing scheme consists of image acquisition through digital camera or web, image pre-processing includes image enhancement and image segmentation where the affected and useful area are segmented, feature extraction and classification. Finally the presence of diseases on the plant leaf will be identified.

In the initial step, RGB images of leaf samples were picked up. The step-by-step procedure as shown below:

1) RGB image acquisition;
2) convert the input image into color space;
3) Segment the components;
4) obtain the useful segments;
5) Computing the texture features;
6) Configuring the Classification for recognition.

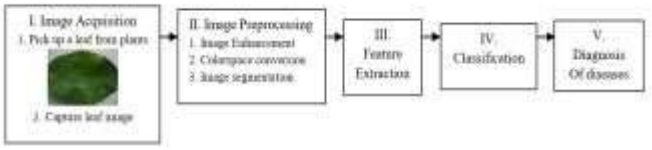


***Fig1: Block diagram of steps involved in plant leaf disease detection.***

Here the Main Step is Classification, the Classification is done using SVM. 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. SVM is designed to work with only two classes. This is done by maximizing the margin from the hyper plane. The samples closest to the margin that were selected to determine the hyper plane is known as support vectors. Multiclass classification is applicable and basically built up by various two class SVMs to solve the problem, either by using one-versus-all or one.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 6, June 2017

## III. PRAPOSED METHODOLOGY

In proposed methodology we will be mainly concentrating on three types of fruits and their leaf diseases ie Banana, Mango, and Grapes. The reason being that these three fruits are found and grown in almost all part of India, the research done on these three fruits will be useful in almost all extension of agriculture and to the wider advantage.

This project involves image processing technique using naive bayes algorithm for probability purpose and determining the amount of effected area in terms of percentage.it also detects the disease and suggests the measuring techniques through scientific facts.

Proposed methodology involves following steps:

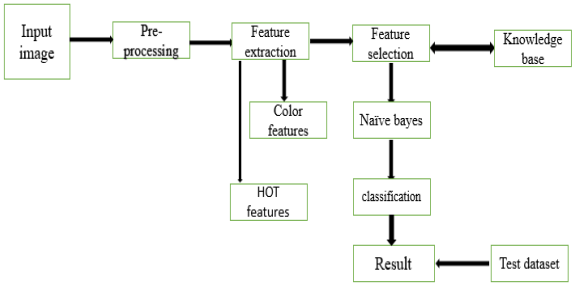


***Fig2:Block diagram of proposed methodology***

The input image will be the sample image of the fruit or leaf which is found to be affected by any disease.

### *A. PRE-PROCESSING:*

The acquired images are dissimilar in dimensions, before future extraction is done the image is needed to be brought into to a certain format with regards to size and shape. The acquired image is re-sized to 256X256. The RGB image is converted to gray and HSI color space. Gray scale images are simple to process, as they contain only intensity value.


***Fig3:Pre-processed image expressed in the form of clusters***

After the pre-processing is done image is represented in the form of clusters based on the histogram values that we obtain in the pre-processing stage, there are three clusters named as 1,2 and 3 region of interest asks for the suitable cluster value that how the input image look like by giving the value further future extraction will be carried out.

### *B. FEATURE EXTRACTION:*

It involves color extraction where many of the characteristics of the image will be extracted using GLCM matrix. GLCM matrix is a statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co- occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix.

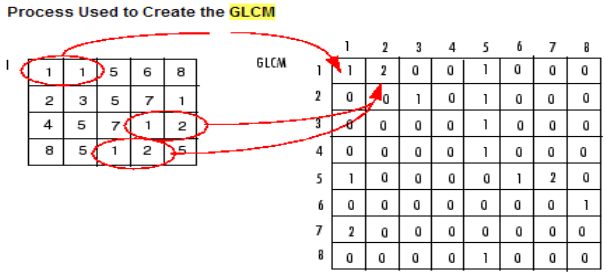


***Fig4: Creation of GLCM***

***Using GLCM we extract the following properties***

### *1. CONTRAST:*

Return a measure of the intensity contrast between a pixel and its neighbor over the whole images. Range= [0 {size(GLCM,1)-1}^2]; Contrast is 0 for constant images.

### *2. CORRELATION:*

Correlation measures the linear dependency of grey levels of neighboring pixels.Return a measure of how correlated a pixel is to its neighbor over the whole image Range=[-1 1]; Correlation is 1 or -1 for a perfectly positively or negatively correlated image. Correlation is null for a constant image.

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)\,p(i,j)}{\sigma_i \sigma_j}$$

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 6, June 2017

### *3. ENERGY:*

Returns the sum of squared elements in the GLCM. Range=[0 1] energy is 1 for constant image.

$$\sum_{i,j} p(i,j)^2$$

### *4. HOMOGENEITY:*

Return a value that measure the closeness of the distribution of element in the GLCM to the diagonal. Range=[0 1] Homogeneity is 1 for diagonal GLCM.

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

Along with four properties explained above other eight properties will also be extracted and there values will be stored in the database.

### *C. FEATURE SELECTION:*

Feature selection involves naïve bayes theorem for prediction of probability that contains in the probability with reference to previous values of the properties maintain in the database. The input values of the image will be compared with the database values, with the help of kernel smoothing function estimate case density with respect to input image values and database values are obtained multiplying these values in response to their rows we obtained the name of the disease.

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

Where A and B are event and P(B)≠0, P(A) and P(B) are the probabbilities of observing A and B without regard to each other. P(A|B) a conditional probability, is the probability of observing event A given that B is trsue. P(B|A) is the probability of observing event B given that A is true. A naive bayes classifier is a simple probabilistic classifier that is based on applying bayes theorem with strong (naive) independence assumptions. A more descriptive term for the probability model would be independent feature mode.

## IV.EXPERIMENTALRESULT

From experiment carried out we will be obtaining the disease part, amount of affected area, measures to cure the affected part. From the naïve bayes algorithm which is much more effective than the artificial neural network we obtain the result such that comparatively more easy with respect to image compression techniques.

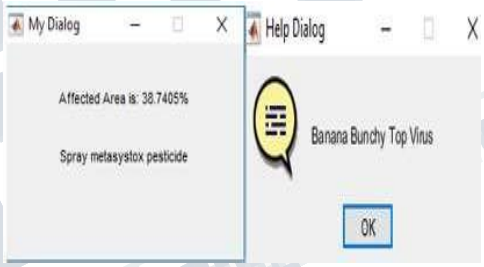


Fig5:Output result of examined leaf.

### *AFFECTED AREA CALCULATION:*

To measure the severity of the disease, grading is assigned based on the affected pixel area. Depending on the percentage of the diseased area, the grades are assigned. This is calculated using the equation.

Total % affected area= (Number of affected leaf pixels)/(total number of leaf pixels)

In the future extraction block from the cluster diagram we will select the cluster which is almost equal to input image so that the cluster images reads the type of disease or virus from the database and suggest the suitable measure for curing of disease.

## V CONCLUSION

In this project we designed a novel methodology to detect the disease of the leaf and fruits

it also specifies the prevention measures for the leaf, the main thing that differs it from the other methodology is the NAVY BAYES algorithm that is efficient from the other algorithm and its found that this methodology works better than other methodologies.

## VI .ACKNOLEDEMENT

I would like to thank Mrs. S SUJATHA, Associate Professor, Dept. of ECE, Bangalore Institute of Technology (BIT) for support in the completion of the project.

## 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] 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.

[3] A.Meunkaewjinda, P. Kumsawat, K. Attakitmongcol and A. Srikaew, "Grape leaf disease detection from color imagery system using hybrid Intelligent system", proceedings of ECTICON, IEEE, PP-513-516, 2008.

[4] 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

[5] Prakash M. Mainkar, Shreekant Ghorpade, Mayur Adawadkar, "Plant Leaf Disease Detection and Classification Using Image Processing Techniques", International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 4, ISSN: 2394- 3343, 2015.

[6] P. Revathi M. Hemalatha, "Homogenous Segmentation based Edge Detection Techniques for Proficient Identification of the Cotton Leaf Spot Diseases", a, International Journal of Computer Applications (0975 – 888) Volume 47– No.2, June 2012.

[7] Tushar H Jaware, Ravindra D Badgujar and Prashant G Patil. "Crop disease detection using image segmentation". World Journal of Science and Technology [2(4):190-194], 2012.

[8]. Dickson, M.A., and Bausch, W.C. "Plant recognition using a neural network classifier with size and shape descriptors", Transactions of the ASAE, vol. 1, pp. [97-102] 1997.

[9] Piyush Chaudhary, A. "Color Transform Based Approach for Disease Spot Detection on Plant Leaf". nand K. Chaudhari, Dr. A. N. Cheeran and Sharda Godara. International Journal of Computer Science and Telecommunications [Volume 3, Issue 6], June 2012.

[10] Sanjay B. Patil, Dr. Shrikant K. Bodhe."Leaf disease severity measurement using image processing". [IEIT Journal of Adaptive & Dynamic Computing, [(4), 25–30], 2012.