International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 4
ISSN: 2454-4248
603 – 606

Predictive Eyes for Disease Detection of Plants using Image Processing and Data Mining

Prof. Kulkarni Geetanjali Ramchandra
Computer Science & Engineering Dept.
Brahmdevdada Mane Institute Of Technology
Solapur.India
Geetanjalikulkarni18@gmail.com

Prof. Pawar Suchitra Chandrakant
Computer Science & Engineering Dept.
Brahmdevdada Mane Institute Of Technology
Solapur.India
suchitrapawar.13@gmail.com

Abstract: In Agriculture field, there are many problems which are faced by the farmers. To solve those types of problems farmers needs expert advice.By developing an Expert system in Mother tongue of farmer and also in English,helps her/him to solve those problems and for increasing crop production .the objective of these paper is to develop expert system in Agriculture field using CANNY's edge detection technique and RGB histogram. We are using C5.0 algorithm for classification and pattern matching technique is used to match images. Depending on the match images expert system gives best advice to the farmer**.**

Keywords : *CANNY Edge Detection,C5.0, RGB Histogram, Pattern Matching, Expert system , Android App.*

*****

I. INTRODUCTION

India is an agricultural country; whereby regarding 70% of the population depends on agriculture. Farmers have wide range of diversity to pick appropriate Fruit and Vegetable crops. There are more problems which are faced by the farmer in agriculture field. Farmer is not able to disease detection and diagnosis of plants by naked eyes. Because of that there may be loss in money, time and decreased in crop production. To overcome these problems we are developing a expert system which is useful to solve these problems. In these expert system initial user capture image using mobile device these capture image is treated as a testing image. testing image is in the form of RGB image.these RGB image then converted into black-white image using gray scale conversion. grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. after that we are using CANNY's edge detection technique to calculate edge histogram of testing image and stored in database. There are two basic objective of CANNY's edge detection first is low error rate, all edges should be found , with a minimum of spurious responses and second is, Edge points should be well localized, The distance between one detected point and the true edge point should be minimum, single edge point response: only one point should be detected for each true edge point. In second phase we are using pattern matching technique to display result. Pattern matching technique is used to compare the testing image histogram and trained image histograms.Trained image histograms arealready classified and stored in database. We are using C5.0 classification algorithm to classify trained image histograms. C5.0 algorithm is the latest algorithm to classify data. C5.0 algorithm is follows the principle of rules of C4.5 and C4.5 algorithm follows the principle of ID3 algorithm. C5 algorithmic rule has several features like: the large decision tree may be viewing as a set of rules that is simple to grasp. C5 algorithm provides the acknowledge on noise and missing information. problem of over fitting and error pruning is solved by the C5 algorithmic rule. In classification technique the C5 classifier can anticipate which attributes are relevant and which aren't relevant in classification compare all the keep results which are classified by the C5.0 algorithmic rule And determine disease infected or not within the plants leaf. If it finds the exact match of image then it provides results associated with that image otherwise the link is provided that connects to the agriculture professional who provides the related info about that disease.

 ISSN: 2454-4248
603 – 606

II. RELATED WORK

In paper [1] they defined the system where the plant diseases identified by using image processing. Generally image processing has several steps to process digital image as per this research advance algorithms used to process the gray scale images and using those the results are generated.

[2] This paper uses the CANNY edge detection Algorithm which is more preferable as they compare other image processing algorithms like sobel edge detection which has less accuracy than the CANNY edge detection. In CANNY edge detection image is first converted to gray scale image then their edges are identified then their histograms are generated and according to those histograms other histograms of other image is compare for the match.

In this paper we are going to make predictions according to previous data. To classify the historical data we uses the classification algorithm i.e. c5.0. This algorithm makes the decision tree for the data that is provided to the algorithm using those decision trees the predictions are done.

In [3] the decision tree algorithm is used which takes the historical data makes classification of that data using different classification algorithm like k-means,c4.5 etc.

II. PROPOSED WORK

PHASE-I

Step1: In First phase of Disease Detection System, RGB image of Diseased plants is capture with the help of mobile application. Resolution of image has been set to capture image. Different mobile applications have different resolution to capture images.

Step2: Afterward these RGB image is convert into the gray scale image.Conversion of a color image to grayscale is not unique.In this Conversion color images are convert into black-white images.

PHASE-II

Step3: The second step of detection of plant diseases start with the Edge detection technique.in this technique, first we separate the layers of tested image into Red, Green and Blue layers and again apply CANNY's edge detection technique to detect the edges of layer's images.

Step4: Once the first phase is finished the histogram is generated for captured images with the help of CANNY's edge detection technique.

Step5: Once the histogram is generated for capture image, pattern matching algorithm is used to compare histograms of capture image and the histogram of trained images which are saved in database. After comparison result is displayed with the help of pattern matching algorithm that is the testing image is diseased or not with the name of disease and diagnosis

Gray Scale Conversion

In this system, users capture RGB images with the help of mobile application. Then to RGB images are converting into the gray scale images. Gray scale images are the images which are in the white and black color.

Grayscale Conversion Algorithm:

Step 1: start

Step 2: Capture image.

Step 3: Converting RGB image to grayscale.

Step 4: stop.

CANNY's edge detection technique

: There are five Steps followed by the CANNY's edge detection technique-

1. Smoothing Process : In the first step of the process smoothing is used to blur the image to remove noise from the image.

2. Finding gradients: In this step, edges of the image should be marked where the gradients of the image have maximum magnitudes.

3.Non-maximum suppression: Only local maxima should be marked as edges.

4. Double thresholding: Potential edges are determined by thresholding.

5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

Canny Detection :

1. The canny operator works in a multi-stage process. Then a simple 2-D first derivative operator is applied to the smoothed image to highlight regions of the image with high first spatial derivatives.

 ISSN: 2454-4248
603 – 606

2. Gradient is the first – order derivatives of image for each direction. Which is non maximal suppression .the gradient can be computed using central difference .

$\partial X(x, y) = [(x+1, y)-(x-1,y)]/2$

$\partial Y(x, y) = [(x, y+1)-(x,y\_1)]/2$

3. Magnitude of horizontal and vertical gradient is used for non maximal suppression process. The magnitude can be computed by :

$Magnitude=(\partial X(x) * \partial X(y)+ \partial Y(x) * \partial Y(y))$

Algorithm:

Step 1: start

Step 2: acquire the leaf image

Step 3: convert color image to grayscale

Step 4: apply CANNY's edge detection technique.

Step 5: calculate histogram

Step6:compare with edge histogram of the database image

Step 7: stop

Edge Histogram

Every leaf is in RGB color with varying intensity of red, green, blue. Depending upon the intensity of red, green, blue histogram is generated. There is different histogram for different colors. The edge histogram of gray scale image is different than RGB image

Algorithm:

Step 1: start

Step 2: acquire the leaf image

Step 3: calculate the histogram of the image.

Step 4: calculate the difference with the database image

Step 5: stop.

Proposed System Algorithm

Step 1: Capture image

Step 2: Set the images into suitable Resolution.

Step 3: Convert Capture image into Grayscale

Step 4: Apply CANNY's edge detection technique on the Captured images.

Step 5: Calculate the edge histogram of the captured images.

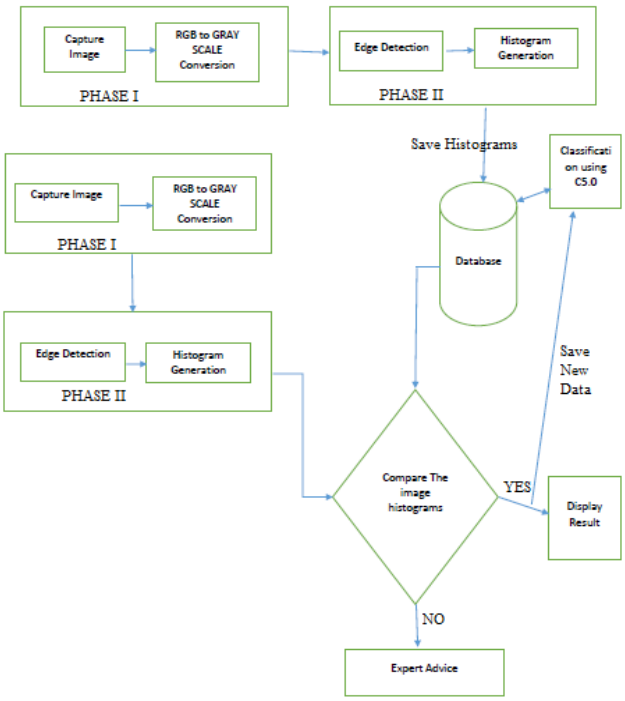
Step 6: Apply pattern matching technique on the histogram of captured images and the histograms of trained images stored in database.

Step 7: Display the result if captured image histogram and trained image histogram matches.

Step 8: Saved Capture images histogram in database as a trained images histogram.

Step 9: If histograms of captured images and trained images is not match then farmer can take expert advice to increase crop production.

Step 10: Stop.



Classification:

Trained images are stored in the database. Trained images are used for the detecting disease on the captured leaf images. There is high dimensionality is presented in the stored images in database. Because of high dimensionality we have to face serious challenges to learning methods. Because of large number of features , a learning model generate the result which is over-fited. So the performance is degraded. To overcome these disadvantages and to increase the performance we are using classification algorithm c5.0.

 ISSN: 2454-4248
603 – 606

C5.0 algorithm:

There are three classification algorithm ID3, C4.5, C5.0. C4.5 algorithm is the improved version of the ID3 algorithm and C5.0 algorithm is the improved version of C4.5 algorithm.

- C5.0 is significantly faster than C4.5.
- C5.0 is more memory efficient than C4.5.
- C5.0 gets similar results to C4.5 with considerably smaller decision trees.
- C5.0 support for boosting improves the trees and gives them more accuracy.
- C5.0 allows you to weight different cases and misclassification types.

Algorithm:

Step 1: Check for the base case.

Step 2: Construct Decision tree using training data.

Step 3: Find the attribute with the highest info gain (A_best).

Step 4:A_best is assigned with Entropy minimization.

Step 5: Partition S into S1, S2, S3…

Step 6: According to the value A_best.

Step 7: Repeat the steps for S1, S2, S3

Step 8: for each ti from D, apply DT.

III CONCLUSION AND FUTURE WORK

This is a accurate and efficient technique for automatically detection of plant diseased and diagnosis. The RGB to gray scale conversion and CANNY's edge detection techniques are applied on testing images of plant and the trained diseased images of the plant. Plant diseased is detected by using edge histogram and pattern matching technique. The histogram matching is based on the color feature and the edge detection technique. The training process includes the conversion of captured RGB image into gray scale conversion. After that edge histogram is calculated using CANNY's edge detection techinique. Once the histograms are generate.Immediately we applied the comparison technique based on the histogram. The comparison is firstly with the testing sample and the trained sample if the testing sample is diseased, it compare testing sample with the diseased sample and these steps take few minute to display the comparison result that is the testing sample is diseased or not and the disease name and diagnosis of disease is display as a result. This is beneficial for us because we are easily understood the processing of implementation phase. The future work mainly concerns with the large database. We can developed expert system in mother tongue language of the farmer area wise **.** we can provide pesticides and more information which is useful for agriculture.

IV. References

Journal Papers:

[1]. Jayamala K. Patil, Raj Kumar, [2011] ―Advances in image processing for detection of plant diseases‖, Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604 Vol 2, Issue 2, pp 135-141

[2]. ShitalBankar, AjitaDube, PranaliKadam, Prof. Sunil Deokule.[2014]- Plant Disease Detection Techniques Using Canny Edge Detection & Color Histogram in Image Processing,International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1165-1168

[3]. Rutvija Pandya, JayatiPandya[2015] C5.0 Algorithm to Improved Decision Tree with Feature Selection and Reduced Error Pruning, International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 16, May 2015

[4]. Swati Zode,R.Ambekar,[2015] Privacy Preserving Clinical Decision Support System Using C5.0 Algoruthm, in cPGCON 2015 held during 13-14 march 2015 at MET's Institute of Engineering,Nasik

[5]. Seema Sharma, Jitendra Agrawal, Shikha Agarwal, Sanjeev Sharma.[2013]" Machine Learning Techniques for Data Mining: A Survey,in proc. 2013 IEEE International Conference on Computational Intelligence and Computing Research.

[6]. Sudeep Marwaha, Subhash Chand, Arjit Saha Disease Diagnosis in Crops using Content Based Image Retrieval, in Proc. 2012 12th International Conference on Intelligent System Design and Application(ISDA).

[7]. Bujlow Tonasz, Riaz Tahir, Peder Sel Zensmyrup A Method for Classification of Network Traffic based on C5.0 Machine Learning Algorithm, in Proc.ICNC'12:2012 International Conference of Computing, Networking and Communication (ICNC)