Use of Image Processing Technique In The Field of Identification of Diseases In Plants

Sachin Kumar
Dept of Computer Science and Engineering
IITM, Murthal, Sonipat, Haryana (131039)

Abstract- *Agriculture has always been the most significant aspects of everyone's life. Agriculture is not just about feed growing inhabitants. It is interesting to note that about forty five percent of total world population relies mainly on agriculture for their livelihood i.e. it feeds a wide range of individuals. There are numerous diseases that impact plants with ability to cause losses. Disease is the primary reason for a reduced amount of yield in terms of quality and quantity. Timely detection and diagnosis of disease in a precise is crucial to avoid agricultural damages. The aim of this work is to build up a software structure which automatically helps in finding and categorizing the disease. The disease detection process involves steps: loading an image, processing, extraction of textural features, segmentation and in then classification. Image processing technique is very useful in agricultural applications to find and classify diseases.*

Keywords- classification, Image Processing, plant disease, Segmentation, Extraction.

I. INTRODUCTION

In today's modern-day era, agriculture has become one of the most vital components of the inhabitants. Agriculture is foundation of every country. Farmers usually have decent variety of reaps based on the ecological situation and geographic position. The crops cultivation for full profit and universal production remains mostly practical. This may perhaps be established through the help of specialist aid. The assistance that persistently cyclical yields need absolute capacity specifically for the disease supervision that might have a consequence lying on aspects of production drastically to get a monetary gain [1-2]. The process of image is the generally admirable practice that come up with a forfeited work in agrarian usage events. Number of diseases are there which severely disturb ecological, economic and social stability. Identifying illness from picture of plants be able to ease with the Backing for agrarian development but a large amount of the initial indicators are miniscule, as a result the detection phase of infection is constrained by human being graphic skills. This approach is dull, time consuming. There is demand meant for quality structure that automatically recognizes, categorizes, and quantitatively identifies infection indications. Mechanical uncovering of plant diseases is an imperative assessment matter as this one may possibly assist in inspection of enormous meadows of yields, and consequently find the warning sign of diseases as soon as they occur on the plant leaves. Infection is spread through micro organisms which is the means of perpetrating disease. Diseases management seems to be a challenging project [3]. Mostly the varieties of diseases are witnessed on plants's leaves, trunks and even on fruits of that plant. Because of complication of visible patterns the precise quantification of these visually diagnosed diseases, pests, and traits has not explored. In most of the cases, diseases are discovered on the leaves or stems of the plants. To prevent a human being's interference, creating a computer vision structure to discover, identify, and classify illness impacted on reaps and therefore results in appropriate balance about disease infection and further assessment. Automatic detection of disease can prove to be advantageous in exploring large field of crops. The detection of disease in leaves of plants has been the most significant assessment subject. The symptoms appearing on plant leaves will help in easy identification of diseases. The process of image processing plays the most important part [4]. The image processing by means of MATLAB begins with securing of digital superior resolution images. Images of nutritious and unhealthy leaves are taken for research. The pre-processing of these images is done for enhancing them. Captured leaf images are segmented utilizing "k" means cluster technique to generate clusters. Choices are extracted before employing K-means and KNN Classifier for training and classification. Lastly, diseases are identified by the approach given.

This work delivers an overall view and significance of disease detection, a brief literature reviews of leaf illness detection practices, methodology of proposed approach established on MATLAB image processing and obtained results are discussed.

II. PLANT DISEASE BASIC

Usually, Plant diseases are of three types. They're categorized in microorganism as Bacterial, viral and fungal. These bacterial's are further have sub categories as soft spots,

spots and wilts. The viral ones are divided as motting, distortion and dwarfing. The fungal ones are categorized further as sooty molds, rusts, mildews, rots, spots and cankers [5].

A. Literature Survey

Vijayi Singh et al [6] discussed regarding an algorithm for image segmentation method for automatic detection and classification of plant leaf diseases andgo through a survey on various diseases classification practices that can be applied for plant leaf disease recognition. Genetic algorithm was employed for leaf color image segmentation method and the classification was performed in two stages. In the first stage MDC with K- means clustering with the precision of93.63% and in the second stage, the precision is enhanced to 95.71% with SVM classifier. A Shankar Reddy et al [2] utilized Image segmentation as the fundamental method for digital image processing. Otsu method is very effective approach for image thresholding due to its straight forward computation. He offered an effective algorithm which comprises of three phases. In the initial phase, contrast enhancement algorithm is employed to generate a high-quality image. Following that created a bimodal view for the resulting image. Lastly, employed the Otsu's binarization on top of the resultant image. Murali Krishnan et al [15] described a method for detection of Bacterial leaf scorch infection in plant. In image segmentation, "K-means clustering algorithm is applied for separating foreground and background images". Clustering in segmentation is based on subtracting the clustered leaf images and intensity mapping for highlighting leaf area. K-means is found to be very operational as well as a simpler method for detecting infected area.

B. Proposed System

The plant disease detection is the practice that has been employed for identification of disease from recorded leaf image of the plant. The plant disease detection practice which is anticipated here is divided in following discussed steps:

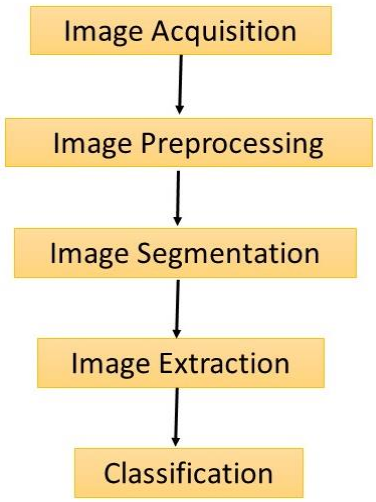


Figure 1. Architecture Diagram

C. Image Preprocessing

In this phase, grayscale conversion of the given plant leaf image is performed. Grey Level Cooccurrence Matrix (GLCM) algorithm is employed that helps in extraction of the textural characteristics of input image. It is a statistical approach of analyzing texture that considers the spatial link of pixels, frequently called as spatial gray level dependence matrix. The GLCM functions clarify the texture of a picture by shrewd how often sets of pixel with values and in a particular spatial relationship arise in the leaf image, creating a GLCM, and after that obtaining statistical measures from the matrix. These statistics provide data about the texture of an image. Statistics like contrast, correlation, energy, homogeneity etc.

D. GLCM algorithm

1. Count all the number of pixels in the matrix in which the data is saved.
2. Store the counted pixels in matrix P [I,j].
3. Check the similarity between pixels in the matrix by applying histogram technique.
4. Calculate contrast factor from the matrix:

$$g = exp\left[\frac{mean(I) - minimum(I)}{maximum(I) - mean(I)}\right]$$

1. The elements of g need to be normalized by dividing the pixels.

$$g = \begin{bmatrix} 0.8\ if\ g < 0.8 \\ 1.2\ if\ g > 1.2 \\ g\ otherwise \end{bmatrix}$$

E. Image Segmentation

The Image segmentation technique will partition an image based on their properties. Image segmentation process

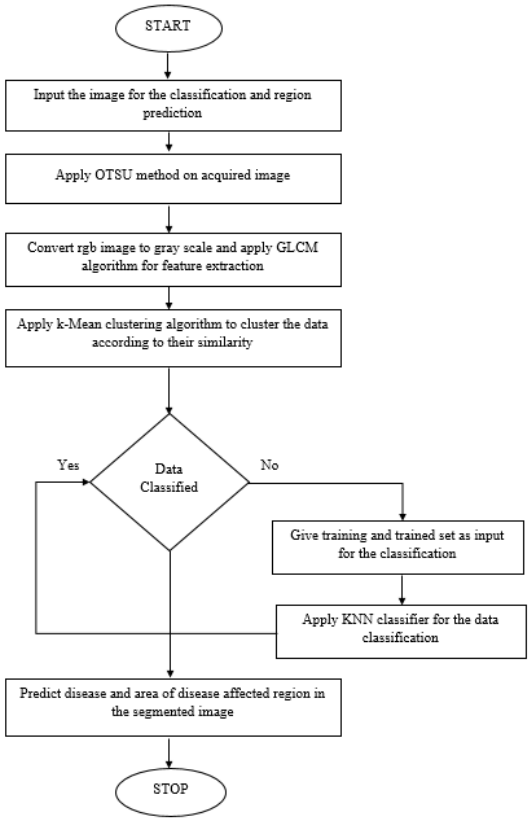
is mostly classified into threshold-based segmentation and region based segmentation. Region-based k-mean segmentations and threshold based Otsu method is applied for the image segmentation in this work. We are using Otsu method only for getting high contrast image to calculate disease affected area and using K mean segmentation for grouping image into different clusters depending upon their features. K-means clustering algorithm is used to classify or to cluster the objects supported attributes/features into K groups. K is assumed as a positive integer number. The grouping is finished by minimalizing the distances between information and the corresponding cluster centroid. The aim of cluster analysis is to cluster information in such a means that like objects are in one cluster and objects of various different clusters are dissimilar.

INPUT: Dataset
OUTPUT: Clustered Data
Start ()
1. Read dataset and dataset has number of rows "r" and number of columns "m"
2. For (i=0 ;i=r; i++) /// selection of medoid point
   i. For (j=0; j=m; j++)
   ii. Select k=data (i, j);
   End
1. Calculation of Euclidian distance()
   i. For (i=0;i=r;i++)
   ii. For (j=9;j=m;j++)
   iii. A(i)=data(i);
   iv. B(i)=data(j);
   v. Distance =squt[(A(i+1)-A(i)^2) –(B(j+1)-B(j)^2);
   End
1. Normalization ()
   i. For (k=0;k=data;k++)
   ii. Swap k(i+1) and k(i);
   End
1. Repeat step 3 to 4 until all points get clustered.

F. Classification

We are using KNN as classification technique in our method which is applied to categorize same and different data into one or more classes. "K-Nearest neighbor classifiers" rely upon learning by analogy. The n-dimensional numeric attributes are used for training samples. A single point in an n-dimensional space is used for representation of a single sample. Along these lines, the larger part of the training samples is accumulated in an n-dimensional pattern space. At the point when given an anonymous sample, a k-nearest neighbor classifier looks at the pattern space for the k training samples that are close to the anonymous sample. "Closeness" is outlined in terms of Euclidean distance. Not in the least like decision have tree induction and back propagation, nearest neighbor classifiers given break even with weight to each attribute. This may bring about uncertainty when there are various irrelevant attributes in the data. Nearest neighbor classifiers can be utilized for prediction, that is, to give back a genuinely valued prediction for a given anonymous sample. In this situation, the classifier provides the average value of the genuine valued linked with the k nearest neighbors of the unknown sample. The k-nearest neighbors' algorithm is among the plainest of all machine learning algorithms.

G. FLOW CHART OF PROPOSED METHOD



III. RESULT & DISCUSSION

Our proposed method is implemented on leaf dataset. We have taken two datasets: training and test dataset. We have used 100 images in our training dataset for training our system to detect diseases by classifying disease on basis of textural features extracted using GLCM. We have used 15 images in test dataset which we used to give as input for classify disease using KNN classifier. We have implemented both SVM and KNN classifier to detect disease in the input image. We are using three parameters accuracy, false acceptance rate and execution time for comparing SVM and KNN classifier.

(i) **Accuracy:** The accuracy is defined as the percentage of correctly classified instances (TP + TN)/(TP + TN + FP + FN). Here; "TP, FN, FP, and TN represent the number of true positives, false negatives, false positives and true negatives, respectively". We can say accuracy is correctly predicted class divided by total testing class and multiplying with 100 if we want percentage. [37]

(ii) **False Acceptance Rate:** False Acceptance Rate (FAR) is stated as, "the ratio of number of false accept instances divided by the number of instances". It means that how many times system accepts wrong answer as a right answer.

FAR= FA /TA where FA is False Acceptances and TA is Total Acceptances. [38]

(iii) **Execution Time:** Execution time can be defined as, "time taken by the system for executing complete task, including the time spent executing run-time or system services on its behalf".

A. Accuracy

As shown in figure 2, the accuracy of the proposed technique is enhanced in comparison to the existing technique. For the latter technique, it is found to lie from 75-85% and that of the proposed technique lies in the range of 90-95%.

Table 1: Comparison of Accuracy

Image No	Existing Method Using SVM(%)	Proposed Method Using KNN(%)
1	83.5	91.5
2	78.6	94.9
3	84.7	91.4
4	77.7	91.6
5	85.2	95.1

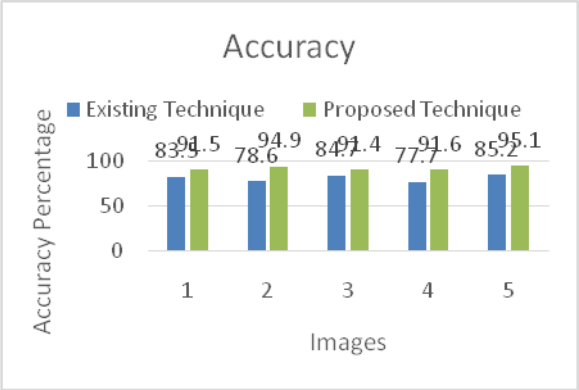


Fig. 2: Accuracy Comparison

B. False Acceptance Rate

As shown in figure 5.1.2, it is clear that in case of present technique, the false acceptance rate has been reduced when compared with an already existing technique. The FAR of the latter technique lies within the ratio of 0.8-0.9 and that of the proposed technique is found to lie in the range 0.7-0.8.

Table 2: False Acceptance Rate Comparison

Image No	Existing Method Using SVM(%)	Proposed Method Using KNN(%)
1	0.935	0.815
2	0.886	0.849
3	0.947	0.814
4	0.877	0.816
5	0.952	0.851

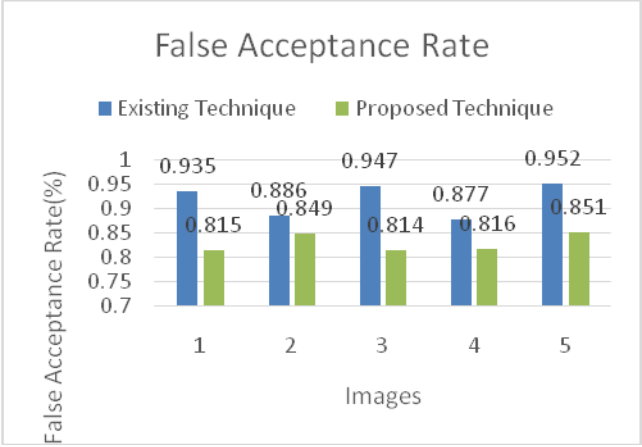


Fig. 3: False Acceptance Rate Comparison

C. Execution Time

As shown in figure 5.1.3, it is analyzed that the proposed technique takes less execution time as compared to existing technique. Proposed method takes approximately 1.1 to 1.5 seconds whereas the existing method takes 2 or more than 2 seconds.

Table 3: Execution Time Comparison

Image No	Existing Algorithm Using SVM(seconds)	Proposed Algorithm Using KNN(seconds)
1	2.45	1.23
2	2.31	1.45
3	2.2	1.89
4	2	1.21
5	2.34	1.13

Fig. 4: Execution Time Comparison

IV. CONCLUSION

The former approach will specially define the nature of disease which impacts the leaf. We are going to provide a result with a lesser amount of time with better accuracy. The structural design of the proposed work is shown in below Figure 3. In this work, "SVM classifier" has been replaced with the "KNN classifier" for classification of data into multiple classes. The efficiency of the algorithm here has been checked in terms of accuracy, execution time and false acceptance rate where accuracy increases up to 10 percent and execution time and false acceptance rate decreases as compared to existing technique.

REFERENCES

[1] R. Chaware, R. Karpe, P. Pakhale and S. Desai"Detection and Recognition of Leaf Disease Using Image Processing"International Journal of Engineering Science and Computing, May 2017.

[2] S.Raut and A. Fulsunge "Plant Disease Detection in Image Processing Using MATLAB" International Journal of Innovative Research in Science, Engineering and Technology, vol. 6, no. 6, pp. 10373-10381, June 2017.

[3] S. Bashir and N. Sharma "Remote Area Plant Disease Detection Using Image Processing" IOSR Journal of Electronics and Communication Engineering (IOSRJECE), vol. 2, no. 6, pp. 31-34,Sep-Oct 2012.

[4] S. Varshney and T. Dalal "Plant Disease Prediction using ImageProcessing Techniques- A Review"International Journal of ComputerScience and Mobile Computing, Vol.5 Issue.5, pp. 394-398, May- 2016.

[5] A. Devaraj, K. Rathan, S. Jaahnavi and K. Indira, "Identification of Plant Disease using Image processing Technique", International Conference on communication and signal processing, vol. 978, pp. 749-753, April 2019.

[6] V. Singh and A.K.Mishra, "Detection of Plant Leaf Diseases Using Image Segmentation and soft computing Techniques" Information Processing in Agriculture, vol. 4, no. 1, pp. 41-49, March 2017.