IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 2, April 2014.

www.ijiset.com

ISSN 2348 - 7968

DETECTION AND CLASSIFICATION OF PLANT DISEASES BY IMAGE PROCESSING

NIKET AMODA[#1], BHARAT JADHAV[#2], SMEETA NAIKWADI[#3]

[#]Department of Electronics & Telecommunication,K.C.College of Engineering, Management Studies & Research
University of Mumbai, India

***Abstract*-- The aim of this project is to design, implement and evaluate an image processing software based solution for automatic detection and classification of plant leaf disease. However studies show that relying on pure naked-eye observation of experts to detect and classify diseases can be time consuming and expensive, especially in rural areas and developing countries. So we present fast, automatic, cheap and accurate image processing based solution. Solution is composed of four main phases; in the first phase we create a color transformation structure for the RGB leaf image and then, we apply color space transformation for the color transformation structure. Next, in the second phase, the images are segmented using the K-means clustering technique. In the third phase, we calculate the texture features for the segmented infected objects. Finally, in the fourth phase the extracted features are passed through a pre-trained neural network.**

Index Terms- Artificial Intelligence, Image Processing

Keywords: Segmentation, neural network, color co-occurrence, plant leaf disease, K-means Method.

I. INTRODUCTION

Plant diseases have turned into a big problem as it can cause significant reduction in both quality and quantity of agricultural products. 70% of the Indian population depends on agriculture. Farmers have wide range of diversity to select suitable Fruit and Vegetable crops. But the cultivation of these crops for optimum yield and quality produce is highly technical. It can be improved by the aid of technological support. The management of perennial fruit crops requires close monitoring especially for the management of diseases that can affect production significantly. Many authors have worked on the development methods for the automatic detection and classification of leaf diseases based on high resolution multispectral and stereo images. The philosophy behind precision agriculture is not only including a direct economical optimization of agricultural production, it also stands for a reduction of harmful outputs into environment and non-target organisms. In particular a contamination of water, soil, and food resources with pesticides has to be minimized in crop production. There are two main characteristics of plant disease detection machine-learning methods that must be achieved, they are: speed and accuracy. In this study an automatic detection and classification of leaf diseases is been proposed which is based on K-means as a clustering.

II. THE PROPOSED APPROACH STEP- BY-STEP DETAILS

The proposed approach starts first by creating device-independent color space transformation structure. Thus we create the color transformation structure that defines the color space conversion. The next step is that we apply device-independent color space transformation, which converts the color values in the image to color space specified in the color transformation structure. The color transformation structure specifies various parameters of transformation. A device independent color space is the one where the resultant color depends on the equipment used to produce it. For example the color produced using pixel with a given RGB values will be altered as brightness and contrast on display device used. Thus the RGB system is a color space that is dependent. To improve the precision of the disease detection and classification process, a device independent color space is required. In device independent color space, the coordinates used to specify the color will produce the same color regardless of the device used take the pictures. The K-means clustering algorithm tries to classify *objects* (pixels in our case) based on a set of features into K number of classes. The classification is done by minimizing the *sum of squares* of distances between the objects and the corresponding cluster or class *centroid* [1; 2]. However, K-means clustering is used to partition the leaf image into four clusters in which one or more clusters contain the disease in case when the leaf is infected by more than one disease. In our experiments multiple values of number of clusters have been tested. Best results were observed when the number of clusters was 3 or 4.

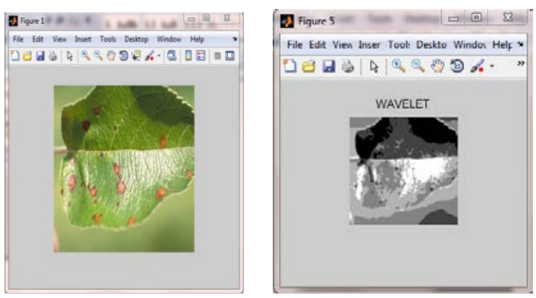

Figure 1: A leaf image infected with tomato plant disease;
a) Original image b) wavelet of image

The overall concept that is the algorithm of image classification and processing is the same. First the digital images are acquired from the field or environment using digital camera. The image processing techniques are applied to the acquired image to extract the needful features.

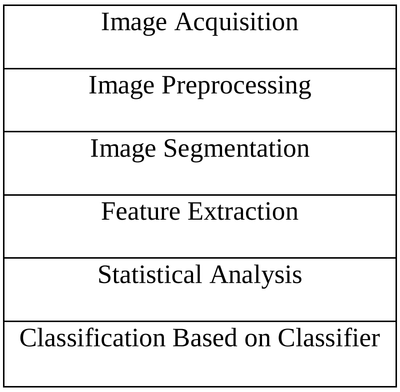

Figure2: The basic procedure of the proposed image processing- based disease detection solution.

The proposed approach step - by - step of the image segmentation and recognition processes is illustrated in Algorithm-

Algorithm: Basic steps describing the proposed algorithm.

1. RGB image acquisition
2. Create the color transformation structure
3. Convert the color values in RGB to the space specified in the color transformation structure
4. Apply K-means clustering
5. Masking green-pixels
6. Remove the masked cells inside the boundaries of the infected clusters.
7. Segment the components
8. Obtain the useful segments
9. Computing the features using color co-occurrence methodology.
10. Configuring Neural Networks for Recognition

Segmentation subdivides an image into its constituent regions or objects. The level to which the subdivision is carried depends on the problem being solved. That is, segmentation should stop when the object of interest in an application have been isolated. For example:, in the automated inspection of electronic assemblies, interest lies in analyzing images of the products with the objective of determining the presence or absence of specific anomalies, such as missing components or broken connection paths. There is no point in carrying segmentation past the level of detail required to identify those elements. Segmentation algorithms for monochrome images generally are based on one of two basic properties of image intensity values: discontinuity and similarity.

K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. *K-means* clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

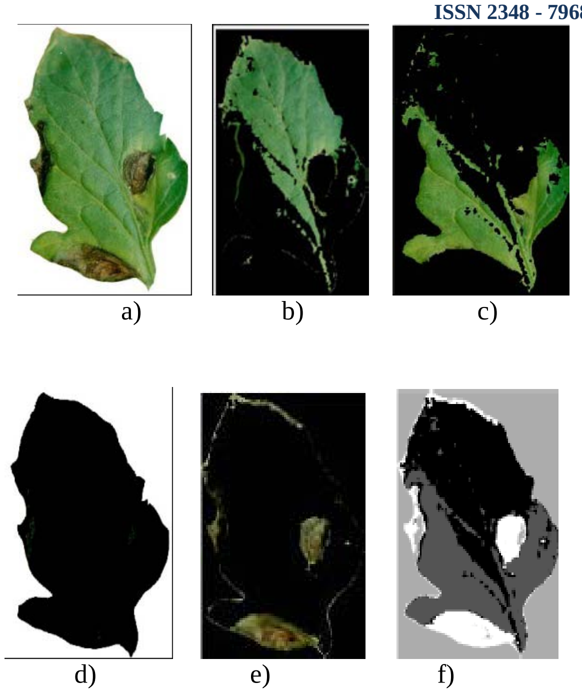

Figure 3: An example of output of K-means Clustering for a leaf infected with early scorch disease. (a)The infected leaf picture. (b, c, d, e) the pixels of the first, second, third & fourth cluster respectively & (f) a single gray-scale image with the pixel colored based on their cluster index.

K-means clustering is used to partition the leaf image into four clusters in which one or more clusters contain the disease in case when the leaf is infected by one or more disease. K-means algorithm was developed by Mcqueen(1967) and then by Hartigan and Wong(1979). The K-means clustering algorithms tries to classify pixels based on a set of features into k-number of classes. The classification is done by minimizing the sum of squares of distance between the objects and the corresponding clusters or class centroid.

In our project, the K-means clustering is set to use squared Eucledean distances. An output of K-means clustering for a leaf infected with early scorch disease is shown in Fig 3f.

The next step in algorithm is the feature extraction. For feature extraction the method used is color co-occurrence method. It is the methodology in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image.

The use of color image features in the visible light spectrum provides additional image characteristic features over the traditional gray-scale representation. The color co-occurrence method consists of three major mathematical processes. First, the RGB images of leaves are converted to HIS color space representation. Once this process is completed , each pixel map is used to generate a color co-occurrence matrix, resulting in three color co-occurrence matrices, one for each of H, S, I. However, we use GLCM function in Matlab to create gray-level co-occurrence matrix; the number of gray levels is set to 8, and the symmetric value is set to “true”, and finally, offset is given a” 0” value.

Input data preparation: - In our experiments, two main files were generated, namely: (i) Training texture feature data and (ii) Testing texture feature data. The two files had 192 rows each, representing 32 samples from each of the six classes of leaves. Each row had 10 columns representing the 10 texture features extracted for a particular sample image. Each row had a unique number (1, 2, 3, 4, 5 or 6) which represented the class (i.e., the disease) of the particular row of data. „1‟ represented early scorch disease infected leaf. „2‟ represented Cottony mold disease infected leaf. „3‟ represented ashen mold disease infected leaf. „4‟ represented late scorch disease infected leaf. „5‟ represented tiny whiteness disease infected leaf, and „6‟ represented normal leaf. Then, a software program was written in MATLAB that would take in .mat files representing the training and testing data, train the classifier using the „train files‟, and then use the test file‟ to perform the classification task on the test data. Consequently, a Matlab routine would load all the data files (training and testing data files) and make modifications to the data according to the proposed model chosen. In the experimental results, the threshold value for each of the above categories is constant for all samples infected with the same disease. This threshold is a global image threshold that is computed using Otsu's method [3; 4]. The architecture of the network used in this study was as follows. A set of 10 hidden layers in the neural network was used with the number of inputs to the neural network (i.e. the number of neurons) is equal to the number of texture features listed above. The number of output is 6 which is the number of classes representing the 5 diseases studied along with the case of normal (uninfected) leaf. Those diseases are early scorch, cottony mold, ashen mold, late scorch, tiny whiteness. The neural network used is the feed forward back propagation with the performance function being the Mean Square Error (MSE) and the number of iterations was 10000 and the maximum allowed error was 10^{-5}.

III. E XPEREMENTAL RESULTS

The experimentation starts with two samples, one for the healthy leaf sample and second for the diseased leaf sample. The results for first step which are input to MATLAB, the following figure shows the diseased leaf sample of tomato leaf.

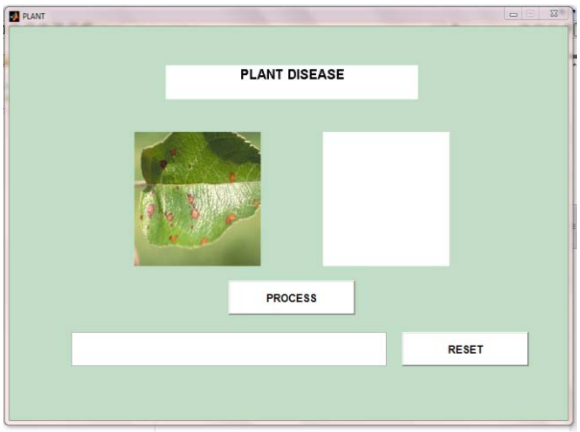

Figure 4: Diseased leaf sample of tomato plant.

The below figure shows the observation result for diseased sample layer. These results show the layers separation of RGB diseased sample into Red, Green & Blue. The layers separation is necessary for the edge detection. We cannot apply edge detection directly on the RGB diseased sample. So we convert the image into gray scale image as shown below.

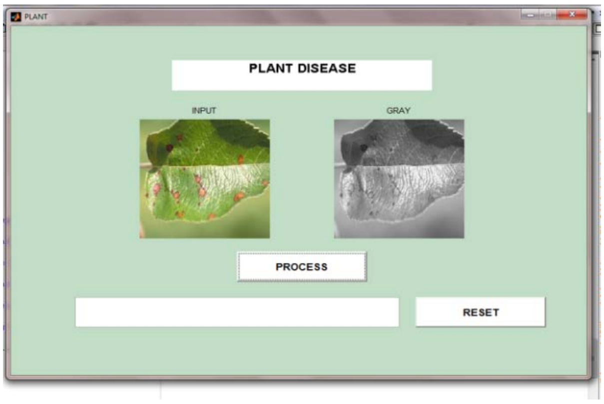

Figure 5: Gray scale image of diseased tomato leaf.

The next step is to renormalized the gray image sample and apply the segmentation method. In this the diseased part is segregated to large extend and the healthy part of the leaf is removed. The segmentation result is as shown below.

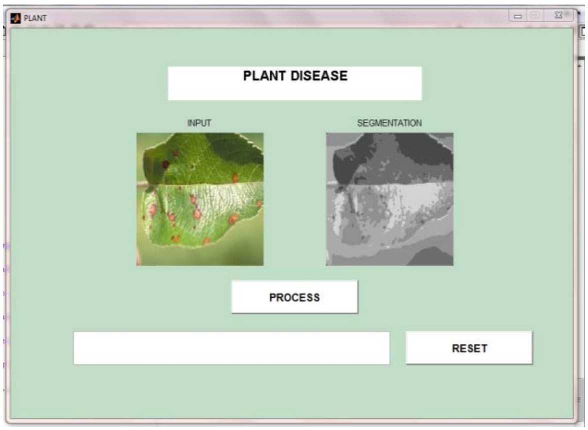

Figure 6: Segmentation result for diseased tomato leaf.

The full screen result and classification of the disease is shown in figure7. This figure shows the name that is the classification of the disease.

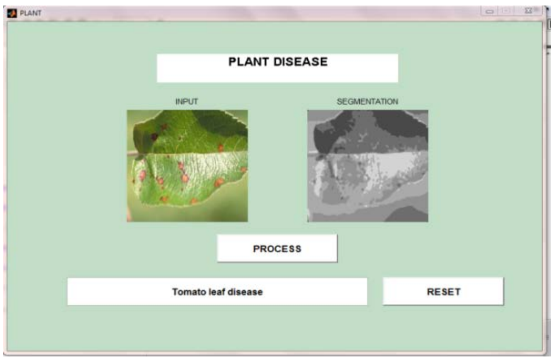

Figure 7: Full screen result of diseased tomato leaf.

IV. CONCLUSION AND FUTURE WORK

To wind up all the information discuss above, I would like to conclude that it is a efficient and accurate technique for automatically detection and classification of plant diseases. In this research, plant diseased is detected and is also classified. The histogram matching is based on the color feature and the edge detection technique. The color features extraction are applied on samples that are contained the diseased leaf of the plant. The training process includes the training of these samples by using layers separation technique which separates the layers of RGB image into red, green, and blue layers and edge detection technique which detecting edges of the layered images. The future work mainly concerns with the large database and advance feature of color extraction that contains a better result of detection. Another work concerns with research work in a particular field with advance features and technology.

REFERENCES

[1]. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In L. M. LeCam and J. Neyman, editors, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, volume 1, pages 281--297, Berkeley, CA, 1967. University of California Press. URL: http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.bsmsp/1200512992.

[2]. Hartigan, J. A.; Wong, M. A. (1979). "Algorithm AS 136: A K-Means Clustering Algorithm". Journal of the Royal Statistical Society, Series C (Applied Statistics) 28 (1): 100–108.

[3]. Otsu, N. (1979). "A threshold selection method from gray-level histograms". IEEE Trans. Sys., Man., Cyber. 9: 62–66. DOI:10.1109/TSMC.1979.4310076.

[4]. Otsu, N.,"A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.

[5]. Brendon J. Woodford, Nikola K. Kasabov and C Howard Wearing [1999]. Fruit Image Analysis using Wavelets, Proceeding of The ICONIP/ANZIIS/ANNES.

[6]. Mix, F.X. Pico and N.J Oubor[2003]. A comparison of stereomicroscope and Image analysis for Quantifying Fruit Traits, seed technology. Vol 1.25 no.1.

[7] Niket Amoda, Ramesh K Kulkarni, "Efficient Image Retrieval using Region Based Image Retrieval", International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA, (ICWAC 2013)

[8] Niket Amoda, Ramesh K Kulkarni, "Image Segmentation and Detection using Watershed Transform and Region Based Image Retrieval", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)-Volume 2, Issue 2, March - April 2013

[9] Niket Amoda, Ramesh K Kulkarni, "Efficient Image Segmentation Using Watershed Transform", INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & TECHNOLOGY (IJCST)-VOL IV ISSUE II, VER. 2, APR. TO JUNE, 2013

[10] Niket Amoda, Smita Naikwadi, "Advances in Image Processing for Detection of Plant Diseases", International Journal of Application or Innovation in Engineering & Management (IJAIEM) – ISSN: 2319 - 4847, Volume 2, Issue 11, November 2013.

AUTHORS PROFILE

Niket Amoda received his B.E. in Electronics & Communication Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India, in 2006 and . He completed his Diploma in Advance Computing from ACTS, Pune, India in 2008 and M.E. in Electronics & Telecommunication Engineering from Vivekanand Education Society's Institute of Technology (VESIT), University of Mumbai, India, in 2013.

He worked as a Software Development Engineer in Ford Motor Company, India in 2008 & 2009. Currently working as a Assistant Professor in Electronics & Telecommunication Engineering Department of KC College of Engineering & Management Studies & Research, Thane, India. He has many publications in international journals and international conferences. His research area is Image processing. area is Image processing.

Smeeta Naikwadi received her diploma in Industrial Electronics from Bharti Vidyapeeths' College of Engineering, Kharghar, India in 2010. She is currently pursuing degree in department of Electronics and Telecommunication from KC College of Engineering & Management & Research, Thane, India.

Her areas of interest are Embedded and Telecommunication. She has published her project in International Journal of Application or Innovation in Engineering & Management.

Bharat Jadhav pursuing his B.E. in Electronics & Telecommunication Engineering from K.C.College of Engineering & Management Studies &

ISSN 2348 - 7968

Research, Thane(E).He completed his A+&N+ Hardware & Networking Certification from Zoom Technology Bangalore's Avera Academy Thane(W) in 2009.He completed his HSC & SSC Exam from Latur Board.

He worked as a Hardware Engineer in Nitin Computers (Pvt.) Ltd Kalyan(W),Thane, India in 2008&09.He has done AIESEC Hacking Certification from IIT Kharagpur in 2010.Also participated in various inter-college technical festivals in specially in Robotics.