Leaf Disease Detection and Classification using Image Processing

Nikita Khurpade[1], Prof. Nutan Dhande[2]
[12]*Department of Computer Science and Engineering*
[12]*Agnihotri College of Engineering Nagthana India*

Abstract- Tea is a popular beverage all around the world, and in India the cultivation of tea plays a vital role. Many diseases affect the proper growth of tea leaves leading to its reduction, thus hindering of the production of tea. However, if the disease is identified at an early age it would solve all the above mentioned problems through the application of appropriate treatment, or through the pruning of the diseased leaves to prevent further spread of the disease. To solve this problem image processing is the best option to detect and diagnose the disease. The main goal of this research is to develop an image processing system that can identify and classify the four most widespread tea leaf diseasesfrom a healthy leaf. Disease identification is the first step; there are many methods that have been used for identifying the leaf disease. In this paper, K-NN classifier (K-NN) is used to recognize the diseases. Thirteen features are analyzed during the classification. These features are then used to find the most suitable match for the disease (or normality) every time an image is uploaded into the K-NN database. When a new picture is uploaded into the system the most suitable match is found and the disease is recognized. The approach is novel since the number of features compared by the K-NN classifier is reduced by three features compared to previous researches, without adversely sacrificing the success rate of the classifier, which retains an accuracy of more than 96%. This also speeds up the identification process, with each leaf image taking 200 ms less processing time compared to previous research using K-NN, thus ensuring a greater number of leaves can be processed in a given time frame. The proposed solution increases in efficiency of the detection, identification, and classification process will enable the tea industry in India to become more competitive globally, by reducing the losses suffered due to diseases of the leaf, and thus increasing the overall tea production rate.

Keywords- Tea Leaf, Detection, Classification

I. INTRODUCTION

Tea is one of the essential beverages in India. Most of the Indian people start their day with a cup of tea. Assam has become an important tea producing country. Today the country has 172 commercial tea estates [1]. The districts that produce tea are Maulvibazar, Habiganj, Sylhet, Chittagong, Panchagarh, Brahmanbaria, and Rangamait [2]. Almost the entirety of the district of Tea production in India is greatly hindered due to a number of pests and diseases, caused by a variety of insects, mites, nematodes, bacteria, algae, fungi, weeds, and other diseases which are caused due to the environmental condition of that particular region [2]. India is an agricultural country where more than 75% population rely on agriculture directly or indirectly [4]. Approximately 20% to 30% of the tea leaves are lost due to various diseases each year [5].

Farmers in the field judge the identification of tea leaf diseases with their naked eye and previous experience. Many a times, experts are needs to be called in to analyze the tea leaves when there is ambiguity in detecting the diseases by local farmers; this process is not only time consuming, but also costly. It is important to catch the spread of the disease in its early stages before they reach epidemic proportions; otherwise the disease can spread quickly throughout the entire plantation, resulting in huge losses for the farmers. To aid the farmers in the crucial task of identifying tea leaf diseases in their infancy, it is practical to have an intelligent system of detection, identification, and classification system in place as a preventative measure.

The first sign that something is wrong with the leaf is usually indicated by a change in color from a healthy dark green hue. When the tea leaf is healthy the color is distinct, but when the leaf is affected by disease, the color of the leaf changes drastically. Each disease usually has a distinguishable leaf color and texture as symptoms. The latest trends of research in agriculture are toward the use of gene technology to develop disease resistant variant of the plant, and to increase food quality and productivity of the plant with reduced expenditure [6]. Numerous technological improvements are responsible for the progress in crop management techniques in recent times; including advances in information technology, remote sensing technology, and image processing and pattern recognition .Therefore, now it is possible to develop and deploy an autonomous system for detection, identification, and classification of diseases in crops in very large fields with minimal manual input.

A search through recent literature have identified research in various types of crop diseases including diseases in rice, citrus, Betel vine, and wheat leaf to name a few [8]. However, research into diseases of tea leaves is one area that has not yet seen any significant efforts. Therefore, there should be a way

to develop tea leaf disease recognition and detection to help the tea industry in India.

II. LITERATURE REVIEW

The main goal of this research is to develop an image processing system that can identify and classify the two most widespread tea leaf diseases in Bangladesh, namely brown blight disease and the algal leaf disease, from a healthy leaf. Disease identification is the first step; there are many methods that have been used for identifying the leaf disease. In this paper, Support Vector Machine classifier (SVM) is used to recognize the diseases.[1]

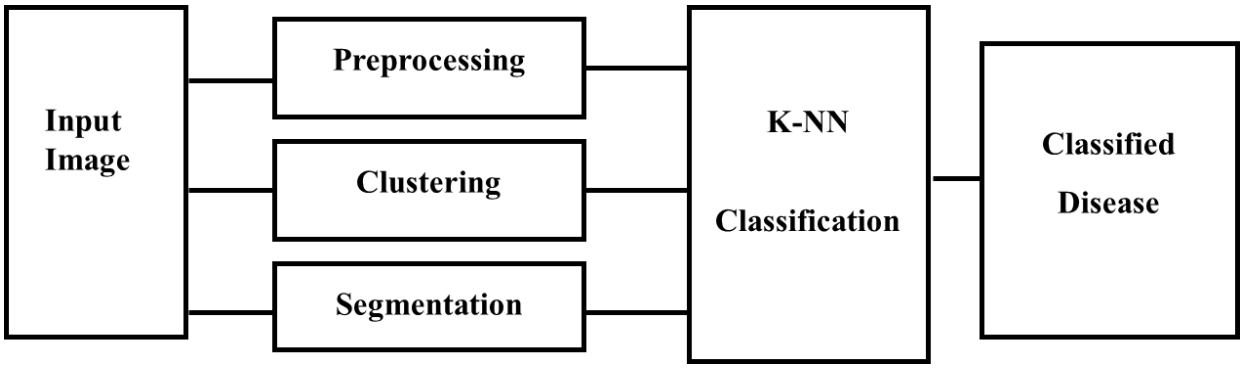
In this paper, the authors evaluate mainly in three well regulated manners: first Detection of Leaf Disease, second Classification i.e. recognizes the type of disease and the last one is Quantification of Disease i.e. measures the severity of Disease.[2]

Provides an autonomous approach of modified SVM-CS is introduces. Here, concept of cuckoo search is considered to optimize the classification parameters. These parameters further help to find more accurate solutions. This autonomous approach also extracts the healthy portion and disease affected leaf portion along with the accuracy of results.[3]

Kholis Majid, et al. [7], has added to a portable application for paddy plant malady identification framework utilizing fuzzy entropy and Probabilistic neural system classifier that keeps running on Android Versa Tile's framework. It includes the identification for all sorts of maladies, in particular brown spot, leaf blast, tungro and bacterial leaf blight. The exactness of paddy sicknesses distinguishing proof is 91.46 percent.

Elham Omrani et al. [10], used Support Vector Regression (SVR) based on radial basis functions to identify and classify diseases of the apple tree. It is a three step process. First, the captured images of the leaves had to be changed into a device independent color space, such as CIELAB, from a device depended format such as Red-Green-Blue (RGB) color space. Then, the image was segmented to extract the infected area from the overall leaf image. The segmentation technique employed was a region-based one using K-means clustering, wavelet, and grey-level co-occurrence matrix. This features extracted using this type of segmentation are the color, shape, and texture. These types of segmentation techniques are normally used for region.

III. PROPOSED SYSTEM FLOW



A. Proposed Methodology

i. Image Pre-processing

Noise gets added during acquisition of leaf images. So we use different types of filtering techniques to remove noise. We create 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 to take the pictures. CIE L*a*b is a device independent color space in which a & b components carry color information

ii. Image segmentation (k-means clustering)

Image segmentation is the process used to simplify the representation of an image into something that is more meaningful and easier to analyse. K-means clustering is a partitioning method. The function 'kmeans' partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Unlike hierarchical clustering, k-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters. The distinctions mean that k-means clustering is often more suitable than hierarchical clustering for large amounts of data. K-means treats each observation in your data as an object

having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible.

iii. Masking green pixels

In this step, we identify the mostly green colored pixels. After that, based on specified threshold value that is computed for these pixels, the mostly green pixels are masked as follows: if the green component of the pixel intensity is less than the pre computed threshold value, the red, green and blue components of the this pixel is assigned to a value of zero. This is done in sense that the green colored pixels mostly represent the healthy areas of the leaf and they do not add any valuable weight to disease identification and furthermore this significantly reduces the processing time.

iv. Removing the masked cells

The pixels with zeros red, green, blue components as well as pixels on the boundaries of infected cluster are completely removed. This is helpful as it gives more accurate disease classification and significantly reduces the processing time. Infected cluster is converted from RGB to HSI color format.

v. GLCM methodology

Gray level Co-occurrence matrix (GLCM) is generated for each pixel map

1. The graycomatrix function creates a gray level co-occurrence matrix by calculating how frequently a pixel with the particular intensity value i occurs in a specified spatial relationship to a pixel with the value j.

2. By default this spatial relationship is the pixel of interest and its immediate right pixel.

3. However we can specify some other spatial relationship between twos. To create multiple GLCMs, specify an array of offsets to the graycomatrix function. These offsets define pixel relationships of varying direction and distance.

4. Calculating statistics from GLCM matrix also known as SGDM

IV. CONCLUSION

This study summarizes major image processing used for identification of leaf diseases are k-means clustering, SVM. This approach can significantly support an accurate detection of leaf disease. There are five steps for the leaf disease identification which are said to be image acquisition, image pre-processing, segmentation, feature extraction, classification. By computing amount of disease present in the leaf, we can use sufficient amount of pesticides to effectively control the pests in turn the crop yield will be increased. We can extend this approach by using different algorithms for segmentation, classification. By using this concept the disease identification is done for all kinds of leafs and also the user can know the affected area of leaf in percentage by identifying the disease properly the user can rectify the problem very easy and with less cost.

V. REFERENCES

[1]. Md. Selim Hossain1 , Rokeya Mumtahana Mou1 , Mohammed Mahedi Hasan2 , Sajib Chakraborty1 , M. Abdur Razzak1 "Recognition and Detection of Tea Leaf's Diseases Using Support Vector Machine IEEE 2018.

[2]. Prabira Kumar Sethy#1, Baishalee Negi*2, Santi Kumari Behera#3,Nalini Kanta Barpanda#4, Amiya Kumar Rath , An Image Processing Approach for Detection, Quantification, and Identification of Plant Leaf Diseases -A Review IJET UAE 2017..

[3]. Vishal Mani Tiwari &Tarun Gupta , Plant Leaf Disease Analysis using Image Processing Technique with Modified SVM-CS Classifier IJEMT 2017.

[4]. M. Ahmed and M.S.A. Mamun,"Distrributional Pattern and SeaonalAbundance of Major PestsofTea in Bangladesh", Tea J. Bangladesh, Vol. 41, pp. 1-10, 2012, ISSN: 0253-5483.

[5]. A.F.M.B. Alam, "Profile of tea industry in Bangladesh" pp.1-22 (Jain NK ed. 1999), Global Advances in Tea Science. New Delhi: Aravali Books. 882pp. Chen Z and Chen X. 1989. An analysis of the world tea fauna. J. Tea Sci. Vol. 9,pp. 13–22.

[6]. Tea Productions, Consumptions and Exports: Bangladesh Perspective Tasnuba Nasir and Mohammad Shamsuddoha Faculty of Business Administration (FBA), University of Science and Technology Chittagong (USTC), 2003.

[7]. Harler, C.R. 1964. The Cultivation and Marketing of Tea. 3d. Ed. Oxford: Oxford University Press.

[8]. Kholis Majid, Yeni Herdiyeni, Annu Rauf, "I-Pedia: Mobile Application For Paddy Disease Identification Using Fuzzy Entropy And Probabilistic Neural Network", ICACSIS, 2013.

[9]. Qing Yao;Zexin Guan;Yingfeng Zhou;Jian Tang; Yang Hu;Baojun Yang. Application of support vector machine for detecting rice diseases using shape and color texture features. International Conference on Engineering Computation, 2009, pp.79-83.

[10]. Elham Omrani, Benyamin Khoshnevisan, Shahaboddin Shamshirband, Hadi Saboohi, Nor Badrul Anuar, Mohd Hairul Nizam Md Nasir, "Potential of radial basis function-based support vector regression for apple disease detection", Measurement, Vol.55, pp.512–519, Sept-2014, (Elsevier).