Research Article

Disease Detection and Classification inCotton Plants using Unsupervised Learning-based Color and Texture Feature Extraction

Anil Moguram[1], Madiapalli Sumalatha[2], Gattu Sandeep[3]
Associate Professor, Dept. of ECE
Siddhartha Institute of Technology and Sciences, Narapally, Hyderabad, Telangana

ABSTRACT

In this, we have used SVM classifier to identify the pest and type of disease in cotton plant. Image acquisition devices are used to acquire images of plantations at regular intervals. These images are then subjected to pre-processing using median filtering technique. The pre-processed leaf images are then segmented using K-means clustering method. Then the color features(mean, skewness), texture features such as energy, entropy, correlation, contrast, edges are extracted from diseased leaf image using gray scale matrix (GSM) in the texture and then compared with normal cotton leaf image.

1. INTRODUCTION

Agriculture is one of the most important sources for human sustenance on Earth. Not only does it provide the much necessary food for human existence and consumption but also plays a major vital role in the economy of the country. But Plant diseases have turned into a dilemma as it can cause significant reduction in both quality and quantity of agricultural products. Nowadays farmers are facing many crucial problems for getting better yield cause of rapid change in climate and unexpected level of insects, in order to get better yield, need to reduce the level of pest insect. Several millions of dollars are spent worldwide for the safety of crops, agricultural produce and good, healthy yield.

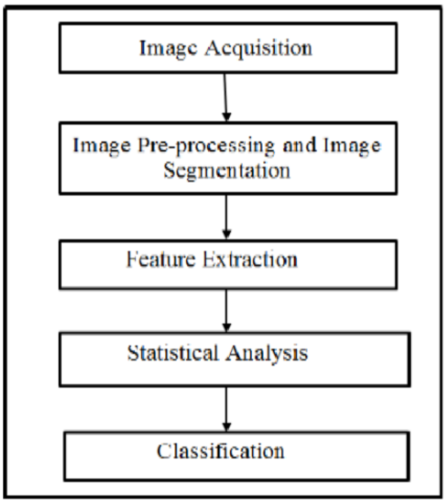

Fig. 1 General flow chart of disease identification

It is a matter of concern to safeguard crops from Bio-aggressors such as pests and insects, which otherwise lead to widespread damage and loss of crops. In a country such as India, approximately 18% of crop yield is lost due to pest attacks every year which is valued around 90,000 million rupees. Conventionally, manual pest monitoring techniques, sticky traps, black light traps are being utilized for pest monitoring and detection in farms. Manual pest monitoring techniques are time consuming and subjective to the availability of a human expert to detect the same. Disease is caused by pathogen which is any agent causing disease. In most of the cases pests or diseases are seen on the leaves or stems of the plant. Therefore, identification of plants, leaves, stems and finding out the pest or diseases, percentage of the pest or disease incidence, symptoms of the pest or disease attack, plays a key role in successful cultivation of crops. In general, there are two types of factors which can bring death and destruction to plants; living (biotic) and nonliving (abiotic) agents. Living agent's including insects, bacteria, fungi and viruses. Nonliving agents include extremes of temperature, excess moisture, poor light, insufficient nutrients, and poor soil pH and air pollutants.

2. RELATED WORK

S. Ananthiet al. proposed a computer aided solution for detection of plant leaf disease using texture statistics and clustering procedure [1]. S. B. Dhaygude et al. presented disease detection in plant leaf using color image segmentation and spatial gray level dependence matrices where they utilizedHSI color space for transforming the input RGB image [3]. Author in [4] addressed the identification of leaf diseases in pepper plants, where they utilized some soft computing approaches. Further, the provided the review of various existing disease detection approaches. Recently, a fuzzy clustering approach is utilized to identify the disease affected plant leaves in earlier stage [5]. In addition, authors employed CIELuv color space for enhanced detection of disease affected area.Poonkuntran et al. [6] proposed an approach which identify that whether there is any disease effected leaf or not by employing gray thresholding algorithm. Authors in [7, 10] utilized image classification techniques to detect and classify illness of plant for agricultural applications. An optimization-based segmentation approach with color histogram is employed in [8]. However, recent advances disclosed that machine learning based approaches providing the accurate detection and classification of diseases in plant leaves in spite of higher computational complexity [9, 11].

3. PROPOSED METHOD

The RGB color images of most frequently encountered Phyto-pathological problems affecting Cotton leaves were captured using camera. Images were stored in.JPG format.

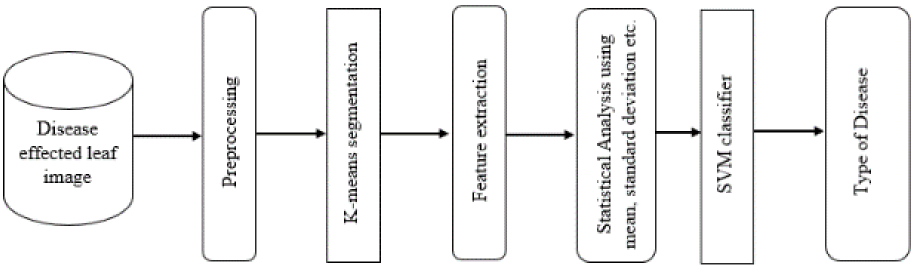

Fig. 1 Proposed framework for disease detection and classification

A. Pre-Processing and Segmentation

The pre-processing involved the procedures to prepare the images for subsequent analysis. The affected leaf images were converted from RGB color format to gray scale images. Segmentation refers to the process of clustering the pixels with certain properties into salient regions and these regions correspond to different faces, things or natural parts of the things. We proposed k-means segmentation technique to fragment goal areas. Target regions are those areas in the image that represented visual symptoms of a fungal disease.

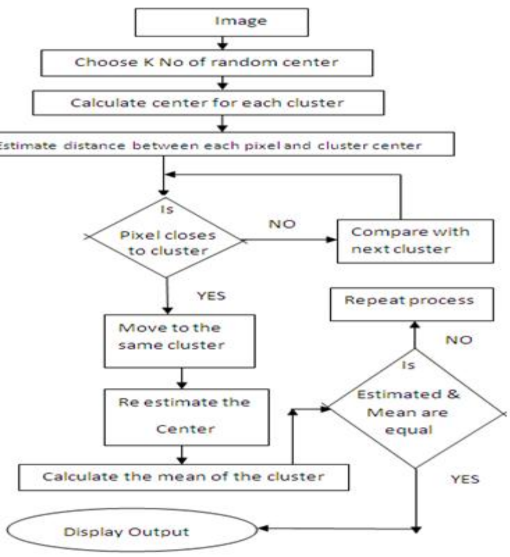

Fig. 2 Flowchart of K-means algorithm

B. Feature Extraction

The symptoms associated with various Phyto-pathological problems of cotton leaves under investigation visible on the affected leaves were extracted from their respective images using K-means. The image analysis was mainly focusing on the extraction of shape features and their color-based segmentation. The image analysis technique is done using Gray-level co-occurrence matrix. The affected areas vary in color and texture and are dominant in classifying disease symptoms. So, we have considered both color and texture features for recognition and classification purpose. The use of color features in the noticeable light spectrum provided additional image characteristic features over traditional gray-scale representation. GLCM is a method in which both color and texture features are considered to arrive at unique features which represent that image.

C. Statistical Analysis

Statistical analysis tasks are completed to choose the best features that represent the given image, thus minimizing feature redundancy. We have found that only 13 features contribute as discriminating features as this is essential for better classification. Magnitudes that are workable to guess via the co-occurrence matrix are: energy, entropy, homogeneity, contrast, Mean, Standard Deviation, RMS, Variance, Smoothness, Kurtosis, Skewness, IDM and correlation.

D. Classification

At present SVM is popular classification tool used for pattern recognition and other classification purposes. Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification or regression. The normal SVM classifier takes the set of involvement data and calculates to classify them in one of the only two separate classes. SVM classifier is trained by a given set of training data and a model is willing to classify test data established upon this model. Most habitual classification models are established on the empirical risk minimization principle. SVM implements the structural risk minimization principle which pursues to reduce the training error and a sureness interval term. A number of submissions showed that SVM hold the superior classification capability in production with minor sample, nonlinearity and high dimensionality pattern identification. Support Vector Machines are based on the concept of decision planes that define decision boundaries. A decision plane is one that splits among a set of objects having different class association. Classifier that separate a set of objects into their corresponding classes with a line. Supreme classification tasks, however, are not that modest, and regularly more difficult structures are needed in order to make an optimal separation, i.e., correctly classify new objects (test cases) on the basis of the examples that are available (train cases). All the evidence from beyond processes is given to multiclass SVM .The Multiclass SVM were used for cotton disease classification.

4. SIMUALTION RESULTS

Identified Phyto-pathological problems experiments modules are developed using MATLAB R2016b, which runs in the environment of Windows7, 8 and 10. Two species of samples are taken for the experiment, whose digital images are obtained by a camera. Figure 3 shows the dataset images utilized for testing the proposed algorithm.

Fig. 3 dataset used for disease identification

Fig. 4 Disease affected leaf from the dataset

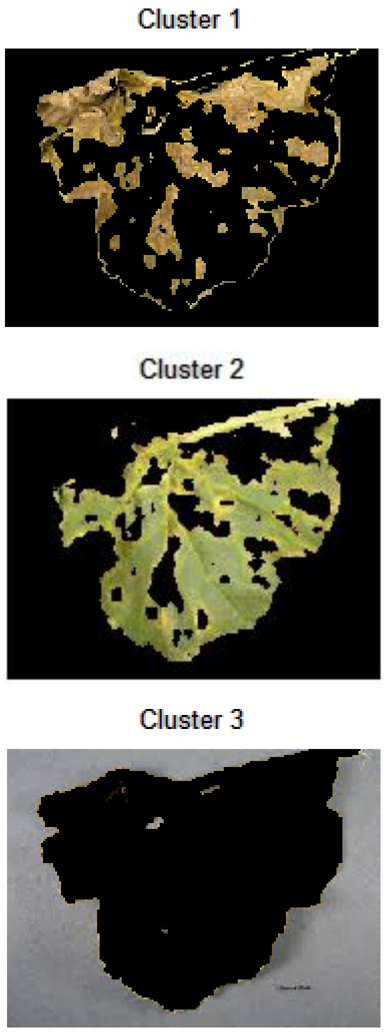

Fig. 5 Cluster indexes after K-means segmentation process

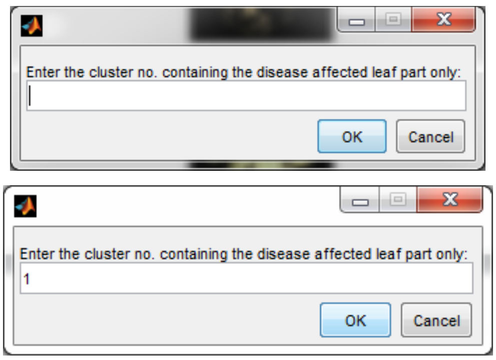

Fig. 6 select the cluster index in which the disease is presented

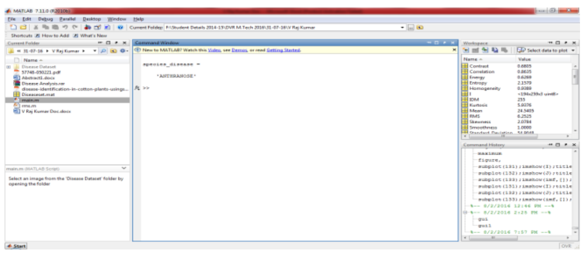

Fig. 7 snapshot of MATLAB environment after the execution of program

5. CONCLUSION

In this paper, we described our work concerned with the discrimination between healthy and diseased to cotton crops using an SVM. In this paper, respectively, the applications of K-means clustering have been formulated for clustering and classification of diseases that effect on plant leaves. Identifying the disease is generally the drive of the proposed method. Thus, the proposed process was tested on 2 diseases which influence on the plants; they are: Leaf spot and Leaf miner. These features are very important for the color and morphology of the leaf spots and they provide critical information about its visual representation. By using segmentation technique, it is easy for us to extract the features of disease leaf of the image. A new approach based on K-mean features extraction was proposed for cotton leaf recognition in this work. The whole process of leaf classification can be implemented using leaf detection, feature extraction and classification. For studying the proposed method, the composed dataset is used. The dataset contains diseased images. Images were preprocessed and cropped to a fixed standard size. Then, features are extracted from all the leaf images in the dataset using K-mean algorithm. For each image leaf more frequent K-mean key points are extracted to identify a unique feature. It permits finding related features for different image. Ultimately, the extracted K-mean features are rendered to a SVM classifier for purpose of classification. There are distinguishing differences between diseased and non-diseased leaf in structure, color, size etc. Therefore, identification is based on these differences. In other words, differences between diseased and non-diseased leaves and the key points which are extracted from leaf are used for classifying. The different method is performed, and the dataset was divided randomly in two parts, 70% for train and 30% for testing.

For future study, different neural network architectures can be used for classification. We can extend this project to classify disease symptoms affected on fruits, vegetables, commercial crops etc., we may work for better application like we develop a site where any person can upload their image, they will find out there diseased and full detail about the disease. What they do for their fields and crops. What is the advantage and disadvantage of this disease and what should do to control it.

REFERENCES

[1] S. Ananthi and S. V. Varthini “Detection and Classification of Plant Leaf Diseases”, International Journal of Research in Engineering andApplied Sciences, vol 2, no. 2, 2012.
[2] R. C. Gonzalez and R. E. Woods, “Digital Image Processing”, Second Edition, Pearson Education.
[3] S. B. Dhaygude, Nitin and P. Kumbhar “Agricultural plant Leaf Disease Detection Using Image Processing”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 1, Jan. 2013.
[4] J. Francis, A. S. DhasD. and B. K. Anoop, “Identification of leaf diseases in pepper plants using soft computing techniques”, IEEE Conference on Emerging Devices and Smart Systems, Namakkal, India, Oct. 2016.
[5] P. Ganesan, G. Sajiv and L. M. Leo, “CIELuv color space for identification and segmentation of disease affected plant leaves using fuzzy based approach”, 3rd International Conference on Science Technology Engineering and Management, Chennai, India, Jan. 2018.
[6] S. Poonkuntran, M. Kamatchidevi, L. S. Poornima and R. Shreeja, “Plant Disease Identification System”, International Research Journal of Engineering and Technology, Vol. 5, No. 3, pp. 2245-2250, Mar. 2018.
[7] A. G. Abdul Wahid, B. Jaison, P. Arun and S. Harish Kuma, “Plant Disease Identification Using Image Classification Techniques”, International Journal of Engineering Science and Computing, vol. 8, no. 4, pp. 16795-16796, Apr. 2018.
[8] M. Mathumitha, K. Jayashree, B. Baron sam and A. Pio sajin, “Survey on disease identification in plant leaves using optimization and segmentation”, International Journal of Pure and Applied Mathematics, vol. 118, no. 24, pp. 1-9, 2018.
[9] S. S. Chouhan, A. Kaul and U. P. Singh, “Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology”, IEEE Access, vol. 6, pp. 8852-8863, Feb. 2018.
[10] A. Devaraj, K. Rathan, S. Jaahnavi and K. Indra, “Identification of Plant Disease using Image Processing Technique”, International Conference on Signal Processing and Communication, Chennai, India, Apr. 2019.
[11] S. Jyothi, R. Patil and A. Pooja, “A survey paper on plant disease identification using machine learning approach”, 3rd International Conference on Communication and Electronic Systems, Coimbatore, India, May 2019.