The Classification and Feature Extraction Technique for Plant Disease Detection

Satjot Kaur, Kiranpreet Kaur

Abstract

The image processing is the technique which is applied to process the digital information from the images. The plant and the crops are ruining because of the excessive use of fertilizers and insecticides. The plant disease detection is the technique which is applied to detect disease from the input images. In this work, technique is applied which is based on textural feature extraction, segmentation and classification. The GLCM algorithm is applied which extract textural features from the image. The k-mean clustering algorithm is applied which segment input image. The SVM classifier is applied in the existing algorithm which will classify the input image into two classes. In the proposed system the GLCM algorithm will be replaced with law textural algorithm for the feature extraction. The multi-class SVM will be replaced with the naive bayes which improve accuracy of plant disease detection. The proposed and existing models are implemented in MATLAB. The results show that proposed model has high accuracy as compared to existing model.

KEYWORDS:

GLCM, Law textural, Multi-class SVM, Naive Bayes, K-mean.

Introduction

Image processing is the technique used for the conversion of the image in digital form and which is used to perform some mathematical operation. This process is used get a good quality of images and extracts some beneficial information from that image. It is the type of process in which image works as the input and the characteristics, features acts as an output of that image. Sometimes, image processing analyzes the images on two-dimensional signals and implements already set signal processing methods [1]. The signals involve the transmission signals, sound signals or voice signals in which the image acts as both the input as well as the output. The processing deals with all the processing of images. It is one of the most rapidly growing and developing technology. It is widely spread throughout the industries and large amount of research is still going on. Digital image processing has several applications and has great importance in day to day life. It has great advancement in the improvement of the pictorial information for human analysis and processing of a scene data for an independent machine observation [2]. It has lot of applications like remote sensing, image and data storage in industrial fields, medical imaging, acoustic imaging, forensic science and many more. The images captured through satellite are used to track earth resources, geographical mapping, and analysis of agricultural crops, calculation of urban population, weather forecasting, flood and other natural disaster [3]. It has another application that is space imaging applications which involve recognition and study of objects that consists of images produced from the space-probe mission. It plays very significant role in the medical applications like processing of X-rays, ultrasonic scanning, electron micrographs, MRI, NMR etc. India is an agricultural country as about 70% of the population earning depends on the agriculture. The agriculture is widely spread and the entire farmers' families relay on their lands and crops for their economic growth. But the plants and the crops are affected by the disease which leads to the reduced amount of quality as well as quantity of the agricultural products. The experts observe the plant disease with their naked eye and detect the type of diseases plant is suffering from [4]. In order to perform this observation, teams with large number of experts are made and they continuously monitor the plant. This procedure is quite expensive when the farms are very large. Sometimes, the farmers do not have the developed and advanced technology, which leads to the delay of disease detection. There are some basic steps to detect the diseases in plants. The first step is the image acquisition, in which the images are captured from the environment by using digital camera. The second step is the image pre-processing, in which the features are extracted from the acquired images for further analysis. After, this many analytical differentiating techniques are employed for the classification of images as per the specific diseases [5]. The plant leaf image is captured using camera and the images are in Red, Green and Blue color form. It creates a color transform of the RGB leaf image and after this

device-independent space transformation is implemented. In order to remove noises from the images various pre-processing techniques are introduced. Image clipping which involves cropping of the leaf images to have a desired image region. Smoothing filter is used for the image smoothing. Image enhancement is responsible for increasing the contrast. Segmentation refers to the partitioning of the images into different parts having same features and characteristics. It can be done using different methods like otsu method, k-means clustering, conversion of RGB images [6]. Feature Extraction plays a very significant role in the detection of the objects. It is used in many applications of the image processing; it includes color, texture, morphology, edges and so on. These are the features of the plant diseases detection. The morphological features are better than the other characteristics. Here, texture signifies the color of the images, also the roughness and hardness of the images. It is also used for the identification of the infected plants and the crops. The last stage of the plant disease detection is the classification of the plants according to the diseases. For this a deep learning algorithm is implemented which is used to classify the specific images into particular diseases. It makes easy to detect the diseases and find the cure for the infected plants. It determines the relevant count of the pixels by comparing the images with data sets [7]. Clustering is the process of dividing data into large number of groups. It is one of the most common classifier, being used by the researcher. It divides the collected data into k-numbers of groups. It further classifies the provided set of data into k-number of disjoint cluster. It contains two separate phases. In first phase, k-centroid values are calculated and in the second phase, every point of cluster is transferred nearest centroid from the respective data point. There are certain methods to calculate the distance of the nearest centroid and most common is the Euclidean distance formula. After grouping, it again calculates the new centroid of each cluster and according to that cluster, a new Euclidean distance is evaluated between center and data point [8]. This assigns the points within the cluster having Euclidean distance. Every cluster is defined by its member objects and also by its centroid point. The centroid is the point of each cluster which represents the sum of distances from all the objects. Hence, it is an iterative algorithm which minimizes the sum of distances from object to its cluster centroid.

Literature Review

M S et al. [9] adopted an approach for the detection and classification of the plant diseases called naked eye observation done by the experts. The decision whether the plant is infected or not is completely depends on the physical conditions of the observer like fatigue, eye sight, work pressure and climatic conditions. In this paper, the symptoms of the diseases are also described. The proposed system is also used to classify and differentiate between the healthy and unhealthy plants. The proposed system reduces the monitoring task of the farmers. The farmers can now easily predict and diagnose the diseases at the very initial stage. The results show that the proposed system is efficient in recognition and identification of the leaf diseases. Also, the diseases can be predicted at early stage and improves the recognition using Artificial Neural Network, bayes classifier and fuzzy logic.

Abed et al. [10] focused on the detection of the two types of diseases caused on the bean leaf i.e. bacterial brown spot and powdery mildew. The procedure of detection includes acquisition, preprocessing, segmentation, feature extraction, and classification. The images are captured for the public database. This methodology successfully detects the diseases at their early stages with an accuracy of 100 %. This proposed system improves its performance by selecting the cluster automatically. In this system the user have to enter inside the infected cluster of the leaves and plant because the value of clustering changes when the testing images are entered in the system. Therefore, the proposed system shows effective and accurate results and detects and classifies the diseases.

Hossain et al. [11] purposed a Support Vector Machine classifier which is used to detect diseases in plants and crops. The author had studied eleven characteristics of this technique. These features are further finds the appropriate match for the diseases whenever an image is uploaded in the system and the diseases is identified. The research is novel as the number of features reduced after compared by the SVM classifier and it attains the accuracy of about 90%. It increases the identification process due to which every leaf require 300ms less processing time in comparison to the previous SVM technique. The proposed technique ensures that the large number of leaves can be processed in a given time period. The researchers concluded that the proposed system

increases the efficiency of the detection, determination, identification and classification by 93% as compared to the other classifiers. This technique also reduces the extract feature which in turn reduces the processing time.

Khan et al. [12] describes the machine vision framework and identifies the manifestation of plant diseases and analyzes the images in CIELab. The main objective of this purpose is to create a procedure for the detection of plant diseases by using cascading unsupervised image segmentation approaches. Additionally, RGB color model for digital images and CIELab color model for performing pre-processing step which increases the each channel timing. The researchers also introduced multilevel segmentation technique which uses expectation maximization with minimum constraint visual information loss. Various experiments were conducted and indicate that the new cascaded design outcomes a superior color segmentation with the confirmation of infected regions.

Mattihalli et al. [13] presented a methodology for the early detection of the leaf diseases, this approach works according to the some extracted features from the images of the images. The approach contains a device called Beagle Bone black; this device is linked digital camera or web camera and used to detect the diseases in leaf. The images of the leaves are captured and then compared with images of the healthy leaves. These images of healthy leaves are stored in the database. Once the image processing is done, the plants with some diseases are turn on the valves by which medicine is automatically supplied the infected leaves. The researcher concluded that the proposed system is effective, less expensive and user friendly.

Jaskaran Singh et al. [14] reviewed various techniques for the detection of plant disease in terms of different parameters. Diseases are identified by using different image processing methods. The image processing is refers as the technique which processes the digital information stored in the images format. On the other hand plant diseases detection is the mechanism having three stages: first, the images are fed into the terminal for pre-processing, and then the images are analyzed according to the segmentation features and at last the image is classified using any of the desired classifier.

Research Methodology

The plant disease detection is the technique which is applied to detect the disease portion from the input leaf image of the plant. The plant disease detection technique is been proposed in this work, which is based on the following steps:-

1. **Pre-processing phase:** In the first phase, the plant leaf image is given which is converted to grayscale. The GLCM algorithm is been applied which will extract the textural features of the input image.

1.1. Laws Textural Method: Feature extraction is the process of obtaining higher level information of an image such as color, shape, texture. It is a method to extract secondary features of the image. This method combines predetermined one-dimensional kernels into various convolution masks. The Laws method uses filter masks to extract secondary features from natural micro-structure characteristics of the image (level, edge, spot, wave and ripple) which can then be used for segmentation or classification. It uses a set of 5x5 convolution masks which is used to compute the texture energy. This texture energy then represented as a vector of nine numbers for each pixel of the image being analyzed.

2. Image Segmentation: The Image segmentation technique is applied which will segment the image on the basis of their properties. The image segmentation techniques are generally applied categorized into region based segmentation and threshold based segmentation. In this work, the region based k-mean segmentations technique is applied for the image segmentation.

3. Classification of Data: The Naïve Bayes is the classification technique which is applied to classify similar and dissimilar data into more than one classes. Naïve Bayes classifiers depend on learning by analogy. The Naive Bayesian classifier is based on Bayes theorem with the independence assumptions between predictors. The representations for naive Bayes is probabilities. Lists of probabilities are stored to file for a learned naive Bayes model. Class probabilities are the probabilities of each class in the training dataset. Conditional probabilities are the conditional probabilities of each input value given each class value.

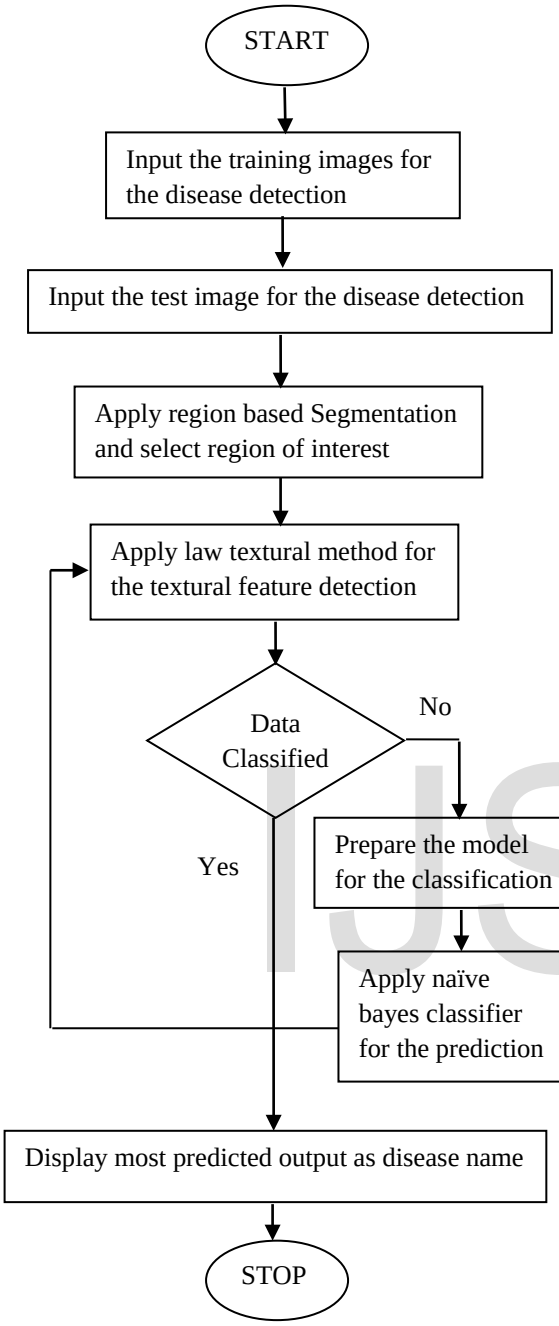

Fig 1: Proposed Flowchart

Experimental Results

The proposed research is implemented in MATLAB and the results are evaluated by making comparisons against proposed and existing approach in terms of existing parameters. The dataset of APS is used for the simulation. The American Phytopathological Society processes thousands of scientifically peer-reviewed images showing disease symptoms, pests, associated with plants and crops through its book- and journal-publishing programs. The APS Image Database is available by personal subscription access. The APS Image Database is curated by the APS PRESS Editorial Board and is expected to build into tens of thousands of disease, pest, and disorder images online.

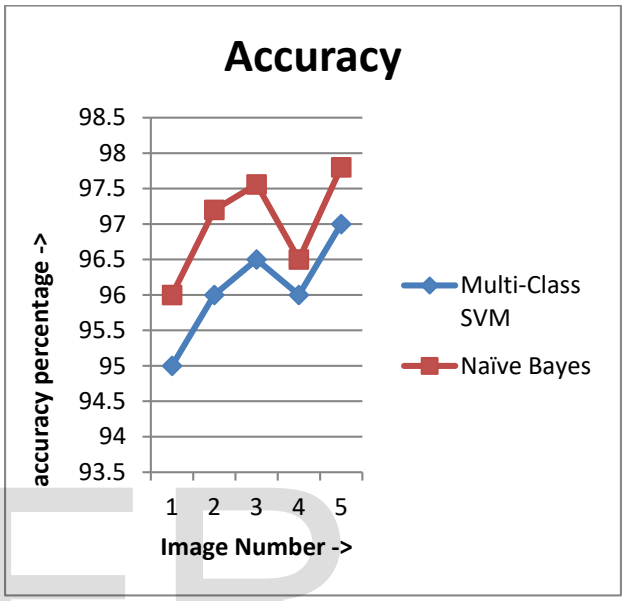

Fig 2: Accuracy Comparison

As shown in figure 2, the accuracy of the proposed technique is enhanced as compared to the existing technique. The accuracy of existing technique is 95 percent and accuracy of the proposed technique is approx 97 percent.

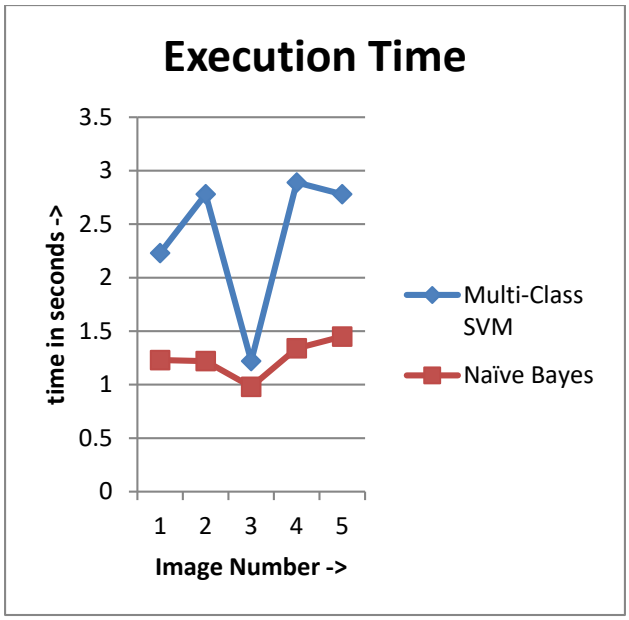

Fig 3: Execution Time Comparison

As shown in figure 3, the execution time of the proposed and existing algorithm is compared for the performance analysis. It is analyzed that proposed algorithm has less execution time as compared to existing algorithm.

Conclusion

In this work, it is been concluded that plant disease detection required three main steps which are feature extraction, segmentation and classification. In the existing the k-mean clustering is applied to segment input images. The GLCM algorithm is applied to extract textural features of the input image. The proposed work utilizes Laws Textural algorithm which extracts more features than GLCM algorithm and provides a better system to detect diseases in plant. The multi-class SVM will be replaced with the naive bayes classifier for the final disease predication. The proposed and existing algorithms are implemented in MATLAB. This novel approach will provide farmers with a more accurate and time saving mechanism for disease identification. Thus, the proposed approach presents a path toward automated plant diseases diagnosis on a massive scale. The results of the proposed model are analyzed in terms of accuracy, precision, recall and execution time. It is analyzed that proposed model performs well as compared to existing in terms of define parameters.

References

[1] Savita N. Ghaiwat, Parul Arora "Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review", International Journal of Recent Advances in Engineering & Technology, ISSN (Online): 2347 - 2812, Volume-2, Issue - 3, 2014

[2] S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features" Agric Eng Int: CIGR Journal, 15(1): 211-217 2013

[3] Prof Sanjay B. Dhaygude Mr. Nitin 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 January 2013.

[4] Smita Naikwadi, Niket Amoda" ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES" , International Journal of Application or Innovation in Engineering & Management , Volume 2, Issue 11, November 2013

[5] Sanjay B. Patil et al. "LEAF DISEASE SEVERITY MEASUREMENT USING IMAGE PROCESSING", International Journal of Engineering and Technology Vol.3 (5), 2011, 297-301

[6] B. Bhanu, J. Peng. "Adaptive integrated image segmentation and object recognition", In IEEE Transactions on Systems, Man and Cy- bernetics, Part C, volume 30, pages 427–441, November 2000.

[7] Keri Woods. "Genetic Algorithms: Colour Image Segmentation Literature Review", July 24, 2007.

[8] S.Beucher, F.Meyer. "The morphological approach to segmentation: The watershed transform", in Mathematical Morphology Image Processing, E. R. Dougherty, Ed. New York Marcel Dekker, 1993, vol. 12, pp. 433–481.

[9] Arya M S, Anjali K and Mrs.Divya Unni, "DETECTION OF UNHEALTHY PLANT LEAVES USING IMAGE PROCESSING AND GENETIC ALGORITHM WITH ARDUINO," 2018, IEEE

[10] Sa'ed Abed and Anwar Ali Esmaeel, "A Novel Approach to Classify and Detect Bean Diseases based on Image Processing," 2018, IEEE

[11] Md. Selim Hossain1 , Rokeya Mumtahana Mou1 , Mohammed Mahedi Hasan2 , Sajib Chakraborty1 , M. Abdur Razzak, "Recognition and Detection of Tea Leaf's Diseases Using Support Vector Machine," 2018 IEEE 14th International Colloquium on Signal Processing & its Applications (CSPA 2018), 9 -10 March 2018

[12] Zia Ullah Khan1,2, Tallha Akram2 , Syed Rameez Naqvi2 , Sajjad Ali Haider2 , Muhammad Kamran2 , Nazeer Muhammad, "Automatic Detection of Plant Diseases; Utilizing an Unsupervised Cascaded Design," 2108, IEEE

[13] Channamallikarjuna Mattihalli 2 Edemialem Gedefaye 3 Fasil Endalamaw 4 Adugna Necho, "Real time Automation of Agriculture Land, by Automatically Detecting Plant Leaf Diseases and Auto Medicine," 2018 32nd International Conference on Advanced Information Networking and Applications Workshops

[14] Jaskaran Singh, Harpreet Kaur, "A Review on: Various Techniques of Plant Leaf Disease Detection," Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018) IEEE Xplore