2011 International Conference on Advancements in Information Technology
With workshop of ICBMG 2011
IPCSIT vol.20 (2011) © (2011) IACSIT Press, Singapore

An Application of K-Means Clustering and Artificial Intelligence in Pattern Recognition for Crop Diseases

Mrunalini R. Badnakhe [1+] and Prashant R. Deshmukh[2]

[1] Lecturer, Dept of IT, Dr. Babasaheb Ambedkar College of Engineering & Research, Amravati (MS) INDIA
[2] Professor, Dept of CSE and IT, SIPNA's College of Engineering & Technology, Amravati (MS) INDIA

Abstract. Crops are very important for the farmers, just like the bread and butter for them. The mainly Indian Economy is depending on the Agricultural productivity. So through this research we can directly help the farmer. By using the automated agricultural inspection, Farmer can give potentially better and accurate productivity .The different products can be yield with better quality. The primary need for the agriculture to predict which the infected crop is. With the help of this work we are indirectly contributing for the Improvement of the Crop Quality. It is a Machine learning based recognition system which will going to help in the Indian Economy. The paper will propose the technique to classify and identify the different disease affected plant.

Digital Analysis of crop color is the important. Now it's becoming popular day by day. It is also of the cost effective method. Because changed in the color are a valuable indicator of crop health and efficiency and survaibility. Then it can be measured with visual scales and inexpensive crop color.

Keywords: Digital Analysis, Crop quality**,** surviability Visual scale.

1. Introduction

The change in the crop color is the important aspect for the notification. When the health of the crop is at good stage then the color of the crop is different but as soon as the crop is affected by some harming pathogens, the color changes automatically. Crop diseases have turned into dilemma because it may cause reduction in productivity. Generally through the naked eyes the observations taken by the Experts ancient time for the detection and identification of crop diseases. But for this the continuous monitoring is required by the Experts and It is too expensive in large fields. So in many under developed countries in agricultural area, farmer needs to take lots of efforts. This work will described that how can we do the automatic detection of Crop diseases as this can gives much benefits in monitoring large fields of crops and detect the symptoms of diseases. Again we can tell the future preventions and treatment for the infected crop. Here in this way want to look for the Fast, Automatic, Less Expensive and Accurate method to detect, classify, identify the crop diseases.

If we are using the machine vision then the identification and the classification of the plant will be done faster at every stage. It affects the performance of the classifier. This is the best application of the digital image processing in Agricultural field. This Study will again explore the image processing technique that visually differentiates between various Crop diseases which are found in different parts of crop.

2. Literature Review & Related Work

[+] Corresponding author. Tel.: + 07212511418.
E-mail address: mrunal.3878@gmail.com.

Basically, for any research work, literatures help the researchers to motivate for the innovation of a new concept for the successful result. Many research works have been published regarding the advancements of image processing for feature extraction and classification. Some researchers has suggest that a fast and accurate new method is developed based on computer image processing for grading of plant diseases. The result gives the technique for detection of plant diseases. The application of K- means clustering and neural network has been formulated for the clustering and classification of disease that affect a plant leaves. This work has been done for the five diseases [1]. Many suggest that Technological strategies using machine vision and artificial intelligence are being investigated to achieve intelligent farming, include early detection of diseases in groves, selective fungicide application. In this work, they have concluded that a study was completed to investigate the use of computer vision for classifying citrus leaf diseases. Algorithms based on image-processing techniques, for feature extraction and classification, were designed. The feature extraction process used color co-occurrence methodology, which uses both the color and texture of an image to arrive at unique features, which represent that image [2]. Some has told during their research work that Changes in foliar color are a precious indicator of plant nutrition and health. Leaf color is measured with visual scales and inexpensive plant color guides that are easy to use Managers and scientists for evaluating foliar nutrition and health in response to environmental stresses. They analyses with sugar maple leaves indicate that digital image analysis provides an accurate means of quantifying foliar color and estimating pigment concentration in multicolored leaves [3].Some of researcher has described their many related things on cotton. The goal of this research is to develop an automated system that can identify the pest or disease affecting a cotton leaf, boll or flower by using image analysis. It helps the farmer to bring in a damaged leaf, boll or flower of a cotton plant. This paper presents a CMYK based image cleaning technique to remove shadows, hands and other impurities from images. The images are subsequently classified using two indigenous techniques RPM and Dis Bin and compared with the classical PCA based technique. The results are tested over a database consisting of 600 images. After the cleaning of the images three techniques have been applied to classify a presented image as a leaf, boll or flower.[4] In next paper, Cotton often suffers from various diseases that can adversely productivity to varying degrees .Proper disease control measures must be undertaken in cotton to minimize losses .Currently, the machine vision and artificial intelligence are being studied to achieve intelligent farming. Early identification of disease symptoms becomes an important aspect of crop disease control. In some cases, disease control actions or remedial measures can be undertaken if the symptoms are identified early. Lastly they concluded that the proposed feature selection approach can automatically and quickly identify independent significant feature for a pattern recognition system. Fuzzy curves can quickly isolate important and significant features subset. Fuzzy surfaces eliminate those features which are dependent on other important features [5].These are the literature survey regarding the digital image processing. By taking the help of this type of researches the new research work is emerging in Image processing

3. Analysis of Problem

Now-a-days, in modern agricultural system, numerous computational methods have been developed to help farmers or agronomists to monitor the proper growth of their crops. In our ancient agricultural system, during harvesting process of the crops, the naked eye observation of farmers or experts is the main approach adopted in practice for detection and identification of crop diseases under microscopic conditions in laboratory. However, this requires continuous monitoring of experts which might be prohibitively expensive in large farms. Further, in some developing countries, farmers may have to go long distances to contact experts, this makes consulting experts too expensive and time consuming. The basic problems regarding with crop is on the field, a fast and accurate recognition and classification of the diseases is required by inspecting the infected leaf images also recognize the severity of the diseases. There are two main characteristics of plant-disease detection machine-learning methods that must be achieved, they are: speed and accuracy. The damage diagnosis of different crops and vegetables has traditionally been done manually. Several efforts have been made to use image processing systems to automate the process through this research work.

3.1. Proposed Work

For a fast and accurate recognition and classification of the diseases infected Crop. It's necessary to work on this concept. The overall concept that is the framework for this vision related algorithm of image classification is as follow. The processing will be done through different steps. Because when any work will be done then at that time it must be divided into different parts. The every part must be self explaining same as the proposed work is divided into subparts or we can say that in the different modules.

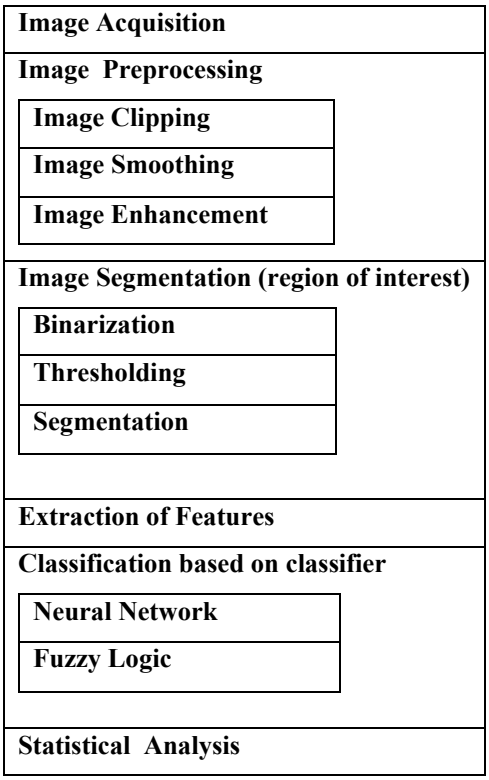

Fig. 1: Digital Image Processing

3.1.1. Image Acquisition: The digital images are acquired from the environment referring different sites. Image pre-processing is a prophase relative to feature extraction and image recognizing. Many way to collect the images means we can visit the agricultural research units. In this research work the help of the research centers is taking which is working in agriculture department.

3.1.2. Image Preprocessing: Image pre-processing is a prophase relative to feature extraction and image recognizing. The images which have input are always not satisfactory regardless of what image acquisition devices are adopted. For e.g., there are noises in the image, the region of interest in the image is not clear or other objects' interference exist in the image and so on. Different pre-processing methods should be chosen for different image applications.

There are three steps included in pre-processing phase: clipping, smoothing and enhancement.

- a. Image clipping: The images of crop leaf were taken with camera, always with complex background. But there was not a segmentation method which can get the region of interest accurately from an image with complete background. With the tool of VC, all the closed areas were clipped down from an image and the regions of interest were gotten from background successfully.
- b. Image smoothing : Noises which may be brought from the process of image collection and lots of information which may be easily leaded from the operating and saving to the image would make the quality of image dropped, thereby affects following recognizing of disease. So the image with low quality must be smoothed. The filter can be used for smoothing.
- c. Image Enhancement: Image enhancement is one of the most interesting and important issues in digital image processing field. The main purpose of image enhancement is to bring out details that are hidden in an image, or to increase the contrast in a low contrast image .Image enhancement produces an output image that subjectively looks better than the original image by changing the pixel's intensity of the input image. Generally, image enhancement enlarges the intensity differences among

objects and background. There are many image enhancement techniques have been proposed and developed

3.1.3. Image Segmentation (Region of Interest) : The image will be segmented into different parts according to the region of interest. Purpose of image segmentation is to divide the image into some meaningful regions. Simply to say, image segmentation means to separate the object from background for following processing in an image. It contains Thresholding, Binarization, and Segmentation.

3.1.4. Extraction of Features: When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Classification based on a classifier: We are using Artificial Intelligence Techniques to solve the problem. We will going to classify it with the original existing work. Artificial Neural Network and Fuzzy Logic can be used to classify the crop diseases.

3.1.5. Statistical Analysis: By comparing the classification results of ANN and Fuzzy Logic Technique. These analyzes , which system is better in sense of Accuracy, Speed, User friendly, easily adaptable topology of the network changes, a new sequence number is necessary before the network re-converges; thus, DSDV between nodes by sending full dumps infrequently and smaller incremental updates more frequently. Whenever there is not suitable for highly dynamic networks. (As in all distance-vector protocols, this does not perturb traffic in regions of the network that are not concerned by the topology change.

Phase 1:-

There are two preprocessing steps important in order to implement the K-means clustering algorithm: The phase starts first by creating device-independent color space transformation structure. We developed the color transformation structure that tells the color space conversion. After that the applied color transformation converts the color values in the image to the color space specified in the color transformation structure. The color transformation structure specifies various parameters of the transformation. A device dependent 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 the brightness and contrast on the display device used. Thus the RGB system is a color space that is dependent. The K-means clustering algorithm tries to classify objects 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. 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. Here multiple values of number of clusters have been tested.

Phase 2: **Masking the Green Pixels**.

This phase consists of two steps: The mostly green colored pixels are identified, and then the global image threshold using different methods has been applied in order to specify the varying threshold value which chooses the threshold to minimize the intra class variance. Next, the green pixels are masked as follows: if the green component of pixel intensities is less than the computed threshold value, then, the red, green and blue components of the this pixel are cleared. In the next step in this phase is focused on deleting both the pixels with zeros components and the pixels on the boundaries of the infected cluster.

Phase 3:- Features Extraction

In the proposed approach, the method adopted for extracting the feature set is called the Color Co-occurrence Method or CCM method in short. It is a method, in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image.

Phase 4:- **Neural Networks**

In this work, neural networks are used in the automatic detection of leaves diseases. Neural network is chosen as a classification tool due to its well known technique as a successful classifier for many real applications. The training and validation processes are among the important steps in developing an accurate process model using NNs. The dataset for training and validation processes consists of two parts; the training feature set which are used to train the NN model; whilst a testing features sets are used to verify the accuracy of the trained NN model.

3.2. Desired Applications

- For detecting crop diseases early and accurately. Here we are using image processing and classification techniques for detecting diseases of Agricultural or Economic crops.
- As agriculture is aimed towards increase of productivity and food quality at reduced expenditure and with yield profit. It is one of the time saving processes for the farmer.
- They can easily identify and get the different parameters related to the crop and its diseases.

4. Conclusion

After analyzing we recognized the different diseases from crop. This proposed work is giving of the better technique to do the classification of crop disease. We can easily develop an application. In future the experimental results indicate that the proposed approach is a valuable approach, which can significantly support an accurate detection of leaf, steam, and root diseases in a little computational effort.

5. References

[1] Fastand Accurate Detection and Classification of Plant Diseases" H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. A L Rahamneh.

[2] Identification of citrus disease using color texture features and discriminant analysis",R.Pydipati, T.F.Burks,W.S.Lee.

[3] An Instructional Guide for Leaf Color Analysis using Digital Imaging Software ", Paula F. Murakami, Michelle R. Turner, Abby K. van den Berg, Paul G. Schaberg.

[4] Automatic Cleansing and classification on Cotton leaves, bolls and flowers using CMYK color splitting ",Syed A. Hayat , Ahsan Abdullah, Muhammad, A.Chaudary , Yasir J. Malik and Waseem Gillani.

[5] Feature Selection of Cotton Diseases Leaves Image Based On Fuzzy Feature Selection Technique", Yan Cheng Zhang, Hang-Ping Mao, Bo Hu Ming-Xi Li.