International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 3, Special Issue 3 (September '2017)
Proceeding of - International Conference on Recent Trends in Computational Life Science and Information Technology,
Conference held at Tiruppur Kumaran College for Women,Tirupur,Tamilnadu, India.

AN IDENTIFICATION OF VARIETY OF LEAF DISEASES USING DATA MINING TECHNIQUES

S.Renuga,
M.Sc (Computer Science).,
PG and Research Department Of Computer Science,
Tiruppur Kumaran College for Women,
Tirupur,Tamilnadu,India.

Abstract: Diseases in plants cause major production and economic losses as well as reduction in both quality and quantity of farming products. Disease management is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Precise quantification of these visually observed diseases, pests, traits has not studied yet because of the complication of image patterns. Hence there has been increasing demand for more specific and sophisticated image pattern understanding. This work presents a method for identifying plant leaf disease based on color. Agrarians are suffering from the issue rising from different types of plant leaf diseases. Sometimes biologists are also unable to identify the disease that leads to need of identification of right type of disease. First the input image is pre-processed. Then input image of leaves is converted as Red Green Blue (RGB) to Hue Intensity Saturation (HIS) or Lab color space. Then leaf disease segmentation is done using Hierarchical clustering. After segmentation the mostly green color pixels are covered based on specific threshold values. The Support Vector Machine (SVM) and Neural Network (NN) is trained for classification. The goal of this study is to provide, different identification techniques for plant leaf. Some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e. when they appear on plant leaves.

Keywords: support vector machine, neural network, hue intensity saturation, disease, agrarians, clustering, classification.

I.INTRODUCTION:

India is an agricultural country where in most of the population depends on agriculture. Farming is one of the most important domains which decides economy of the nation. Farming creation system is an outcome of a complex interaction of soil, seed, and agro chemicals. Disease detection is a series task and mostly diseases are seen on the leaves or stems of the plant. Diseases are destruction to the normal state of the plant that modifies or interrupts its vital functions such as photosynthesis, transpiration, pollination, fertilization, germination etc. These diseases are caused by pathogens viz., fungi, bacteria and viruses, and due to difficult environmental conditions. Therefore, the early stage analysis of plant disease is an important task. Farmers have need of continuous monitoring of experts which might be prohibitively expensive and time consuming.

This research work is disclosure to automatic detection of disease on all variety of plant leaves.

II. TYPES OF DISEASES ON PLANT LEAVES

The diseases on the leaves are classified as,

a) **Viral disease**: e.g. Leaf Curl, Leaf Crumple, Leaf Roll.
b) **Fungal diseases**: e.g. Anthracnose, Black Spot.
c) **Bacterial disease**: e.g. Bacterial Blight, Crown Gall, Lint Degradation.

2.1 VIRAL DISEASE SYMPTOMS:

Among all plant leaf diseases, those caused by viruses are the most challenge to diagnose. Viruses produce no telltale signs that can be readily observed and often easily confused with nutrient deficiencies and herbicide injury. Aphids, leafhoppers, white flies and cucumber beetles insects are common carriers of this disease, e.g. Mosaic Virus, Look for yellow or green stripes or spots on foliage. Leaves might be wrinkled, curled and growth may be stunted.

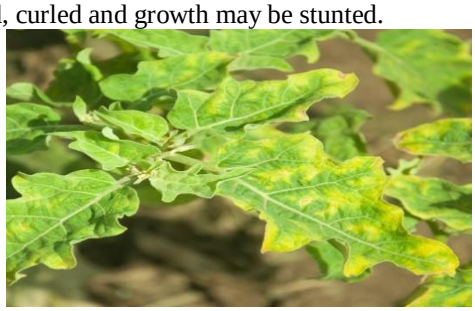

Figure1: Mosaic virus leaf disease

2.2 FUNGAL DISEASE SYMPTOMS

Among all plant leaf diseases, those caused by fungus some of them are discussed below and e.g. Late blight caused by the fungus Phytophthora infesters. It first appears on lower, older leaves like water-soaked, gray green spots. When fungal disease matures, these spots darken and then white fungal growth forms on the undersides. Early blight is caused by the fungus Alternariasolani. It first appears on the lower, older leaves like small brown spots with concentric rings that form a bull's eye pattern. When disease matures, it

spreads outward on the leaf surface causing it to turn yellow. In downy mildew yellow to white patches on the upper surfaces of older leaves occurs. These areas are covered with white to grey on the undersides.

Figure2: Fungal leaf disease

2.3BACTERIAL DISEASE SYMPTOMS The disease is characterized by tiny pale green spots which soon come into view as water- soaked. Then it appear as dry dead spots as e.g. bacterial leaf spot have brown or black water-soaked spots on the plants, sometimes with a yellow halo, generally identical in size. Under dry conditions the spots have a flecked appearance.

Figure3: Bacterial disease

III. IDENTIFICATION OF LEAF DISEASE PROCESS:

The following five process are used to indentifying the leaf disease.

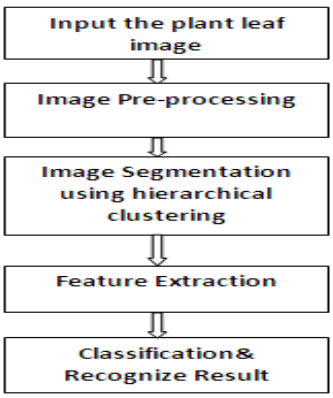

Figure4: System Overview Architecture

3.1 INPUT THE PLANT LEAF IMAGE :

Figure5: Affected leaf

3.2 IMAGE PRE-PROCESSING:

First the input image is preprocessed. Image Pre-processing Noise gets added during acquisition of leaf images. So various types of filtering techniques are used to remove noise. Create device independent color space transformation structure. Thus create the color transformation structure that defines the color space conversion. After that handle 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 different 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 changed 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. L*a*b is a device independent color space in which a & b components carry color information.

3.3 IMAGE SEGMENTATION

In this research hierarchical clustering is used for segmentation. Convert Image from RGB Color Space to L*a*b* Color Space. The L*a*b* space consists of a luminosity layer 'L*', chromaticity-layer 'a*' and 'b*'. All of the color information is in the 'a*' and 'b*' layers. Classify the colors in a*b* color space using hierarchical clustering. Since the image has 3colors create 3clusters.Measure the distance using Euclidean Distance Metric.

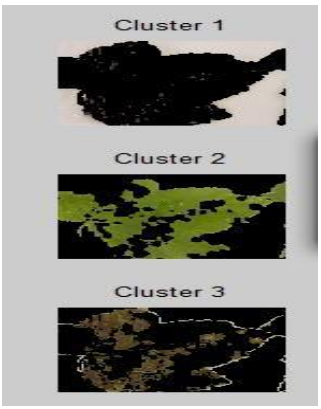

Figure6: Image Segmentation

3.4. FEATURE EXTRACTION

The diseases can be controlled by proper Disease management which is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Because of the complexity of visual patterns of the diseases there has been increasing demand for development of more specific and sophisticated image pattern understanding algorithms which can be used for studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc.
Hence to conduct high through put experiments, plant biologist need efficient computer software to automatically extract and analyze important features. As far as the leaf of the plant is considered, the significant features can be obtained by,

1. Color of the leaf
2. Texture of the leaf
3. Shape of the leaf

Color is one of the most widely used features. Color features can be obtained by various methods like Color histogram, Color correlogram, Color Moment, Color structure descriptor. The Color moment method has the lowest feature vector dimension and lower computational complexity. Hence it can be considered as suitable parameter to generate feature vectors which can be further used for classification purpose or for image retrieval.

3.5. SUPPORT VECTOR MACHINE

]Support vector machine (SVM) is a non-linear classifier. The idea behind the method is to non-linearly map the input data to some high dimensional space, where the data can be linearly separated, thus providing great classification performance. Support Vector Machine is a machine learning tool and has appear as a powerful technique for learning from data and in particular for solving binary classification problems. The main concepts of SVM are to first transform input data into a higher dimensional space by means of a kernel function and then construct an OSH (Optimal Separating Hyper Plane) between the two classes in the transformed space. For plant leaves classification it will transform feature vector extracted from leaf's contour. SVM finds the OSH by maximizing the margin between the classes. The SVM estimates a function for classifying data into two classes. Using a nonlinear transformation that depends on a regularization parameter, the input vectors are placed into a high-dimensional feature space, where a linear separation is employed.

3.6. CLASSIFICATION

A classification problem deals with associating a given input pattern with one of the distinct classes. Patterns are specified by a number of features (representing some measurements made on the objects that are being classified) so it is natural to think of them as d-dimensional vectors, where d is the number of different features. This representation gives rise to a concept of feature space. Patterns are points in this d-dimensional space and classes are sub-spaces. A classifier charge one class to each point of the input space. The problem of classification mostly establishes a transformation between the features and the classes. The optimal classifier is the one expected to produce the least number of misclassifications. In this research work SVM classification technique is used. SVM classifier is used to identify the right type of disease affected on input leaf image.

IV.CONCLUSION

Some of the challenges in these techniques are optimization of the technique for a specific plant, effect of the background noise in the acquired image and automation technique for a continuous automated monitoring of plant leaf diseases under real world field conditions. Here extended by developing better segmentation technique; selecting better feature extraction and classification algorithms are used. Also by computing severity and amount of disease present on the crop, only necessary and sufficient amount of pesticides can be used making agriculture production system economically efficient.

V.REFERENCES

[1] 1.Vyshnavi.G.K.P,Sirpa.M.R, Padmapriya.B Chandramoort.M, "Healthy and unhealthy plant leaf identification and classification using hierarchical clustering" International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 03 Mar-2016.
[2] Kshitij Fulsoundar, Tushar Kadlag, Sanman Bhadale, Pratik Bharvirkar, "Detection and classification of plant leaf diseases" International Journal of Engineering Research and General Science Volume2, Issue 6, October-November, 2014.
[3] Bed Prakash, Amit Yerpude,"A Survey on Plant Leaf Disease Identification" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 3, March 2015.
[4] S.A.Ramesh Kumar,Dr.K.Ramesh Kumar "A Study On Paddy Crops Disease Prediction Using Data Mining Techniques", International Journal of Data Engineering (IJDE) Singaporean Journal of Scientific Research(SJSR) Vol.7.No.1 2015.

