I J C T A, 9(17) 2016, pp. 8391-8397

Plant Leaf Disease Detection based on Unsupervised Learning

Manpreet Kaur* **and Sanjay Singla***

ABSTRACT

Disease occurred in plants and leaves reduce the adequacy and magnitude of crops productions. The identification of disease prevents damage of crops during growth, harvest and post-harvest. The disease is diagnosed in direct and indirect methods. In the direct methods the diseases caused by micro-organisms such as bacteria, fungi and viruses are directly recognised to provide exact identification of disease. In case of indirect method, the different features such as morphological changes, temperature change and organic compounds are identified for plant disease identification. Some of the major problems with traditional process were the constant observation of specialists which becomes costly in case of large farms. Moreover miles have to be travelled by farmers for contacting the experts with expensive cost. Automatic detection of plant disease covers large area of crops for damage and disease identification. The monitoring of crops can be done properly keeping the constraints for changes in temperature and organic compounds. The methodology makes the identification of the disease by extracting the important features for categorizing different disease occurred in the plant leaf.

Keywords: Morphological changes, Automatic detection, Fuzzy C-Means (FCM), Support Vector Machine (SVM).

I. INTRODUCTION

Plants have become an important resource. The diseases caused in plants have affected the plants in both quantity and quality. The growth of the crops requires continuous monitoring for finding the disease and its management through the experts. There are several diseases that affect the plants devastating economic, social and ecological losses. Among different kinds of diseases, some plants have disease symptoms that are not visible, whereas for other cases the signs can be detected with electromagnetic spectrum. Reduction in agricultural fertility can be due to a many factors among which damage caused by pests and micro-organisams plays a powerful role in crop losses. The failure in crop yield ranges between 20% and 40% due to pathogen infections. The advanced disease detection and prevention in crops is necessary in order to shrink the disease induced damage in crops during growth, harvest and post harvest processing, as well as to increase productivity and ensure agricultural sustainability. Detection and identification of disease can be done using the two direct and indirect methods. The direct method identifies the disease with molecular and serological methods for high throughput analysis when wide number of samples are analysed. This method includes identification of the disease caused by micro-organisms such as bacteria, fungi and viruses. Whereas the parameters such as morphological change, transpiration rate change, climate change, and volatile organic compounds released by infected plants are recognised by the indirect method. Automatic detection of plant disease covers large area of crops for damage and disease identification. The monitoring of crops can be done properly keeping the constraints for changes in temperature and organic compounds. The methodology makes the identification of the disease by extracting the important features for categorizing different disease occurred in the plant leaf.

The rest of the paper is organized as follows. Disease detection symptoms are are explained in section II. Plants and their diseases are explained in section III. Related work is discussed in section IV. Problem

* Department of Computer Science Engineering, IET Bhaddal, Ropar, Punjab, India, *E-mail: mkaur6711@gmail.com; hodcse@ietbhaddal.edu.in*

definition is discussed in section V. Objectives are discussed in section VI. Proposed system is explained in section VII. Results are presented in section VIII. Conclusion is given in section IX. Future work is mentioned in section X

II. DISEASE DETECTION SYMPTOMS

2.1. Underdevelopment of tissues or organs –

Symptoms for the underdevelopment of plants are the faulty root development, leaf abnormality, inadequate chlorophyll production that leads to improper development of flowers and fruits.

2.2. Overdevelopment of tissues or organs

Some of the overdevelopment of tissues include galls on roots, stems and lavish flowers.

2.3. Death of plant tissues or organs

These symptoms affect the whole plant having shoot or leaf blights.

2.4. Variation of normal appearance

In this examples include mosaic patterns of light and dark green color on leaves, and varied coloration in leaves and flowers.

III. PLANTS AND THEIR DISEASES

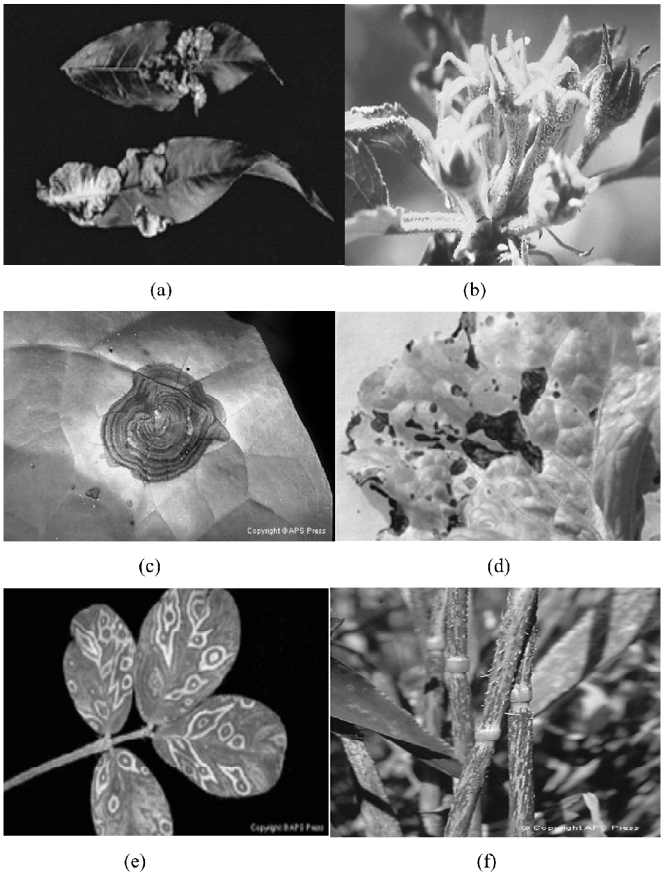


Figure 1. (a) Leaf Deformation (b) Powdery Mildew (c) Fungous leaf spots (d) Bacterial leaf spots (e) Mosaic and ringspot (f) Preesnce of spores

IV. RELATED WORK

S.Arivazhagan et.al [1] proposed method for automatic classification and discovery of plant leaf diseases. The process initiates with converting the RGB image of the plant leaves into HIS color space representation for facilitating color specifications. H component is considered for the analysis. In the next step the masking of the green pixels are done with the thresholding value. After masking the infected portion of the leaf are segmented into equal size patches. Patch having greater than fifty percent of information among the segmented patches are considered for further analysis. The color co-occurrence is done with the Spatial Gray-level Dependence Matrix (SGDM). SGDM matrices measure the probability of pixel at one appropriate gray level will occur at a distinct distance and orientation from any pixel given that pixel has a second appropriate gray level. The texture features from the H component is computed for classification. The classification is one with minimum distance criterion and support vector machine.

John William et.al [2] proposed the method that performs the image acquisition at initial step. The image contrast enhancement and noise removal is done with the HSV transformation. For adjusting the image intensity the Otsu's thresholding is done from gray level histogram. The image segmentation is done with thresholding to convert binary level image. The image is converted into histogram equalization using bins. The converted image performs masking to remove disease part of leaves. Different features are extracted based on fraction of leaf covered by disease, arithmetic mean of RGB color components, and standard deviation of RGB components and mean of HSV component. All the four components are analyzed with back propagation ANN.

Patil Sanjay B. et.al [3] proposed a method that works in two stages: Feature Extraction and Feature Matching. In feature extraction the feature vectors from images are extracted and stored in feature database. If the distance between the feature vector of the query image and database image is small enough compared with predefined threshold, then the corresponding image in the database is considered as a match to the query image. As a color feature mean value of hue histogram of the leaf is extracted. The shape features are extracted in the form of matching key points of the leaf image. These key points are extracted using SIFT algorithm. The texture features are extracted using our newly proposed LGGP (Local Gray Gabber Pattern) algorithm. The mean value of the histogram of the LGGP image is used as texture feature. Euclidean distance is performed for the feature retrieval from the database. The segmentation method is done for the identification of the disease, this is done with thresholding and region growth that is done with the K-means clustering that clusters all contains of green portion (Healthy) of the leaf into one cluster and another cluster contains diseased portion of the leaf.

Elham Omrani et.al [4] proposed a method that initiates with the removal of noise from the image and transformation of RGB to L*a*b* color space transformation. To extract the affected region, K-means clustering is done. From the extracted infectious area of the leaf, color, shape and texture features are obtained for classification. Texture features are obtained with the wavelet and GLCM matrix. For the classification the ANN and SVM are applied for detecting the disease of the plant leaf.

Sannakki Sanjeev S et.al [5] proposed a method that discusses different drawbacks of the existing methods for finding the leaf diseases with accuracy percent. The proposed method is initiated with the removal of noise with the Gaussian filter. The K-Means clustering technique is performed for computing the segmentation work. The result of the segmentation the diseased regions are extracted as clusters. From the clustered result, the total leaf area and the diseased leaf area are calculated for further computation of the percentage infected. A triangular fuzzy membership function is defined for grading the disease and find the percentage of disease occurred.

H. Al-Hiary et.al [6] proposed a method that is initiated by performing the RGB images as input to the system. The color transformation is done followed with K-Means clustering. The Ostu's threshold value is compared with the pixel intensities of the RGB components. The thresholding is done to represent the healthier parts of the leaf. The pixel values that are set to zero represents the edges of the infected leaf boundary. The affected cluster was then converted from RGB to HSI format. In the next step, the Spatial Gray Level Dependence Matrix (SGDM) were then developed for each pixel map of the image for only H and S images. The SGDM is a measure of the probability that a given pixel at one particular gray-level will appear at a distinct distance and orientation angle from another pixel, given that pixel has a second particular gray-level. From the SGDM matrices, the texture statistics for each image were achieved. Based on the features obtained the leaves diseases detection is done with automatic neural network.

Sanjay B. Patil et.al [7] proposed a method that is is initiated with conversion of the image into gray image. The RGB is transformed into HSI color space for proper identification of the plant leaf disease. For choosing a thresholding value of gray image, a triangle thresholding technique is used. To choose the value, triangle is created by constructing a line between the max value of the histogram at brightness bmax and the min value bmin in the image. The distance 'd' within the line and the histogram h[b] is computed for all values of 'b' from b=bmin to b max. The severity of the disease is categorized into different levels based on the Horsfall and Heuberger scale. The validity of the algorithm is obtained with the accuracy of estimation of the leaf disease region of the plant leaf.

V. PROBLEM DEFINITION

Apart from number of methods applied for the plant disease identification there are some methods which are too much specific. However, in some papers the proposed methods can deal with only one kind of plant species. The results for validation are not applied for finding the accuracy.

Another important point is the lack of technical knowledge about the methods. The problems having simple approaches is always favourable but the complex solutions give better results and require high technical knowledge. However some problems for over fitting and under fitting remains in the sample set having low representations and bias value.

The new methods have to be developed that would perform the identification of plant disease addressing the different types of plant diseases. Moreover, the image processing techniques needs to be applied in various dimensionality in order to find better results.

VI. OBJECTIVES

To study the existing methods on plant disease detection.

To design and implement a proposed methodology to proper identification of the plant disease.

Analysing the result of proposed methodology on different types of plant species.

Performance evaluation of proposed algorithm.

VII. PROPOSED SYSTEM

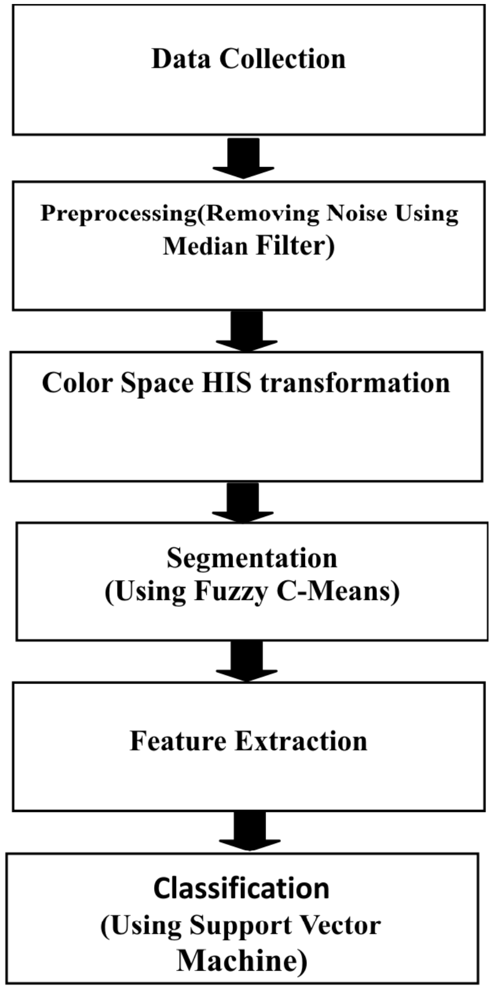
The step-by-step methodology is to be followed for plant leaf disease detection. Following diagram represent step by step procedure of implementation steps.

7.1. Data Collection

Firstly we need to collect the data. So the foremost step is the collection of data which is very necessary step.

7.2. Preprocessing

Preprocessing is done using the median filter.



7.3. Color space HIS transformation

The color space transformation is performed with HIS transformation that will have H component for analysis as it has more dominance to the user observance.

7.4. Segmentation Using FCM (Fuzzy C means)

Segmentation divides the image into meaningful parts for improved analysis and understanding of the image. FCM is used as segmentation method as it makes a decision for clustering the image pixel value depending on its degree of membership obtained from its belongingness to the diseased on non-diseased part of the plant leaf.

7.5. Feature Extraction

Feature extraction is done from the segmented region obtained after performing FCM. The different mean and variance of the pixels are considered for finding the features of the segmented portion.

7.6. Classification

Based on the features obtained, the Support Vector Machines (SVM) execute the classification for a given training set.

The step-by-step methodology is to be followed for plant leaf disease detection. Following diagram represent step by step procedure of implementation steps.

VIII. RESULTS AND DISCUSSIONS

The experimental analysis is done on six different types of plant leaf disease images. A training data set of 6 images is considered for performing the experiment on given test image. The result from the graph shows that the proposed algorithm performs better than the existing method. In the existing method Ostu's thresholding is done for segmenting the disease regions and extraction of the features. The proposed method performs the HSI transformation followed by calculating membership based Fuzzy C-Means algorithm performs the segmentation of the diseased regions resulting in better feature extraction. The feature values for the diseased images are computed with the mean and variance that identifies the plant disease properly. The graph plot represents the features values of the plants with respect to the existing and proposed method. In Fig.7 (c) the graph shows an average value for all the different plant diseases for the proposed method whereas there is a steep ascend and descend in the existing method for the same plants. The FCM algorithm optimizes the objective function resulting in better identification of the plant disease. However, the Ostu's thresholding in the existing method is incapable of identifying proper disease regions. Fig.7 (d) shows the variation as the feature value which is the deviation from mean value of segmented regions. As the FCM clustering is done for finding the diseased region the grouping of data would show less deviation than thresholding of the data that would require tedious work for finding the exact threshold value. Hence, from the analysis we can conclude that the proposed methodology gives better result than the existing ones.

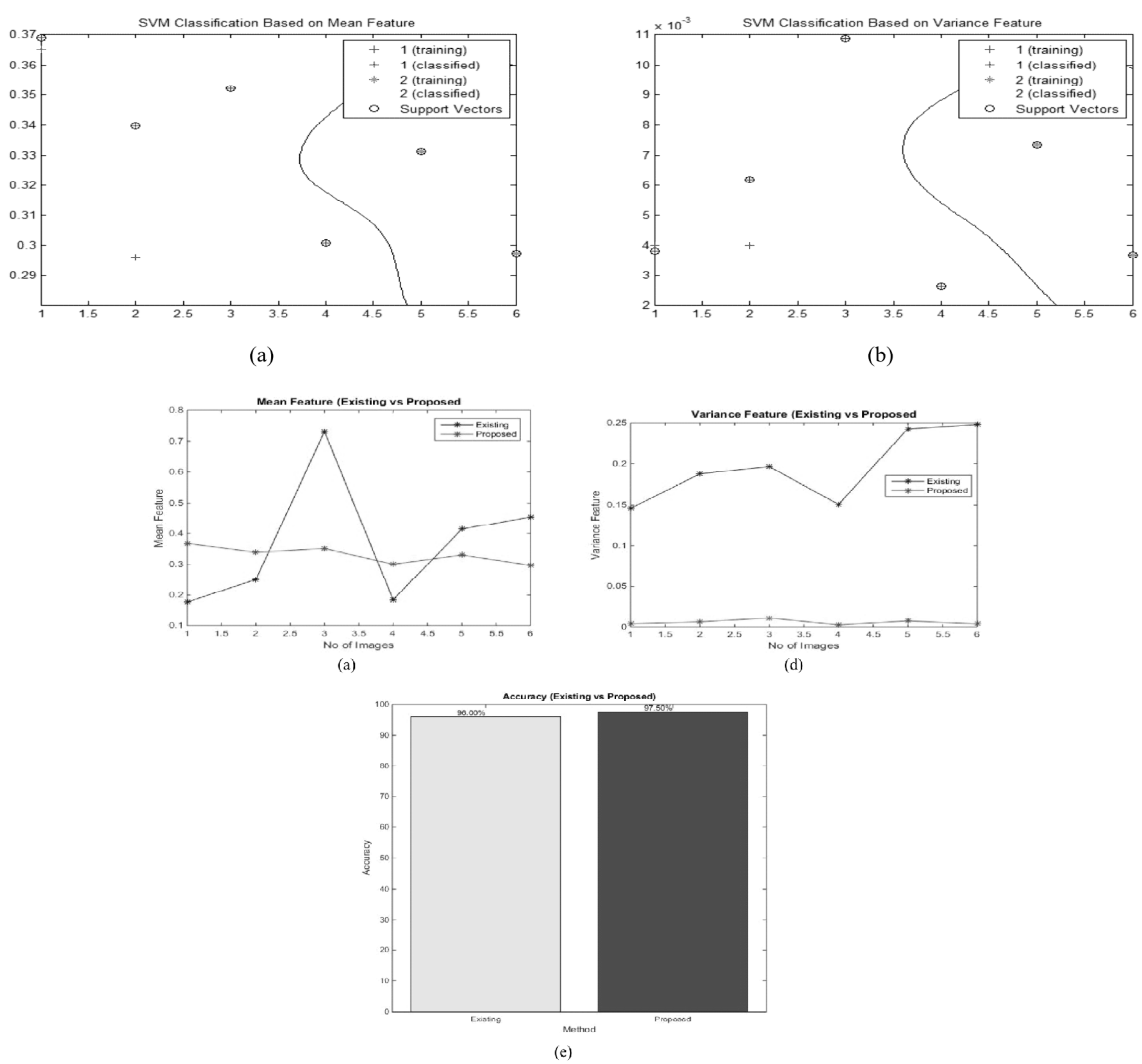


Figure 3: (a) SVM classification based on mean feature (b) SVM classification based on variance feature (c) Mean feature (Existing Vs Proposed) (d) Variance feature (Existing Vs Proposed) (e) Accuracy (existing Vs proposed)

IX. CONCLUSION

This work is the study of plant leaf diseases and various techniques for the finding the disease region. The various image processing techniques are deployed for finding in less computation time. The best techniques will be evaluated who have the maximum benefits from all these techniques. Inspite of the maximum benefits every technique has certain disadvantages. Automatic detection of plant disease covers large area of crops for damage and disease identification. The monitoring of crops can be done properly keeping the constraints for changes in temperature and organic compounds. The proposed methodology has been implemented successfully. The performance for the same is tested on a real set of leaf data. The result is quite convincing and immense adaptability in developing countries, where such information plays an essential role for enrichment in yield. The methodology makes the identification of the disease by extracting the important features for categorizing different disease occurred in the plant leaf. The proposed method is application of FCM with SVM classification for detection of plant leaf. Automatic detection and classification of plant leaf disease can be significantly supported by our experimental results. The proposed system provides 97.50% accuracy.

X. FUTURE WORK

The future scope of this work is to optimise the function with Evolutionary Computation.

REFERENCES

[1] Arivazhagan S, Shebiah R Newlin, Ananthi S Vishnu, 2013 " Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features".

[2] Willam John Orillo, Dela Jennifer Dela, Leobelle, Jensen Sa Paul, Ira Venenzula, 2014 "Identification of diseases in rice plant using Back Propogation Artificial Neural network".

[3] Patil Sanjay B., Dudhane T.M., 2015 "Leaf disease severity Detection using K-means clustering and CBIR".

[4] OmraniElham, Khoshnevisan Benyamin, ShamshirbandShahaboddin, Saboohi c Hadi, Anuar Nor Badrul , Md Nasir MohdHairulNizam, 2014 "Potential of radial basis function-based support vector regression for apple disease detection".

[5] SannakkiSanjeev S, Vijay S Rajpurohit Vijay S, Nargund V B, Kumar Arun R, YallurPrema S, 2011 "Leaf Disease Grading by Machine Vision and Fuzzy Logic".

[6] Hiary H. Al, Ahmad S. Bani, Reyalat M., Braik M. and ALRahamneh Z., 2011 "Fast and Accurate Detection and Classification of Plant Diseases".

[7] Patil Sanjay B., Bodhe Shrikant k., 2011 "Leaf disease severity Measurement Using Image Processing".

[8] Gurjar Ajay A., GulhaneViraj A.,2015 "Disease Detection On Cotton Leaves by Eigenfeature Regularization and extraction techniques".