

DETECTION OF LEAF DISEASES AND CLASSIFYING THEM USING MULTICLASS SVM

A. Santhosh¹, K. Saravanakumar², P. Saravanakumar³, S. Vimalraj⁴, R. Saranya⁵

^{1,2,3,4}B.Tech, Department of Information Technology, Valliammai Engineering College, Kanchipuram.

⁵Asst. Prof, Department of Information Technology, Valliammai Engineering College, Kanchipuram.

Abstract - Plant plays a major role in human life cycle. Plant intakes carbon-di-oxide (CO₂) and gives out Oxygen (O₂) which is more essential for the human for respiratory process. This process is done by the use of leaves. If the leaf gets infect, then this conversion process will be difficult to carryout. The scope of this project is to detect those deficiency/defect in the leaves. This project mainly involves four parts. They are (i)Image processing (or) Acquisition (ii)Image Segmentation using Lloyd's (or) K-means Algorithm (iii)Feature Extraction using GLCM (iv)Classification using Multiclass SVM. In this project, median filtering is used for removal of noise. The advantage of using median filter is that it avoids fading of edges. Segmentation process is carried out to extract interested portions of the leaf. For Segmentation process, K-means clustering algorithm is used. After segmenting, feature extraction process is carried out. In feature extraction, statistical feature of the leaves are extracted using GLCM (Grey-Level-Co-occurrence Matrix). Classification is done using Multiclass SVM. Classification is based on trained dataset and testing dataset. By comparing the classes of dataset the disease is identified.

Key Words: Image Acquisition, MF-Median Filtering, Lloyd's algorithm (or) K-means, GLCM- Grey Level Co-occurrence Matrix, Multiclass SVM (Support Vector Machine)

1. INTRODUCTION

India is fast developing country and Agriculture is the back bone for the countries development in the early stages. Agriculture is the backbone of Tamil Nadu since it provides a major contribution towards the GDP of the state. It is one of the predominant areas of the state economy since 70 percent of the state population is involved in agricultural activities for their livelihood. Tamil Nadu holds the second position in the large-scale production of agricultural products. The technology in the field of agriculture is developing day-by-day. Also, a large number of software is being simultaneously developed to educate the farmers with this technological information. Due to industrialization and globalization concepts the field is facing hurdles. On top of that the awareness and the necessity of the cultivation need to be installed in the minds of the younger generation. Now a day's technology plays a vital role in all the fields but till today we are using some old methodologies in agriculture. Identifying plant disease wrongly leads to huge loss of yield, time, money and quality of the food products. The condition of plant can be analyzed because it plays an important role for successful cultivation. In olden days identification is done manually by the experienced people but due to the environmental changes the prediction is becoming tough and expensive. So we can use image processing techniques for identification of plant and leaf diseases at a higher accuracy rate. In common, we can observe the symptoms of disease on leaves, stems, flowers etc. In our proposed system, we use leaves for identification of disease affected plants.

2. OBJECTIVE

This project is proposed to detect leaf diseases using multiclass SVM.

1. Initially, the given input image is converted from RGB to grey scale image for image processing.
2. Then the image is enhanced by resizing and removal of noise.
3. Then, segmentation of the components is done using K-means clustering algorithm (or) Lloyd's algorithm.
4. For feature extraction, texture features are obtained using statistical Grey-Level Co-occurrence Matrix.

3. RELATED WORK:

[1] R.Meena Prakash "Detection of Leaf Diseases and Classification using Digital Image Processing" [2017]IEEE, Leaf is the most important factor in performing photosynthesis in plants which results in the manufacture of oxygen. If the leaf is affected by any disease then the process of photosynthesis may affect and the leaf will find it difficult in conversion of carbon-di-oxide and oxygen. If there is more number of leaves the defect cannot be noticed and corrected. So image processing technique is used to

detect the plant leaf disease. To implement image analysis and classification techniques for detection of leaf disease and classification this system uses a framework consists of four parts. They are Image pre-processing, Image segmentation, Feature Extraction, Classification. Classification is done using Support Vector Machine.

[2] Aparajita, Rudransh sharma, Anushikha Singh, Malay Kishore Dutta" Image processing based automated identification of late blight disease from leaf images of potato crops"[2017]IEEE, Late Blight is one of the most common and devastating disease for potato crops in all over the world. For less use of pesticide and to minimize loss of potato crops, identification of late blight disease is necessary. The conventional method of disease identification based on visual assessments which is time consuming process and involves manpower. The proposed work presents image processing based automated identification of late blight disease from leaf images. In the proposed method, adaptive thresholding is used for segmentation of disease affected from leaf image. The threshold value is calculated using statistical features of image which makes the proposed system fully automatic and invariant under environment conditions. 96% accuracy can be obtained using this method.

[3] Chaitali G.Dhaware, K.H.Wanjale" A modern approach for plant leaf disease classification which depends on leaf image processing"[2017]IEEE, Agrarian production is that trait on which our nation's economy immensely depends. This is the motivation that recognition of crops and productivity. It requisite enormous amount of work, mastery in the leaf diseases, and additionally need the extreme amount of time. Thus, image processing techniques are applied for the discovering and recognition of plant leaf unhealthiness. Recognition of plant diseases along some automatic method is useful as it decrease a huge effort of observing in large farms, and at initial phase itself it identify the signs of diseases.

[4]R.P.Narmadha, G.Arulvadivu" Detection and measurement of paddy leaf disease symptoms using image processing"[2017]IEEE, Plant are one of the major resources to avoid the global warming in the world. But the plants are affected by the diseases like Blast, Canker, etc... Paddy disease can be detected using this method. Disease can infect paddy at different stages of growth and all parts of the plants as the leaf neck and the node. This methodology was designed to remove the noise automatic, error by human and minimizing the time taken to measure the affect of paddy leaf disease. This proposed method increases accuracy.

[5] Jingzhu li, Peng Wang, Changxing Geng" The disease assessment of cucumber downy mildew based on image processing"[2017]IEEE, Cucumber downy mildew is a kind of disease which spreads very fast and is dangerous, in order to prevent the disease, people always spray plenty of pesticides indiscriminately. Accurate assessment of the level of cucumber downy mildew is very important to the disease prevention and control. In a cucumber growing season, this paper collected the typical cucumber downy mildew leaf samples, and developed the downy mildew spot extraction algorithm by using leaf image scanning method, calculated the index of the disease. The average identification accuracy of downy mildew image reaches 98.3%, and average image processing takes 10.9 ms/picture. By compared with human eyes assessment method have strong subjectively, dramatic changes and bigger error, while the image analysis method get the correlation coefficient for disease index and basic value of 0.9417, has obvious linear correlation.

[6] Anand R, Veni S, Aravinth J" An application of image processing techniques for detection of diseases on brinjal leaves using K-means clustering method"[2016]IEEE, This method identifies leaf disease and it gives approach for careful detection of diseases. The main goal of proposed work is to diagnose the disease of brinjal leaf using image processing and artificial neural techniques. The study of interest is the leaf rather than whole brinjal plant because about 85-95% of diseases occurred on the brinjal leaf like Bacterial Wilt, etc... In this method, they use k-mean for segmentation and neural network for classification. The proposed detection model based artificial neural networks are very effective in recognizing leaf diseases.

[7] Dheeb Al Bashish, Malik Braik, Sulieman Bani-Ahmad" A framework for detection and classification of plant leaf and stem diseases" [2010] IEEE, They proposed framework for detecting plant leaf/stem diseases. It provides fast, automatic, cheap and accurate image-processing based solutions for that task can be of great realistic significance. The proposed framework is image-processing based and is composed of the following steps. Images are segmented using K-means and in second step the segmented images are passed through a pre-trained neural network. The developed Neural Network classifier that is based on statistical classification perform well and could successfully detect and classify tested diseases with a precision of 97%.

4. PROPOSED SYSTEM:

In this system, the accuracy rate is improved with the usage of Multiclass Support Vector Machine (M-SVM). Initially the RGB image acquisition is performed and then Resizing is done. Then, Median filtering method is used to detect and remove noises in each and every pixel thus making it effective and it replaces the existing Gaussian filtering method. Then the image is converted from RGB to Greyscale image for the purpose of Preprocessing. Then segmentation of the image is done using K-means Clustering algorithm which is one of the best segmentation algorithm. Feature Extraction is done with the help of segmented

image with the help of GLCM algorithm which extracts statistical features. Then for the diseased leaf, the classification is done using Multiclass SVM which produces an accuracy rate of 97%.

5. SYSTEM ARCHITECTURE:

The input image is being photographed with the help of using photo camera. The captured is being given as input to the image acquisition process. It helps to resize the image resolution and median filter is applied to remove the noises from the image. Then image enhancement takes place. After this process, image segmentation takes place. With the help of K-means clustering, segmentation is made efficient. Then GLCM is used for feature extraction. Finally classification takes place by using Multiclass SVM classifier. Then the predicted disease is produced as a result.

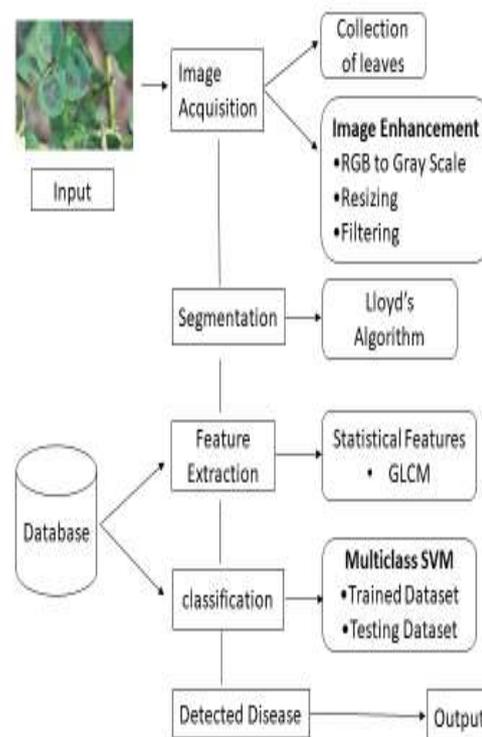


Fig -1: Overall Architecture Diagram

6. METHODOLOGY:

A. IMAGE ACQUISITION:

Suspected plant leaves are used to carry out this project. The digital camera was used to capture images in controlled environment with the dark background. The images were stored in either JPEG or PNG format. This was to eliminate any reflection and get light evenly distributed everywhere in order to get a better view and brightness of the image to be processed. The object (leaf) was properly zoomed using the digital camera to ensure that the picture taken contained only the leaf and black background with the rightful image size.

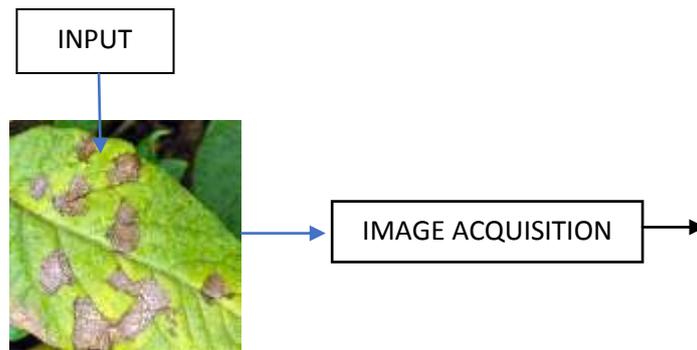


Fig 2: Image acquisition process

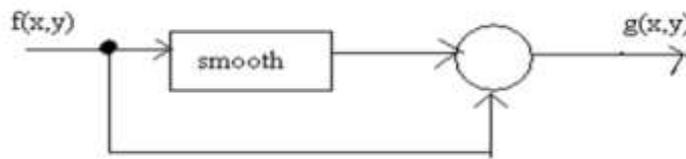


Fig 3: Image Enhancement process

B. IMAGE SEGMENTATION:

Image segmentation is the process used to simplify the representation of an image into meaningful form, such as to highlight object of interest from background. The K-means clustering algorithm performs segmentation by minimizing the sum of squares of distances between the image intensities and the cluster centroids. K-means clustering algorithm, or Lloyd's algorithm, is an iterative algorithm that partitions the data and assigns n observations to precisely one of k clusters defined by centroids. The steps in the algorithm are given below.

1. Choose k initial cluster centers (centroid).
2. Compute point-to-cluster-centroid distances of all observations to each centroid.
3. Assign each observation to the cluster with the closest centroid.
4. Compute the mean of the observations in each cluster to obtain k new centroid locations.
5. Repeat steps 2 through 4 until there is no change in the cluster assignments or the maximum number of iterations is reached.

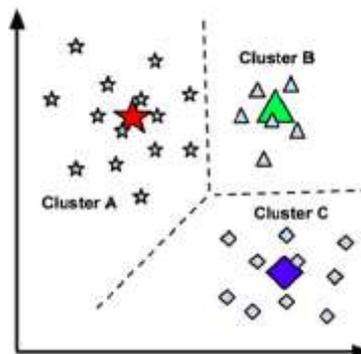


Fig 4: k means clustering algorithm for k clusters

C. FEATURE EXTRACTION:

After segmentation, the GLCM features are extracted from the image. Gray-Level Co-Occurrence Matrix (GLCM) is the statistical method of investigating texture which considers the spatial relationship of pixels. The GLCM functions characterize the texture of images by computing the spatial relationship among the pixels in the images. The statistical measures are extracted from this matrix. In the creation of GLCMs, an array of offsets which describe pixel relationships of varying direction and distance have to be specified. In the proposed method, four features are extracted which include contrast, energy, homogeneity and correlation. Let P_{ij} represents the (i,j) th entry in the normalized Gray-Level Co-Occurrence Matrix. N represents the number of distinct gray levels in the quantized image.

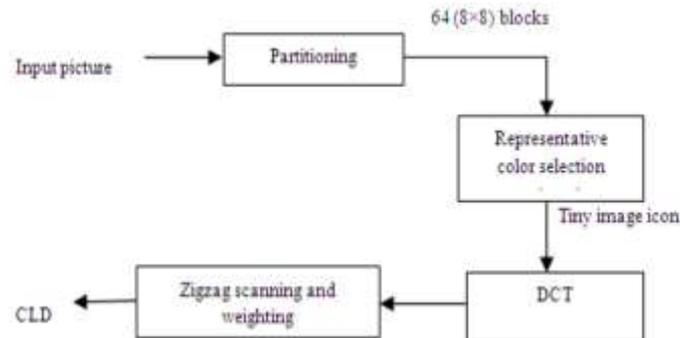


Fig 5: Feature Extraction process

D. CLASSIFICATION USING M-SVM:

Support Vector Machine is kernel-based supervised learning algorithm used as a classification tool. The training algorithm of SVM maximizes the margin between the training data and class boundary. The resulting decision function depends only on the training data called support vectors, which are closest to the decision boundary as shown in Figure 2. It is effective in high dimensional space where number of dimensions is greater than the number of training data. SVM transforms data from input space into a high-dimensional feature space using kernel function. Nonlinear data can also be separated using hyper plane in high dimensional space. The computational complexity is Reduced by Kernel Hilbert Space(RKHS). The idea of support vector machine is to create a hyper plane in between data sets to indicate which class it belongs to. The feature vector is given as input to the classifier. The feature vectors of the database images are divided into training and testing vectors. The classifier trains on the training set and applies it to classify the testing set. The performance of the classifier is measured by comparing the predicted labels and actual values.

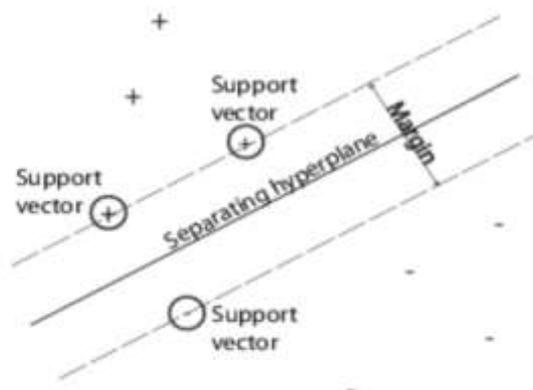


Fig 6: Classification using Multiclass SVM

7. CONCLUSION

A method for detection and classification of leaf diseases is implemented with the help of Multiclass SVM. The segmentation of the infected part of the leaf is done using K-Means segmentation. Then, GLCM texture features are extracted and classification is done using Multiclass SVM. The method is tested for detection of diseases in plant leaves. Future work is to be carried out for classification of diseases in different leaves and species of plants and to improve the classification accuracy.

8. REFERENCES

- [1] R.Meena Prakash"Detection of Leaf Diseases and Classification using Digital Image Processing"[2017]IEEE.
- [2] Aparajita, Rudransh sharma, Anushikha Singh, Malay Kishore Dutta"Image processing based automated identification of late blight disease from leaf images of potato crops"[2017]IEEE.
- [3] Chaitali G.Dhaware, K.H.Wanjale"A modern approach for plant leaf disease classification which depends on leaf image processing" [2017]IEEE.
- [4] R.P.Narmadha,G.Arulvadvu"Detection and measurement of pady leaf disease symptoms using image processing"[2017]IEEE.
- [5] Jingzhu li, Peng Wang, Changxing Geng" The disease assessment of cucumber downy mildew based on image processing"[2017]IEEE.
- [6] Anand R,Veni S,Aravinth J"An application of image processing techniques for detection of diseases on brinjal leaves using K-means clustering method"[2016]IEEE.
- [7] Dheeb Al Bashish,Malik Braik,Sulieman Bani-Ahmad"A framework for detection and classification of plant leaf and stem diseases"[2010]IEEE.