

A Survey of Identification of Plant Disease using Image Processing Techniques and Chameleon Clustering Method

Shailja M. Maniya

PG Student

Department of Information & Technology

L.D. College of Engineering Gujarat Technological University Ahmedabad, Gujarat, India

Abstract— This work presents overview of different methods to find plant diseases based on leaf characteristics. If these plant diseases are not identified at appropriate time then that can be result into less agriculture productivity. Sometimes even Farmers find it difficult to detect disease in plant. Using image processing, we can detect disease easily using various clustering methods. These methods are only used for detection of diseases using leaves and 65-75 % diseases are on leaves of plant. Here, analysis on different clustering methods i.e. Known Nearest Neighbor (K-NN), K-means and K-Medoid is done which are being used in process for detection of disease. Basic steps includes image acquisition, image preprocessing, feature extraction and statistical analysis and classification based on classifier. All three classification methods detects diseases successfully, But accuracy is different.

Key words: Known Nearest Neighbor (K-NN), K-means, K-Medoid, Classification, Image Processing

I. INTRODUCTION

India is such an agricultural country where most of people earn through agriculture. Vegetables, seeds, fruits etc. are most important products of agriculture. These products are necessity of human life. Any loss of production in these products can cost too much to human being. It is considerable that this loss of productions is because of plant diseases. In order to overcome this issue, plant disease should be found correctly and efficiently. For that the best and accurate way to find plant disease is done in this research work by analysing all three clustering methods.

II. LITERATURE SURVEY

This research explores various clustering methodologies. The diseases can be in various parts of plant like stem, root, leaf or fruits. But, here the focus is only on leaves of plant.

Anand R et al. [5], describes a technique by which disease on brinjal leaves can be identified using k-mean clustering method and also calculates features by calling gray-level co-occurrence matrix (GLCM) parameters as well as recognises by artificial neural network(ANN).

Priya Soni et al. [6], discussed a methodology to detect 9 types of different plants which are affected by diseases. This methodology uses k - Nearest Neighbor classifier and neural network for precise- ness. It gives accuracy of 90-100%. The paper by Arti N. Rathod et al. [9], describes algorithm to find leaf disease using k-medoid clustering technique, SGDM matrix, GLSM function as well as Neural network for disease recognition with accuracy 96%.

Sanjay B. Dhaygude et al. [1], proposed a methodology which detects plant leaf disease using colour co-occurrence matrix and texture analysis using contrast, energy, homogeneity, correlation. Yogesh Dandawate et al. [2], proposed a technique to identify the disease using SVM classifier. Basic steps involved in this technique are image acquisition, image pre-processing, image segmentation, statistical analysis, feature extraction, classifier, predicted disease class, decision support system. It gives accuracy of 93.79%.Rajleen Kaur et al. [3], discussed methodology to detect disease on plant leaves using some basic steps such as are image acquisition, image pre-processing, image segmentation, feature extraction, statistical analysis, classifier. But, here main focus is on feature extraction which is in three subparts that is shape, colour and texture. Using this Enhanced SVM, better results are obtained than Existing SVM.

P.B. Chopade et al. [4], describes a methodology for disease detection using image processing. In this methodology, segmentation is done using histogram thresholding and some morphological operations. Zulkifli Bin Husin et al. [7], discusses a technology for detection of plant disease using color features of an image. It is specially experimented on chilli plants. Based on amount of yellow, cyan and green color it gives result whether the plant is healthy or risky. The study by Kiran R. Gavhale et al. [8], using k-mean clustering algorithm for image segmentation and GLCM for feature extraction. In SVM, training dataset contains 200 samples for both citrus leaves (canker and Anthracnose).

In order to concern above mentioned work and their results, I have put forward a new technique to identify and analyse the plant disease based on leaves. My proposed work can be applied to all types of images can give better result in terms of accuracy for diseased area.

III. PROPOSED METHODOLOGY

In this section, I am considering the general flow of steps which should be considered in order to achieve the desired output. There are basically four steps: Image acquisition, feature extraction, statistical data and classification by SVM, chameleon classifier, Neural Network. Proposed Methodology for "Identification of Plant Disease Using Image Processing Techniques and Chameleon Clustering Method" is shown below:

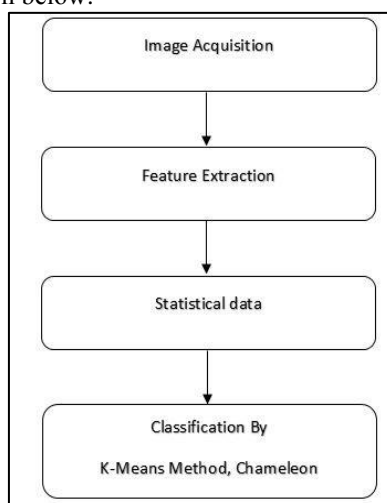


Fig. 1: Proposed Methodology

IV. CLASSIFICATION METHODS USED FOR IMAGE PROCESSING

In classification process, here I am going to use three methods:

- K-Means Clustering
- K-Medoid Clustering
- Chameleon Clustering

A. K-Means Clustering

In this approach, K number of clusters are made and mean of some points (co-ordinates of image) is considered as centroid of the cluster. The steps in k-means clustering method is described as follow:

- Each point is assigned to group for which distance from centroid is less.
- When all points are assigned, recalculate positions of K centroids.
- Repeat step 1 and 2 until the change of position of points remain unchanged. Thus, clusters are formed in such a way that each points are assigned to the groups in order to minimize the size of cluster.

Advantage of this method is it produces strong clusters if there are large number of variables. It is computationally faster than hierarchical clustering techniques. Although there are some disadvantages of this method such as it is sometimes difficult to compare the quality of cluster and also to predict the value of K. If non-globular data (Outliers) then it does not work well with it.

B. K-Medoid Clustering

In this approach, same as k-means method, k number of clusters are made. But here one change is that the centroid of that cluster is medoid of points, not mean of those points. Thus, the centroid is now one of the points from image itself. Otherwise algorithm for K-Medoid is same as K-Means algorithm.

Advantage of this clustering method is we get here more accuracy than K-Means method as K-Medoid is not sensitive to Outliers. This K-Medoid method includes three sub-methods called PAM (Partitioning around Medoid), CLARA (Clustering Large Applications) and CLARANS (A Clustering Algorithm based on Randomized Search). Basically PAM works for small data and CLARA works for large data. CLARANS is 5 times faster than PAM otherwise cluster quality is same.

C. Chameleon Clustering

This approach is using hierarchical clustering. It is one of K-NN (Known Nearest Neighbour) algorithms which finds the point in a given set that is closest to the given point. On this basis clustering algorithm works. It uses dynamic modelling. This method does not generate any clusters but it merges two clusters using interconnectivity and closeness (proximity). Algorithm for this clustering is divided into two phase:

- Graph partitioning algorithm
- Agglomerative hierarchical clustering algorithm

Here in second phase we are using Agglomerative hierarchical clustering which is bottom up approach. This approach joins two points based on dissimilarity. Less the dissimilarity they will be part of a cluster.

Dissimilarity can be measured by given equation

$$E = \sum_{j=1}^k \sum_{p \in C_j} |p - O_j|$$

Where C_j is the centroid of the clusters, k is total number of clusters, O_j is object belonging to any cluster. This process is repeated until there is no change in clusters from previous case to present case. Thus, Cluster formation process takes place in second phase using agglomerative approach after partitioning.

The flow of this algorithm is shown in below figure

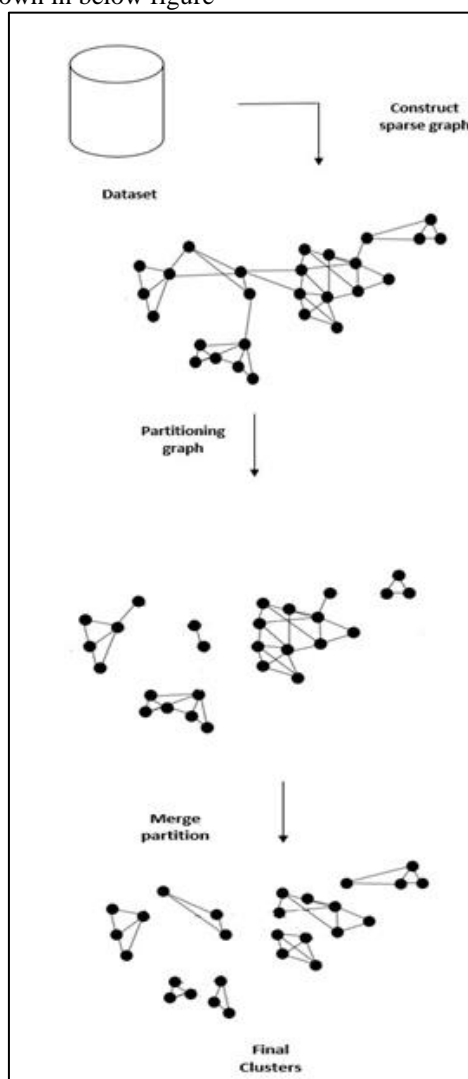


Fig. 2: Flow of Algorithm

V. CONCLUSION

In this article, I have analysed three clustering methods for classification in image processing. Among these methods, chameleon seems better because it can give most efficient result in terms of accuracy for finding percentage of disease in plant. As chameleon clustering algorithm focuses on partitioning and joining those clusters based on dissimilarity, it gives better cluster formation than other two methods.

REFERENCES

- [1] Sanjay B. Dhaygude, Nitin P. Kumbhar, "Agricultural plant Leaf Disease Detection using Image Processing", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol.2, pp.599-602, 2013.
- [2] Yogesh Dandawate, Radha Kokare, "An Automated Approach for Classification of Plant Diseases Towards Development of Futuristic Decision Support System in Indian Perspective", International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.794-799, 2015.
- [3] Rajleen Kaur, Sandeep Singh, "An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection", IEEE 3rd International Conference on MOOCs, Innovation and technology in Education (MITE), pp.135-140, 2015.

- [4] Chopade, P.B. and Katarak, Bhagyashri, "Image processing Based Detection and classification of leaf disease on fruits crops", International Journal of Electrical, Electronics and Computer Systems, ISSN.vol.4, pp.33-36,2016.
- [5] Anand R.,Veni S. and Arvinth J, "An Application of image processing techniques for Detection of Diseases on Brinjal Leaves Using K-Means Clustering Method", 2016 Fifth International Conference on Recent Trends in Information Technology, DOI: 978 – 1 – 4673 – 9802 – 2/16/\$31.00, 2016.
- [6] Priya Soni,"Identification of Plant Disease Using Image Processing Techniques", International Journal of Research in Medical & Applied Sciences IJRMAS ISSN: 2454 – 3667,vol.1, Nov 2015.
- [7] Zulkifli Husin, Abdul Aziz, Ali Shakaff, Rohani Farook,"Feasibility Study on Plant Chili Disease Detection using Image Processing Techniques. Third International Conference on Intelligent Systems Modelling and Simulation".pp.292-296, 2012.
- [8] Kiran Gavhale, Ujwalla Gawande, Kamal Ha- jari, "Unhealthy Region of Citrus Leaf Detection Using Image Processing Techniques", International Conference for Convergence of Technology- 2014 .DOI:978-1-4799-3759-2/14/\$31.00,pp.1-6, 2014.
- [9] Arti N. Rathod,Bhavesh A. Tanawala, Vatsal H. Shah, "Leaf disease detection using image processing and neural network", International Journal of Advance Engineering and Research Development (IAERD),pp.1-10,2014.