

Identification of Disease in Leaves using Genetic Algorithm

K. Beulah Suganthi

Associate Professor, Department of Electronics and Communication Engineering,
Thanthai Periyar Govt. Institute of Technology, Vellore, Tamil Nadu, India

How to cite this paper: K. Beulah Suganthi "Identification of Disease in Leaves using Genetic Algorithm" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-3, April 2019, pp.1264-1267, URL: <https://www.ijtsrd.com/papers/ijtsrd22901.pdf>



IJTSRD22901

ABSTRACT

Plant disease is an impairment of normal state of a plant that interrupts or modifies its vital functions. Many leaf diseases are caused by pathogens. Agriculture is the mainstay of the Indian economy. Perception of human eye is not so much stronger so as to observe minute variation in the infected part of leaf. In this paper, we are providing software solution to automatically detect and classify plant leaf diseases. In this we are using image processing techniques to classify diseases & quickly diagnosis can be carried out as per disease. This approach will enhance productivity of crops. It includes image processing techniques starting from image acquisition, preprocessing, testing, and training.

KEYWORDS: Leaf disease detection, image processing techniques, Genetic algorithm, SVM Classifier

Copyright © 2019 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



I. INTRODUCTION

In developing countries, farming land can be much larger and farmers cannot observe each and every plant, every day. Farmers are unaware of non-native diseases. Consultation of experts for this might be time consuming & costly. Also unnecessary use of pesticides might be dangerous for natural resources such as water, soil, air, food chain etc. as well as it is expected that there need to be less contamination of food products with pesticides. There are two main characteristics of plant disease detection machine-learning methods that must be achieved, they are: speed and accuracy. There is need for developing technique such as automatic plant disease detection and classification using leaf image processing techniques. This will prove useful technique for farmers and will alert them at the right time before spreading of the disease over large area. Solution is composed of four main phases; in the first phase we create a color transformation structure for the RGB leaf image and then, we apply color space transformation for the color transformation structure. Most of the disease on plant is on their leaves and on stem of plant. The diseases are classified into viral, bacterial, fungal, diseases due to insects, rust, nematodes etc. on plant. Early detection of diseases is a major challenge in horticulture/agriculture science. Many disease produce symptoms which are the main tools for field diagnosis of diseases showing external symptoms out of a series of reactions that take place between host and pathogen.

II. LITERATURE REVIEW

Various techniques of image processing and pattern recognition have been developed for detection of diseases occurring on plant leaves, stems, lesion etc. by the researchers. The sooner disease appears on the leaf it should be detected, identified and corresponding measures should be taken to avoid loss. Hence a fast, accurate and less expensive system should be developed. The researchers have adopted various methods for detection and identification of disease accurately. One such system uses thresholding and back propagation network. Input is grape leaf image on which thresholding is performed to mask green pixels. Using K-means clustering segmented disease portion is obtained. Then ANN is used for classification [1]. The other method uses PCA and ANN. PCA is used to reduce the dimensions of the feature data. to reduce the no. of neurons in input layer and to increase speed of NN[2]. Sometimes threshold cannot be fixed and object in the spot image cannot be located. Hence authors proposed LTRSG-algorithm for segmentation of image [3]. In cucumber leaf disease diagnosis, spectrum based algorithms are used [4].

In the classification of rubber tree disease a device called spectrometer is used that measures the light intensity in electromagnetic spectrum. For the analysis SPSS is used [5]. In citrus canker disease detection uses three level system. Global descriptor detects diseased lesion. To identify disease from similar disease based regions zone based local

descriptor is used. In last stage two level hierarchical detection structure identifies canker lesion [6]. For identification of disease on plant and stems first segmentation is carried using K-means clustering. Feature extraction is done by CCM method. Identification is done by using BPNN [7].

III. PROPOSED SYSTEM

➤ Image Acquisition

In the proposed method collected the images from the dataset like pomegranate leaf Image Database Consortium. The dataset contains two types of images such as disease affected leaf images and healthy leaf images.



Figure 1: Input leaf images

➤ Enhancement

Enhancement technique enhances the contrast of images. The contrast enhancement can be helpful to remove the noise, which is present in the image.

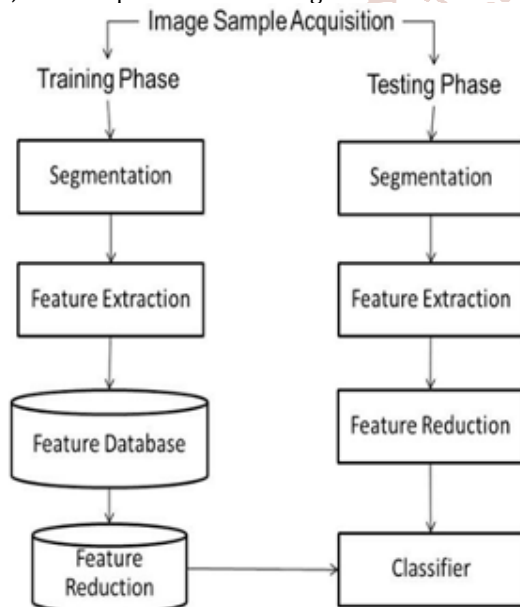


Figure 2: Flow of system

➤ Segmentation

Segmentation means it subdivides the image region into small regions. In our proposed method we have used genetic algorithm for the segmentation. Genetic algorithm is used for classification of object based on a set of features into number of classes.

A **genetic algorithm** (or **GA**) is a search technique used in computing to find true or approximate solutions to optimization and search problems.

➤ Basic principle

The searching capability of GAs has been used in this article for the purpose of appropriately determining a fixed number K of cluster centers in RN; thereby suitably clustering the set

of n unlabelled points. The clustering metric that has been adopted is the sum of the Euclidean distances of the points from their respective cluster centers.

```

Begin
1.  $t=0$ 
2. initialize population  $P(t)$ 
3. compute fitness  $P(t)$ 
4.  $t = t+1$ 
5. if termination criterion achieved go to step 10
6. select  $P(t)$  from  $P(t-1)$ 
7. crossover  $P(t)$ 
8. mutate  $P(t)$ 
9. go to step 3
10. Output best and stop
End
  
```

Basic steps in GA

➤ GA-clustering algorithm

The basic steps of GAs, which are also followed in the GA-clustering algorithm, are shown in Fig. 1. These are now described in detail.

➤ String representation

Each string is a sequence of real numbers representing the K cluster centers. For an N -dimensional space, the length of a chromosome is $N \times K$ words, where the first N positions (or, genes) represent the N dimensions of the first cluster centre, the next N positions represent those of the second cluster centre, and so on. As an illustration let us consider the following example.

Example1. Let $N=2$ and $K=3$, i.e., the space is two-dimensional and the number of clusters being considered is three. Then the chromosome 51.6 72.3 18.3 15.7 29.1 32.2 represents the three cluster centers (51.6, 72.3), (18.3, 15.7) and (29.1, 32.2). Note that each real number in the chromosome is an indivisible gene.

➤ Population initialization

The K cluster centers encoded in each chromosome are initialized to K randomly chosen points from the data set. This process is repeated for each of the P chromosomes in the population, where P is the size of the population.

➤ Fitness computation

The fitness computation process consists of two phases. In the first phase, the clusters are formed according to the centers encoded in the chromosome under consideration. This is done by assigning each point x_i , $i=1, 2, 3, \dots, n$, to one of the clusters C_j with centre z_j

➤ Selection

The selection process selects chromosomes from the mating pool directed by the survival of the fittest concept of natural genetic systems. In the proportional selection strategy adopted in this article, a chromosome is assigned a number of copies, which is proportional to its fitness in the population, that go into the mating pool for further genetic operations. Roulette wheel selection is one common technique that implements the proportional selection strategy.

➤ Crossover

Crossover is a probabilistic process that exchanges information between two parent chromosomes for generating two child chromosomes. In this article single

point crossover with a fixed crossover probability of k_c is used. For chromosomes of length l , a random integer, called the crossover point, is generated in the range $[1, l-1]$. The portions of the chromosomes lying to the right of the crossover point are exchanged to produce two offspring.

➤ Mutation

Each chromosome undergoes mutation with a fixed probability k_m . For binary representation of chromosomes, a bit position (or gene) is mutated by simply flipping its value. Since we are considering floating point representation in this article, we use the following mutation.

A number d in the range $[0, 1]$ is generated with uniform distribution.

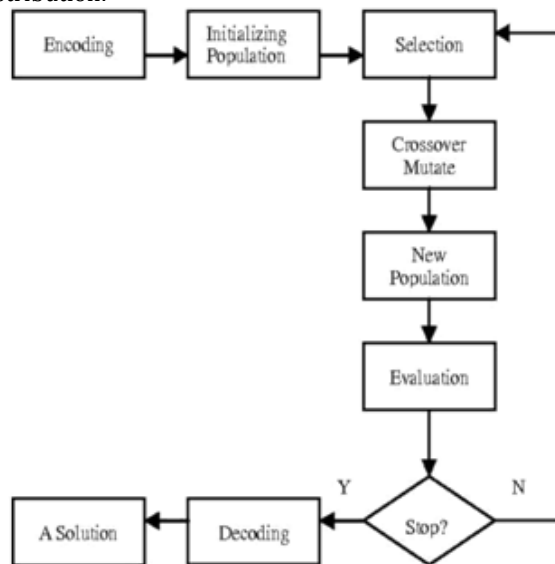


Figure3: Genetic algorithm

➤ Gray level co-occurrence matrix Features

Feature extraction is very important and essential step to extract region of interest. In our proposed method the basic features are mean, standard deviation, entropy, IDM, RMS, variance, smoothness, skewness, kurtosis, contrast, correlation, energy and homogeneity are calculated and considered as feature values. Then we have created the feature vector for these values. The segmented method shows different values for images.

In feature extraction desired feature vectors such as color, texture, morphology and structure are extracted. Feature extraction is method for involving number of resources required to describe a large set of data accurately. Statistical texture features are obtained by Gray level co-occurrence matrix (GLCM) formula for texture analysis and texture features are calculated from statistical distribution of observed intensity combinations at the specified position relative to others.

Numbers of gray levels are important in GLCM also statistics are categorized into order of first, second & higher for number of intensity points in each combination. Different statistical texture features of GLCM are energy, sum entropy, covariance, information measure of correlation, entropy, contrast and inverse difference and difference entropy.

➤ Classification of disease

The binary classifier which makes use of the hyper-plane which is also called as the decision boundary between two of

the classes is called as Support Vector machine(SVM). Some of the problems of pattern recognition like texture classification make use of SVM. Mapping of nonlinear input data to the linear data provides good classification in high dimensional space in SVM.

SVM is basically binary classifier which determines the hyper plane in dividing two classes. The boundary is maximized between the hyper plane and the two classes. The samples that are nearest to the margin will be selected in determining the hyper plane are called as support vectors. Multiclass classification is also possible either by using one-to-one or one-to-many. The highest output function will be determined as the winning class. Classification is performed by considering a larger number of support vectors of the training samples. The standard form of SVM was intended for two-class problems. However, in real life situations, it is often necessary to separate more than two classes at the same time.

SVM can be extended from binary problems to multi classification problems with k classes where $k > 2$. There are two approaches, namely the one- against-one approach and the one-against-all approach. In fact, multi-class SVM converts the data set to quite a few binary problems. For example, in one-to-one approach binary SVM is trained for every two classes of data to construct a decision function. Hence there are $k(k-1)/2$ decision functions for the k -class problem. Suppose $k=15$, 105 binary classifiers need to be trained. In the classification stage, a voting strategy is used where the testing point is designated to be in a class having the maximum number of votes.

IV. RESULT ANALYSIS



Figure 4:Output

V. CONCLUSION

The goal to identify leaf diseases was accomplished. The developed system is used for leaf disease identification; there is a need for the development of high-quality classification methods and accurate feature extraction, which is very significant to execute the system in actual operating environment. The accurate detection and classification of the disease in a particular vegetable is very important for the successful cultivation and this can be done

using image processing. This paper also discussed some feature extraction using texture and classification techniques to extract the features and can also detect the affected area, perimeter, eccentricity, entropy, etc., Genetic algorithm is used for segmentation and classification is done by SVM classifier to identify the condition of the leaf.

REFERENCES

- [1] Gouri C. Khadabadi Prof. Arun Kumar Dr. Vijay S. Rajpurohit Dept. of computer science & Eng G. "Identification and Classification of Diseases in Carrot Vegetable Using Discrete Wavelet Transform", International Conference on Emerging Research in Electronics, Computer Science and Technology-2015.
- [2] Gouri C. Khadabadi#1, Vijay S. Rajpurohit*2, Arun Kumar*3, V. B. Nargund*4 Computer Science and Engineering Department, Gogte Institute of Technology, "Disease Detection in Vegetables Using Image Processing Techniques" International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 14 Issue 2 -APRIL 2015
- [3] Mukesh Kumar Tripathi, Dr. Dhananjay D. Maktedar "Recent Machine Learning Based Approaches for Disease Detection and Classification of Agricultural Products" Conference of Department of Computer Engineering, August 2016
- [4] Johnny L. Miranda, Bobby D. Gerardo, and Bartolome T. Tanguilig III "Pest Detection and Extraction Using Image Processing Techniques" International Journal of Computer and Communication Engineering, Vol. 3, No. 3, May 2014
- [5] Hulya Yalcin, Visual Intelligence Laboratory Electronic & Communication Engineering Department "Phenology Monitoring Of Agricultural Plants Using Texture Analysis"
- [6] Min Zhang, Qinggang Meng "Citrus canker detection based on leaf images analysis" 2010 IEEE
- [7] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad "A Framework for Detection and Classification of Plant Leaf and Stem Diseases" 2010 IEEE.
- [8] Jun Pang, Zhong-ying Bai, Jun-chen Lai, Shao -kun Li "Automatic Segmentation of Crop Leaf Spot Disease Images by Integrating Local Threshold and Seeded Region Growing" 2011 IEEE
- [9] Hong-ning Li, Jie Feng, Wei-ping Yang, Xiang-sheng Wu, Ze-dong Li, Wei Liu "Spectrum-based Method for Quantitatively Detecting Diseases on Cucumber Leaf" 2011 IEEE
- [10] Hashim H.; Haron M. A.; Osman F. N; Al Junid, S. A. M "Classification of Rubber Tree Leaf Disease Using Spectrometer" 2010 IEEE

