

Classification of Fungal Disease Symptoms affected on Cereals using Color Texture Features

Jagadeesh D. Pujari¹, Rajesh Yakkundimath² and Abdulmunaf S. Byadgi³

¹*S.D.M.College of Engineering & Technology Dharwar – 580 008, INDIA*

²*KLE.Institute of Technology Hubli – 580 030, INDIA*

³*University of Agricultural Sciences, Dharwar – 580005, INDIA*

¹*jaggudp@yahoo.com, ²rajeshymath@gmail.com, ³byadgi_uas@rediffmail.com*

Abstract

This paper describes Support Vector Machine (SVM) and Artificial Neural Network (ANN) based recognition and classification of visual symptoms affected by fungal disease. Color images of fungal disease symptoms affected on cereals like wheat, maize and jowar are used in this work. Different types of symptoms affected by fungal disease namely leaf blight, leaf spot, powdery mildew, leaf rust, smut are considered for the study. The developed algorithms are used to preprocess, segment, extract features from disease affected regions. The affected regions are segmented using k-means segmentation technique. Color texture features are extracted from affected regions and then used as inputs to SVM and ANN classifiers. The texture analysis is done using Color Co-occurrence Matrix. Tests are performed to classify image samples. Classification accuracies between 68.5% and 87% are obtained using ANN classifier. The average classification accuracies have increased to 77.5% and 91.16% using SVM classifier.

Keywords: *Fungal disease, Color Co-occurrence matrix, Segmentation, Pattern Recognition, Early detection*

1. Introduction

Today India ranks second world wide in farm output. Agriculture is still the largest economic sector and plays a major role in socio-economic development of India. Agriculture in India is the means of livelihood of almost two thirds of the workforce in India. India has over 210 million acres of farm land. Jowar, wheat, maize, are the major cereals. Apple, banana, sapota, grapes, oranges are the most common fruits. Sugarcane, cotton, chili, groundnuts are the major commercial crops. Crop cultivation depends on rainfall, quality of the soil and climatic conditions and short of any one of these leads to loss of crop. Diseases are major for loss of crop every year and really it is a challenge to control the diseases.

Plant disease diagnosis is an art as well as science. The diagnostic process (*i.e.*, recognition of symptoms and signs), is inherently visual and requires intuitive judgment as well as the use of scientific methods. Plant diseases reduce both quantity and quality of plant products. The prime objective of plant pathology is to prevent epidemic which are widespread outbreak of destructive diseases. Knowledge of different disease causing pathogen and their control is very essential in order to prevent the epidemics of the disease. Farmers are very much concerned about the huge costs involved in disease control activities and it causes severe loss. The cost intensity, automatic correct

identification and classification of diseases based on their particular symptoms become essential and very useful to farmers and also agriculture scientists. 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. As such, several important decisions regarding safe practices, the production and processing of plant have been made in the recent past. One of the main concerns of scientists is the automatic disease diagnosis and control [3, 7].

Computer vision systems would help to tackle the problem. Computer vision systems developed for agricultural applications, namely detection of weeds, sorting of fruits in fruit processing, classification of grains, recognition of food products in food processing, medicinal plant recognition, *etc.* In all these techniques, digital images are acquired in a given domain using digital camera and image processing techniques are applied, on these images to extract useful features that are necessary for further analysis.

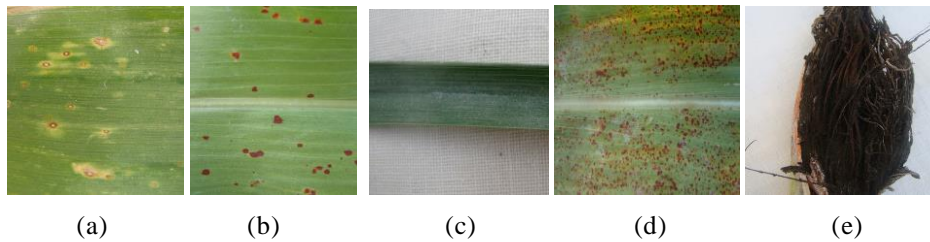


Figure 1. Images Showing the Visual Symptom caused by: (A) Leaf Blight; (B) Leaf Spot; (C) Powdery Mildew; (D) Leaf Rust; (E) Smut

Plant diseases are caused by bacteria, fungi, virus, *etc.*, of which fungi are responsible for a large number of diseases in plants. In the proposed work, we have focused on recognition of fungal disease from the visual symptoms and classify them using image processing and pattern recognition techniques. Figure 1 shows image samples affected by fungal disease symptoms. The paper is organized into four sections. Section 2 gives materials and methods used. Section 3 describes results and discussion. Section 4 gives conclusion of the work.

2. Materials and Methods

This work implements a machine vision system for the classification of the visual symptoms of fungal disease. In the present work, tasks like image acquisition, segmentation, feature extraction and classification are carried out. The classification tree is shown in Figure 2. The detailed block diagram of adopted methodology is shown in Figure 3.

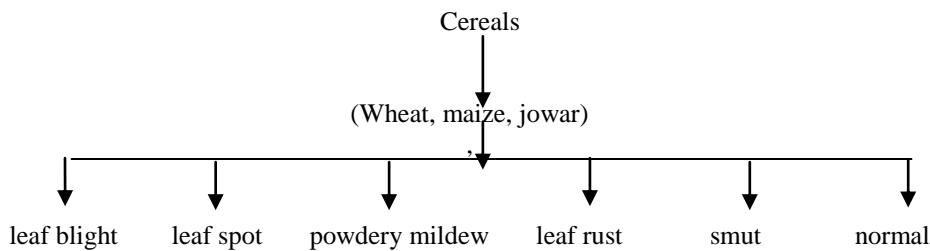


Figure 2. Classification Tree

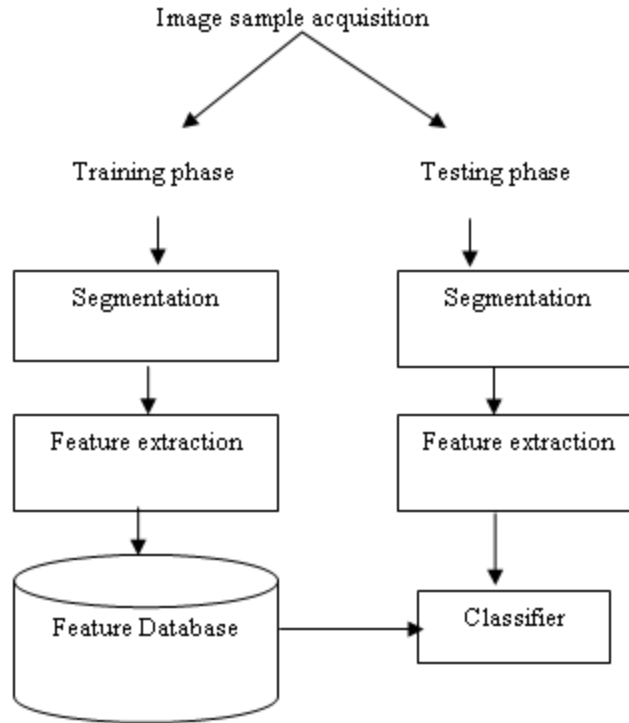


Figure 3. Proposed Block Diagram for Classification of Fungal Disease

2.1. Image Set

The set of 750 images both fungal affected and normal, used in this work was obtained from department of plant pathology, at the University of Agricultural Sciences, Dharwad, INDIA. We have considered 125 image samples of normal type and 125 image samples of each fungal affected symptoms namely leaf blight, leaf spot, powdery mildew, leaf rust, and smut amounting to a total of 750 image samples. The fungal symptoms affected on cereals like jowar (*Sorghum bicolor*), wheat (*Triticum Spp.*), maize (*Zea mays*) are considered for the study. In all cases, the image format used was JPEG, 24 bits

2.2. Image Preprocessing

The single cereal image is captured by analog camera. Then preprocessing steps are applied over image. The preprocessing of image includes shade correction, removing artifacts, formatting. Formatting deals with storage representation and setting the attributes of the image.

2.3. Segmentation

In computer vision, segmentation refers to the process of clustering the pixels with certain properties into salient regions and these regions correspond to individual surfaces, objects or natural parts of the objects. We proposed k-means segmentation technique to segment target regions [8]. Target regions are those areas in the image that represented visual symptoms of a fungal disease. Figure 4 shows segmented images obtained from k-means segmentation technique. The Algorithm.1 gives the steps involved in k-means segmentation technique.

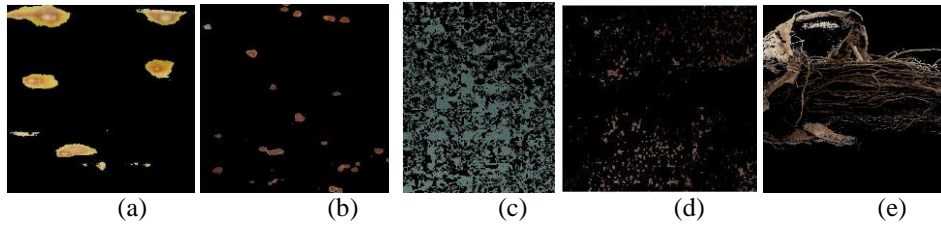


Figure 4. Fungal Disease Symptom Segmented using K-means: (a) Leaf Blight; (b) Leaf Spot; (c) Powdery Mildew; (d) Leaf Rust; (e) Smut

Algorithm 1: k-means clustering technique based segmentation

Input: RGB image.

Output: Segmented image.

Start

Step 1: Select the value of k depending upon the types of image samples being clustered.

Step 2: Assign each pixel to the group that has the closest centroid i.e. nearest cluster center.

Step 3: When all items have been assigned, recalculate the positions of the k centroids.

Step 4: Repeat steps 2 and 3 until the convergence is attained i.e. no pixels change the clusters.

Stop.

2.4. Feature Extraction

The image analysis technique is done using Color Co-occurrence Matrix (CCM). The fungal affected areas vary in color and texture and are dominant in classifying fungal symptoms. So, we have considered both color and texture features for recognition and classification purpose. Image texture, defined as a function of the spatial variation in pixel intensities (gray values). The use of color features in the visible light spectrum provided additional image characteristic features over traditional gray-scale representation. CCM is a method in which both color and texture features are taken into account to arrive at unique features which represent that image. The CCM method involved three major steps. First, transformations of an RGB (Red, Green, Blue) color representation of an image to an equivalent HSI (Hue, Saturation, Intensity) color representation. Once this process is completed Color Co-occurrence Matrices from the HSI pixels is generated. Lastly, texture features from the CCM matrices are generated [8, 14].

From the original image, RGB components are separated and the Hue (H), Saturation(S) and Intensity (I) components are extracted. The Equations (1), (2), (3) and (4), are used to obtain Hue, Saturation and Intensity parameters of the image samples.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \quad (1)$$

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \quad (2)$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)] \quad (3)$$

$$I = \frac{1}{3} (R+G+B) \tag{4}$$

Each pixel map is used to generate a Color Co-occurrence Matrix, resulting in three CCM matrices, one for each of the H, S and I pixel maps. The Color Co-occurrence texture analysis method was developed through the use of spatial gray level dependence matrices (SGDM's). The Gray Level Co-occurrence Method (GLCM) is a statistical way to describe shape by statistically sampling the way certain grey-levels occur in relation to other grey-levels[15]. The GLCM is based on the repeated occurrence of some gray-level configuration in the texture. This method measures occurrence of gray levels between a specific position $P(i, j)$ in the image and a neighboring pixel, according to a given distance d and direction θ . The CCM matrices are then normalized using the Equation 5.

$$P(i, j) = \frac{p(i, j, 1, 0)}{\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} p(i, j, 1, 0)} \tag{5}$$

Where,

$P(i, j)$ = the image attribute matrix

$P(i, j, 1, 0)$ represents the intensity co-occurrence matrix

Ng = total number of intensity levels

The hue, saturation and intensity CCM matrices are then used to generate the texture features [16]. The resulting 13 texture features for each HSI component and thereby a total of 39 CCM texture statistics. We have found that only 5 features contribute as discriminating features as this is essential for better classification. Measurements that are possible to estimate via the co-occurrence matrix are: energy, entropy, homogeneity, contrast and correlation. The Equations (6), (7), (8), (9) and (10) are used to evaluate the textural features. The color texture feature values for each fungal affected symptom are shown in Figure 5. The procedure involved in texture feature extraction is given in Algorithm.2.

$$\text{Energy} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p^2 d(i, j) \tag{6}$$

$$\text{Entropy} = - \sum_{i,j} P(i, j) \log P(i, j) \tag{7}$$

$$\text{Homogeneity} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{p_d(i, j)}{1 + |i - j|} \tag{8}$$

$$\text{Contrast} = \sum_{n=0}^{Ng-1} n^2 \sum_{|i-j|} P d(i, j) \tag{9}$$

$$\text{Correlation} = \frac{\sum_i \sum_j p^2_d(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (10)$$

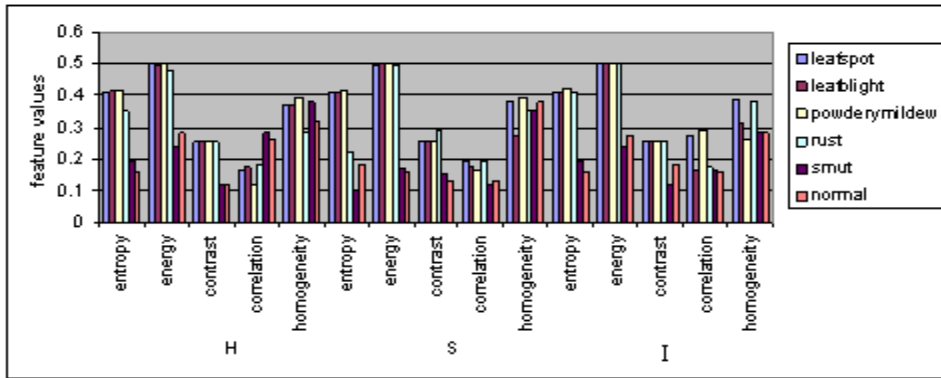


Figure 5. Texture Features

Algorithm 2: Color texture feature extraction

Input: Segmented image (output of Algorithm.1)

Output: Color texture features

Step 1: Convert RGB to HSI translation using Equations (1) thru. (4)

Step 2: SGDM Matrix Generation for H, S and I

Step 3: Calling the GLCM function to calculate the texture features using Equations (6) thru. (10)

Step 4: Texture Statistics Computation

Stop.

2.5. Classifiers

2.5.1. Support Vector Machine: Support Vector Machine comprise of a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors in an n -dimensional space, SVM constructs a separating hyperplane in the space, one which maximizing the margin between the given two data sets. To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the neighboring data points of both classes, since in general the larger the margin the better the generalization error of the classifier. Classifying data is a common need in machine learning. Suppose some given data points belong to one of two classes and the goal is to decide, which class a new data point will be in. In the case of support vector machines, a data point is viewed as a p -dimensional vector (a list of p numbers), and we want to know whether we can separate such points with a $(p-1)$ dimensional hyperplane. This is called a linear classifier. The classifier is also known as a maximum margin classifier.

2.5.2. Artificial Neural Network: An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Multilayer feed-forward neural networks are the most commonly used neural networks for object identification and classification. The layers of neurons between the inputs and the "output layer" are called

“hidden layer”. Back propagation neural network (BPNN) is the most important algorithm for the supervised training of multilayer feed-forward ANN. The BPNN are simple and effective to implement and found suitable for a wide range of machine learning applications. The number of neurons in the input layer corresponds to the number of input features and the number of neurons in the output layer corresponds to the number of classes. The number of nodes in the hidden layer is calculated using the Equation 11.

$$n = \frac{(I+O) + y^{0.5}}{2} \quad (11)$$

Where n = number of nodes in hidden layer,
 I = number of inputs feature,
 O = number of outputs, and
 y = number of inputs pattern in the training set.

The steps involved in recognition and classification of fungal symptoms affected on cereals is given in Algorithm 3.

Algorithm 3: Recognition and Classification of fungal affected symptoms

Input: Color texture features

Output: Classified fungal affected image samples

Start

Step 1: Accept the fungal affected cereal images

Step 2: Extract color texture features using Algorithm 2.

Step 3: Train the SVM and ANN with extracted features

Step 4: Accept test images and perform Step 2

Step 5: Recognize and classify image samples using SVM and ANN classifiers.

Stop.

4. Results and Discussion

All the algorithms used in this work are implemented in MATLAB. In this work, we have used the OSU-SVM toolbox available in MATLAB 7.0. The core of this toolbox is based on Dr. Lin's Lib SVM version 2.33. It is developed by Junshui Ma, Los Alamos National Lab and Yi Zhao, EE department, Ohio State University. For classification using Neural Networks, the Neural Network Toolbox available in MATLAB 7.0 was used. The SVM and ANN are trained with 50 images of each type. The remaining 75 images are used for testing. Around 15% image samples are used for validation of the designed classifier model.

The percentage accuracy is defined as the ratio of correctly recognized image samples to the total number of test image samples. The Percentage accuracy is given by Equation 12.

$$\text{Percentage accuracy (\%)} = \frac{\text{Correctly Recognized Image Samples}}{\text{Total Number of Test Image Samples}} * 100. \quad (12)$$

From Table 4 the highest recognition and classification accuracy of 90.16% is observed with Hue, Saturation (HS) and the lowest of 77.5% is observed with Intensity (I) using SVM classifier.

Table 4. Identification Efficiency based on each Feature Set using SVM

Color features	Fungal disease Symptoms						Average
	Leaf Spot	Leaf Blight	Powdery Mildew	Leaf Rust	Smut	Normal	
HS	87	85	89	93	96	91	90.16
H	82	73	85	89	96	94	86.5
S	75	71	81	83	86	79	79.16
I	79	73	74	81	75	83	77.5
HSI	80	78	82	91	93	88	85.33

From Table 5 the highest recognition and classification accuracy of 87% is observed with Hue, Saturation (HS) and the lowest of 68.5% is observed with Intensity (I) using ANN classifier.

Table 5. Identification Efficiency based on each Feature Set using ANN

Color features	Fungal disease Symptoms						Average
	Leaf Spot	Leaf Blight	Powdery Mildew	Leaf Rust	Smut	Normal	
HS	81	89	79	90	89	94	87
H	80	78	82	89	90	90	84.83
S	69	67	83	79	81	72	75.16
I	80	65	56	76	70	64	68.5
HSI	82	80	74	85	90	82	82.16

The classification results using SVM and ANN classifiers are given in Table 6. It indicates that SVM achieves a significant improvement in classification accuracy over ANN. SVM proved to be a powerful tool for automatic classification of fungal affected symptoms considered in this work.

Table 6. Classification Results of Fungal Disease Symptoms

Classifier	Minimum Accuracy	Maximum Accuracy	Average Accuracy
SVM	77.5	90.16	83.83
ANN	68.5	87	77.75

Confusion matrix of fungal disease symptoms for HS features is given in Table 7 along with precision, recall and F-measure respectively.

Table 7. Confusion Matrix across Fungal Disease Symptoms using HS Color Features with SVM

Fungal disease Symptoms	Leaf Spot	Leaf Blight	Powdery Mildew	Leaf Rust	Smut	Normal	Precision	Recall	F Measure
Leaf Spot	56	2	1	6	0	0	0.85	0.74	0.79
Leaf Blight	0	55	0	2	0	9	0.83	0.73	0.77
Powdery Mildew	7	0	59	2	0	0	0.89	0.78	0.83
Leaf Rust	3	1	1	60	0	0	0.90	0.80	0.84
Smut	0	0	0	0	63	1	0.95	0.84	0.89
Normal	0	2	0	4	0	62	0.93	0.82	0.87

Table 8 shows computation time required for the proposed work using the SVM and ANN classifiers. It shows that SVM has 9% speedup over the ANN classifier.

Table 8. Average Computation Time in Seconds

Color features	ANN	SVM
HS	432.13	383.85
H	359.62	331.53
S	377.24	356.98
I	401.12	377.45
HIS	450.67	439.86
Average	404.156	377.934

5. Conclusion

The technology leverage farmers can take up to assess the crop, look at the possibility of diseases at early stages take decision on possible treatment, and the like. The identification of the symptoms affected by fungal disease, by means of a machine vision system may support farmers in proper assessment of crops. Here we used image samples of cereals that showed visual symptoms of a fungal disease. These diseased regions were identified and segmented using k-means segmentation. Color texture features were extracted from each segmented region and used as inputs to a SVM and ANN classifiers. The performance of SVM classifier found to be better than ANN classifier for the work done

The work carried out has relevance to the real world classification of crop disease and it involves both image processing and pattern recognition techniques. For future study, different neural network architectures can be used for classification. We can extend this project to classify disease symptoms affected on fruits, vegetables, commercial crops *etc.*

Acknowledgement

We would like to express our thanks to Dr.B.S.Anami, principal, KLE.Institute of Technology, Hubli, INDIA, for his valuable suggestions.

References

- [1] J. D. Pujari, R. Yakkundimath and A. S. Byadgi, "Grading and Classification of anthracnose fungal disease in fruits", International Journal of Advanced Science and Technology, vol. 52, (2013).
- [2] H. Patel and S. Dave, "An application of Radon and Wavelet Transforms for Image Feature Extraction", International Journal of Electronics and Communication Engineering, vol. 1, no. 2, (2012), pp. 1-8.
- [3] J. K. Patil and R. Kumar, "Advances in image processing in image processing for detection of plant diseases", Journal of Advanced Bioinformatics Applications and Research, vol. 2, no. 2, (2011), pp. 135-141.
- [4] N. A .Lili, F. Khalid, and N. M. Borhan "Classification of Herbs Plant Diseases via Hierarchical Dynamic Artificial Neural Network after Image Removal using Kernel Regression Framework", International Journal on Computer Science and Engineering, vol. 3, no. 1, (2011).
- [5] D. Moshou, C. Bravo , R. Oberti , J.S. West , H. Ramon , S. Vougioukas , and D. Bochtis, "Intelligent multi-sensor system for the detection and treatment of fungal diseases in arable crops", Biosystems Engineering, pp.311-322, (2011).
- [6] D. S. Guru, P. B. Mallikarjuna, S. Manjunath, "Segmentation and Classification of Tobacco Seedling Diseases", COMPUTE '11 Proceedings of the Fourth Annual ACM Conference, Bangalore, (2011).
- [7] B. S. Anami, J. D. Pujari and R. Yakkundimath, "Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction", International Journal of Computer Applications in Engineering Sciences, vol. 1, no. 3, (2011).
- [8] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, vol. 17, no. 1, (2011).

- [9] M. Miciak, "Radon Transformation and Principal Component Analysis Method Applied in Postal address recognition task", *International Journal of Computer Science and Applications*, vol. 7, no. 3, (2010), pp. 33 – 44.
- [10] A. Camargo, and J.S. Smith, "Image pattern classification for the identification of disease causing agents in plants", *Computers and Electronics in Agriculture*, vol. 66, (2009), pp. 121–125.
- [11] Q. Yao, Z. Guan, Y. Zhou, J. Tang, Y. Hu and B. Yang, "Application of support vector machine for detecting rice diseases using shape and color texture features", *International Conference on Engineering Computation*.
- [12] G. Ying, L. Miao, Y. Yuan and H. Zelin, "A Study on the Method of Image Pre-Processing for Recognition of Crop Diseases", *International Conference on Advanced Computer Control*, (2008).
- [13] K. Y. Huang, "Application of Artificial Neural Network for detecting Phalaenopsis seedling diseases using color and texture features", *computers and electronics in agriculture*, vol. 57, (2007), pp. 3–11.
- [14] R. Pydipati, T. F. Burks and W. S. Lee, "Identification of citrus disease using color texture features and discriminate analysis", *Computers and Electronics in Agriculture*, vol. 52, no. 1-2, (2006), pp. 49-59.
- [15] S. A. Shearer, and R. G. Holmes, "Plant identification using color co-occurrence matrices", *Transactions of the ASAE*, vol. 33 no. 6, (1990), pp. 2037-2044.
- [16] R. M. Harlick, K. Shanmugam and H. Dinstein, "Texture features for image classification", *IEEE Transactions on systems, man and cybernetics*, vol. 3, (1973), pp. 610-621.
- [17] A. K. Jain, "Fundamentals of Digital Image Processing", Englewood Cliffs, NJ: Prentice-Hall, (1989).
- [18] R. C. Gonzalez," *Digital Image Processing using MATLAB*", Second Edition. Prentice Hall, New Jersey, (2003).