



## Medicinal Plants Disease Identification Using Canny Edge Detection Algorithm, Histogram Analysis and CBIR

Vinita Tajane

Computer Science Engineering  
Godwana University, India

Prof. N.J. Janwe

Computer Technology  
Gondwana University, India

*Abstract— Ayurveda certainly brings substantial revenue to India by foreign exchange through export of ayurvedic medicines. Plants diseases cause major production and economic losses in agricultural and medicinal industry worldwide. Monitoring of health and detection of diseases in plants and trees is critical issue. This paper presents a method for identification of medicinal plants based on some important features extracted from its leaf images. The most significant part of research on plant disease to identify the disease based on CBIR (content based image retrieval) that is mainly concerned with the accurate detection of diseased medicinal plant. This paper presents an approach where the plant is identified based on its leaf features such as color histogram and edge histogram. Canny edge detection is also very useful to find the strong edges of leaf of plants and that is used to draw the edge histogram which is one of the parameter for testing. Edge Histogram is plotted on strong edges after applying canny edge Detection Algorithm. Color Histogram separate the layers to plot the red, green and blue layer histogram to check the intensity of each color pixels in the sample image which is another parameter for testing that image is healthy or infected.*

*Keywords— Image processing, CBIR, Canny Edge Detection Algorithm, Ayurveda Medicinal System, Edge histogram, Color Histogram*

### I. INTRODUCTION

Medicinal plants form the backbone of a system of medicine called Ayurveda and is useful in the treatment of certain chronic diseases. Ayurveda is considered a form of alternative to Allopathic medicine in the world. This system of medicine has a rich history. Ancient epigraphic literature speaks of its strength. Ayurveda certainly brings substantial revenue to India by foreign exchange through export of ayurvedic medicines, because of many countries inclining towards this system of medicine. There is considerable depletion in the population of certain species of medicinal plants. Hence we need to grow more of these plant species in India. This rejuvenation work requires easy recognition of medicinal plants. Plant diseases have turned into a dilemma as it can cause significant reduction in both quality and quantity of agricultural products. Diseases in leaves cause major production and economic losses in agricultural industry. Monitoring of health and detection of diseases in leaves is critical for agriculture and medical system. The most significant part of research on plant disease to identify the disease based on CBIR (Content based image retrieval) that is mainly concerned with the accurate detection of diseased plant. It is important for Ayurveda practitioners and also traditional botanists to know how to identify the medicinal plants through Computers. Based on the color space, histogram, and edge detection techniques, we can able to find the disease of plant. Hence here is a proposal of identification of these plants using leaf edge histogram, color histogram.

The methodology here gives the identification of medicinal plants based on its edge features. The color image is converted to its gray scale equivalent image. From this gray scale image, calculate the edge histogram. Apply canny edge detection algorithm for this purpose. The next information i.e., the area is calculated by the proposed algorithm. The next information is the color of the image which is extracted in the form of the histogram for the overall image. These algorithms are applied for the test image and the database image and difference in edge histogram and color histogram is calculated. Obtain the average value of these parameters. Repeat this process for all the leaf images in the database and calculate the difference in the average value parameter between the test and database image. The test image and database image pair which gives the least values is the correctly identified image in turn the plant. Content based image retrieval (CBIR) offers efficient search and retrieval of images based on their publishing, medicine, architecture, etc.

#### A. Types of Diseases

Plant diseases may be broadly classified into three types. They are bacterial, fungal and viral diseases.

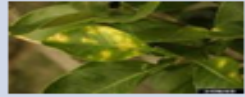


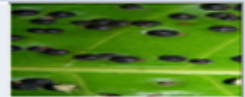
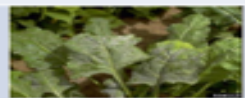

Sr. No	Name Of Plant	Disease	Description	Image
1	Citrus Canker ( <i>Xanthomonas Citri</i> )	Bacteria	Citrus Cankers on Leaves	
2	Melon ( <i>Cucumis melo</i> L.)	Powdery mildew	Powdery mildew infection on cantaloupe leaves.	
3	Banana Freckle	Freckle	It is caused by a fungus, which has three names, <i>Phyllosticta maculata</i> , <i>Phyllosticta musarum</i> and <i>Phyllosticta cavendishii</i> .	
4	Mango Leaf gall midge	Insect	It is caused by insect -Galls on leaves caused by mango leaf gall midge	
5	Sugar Beet	Powdery mildew	Sugar beet leaves showing signs of infection with powdery mildew.	
6	Tomato	Virus	Early infection of tomato yellow leaf curl virus	

Fig 1: Types of Diseases on Plants

#### B. Parameters for Identifying Healthy or Infected Plant

Edge histogram is chosen to find the given plant shape. Since edge detection is used to characterize the plant boundaries, it is considered an important stage in feature extraction and classification.

Variables involved in the selection of an edge detection operator include:

- **Edge orientation:** The geometry of the operator determines a characteristic direction in which it is most sensitive to edges. Operators can be optimized to look for horizontal, vertical, or diagonal edges.
- **Noise environment:** Edge detection is difficult in noisy images, since both the noise and the edges contain high-frequency content. Attempts to reduce the noise result in blurred and distorted edges. Operators used on noisy images are typically larger in scope, so they can average enough data to discount localized noisy pixels. This results in less accurate localization of the detected edges.
- **Edge structure:** Not all edges involve a step change in intensity. Effects such as refraction or poor focus can result in objects with boundaries defined by a gradual change in intensity.

Color histogram helps for finding color distribution of pixels in a plant image. All of the different colors and shades of colors in RGB color are derived from varying combinations of RED, GREEN and BLUE. All of the different colors and shades of colors in RGB color are derived from varying combinations of RED, GREEN and BLUE. The color of each pixel in an RGB digital image is determined by the tonal value (0-255) assigned to each color channels for each pixel. In other words every pixel has a numerical tonal value assigned for each of the three color channels R, G and B.

#### C. Content Based Image Retrieval (CBIR)

The term [CBIR] describes the process of retrieving desired images from a large collection on the basis of features (such as color, texture and shape) that can be automatically extracted from the images. Content based image retrieval (CBIR) offers efficient search and retrieval of images based on their content. In many current applications with large image databases, traditional methods of image indexing have proven to be insufficient. Now days Users are not satisfied with the traditional information retrieval techniques. so content based image retrieval technique is a source of exact and fast retrieval. Conventional databases allow for textual searches on Meta data only. Content Based Image Retrieval (CBIR) is a technique which uses visual contents, normally called as features, to search images from large scale image databases according to users' requests in the form of a query Edges convey essential information to a picture and therefore can be applied to image retrieval. The edge histogram descriptor captures the spatial distribution of edges. The retrieval may involve the relatively simpler problem of finding images with low level characteristics (e.g. finding images of sunset) or high level concepts (e.g. finding pictures containing bicycles). With the development of the Internet, and the availability of image capturing devices such as digital cameras, image scanners, the size of digital image collection is increasing rapidly. This paper is organized as follows; Section 2 introduces canny edge detection algorithm; Section 3 gives the literature review; Section 4 discusses our proposed methods; Section 5 gives the experimental results and section 6 gives the conclusion to the problem

## II. CANNY EDGE DETECTION ALGORITHM

The Canny edge detection algorithm is known to many as the optimal edge detector. Canny's intentions were to enhance the many edge detectors already out at the time he started his work. The canny edge detector first smoothes the image to eliminate and noise. It then finds the image gradient to highlight regions with high spatial derivatives. The algorithm then tracks along these regions and suppresses any pixel that is not at the maximum (nonmaximum suppression). The gradient array is now further reduced by hysteresis. Hysteresis is used to track along the remaining pixels that have not been suppressed. Hysteresis uses two thresholds and if the magnitude is below the first threshold, it is set to zero (made a nonedge). If the magnitude is above the high threshold, it is made an edge. And if the magnitude is between the 2 thresholds, then it is set to zero unless there is a path from this pixel to a pixel with a gradient above T2.

The methodology here is that first picked up the images of healthy and infected plants. Then apply the canny edge detection algorithm on samples. Canny edge detection algorithm preserving the structural properties to be used for further image processing. The purpose of edge detection in general is to significantly reduce the amount of data in an image. The aim of this algorithm with regards to the following criteria:

- i) **Detection:** The probability of detecting real edge points should be maximize while the probability of falsely detecting non edge points should be minimized. This corresponds to maximizing the signal to noise ratio.
- ii) **Localization:** the detected edges should as close as possible to the real edges.
- iii) **Number of responses:** one real edge should not result in more than one detected edge.

### *The Canny Edge Detection Algorithm*

The algorithm runs in 5 separate steps:

- 1. Smoothing:** Blurring of the image to remove noise. Implemented through Gaussian Filtering with Specific Kernel Size (N) and Gaussian Envelope Parameter Sigma.
- 2. Finding gradients:** The edges should be marked where the gradients of the image has large magnitudes.
- 3. Non-maximum suppression:** Only local maxima should be marked as edges. Find gradient direction and using these directions perform non maxima suppression.
- 4. Double thresholding:** Potential edges are determined by thresholding.
- 5.Edge tracking by hysteresis:** Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

## III. LITERATURE REVIEW

H. Al-Hiary, S Bani-Ahmad, M Reyalat, M Braik and Z ALRahamneh presents the plant diseases are detected for few sample only and using the k-map techniques the applications of K-means clustering and Neural Networks (NNs) have been formulated for clustering and classification of diseases that effect on plant leaves. Recognizing the disease is mainly the purpose of the proposed approach. The experimental results indicate that the proposed approach is a valuable approach, which can significantly support an accurate detection of leaf diseases in a little computational effort. Not involving the advantages of hybrid technique Dedicate the future works on automatically estimating the severity of the detected disease [1].

N. Valliammal presents this paper focused on the preprocessing step of CAP-LR. Leaf images normally changes to blurred images by the presence of noise and low or high contrast both in the edge area and image area. In this paper, an approach that simultaneously adjusts contrast and enhances boundaries is presented. As multiple noise damages the quality of nature images, improved enhancement technique is required for improving the contrast stretch in leaf images [2].

S. Ananthi S. Vishnu Varthini presents an application of texture analysis in detecting the plant diseases has been explained. Recognizing the disease is mainly the purpose of the proposed approach. The experimental results indicate the proposed approach can recognize the leaf diseases with little computational effort. Automatically detect the diseases from the symptoms that appear on the plant leaves are the research topic is mentioned in this paper. Extension of this project is to increase the recognition rate of classification process [3].

Arunkumar beyyala and Sai Priya beyyala presents Crop diseases are caused by pathogens fungi, virus and bacteria but there is a problem of complexity of visual pattern. Disease is caused by pathogen which is any agent to cause disease. In most of the cases traits or diseases are seen on the leaves or stems of the plant. Therefore identification of plants, leaves, stems and then finding out the diseases, percentage of the disease incidence. Hence there was an enormous demand for specific and sophisticated image patterns understanding [4].

K. Padmavathi present a study was completed to investigate the use of computer vision and image processing techniques in agricultural applications. In biological science, sometimes thousands of images are generated in a single experiment. These images can be required for further analysis of plant diseases [5].

Pranjali Vinayak Keskar, Shubhangi Nimba Masare, Manjusha Suresh Kadam gives the actual database of the project is stored in the PC The disease affected leaf to be tested is compared with database in the PC which gives result and displays. The developed system consists of four stages which includes HSI transformation, histogram analysis and intensity adjustment. Feature extraction is the third stage which deals with three features namely; color, size and shape of the spots. In precious papers they are not using the CBIR system and simple method for detection of leaf diseases [6].

#### IV. PROPOSED METHODOLOGY

The methodology here gives the identification of medicinal plants based on its edge features. The color image is converted to its gray scale equivalent image. From this gray scale image, calculate the edge histogram. Apply canny edge detection algorithm for this purpose. The next information is the color of the image which is extracted in the form of the histogram for the overall image. Plant diseases cause major production and economic losses in agriculture and forestry. The bacterial, fungal, and viral infections, along with infestations by insects result in plant diseases and damage. The medicinal plants are used in ayurvedic medicines. Manual identification of medicinal plants requires a priori knowledge. The color histograms are obtained in RGB color spaces. This system of medicine is useful in the treatment of certain chronic diseases such as cancer, diabetes, blood pressure, skin problems etc. But, the knowledge of these plants dies with the experts, because of the fact that the experts do not share with others. Hence it is necessary to use technology and develop tools for the recognition and use of medicinal plants from their image. The RGB images of healthy and infected plants will pick up. Apply the CANNY's edge detection technique on RGB image to detect the edges of layered images. Choose the test sample of plant. When the testing sample(healthy or infected) is selected, the training process will started again on the testing image and plot the histogram for the strong edges of testing samples. Histogram is a tool used to show central tendencies. For any statistical data can use the good old parameters of mean, median and mode and can directly find them. Color histogram separate the layers of RGB image into red, green and blue color histogram to check the intensity of each color pixels in that testing image which is helpful for identification of healthy and infected sample.

##### A. System Algorithm

- Step 1: Read test image and database image.
- Step 2: Apply Canny Edge Detection Algorithm on RGB image
- Step 3: Retrieve Strong Edges of image (Healthy or Infected)
- Step 4: Plotted Edge Histogram.
- Step 5: Plotted Color Histogram.
- Step 6: Separate the layers of testing image into Red, Green and Blue Histogram.
- Step 7: Calculate the difference in histograms of both the images (Testing and Database Images).
- Step 8: Results shows the disease of plant by comparing the testing image with database image.

##### B. Proposed Work Flow Diagram

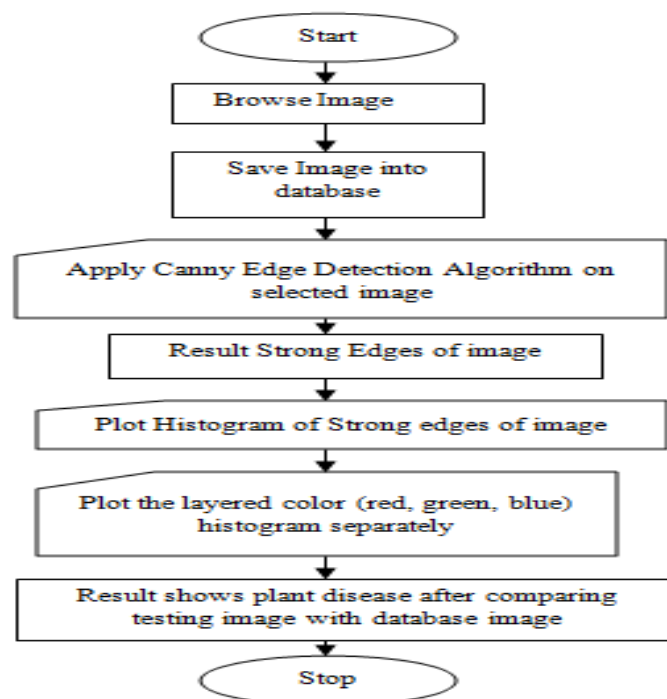


Fig 2: Flow Diagram

##### C. Image Acquisition

The images of leaves of medicinal plants were obtained from a 5 to 12 mega pixel camera and it can be used as per our requirement. The distance between camera and the leaf was maintained to be 15cms and the image was taken from the top view. All the images were taken in natural day light with white background.

##### D. Images Saved to Database:

After capturing the image from camera it has to save in database and apply canny edge detection algorithm.

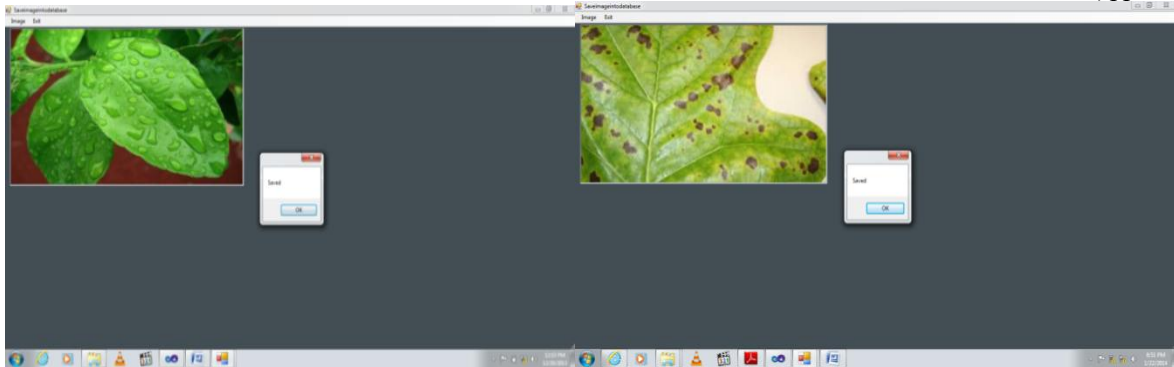


Fig3: Healthy Image Saved into Database



Fig4: Infected Image Saved into Database

The methodology here gives the identification of medicinal plants based on its edge features. The color image is converted to its gray scale equivalent image. From this gray scale image, calculate the edge histogram. Apply canny edge detection algorithm for this purpose.

### E. Histogram

#### i) Edge histogram

Every leaf is having its own edge features. Some leaf boundaries are saw tooth, some are smooth and some are wavy so on. Also midrib alignment and vein pattern of leaves are different.

#### ii) Algorithm:

- Step 1: Start
- Step 2: Acquire the leaf image
- Step 3: Convert Color image to grayscale
- Step 4: Apply Canny edge detection algorithm
- Step 5: Calculate histogram
- Step 6: Stop

#### iii) Color Histogram

Every leaf is having its own color with varying intensity. Some are green; some are blue and red. To check the intensity of layered image color histogram separate the layers of image into red, green and blue layers.

#### iv) Algorithm

- Step 1: Start
- Step 2: Acquire the testing leaf image.
- Step 3: Calculate the green histogram, red histogram and blue histogram separately of the image.
- Step 4: Stop

## V. EXPERIMENTS RESULTS:

The Canny Edge Detector first smoothes the image to eliminate the noise. It then finds the image gradient to highlight regions with high spatial derivatives. The algorithm then tracks along these regions and suppresses any pixel that is not at the maximum (nonmaximum suppression). The gradient array is now further reduced by hysteresis. Hysteresis is used to track along the remaining pixels that have not been suppressed. Hysteresis uses two thresholds and if the magnitude is below the first threshold, it is set to zero (made a nonedge). If the magnitude is above the high threshold, it is made an edge. And if the magnitude is between the 2 thresholds, then it is set to zero unless there is a path from this pixel to a pixel with a gradient above T2.

Draw the Edge histogram of strong edges after applying Canny Edge Detection Algorithm and then draw the color Histogram to separate the red, green blue layer histogram for pixels intensity of each color to find the mean and median value of testing sample which is stored in database.

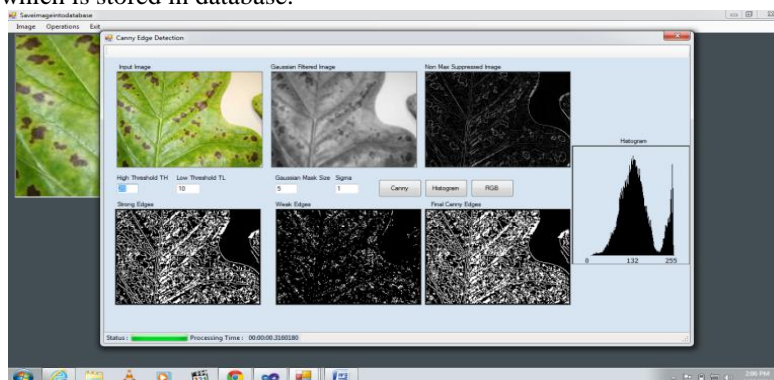


Fig 5: Edge Histogram of Infected Image

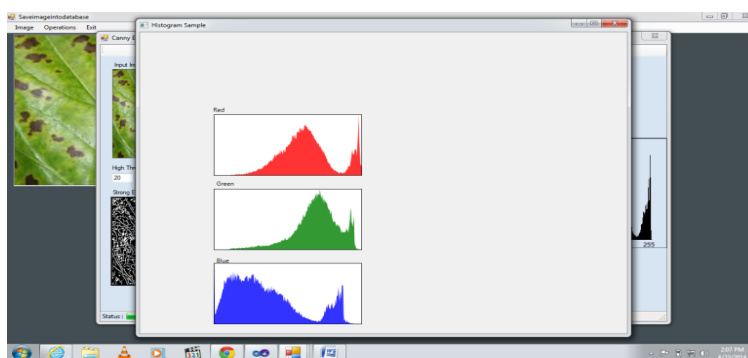


Fig 6: Color Histogram of Infected Image

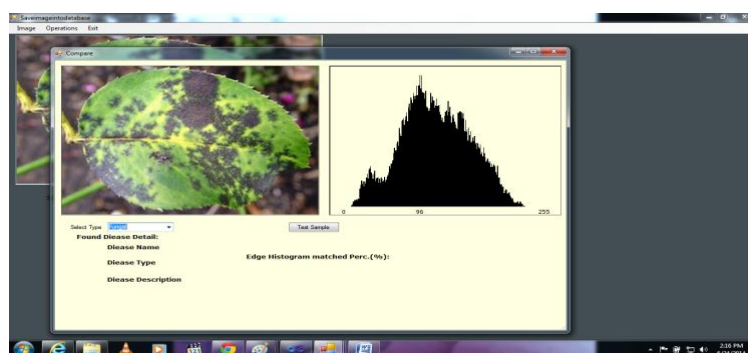


Fig 7: Compared the sample image with Database image for testing

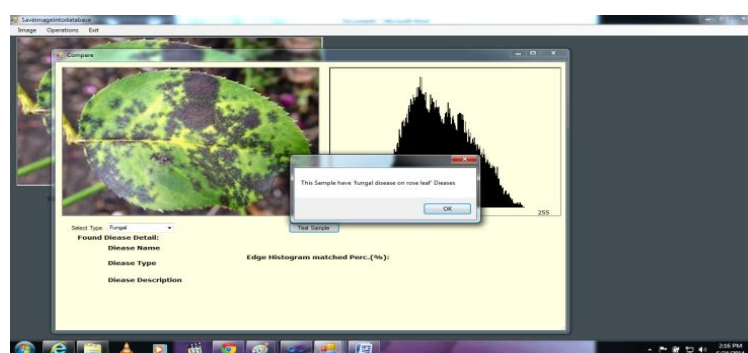


Fig 8: Showing the Fungal disease result after comparing the test sample with Database sample

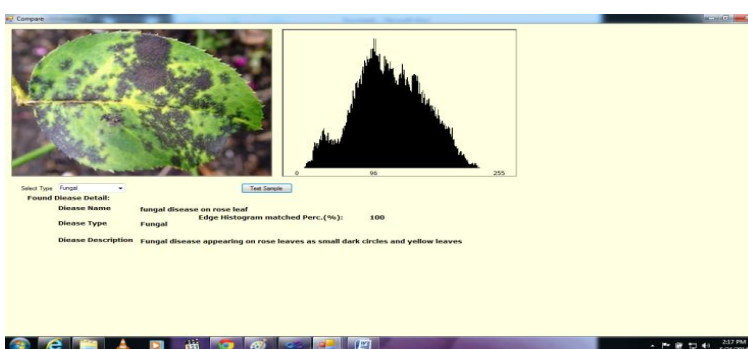


Fig 9: Showing the tested result with 100% accuracy of fungal disease on Rose leaves

## VI. CONCLUSION

To wind up all the information discuss above, I should like to conclude that Canny Edge Detection Algorithm is an efficient and accurate technique for detecting edges of healthy and infected plants and significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an images. Plot the histogram of strong edges of healthy and infected plant after applying Canny edge Detection Algorithm. In training process separate the layers of testing image into Red, Green and Blue layers histogram which is one of the parameter for testing the image with database image. In this paper compare testing sample with the diseased sample and these steps will take few minute to display the comparison result on the basis of mean median values of image histogram and Display which type of disease. The work assists human beings in classification of medicinal plants in the real world and considered an essential task in pharmaceutical industry, Ayurvedic practitioners and botanists.

**ACKNOWLEDGMENT**

Sincere thanks and recognition goes to my advisor, Professor, N.J. Janwe, who guided me throughout this research, inspired and motivated me. We also thank universities for supporting in this research

**REFERENCES**

- [1] H.Al-Hiary, S Bani-Ahmad, M Reyalat, M Braik and Z ALRahamneh ,“Fast and Accurate Detection and Classification of Plant Diseases”, International Journal of Computer Applications, Vol. 17, Issue 1, pp. 31-38, Published by *Foundation of Computer Science* March 2011.
- [2] N.Valliammal , Dr.S.N.Geethalakshmi ,”Hybrid Method for Enhancement of Plant Leaf Recognition”, *World of Computer Science and Information Technology Journal (WCSIT)*, Vol. 1, Issue 9, pp. 370-375, INDIA. 2011.
- [3] S. Ananthi ,S. Vishnu Varthini, “Detection and classification of plant leaf diseases”, *IJREAS* ,Volume 2, Issue 2 , February 2012.
- [4] Arunkumar beyyala1, Sai priya beyyala1, “Application for diagnosis of diseases in crops using image processing”, *International Journal of life Sciences Biotechnology and Pharma Research*, Vol. 1, Issue 2, April 2012.
- [5] K. Padmavathi, “Investigation and monitoring for leaves disease detection and evaluation using image processing”, *International Research Journal of Engineering Science, Technology and Innovation (IRJESTI)*, Vol. 1, Issue 3, pp. 66-70, June 2012.
- [6] Pranjali Vinayak Keskar, Shubhangi Nimba Masare, Manjusha Suresh Kadam , Prof. Mrs. Seema U. Deoghare, ”Leaf Disease Detection and Diagnosis”, *International Journal of Emerging Trends in Electrical and Electronics*, Vol. 2, Issue 2, April-2013.
- [7] Jayamala K. Patil, Raj Kumar,” Advances in image processing for detection of plant diseases”, *Journal of Advanced Bioinformatics Applications and Research* ,Vol. 2, Issue 2, pp 135-141, June-2011.
- [8] Vinita M. Tajane, Prof.N.J.Janwe, “*CBIR in Image Based Medicinal Plant Diseases Retrieval*”, *International Journal of Engineering Research & Technology (IJERT)*, Vol. 2 Issue 11, November 2013.