LEAF FEATURE EXTRACTION AND RECOGNITION

B. M. Azharuddin, R. Revathy, K. Srinivasan, T. Devibala, P. Harish
Asst. Prof, Asst. Prof, III-MCA, II-MCA, II-MCA
Department of MCA,
Vasavi Vidya Trust Group of Institutions, Salem, Tamil Nadu, India

Abstract : Abstract - Plants assume a vital part in our condition. Without plants there will be no presence of the world's biology. Be that as it may, as of late, numerous kinds of plants are at the danger of termination. There are an immense number of plant species around the world. It is likewise important to indentify the plants for their usage in restorative properties and utilizing them as wellsprings of elective vitality sources like bio-fuel. There are a few approaches to perceive a plant, similar to blossom, root, leaf, organic product and so forth. Watching the leaf of a tree or bush is an amazing method to decide the species. Qualities, for example, shape, size, surface and how the leaf is orchestrated can help recognize the kind of plant when natural product has not yet framed. Leaf include extraction on organic product leaf picture is as yet being an issue on programmed plant leaf recognizable proof. In this investigation three diverse leave species to be specific Plectranthus, Scared fig and Bauhinia tomentosa are tested by utilizing picture handling calculations. The plant leaves are procured by advanced camera. The procured pictures are improved by applying middle channel at that point changed over to parallel picture. In the twofold picture edges are recognized utilizing Sobel operator and further the distinguished edges are diminished. Leaf highlights, for example, height, width, centroid, region, border, equli-separation and roundness are ascertained. The last stage is to perceive the leaves highlight utilizing Classifiers. The three classifiers from WEKA apparatus in particular Navie Bayes, Multi Layer Perceptron and J48 are connected. The examination reports 90.32% exactness for characterization under J48.

***IndexTerms* – Feature Extraction, Recognition, Sobel Operator, Naïve Bayes.**

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operation on it, in order to get an enhanced image or to extract some useful information from it. Plants play an important role in our environment. Without plants there will be no existence of the earth's ecology. But in recent days, many types of plants are at the risk of extinction. To protect plants and to catalogue various types of flora diversities, a plant database is an important step towards conservation of earth's biosphere. There are a huge number of plant species worldwide. There are several ways to recognize a plant, like flower, root, leaf, fruit etc. In recent times computer vision methodologies and pattern recognition techniques have been applied towards automated procedures of plant recognition[1].

II. LITERATURE SURVEY

a. REVIEWS ON FEATURE EXTRACTION OF LEAF

Mahdi et al. [2] proposed a new method for measuring a number of leaf dimension parameters including height, width, average width, perimeter and area. Edge detection, filtering and thresholding algorithms are applied to identify the leaf section of the image against the background. Forty leaves that differ in shape and size were used to validate the estimated parameters against the true values and parameters Data indicated that the proposed method achieved a constantly high accuracy.

Sanjay B Patil et al. [3] stated that the size and color of leaf is important factors to categorize the product in market. This includes the area measurement method for Betel leaf based on image processing technique. The results are compared with the results of graphical area measurement technique. To evaluate image processing method and graphical methods known area object is used as reference. It is experimentally proved that this method for measuring sugarcane leaf area is accurate with smallest relative error.

b. REVIEWS ON IDENTIFICATION OF LEAF

Pande Ankita V. et al. [4] reported that Fruit and Flower trees are broadleaf and usually deciduous, meaning they lose their leaves annually in the fall. Knowing what types of fruit trees occur naturally will narrow the list of possibilities when trying to identify a tree. Characteristics such as shape, size, texture and how the leaf is arranged can help to identify the type of plant when fruit has not yet formed and identification of local fruit trees through leaf structures using image processing techniques.

Zalikha et al. [5] stated that the data set consisted of images of 10 different plant species, with different sized leaves. Using grayscale conversion followed by thresholding, the images were converted into binary images, from which the descriptors could be derived. The incorporation of variant images allowed the system to be tested for rotation and scale invariance. A Generalized Regression Neural Network (GRNN) was used for the classification, with classification results showing that features from the TMI were the most effective.

Ananthi et al. [6] proposed the recognition approach that identifies the shape and texture features of the medicinal leaves.

III. IMPLEMENTATION

Automation is an interdisciplinary concept that uses technologies in the computer world to simplify complex issues in other disciplines or in everyday life. This research focuses on using Image processing to automate classification and perform plant recognition based on the images of the leaves.

a. IMAGE ACQUISITION

The database utilized as a part of our investigation is gathered by our self. We taken the leaf from the plant close to our grounds that are 62 leaves of three distinctive plant species is given in Table 1.

Table 1 Leaf Data Set

Leaf Name	Number of Leaves
Bauhinia tomentosa	18
Plectranthus	25
Sacred fig	19

b. GRAY SCALE CONVERSION

In this process, RGB images are converted to grayscale image (grayscale). The formula used to convert the RGB pixel values to gray scale pixel values are as follows:

Gray = 0.2989 * R + 0.5870 * G + 0 .1140 * B.

Where R is red, G is Green and B is blue represent each color pixel.

c. IMAGE FILTER

The fundamental objective of picture upgrade is to process a picture so it seems more adequate or satisfying way. The removal of noise, the sharpening of image edges and blurring effects are all some popular enhancement techniques. These operations can be achieved through the process of filtering.

Filters act on an image to change the values of the pixels depends on the method used. Each of the pixels in an image, the pixel under consideration at a given movement is called as target pixel and it is successively addressed. Filtering operations over an image are performed as a series of local neighborhood operations using a sliding-window-based principle. Filters are generally classified as linear and non-linear. One of the primary uses of both linear and non-linear filtering in image enhancement is for noise removal.

Coding:

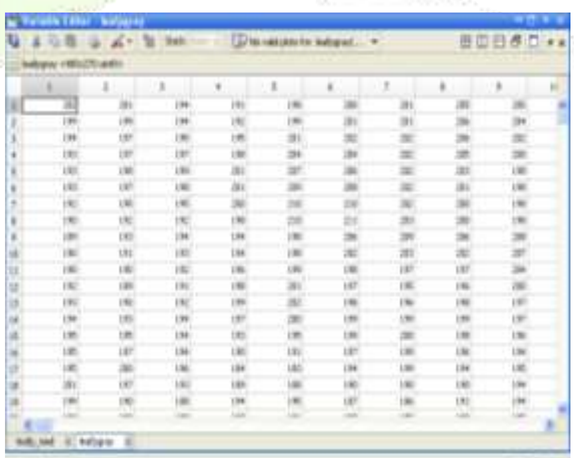
```
leafp_med=medfilt2(leafpgray,[3 3]);
    subplot(1,1,1),imshow(leafp_med),title
            ('median filter')
```

i. Median filtering

Replace each pixel value with the median of the gray values in the region of the pixel:

1. Take a 3x3
2. Sort the intensity values of the pixels in the region into ascending order
3. Select the middle values as the value of pixel (i , j)

The median value is



d. GRAY SCALE IMAGE TO BINARY IMAGE

Taking the threshold as a level to separate the background from the leaf, the values less than threshold is taken as white and the values greater is taken as black. The leaf image is now white color and the background black.

Coding:

```
leafpbw=im2bw(leafp_med,0.3);
subplot(1,1,1),imshow(leafpbw),title('convert to binary');
```

e. EDGE DETECTION

The **Sobel operator** is used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations.

On the other hand, the gradient approximation that it produces is relatively crude, in particular for high frequency variations in the image. The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical.

Coding:

```
leafpes=edge(leafpbw,'sobel');
subplot(1,1,1),imshow(leafpes),title('sobel')
```

f. EDGE THINNING

Line thinning process is a morphological operation used to remove foreground pixels in a binary image. After the edge detection process is carried out, the resulting lines are thick which contains several pixels. Therefore, the line thinning process is carried out to get only one line of pixels. This is required to facilitate the feature extraction process to be carried out later. If the line has a number of pixels, the feature extraction process will experience problems.

Coding:

```
leafpth=bwmorph(~leafpes,'thin','inf');
set(gcf,'position',[2 2 700 700])
subplot(1,1,1),imshow(~leafpth),title('thinning process')
```

g. FEATURE EXTRACTION

After the preprocessing the feature extraction is done. It takes into account only the shape of the leaf and biometric features of the leaf. The following features are extracted from the leaves.

1. **Height and width**

Height and width is calculated under major axis and minor axis of image pixel values.

2. **Area**

The total number of 'ON' pixels in the image.

3. **Centroid**

Find the row and column having pixel value one.

4. **Bounding Box**

- Starting position (x, y), length and breadth are needed.
- Minimum value of row and column minus 0.5 gives starting position(x, y) respectively.
- Minimum value of row=1-0.5=0.5
- Minimum value of column=1-0.5=0.5
- Maximum value of column – minimum value of column+1 gives breadth of the box
- Maximum value of column=4
- Max value-min value of column=3+1
- Maximum value of row- minimum value of row +1gives length of the box
- Maximum value of row=3
- Max value – Min value=2+1

5. **Perimeter**

To find the Perimeter find the boundary of the labeled component

6. **Equiv Diameter**

To find the Equivdiameter**:** sqrt(4*Area/pi).

7. **Roundness**

To find the Roundness: (4*Area*pi) / (Perimeter. ^2)

h. LEAF RECOGNITION

The extracted features are classified using three algorithms in WEKA tool. They are **Naive Bayes, Multilayer Perceptron** and **J48.**

IV. RESULTS AND DISSCUSSION

a. Results using Naive Bayes

The results obtained using Naive Bayes are reported in Table 2 and graphically presented in Figure 1.

Table 2 Results of Naive Bayes

Data set	No. of Leaves	Correctly Classified Leaves	Incorrectly Classified Leaves
Plectranthus	25	17	8
Scared fig	19	4	15
Bauhinia tomentosa	18	16	2

The 'Naïve Bayes' classification shows the overall accuracy for correctly classified instances is 37 with 59.68%, (and incorrectly classified instances is 25 with 40.32%).

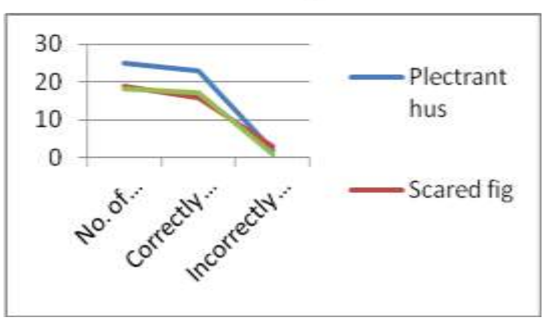


Figure 1 Results of Naive Bayes

b. Result using Multilayer Perceptron

The results obtained using Multilayer Perceptron is reported in Table 3 and is graphically presented in Figure 2.

Table 3 Results of Multilayer Perceptron

Data set	No of Leaves	Correctly Classified Leaves	Incorrectly Classified Leaves
Plectranthus	25	21	4
Scared Fig	19	13	6
Bauhinia Tomentosa	18	18	0

The 'Multilayer Perceptron' classification shows the overall accuracy for correctly classified instances is 52 with 83.87 %.(and incorrectly classified instances are 10 with 16.13 %.)

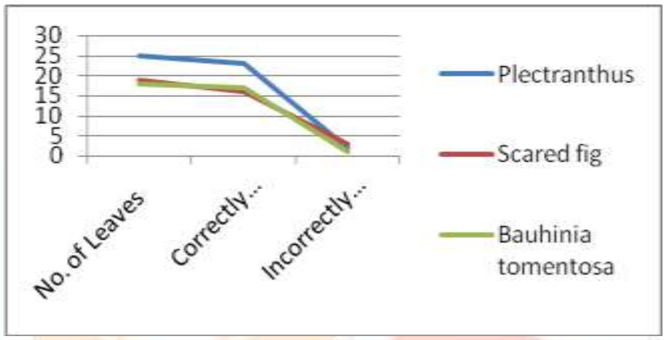


Figure 2 Results of Multilayer Perceptron

c. Results using J48

The results obtained using J48 are reported in Table 4 and is graphically presented in Figure 3

Table 4 Results of J48

Data set	No of Leaves	Correctly Classified Leaves	Incorrectly Classified Leaves
Plectranthus	25	23	2
Scared fig	19	16	3
Bauhinia tomentosa	18	17	1

The 'J48' classification shows the overall accuracy for correctly classified instances is 56 with 90.32 %.(and incorrectly classified instances are 6 with 9.68 %.)

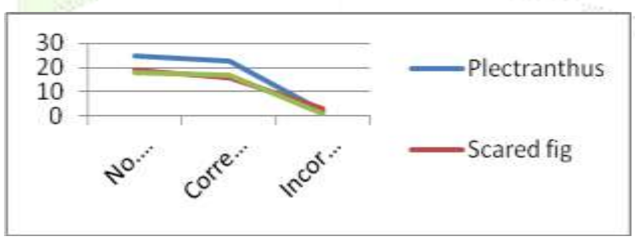
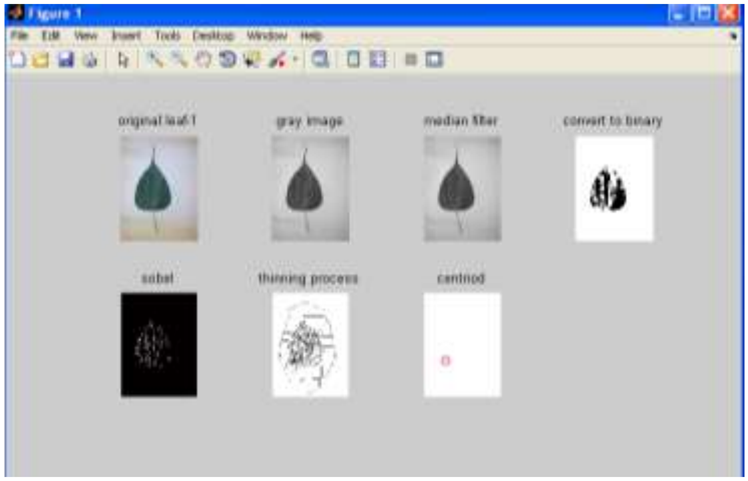


Figure 3 Results of J48

V. COMPARATIVE ANALYSIS

S. No.	Algorithm	Classification Accuracy
1	Naive Bayes	59.68%
2	Multilayer Perceptron	83.87%
3	J48	90.32%

VI. SCREEN SHOTS



VII. CONCLUSION

The photographed images served as a database for this study. The leaves are plectranthus, scared fig and bauhinia tomentosa. The leaf images are enhanced by applying median filter. Then by using thresholding binary image is generated. Edges are detected by sobel operator. Later the edges are thinned. The features available in the literature such as Leaf Height, Leaf Width, Area, Centroid, Perimeter, Equiv-diameter and Roundness are extracted from the thinned leaf images. Three classifiers available under WEKA are used to classify leaves. They are Naïve Bayes, Multilayer Percetron and J48.The highest accuracy obtained is 90.32% using J48, and then the better result is obtained by Multilayer Perceptron at 83.87%. The accuracy at Naïve Bayes is 59.68%.

VIII. BIBILIOGRAPHY

1) Miss. Pande Ankit V, Prof. Shandilya V.K, "Digital Image Processing Approach for Fruit and Flower Leaf Identification and Recognition", Department of computer science and engineering Spina college of engineering and technology Amravati: International Journal of Engineering and Computer Science, Vol. 2, pp. 1280 - 1285, 2013.
2) Mahdi M.Ali, Ahmed Al-Ani, Derek Eamus and Daniel K.Y. Tan, "A New Image-Processing-Based Technique for Measuring Leaf Dimensions", Faculty of engineering and Information Technology, Vol. 12, pp . 1588-1594, 2012.
3) Sanjay B. Patil, "Betel Leaf Area Measurement Using Image Processing", Department of Electronics Engineering: International Journal on Computer Science and Engineering (IJCSE), Vol. 3, pp. 2656 - 2660, 2011.
4) Sandeep Kumar.E, "leaf color, area and edge feature based approach for identification of Indian medicinal plants", Department of Telecommunication Engineering: Indian Journal of Computer Science and Engineering, Vol. 3, pp. 436-442, 2012.
5) Zalikha.Z, Puteh.S, Itaza.S, and Mohtar.A, "Plant identification using moment invariants and general regression neural network." 11th International Conference on Hybrid Intelligent Systems (HIS), pp. 430 - 435, 2011.
6) C.Ananthi, Azha Periasamy and s. Muruganand, "Pattern Recognition of Medicinal Leaves Using Image Processing Techniques", Dept. of Electronics and Instrumentation: journal of nanoscience and nanotechnology, Vol. 2, pp. 13 - 25, 2011.