IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 3, March 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

Plant Identification Using Leaf Images

J.Jeyaseeli, N.Nandhini,Msc.,M.Phil

Student, Dept. of Computer Science, Sakthi College of Arts and Science for Women, Oddanchatram, TamilNadu, India

Assistant professor, Dept. of Computer Science, Sakthi College of Arts and Science for Women, Oddanchatram, TamilNadu, India

ABSTRACT: This paper presents a computer based automatic plant identification system. Out of all available organs of plant, leaf is selected to obtain the features of plant. Five geometrical parameters are calculated using digital image processing techniques. On the basis of these geometrical parameters six basic morphological features are extracted. Vein feature as a derived feature is extracted based on leaf structure. At the first stage leaf images are obtained using digital scanner. Then above mentioned morphological features are extracted which act as input to the classification stage. Recognition accuracy of the proposed algorithm is tested. Accuracy of this algorithm is tested on two different databases and compared. False acceptance ratio and false rejection ratio for both databases is calculated. Total 12 kinds of plants are classified using this algorithm. Dataset consists of 92 images of total 12 plants. This method implements effective algorithm used for plant identification and classification as it is independent of leaf maturity. Proposed method is easy to implement and fast in execution.

I. INTRODUCTION

Plants are important part of our ecosystem. There is huge number of plants existing in nature. Many of them are at the risk of extinction. So it is very necessary to catalogue their identity, features and useful properties. Also there are many medicinal plants which play an important role in Ayurveda. Now a day, whole world is facing various problems like global warming, biodiversity loss, effects of fast urban development, and various environmental damages. Hence there is an urgent need to apply advanced computer vision techniques to obtain the botanical knowledge like plant taxonomy, various features of plant and make this information accessible and useful to different kinds of people like researchers, farmers, botanists, and students [1].

Hence plant identification is the first and important task. There are many plant organs like leaves, flowers, fruits, seeds which can be used for plant identification. In this paper leaves are selected to obtain the features of plant. Because leaf can be easily obtained and scanned and also it consists of more excluding information which is useful for plant classification. These leaf images are sent to computer and then by using image processing tools, leaf features are extracted to identify the plant. Several methods have been introduced for plant identification.

Many of them have selected leaf as characterizing organ. Some of them used shape description method while other deal with colour feature, texture features, vein structure, geometrical features etc. [2]. B.Wang et al. [3] propose shape descriptor method for online plant leaf identification working on mobile platform. They measured the convexity and concavity property of the arches of various levels and obtained the multilevel shape descriptor. S.G.Wu et al. [4] extracted 12 leaf features and orthogonalized them into 5 principal variables using Principal Component Analysis (PCA) and Probabilistic Neural Network (PNN) as a classifier.

To recognize the plant species J. Chaki et al. used Gabor Filter. Convolving Gabor filter with leaf images, real and imaginary parts of resultant signal are obtained [5]. Absolute difference between them is calculated which is used as feature value. In this paper leaf identification algorithm is implemented by extracting the leaf features. Five geometrical parameters of each leaf image are calculated. Using these geometrical parameters six basic features is obtained. One derived feature which is vein feature is extracted. This paper uses Euclidean classifier to classify the leaf images. Advantage of using Euclidean classifier is that it is fast and easy to implement.

Plants are essential for mankind. In particular, herbs have been used as folk medicines by indigenous people since ancient times. Herbs are usually identified by practitioners based on years of experiences through personal sensory or olfactory sense [1]. Recent advances in analytical technology have significantly assisted in herbal recognition based on scientific data. This eases many people, especially those that are lacking experience in herbal recognition. Laboratory-based testing requires skills in sample treatment and data interpretation, in addition to time consuming procedures [2]. Therefore, a simple and reliable technique for herbal recognition is needed. Computation combined with statistical analysis is likely to be a powerful tool for herbal recognition. This nondestructive technique

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

shall be the method of choice to rapidly identify herbs, particularly for those who cannot afford to apply expensive analytical instrumentation.

II. LITERATURE SURVEY

Image Processing in Leaf Pattern Recognition

Leaf pattern recognition usually follows the steps as shown in Figure 1. The most challenging part of this study is to extract distinctive features of leaves for plant species recognition. In this case, different classifiers using high performance statistical approaches have been used to perform leaf features extraction and classification. The advancement in computer vision and artificial intelligence have greatly assisted researchers to classify plants through statistical modeling. The pre-processing step consists of image reorientation, cropping, gray scaling, binary thresholding, noise removal, contrast stretching, threshold inversion, and edge recognition.

Image aligning the input image to a standardized position, with the leaf aligned to either the x-axis or y-axis. For leaves that have the greater width: Length ratio, the length is preferably placed in the vertical or upright position [10]. To decrease the amount of computational load that is exerted upon the graphic processing unit, cropping the image is a necessary step to reduce the unnecessary foreground region of the prompt image. Turkoglu and Hanbay [11] suggested that leaf feature extraction could be done by dividing the leaf image into two or four parts, instead of extracting for the whole leaf. The proposed image processing techniques using color, vein, Fourier descriptors (FD), and gray-level co-occurrence matrix (GLCM) methods had proven to achieve 99.1% accuracy using the Flavia leaf dataset. Gray scale conversion of the image into geometrical data is implemented to optimize the contrast and intensity of images.

Later, the thresholding process creates a binary image from the gray scaled image to translate the value of the image to its closest threshold, and therefore having either one of two possible values for each pixel, as presented in Figure 2. Different types of noises, such as grains, and holes, could affect digital images, therefore erosion and dilation are a series of operations implemented in order to remove the background noises. The images are considered homogenous if they do not exhibit substantial differences between one another in terms of contrast stretching. These images, when shown in histogram representation, exhibit very narrow peaks. Inhomogeneity is caused by the lack of uniform lighting upon the image. The image is normalized in order to stretch the narrow range to a more dynamic range. The binary images from the process are inverted during threshold conversion, to convert the background into black. Suzuki algorithm can be utilized to extract the contours of images and further refined by diminishing the contours with small lengths with regards to its largest contour [10]. This process is known as edge recognition.

Ma et al. [12] suggested that their algorithm was better than the conventional back-propagation algorithm in terms of efficiency and accuracy. They analyzed soybean leaf image to evaluate the nitrogen content by introducing median filter in their preprocessing stage. A hazier image is collected to remove grain noise, which could disrupt image processing due to high frequency properties as a result of grey difference [13]. The subsequent step would be emphasizing the leaf edge to obtain a clear image. In this case, the grey linear transformation technique is used to further add the difference in grey saturation between the leaf image and the background, and thus enhance the image by gaining a comparable threshold value of sample image and background to decrease error rate [14]. Since the image background still has an undesired value, the image is binarized to remove the background value completely before the original image is imposed to the processed image [15].

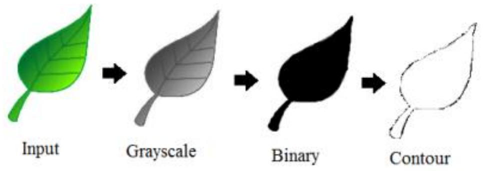


Figure 2. Image pre-processing stage.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

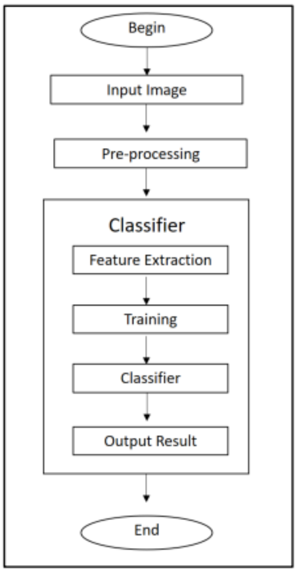


Figure 1. Fundamental of leaf pattern recognition.

III. SYSTEM ANALYSIS

3.1 Existing System

The weight passes between the layers of nodes are established in the early stage, the pre-existing weights will not be manipulated. Nevertheless, new vectors are placed into weight matrices during the training phase which results in real-time viability. During the process of recognizing the leaf pattern, the feature vector of leaf is classified by the network into a particular class since the assigned class is assumed to have the highest probability to be accurate. Classifiers with several features extraction of the leaf images have been shown to yield maximum accuracy output as compared to them individually and all existing methodologies with accuracy of 93.11%

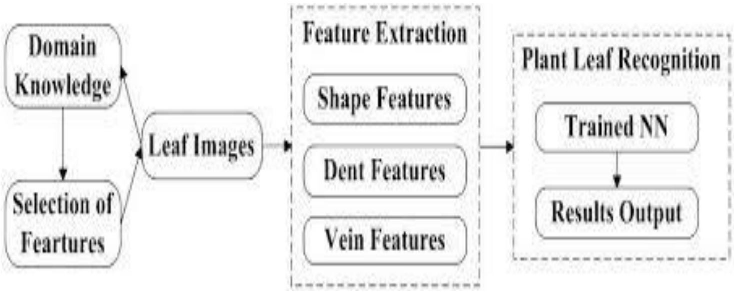
3.2 PROPOSED SYSTEM

The proposed method used 70 samples taken from the Flavia dataset and their geometrical and shape features were extracted and yielded an accuracy of 97.7% which is the highest achievement reported in literature. The results found that the proposed method was able to improve for 11.7% and 4.2% in the scan category and 4.06% and 5.87% for the scan-like category in the average classification score relative to the best results reported in the literature for Image CLEF 2011 and Image CLEF 2012 datasets, respectively. Hence, the multiple classifier system overcame the performance of monolithic methodologies. The computational model was proposed using digital plant images based on biometric features such as shape and vein patterns.

Advantages

- The classifier of SVM also proved to be effective to detect plants in their natural habitat.
- The model could achieve an average accuracy, 97% of leaf identification based on 10 different biometric information extracted from 1907 leaf images of 32 plant species taken from the Flavia dataset.
-

3.3 ARCHITECTURE



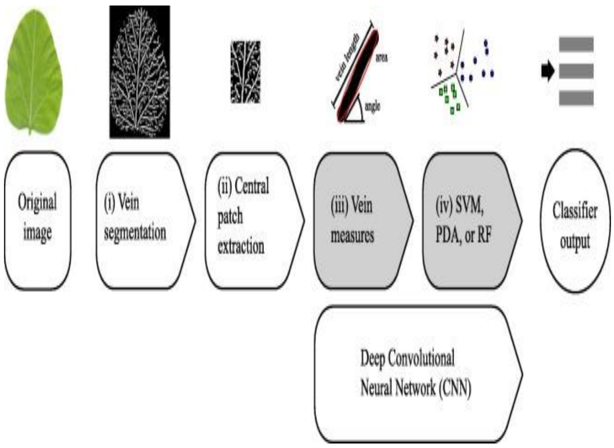
International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

Work Process



IV. SYSTEM IMPLEMENTATION

4.1 SYSTEM METHODOLOGY

A. Leaf Image Acquisition/ Database selection:

In this paper two database of leaf images are used. One directly obtained from [6]. Other one is created by taking leaves from trees available in college campus. All these leaves are scanned using digital scanner. Back side of leaves are used for scanning. As leaf images are scanned rather photographed, there is no problem of angle of photoing and lighting conditions. The overall flow diagram of proposed work is shown in Figure 1.

B. Preprocessing:

The leaf images obtained in database2 are color images. These RGB images are converted into gray scale. Obtained gray scale images are converted into binary by using appropriate level of thresholding. Median filtering is used to reduce the effect of noise. Leaf contour is obtained by using boundary extraction algorithm. The flow diagram of preprocessing step is shown in Figure 2.

C. Geometrical parameter calculation:

In this technique five basic geometrical parameters are calculated. These are length, width, diameter, area and perimeter of leaf. 1) Length of leaf: It is defined as the distance between two extreme points on the main vein. 2) Width of leaf: It is defined as the length of minor axis of leaf contour. 3) Diameter: The diameter is defined as the longest distance between any two points on the margin of the leaf. [4] 4) Leaf area: Leaf area is calculated as counting the number of pixels having binary value 1 inside the leaf margin. [4] 5) Leaf perimeter: It is obtained by counting the number of pixels contributing margin of leaf.

D. Basic feature extraction:

Using five geometrical parameters which are defined above, six basic morphological features are calculated.
1) Aspect ratio: It is defined as the ratio of length to the width of leaf. It is given as, [4] Aspect Ratio = Length/ Width
2) Form factor:
Form factor is used to obtain the difference between circle and leaf. It is calculated as, [4]
Form Factor = 4 π Area/ Perimeter
3) Rectangularity:
Rectangularity feature is used to obtain the similarity between leaf and rectangle. It is calculated as, [4]
Rectangularity= Length* Width/ Area
4) Narrow factor: It is defined as ratio of diameter to the length. It is given as, [4]
5) Perimeter ratio of diameter: It is given as the ratio of leaf perimeter to its diameter. It is given as, [4]
6) Perimeter ratio of length and width: It is given as the ratio of perimeter of leaf to the sum of its length and width. It is given as, [4]

E. Derived feature extraction:

One derived feature is extracted named as vein feature. In this feature vein structure of leaf is obtained. Morphological opening is done on input image with disk-shaped structuring element of radius 1,2,3,4 sequentially.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

Then final obtained image is subtracted from input image. Here vein structure of leaf is obtained. Total number of white pixels in the vein structure is counted. Vein feature is obtained as the ratio of total number of white pixels in the vein structure to the area of leaf image .Vein feature is given as,

Vein Feature = Number of White Pixels/ Area

In this way total seven features are extracted. These seven features are calculated for all training images and training feature set is created

F. Detection:

In detection step input query image that is test image is classified to appropriate class of tree. First query image is acquired then pre-processed it. Five geometrical parameters for this query image are calculated. Using these parameters six morphological features is obtained. Vein feature is calculated by obtaining vein structure of query image. Test feature set is created using extracted features. Euclidean classifier is used which is very simple to implement and fast as compare to other classifiers. Distance between training feature set and testing feature set is calculated. The tree for which minimum specified distance obtained, query image is classified to that particular tree.

4.2 EXPERIMENTATIONS & RESULTS

Experiments are done on two types of database. One obtained from [6] and other which is created from plants available in college campus. First database contains 32 samples of 6 plants, named as b1 to b6. Each image is saved in a TIFF format. Second database contains 60 samples of 6 plants, named as c1 to c6. Each image is saved in a JPEG format. Training and testing data set of first database are shown below in Figure 1 and Figure 2. And of second database is in figure 3 and figure 4.

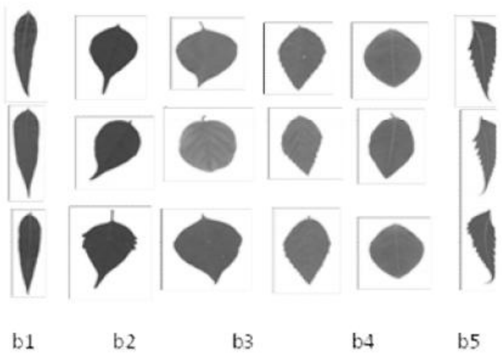


Fig. 3.1 Samples of Dataset1.

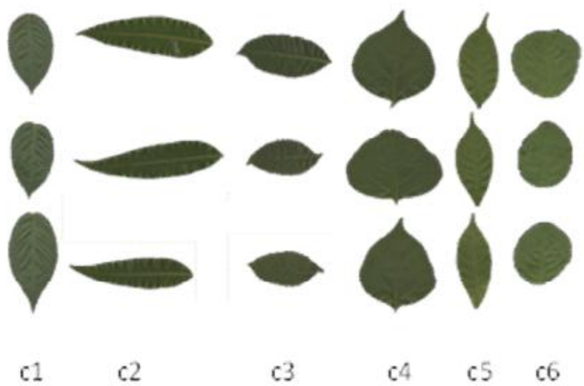


Fig. 3.4 Samples of Dataset2.

G. Analysis:

From first database 2 leaf images per plant and from second database 3 leaf images per plant are used to obtain the accuracy of this algorithm. The accuracy obtained for first database is 78.12 % while for second database it is 85%. Some plants get less accuracy due to inter-species similarity and the intra-species variability and also due to the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

simplicity of algorithm. False Acceptance Rate and False Rejection Rate for Database1 and Database2 are as shown in TABLE I and TABLE II respectively. By observing both tables it is cleared that FAR and FRR values of Database2 are less than Database1.

V. CONCLUSION

This paper introduces statistical method approach for plant identification. Proposed method is verified for two databases consisting of total 12 plants. It uses Euclidean classifier for classifying leaf images. This method is easy to implement and fast in execution compared to other methods. Accuracy is moderate. Accuracy can be increased by adding more statistical features and using other advanced classifier. But due to this complexity of algorithm will increase and it will slow down the execution speed. So this method gives the optimum solution in between speed of recognition and accuracy.

VI. FUTURE WORK

Since the essential of the competitive function is to output the index of the maximum value in an array, we plan to let our algorithm output not only the index of maximum value, but also the indices of the second greatest value and the third greatest value. It is based on this consideration that the index of the second greatest value corresponds to the second top matched plant. So does the index of the third greatest value. Sometimes, maybe the correct plant is in the second or the third most possible plant. We are going to provide all these three possible answers to users. Further more, users can choose the correct one they think so that our algorithm can learn from it to improve its accuracy. Other features are also under consideration. Daniel Drucker from Department of Psychology, University of Pennsylvania, suggested us to use Fourier Descriptors so that we can do some mathematical manipulations later. We are also trying to use other features having psychology proof that is useful for human to recognize things like the leaf, such as the surface qualities [30]. Our plant database is under construction. The number of 6 plants that can be classified will be increased.

REFERENCES

1. O. Mzoughi, I. Yahiaoui, N. Boujemaa and E. Zagrouba, “Advanced tree species identification using multiple leaf parts image queries,” IEEE ICIP 2013.
2. Bhardwaj, M. kaur, “A review on plant recognition and classification techniques using leaf images,” International Journal of Engineering Trends and Technology, Volume 4, Issue 2, 2013.
3. Wang, D. Brown, Y. Gao and J. L. Salle, “Mobile plant leaf identification using smart-phones,” IEEE ICIP 2013.
4. S. G. Wu, F. S. Bao, E. Y. Xu, Y. X. Wang, Y. F. Chang, and Q. L. Xiang, “A leaf recognition algorithm for plant classification using probabilistic neural network,” IEEE ISSPIT, pp. 11– 16, 2007.
5. J. Chaki, and R. Parekh, “Plant leaf recognition using Gabor filter,” International Journal of Computer Applications Volume 56, No.10, October 2012.

[6] Miranda J .L. & Gerardo B. D., & Tanguilig III B. T. (2014). Pest Detection and Extraction Using Image Processing Techniques. International Journal of Computer and Communication Engineering, Volume. 3, No. 3, 189-192.

[7] Guzman J, D, & Peralta E. K. (2008). Classification of Philippine Rice Grains Using Machine Vision and Artificial Neural Networks. IAALD AFITA WCCA 2008 World Conference on Agricultural Information and IT, Volume 2, 241-481.

[8] Wang, Z., Chi, Z., Feng, D., & Wang, Q. (2000). Leaf Image Retrieval with Shape Features. 4th International Conference on Advances in Visual Information Systems. Volume 3, 477-487.

[9] H. Fu, Z. Chi, D. Feng, & J. Song. (2004). Machine learning techniques for ontology-based leaf classification, in IEEE 2004 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China. Volume 2, 201-348.

[10] Timmermans.A. J. M & Hulzebosch.A. A. (1996). Computer vison system for on-line sorting of pot plants using an artificial neural network classifier. Computers and Electronics in Agriculture. Volume 1, 41-55.

[11] Beghin.T, Cope.J. S., Remagnino.P., & Barman.S. (2010). Shape and texture based plant leaf classification. International Conference on Advanced Concepts for Intelligent Vision Systems (ACVIS). Volume 3, 345-353.

[12] Smith T. S. (2006). MATLAB: Advanced GUI Development. Dog Ear Publishing 4010 W.

[13] Higham D. J & Higham N. J. (2005). MATLAB Guide: Second Edition. Library of Congress Cataloging-inPublication Data, Datta,Biswa Nath.

[14] Kattan P. (2009). MATLAB for Beginners: A Gentle Approach: Revised Edition. CreateSpace Independent Publishing Platform.

[15] Singh S. (2005).Pattern Recognition and Image Analysis: Third International Conference on Advances in Pattern Recognition. Scientific Publishing Services, Chennai, India.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903040|

[16] Lukas R., Capungay K. N. (2006). Color Image Processing: Methods and Applications. Bombay,India,Narsosa Publishing House.
[17] John W. & Sons. (2011). Practical Image and Video Processing Using MATLAB. Kluwer Academic Publishers.
[18] Poon T. C., Banerjee P.P. (2001). Contemporary Optical Image Processing with MATLAB. Elsevier`s Optical, Electronic And Control Publishing
[19] Parekh J. C. (2012). Plant Leaf Recognition using Gabor Filter, International Journal of Computer Applications, Volume 56 - Number 10, 89-96.
[20] Chollet, Franc¸ois, 2015, https://blog.keras.io/ building-powerful-image-classification-models-using-very-little-data. html
[21] Andrej Karpathy et al. http://cs231n.github.io/transfer-learning/
[22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition arXiv:1512.03385, Dec. 2015.
[23] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recogni- tion. arXiv preprint arXiv:1409.1556, 2014.