RESEARCH ARTICLE OPEN ACCESS

A Review on Recognition of Plant Species Using Leaf Image through Digital Image Processing

Gautam Saikia [1], Mutum Bidyarani Devi [2]

[1] Computer Science Department, Manipur International University, Manipur

[2] Electronics and Computer Science Department, Manipur International University, Manipur

ABSTRACT

Identification of new plant species is an important aspect for conservation on Earth. Among many morphological features of a plant the leaf morphology is the most popular and effective way to identify a plant species using digital image processing. This paper aims to review and discuss some of the significant literatures and their methodologies used for recognizing plant species using leaf image. Also, the performance of the different techniques used for plant species extraction and classification has been reviewed in this paper.

Keywords — Digital Signal Processing, Leaf Recognition, Feature Extraction, Neural Network.

I. INTRODUCTION

Plants play an important rule on maintaining the balance on Earth's atmosphere. Around 3,10,000–4,20,000 known and unknown plant species are available all over the world [1]. Finding out the correct nomenclature plays a vital role in recognizing where an unknown plant belongs to the ecosystem. But the process of identifying the correct nomenclature is limited among taxonomists. Due to global shortage of taxonomists and the increasing ability of computer vision and machine learning, the demand of automated software tool for recognizing of correct nomenclature of plants has increased. For identification of plants, morphological characters are primary based before molecular diagnostic. Morphological features of leaves regarding perfect nomenclature signified permanent charter when it is digitalized. Therefore, among several traditional ways for recognizing a plant morphological character, most of the scholars and researchers applied morphological based character as leaf in respect to its base, margin, apex, blade, texture and venation system etc. This paper presents a brief overview and literature review on various plant recognition techniques based on leaf image.

II. ANALYTICAL COMPUTER-BASED TECHNIQUES APPLIED IN PLANT SPECIES RECOGNITION

An enormous application has been done on leaf morphology based on plant classification and recognition, but these are still limited approaches for investigation of plant identification system. Relating to various applications a few methods perform satisfactory results in respect to specific samples of leaves. Sometimes it has been observed that two or more plant have similar leaf natures that possess a critical illness during reorganisation of correct nomenclature. In such cases, shape and size of leaf may be rejected but the colour and margin character of the leaf cannot be neglected. The following Fig. 1 represents a typical model of leaf base plant identification system.

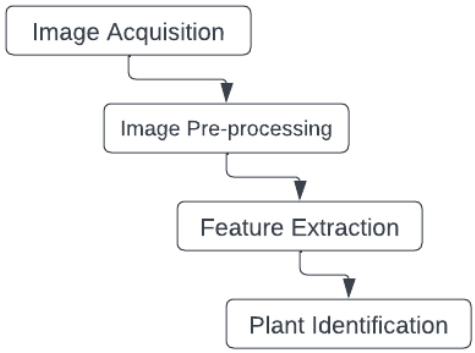

Fig. 1 A flowchart representation a typical leaf base plant identification system

Hati, et al. [2] use Artificial Neural Network (ANN) for recognition of plant based on leaf shape, where new input features and image processing approach has been introduced. During data acquisition phase, for preparation of leaf database of 534 leaves from 20 plants, each leaf images were captured on a clear white background. Due to this new input features, the colour to grayscale conversion of leaf image can be done only by processing the blue band instead of converting RGB value of each pixel of the image to grayscale value by using the following formula [3].

$$grey = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B$$

Aspect ratio, width ratio, apex angle, apex ration, base angle, moment ration and circularity were the seven parameters, which were extracted during feature extraction phase, and constitute the input layer of the artificial neural network. A software tool namely LEAFilia has been developed in Java with three modules for image processing, training of neural network and matching of unknown leaf image with a recognition accuracy of 92%.

Sandeep Kumar. E [4] proposed a simple approach where indian medicinal plant has been identified using area, color histogram and edge histogram of leaf. The leaf image was

obtained using 5 mega pixel camera where 15cms distance has been maintained between leaf and the camera and all the images were captured with white background in natural day light. While considering the edge features of the leaf image, RGB to grayscale conversion has been done and by using Canny edge detection algorithm the edge histogram has been calculated. Also, the area of the leaf was calculated using the same algorithm as well as the colour of the image was also extracted in form of histogram. Three different algorithms have been proposed for image acquisition phase along with a system algorithm which was implemented and tested on MATLAB 7.0.

Satti, V. et al. [5] proposed a leaf recognition system where they considered the leaf color as well as the shape feature for improving the efficiency of recognition which is tested on Flavia dataset consist of 1907 RGB leaves of 33 plants. Neural Network and Euclidean Distance Method has been used to classify the dataset where 93.3% and 85.9% of accuracy has been found, respectively.

Kaur, G. et al. [3] introduce a system for classification of biological species based on leaf architecture which consist of four modules namely – image acquisition, image preprocessing, image recognition and display result. In this work author used the Gray-level co-occurrence matrix (GLCM) for texture feature extraction as well as Artificial Neural Network (ANN) and Multilayer Perceptron Neural Network (MLPNN) for classification, which improved the system accuracy up to 97.9%.

Anubha Pearline S. et al. [6] has conducted a compression study by applying both traditional image processing method and deep learning methods for plant recognition. In traditional approach, local binary patter (LBP), color channel statistics etc. were used during features extraction and during classification, logistic regression, k-nearest neighbour etc. classifiers were used. In deep learning approach, Convolutional Neural Network (CNN) architecture with logistic regression had been used. From the study it has been observed that 82.38% of plant recognition accuracy for Leaf12 dataset has been found while applying traditional approaches whereas 97.14% of accuracy result has been obtained from deep learning approach.

Ji-Xiang Du et al. [7] proposed an efficient computer-aided plant species identification approach based on leaf image where Douglas-Peucker approximation algorithm has been used to form the sequence of invariant attributes from the original leaf image input. Then for recognition of plant leaf, unlike traditional methods, the research proposed a modified dynamic programming (MDP) algorithm for image shape matching. From the experimental result the MDP algorithm has been found more suitable not only for intact leaves but also for distorted, partial and overlapped plant leaves over traditional approaches.

Bisen, D. et al. [8] use deep convolutional neural network for identification of plant species through leaf image. Three main identification steps were undertaken in the proposed method namely image pre-processing, feature extraction and recognition where the convolutional neural network (CNN) classifier consist of convolutional layer, maxpooling layer, dropout layer and fully connected layer. The proposed model found a prediction accuracy of 97% on unknown leaf images.

Hossain J. et al. [9] proposed a plant species recognition technique based on leaf image using probabilistic neural network. During data acquisition phase only the plant with broad flat leaves which were two dimensional in nature were considered, which may be a limitation of the proposed method as it is unable to recognise all type of plant leaves. 91.41% of average recognition accuracy has been found during testing using ten-fold cross-validation technique.

Quach, Boi M. et al. [10] proposed a method to recognise leaf using convolutional neural network, where during pre-processing of leaf image, a refined color image, vein image, xy-projection histogram, handcrafted shape, texture features and Fourier descriptors were extracted and using neural network based encoder all the extracted attributes were transformed into a better representation. Support vector machine (SVM) has been used for classification of pre-processed image. During testing on Flavia leaf dataset, 99.69% of recognition accuracy has been achieved by the proposed method.

Ehsanirad, A. et al. [11] conducted excremental research to illustrate the performance on extraction of leaf texture features of two algorithm namely grey-level co-occurrence matrix (GLCM) and principal component analysis (PCA). Both the algorithms were trained to classify 13 kinds of plant based on 390 leaves where GLCM got a 78.46% of accuracy while PCM accuracy achievement is 98.46%.

Selda, J. D. S. et al. [12] build a portable device with a capability to identify the plant species based on leaf vein, where scale invariant feature transform (SIFT) algorithm has been used for extracting the matrix from the leaf vein image and the support vectors machines (SVM) has been used for classification.

Mouine, S. et al. [13] proposed a new multiscale shape-based approach for leaf image classification where they studied four types of multiscale triangular shape namely Triangle side length representation (TSL), Tringle area representation (TAR), Tringle side lengths and angle representation (TSLA) and Tringle oriented angles (TOA). While testing their system on Swedish image dataset the classification accuracy rate of 95,73%, 90.4%, 96.53% and 95.20% has been found in respect of Triangle side length representation (TSL), Tringle area representation (TAR), Tringle side lengths and angle representation (TSLA) and Tringle oriented angles (TOA) respectively.

Aakif, A. et al. [14] used artificial neural network along with back propagation for develop an algorithm which can be able to automatically classify the plant species based on leaf image. A vector of morphological features along with Fourier descriptors (FD) extracted during feature extraction phase were fed into the artificial neural network classifier for accurate classification of plant species. During testing on Flavia and ICL dataset, the proposed system achieved a classification accuracy of 96% in respect of both datasets.

Carranza-Rojas J. et al. [15] proposed a remarkably interesting method where they created a clean dataset as well as a noise dataset and implement the histogram of curvature over scale (HCoS) along with local binary pattern variance (LBPV) algorithm on both datasets to extract contour information and texture information of leaf image. It has also observed that the clean dataset performs only 7.3% better result than the noise dataset.

Herdiyeni Y. et al. [16] developed a mobile base application for identification of medicinal plant of Indonesia where they use fusion of fuzzy local binary pattern, fuzzy colour histogram and a probabilistic neural network (PNN) classifier for identification of plant based on leaf features. The proposed application has been tested on 2448 Indonesian medical plant's leaf images, where each image was of 270*240 pixel, and a 74.5% of classification accuracy has been found.

Hernandez-Serna A. et al. [17] proposed a new identification system capable for identification of fish, plant as well as butterfly species where multiple image processing techniques were used for extracting 6 geometrical, 2 morphological and 8 texture features of the image, and these extracted features were feed into an artificial neural network for recognising of pattern. The proposed system has been tested on Flavia dataset where 92.87% of accuracy has been obtained in respect of plant species identification.

Begue A. et al. [18] introduced an automated method for recognition of medicinal plant using computer vision and five different machine learning classifier namely random forest, multilayer perceptron neural network, support vector machine, naïve Bayes and k-nearest neighbour. Among all the five classifier the best classification accuracy of 90.1% has been achieved while using random forest classifier. Authors also clarified that due to constraints on resource the function of multilayer perceptron neural network classifier could not be fully exploited therefore the accuracy of this classifier may be more than they opiated in their research which is 88.2%.

Kaur S. et al. [19] proposed a plant species identification system using computer vision and machine learning approach based on leaf images of a plant. The operational flow chart of the proposed system is shown in Fig.2.

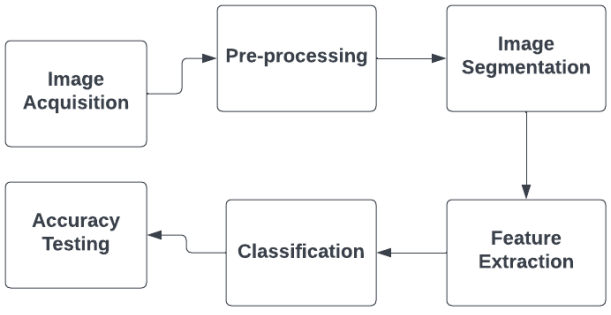

Fig. 2 Operational flowchart

During image acquisition phase the author used the Swedish dataset which consist of total 1125 images of 15 species (75 images from each species) which was followed by image pre-processing where noise handling, image resizing and image enhancement were done. Gaussian filtering has been used for handing of noise and the image has been resized to 330*400 pixel. To extract the region of interest (ROI) during image segmentation phase author apply k-means color based clustering techniques by using MATLAB R2018b. Texture and colour feature were extracted on feature extraction phase using grey-level co-occurrence matrix (GLCM) and first order statistics from colour image respectively. During classification phase of the proposed system multiclass support vector machine (MSVM) classifier has been used. And aggregate accuracy of 93.26% has been achieved by the system during accuracy testing phase.

Wang Z. et al. [20] proposed an approach for retrieval of image leaf based on shape of the leaf which includes two steps. To reduce the search space in the first step, all dissimilar leaf images with the query image were filtered out using eccentricity. In the next step, a fine retrieval in the reduced search space has been carried out using simple shape features including centroid contour distance (CCD), eccentricity and angle code histogram (ACH). During testing phase 1400 leaf images were taken from 140 plant species where a better retrieval accuracy has been found.

Ibrahim Z. et al. [21] conducted a comparative study on recognition of leaf images among three popular texture features namely histogram of oriental gradients (HOG), local binary pattern (LBP) and speeded-up robust features (SURF) with one-against-one (OAO) approach multiclass support vector machine (MSVM). The recognition accuracy of three texture features HOG, LBP and SURF with MSVM classifier has been tested on Flavia dataset where HOB and LBP achieved the same accuracy rate of 99% which is found better then SURF.

Chaki J. et al. [22] developed a method to detect 3 classes of plant category by analysing leaf shape using moment-invariants (M-I) and centroid-radii (C-R) approach individually with some neural network algorithm for classification. Also, a study has been conducted by the author to improve the reorganization accuracy rate by introducing a hybrid representation with combination of M-I and C-R features. 90%-100% range of accuracy has been reported during testing phase of the study.

Wu Q. et al. [23] introduce an approach on 2006 for features extraction and recognition of plant based on leaf shape using artificial neural network, where four visual features were extracted for representing the leaf shape that were slimness, roundness, solidity and moment invariants of the leaf. To represent the leaf margin and dent four features namely coarseness, size, angle and sharpness of the leaf were extracted. Moreover, ramification and camber were two another features, which were also extracting to represent the leaf vein. Further, leaf shape, dent and vein features that were extracted during feature extraction phase were feed into the trained neural network for recognition of the plant. 180 no of leaf images has been collected from 6 plant species (30 leaf image from each) for identifying the extraction and

recognition performance of the system which was found to be satisfactory according to the authors.

Using leaf shape features and minimum Euclidean distance Haque, F. et al. [24] proposed a plant recognition system consisting of three main steps namely image pre-processing, feature extraction and matching. During image pre-processing the following operation were performed on the input leaf image:

- Resize the image,
- Convert from RGB to grayscale,
- Convert from grayscale to binary,
- Swap between the white pixels to black pixels,
- Perform morphological opening,
- Use average filter,
- Extract the region of interest (ROI),
- Obtain the properties of ROI,
- Convert the image into convex hull image,
- Obtain the properties of convex hall image,
- Obtain the length and width of bounding box.

The following ratios were extracted during feature extraction phase:

- Aspect Ratio
- Rectangularity
- Area convexity
- Perimeter convexity
- Circularity
- Eccentricity
- Form factor

All the ratios extracted from feature extraction phase were considered as input for matching phase. On matching phase, the closest match to the leaf image has been computed by using minimum Euclidean distance. The system has been tested on 50 leaf images from 10 plant species which were collected from flavia dataset, and 94% accuracy rate was achieved.

The following Table-I illustrate an overview of various plant feature extraction and classification method used for plant recognition based on leaf image introduced by different researchers.

TABLE I
AN OVERVIEW OF VARIOUS PLANT FEATURE EXTRACTION AND CLASSIFICATION METHOD USED FOR PLANT RECOGNITION

Sl. No.	Researchers	Research Based on	Method Used	Reference
1	Shayan Hati and Sajeevan G.	Leaf shape	ANN, LEAFilia	[2]
2	Gurpreet Kaur and Gurpinder Kaur	Leaf shape	GLCM, ANN, MLPNN	[3]
3	Sandeeo Kumar E.	Leaf Shape and colour	Canny Edge Detection, MATLAB	[4]
4	Vijay Satti, Anshul Satya and Shanu Sharma	Leaf color and shape	ANN, Euclidan Distrance Method	[5]
5	S. Anubha Pearline, V. Sathiesh Kumar, S. Harni	Leaf color and shape	Local Binary Pattern (LBP), Color Channel Statistics, Logistic Regression, k-nearest neighbour, CNN	[6]
6	Ji-Xiang Du, De-Shuang Huang, Xiao-Feng Wang,Xiao Gu	Leaf Shape	Douglass-Peucker algorithm, Modified Dynamic Programme (MDP)	[7]
7	Dhananjay Bisen	Leaf Shape	Deep Conversational Neural Network, CNN	[8]
8	Javed Hossain and M. Ashraful Amin	Leaf shape	PNN	[9]
9	Boi M. Quach, V. Cuong Dinh, Nhung Pham, Dang Huynh, Bint T. Nguyen	Leaf Shape, Colour and Vein	CNN, SVM	[10]
10	Abdolvahab Ehsanirad and Y.H. Sharath Kumar	Leaf Shape	GLCM, PCA	[11]
11	Jesse Dave S. Selda, Roi Martin R. Ellera, Leandro C. Cajayon, Noel B. Linsanga	Leaf Vein	Scale-Invariant Feature Transform (SIFT), SVM	[12]
12	S. Mouine, I. Yahiaoui, A. Verroust-Blondet	Leaf shape	Triangle side length representation (TSL), Tringle area representation (TAR), Tringle side lengths and angle representation (TSLA) and Tringle oriented angles (TOA)	[13]
13	A. Arkif and M. F. Khan	Leaf shape	ANN with back propagation	[14]
14	J. Carranza-Rojas and E.	Leaf shape	Histogram of curvature over	[15]

	Mata-Montero		scale (HCoS), local binary pattern variance (LBPV)	
15	Yeni Herdiyeni and Ni Kadak Sri Wahyuni	Leaf Shape	Fuzzy local binary pattern, Fuzzy colour histogram, PNN	[16]
16	Andres Hernandez-Serna and Luz Fernanda Jimenez-Segura	Leaf Shape	ANN	[17]
17	Adams Begue, Venitha Kowlessur, Fawzi Mahomoodally, Upasana singh, Sameerchand Pudaruth	Leaf Shape	Random forest, Multilayer perceptron neural network, Support vector machine, Naïve bayes, k-nearest neighbour	[18]
18	Surleen Kaur and Prabhpreet Kaur	Leaf color and shape	Gaussion filtering, k-means color based clustering techniques, GLCM, MSVM	[19]
19	Z. Wang and Zheru Chi	Leaf shape	Centroid Contour Distance (CCD), Eccentricity and Angle Code Histogram (ACH)	[20]
20	Zaidah Ibrahim, Nurbaity Sabri, Nur Nabilah Abu Mangshor	Leaf Shape	Histogram of Oriental Gradients (HOG), Local Binary Pattern (LBP) and Speeded-Up Robust Features (SURF), one-against-one MSVM	[21]
21	Jyotismita Chaki and Ranjan Parekh	Leaf Shape	Moment-Invariants (M-I) and Centroid-Radii (C-R)	[22]
22	Qingfeng Wu, Changle Zhon, Chaonan Wang	Leaf Shape, Dent and Vein	ANN	[23]
23	Farhana Haque and Safwana Haque	Leaf Shape	Minimum Euclidean Distance	[24]

III. CONCLUSIONS

To address the environmental challenges such as loss of plant habitats and efforts towards conservation of plant biodiversity, recognition of plant species became an important aspect. In the meanwhile, the significant growth of computer vision and machine learning has led to increasing demand for automated systems for proper prediction of plant species. In this paper we have discussed and explained several techniques introduced in recent years by different researchers on plant species recognition based on leaf image. While analysis the accuracy rate of the proposed systems of different literatures it has been observed that system using neural network classifiers like probabilistic neural network (PNN), convolutional neural network (CNN), Deep Conversational Neural Network achieved a high classification accuracy rate comparing to other methods. Table-1 summarizes an overview of various plant feature extraction and classification method used for plant recognition based on leaf image proposed in each literature that are reviewed in this paper.

REFERENCES

[1] Ji-Xiang Du, Xiao-Feng Wang and DS Huang, "Automatic plant leaves recognition system based on image processing techniques," Technical Report, Institute of Intelligent Machines, Chinese Academy of Science, 883-893, 2004.

[2] Shayan Hati and Sajeevan G., "Plant recognition from leaf image through artificial neural network," International Journal of Computer Application, Vol 62, No. 17, 2013.

[3] Gurpreet Kaur and Gurpinder Kaur, "Classification of biological species based on leaf architecture," International Journal of Engineering Research and Development, 2012.

[4] Sandeeo Kumar E., "Leaf color area and edge features based approach for identification of indian medicinal plants," Indian Journal of Computer Science and Engineering, 2012.

[5] Vijay Satti, Anshul Satya and Shanu Sharma, "An automatic leaf recognition system for plant identification using machine vision technology," International Journal of Engineering Science and Technology, 2013.

[6] S. Anubha Pearline, V. Sathiesh Kumar, S. Harni, "A study on plant recognition using conventional image processing and deep learning approaches," Journal of Intelligent and Fuzzy System 36(6):1-8, 2019.

[7] Ji-Xiang Du, De-Shuang Huang, Xiao-Feng Wang,Xiao Gu, "Computer-aided plant species identification (CAPSI) based on leaf shape matching technique," Transaction of the Institute of Measurement and Control 28(3), 2006.

[8] Dhananjay Bisen, "Deep convolutional neural network based plant species recognition through features of leaf," Multimedia Tools and Applications 80(8):1-14, 2021.

[9] Javed Hossain and M. Ashraful Amin, "Leaf shape identification based plant biometrics," DOI:10.1109/ICCITECHN.2010.5723901, 2010.
[10] Boi M. Quach, V. Cuong Dinh, Nhung Pham, Dang Huynh, Bint T. Nguyen, "Leaf recognition using convolutional neural networks based features," Multimedia Tools and Applications 82(1):1-25, 2022.
[11] Abdolvahab Ehsanirad and Y.H. Sharath Kumar, "Leaf recognition for plant classification using GLCM and PCA," Oriental journal of computer science and technology 3-6, 2010.
[12] Jesse Dave S. Selda, Roi Martin R. Ellera, Leandro C. Cajayon, Noel B. Linsanga, "Plant identification by image processing of leaf veins," The international conference, DOI:10.1145/3132300.3132315, 2017.
[13] S. Mouine, I. Yahiaoui, A. Verroust-Blondet, "A shape-based approach for leaf classification using multiscale triangular representation," In proceeding of the 3rd ACM conference on International Conference on multimedia retrieval, pp. 127-134, 2013.
[14] A. Arkif and M. F. Khan, "Automatic classification of plants based on their leaves," Biosystems Engineering, Vol. 139, pp. 66-75, 2015.
[15] J. Carranza-Rojas and E. Mata-Montero, "Combining leaf shape and texture for costa rican plant species identification," CLEI Electronic Journal, 19(1), paper 7, 2016.
[16] Yeni Herdiyeni and Ni Kadak Sri Wahyuni, "Mobile application for indunesian medicinal plants identification using fuzzy local binary pattern and fuzzy color histogram," International Conference on Advance Computer Science and Information System, 2012.
[17] Andres Hernandez-Serna and Luz Fernanda Jimenez-Segura, "Automatic identification of species with neural networks," PeerJ 2: e563, DOI 10.7717/peerj.563, 2014.
[18] Adams Begue, Venitha Kowlessur, Fawzi Mahomoodally, Upasana singh, Sameerchand Pudaruth, "Automatic recognition of medicinal plants using machine learning techniques," International Journal of Advance Computer Science and Applications, Vol. 8, No. 4, 2017.
[19] Surleen Kaur and Prabhpreet Kaur, "Plant species identification based on plant leaf using computer vision and machine learning techniques." Journal of Multimedia Information System, Vol. 6, No. 2, 2019.
[20] Z. Wang and Zheru Chi, "Shape based leaf image retrieval," IEE Proceedings – Vision Image and Signal Processing, Vol. 150, 2003.
[21] Zaidah Ibrahim, Nurbaity Sabri, Nur Nabilah Abu Mangshor, "Leaf recognition using texture features for herbal plant identification," Indonesian Journal of Electrical Engineering and Computer Science, Vol. 9, No. 1, 2018.
[22] Jyotismita Chaki and Ranjan Parekh, "Plant leaf recognition using shape based features and neural network classifiers," International Journal of Advance Computer Science and Applications, Vol. 2, No. 10, 2011.
[23] Qingfeng Wu, Changle Zhon, Chaonan Wang, "Featuire extraction and automatic recognition of plant leaf using artificial neural network," Advance en Ciencias de la Computacion, pp. 5-12, 2006.
[24] Farhana Haque and Safwana Haque, "Plant recognition system using leaf shape features and minimum Euclidean distance," ICTACT Journel on Image and Video Processing, Vol: 09, Issue: 2, 2018