Classification of Biological Species Based on Leaf Architecture–A review

Gurpreet kaur
Dept.of Computer Science
SVIET-Banur,India.

Himanshu Monga
Dept.of Computer Science
SVIET-Banur,India

***Abstract*-Plants play an important role in human life and provide required information for the development of human society. The urgent situation is that due to environmental degradation and lack of awareness, many rare plant species are at the risk of extinction so it is necessary to keep record for plant protection. We believe that the first step is to buildup a database for protecting plants. So the need arises to teach a computer how to classify plants. Despite the great advances made in Botany, there are many plants which are still unknown.This research focuses on using digital image processing for the purpose of automate classification and recognition of plants based on the images of the leaves. In this paper we review leaf architecture and various techniques for automated plant classification and recognition.**

Keywords: automatic classification; classification algorithms; image processing.

I. INTRODUCTION

Leaf analysis is beneficial for those who are interested in preserving biodiversity. Leaf analysis can become much easier with the help of computer aided technologies. For doing leaf recognition botanists uses morphological features of leaves. These features help in automated classification tool. For example, Digital images of plant leaves are processed using various algorithms to extract shape related features that could be used in classifying different species. Finally, the classification is done based on the extracted features. This will make studying plant taxonomy and ecology easier, raise biology skill standards and it will be helpful for the management of natural reserve parks and forest plantation. Plant Kingdom.[13] is mainly classified on the basis of how they reproduce. Sometimes plants are classified according to their size and complexity. The bryophytes are small in size and perform the function of leaves, stems and roots. Further there are deseminated byspores plants like liverworts and mosses. Ferns (pteridophytes), gymnosperms and angiosperms have vascular tissue which helps in transportation of the nutrients and water through the plant. They are called tracheophytes. Mosses, liverworts, hornworts (bryophytes) are non-vascular i.e. they do not help in transportation of sugar, water and nutrients.

The review paper is organized as Section I includes Introduction, Section II includes methodologies used for plant classification and recognition, Section III includes related work and Section IV includes Algorithms used for plant species classification, Section V includes Conclusion.

II. METHODOLOGIES USED FOR PLANT CLASSIFICATION AND RECOGNITION

Plant species identification requires recognizing the plant by various characteristics, such as size, form, leaf shape, flower color, odor, etc., and linking it with a common or so-called scientific name. Correct identification provides basic information about size, shape and texture of a plant and can be helpful in protecting it from various types of pests and diseases. Plant species classification can be done through various ways like flower, root, leaf, fruit etc. Plant classification can be done according to the structure of their leaf, bark, flower shape, color, texture and seedling and morph. But if the plant classification is based on only two dimensional images then shapes of flowers, seedling and morph of plants is unsuitable because of their complex three dimensional structures [1]. Since the plant leaves are two dimensional in nature they are well suited for classification of various plant species. Sampling leaves and capturing leaves is low cost and convenient. Leaf image can be easily transferred to computer that can automatically extract features using various image processing techniques. In this paper we review various image processing techniques for automated plant classification and recognition.

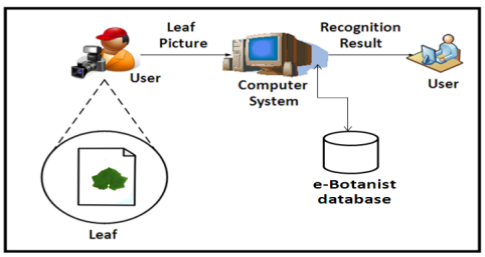

Figure 1. The system conceptual diagram

III. RELATED WORK

Studies have been conducted in the past decade on the automation of plant classification and recognition. Some of these studies were based on

the extraction of a single feature from the image of a plant part such as the leaf, or the flower. Some studies were based on the extraction of multiple features but from a single family of plants. Some studies worked on image-based plant classification, while others focused on image-based plant recognition.

Jyotismita Chaki *et al* [1].In this paper leaf recognition has been done by using shape analysis and feature extraction. The features are given as input to the neural network which is based on supervised learning algorithm having multilayer preceptor with feed forward backward back propagation architecture. The result of this are tabulated for each feature category and are of the range from 90-94 having mean square error in the range of 0-0.7.In certain test folds foe classifying the leaf types the classification rate even goes as high as 100 but in his conclusion he is mentioning an accuracy of 94%.

Brendon J. Woodford *et al* [2].In this paper they have used wavelet based analysis of fruit/leaf images for doing classification. The fruit/leaf images consist of the fruit and the insect which is damaging it; therefore it is a classification problem which tries to identify fruit/leaf having particular pests damaging it. The objective of this thesis have been to take advantage of taking images of the fruit /leaf without doing manual labor in terms of inspection and climbing trees and manually checking the pest infected areas .This person is also using neural network based classifier for identifying different types of pests on particular set of fruits/leafs and the accuracy result is greater than 90%.

L. Tang [3].In this paper low level features (texture features)are used to classify different types of grass weeds leaves The image is subjected to Gabor wavelet transformation and its features are extracted and given to multilayer preceptor neural network. The weeds are classified into two major categories –broad leaves and grass category. The results of paper for each type of broad leaf and grass weed ranges from 88-92%.

Stephen Gang Wu *et al* [4].In this paper PNN (Probabilistic Neural Network) has been used for plant species classification based on leaf architecture They have extracted 12 features of the leaf which are further orthogonal into five principle variables which consist of the input vectors .PNN has been trained on very large number of leaves and has classified 32 species of plants. These papers also give an accuracy which is greater than 90% for most of the classification.

N.Valliammal *et al* [5].In this paper the researcher is discussing the importance of discovering new plants to mankind and purposes leaf recognition and classification based on hybrid image segmentation algorithm. He develops a histogram of each leaf image segmented. In the end his results include calculation of energy mean square error and PSNR (peak to signal ratio).These results basically give the quality of segmentation.

IV. ALGORITHMS USED FOR PLANT SPECIES CLASSIFICATION.

The classification algorithm implemented for accurate identification of the plants based on Leaf images

A. *K-means clustering*

K-means clustering is an algorithm to classify the objects based on attributes/features into *K* number of groups where *K* is a positive integer. K-means clustering is a supervised learning algorithm and utilized a prior knowledge of the number of clusters. Biva [8] introduced K-means clustering for the purpose of leaf image classification.

B. *General regression neural networks*

General regression neural networks perform regression where the target Variable is continuous. The basic idea is that a predicted target value of an item is likely to be the same as other items that have close values of the predictor variables. Jyotismita [1] used GRNN for classification scheme.

C. *The k-nearest neighbor algorithm*

The k-nearest neighbor algorithm is the simplest of all machine learning algorithms: an object classification depends on the majority vote of its neighbors, with the object being assigned to the class most common amongst its *k* nearest neighbors. Gu *et al*.[6] used the k-nearest neighbor algorithm for doing plant species classification.

D. *Move median centers (MMC) hypersphere* .

Du *et al*.[7] used move median centers (MMC) hypersphere classifier to classify plants based on shape-related features of leaf such as aspect ratio, rectangularity, area ratio of convex hull, perimeter ratio of convex hull, sphericity, circularity, eccentricity, form factor, and invariant moments.

E.*Radial basis probabilistic neural network (RBPNN).*

Gu *et al*. [6] introduced RBPNN .The plants were classified based on the extracted and derived features using and radial basis probabilistic neural network (RBPNN).

F.*Probabilistic neural network.*

Wu et al.[9] used a probabilistic neural network for classification of leaf images on the basis of 12 leaf features.

G. *A Gabor wavelet–based algorithm.*

Brendon et al.[2] developed The Gabor wavelet/ANN system to classify weed images into broadleaf and grass categories on the basis of texture features.

H.*A radial basis function network.*

A radial basis function network is an artificial neural network that uses radial basis functions as activation functions. It is a linear combination of radial basis functions. Hongfei *et al*.[10] utilized RBF to classify Camellia (Theaceae) species.

I. SVM (Support Vector Machines)

SVM (Support Vector Machines) is based on the concept of decision planes that define decision boundaries. Hongfei [10] utilized SVM for identification of Camellia species.

J. *Probabilistic Neural Network (PNN).*

Stephen [4] employed PNN with image and data processing techniques to implement general purpose automated leaf recognition for plant classification.

V. CONCLUSION AND FUTURE WORK

There are various algorithms used for classification as per literature survey. Most of these researches have been done on standard 32 species of plants. However these species have been tested on at the most 2 techniques of classification algorithms. There is no research been done in terms of conducting a comparative analysis of those algorithms with these species. *For Future scope* we suggest that a full length comparative analysis of techniques like SVM, KNN, Bayesian classification and neural networks must be done, so as to find the most optimal classifier which gives output in terms of very high accuracy using minimal computational resources

ACKNOWLEDGEMENT

I would like to thanks my parents and my friends for showing trust and giving their support which helps me successful in completing this review process.

REFERENCES

[1] Jyotismita Chaki, Ranjan Parekh.: Plant Leaf Recognition using Shape based Features and Network classifiers,(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, 2011.

[2] Brendon J. Woodford, Nikola K. Kasabov and C. Howard Wearing.: Fruit Image Analysis using Wavelets.

[3] L. Tang, L. Tian, B. L. Steward. : Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network, Transactions of the ASAE Vol. 46(4).: 1247–1254 _ 2003 American Society of Agricultural Engineers ISSN 0001–2351.

[4] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang.:A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network,arXiv:0707.4289v1 [cs.AI] 29 Jul 2007.

[5] N.Valliammal,Dr.S.N.Geethalakshmi (2011).:Hybrid Image Segmentation Algorithm for Leaf Recognition and Characterization, ©2011 IEEE

[6] X. Gu et al.,Wang.: Leaf recognition based on the combination of wavelet transform and Gaussian interpolation, *ICIS*, vol. 3644/2005, pp. 253-262, 2005.

[7] J.X. Du et al.: Leaf shape based plant species recognition, Applied Mathematics and Computation, vol. 185, issue 2, Amsterdam, Elsevier, 2007, pp.883-893.

[8] Biva shrestha" classification of plants using images of their leaves"

[9] S.Wu et al: A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network, in Proc. Signal Process. And Inform.Technology, IEEE Int. Symp., Cairo, Egypt, 2007.

[10] Hongfei et al., Classification of Camellia (Theaceae) Species Using Leaf Architecture Variations and Pattern Recognition Techniquies.

[11] http://en.wikipedia.org/wiki/Plant.

AUTHORS PROFILE

Ms. Gurpreet kaur was born in the small village of khanna, punjab. After finishing high school in khanna ,she moved to BBSBEC fatehgarh sahib, Punjab to pursue a Bachelor's degree in Information technology. After graduating with a Bachelor of technology in Information Technology from BBSBEC Engineering College-Fatehgarh sahib in 2010,she started her M.tech in computer science engineering from SVIET,banur(Punjab).

*Mr.Himanshu Monga***.** He Was born in Joginder Nagar, Himachal Pradesh, India, on 06th November 1974. He obtained his bachelor's degree in electronics & Communication Engineering from Amravati University, Maharastra, India and Master's degree in Electronics and Telecommunication Engineering & Masters degree in management from IGNOU in Human resource management. He is pursuing Ph.D.from Thapar Institute of Engineering and Technology, Patiala, Punjab, India.. Presently, He is working as Associate Professor in computer science Engineering department at SVIET Banur India. He has more than 15 Research papers in international/national Journals.