www.ijecs.in

International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 5 Issue 8 August 2016 Page No. 17658-17661

A Review on Plant Leaf Classification and Segmentation

Lakhvir Kaur[1], Dr. Vijay Laxmi[2]

[1]Research Scholar M-Tech, Computer Science and Engineering, Guru Kashi University, India
[2]Dean, UCCA, Guru Kashi University, India

[1]lakhvirgrewal92@gmail.com, [2]cse_vijay2003@yahoo.co.in

Abstract: A leaf is an organ of vascular plant and is the principal lateral appendage of the stem. Each leaf has a set of features that differentiate it from the other leaves, such as margin and shape. This paper proposes a comparison of supervised plant leaves classification using different approaches, based on different representations of these leaves, and the chosen algorithm. Beginning with the representation of leaves, we presented leaves by a fine-scale margin feature histogram, by a Centroid Contour Distance Curve shape signature, or by an interior texture feature histogram in 64 element vector for each one, after we tried different combination among these features to optimize results. We classified the obtained vectors. Then we evaluate the classification using cross validation. The obtained results are very interesting and show the importance of each feature. the classification of plant leaf images with biometric features. Traditionally, the trained taxonomic perform this process by following various tasks. The taxonomic usually classify the plants based on flowering and associative phenomenon. It was found that this process was time consuming and difficult. The biometric features of plants leaf like venation make this classification easy. Leaf biometric feature are analyzed using computer based method like morphological feature analysis and artificial neural network based classifier. KNN model take input as the leaf venation morphological feature and classify them into four different species.

Keyword - Artificial neural network, Canny edge detection, k-NN Classification, Leaf venation pattern, Morphological Features

I. INTRODUCTION

Plant diseases have turned into a dilemma as it can cause significant reduction in both quality and quantity of agricultural products. In India 70% of the population depend on agriculture. Farmers have wide range of diversity to select suitable Fruit and Vegetable crops. However, the cultivation of these crops for optimum yield and quality produce is highly technical. It can be improved by the aid of technological support. The management of perennial fruit crops requires close monitoring especially for the management of diseases that can affect production significantly and subsequently the postharvest life. In [2] the authors have worked on the development of methods for the automatic classification of leaf diseases based on high resolution multispectral and stereo images. Leaves of sugar beet are used for evaluating their approach. Sugar beet leaves might be infected by several diseases, such as rusts (Uromyces betae), powdery mildew (Erysiphe betae).

Disease is caused by pathogen which is any agent causing disease. In most of the cases pests or diseases are seen on the leaves or stems of the plant. Therefore identification of plants, leaves, stems and finding out the pest or diseases, percentage of the pest or disease incidence , symptoms of the pest or disease attack, plays a key role in successful cultivation of crops. It is found that diseases cause heavy crop losses amounting to several billion dollars annually.

Fig1: Potato leaf affected by the late blight

Fig2: Leaf symptoms of canker on top and bottom of leaf

Disease management is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Precise quantification of these visually observed diseases, pests, traits has not studied yet because of the complexity of visual patterns. Hence there has been increasing demand for more specific and sophisticated image pattern understanding. In biological science, sometimes thousands of images are generated in a single experiment. These images can be required for further studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc. Almost all of these tasks are processed manually or with distinct software packages. It is not only tremendous amount of work but also suffers from two major issues: excessive processing time and subjectiveness rising from different individuals .Hence to conduct high throughput experiments, plant biologist need efficient computer software to automatically extract and analyze significant content. Here image processing plays important role.

II. IMPORTANCE

In object recognition research a lot have been done about general features extraction or recognition between different classes of objects. In case of a species domain recognition, 10 taking into account the unique characteristics that belong to this category, improves the performance of the system. Despite the high technical aspect of this project, dealing with leaves, gives it a biological connotation. Some basic knowledge about leaves have to be learned and concepts about how the biologists themselves recognize leaves has to be studied. The next two paragraphs are devoted to these experiences. Precision Botany (PB) refers to the application of new technologies in plant identification. Computer vision can be used in PB to distinguish plants from its species level, so that an identification can be applied on the size and number of plants detected for the classification purpose. This is focused on the application of computer vision for identification purposes of species in Stemonoporus genus. Surveys reveal that there are 3711 flowering plant species in Sri Lanka [4]. Out of these, 926 are endemic [5, 6]. Since some of these have minute variations, identification of these species has become difficult. Accurate and speedy identification of plants has become a time consuming and a fuzzy work due to nonavailability of a computerized scientific plant identification system. Design and implementation of image-based plant classification system is a long felt. Biologists receive a large number of requests to identify plants for people, many species of plants look very similar on their leaves, and botanists will turn to identifying the species based on their structure or other morphologies.

There are three main parts to a leaf:

1. The base which is the point at which the leaf is joined to the stem.
2. The stalk or petiole is the thin section joining the base to the lamina - it is generally cylindrical or semicircular in form.
3. The lamina or leaf blade is the wide part of the leaf Leaves can be of many different shapes: Primarily, leaves are divided into simple - a single leaf blade with a bud at the base of the leafstem; or compound - a leaf with more than one blade.

All blades are attached to a single leafstem. Where the leafstem attaches to the twig there is a bud. Leaves may be arranged on the stem either in an alternate arrangement – leaves that are not places directly across from each other on the twig; or in an opposite arrangement – 2 or 3 leaves that are directly across from each other on the same twig.

The margin (the edge of a leaf) as seen in Figure 3.2 may be entire, singlytoothed, doubly-toothed, or lobed. Compound leaves may be palmate – having the leaflets arranged round a single point like fingers On the palm of a hand; or pinnate – when the leaves are joined on the two sides of the stalk, like the vanes of a feather as seen in the Figure 1.3. Leaf arrangements are pretty straightforward to figure out. Need to look for the nodes and then determine how many leaves are coming off each node. If there's only one leaf per node, then need only to determine whether the arrangement is alternate or spiral, and it's usually pretty obvious.

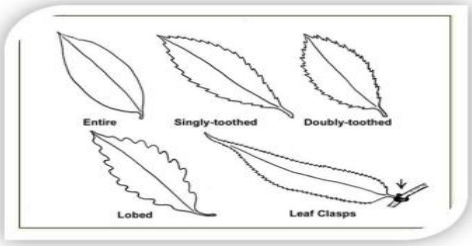

Figure 1. Simple leaves – Margin structure.

III. LITERATURE SURVEY

Smita Naikwadi et.al.[2013] We propose and experimentally evaluate a software solution for automatic detection and classification of plant leaf diseases. Studies of plant trait/disease refer to the studies of visually observable patterns of a particular plant. Nowadays crops face many traits/diseases. Damage of the insect is one of the major trait/disease. Insecticides are not always proved efficient because insecticides may be toxic to some kind of birds. It also damages natural animal food chains. The following two steps are added successively after th segmentation phase. In the first step we identify the mostly green colored pixels. Next, these pixels are masked based on specific threshold values that are computed using Otsu's method, then those mostly green pixels are masked. The other additional step is that the pixels with zeros red, green and blue values and the pixels on the boundaries of the infected cluster (object) were completely removed. The experimental results demonstrate that the proposed technique is a robust technique for the detection of plant leaves diseases. The developed algorithm's efficiency can successfully detect and classify the examined diseases with a precision between 83% and 94%, and can achieve 20% speedup over the approach proposed in [1].

Abdul Kadir et.al. [2012]This paper reports the results of experiments in improving performance of leaf identification system using Principal Component Analysis (PCA). The system involved combination of features derived from shape, vein, color, and texture of leaf. PCA was incorporated to the identification system to convert the features into orthogonal features and then the results were inputted to the classifier that used Probabilistic Neural Network (PNN). This approach has been tested on two datasets, Foliage and Flavia, that contain various color leaves (foliage plants) and green leaves respectively. The results showed that PCA can increase the accuracy of the leaf identification system on both datasets.[2]

Samuel E. Buttrey et.al. [2002] We construct a hybrid(composite) classi&er by combining two classi&ers in common use— classi&cation trees and k-nearest-neighbor (k-NN). In our scheme we divide the feature space up by a classi&cation tree, andthen classify test set items using the k-NN rule just among those training items in the same leaf as the test item. This reduces somewhat the computational loadassociatedwith k-NN, andit produces a classi&cation rule that performs better than either trees or the usual k-NN in a number of well-known data sets. [3]

Prof. Meeta Kumar et al. [2013]In this paper we present survey on various classification techniques which can be used for plant leaf classification. A classification problem deals with

associating a given input pattern with one of the distinct classes. Plant leaf classification is a technique where leaf is classified based on its different morphological features. There are various successful classification techniques like kNearest Neighbor Classifier, Probabilistic Neural Network, Genetic Algorithm, Support Vector Machine, and Principal Component Analalysis. Deciding on the method for classification is often a difficult task because the quality of the results can be different for different input data. Plant leaf classifications has wide applications in various fields such as botany, Ayurveda, Agriculture etc. The goal of this survey is to provide an overview of different classification techniques for plant leaf classification.[4]

Gurpreet kaur et.al.[2012] 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. [5]

Comparison Table :

Methods of classification	**Advantages**	**Disadvantages**
K-NN	1. Simpler classifier as exclusion of any training process. 2. Applicable in case of a small dataset which is not trained.	1. Speed of computing distance increases according to numbers available in training samples. 2. Expensive testing of each instance and sensitive to irrelevant inputs.
SVM	1. Simple geometric interpretation and a sparse solution. 2. Can be robust, even when training sample has some bias.	1. Slow training. 2. Difficult to understand structure of algorithm. 3. Large no. support vectors are needed from training set to perform classification task
ANN	1. Tolerant of noisy inputs. 2. Instances classified by more than one output.	1. Long training time. 2. Large complexity of network structure. 3. Need lot of memory for training data
Backpropagation Network (BPN)	1. Easy to implement. 2. Applicable to wide range of problems.	1. Learning can be slow. 2. It is hard to know how many neurons as well as layers are required.
	3. Able to form arbitrarily complex nonlinear mappings	
Radial Basis Function (RBF) or Naive Bayesian	1. Training phase is faster. 2. Hidden layer is easier to interpret.	1. It is slower in execution when speed is a factor.

IV. PROBLEM STATEMENT

Plant leaf classification is the major problem and also their diseases are also major problem. Most of plant leaf have attacks by snail, worm and fungi. Furthermore, when the leaf had been infected or attacked, the others areas had been exposed to be infected. Thus, it will decrease the life of plant. There is the overlapping of the plant diseases problem. Plants play an important role in our lives, without plants there will not be the existence of the ecology of the earth. The large amount of leaf types now makes the human being in a front of some problems in the specification of the use of plants, the first need to know the use of a plant is the identification of the plant leaf. The problem with a number of these techniques is that they require some manual intervention such as correctly orienting the image or identifying the end points of the leaf's main vein. S. R. Deokar (2013) has worked on leaf recognition by extracting 28 and 60 Feature point. These features extracted by vertical and horizontal splitting the leaf images. ANN is used to compare performance of leaf recognition.

V. SCOPE

- The users of this system are paddy farmers.
- The prototype will be develop by using MATLAB 2013 a.
- 10 samples each of the normal, brown spot disease, narrow brown spot disease and blast disease will be used in this project.

VI. CONCLUSION

Plants play an important role in our lives, without plants there will not be the existence of the ecology of the earth. The large amount of leaf types now makes the human being in a front of some problems in the specification of the use of plants, the first need to know the use of a plant is the identification of the plant leaf. This work proposed a comparison of supervised classification of plant leaves, where we used to represent species in seven different representations, using three features extracted from binary masks of these leaves: a finescale margin feature histogram, by a Centroid Contour Distance Curve shape signature, and by an interior texture feature histogram. Results were very interesting in a way that gives as clear ideas: In term of representation: we can differentiate leaves by its margin better than shape or texture, but, experiments shown in this study prove our idea: the more we combine these features, the more precise the difference between samples is and that is what gives better results in classification. In term of classification: distance based algorithms give the best result for plant leaves classification. So, we can conclude that these algorithms are the most suitable for that task. On the other hand, the approach based on decision tree

gives the worst results because of the overfitting problem. In general, a learning algorithm is said to overfit relative to a simpler one if it is more accurate in fitting known data but less accurate in predicting new data. Use of the three features proved that there is some information more important than other. We discovered that margin representation can affect results more than the shape of the leaf. However, the combination of the three features gives the best result. To solve this problem, we plan, as future work, use of feature extraction algorithms, like PSO, to clean dataset and keep the important information in order to optimize the obtained results and avoid overfitting problem posed by decision tree algorithm. We plan also to use bio-inspired algorithms. They are part of a new research domain that is becoming more important due to its results in different areas.

REFERENCES

[1]. Smita Naikwadi et.al."ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES" International Journal of Application or Innovation in Engineering & Management (IJAIEM) Web Site: www.ijaiem.org Email: editor@ijaiem.org, editorijaiem@gmail.com Volume 2, Issue 11, November 2013.

[2]. Abdul Kadir et.al. "Performance Improvement of Leaf Identification System Using Principal Component Analysis" International Journal of Advanced Science and Technology Vol. 44, July, 2012

[3]. Samuel E. Buttrey et.al. "Using k-nearest-neighbor classi&cation in the leaves of a tree" Computational Statistics & Data Analysis 40 (2002) 27 – 37 www.elsevier.com/locate/csda.

[4]. Prof. Meeta Kumar et al. "Survey on Techniques for Plant Leaf Classification" International Journal of Modern Engineering Research (IJMER) www.ijmer.com Vol.1, Issue.2, pp-538-544 ISSN: 2249-6645.

[5]. Gurpreet kaur et.al. "Classification of Biological Species Based on Leaf Architecture–A review" IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol. 2, No.2, April 2012.

[6]. Amlekar Manisha, Manza R.R, Yannawar Pravin,(2013), Leaf classification based on leaf dimension biometric features of leaf shape using k-means classifier, NCAC, Jalgoan.

[7]. Amlekar Manisha, Manza R.R, Yannawar Pravin, Gaikwad B.P,(2013), Image data mining for classifying leaf dimension biometric features of leaf shape using KNN classification technique, CMS,

[8]. Aurangabad. Beghin T., Cope J. S., Remagnino P. and Barman S., (2010), Shape and texture based plant leaf Classification, ACIVS, 2, 345–353.

[9]. J. S., Remagnino P., Barman S., and Wilkin P., (2010), Plant texture classification using gabor co-occurrences," in Proceedings of the 6 th international conference on Advances in visual computing,2, 669– 677.

[10]. Deokar S. R., Zope P. H. Suralkar S. R ,(2013), Leaf Recognition Using Feature Point Extraction and Artificial Neural Network, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, 2(1). Ehsanirad, (2010), Plant classification based on leaf recognition, International Journal of Computer Science and Information Security, 8(4), 78–81.

[11]. C. Mallah, J. Cope, and J. Orwell ,"Plant leaf classification using probabilistic integration of shape, texture and margin features," Signal Processing, Pattern Recognition and Applications, 2013

[12]. S. Zhang, and K. W. Chau, "Dimension reduction using semisupervised locally linear embedding for plant leaf classification," Emerging Intelligent Computing Technology and Applications, Springer Berlin Heidelberg, 2009, pp. 948-955

[13]. A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Leaf classification using shape, color, and texture features , arXiv preprint arXiv:1401.4447, 2013

[14]. A. H. M. Amin, and A. I. Khan, "One-shot Classification of 2-D Leaf Shapes Using Distributed Hierarchical Graph Neuron (DHGN) Scheme with k-NN Classifier," Procedia Computer Science, 24, 2013, pp.84-96