ISSN No:-2456-2165

Plant Discovery from Leaf

[1]Blessy Benny
Dept.Computer Science and Engineering)
Sahrdaya College of Engineering and Technology
Kodakara

[2]Helna Jose M
Dept.Computer Science and Engineering)
Sahrdaya College of Engineering and Technology
Kodakara

[3]Anjana Francis
Dept.Computer Science and Engineering)
Sahrdaya College of Engineering and Technology
Kodakara

[4]Ann Rija Paul
Dept.Computer Science and Engineering)
Sahrdaya College of Engineering and Technology
Kodakara

Abstract:- The human visual system has no problem in interpreting the subtle variations in colour and shading in a photograph and correctly recognize the object from its background. Suppose if a person takes a field trip, and when he or she sees a plant on the ground, that person would like to gain information about the plant. With the help of a mobile camera and a recognition program, we enable our users to get useful information of the plant's leaf which they have taken. The sole purpose of this project is to develop a leaf recognition algorithm based on its specific characteristics. For this photograph of the leaf is taken in a mobile camera and send it to the image processing an application. After processing the image result is send back to the mobile. Result contains the name, species, life span, industrial and medicinal usages of the desired image. This is done by preparing dataset of the plants beforehand. Proposed algorithm is performed as three stages viz preprocessing, extraction of features, classification. In Preprocessing one usually process the image data so it should be in suitable form which means one could single out each objects after this step. In second step measure the features of intended objects. Class of object based on features are determined in the final step.

Keywords:- *Random Forests, Preprocessing, Feature Extraction.*

I. INTRODUCTION

Plants plays a pertinent role to human welfare because they are the root of clothing, medicines, food, housing materials and more. Back in the day, plant identification was the main domain of taxonomists, botanists and even ordinary people who used to identify the plants for which they search about, by comparing them with previously collected specimens or by using books or web surfing. Computers have evolved to be a vital device in many applications like agriculture, medical, defense, engineering and so forth. It has the ability to process multimedia information captured from computing devices. An image consists of important information that are retrieved using some computational method. Image segmentation partitions an image into smaller portions that are meaningful. In other words it can be stated as identification and classification of region needed. The segmentation is performed based on some common features .

A. Problem Statement

We human have no problem in interpreting the minute dissimilarity in colour and shade in an image and rightly recognize the leaf from its background. Taking an example of a person walking in a field , and on seeing a plant on the ground, he or she would like to know about the plant. By a good mobile camera and a recognition program, we enable user to get to know about the plant of interest. Development of a leaf recognition algorithm based on its specific features is the sole purpose of our project.

B. Motivation

Plants plays a pivotal role in our environment. Earth's ecology is dependant on plants. Of late, risk of annihilation has incurred to many types of plants. Inorder to protect these endangered plants and to categorize various sort of diversities, a leaf database of plant is a significant step which leads to conservation of earths biosphere. In worldwide there are a uncountable number of plant species. To handle such volumes of information, there is a need to develop an efficient classification mode . Other than conservation aspect, recognition of plants is also necessary to know about their medicinal characteristics and to use them as sources alternative energy sources. There are several ways to recognize a plant, like root, leaf, flower, fruit etc. In this project leaf is used to identify the plant.

C. Objectives

The target to achieve in this project is to develop a leaf recognition program based on specific features extracted from photography. The recognition of leaves from photographs implies several steps viz Image preprocessing, Feature extraction, Classification and Obtaining the results. User can upload the image of leaf of corresponding plant and on searching in the app will display the name, species, medical status, lifespan etc.

II. LITERATURE SURVEY

Up to the time of few years ago, there was a limited number of publications were concerning automatic plant identification, comprised of botanical applications targeting leaf-based plant identification or agricultural applications, such as for detecting diseased plants. This section has presented an overview of the existing approaches for plant identification. The plant identification campaign organized by Image CLEF in 2011, and repeated in 2012 and 2013, has marked the turning point in this area [3].In systems

ISSN No:-2456-2165

geared towards botanical applications segmented leaf images are used to identify unknown plant varieties, often using features obtained from the leaf contour. Among these, Yahiaoui etal. proposed an image retrieval system for identifying plants by using contour-based shape features in [1]. The extracted shape descriptors in this study include the length histogram of contour segments in different directions. Similarly, Wang and his group [4] focused on the leaf image retrieval problem using features such as centroid contour distance curve, eccentricity and angle code histograms, while later on investigating shape context [7].Concentrating on shape, [15] extract elliptic Fourier harmonic functions from the leaf boundary and Nam et al. [10] use minimum perimeter polygons for shape description in combination with a matrix of interest points for similarity computation.[17] focused on the combination of shape and texture, by employing contour signatures along with edge gradient orientations, while Hussein et al. exploit texture by means of the discrete wavelet transform. Similarly, Man et al combines color and texture features after rotating each leaf so as to align its central axis with the horizontal [11]. Both [12] and [13] rely on simple geometric and morphological features, while [3] additionally combines them with a powerful moving median centers hypersphere classifier. PlantSnap : PlantSnap is powered by deep learing and artificial intelligence and is changing the game for plant identification. PlantSnap Features: Identify trees and other plants by snapping a picture and Identify flowers, trees and other plants by plugging their picture into their Plant Database. Use PlantSnap for Flower identification, Identify trees, Identify leaves, Mushroom Identification and Identify Succulents, Cactus more. 2. Pl@ntnet : Pl@ntNet is a picture revea1ing and retrieva1 app1ication for recognition of p1ants.It is deve1oped by professiona1s from four French study estab1ishments (Cirad, INRA, Inria and IRD), and a1ong with Te1a Botanica system, using the financia1 support invo1ving Agropo1is foundation. Invo1ving additiona1 capabi1ities, this particu1ar free app he1ps determining p1ant types through pictures, by using a visua1 identification app1ication.

III. DESIGN MODULES

Image is captured and binarization is done. Noise is removed by filtering. Region and texture properties are extracted and added to feature set. Neural network is trained and matching is performed to display the output. Each leaf has its own features and carries significant information that can help people to recognize and classify the plant by looking at it. Initially parameters are set up to capture image. Inorder to remove noise, filter is applied. Then image is binarized. Leaf is segmented. After this, region properties are extracted. This features are added to feature set. Original image is converted to grey scale image. GLCM matrix is calculated. These features are added to feature set. Neural network is trained. Features of the test images are extracted . Finally matching is done and output is displayed. Initially preprocessing is applied. Where it filters the noise of the image caused by many external factors. Then image is converted into gray scale. Followed by the extraction of region, color and texture properties. Feature extraction is reduction of attribute. Dataset is trained beforehand. Then classification is performed. Classifier used is random forest.

- *Algorithm*

Step 1: Setup camera parameters.
Step 2: Capture image.
Step 3:Apply noise filtering
Step 4: Binarize the image.
Step 5:Segment leaf region.
Step 6:Extract region properties.
Step 7:Add region properties to feature set.
Step 8:Convert original image to gray scale.
Step 9:Compute glcm matrix.
Step 10:Find texture features from glcm.
Step 11:Add features to feature set.
Step 12:If all data are acquired, goto step 13 else Step 2.
Step 13:Train the neural network.
Step 14:Capture test image.
Step 15:Extract regional and textural features.
Step 16:Input features to neural network.
Step 17:Check output from neural network and label.
Step 18:Display result.

A. Image Acquisition

Image Acquisition is the first step in any image processing system. The images are formed by integration of an illumination source and the absorption or reflection of the energy by the elements being imaged of scene. Illumination maybe originated by radar, infrared energy source, computer generated energy pattern, ultrasound energy source, X-ray energy source etc. Inorder to sense an image, sensor is used according to the illumination by its nature. Image sensing as a process is known as image acquisition stage. Geometric model or mathematical formula is used to compute a digital image. Image synthesis is more suitable term, and it is also known as rendering. Here through filtering image data is free from unwanted distortions or enhances important image features that are useful for further processing. Brightness regardless to position in the image is changed in gray scale transformation. If the results are viewed by a human gray scale transformation is used. Image processing filters are mainly used to eliminate either the high frequencies in the image which means smoothing the image, or the low frequencies which detects edges in the image. An image can be filtered by frequency.

B. Preprocessing

Preprocessing is lowest level of abstraction process in case of both input and output images intensities. Images are original data captured by the sensor and intensity image that are represented by image function values as matrix . The focus of preprocessing lies in the improvement of the image by eliminating distortions or enhancing features important for forthcoming processing, even though geometric transformations of images like rotation, scaling, translation are categorized among preprocessing methods. Redundancy in images are used in image preprocessing. There are many filters like Gaussian filter, median filter,

ISSN No:-2456-2165

mean filter. Certain filters, such mean filter is appropriate for this purpose. For example, mean filter is used for removing grain noise from a photo captured. Here each pixel is set to the average of the pixels in its neighboring pixel so that variations caused by grain are reduced.

The mean filter is a filter that uses a mask over each pixel in image. It is also known as average filter. But it has a drawback of edge preserving. Preprocessing is a process which does the initial processing of input leaf image captured by the user that includes elimination of noise and change the distorted data.

C. Feature Extraction

Feature extraction is the next process which defines a set of features or characteristics of image that will efficiently represent the information that are important for its classification. Feature extraction is mainly about reducing the amount of characteristics required to describe each one in the large dataset. Collecting dataset is a difficult task. Quantifying its texture content is important to describe region . In texture description there will be measures such as regularity, coarseness and smoothness. In this work statistical techniques are used to describe the textures. In this stage Grey Level Co-occurrence matrix (GLCM) of the leaf images are calculated. GLCM is a matrix of image. Features extracted from three basic characteristics of a leaf: region, color, and texture.

*1) Region :*There are 10 features used to identify the region of a leaf.

a) Roundness : Ratio of perimeter to the area. Area A and perimeter P, roundness R

$$R = \frac{7\pi A}{P^2}$$

b) Perimeter ratio of diameter : Perimeter is the ratio of leaf perimeter P to its diameter D.

$$P_D = \frac{P}{D}$$

c) Compactness :

Ratio of leaf perimeter to leaf area. Complexity and form factor of a leaf is assesed by compactness.

$$C = \frac{P}{\sqrt{A}}$$

d) Major Axis: The major axis(L) of a leaf is its longest diameter.In other words it is a line that passes through the center and both foci.

*e) Minor Axis:*The minor axis(W) of a leaf that is a point equidistant from the foci perpendicular to themajor axis.

f) Eccentricity : Deviation of a leaf of a circle. Ratio of foci(f)f of a leaf and its major axis length(a).

$$E = \frac{f}{a}$$

g) Aspect ratio : Ratio of major axis length (L) to minor axis length (W). It shows narrowness or wideness of the leaf.

$$AR = \frac{L}{W}$$

h) Rectangularity : Ratio of area(A) of a leaf to minimum bounding rectangle(LW).

$$N = \frac{A}{LE}$$

i) Narrow factor : Ratio of diameter(D) of leaf to the major axis(L).

$$NF = \frac{D}{L}$$

j) Perimeter ratio of major and minor axis length : Ratio of leafs perimeter(P) to the sum of major(L) and minor(W) axis length ie,L+W.

$$P_{WL} = \frac{P}{L+W}$$

2) Color : There are 4 features which are extracted to describe color characteristics of a leaf.

a) Mean :

$$\bar{X} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}}{NM}$$

b) Standard deviation :

$$\sigma_i = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{x})^2}$$

c) Skewness :

$$S_i = \sqrt[3]{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{x})^3}$$

d) Kurtosis :

$$K_i = \sqrt[4]{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{x})^4}$$

3) Texture :

Texture features distinguishes leaf based on coarse or smooth features. Gray level co- occurrence matrix (GLCM) is a textural measure which shows thespatial distribution of pixels in an image [14]. GLCM gives four features :

a) Contrast : Variation in GLCM is measured.
b) Correlation : Calculates probability of occurring pair-wise pixel.
c) Energy : Gives sum of squared elements in the GLCM.
d) Homogeneity : It is the closeness of element in the GLCM and the GLCM diagonal.

ISSN No:-2456-2165

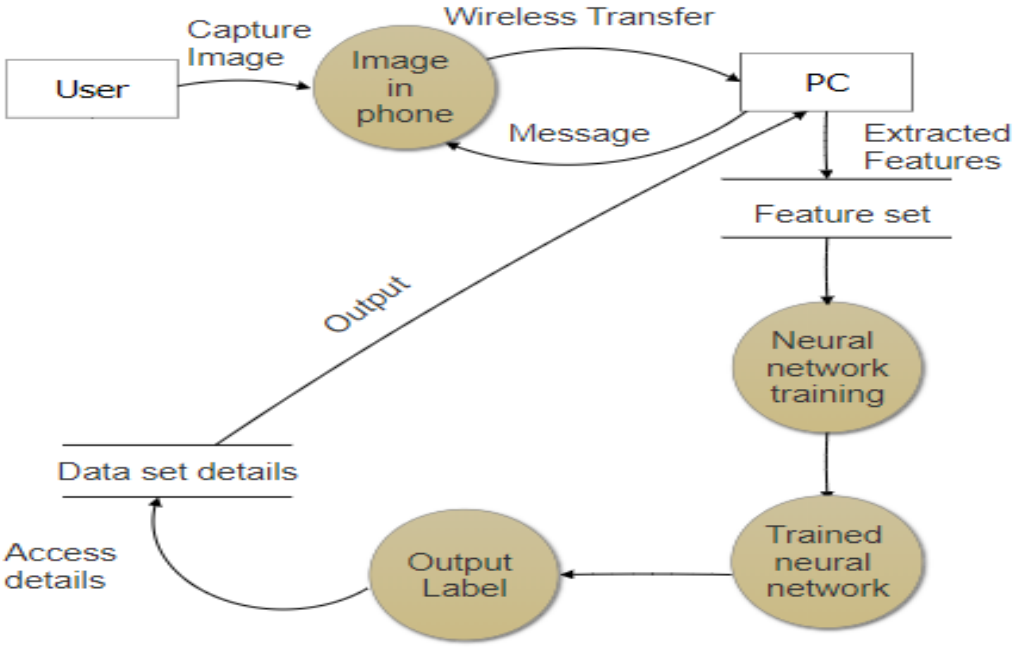

Fig 1

D. Classification

There are many species that are almost similar to one another and there are large dissimilarities within a a species due to maturity changes or seasonal changes.Plant identification is one of the challenging problems in machine learning.Sources of variation common in many object recognition problems are scale pose, and lighting variations are present in plant identification problem.Random forest is the classifier used.From randomly selected subset it creates a set of decision trees of training set. From different decision trees it collects vote to determine the class of the test leaf image.The random forest can apply concept of weight to consider the impact of result in the decision tree. High error rate trees are assigned with low weight value and low error rate trees are assigned with high weight value. Impact of decision is increased when of trees are having low error rate.

IV. FUTURE ENHANCEMENT

Our system has few limitations which will be enhanced in future. One of them is preprocessing stage. It is the main drawback. Inability to process with images with complicated background is its another drawback. We hope to overcome these limitations. We initially planned to achieve image processing over the mobile phone in which the user captures leaf image. Because of some constraints which involve high cost, processing algorithms on a smart-phone are often difficult to do. So our future goal is to implement image processing which we are currently desktop application to an android app.

V. CONCLUSION

This project is about developing plant recognition program. It mainly has three stages viz image preprocessing, feature extraction and classification. Initially camera parameters are set up. User is supposed to take image of the plant leaf of which they want to get details and send to the desktop based application. In image preprocessing noise and distortion are eliminated. Now the image is ready for further processes. Feature extraction stage takes the significant features of the image like texture, color and region properties. Eventually the leaf is correctly classified. Details of the plant of which user has taken is send back to the mobile from which the picture was send.

VI. REFERENCES

[1]. **Asem Khmag , S.A.R Al-Haddad and Noraziahtulhidayu Kamarudin**, (2017) "Recognition System for Leaf Images Based on Its Leaf Contour And Centroid",IEEE 15th Student Conference on Research and Development (SCOReD).

[2]. **Haixi Zhang, Guiqing He, Jinye Peng, Zhenzhong Kuang, Jian ping Fan** ,(2018), "Deep Learning of Path-based Tree Classifiers for Large-Scale Plant Species Identification "2018 IEEE Conference on Multimedia Information Processing and Retrieval.

[3]. **Roemi Fernandez, Hctor Montes, Jelena Surdilovic , Dragoljub Surdilovic ,Pablo Gonzalez-de-Santos and Manuel Armada**,(2018) "Automatic Detection of Field-Grown Cucumbers for Robotic Harvesting ".

[4]. **Z. Wang, Z. Chi, and D. Feng**,(2003) "Shape based leaf imageretrieval. "IEEE P-Vis Image Sign. 150:3443, 2003.

[5]. **S. Wu, F. Bao, E. Xu, Y. Wang, Y. Chang, and Q. Xiang**,(2007) leaf recognition algorithm for plant classification using probabilistic neural network. 7th IEEE International Symposium on Signal Processing and Information Technology, Cairo, Egypt.
[6]. **Tong, Simon and Edward Chang** ,(2012), "Support vector machine
[7]. active learning for image retrieval "
[8]. **Pankaja K and Dr. Thippeswamy G**, "Survey on Leaf Recognization and Classification ".
[9]. **Sigit Adinugroho and Yuita Arum Sari** , "Leaves Classification Using Neural Network Based on Ensemble Features ".
[10]. **Vijai Singh, and A.K. Misra**,(2016) "Detection of plant leaf diseases using image segmentation and soft computing techniques".
[11]. **Vishal Mani Tiwari and Tarun Gupta**, (2017) "IPlant Leaf Disease
[12]. Analysis using Image Processing Technique with Modified SVM-CS Classifier", IJEMT.
[13]. **James, Gareth**, (1996) "An Introduction to Statistical Learning: with
[14]. Applications in R", Springer. p. 176. ISBN 978-1461471370.
[15]. **Ripley, Brian**, (1996) "Pattern Recognition and Neural Networks",Cambridge University Press. p. 354..
[16]. Yanbo Huanga, Yubin Lanb, Steven J. Thomsona, Alex Fangc, Wesley C. Hoffmannb,, (2010) "Development of soft computing and
[17]. applications in agricultural and biological engineering", Elsevier B.V.
[18]. Zhun Fan,Jiewei Lu,Maoguo Gong, Honghui Xie , Erik D.Goodman, (2018) "Automatic Tobacco Plant Detection in UAV Images via Deep Neural Networks", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing .
[19]. **Maria-Elena Nilsback and Andrew Zisserman**, (2010) "Automated
[20]. flower classification over a large number of classes", Visual Geometry Group, Department of Engineering Science University of Oxford, United Kingdom .
[21]. **Stephen Gang Wu,Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang Yi-Fan Chang and Qiao-Liang Xiang** ,(2007) "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network"