An Application to Identify the Availability of Medicinal Plants

MAGIN SUGI. S [1], SUPRIYA. M [2], S. OSWALT MANOJ [3]
[1,2,3] *Sri Ramakrishna Institute of Technology, Coimbatore*

Abstract -- Most of the people living in cities are not aware of medicinal plants around them and visits hospitals for any mild symptoms, but people in villages visits hospital rarely they treat themselves using the medicinal plants around them. The project focuses on medicinal plant identification based on leaf image classification and listing out their medicinal uses/values using an mobile application. To create awareness among the people this application is quite useful and long discussed topic to extract or measure leaf features. The urgent need is that many plants are at risk of extinction. There are thousands of plants known and unknown that yield to medicine or drugs and is of great use to men. We believe that the first step is to teach a Computer how to classify plants. The recognition of the tens of thousands of plants on earth is an important and difficult task, this makes an application in this field as a new challenge. Designing a convenient and automatic recognition system of plants is necessary and useful since it can facilitate fast classification of plants, understanding and managing them. Our project is based on identification and knowing the uses/values of medicinal plants in Tamil Nadu which runs as an android application. Major steps followed here are image processing- RGB to Grey scale and binary conversion, feature extraction, classification- SVM algorithm. Finally, it becomes our duty to know the plants and preserve it and make the people aware of it. The work done here is implemented using Image Processing and Support Vector Machine in MATLAB.

Indexed Terms: Support Vector Machine, Automatic Recognition, Image Pre-processing, Classification, Feature Extraction, MATLAB

I. INTRODUCTION

The world bears thousands of medicinal plants which have its own medicinal values, some are close to extinction and some are harmful to human. Plants are essential resource for human beings and form the base of all food chains. In India, there are 7200 medicinal plants of which its leaves may have medicinal properties. Identifying the unknown plants is a great challenge. To identify those plants correctly we should classify them based on the leaves, fruits or flowers images and feature extraction. Thus many processes involved in classifying these medicinal plants depend on human beings skill, knowledge and accumulation. Many researchers have been proposed the system for identifying and classifying plant species by its physical characteristics [1]. Our idea is to propose a system of automated classification by extracting plant features based on colour, shape and texture.

The collection of medicinal plants images its medicinal values by extracting features undertaking some pre-processing to identify their attributes, classification of leaves, fruits, flowers, formation of databases, training the dataset for recognition and finally evaluating the results. An automated Mobile Application to identify plant species and knowing their medicinal values can also be used by non-botanical experts for quick identification. There are more methods for classification, still macroscopic observation of leaf image classification is the first choice, here classification is based on leaf shape, size and morphological characteristics. The classification also includes shape of the leaves, arrangements, margin used for differentiation [2].

Leaves can be collected easily using mobile phones and image acquisition may be carried out. Therefore, our project is implemented using image processing, support vector machine in MATLAB and mobile application runs in Android studio.

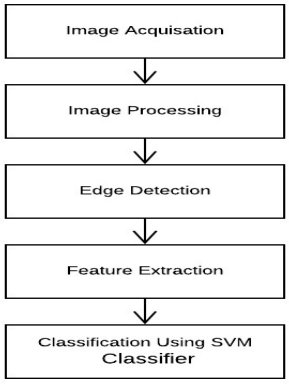

Fig. 1

II. RELATED WORKS

Plants identification systems have been developed by several researchers.

- R. Janani, A. Gopal:

Identified selected medicinal plant leaves using image feature extraction such as its shape, colour and texture and training an ANN classifier to identify the exact leaf class [1] with an accuracy of 94.4% with a minimum of 8 input features and implemented in MATLAB.

- Adams Begue, Venitha Kowlessur, Fawsi Mohamoodally, Upasana Singh, Sameerchand pudaruth:

From university of Mauritius proposed an automated plant identification system using computer vision and machine learning algorithms Random Forest (RF) which is better than other machine learning algorithms such as KNN, NB, SVM and NN with an accuracy of 90.1% by its feature extraction of each leaf.

- D. Venkatram and N. Mangayarkarasi:

From Amirta University created an automated system which identifies the plants and provides its medicinal values thus helping even a common man to be aware of medicinal plants around them using a computer vision based feature extraction. Thus a survey on various algorithm like PNN, SVM, PCA and TBA is to detect a plant.

- Ionel-Bujorel, Robert, Claudia-Maria, Andrei**:**

From Carol Davila University of Medicine and Pharmacy, Romania made a feasibility study of Romanian Medicinal plant recognition from leaf images and perform a feature selection procedure on the parameters which are used for the classification.

- Manoj Kumar P, Surya C.M., Varun P. Gopi:

From Government Engineering College, Kerala explores feature vectors from both the front and back side of a green leaf along with morphological features which has unique, optimum features combinations that maximizes the identification rate using leaf image dataset. Classification of leaves is based on unique feature combination. When tested over a wide spectrum of classifier identification rates up to 99% and when identification by dry leaves and feature vector combination then identification rates exceeding 94%.

- Vinita Tanjane and prof. N. J. Janwe:

Of Godwana University identified Medicinal Plant diseases using Canny Edge Detection Algorithm, Histogram Analysis such as Edge Histogram, Colour Histogram, CBIR, Image Processing and by feature extraction of leaf images with the accurate detection of diseased medicinal plants.

- Mohammed Shamrie, Taqiyah Khadijah and Rayner Alfred:

From University of Malaysia proposed a study on a framework to identify and classify tropical medicinal plants in Malaysia based on the extracted patterns of the leaf such as several angle features. Five classifier from WEKA and ensemble classifier called DECIML are used to compare the accuracies performance over this data.

III. METHODOLOGY

Simply, our project is based on 4 divisions that is: image acquisition, image processing, Edge Detection, Feature Extraction and Classification.

1. IMAGE ACQUISITION:

It is defined as the action of retrieving an image from some source, usually a hardware based source for processing. It is the first step in the workflow sequence because, without an image, no processing is possible. Here it is used for capturing medicinal plant.

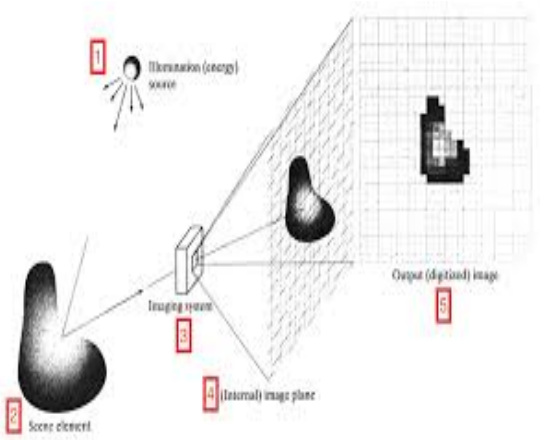

Fig. 2

2. IMAGE PROCESSING:

Image pre-processing is a program that processes its input image to produce output and that is used as input to another program like a compiler.

- IMAGE GREY PROCESSING:

Usually, the medicinal plant colour is green. However, the leaves colour on the same plant could be different due to various factors such as growth period and environment. To identify the type of plants, the leaves are considered as dominant features. The input is given as the digital image of the leaf part of any plant. The image pre-processing is done to remove any kind of external noise and shadow present in an image. The noise and shadow removed leaves then converted from RGB to grey-scale image which will be useful for feature extraction process.

- IMAGE DISCRETIZATION:

To perform threshold operation, the image is converted into binary image with only two values: black and white pixels. The threshold segmentation is used to separate the leaves from the background in order to extract the leaf contour. Finally to extract the leaf contour of binary image, the edge tracing method is used.

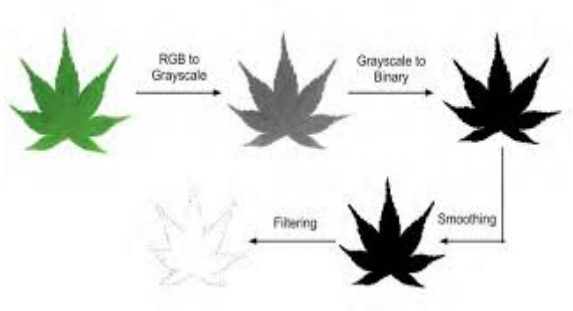

Fig. 3

3. FEATURE EXTRACTION:

Each leaf possess unique feature that is different from the other leaves.

- SHAPE FEATURE:

Some of the geometric shape features namely compactness, eccentricity, aspect ratio, etc., are extracted.

Length= Euclidean distance between two tip points on either side long the axis.

Breadth= length of the minor axis.

Aspect ratio= ratio between length and breadth of the leaf.

Area= area of pixel* total number of pixels.

Perimeter= count of the pixels having the leaf margin.

Rectangularity= (length*width)/area.

Compactness= 4 pi*area/perimeter^2

Eccentricity= sq.root (1-(1/aspect ratio)^2)

- COLOUR FEATURE:

Color features include mean and standard deviation which is used to characterize a color image.

Mean (u) = average sum of the total pixel in the leaf.

Standard deviation (sd) = 1/[mean-u]^2.

- TEXTURE FEATURE:

The texture provides information about the spatial arrangement of colour or intensity in a leaf image.

Features	Techniques	Accuracy	Dimension	Advantage	Disadvantage
Color Feature	Color Moment	Low	Low	Lower computational complexity	Precision is low
	HSV histogram	High	Medium	Simple ,Fast computation	No spatial information
	Color Correlogram	High	High	Includes the spatial correlation of colors ,Simple to compute	Very slow computation
Texture Feature	Gabor Filter	High	High	Achieves highest retrieval results	Computationally intensive
	Gabor Moment	Low	Low	Lower Dimensionality	Low retrieval result compare to Gabor filter
	Gray level co-occurrence matrix	High	High	Include positions of pixels having similar gray level values	High Dimensionality
Shape Feature	Moment Invariant	High	Low	Invariable to translation, rotation and scale	Limited recognition power
	Zernike moments	High	Low	Invariable to translation, rotation and scale	Computational Complexity is High

Fig. 4

IV. CLASSIFICATION

- SUPPORT VECTOR MACHINE:

Support vector machine (SVM) is a classification method that maps the input data to a high dimensional feature space through some nonlinear transformation that separated by optimal hyper plane, which maximizes the gap of positive samples and negative samples. SVM used the kernel function to transform the input data into a higher dimensional space and optimal hyper plane is constructed with maximum margin. The classification involved image features in this approach, therefore SVM classifier was trained using the one versus all approach. In this method, every class was trained with test cases of that class as positive and all other as negative. An open source android AAR library of the famous LibSVM is used to run the application.

Fig. 5(a)

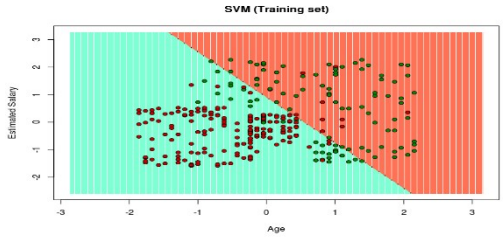

Fig. 5(b)

V. RESULT AND DISCUSSION

Our application run in android studio, application named as Medico-sap. Some of the snapshots are attached below:

Fig. 6(a)

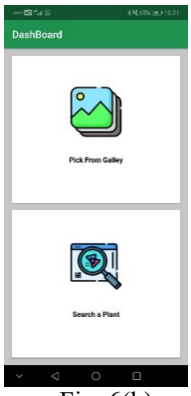

Fig. 6(b)

Front page of our mobile application. The page indicating how to get details of the medicinal plants and knowing their medicinal values. Here gallery indicates getting images from gallery and fetching its details and search indicates searching the plants by typing its names.

Fig. 6(c)

Fig. 6(d)

When we click gallery, our mobile gallery opens with saved images. By selecting a plant image it gives the details of the plant in the result page.

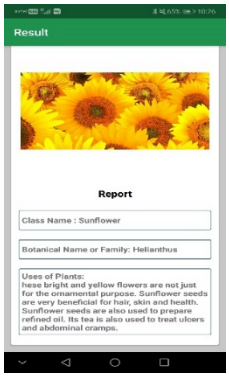

Fig. 6(e)

Parsing and fetching details from the datasets using SVM classifier. When search a plant button is clicked, search box appears. When we typed some plants name Result page shows searched plant details and medicinal values.

VI. CONCLUSION AND FUTURE WORK

Thus the plant images are stored in the dataset which can be viewed in gallery, the medicinal plant is chosen from gallery or searched by giving its name and its medicinal values are displayed in result page. We created both web application and mobile application, the snapshot for mobile application is inserted. The mobile application implemented in android studio and web application implemented in MATLAB. The classification is done using machine learning classifier known as Support Vector Machine. Image processing is used to covert the RGB images to grey scale image to extract its features such as colour, shape and texture and also used for boundary detection of the images.

In future, the work will be geared towards using huge dataset and high performance computing facilities. To the best of our knowledge the work presented here has unique image dataset of medicinal plants that are available in Tamil Nadu. In future, the dataset will consist of 2000 images of Tamil Nadu medicinal plants and camera option will be added.

REFERENCES

[1] R.Janani, A.Gopal," Identification of selected medicinal plants leaf using Image Features and ANN", in proceedings International Conference on Electronic Systems, 2013.

[2] Adams Begue, Venitha Kowlessur, Fawsi Mahomoodally, Upasana Singh, Sameerchand Pudaruth "Automatic Recognition of Medicinal Plants using Machine Learning Techniques", in proceedings International Journal of Advanced Computer Science and Applications (IJACSA), 2017.

[3] D Venkataraman, Mangayarkarasi N, "Computer Vision Based Feature Extraction of Leaves for Identification of Medicinal Values of Plants", IEEE International Conference on Computational Intelligence and Computing Research, 2016.

[4] Ionel-Bujorel Păvăloiu1, Robert Ancuceanu2, Claudia-Maria Enache3, Andrei Vasilăţeanu1, "Important Shape Features for Romanian Medicinal Herb Identification Based on Leaf Image", The 6th IEEE International Conference on E-Health and Bioengineering - EHB 2017.

[5] Manojkumar P., Surya C. M., Varun P. Gopi "Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples", Third International Conference on Research in Computational Intelligence and Communication Networks(ICRCIC), 2017.

[6] Vinita Tajane, prof. N.J. Janwe, "Medicinal Plants Disease Identification Using Canny Edge Detection Algorithm, Histogram Analysis and CBIR ", International Journal of Advanced Research in Computer Science and Software Engineering, 2014.