Predicting Medicinal Plant Based On Diagnostic Keys and Location Information

Shyam Vijayrao Pundkar (ME-IT II)
Dept. of Computer Engineering
Dattakala Group of Institutions,
Faculty of Engineering,
Swami-Chincholi, Pune, India.
Shyam.pundkar@gmail.com

Prof. M. M. Waghmare
Dept. of Computer Engineering
Dattakala Group of Institutions,
Faculty of Engineering,
Swami-Chincholi, Pune, India.
Monawaghmare25@gmail.com

***Abstract*—Ayurved is an oldest medicine System. The foremost ingredients of most of the Ayurved medicines are the medicinal plants. Botanist or researchers or students, identifies these plants using its morphological characteristics. The identification of herbs becomes difficult when they get partial access to plant's morphological elements (leaves, stem, flowers, fruit, seed, root and inflorescence). In this paper author put forward computer software, which can be used to predict medicinal plant based on its morphological characteristics along with its geographical location.**

Keywords— Medicinal Plants, Taxonomic Keys, Geographical locatio, Diagnostic keys

I. INTRODUCTION

Plant identification is crucial in research and business related to herbarium, medicine and cosmetics to name a few. Not all plants have medicinal value, some are poisonous, and some are carnivorous. Allopathic, Ayurvedic and Homeopathic Medicine uses Medicinal plant extract to some extent. To identify medicinal plant knowledge of morphological characteristics of the plants are required.

The morphological characteristics are the traits of plant that is visible and botanist can identify these characteristics by looking at the plants' leaves, fruit, root, steam, flower, inflorescence and stem. They need to match these characteristics against the database to identify the plants. More the characteristics available for matching, more accurate will be the results.

The research piloted in the area of plant prediction has used leave's morphological characteristics. The presumption is that the plant leaves availability all the time for its identification. But in reality only limited access may be available to user such as only fruit, only flower or the assortment of various elements of the plants. The proposed system not only uses these various morphological characteristics but also the location information to predict the medicinal plant.

The intention of including location information in the prediction process is because it is observed that some plants grow in particular geographical location only as the location supports its growth. Any location has its own characteristics such as temperature, humidity and altitude which may satisfy need of a plant. It would be difficult for a plant to survive in disfavoring location.

II. LITERATURE SURVEY

The research involving medicinal plant prediction has spanned into many domains of technologies. These are the medium using which the medicinal plant prediction has been attempted so far. These areas are Artificial Neural Network, Image Processing, Near Infrared Spectroscopy, DNA Barcoding.

A. Image Processing

Many researchers have used Image processing technique to identify medicinal plant using a particular part of it (typically leaf). It is not always likely for a researcher or botanist or student to get access to leaves; some time only stem or flower can be found. Even building such system of searching on several characteristics of the medicinal plant using image processing requires a huge database of imageries and enormous processing power to process this dataset.

Sathwik, T. et al[1]suggested a texture for processing of leaf's image method to identify medicinal plants.

Kusmana, I; Herdiyeni, Y.[2]. Suggested for identification of the medicinal plant the texture feature can be used.

Prudhveeswar Reddy; S.Gopal, A.; Gayatri, V[3] suggests leaf images for identification. Accuracy of 92% on 50 plant leaves has been achieved.

Wahyuni, N.K.S; Herdiyeni, Y.; [4] Suggests Fuzzy logic for identification with the accuracy of 74.51%.

Prof. N.J.Janwe; Vinita Tajane; [5] suggests use of image retrieval system for disease identification in plants based on its leaf's color and leaf's edge histogram.

Bhandarkar, P [6] suggests the uses of the boundaries of the leaves. Plant leaves, total 40, are used with accuracy of 67.5% is gained.

Charters, J. et al [7] suggested the venation structure of a plant's leaf for plant prediction. The performance gain of 6% is achieved there.

Harish, B.S. et al [8] suggests two different methods such as Zernike moments along with Morphological features for plant leaves classification.

Prasad, S. et al [9] suggested a squashed shape and hue feature extraction for plant identification.

Che Hussin, N.A et al [10] suggested Colour Moment Feature Transform for identification purpose.

Prasad, S. et al [11] suggested cut leaf images and transform on it to extract leaf information. The accuracy Obtained is of 95.6%.

Almeida, J. et al [12] suggested image processing to extract the features from images of leaf to identify the plants.

Verroust-Blondet A; Mouine, S.; Yahiaoui, I.; [13] suggested the leaf's main points and the leaf's side-line for identification.

Quang-Khue Nguyen [14] suggested the classifier frequencies at which the word appears for medicinal plant identification purpose.

B. DNA Barcoding

Analysis of DNA can be a complex and requires hours of time for obtaining the results. As per [15] the laboratory procedure costs between US$2 and $5 in most well-equipped laboratories, and less than US$2 in some cases. It takes a few hours in most laboratories and can be done in as little as 90 minutes. The results take time and requires laboratory for processing the sample. Even if we develop a database of the DNA barcodes and provide open access to it, it will still not be suitable in the real life scenario of field study as it requires quick results with minimum practical resources.

Natascha Techen[16] has appraised DNA barcoding. They come across the use of eminent genomic sections for prediction.

Andrea Galimberti[15] suggested barcode markers. For experiments, dissimilar researchers are using diverse sections for barcode markers.

C. Near-Infrared Spectroscopy

Near-infrared Spectroscopy is mostly used for analysing agriculture goods. This technique uses signature from spectroscope of the leaves for identification of medicinal plant. This signature is electromagnetic radiation from analyser. The challenge is obtaining these signatures using analyser when on field study. The additional device will definitely could be an overhead for medicinal plant prediction.

Kelina Sahaya Rajesh, P. [17] suggests use of various leaves to obtain signature to identify medicinal plant.

D. Neural Network

The neural network stores graph of connections to neurons in the database and simulate transmission of signals among them to simulate the real human brain intelligence. So to make artificial neural network work enormous amount of storage space and large processing power is required. The will results take time and will require large processing and storage. Again it won't be suitable for field study. There the resources will be limited with the constraints wanting results quickly.

Janani, R.; Gopal, A. [18] suggests the artificial neural network to predict the leaf of the plant.

Herdiyeni, Y.; Santoni, M.M. [19] suggest the use of texture, shape, and colour for plant identification. The experiment used 2448 colour images. Neural network is also used here to identify the plants.

Lei Zhang [20] uses images of venation and neural network to predict the medicinal plant.

Zulkifli, Z. et al [21] suggested neural network for classification purpose for medicinal plant.

The image processing along with neural network does hit the road block for storage, timing and training.

E. Other Areas

Lang, M. et al [22] suggests various number of traits of the medicinal plant for its prediction.

Nordin, Sharifalillah [23] suggest morphological keys repository for improved performance in the visualization, identification, and retrieval of medicinal plants.

However, to the best of the information examined by author, not a single existing application uses seven diagnostic keys or taxonomic keys such as flower, steam and leaves of plant along with terrestrial location information for prediction commonly used medicinal plants. This proposed system is based on the study of prediction using morphological individualities of medicinal plant, morphological individualities along with its position information (state level) and morphological characteristics along with its GPS location information.

The proposed system has the database of the medicinal plant's characteristics long with its state level geographical location information and GPS location information.

The experiments were carried out on a small dataset of medicinal plant species, the agenda of paper is: section 3 of the paper describes the proposed system; section 4 describes the experimental dataset; section 5 describes the algorithm performed over the dataset, section 6 describes the mathematical model, section 7 describes the experiment, section 8 analyze results, section 9 tells about technical analysis and finally section 10 discusses the future road-map and concludes the research.

III. PROPOSED SYSTEM

The architecture of the application is shown in Figure 1. It is separated in three layers:

User interface layer: Responsible for taking input and providing output to the user. It's the layer which is most close to the user.

Business Layer: The business of this application is to foretell the species based on the information provided by the user and the location information fetched by the GIS

extractor component itself. This layer does the same it uses the actual prediction algorithm to predict the medicinal plant.

Data Access Layer: This Layer is responsible for the data access activities. Such as providing an interface to business layer to fetch the data and store the data into the database.

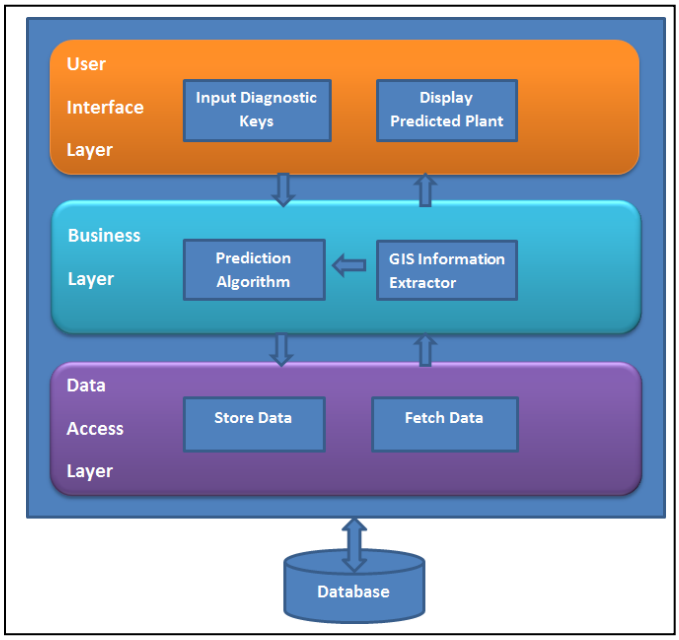

Figure 1. Architecture of the proposed system

Database: It is the actual location where the data is stored in the relational or document format.

IV. Experimental Dataset

The dataset is the 40 commonly used medicinal plants. These are mostly used by the Physician as well as the Ayurvedic drug manufacturers for drug preparation. Students studying Dravyaguna – a branch of Ayurveda – also study the medicinal plants.

The initial location information databases of these medicinal plants are created and restricted to states of India. GPS information used for this experiment is mocked and taken as a GPS location of a city of the state. The Algorithm for predicting the medicinal plants uses the technique of searching the dataset by diagnostic keys entered by the user and the current location information extracted by application or state information entered by the user.

V. Algorithm

The Figure 2. shows the proposed algorithm of the system. The database is created having all information related to diagnostic keys. The location information is also maintained in the same database. When a Researcher is at a particular Geographical location finds a fruit or a flower or a leaf, he/she can feed its diagnostic keys information into the application. The Current GPS location of researcher/students is taken by mobile device.

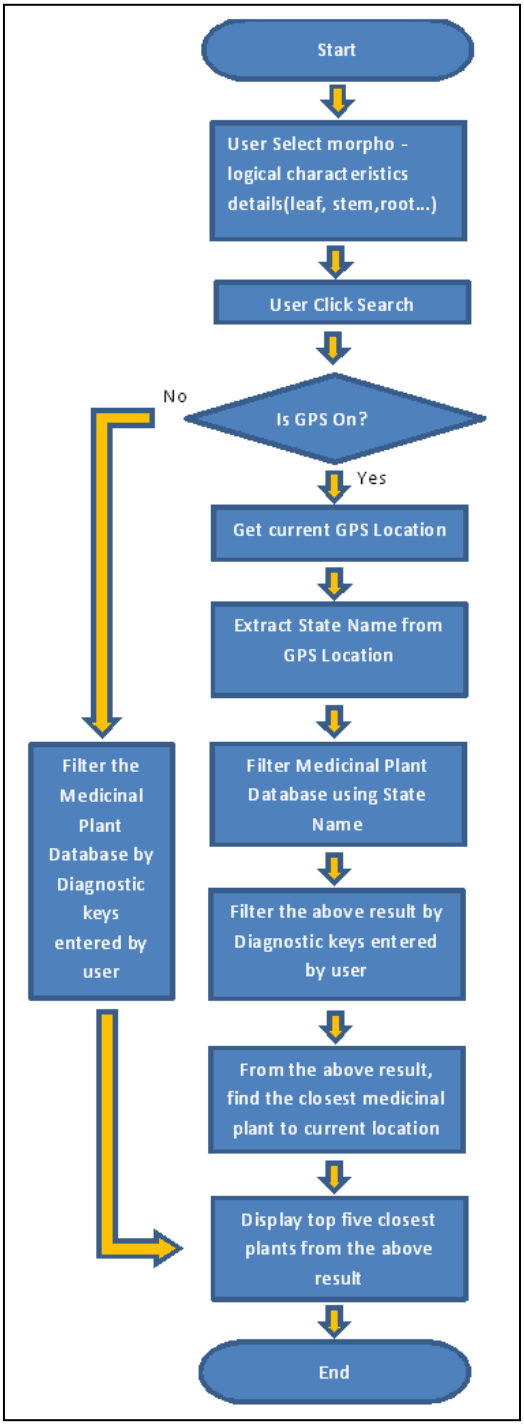

Figure 2. Proposed Algorithm

The application takes these diagnostic keys with location information and search the database with matching criteria. If user wants to search based on the state of the location then diagnostic keys and the state is taken as an input and search the database with matching criteria.

Using the location information the application suggests the medicinal plant by checking if it has the same characteristics and location information. In the absence of geographical location data application predicts the

medicinal plant using diagnostic keys only. Only Indian state level location Information is provided.

The algorithm uses the location information to limit the number of species on which the rest of the criteria such as diagnostic keys are applied. This approach improves the efficiency of the algorithm.

VI. MATHEMATICAL MODEL

The mathematical model for proposed work is as follows

Let, A Medicinal plant can be identified using eight characteristics.

These features of the plants are Feature Set (FS)

FS = {fruit, stem, root, seed, flower, leaves, inflorescence, location}

There is a Dataset of medicinal plants

DS = {p1, p2, p3, . . .}

R is the prediction algorithm it uses the FS and DS to predict the medicinal plant

FSi is the input feature vector to the algorithm R

xi is the predicted medicinal plants { x1,x2,x3,… }

xi = R(FSi;DSi)

VII. EXPERIMENTS

A. Experimental Setup

The experiment is performed on following hardware and software:

Hardware Requirement: Intel Core i5, 2.8 GHz processor with 3 GB RAM, Media Tek Dual Core Processor,

Software Requirement: Windows 8 64 bit, .NET4.0 Visual Studio, Android Ice Cream Sandwich 4.0, Google Maps, AppInventor and Android Studio.

B. Experiment Details

The experiment is done using three different methodologies, identifying using only three morphological characteristics, identifying using three morphological characteristics and geographical location information, lastly, identifying using seven morphological keys and spatial GIS Location information.

1) Using Morphological characteristics only (Existing Approach):

In this approach, only morphological characteristics of the observed part of a plant are used to predict the medicinal plant. This approach is based on the input of external morphological characteristics researcher or botanist provides to the system. User should enter characteristics keys of leaves and stem. Both of these keys should be entered into the system to increase the prediction accuracy. The results of this experiment are recorded.

2) Using Morphological characteristics and location information:

In this approach the search are made as specified in approach 1) but along with the location information in the form of the state of India.

3) Using Morphological characteristics and GPS Location Information:

In this approach the search results are constructed based on current GPS location and plants around it. The aim is to refine the results of approach 2).

The accuracy could be best in case of approach 3) because the possibility of finding the plant which is of same type nearby is high.

We created a windows based application having the required database of 40 species with their location information like the states in which they are found in India. Then we applied existing approach 1) and a new approach 2) as mentioned above.

Another approach 3) uses android based application. This application is required to capture the GPS location and it will also be convenient and helpful for the botanist or researcher on the field study.

VIII. RESULT ANALYSIS

The results of experiments are exemplified in the table below.

TABLE I. SEARCH USING MORPHOLOGICAL CHARACTERISTICS ONLY (EXISTING APPROACH 1)

Morphological Characteristics			**Accuracy (%)**
Leaves	***Stem***	***Flowers***	
Alternate	Unbranched	Incomplete	50
Alternate	Branched	Regular	11
Radical	Tall	-	33
Alternate	Branched	Complete	10
-	Underground	Unisexual	50
Petiolate	Straight	Complete	50
Alternate	Straight	Complete	50

In Table I above we can see that three Diagnostic Keys Leaves, Steam and Flowers are used to predict a medicinal plant. User has entered these three entries in the application and last column shows the experiment accuracy in percentage.

TABLE II. SEARCH USING MORPHOLOGICAL CHARACTERISTICS AND LOCATION INFORMATION (APPROACH 2)

Morphological Characteristics				**Accuracy (%)**
Leaves	***Stem***	***Flowers***	***State***	
Alternate	Un-branched	Incomplete	Maharashtra	60
Alternate	Branched	Regular	Goa	25
Radical	Tall	-	Arunachal Pradesh	100
Alternate	Branched	Complete	Andhra Pradesh	20
-	Underground	Unisexual	Jammu and	70

Morphological Characteristics				Accuracy (%)
Leaves	***Stem***	***Flowers***	***State***	
			Kashmir	
Petiolate	Straight	Complete	Maharashtra	60
Alternate	Straight	Complete	Mizoram	65

In Table II above we can see that three Diagnostic Keys Leaves, Steam, Flowers as well as location information are used to predict a medicinal plant. User has entered these four entries in the application and last column shows the experiment accuracy in percentage.

In Table III results, the third approach is taken into consideration where the Plant's GPS location is deciding factor. There would be seven morphological keys and the GPS Location information as an input. The plants which are nearby will be predicted as the plant in the query having its morphological characteristics entered into the system for identification.

TABLE III. SEARCH USING SEVEN MORPHOLOGICAL CHARACTERISTICS AND GPS LOCATION INFORMATION (APPROACH 3)

Experiments	1	2	3	4	5	6	7
Leaf	Alternate	Alternate	Radical	Alternate	-	Petiolate	Alternate
Stem	Un-branched	Branched	Tall	Branched	Undergroun d	Straight	Straight
Flowers	Incomplete	Regular	-	Complete	Unisexual	Complete	Complete
Root	-	-	Rhizomes	-	-	-	-
Inflorescence	Compound spadix	Terminal	Stiff	female	Spadix	Terminal	Axillary
Seed	-	Numerous	Obpyramidal	Endosperm	-	-	-
Fruit	Drupe	elliptical	Minutes	Drupe	Berry	obliquely	Woody
GPS Location	18.62962 8, 73.77716 2	15.49129 3, 73.82747 3	27.08 4681, 93.60 4696	16.507 783, 80.648 924	34.0832 24, 74.7967 24	18.62962 8, 73.77716 2	23.72724 9, 92.72051 1
Location Name	Pune, Maharashtra	Panjim, Goa	Itanagar, Arunachal Pradesh	Vijaywada, Andhra Pradesh	Srinagar, J&K	Pune, Maharashtra	Aizawl, Mizoram
Accuracy (%)	75	45	100	35	73	61	68

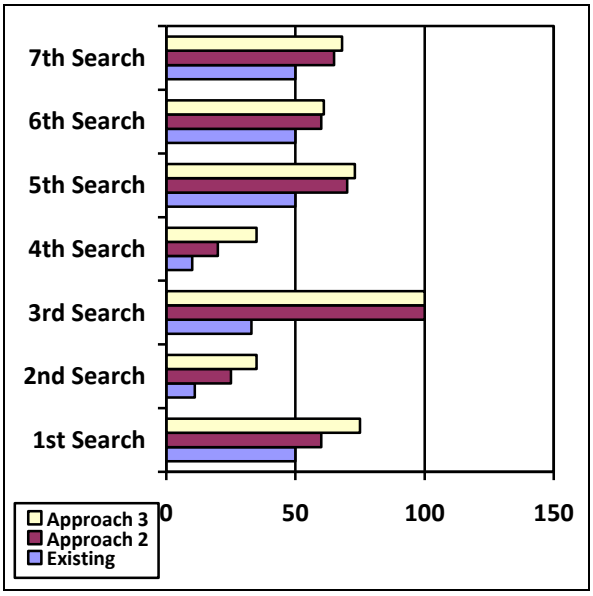

Figure 3. Comparison between approach 1,2,3

Figure 3 above shows the comparison between approaches existing approach 1, approach 2 and approach 3. We can clearly see that adding location information in medicinal plant prediction improves the accuracy of the results. In the experiment, 20% of improvement is recorded for 40 species of medicinal plant. Approach 3) had enhanced the result further by 8%.

IX. TECHNICAL ANALYSIS

The total size for 40 medicinal plant dataset is 74 KB. That means around 2 KB per medicinal plant. This includes seven morphological characteristics along with its location information both state level and spatial. Currently only one GPS location per medicinal plant is saved.

The total time required for prediction is in milliseconds in case of state level location and 30 seconds to a minute in case of GPS location using mobile device.

X. CONCLUSION AND FUTURE WORK

The medicinal plant prediction can be done using several different techniques. Adding Geographical location information is helpful in increasing the accuracy up till 20% based on experiment mentioned in this paper. As per this paper, the spatial information further enhances this result to 8%. The Geographic Information System (GIS) information is used to mark the location of the plant and store this information in database for future predictions.

To further experiment on this, we will be taking a much larger dataset of medicinal plants. More precise location information may also improve the accuracy of prediction.

ACKNOWLEDGMENT

We wish to express our sincere thanks to our Principal, HOD, Professors and staff members of

Computer Engineering Department at Dattakala Faculty of Engineering, Swami Chincholi, Pune. We are very grateful to the authors of various articles available, helping us become cognizant of the research currently ongoing in this field. Last but not the least, We would like to thank all our Friends and Family members who have always been there to support and helped us to complete this research work.

The author would also like to thank Dr Yogini Kulkarni (MD, Phd in Ayurved) and Dr. Poonam Thakare (MD-PG Scholar) for the efforts taken by them for validating the results.

REFERENCES

[1] Sathwik, T. Yasaswini, R. ; Venkatesh, R. ; Gopal, A."Classification of selected medicinal plant leaves using texture analysis", Chennai Centre, Central Electron. Eng. Res. Inst., Chennai, India ; Pages 1-6 July 2013

[2] Herdiyeni, Y. ; Kusmana, I. ; Dept. of Comput. Sci., Bogor Agric. Univ. (IPB), Bogor, Indonesia ; Kusmana, I."Fusion of Local Binary Patterns features for tropical medicinal plants identification" Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference Sept. Page 353 - 357, 2013

[3] Gopal, A. ; Prudhveeswar Reddy, S. ; Gayatri, V. ―Classification of selected medicinal plants leaf using image processing" Machine Vision and Image Processing (MVIP), 2012 International Conference, Page(s): 5 – 8, 2012

[4] Herdiyeni, Y. ; Wahyuni, N.K.S.,"Mobile application for Indonesian medicinal plants identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram"Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference,Page(s): 301 - 306,2012

[5] Vinita Tajane Prof.N.J.Janwe, "Medicinal Plants Disease Identification Using CannyEdge Detection Algorithm, Histogram Analysis and CBIR",Computer Science Engineering, Godwana University,, India, June,2014

[6] Bhandarkar, P. ; Doshi, H. ; Kadakia, M. ; Ahmed, R. "Leaf identification using morphology and structural decomposition", Signal Processing and Integrated Networks (SPIN), International Conference on DOI: 10.1109/SPIN.2014.6776934 Page(s): 125 – 129, 2014

[7] Charters, J. ; Zhiyong Wang ; Zheru Chi ; Ah Chung Tsoi ; Feng, D.D.―EAGLE: A novel descriptor for identifying plant species using leaf lamina vascular features" Multimedia and Expo Workshops (ICMEW), 2014 IEEE International Conference on DOI: 10.1109/ ICMEW .2014.6890557 , Page(s): 1 – 6, 2014

[8] Harish, B.S. ; Hedge, A. ; Venkatesh, O. ; Spoorthy, D.G. ; Sushma, D―Classification of plant leaves using Morphological features and Zernike moments" Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on DOI: 10.1109/ ICACCI.2013.6637459, Page(s): 1827 – 1831, 2013

[9] Prasad, S. ; Peddoju, S.K. ; Ghosh, D. ―Mobile plant species classification: A low computational aproach" Image Information Processing (ICIIP), 2013 IEEE Second International Conference Page(s): 405 – 409, 2013

[10] Che Hussin, N.A. ; Jamil, N. ; Nordin, S. ; Awang, K. ―Plant species identification by using Scale Invariant Feature Transform (SIFT) and Grid Based Colour Moment (GBCM)" IEEE Conference, Page(s): 226 – 230, 2013

[11] Prasad, S. ; Kumar, P. ; Tripathi, R.C. ―Plant leaf species identification using Curvelet transform" Computer and Communication Technology (ICCCT), 2nd International Conference, Page(s): 646 – 652, 2011

[12] Almeida, J. ; dos Santos, J.A. ; Alberton, B. ; Morellato, L.P.C. ; Da S Torres, R. ―Plant Species Identification with Phenological Visual Rhythms" eScience (eScience), 2013 IEEE 9th International Conference, Page(s): 148 – 154, 2013

[13] Mouine, S. ; Yahiaoui, I. ; Verroust-Blondet, A. ―Plant species recognition using spatial correlation between the leaf margin and the leaf salient points" 2013 20th IEEE International Conference, Page(s): 1466 – 1470, 2013

[14] Quang-Khue Nguyen ; Thi-Lan Le ; Ngoc-Hai Pham ―Leaf based plant identification system for Android using SURF features in combination with Bag of Words model and supervised learning" Advanced Technologies for Communications (ATC), 2013 International Conference, Page(s): 404 – 407, 2013

[15] I Bruni, A Galimberti, L Caridi, D Scaccabarozzi, F De Mattia, M Casiraghi, M Labra ―A DNA barcoding approach to identify plant species in multiflower honey" , 2015

[16] Natascha Techen, Iffat Parveen, Zhiqiang Pan, Ikhlas A Khan," DNA barcoding of medicinal plant material for identification" Pages 103-110, 2014

[17] Kelina Sahaya Rajesh, P. ; Kumaravelu, C. ; Gopal, A. ; Suganthi, S.," Studies on identification of medicinal plant variety based on NIR spectroscopy using plant leaves", Advanced Computing Technologies (ICACT), 2013 15th International Conference, Page(s): 1 – 4, 2013

[18] Janani, R. ; Gopal, A. ,"Identification of selected medicinal plant leaves using image features and ANN", Advanced Electronic Systems (ICAES), 2013 International Conference, Page(s): 238 – 242, 2013

[19] Herdiyeni, Y. ; Santoni, M.M.," Combination of morphological, Local Binary Pattern Variance and color moments features for Indonesian medicinal plants identification", Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference, Page(s): 255 – 259, 2012

[20] Lei Zhang ; Jun Kong ; Xiaoyun Zeng ; Jiayue Ren, ―Plant Species Identification Based on Neural Network" Natural Computation, 2008. ICNC '08. Fourth International Conference, Page(s): 90 – 94, 2008

[21] Zulkifli, Z. ; Saad, Puteh ; Mohtar, I.A. ―Plant leaf identification using moment invariants & General Regression Neural Network", Hybrid Intelligent Systems (HIS), 2011 11th International Conference, Page(s): 430 – 435, 2011

[22] Lang, M. ; Kosch, Harald ; Stars, S. ; Kettner, C. ; Lachner, J. ; Oborny, D.,Recognition of Botanical Bloom Characteristics from Visual Features", Image Analysis for Multimedia Interactive Services, 2007. WIAMIS '07. Eighth International Workshop, Page(s): 27, 2007

[23] Nordin, Sharifalillah ; Mohd Sapiyan Baba ; Itam, Khairuddin," Plant taxonomic keys repository", Research and Innovation in Information Systems (ICRIIS), 2011 International Conference, Page(s): 1 – 6, 2011