May – June 2020
ISSN: 0193-4120 Page No. 5770-5775

Pattern Recognition and Classification of Herbal Plants Using Leaf Images

C.S. Udaya[1], M. Usha Rani[2]
[1]Dept. Computer Science, Smt N.P.S Govt. College(W), Chittoor, India,
mudayag@gmail.com)
[2]Dept. Computer Science, SPMVV, Tirupati, India
musha_rohan@gmail.com)

Article Info
Volume 83
Page Number: 5770 - 5775
Publication Issue:
May - June 2020

Article History
Article Received: *19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *17 May 2020*

Abstract
Healing by using medicinal plants is old method of treatment. Natural products derived from plants for the treatment of diseases have proved that there is no existence of life without plants. Plants are essential foundation of medicine. Ayurvedic treatments are used by people all over the world for centuries. These treatments are made by using parts of medicine plants or herbal species. The use of herbs is incredibly important for human and animals. The use of herbal medicines has many advantages compared to pharmaceutical products, such as reduce risk of side effects, lower cost, large availability and it also very effective in chronic conditions. So the Researches and utilization of herbal medicine in the treatment of diseases increased every day. But herbs are very similar in their shape, color and texture. So it is difficult to a normal people to identify the herbs. if there are no proper training, there may be chances to misidentification, it may leads to serious problems. So identification of the herbs is digitally challenging object for better classification. In this paper we present the current methods for plant identification, using leaf images.

Keywords: *Medicinal herbs, Pattern recognition, Ayurvedic treatment, Pharmaceutical products, classification.*

I. INTRODUCTION

For our Natural Wealth, Medicinal plants are very important and main source. There is no existence of life without plants and trees. Herbal plants naturally incorporate and gather some metabolites. Medicinal plant posses healing properties or applies beneficial psychopharmacology medicine effects on the human being body. Those plants that possess therapeutic value along with the nutritional value can be considered as Neutraceuticals. The conservation of medicinal plants are not only a good tool for biodiversity conservation, but also has great prospects for livelihood enhancement [1].

Chinese, Egyptians, Indigenous culture used herbs in their medical system. Medicinal plants are common source of medicine. 600 to 700 plants based Medicines are available today. People are get so much awareness of medicinal plants that is what during past few years, worldwide utilization of medicinal plants and demand of medicinal plants has been increased. Because of the awareness, the importance of the medicinal plants also increased phenomenally [1].

The dependence of human being on plants dates back to the start of the human race. Herbal remedies that have the ability to heal and boost physical and mental well-being. Majority of the people in the world is struggling to raise living standards and improve their health due to increasing poverty and population. Peoples who are living in the remote areas or hilly areas are mostly relying on Medicinal plants. Medicinal plants are being used in the treatments of diseases in ancient system of medicine such as Ayurvedic, Unani, and Chinese traditional medicine [2].

The estimation of the World Health Organization is that 80% of the people in Worldwide are dependent on Medicinal plants for their Health Issues. The hunt for new medicines has engaged ethnobotony and ethnopharmaiology. The use of herbal medicines has many advantages compared to pharmaceutical products, such as reduce risk of side effects, lower cost, large availability and it also very effective in chronic conditions.

But herbs are very similar in their shape, color and texture. So it is difficult to a normal people to identify the herbs. if there are no proper training, there may be chances to misidentification, it may leads to serious problems. So identification of the herbs is really very challenging object in the form of digital world, for better classification.

Data mining helps to recognize the important attributes of herbal leaves with data sets of different higher dimensions. But Identification of the right plant is very important for the we should know the types of plants in detail and complete and scientifically accountable. Reorganization can be done in various ways through herbarium, flower, fruit, roots, stem, leaves etc,. but identification of the leaves has to consider.

II. PATTERN REORGANIZATION OF THE MEDICINAL LEAF

Pattern Reorganization of the Medicinal Leaf usually fallows few phases.

Step 1: In this first need to take leaf images as input. Leaf Images can be captured from Digital Cameras, Datasets, and Smart Phones. These collective images are called Image Input. This can be used for classification [2].

Step 2: After collecting the leaf images, then pre-processing step is to be taken. Here we have to do image rejuvenate, pruning, gray scaling, binary thresholding, removing the noise, contrast stretching, threshold inversion, and edge recognition. These can be done by using K Means clustering, Gaussian, Filter, Graphcut, Gradcut.

Step 3: Most challenging part is Feature Extraction. Here we need to use different feature extraction techniques like

For Texture: GLCM, 2D-DWT, GTSDM, LBG, HOG, SURF, MRS

For Shape: CAPSI, SIFT

For color: RGB, HSV, YCrCb

For Edge: Prewitt Edge, Canny, Edge Detection, Laplace filtering, Sobel operator

Step 4: After Feature Extraction, we need to classify the herbal plants by using Training dataset. For classification we can use the following algorithms like SVM, ANN, PNN, PCA, K-NN, DT, ET, ED, MLP, Binary Classification Tree, Multiboost Classifier, Random forest.

Step 5: After following the above steps finally we will get output results [2,4].

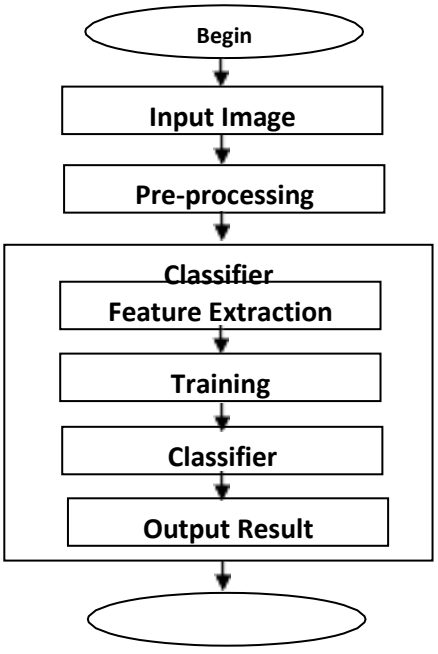

Fig 1. General Model for Identification and Classification of Herbal Plants [2]

May – June 2020
ISSN: 0193-4120 Page No. 5770-5775

III. PRE PROCESSING ALGORITHMS

K-means Clustering: The main goal of clustering is to get meaningful intuition from data to separate them based on their appearances. Based on their characteristics we need to segregate all the objects into groups. K-mean Clustering performs well for this purpose. Because it is very easy to understand, it is faster than hierarchical clustering because it has number of variables. It can be used in the fields like Image compression, Image segmentation, Image recognition, pattern analysis etc.

Need to group the based on the similarities, for the we can use K-mean Clustering algorithm. It is an unsupervised algorithm and an iterative algorithm that first it partitions a group of data of n values into K subgroups, with the nearest mean all of the n values are belongs to the k cluster. It computes centroids and iterates until it finds the optimal centroid. 1. By using algorithm, first needs to specify how many number of clusters should be generated known as 'K'. 2. Next assign each data point to the cluster by randomly selecting k data points. To find the data points, we can use Euclidean Distance, that are closest to their centerW of the cluster. 3. After that compute the mean value of all points, which finds clusters centroids. 4. Until all the points are assigned to their clusters, Need to repeat the steps 1,2, and 3.

1. Gaussian filter: it is studied in image processing and computer vision. Gaussian filter can be used for noise suppression, from that the noise is smoothed out and the signal is distorted. By using Gaussian filter we can develop and analyze real-time and accurate filters for nonlinear filtering problems.

IV. FEATURE EXTRACTION ALGORITHMS

Texture feature extraction: The texture extraction algorithms performs two major operations that is first it extracts the features , for that it generates image signature or feature vectors accurately based on the content of each image in the database. A feature vector is much smaller in size than the original image. Second task is similarity measurement, here we need to compute the distance between the query image and each image in the database, so that the top closest images can be retrieved. To perform these tasks we can use CBIR, GLCM, LBG, KPE etc. algorithms. For texture feature extraction CBIR is 9 times faster than GLCM method, and also the precision and recall values are much higher as compared to conventional GLCM. In case of LBG and KPE, KPE performs marginally better than LBG in terms of slightly higher precision and recall values [3].

Shape feature extraction: For Shape feature extraction we can use SIFT algorithm, it is used to achieve the desired fault tolerance. SIFT descriptors comprising the local distribution function of the image gradient orientations are extracted at each sampling key point location over a local measurement aperture. It follows four steps **1. Key point detection -** need to select key points as local interest points and characterized by their localization (x, y), scale σ and orientation θ, p(x, y, σ, θ) **2. Orientation assignment** – Calculae a local histogram of gradient orientation, weighted by the gradient magnitude to determine the main orientation association with key point. The histogram is calculated on scale dependent neighborhood. **3. Descriptors extraction** – should arrange each key point p(x, y, σ, θ) to describe its local geometry. To obtain translation and scale invariance, a square neighborhood id defined around each key point with size depending on σ. To ensure rotation invariance, need to rotate it by angleθ . The SIFT descriptor is obtained by combing and normalizing these histograms. **4. Key point matching** – key points of two different images are matched according to their respective descriptors. Euclidean distance is calculated, nearest neighbor is chosen [4,12].

Color feature extraction: HSV (Hue, Saturation and Value) can be used for color space and for color histogram YCrCb (Luminance/red Chrominance/ blue Chrominance) can be used.

Usually color images are converted to gray scale for

May – June 2020
ISSN: 0193-4120 Page No. 5770-5775

computational reasons. In case of shape, texture feature extraction the color information is not needed, so we can convert the image into gray scale. When color extraction is need, then the captured image is considered. Color histogram helps to find the color distribution of pixels in the image, it gives variance. For this we need to follow few steps 1) first we need to capture image in RGB color space as input. 2. Convert into HSV and YCrCb color space 3. Extracts the color features, Hue (H), Saturation (S), Intensity Value (V), Luminance (Y), red Chrominance (Cr), Blue Chrominance (Cb) components. 4. By using HSV we can obtain the histogram 5. By using YCrCb we can obtain color histograms 6. Finally normalize the histogram [12].

Edge Feature extraction: Recognizing the edges is very importance in pattern recognition of the Medicinal plants.. Before extracting the edge feature first to remove noises, for that we can use Wavelet transform algorithm. Edges are boundaries of the leaf and the background. By getting information of edges we can differentiate the target and background. By using the Edge Enhancement operator we can obtain outstanding partial edge of the image. After that by setting threshold, we need to define pixels the edge intensity, then we get edge points. Some edge points are removed because it may produce interruptions from which we may get existing noise and image dark. Then obtained edge points are connected to be a line. For edge feature extraction, commonly used operators are prewitt, canny, Log, Sobel, Binary morphology etc., The Prewitt operator not only detects edge points, but also restrain the noise. The Sobel operator has similar function as the Prewitt operator, but the edges detected by the Sobel operator are much wider than prewitt. Roberts operator is a kind of the simplest operator, the borderline of the extracted image is quite thick, so edge location will not that much accurate. The Log operator is used to filter and counting the differential coefficient of the image. By using the convolution of revolving symmetrical Log template and the image, the log operator determines the zero overlapping position of filter output [6].

V. CLASSIFIER ALGORITHMS

After Feature Extraction, we need to classify the herbal plants by using Training dataset. Classification of the data is one of the major tasks in machine learning where the new data has to be classified to one of the classes to which it belongs. For classification we can use the following algorithms like SVM, ANN, PNN, PCA, K-NN, DT, ET, ED, MLP, Binary Classification Tree, Multiboost Classifier, Random forest [4,5].

Support Vector Machine (SVM) algorithm: SVM has great generalization potential and is exceptionally robust. SVM Contains an algorithm, which deducing a function from given data. This method is used to classification and reverting analysis of the data. SVM is a useful technique for data classification. Even though it's considered that Neural Networks are easier to use than this, however, sometimes unsatisfactory results will be obtained. Generally classification can be done in between training and testing data which consist of some instances of data [7,8].

Every instance of the training set contains many attributes and one target values. Need to produce a model that must predict target value of the data instances in the testing data set. The target values are given by the attributes. This is the main goal of SVM [8].

We need to check whether data points are separated with a dimensional hyper plane, this can be done in SVM by viewing the data points as dimensional vector. A hyper plane decides, the representation has the largest separation between the two categories or class, and then it may be considered as the best [5].

Even when the prediction of unknown samples is not necessary, then also Feature selection and SVM classification together may be used. They can be used to identify key sets which are involved in

May – June 2020
ISSN: 0193-4120 Page No. 5770-5775

whatever processes distinguish the classes The Main property of SVM id Duality [5,8].

Convolution Neural Network (CNN): In recent years, CNN has gained popularity in computer vision applications including plant and leaf recognition. The application of deep layers in CNN is possible due to the availability of vast amount of data and powerful computer devices. Information on feature and classifier is not needed when applying CNN. Large amount of data consumes a long training time. To overcome this limitation, utilizing pre-trained CNN [9-11].

VI. CONCLUSION

History of Ayurveda says every plant has a medicinal value, so the identification of which part of the plant has medicinal value and for which disease this medicine is going to use is very essential to mankind. Medicinal plant parts like leaves, flowers, bark, seeds, fruits, roots, and stem are used in many disease diagnosis. Through these plants parts botanists and herbal practitioners are identifying the medicinal plants manually which is time consuming process and also herbs are very similar in their shape, color and texture. So it is difficult to a normal people to identify the herbs. if there are no proper training, there may be chances to misidentification, it may leads to serious problems. The objective of this study is to reduce manual work and increase the efficiency by the automating the identification of medicinal plants using image processing techniques. Automatic identification and classification of medicinal plants will give knowledge to identify the herbs, which is very helpful to common people and farmers to increase production of such essential plants. This automatic classification system also helps botanists, consumers, forestry services, taxonomists, pharmaceutical companies and Ayurveda practitioners to identify and classify the medicinal plants without any human assistance.

REFERENCES

[1] Mohammad Abdul Motaleb, "Selected Medicinal Plants of Cittagong Hill Tracts" in "IUCN International Union for Conservation of Nature" - Bangladesh Country Office, 2011.

[2] Rajani S, Veeba M.N "STUDY ON IDENTIFICATION AND CLASSIFICATION OF MEDICINAL PLANTS" in "International Journal of Advances in Science Engineering and Technology" ISSN(p): 2321 –8991, ISSN(e): 2321 –9009 Vol-6, Iss-2, Spl. Issue-2 Jun.-2018, http://iraj.in

[3] H.b.Kekre, Sudeep D, Tanuja K. Sarode and Vashali Suryawanshi "Image Retrieval using texture Features Extracted from GLCM, LBC and KPE" in "International Journal of Computer Theory and Engineering". Vol. 2, No. 5, October, 2010 – 1793-8201

[4] Muhammad Azfar Firdaus Azlah, Lee Suan Chua, Fakhrul Razan Rahmad, Farah Izana Abdullah and Sharifah Rafidah Wan Alwi - "Review on Techniques for Plant Leaf Classification and Recognition" - 21 October 2019

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

[6] Feng-ying Cui and Li-jun Zou Bei Song "Edge Feature Extraction Based on Digital Image Processing Techniques" Proceedings of the "IEEE International Conference on Automation and Logistics Qingdao", China September 2008

[7] D. Venkataraman; N. Mangayarkarasi "Support Vector Machine Based Classification of Medicinal Plants Using Leaf Features" in International Conference on Advances in Computing , Communication and informatics (ICACCI) – 2017.

[8] Vikramaditya Jakkula, School of EECS - "Tutorial on Support Vector Machine (SVM)", Washington State University, Pullman - 2011

May – June 2020
ISSN: 0193-4120 Page No. 5770-5775

[9] Zaidah Ibrahima, Nurbaity Sabrib and Dino Isac, "Multi-maxpooling Convolutional Neural Network for Medicinal Herb Leaf Recognition" Proceedings of the 6th "IIAE International Conference on Intelligent Systems and Image Processing" - 2018

[10] Agus Ambarwari Yeni Herdiyeni, Taufik Djatna – "Combination of Relief Feature Selection and Fuzzy K-Nearest Neighbor for Plant Species Identification" – in "International Conference on Advanced Compute Science and Information System (ICACSIS)" - 2016.

[11] Y. R. Azeez, C. Rajapakse, "An Application ofTransfer Learning Techniques in Identifying Herbal Plants in Sri Lanka" in "International Research Conference on Smart Computing and Systems Engineering (SCSE)" - 2019.

[12] Basavaraj S. Anami, Suvarna S. Nandyal, A. Govardhan – "A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants" in "nternational Journal of Computer Applications" (0975 – 8887) Volume 6– No.12, September 2010

May – June 2020
ISSN: 0193-4120 Page No. 5770-5775