Shodhshauryam, International Scientific Refereed Research Journal
Available online at : www.shisrrj.com
OPEN ACCESS
© 2024 SHISRRJ | Volume 7 | Issue 2
ISSN : 2581-6306
doi : https://doi.org/10.32628/SHISRRJ

# Medical Herbs Identification Using Machine Learning

**K Naresh[1], M Teja[2]**

[1]Assistant Professor, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India

[2]Post Graduate, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India

**Article Info**

**Article History**
Received : 02 April 2024
Published : 13 April 2024

**Publication Issue :**
March-April-2024
Volume 7, Issue 2

**Page Number :** 519-525

**ABSTRACT**

In Ayurveda, plants have been utilised as a source of medicine since the Vedic era. The most crucial step in making ayurveda medicine is correctly identifying the plants. step, which was completed by hand. Because of the need for mass production, it's critical to automatically identify these plants. In this research, we implement an ensemble supervise machine learning system based on colour, texture, and geometrical aspects called the random forest algorithm for medicinal plant identification.

**Keywords :** Medicinal Plant, Supervise learning, Image Processing, Identification, Automatic, Ayurveda.

## I. INTRODUCTION

Approximately 5000 years ago, during the Vedic era, India developed the ancient medical system known as Ayurveda. The primary ingredients found in ayurvedic medications are plant leaves as well as additional plant elements like the bark, roots, etc. It has been discovered that over 8000 plants with Indian origins have therapeutic properties. Herbal remedies from various Indian systems use combinations of a tiny subset of these plants, totaling 1500. Specifically, 500 of these herbs are used in commercial Ayurvedic treatments. The majority of plants used in ayurvedic formulations—more than 80%—are gathered from forests and wastelands, with the remainder being grown on agricultural land.The Ayurvedic doctors of old selected the therapeutic plants and made the remedies for their patients themselves. The majority of plants are recognised bytheir foliage Common procedures for classifying a plant's leaf include taking a picture, adjusting the size and noise level, extracting features, applying the suggested methodology, and ultimately identifying or recognising the plant.

## II. EXISITING SYSTEM AND PROPOSED SYSTEM

### A. Existing system

The Artificial Neural Network (ANN) model was the foundation of the current medicinal plant

encourages significant contributions to the study and practice of herbal medicine.

**Advantages:**

1. High accuracy
2. Larger and Diverse dataset
3. Enhanced Robustness
4. Potential for scalibility
5. Automation and time efficiency

## III. LITERATURE SURVEY

1. An Android-based computerized plant identification system H. A. Chathura, Priyankara, and D. K. Withanage are authors of the paper.

The characteristics seen on plant leaves are sufficient to set them apart from other species. One of the fundamental problems in digital image processing is the identification of plants from leaf photographs.
These image processing methods frequently use shape-based digital morphological cues to detect leaves. Extremely a few studies has focused on handheld devices, despite the fact that several studies were done on finding plants via leaves. In this study, we present a plant identification system that uses leaf images and blends the classifier developed by the Support Vector Machine (SVM) with the Pack of The user (BOW) model.

with SIFT characteristics. After 20 species were trained to be classified, the accuracy level of the system was 96.48%. Based on the findings, we created an Android app that connects to the server and makes it possible for users to recognize a variety of plants

by utilising smartphone photos of plant leaves. Ayurvedic medicinal plant identification using processing leaf sample images The authors of this work are P. M. Krishna, C. M. Surya, and V. P. Gopi. In the ayurvedic medical field, Finding the right botanical medicines for utilization in the production of drugs is essential. The shape, hue, & feel of a leaves
are the primary characteristics needed to identify a medicinal plant. Deterministic characteristics for species identification are present in the colour and texture of the leaf on both sides. This work investigates morphological characteristics and feature vectors from the front and rear of a green leaf to determine the optimal feature combination that optimises the identification rate.Scanned photos of the front and rear surfaces of leaves from popular ayurvedic medicinal plants are used to construct a database of medicinal plant leaves. The distinctive feature combination is used to categorise the leaves. Experiments spanning a broad range of classifiers have yielded identification rates as high as 99%. By including identification by dried leaves into the previously mentioned study, a combination of feature vectors is created that allows for identification rates to surpass 94%.

3.Artificial neural network identification of a leaf from a Philippine herbal medicine plant CO-AUTHORS: R. G. de Luna and others

The work described in the above article employs a method which identifies and recognizes specific Philippine herbs.

combines the use of artificial neural networks with image processing techniques to extract pertinent leaf attributes. Twelve distinct herbal medicine plant leaves are sampled in real life, with each leaf

captured in a single photograph. Several image processing techniques are used to extract various aspects. The technique can identify the type of leaf used in medicinal products plants.

under examination by using an artificial neural network that functions as an independent Neural network design. Furthermore, the device can provide information concerning the ailments that a

herbal plant can treat. A 600-image features dataset, with 50 photos from each plant utilized as a herbal remedy for training. With Python's assistance, a neural network framework with tailored settings is developed, resulting in 98.16 % id for the whole dataset. To evaluate the true effectiveness of the system, a neural network simulation created in MATLAB is applied to a distinct set of 72 example images of herbs. The

findings of the experiment show that the accuracy of herbal plant identification is 96.61%.

4.Identifying plants with new geometric traits using conventional data mining techniques

AUTHORS: AsdrÃºbal LÃ³pez-Chau and Rafael Rojas-HernÃ¡ndez

Plant identification is inside a certain data mining application domain. Plant leaves are typically the primary feature that set one plant apart from another. Feature extraction is required for accurate identification. The majority of plant recognition systems reported in the literature combine characteristics with a classification algorithm that has been adjusted or modified for usage in this kind of situation. Three novel geometric properties that explain In this paper, the skyline and diagonal harmony of results in is suggested.

It is easy to extract these features from photos. Experimental results show that when these qualities are combined with additional well-known geometric properties, The effectiveness of traditional categorization techniques has significantly increased.

We evaluate seven classifiers using publicly accessible leaf photos from the Internet to demonstrate the efficacy of the approach.

## IV. METHODLOGY

### DESCRIPTION OF MODULES:

**Dataset**: We created the system to obtain the input dataset in the first module of the Medicinal Herbs Identification course. The process of gathering data is the initial step in the process of actually creating an automated learning model. The quality of the simulation is going to impacted during this essential phase within an increasing approach; your more and better information I discover, the more powerful it will be. There are other ways to collect the data, including as manual interventions and online scraping. Our dataset may be found in the model folder of the project. The dataset is sourced from Kaggle, a widely used standard dataset repository that is used by academics worldwide. There are 17,973 photos of medicinal plants in the dataset. The dataset referenced from Kaggle may be found at this This link will take you to a dataset for medicinal plants created by Jaya Prakash Pondy on Kaggle.com. importing the necessary libraries; you will require Python. The necessary libraries, such as the pandas, numpy, matplotlib, and a tool such as will first be loaded. Keras, also will be used to build a primary hypothesis, sklearnn will divide up the test and training information, so PIL is going to transform the photos into an array of numbers.

**Getting the photos**: In this module, we will take the images out of the dataset and transform them into a format that will allow us to use them for both model testing and training. Reading the photos, scaling them, and normalising the pixel values are all necessary for this. The pictures and their labels will be retrieved.Next, resize each image to (180,180) so that they are all the same size for identification. Next, turn the pictures into a numpy array.

**Dataset division**: The picture dataset will be split into training and testing sets for this module. Divide the dataset into test and train sets. There are 20% test and 80% train data. This will be carried out in order to test the model on untested data in order to assess its correctness, validate the model's performance, and train the model on a subset of the data. Divide the dataset into test and train sets. There are 20% test and 80% train data.

**Accuracy on test set**: The model's accuracy will be evaluated on the test set subsequent to its training and assessment on the validation set. The model's performance will be assessed in part by measuring the accuracy on the test set. On the test set, we achieved 96.79% accuracy.

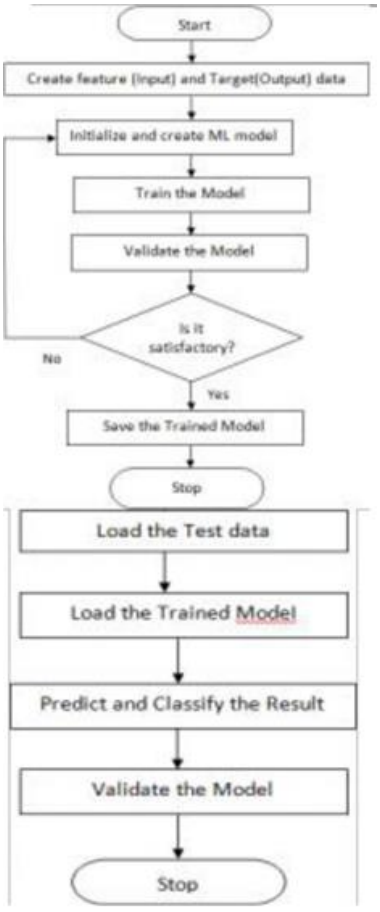
Storing the Trained Model: Save your tested and created model as an.h5 or.pkl file via a framework such as pickle.

as soon as you are comfortable enough to introduce it into a production-ready environment.Verify that pickle is installed in your setting.The next step is to

Save the design to an.h5 document after importing it into the module.


Fig. 1 : Elodie, Alovera, Coriander, Drumstick, Hibiscus, Mint, Neem, Papaya, Palem, Rui, and Tulsi are medicinal plants._Basil Model of training

**Flow Chart**



## V. EXPERIMENTAL SETUP

### 5.1 Algorithm used

Multiple array data, neural networks are able to handle challenges like processing a color image composed of three two-dimensional arrays containing pixel intensities in every single one of the three color channels. Numerous arrays can represent a wide range of data modalities,

including: 3-dimensional video or holographic photographic; two-dimensional (D) pictures or music spectral images; & one pixel messages and scenes, such as speech

. Convolutional neural networks (ConvNets) leverage the characteristics of natural signals through four main concepts: multiple layer usage, Geographic links, combining, and sharing weights. A structure of a standard ConvNet is broken up into multiple stages. Layers made up of pooling and layers with convolution are the layers that comprise the first phases. The units of the convolutional layer, as the name suggests, are split into map patterns, with all of them connected to specific patches in the data maps of the layer above by an adjacency bank, which is a set of weights. Subsequently, a non-linearity such as a Reynolds Linear Option (ReLU) applies on the localised calculated result. The units of a feature map share a common filter bank.

Within a layer, distinct feature maps employ various filter banks. This architecture serves two purposes. First, local clusters of values in array data—like images—are frequently highly connected, resulting in recognisable, easily identifiable local themes.

### 5.2 Image understanding with deep convolutional networks

ConvNets and recurrent net modules are combined in a recent amazing presentation to generate image captions. Current ConvNet topologies consist of billions of connections between units, tens to 20 layers of ReLUs also and tens of billions of weights. Training such a multitude a few years ago could've needed fortnight; with to advances in software, technology, and technique speed, it can now only take a few hours.ConvNet-based imaging algorithms execute sufficiently successfully that most major tech companies, which includes Facebook, Google, and others. Microsoft, Oracle, IBM, Yahoo!, Twitter, or Adobe, along with an increasing quantity of start-ups, have initiated R&D projects and deployed ConvNet-based visual understanding solutions and services. ConvNets are easily and effectively implemented in field-programmable gates or circuits.

## VI IMPLEMENTATION

An output that meets the demands of the end user and effectively conveys the material is considered high quality. The way that processing results are communicated to users and other systems is through any system's outputs. The output design determines how data is moved for instant usage and the physical output. It is the most important and accessible source of information for the user. Well-considered and efficient output design improves system integration and streamlines user decision-making.

1. The process of designing computer output should be methodical and well-planned; appropriate output should be created while making sure that

every output component is made in a way that makes the system simple and efficient for users to utilise. When analysing computer output, designers should pinpoint the precise output required to satisfy specifications.

2. Decide how information will be presented.
3. Create documents, reports, or other formats with data that the system has created. The output form of an information system should accomplish any or all of the following objectives. Share information about what was done in the past, the current state of affairs, or what is anticipated in the future. 1. List important events, opportunities, problems, etc

Utilize the knowledge as the system collected to develop written materials, reports, or other types of information. An data system's final type should achieve every one of a few things goals. Talk about what has been performed using the past, how things are going now, or what's planned for the future.

1. Enumerate significant occurrences, chances, issues, etc.
1. cautions.
2. Set off an event.
3. Verify a course of action.

**Outcome**

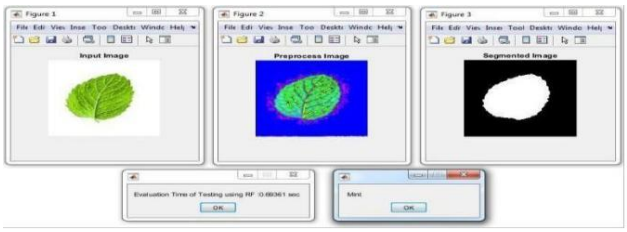


Fig. 2 : Understanding an unidentified image as Tulsi

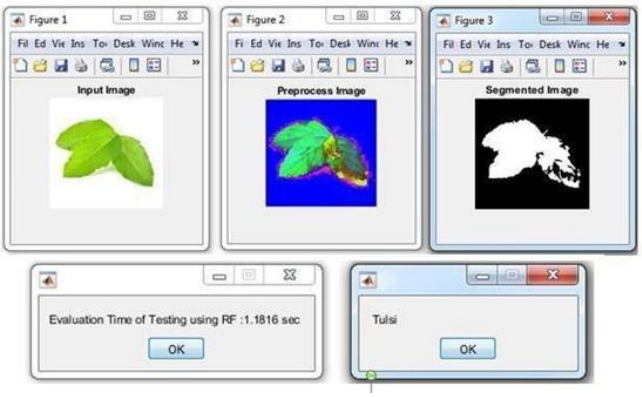
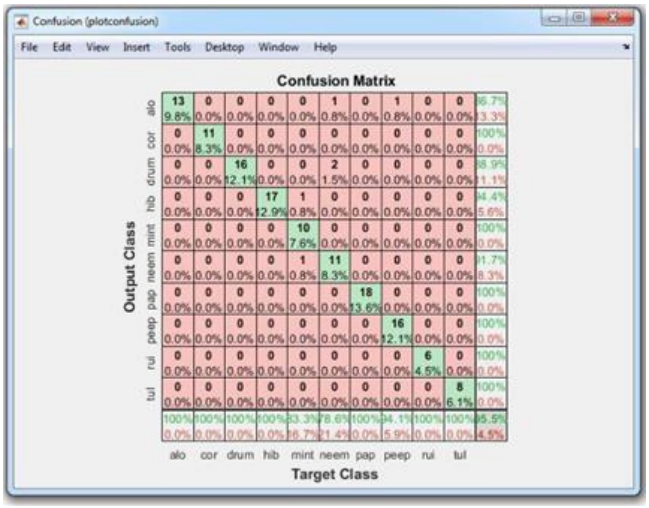



Fig. 3 : Recognizing an unidentified image as Mint




Picture -4: the System's Context Matrix Testing

**Outcome Table -1**

| **Parameters** | **Performance-testing parameters (percentage)** |
|---|---|
| Percent accuracy | 94.54 |
| Indicates the sensitivity | 96.23 |
| Specificity | 50.00 |
| F-score | 97.70 |

## CONCLUSION

The To determine the proper variety of medicinal plant, the random forest algorithm is an ensemble managing machine learning process based on hue, texture, and geometrical properties.

method for medicinal plant identification that we have applied in this study. With the help of shape, colour, and texture traits, leaves may be identified with 94.54 percent accuracy.The technique's highly encouraging findings highlight how suitable this algorithm is for systems that identify therapeutic plants. In the future, this approach can be expanded to include more plant species with higher precision.

## FUTURE ENHANCEMENT

The suggested method can develop further by investigating these directions for future research, increasing its precision, resilience, and usefulness for the goal of finding medicinal herbs. These developments will support research on herbal medicine, conservation initiatives, and the identification of novel uses and beneficial qualities of medicinal plants.

## REFERENCE

[1]. "Computer assisted plant identification system for Android," 2015 Moratuwa Engineering Research Conference (MERCon), Moratuwa, 2015, pp. 148-153.

[2]. P. M. Kumar, C. M. Surya and V. P. Gopi, "Identification of ayurvedic medicinal plants by image processing of leaf samples," 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, 2017, pp. 231-238.

[3]. R. G. de Luna et al., "Identification of philippine herbal medicine plant leaf using artificial neural network," 2017IEEE

[4]. 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, 2017, pp

[5]. Rafael Rojas-HernÃ¡ndez and AsdrÃºbal LÃ³pez-Chau, "Plant identification using new geometric features with standard data mining methods", Networking Sensing and Control (ICNSC) 2016 IEEE 13th International Conference on, pp. 1-4, 2016.

[6]. Wang, X., D. Huang, J. Du, H. Xu, & L. Heutte, 2008. Classification of plant leaf images with complicated background.