Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Medicinal Plant Identification and Care for Disease

Priyanka Indoriya[1], Dr. Asha Ambhaikar[2]

[1]*Research scholar, MATS University, Raipur, M.P., India, priyanka10sonkar@gmail.com*
[2]*Director, MSEIT, MATS University, Raipur, (C.G.), India, drambhaikar@gmail.com*

In Ayurveda, plants were highly valued sources of medicine during the Vedic era. However, when demand for mass production increases, it becomes impractical to identify these plants by hand for medical purposes. In this research, an automated method for CNN Classifier-based medicinal plant identification is proposed. Through the examination of color, texture, and geometric characteristics taken from plant photos, this technique seeks to precisely categorize medicinal plants and offer additional details regarding the disease's properties and course of treatment. The precision and efficiency are further enhanced by the use of image processing techniques. Meeting the growing demand for herbal medicines and streamlining production procedures are benefits of automating this crucial phase in the preparation of Ayurvedic medication.

Keywords: Ayurveda, Medicinal plants, Disease treatment, Machine learning, CNN.

1. Introduction

The goal of this project is to identify plants by using digital image processing and machine learning techniques to recognize their leaves' shape, color, and texture. Plants were considered to be extremely important sources of medicine in Ayurveda throughout the time of the Vedas. However, with the increasing demand for mass production, the manual identification of medicinal plants becomes problematic and demands extensive prior knowledge. Consequently, this research suggests an automated method that combines conventional Ayurvedic knowledge with contemporary technology. Based on input leaf photos, the study uses a variety of image-processing approaches to identify different plant species.

Our method attempts to reliably identify the plant and provide information about its medicinal characteristics and ailments connected with its leaves. It does this by analyzing the color, texture, and geometrical properties retrieved from plant photos.

Image processing is the main technique used in our work to help classify plant leaves for identification. Results from the experiment show that identifying plant species based on leaf traits may be done with more accuracy, which bodes well for a quicker and more effective plant identification procedure. By combining image processing techniques, plant identification becomes even more precise and effective, closing the gap between conventional Ayurvedic medicine and contemporary technological developments.

The overall goal of this research is to improve automated plant identification methods that could find use in a variety of fields, including characteristics. By fusing ancient knowledge with cutting-edge technology, it promises to address the growing demand for herbal treatments, expedite manufacturing processes in the preparation of Ayurvedic medicine, and promote breakthroughs in healthcare practices.

In the research paper, section II gives a brief description of previous research work performed in this field.section III describes the methodology and result analysis shown in Section IV and finally conclusion is shown in section V.

2. LITERATURE REVIEW

Many researchers have contributed to this field. Using established structural features, ParagBhandarkar, Rizwan Ahmed, and colleagues [3] dissectedthemorphologyofleafedgesand produced a structural signature that quantifies the leaf. characteristic of the form. For calculating the identity, they employed the root mean square error between the feature vectors of the input image and the image in the database. The authors' database is made up of 40 leaf samples from 10 different species. They acquired a 66.5% categorization rate overall, which is unaffected by the size or orientationoftheleaves.Theidentification rateis not high enough to beusefulin real-world applications. For NadiPariksha, Roopini et al. [2015] designed a device. It preprocessesthethreesignalsusingoptical pulse sensors using an eighth-order Butterworthfiltertoreducenoise.Thethree signals are visualized using three pulse sensors. Then, using artificial neural networks,thepulsedataisfurtherclassified intoVata,Pitta,andKapha.Fourgeometric features—convexity, solidity, eccentricity, and circularity—as well as three RGB color features—redness, greenness, and blueness indices—were used in the experiment by Pavan Kumar Mishra, SanjayKumarMaurya,etal.[7].adatabase of1000photosofleavesthattheauthors have assembled. To accelerate the identification process, feature vector comparisonsaremadeinthreestages.They attained an 85% rate of identification.

Alam et al., [2017]. A study on three machine learning techniques focused on selecting between two learning algorithms basedoncalibratedtesting.Thetopicofthe Comparative Study of J48, Naive Bayes, and One-R Classification Techniques for Credit Card Fraud Detection using WEKA is the use of the OneR, Naive Bayes, and J48 machine learning approaches to mould a tool for the indicators of potential fraud occurrences. Contribution from Arthur S., etal.,[2020]everythingthatincorporatesor encompasses machine learning is, in essence, machine learning theory. The applicationofmachinelearningtheorypits real-world applications against the illimitablepresumptionsthatareassociated with its abstract concepts. Barde et al, [2022]presentedtheViolaJonesalgorithm method for face detection. It is also useful inthe caseofillumination-invariantfaceimages [3-5].SnehlataBarde et al., [2022] described thetraditional method for the prediction of suitable herbs where the Rajvedhya examines the pulse is known as NadiPriksha in the current scenario is changed many technologies is developed, andmachinesareabletoperformsucha task. They designed a system for all the people who prefer to use herbal medicine fordiseases.Thatpredictsthesuitableherbs for the disease by the characteristics of Ayurveda herbs by using the classification technique of machine learning. For this, theycollected200herbssamplesanddefine the suitability of herbs for disease.

3. Methodology

The research focuses on image identification, finding the features of plants and the remedies for the suitable diseases shown in the block diagram.

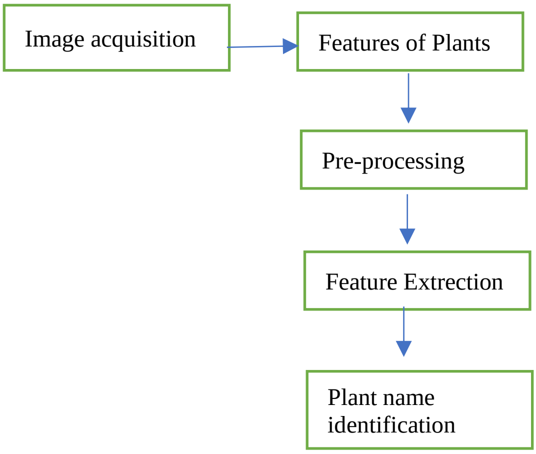

Fig.1: Block diagram for the research work

a) Image acquisition: -

With over 150 photos per plant and over 40 different types of plant species, the Indian medicinal plant dataset we gathered is nearly evenly distributed throughout all classes.

Fig.2 Sample Images

Source:Indian Medicinal ⚕ Leaves 🌿 Dataset | Kaggle

To better comprehend the type of photos we have and the necessary operations and transformations to develop the model that will identify the plant, Fig. 2 displays representative images from the dataset.

b) Properties of Plants

The characteristics of plants and how they are used in different ailments. Mango, Tulsi, Amla, aloe vera, ashwagandha, Ashoka, and many more are among the plants we have included.

These photos are taken from a variety of perspectives and angles, capturing a greater range of images to optimize our model's training. The collection has been constructed to capture the natural diversity and complexity found in plant pictures. Multiple viewpoints are used to capture a range of views, angles, and environmental variables in the images. This methodical methodology seeks to capture a broad spectrum of visual characteristics, such as leaf forms, colors, textures, and structural details, which are essential for thorough model training and precise classification. The Pixel Intensity Distribution for a specific image within the sample dataset is depicted in Figure 3. This aids in gaining a better understanding of the machine's representation of the image and how it interprets the image (Discrete image).

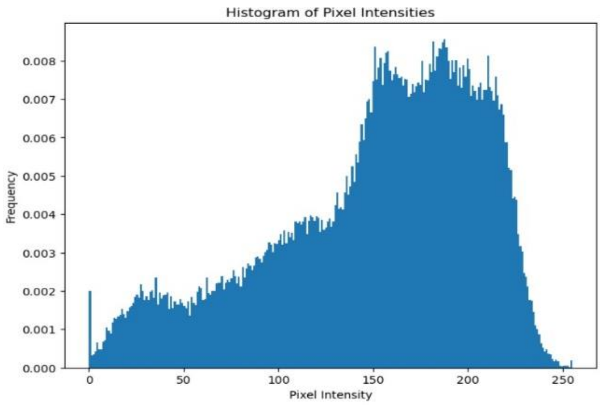

Fig.3Discrete Image distribution

c) Pre-Processing for images

Because the input photos can be noisy or contain irrelevant features, they might not be the greatest match to produce the best performance model for confidently identifying the plant species. Thus, it is imperative to employ efficient methods for image cleaning and enhancement before using this dataset to generate a classification model representation [2].

The following are the changes we made to the input image dataset:

1. Resize the original image to (180,180,3) to lower its high resolution, whichrequires more calculation.

2. Enhancement: The quality of the characteristics we extract from the input image will be impacted by the image's smoothing.

d) Feature extraction and representation

1. Feature vectorization is the process of representing a picture into an array of intensity-based pixels ranging from 0 to 255.
2. Normalisation: To lessen the variance in the pixel values, this is done after the image's pixels are retrieved. This is accomplished by dividing the image's or vector collection's maximum intensity by every pixel value.
3. Train, Validation, and Test Split: To test our model impartially on the newly discovered data, we set aside some data in addition to the ones that will be used to train the model and

update the weights in the backward propagation.

About 70% of the data are utilized for training, 20% are used for validation, and the remaining data are used to assess the model's performance[6].

4. RESULT AND ANALYSIS

a) ML Model Selection and Training

A sequential model with an input shape of (180,180,3) has been employed. Three convolution layers are added to the design, which successfully captures the patterns that the model utilizes to identify the plant and determine its appropriate applications. Following the extraction of a pattern from the input training images' features, the vectors are transformed into a single 1-D array of 3-D images (the color channel serves as the third dimension in this case). To incorporate resilience in the model, a random dropout is implemented before flattening the input picture pixel matrix. This dropout drops the neuron with a likelihood of 20%.Subsequently, the ANN is designed to acquire new features, validate them against a validation dataset, and initialize both forward and backward weight adjustments based on the errors. Approximately 27 is the ideal hyperparameter for the number of iterations on the training data images; this value can be adjusted based on the model's complexityand learning rate.The model architecture, or summary of the model we created to categorize the 40 plant types using the TensorFlow framework, is shown in Figure4.

Model: "sequential"

Layer (type)	Output Shape	Param #
rescaling_1 (Rescaling)	(None, 180, 180, 3)	0
conv2d (Conv2D)	(None, 180, 180, 16)	448
max_pooling2d (MaxPooling2D)	(None, 90, 90, 16)	0
conv2d_1 (Conv2D)	(None, 90, 90, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 45, 45, 32)	0
conv2d_2 (Conv2D)	(None, 45, 45, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 22, 22, 64)	0
flatten (Flatten)	(None, 30976)	0
dense (Dense)	(None, 128)	3965056
dense_1 (Dense)	(None, 40)	5160

Total params: 3993800 (15.24 MB)
Trainable params: 3993800 (15.24 MB)
Non-trainable params: 0 (0.00 Byte)

Fig.4 Architecture of the CNN model

b)Model Evaluation and Performance

Out of the entire training validation dataset, we obtain about 0.8 of the plant photos that are correctly identified after fine-tuning the hyperparameters improving the model network and adding the dropout of 0.2. The iterative technique shown in Figure 5 determines the ideal number of training iterations on the training image set to achieve the best results on the testing images.

The most crucial KPI matrix for assessing how well the model performs when evaluating performance on unseen data is how well the model recognizes fresh plant photos figure 6.

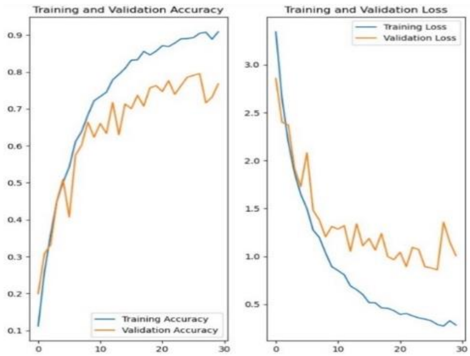

Fig.5Training History

Precision: - The proportion of plants that truly belong to each category class, out of the total number of plants classified in that class. We may state that, for a given plant image, we projected that the predicted class would match the actual plant image 89% of the time because the weighted average precision value for all the classes is 0.89.

Recall: - It displays the percentage of the times we correctly identify a certain plant based on its image, derived from a real class of that image. For every class, the weighted average of the recallsis 0.88.

The F1-score indicates when recall and precision are balanced. The problem's objective willdetermine the focus.

	precision	recall	f1-score	support
Aloevera	0.86	0.90	0.88	164
Amla	0.90	0.97	0.93	146
Amruta_Balli	0.85	0.79	0.82	146
Arali	0.99	0.98	0.99	146
Ashoka	0.88	0.97	0.92	146
Ashwagandha	0.98	0.87	0.92	146
Avacado	0.94	0.93	0.93	146
Bamboo	0.99	0.95	0.97	146
Basale	0.97	0.92	0.94	146
Betel	0.95	0.82	0.88	151
Betel_Nut	0.85	0.99	0.91	146
Brahmi	0.82	0.90	0.86	146
Castor	0.59	0.88	0.70	160
Curry_Leaf	0.78	0.95	0.86	146
Doddapatre	0.96	0.94	0.95	146
Ekka	1.00	0.80	0.89	146
Ganike	0.91	0.83	0.87	115
Gauva	0.78	0.82	0.80	146
Geranium	0.98	0.88	0.92	146
Henna	0.64	0.94	0.76	150
Hibiscus	0.83	0.95	0.89	165
Honge	0.95	0.87	0.91	146
Insulin	0.92	0.95	0.94	146
Jasmine	0.90	0.74	0.81	187
Lemon	0.98	0.92	0.95	146
Lemon_grass	0.93	0.86	0.90	146
Mango	0.74	0.92	0.82	146
Mint	0.89	0.99	0.94	153
Nagadali	0.94	0.88	0.91	152
Neem	0.87	0.73	0.79	146
Nithyapushpa	0.95	0.95	0.95	146
Nooni	0.86	0.90	0.88	146
Pappaya	0.98	0.88	0.93	146
Pepper	0.92	0.82	0.87	146
Pomegranate	0.95	0.71	0.81	146
Raktachandini	0.96	0.72	0.82	146
Rose	0.89	0.83	0.86	168
Sapota	0.98	0.86	0.91	146
Tulasi	0.89	0.86	0.87	146
Wood_sorel	0.90	0.92	0.91	146
accuracy			0.88	5945
macro avg	0.90	0.88	0.88	5945
weighted avg	0.89	0.88	0.88	5945

Fig.6Classification Report

Our problem's primary goal is to concentrate on the Recall. The success of our case will be determined by how many plants are correctly classified into the actual class to which they belong.

c) Robustness Testing:

Manual method vs Automated Identification

Anyone cannot perform the manual approach of identifying plant leaves and their medicinal characteristics since it requires specialized knowledge of numerous Ayurvedic plants and medication. With the automated identification approach, anyone may reveal the name, health properties, and uses of medicinal plants for treating a wide range of illnesses, from cough and cold to gum disease, stomach issues, and more. This means that no prior knowledge or expertise in identifying medicinal plants is required.

After being educated on hundreds of photos, the automated system successfully recognizes the type of plant each photograph belongs to. This system is reliable and unaffected by noise in real-world situations. In practical applications, this can identify the plant with ease.

The model's outcome is very unexpected because traditional identification requires knowledge of the plant, takes a long time, and cannot be done on a big scale. 88% of the time, the identification of the unseen data across all plants is done correctly. In terms of accessibility, scalability, and efficiency, the automated identification approach fared better than the manual method. It made it available to a larger audience by doing away with the requirement for specialized knowledge. The automated identification model demonstrated resilience to fluctuations and noise, demonstrating its dependability in practical situations.

Here are a few examples of test image identifications shown in Figure 7 to9:

1. Amla:-

Fig.7 Amla

2. Bamboo: -

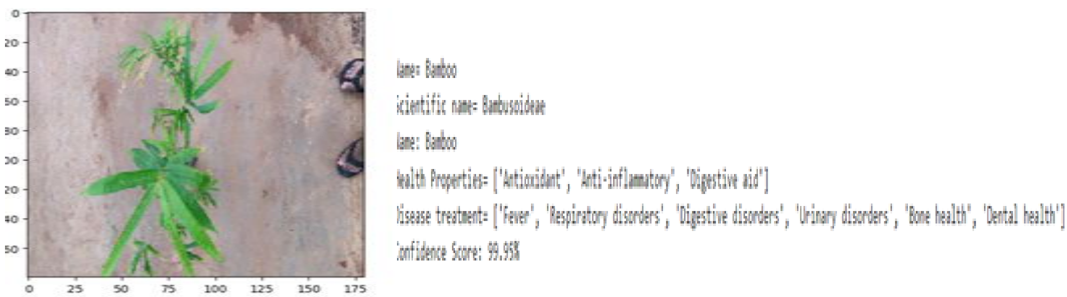

Fig.7 Bamboo

Carry Leaves: -

Fig.8 Carry Leave

This image most likely belongs to Betel with a 4.99% confidence.

{'Name': 'Betel', 'Scientific Name': 'Piper betle', 'Health Properties': ['Antibacterial', 'Digestive aid', 'Oral health'], 'Proper Diseases': ['Oral health', 'Digestion', 'Respiratory ailments', 'Diabetes', 'Joint pain']}

Fig.9Aloe vera

5. CONCLUSION

In summary, an automated CNN-based model for detecting medicinal plants based on leaf attributes was effectively built and evaluated in this study. The study proved the model's efficacy in correctly categorizing plant species and offering insightful details about their therapeutic qualities and applications. Through the integration of contemporary machine-learning methods with conventional Ayurvedic knowledge, this study advances the development of herbal medicine and healthcare procedures. The automated identification approach is a useful tool in the field of Ayurvedic medicine since it is scalable, efficient, and accessible.

References

1. Azizi, A.; Gilandeh, Y.A.; Mesri-Gundoshmian, T.; Saleh-Bigdeli, A.A.; Moghaddam, H.A. Classification of soil aggregates: A novel approach based on deep learning. Soil Tillage Res. 2020, 199, 104586.

2. B.R. Pushpa, N.S. Rani, Ayur-PlantNet: an unbiased lightweight deep convolutionalneural network for Indian Ayurvedic plant species classification, J. Appl. Res. Med. Aromat. Plants 34 (2023) 100459.
3. P. Indoriya and S. Barde, "Prediction of Ayurvedic Herbs for Specific Diseases by Classification Techniques in Machine Learning," 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 2022.
4. Azadnia, R.; Kheiralipour, K. Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier. J. Appl. Res. Med. Aromat. Plants 2021, 25, 100327.
5. DhanalaxmiGaddam, A Survey on NadiPareeksha for Early Detection of Several Diseases Computational Models using Nadi Patterns, International Journal of Computer Science and Information Technologies, 2015.
6. Baliga, M.S.; Bhat, H.P.; Joseph, N.; Fazal, F. Phytochemistry and medicinal uses of the bael fruit (Aeglemarmelos Correa): A concise review. Food Res. Int. 2011, 44, 1768–1775.
7. Dileep, M.R., &Pournami, P.N. (2019). AyurLeaf: A Deep Learning Approach for Classification of Medicinal Plants. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), 321-325.
8. J.G. Thanikkal, A.K. Dubey, M.T. Thomas, An efficient mobile application for identification of immunity-boosting medicinal plants using shape descriptor algorithm, Wirel. Pers. Commun. (2023) 1–17.
9. Rajesh, K.V.N., Bhaskari, D.L. (2021). Optimum Dataset Size for Ayurvedic Plant Leaf Recognition Using Convolution Neural Networks. In: Sharma, N., Chakrabarti, A.,Balas, V.E., Bruckstein, A.M. (eds) Data Management, Analytics and Innovation. Lecture Notes on Data Engineering and Communications Technologies, vol.70 Springer, Singapore.
10. P.Manoj Kumar, C. M. Surya, Varun P. Gopi, Identification of ayurvedic medicinal plants by image processing of leaf samples, 2017 Third International Conference on Research in Computational Intelligence and CommunicationNetworks (ICRCICN), 2017