IDENTIFICATION OF AYURVEDIC MEDICINAL PLANT FOR SPECIFIC DISEASES BY CLASSIFICATION TECHNIQUES IN MACHINE LEARNING

Priyanka Indoriya[1*], Vijayant Verma[2], Snehlata Barde[3]

Article History: Received: 04.05.2023 Revised: 14.06.2023 Accepted: 08.07.2023

Abstract

India is significant in Ayurveda medicine and treatment since it is a repository for numerous medications. Until recently, only rishi-Munis or Rajvidhyas were aware of the old Ayurvedic medical system. However, today, many individuals are eager to learn more about the herbs used in Ayurvedic therapy. Nadi Priksha is a traditional way of predicting appropriate herbs in which the Rajvedhya studies the pulse, however as technology has advanced, machines are now capable of carrying out such work. In the pharmaceutical sector, it is crucial to identify the proper ayurvedic botanicals used in a medicine's manufacture. The primary characteristics needed to recognize a medicinal plant are found in the shape, color, and texture of its leaves. This study examines morphological traits and feature vectors from the front and back of a green leaf to find the optimal mix of features that maximizes the identification rate. Scanned photos of the front and back of the leaves of frequently used ayurvedic medicinal plants are used to build a database of medicinal plant leaves. Using a special feature combination, the leaves are categorized.

Keywords: Artificial Intelligence, Ayurveda Herbs Classification Technique, Weka Tool.

[1*,2,3] Research scholar, Assoc.Professor, MSEIT Professor, MSIT MATS University, Raipur

Email: [1*]priyanka10sonkar@gmail.com**,** [2]drvijyantverma@matsuniversity.ac.in, [3]v.snehabarde@gmail.co

DOI: 10.31838/ecb/2023.12.s3.629

1. Introduction

Ayurveda is a traditional medical system with roots in the Vedic era, or roughly 6000 years ago. It is still used today in India. Ayurvedic remedies primarily contain plant leaves and other plant parts like roots, bark, and so forth. Over 7000 plants with Indian ancestry have been discovered to have medicinal properties. 2000 of these species, in combination, are employed in the herbal remedies of several Indian medical systems. 600 of these plants specifically are used in commercial Ayurvedic medicines. The majority of plants used in Ayurveda formulations are gathered from wastelands and woods, with the remainder being grown on agricultural land [1]. Without the medicinal plant, ayurvedic treatment is not possible. Today, the person who have not had formal training in the identification of appropriate medical herbs pick the plants from forests. Medicinal plants are frequently swapped or incorrectly delivered to manufacturing units. The majority of these units lack sufficient quality control systems to inspect these plants. Additionally, there is a great deal of uncertainty brought on by regional name variations. Some plants are delivered in dry state, which makes manual plant identification much more challenging. The ineffectiveness of Ayurvedic medicine is caused by improper usage of medicinal plants. Additionally, it could have unforeseen adverse effects. Numerous plants that have ayurvedic properties are all around us, but we are unable to use them without consulting a specialist. As a result, we need technology that goes beyond identification. Due to their shapes, colours, and structures, plants are very difficult to classify in two dimensions [2]. As a result, we have decided to concentrate on the plant's leaf because it has traits unique to plants and can be easily distinguished from other items. Plant leaves, on the other hand, are 2D objects and effectively contain information to identify the plant. It is simple to gather leaves, and image acquisition can be done using low-cost digital cameras, cell phones, or document scanners. Unlike flowers and seeds, it can be found at any time of the year. When leaves grow, they take on a certain color, texture, and shape; these changes are minor. In order to recognize plants based on their leaves, precise descriptors must be found and feature vectors must be extracted from them. Then, a suitable classifier is used to compare the feature vectors of the training and test samples in order to determine how similar they are.

Related Work

There are many researchers who have contributed to this field. Using established structural features, Parag Bhandarkar, Rizwan Ahmed, and colleagues [3] dissected the morphology of leaf edges and 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 orientation of the leaves. The identification rate is not high enough to be useful in real-world applications. For Nadi Pariksha, Roopini et al. [2015] designed a device. It preprocesses the three signals using optical pulse sensors using an eighth order Butterworth filter to reduce noise. The three signals are visualised using three pulse sensors. Then, using artificial neural networks, the pulse data is further classified into Vata, Pitta, and Kapha. Four geometric features—convexity, solidity, eccentricity, and circularity—as well as three RGB colour features—redness, greenness, and blueness indices—were used in the experiment by Pavan Kumar Mishra, Sanjay Kumar Maurya, et al. [7].a database of 1000 photos of leaves that the authors

have assembled. To accelerate the identification process, feature vector comparisons are made in three stages. 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 based on calibrated testing. The topic of the 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., et al., [2020] everything that incorporates or encompasses machine learning is, in essence, machine learning theory. The application of machine learning theory pits real-world applications against the illimitable presumptions that are associated with its abstract concepts. Barde et al, [2022] presented that viola Jones algorithm method for face detection. It is also useful in case of illumination invariant face image [3-5]. Snehlata Barde et al., [2022] described the traditional method for the prediction of suitable herbs where the Rajvedhya examines the pulse is known as Nadi Priksha in the current scenario is changed many technologies is developed, and machines are able to perform such a task. They designed a system for all the people who prefer to use herbal medicine for diseases. That predict the suitable herbs for the disease by the characteristics of Ayurveda herbs by using the classification technique of machine learning. For this, they collected 200 herbs sample and define the suitability of herbs for disease.

2. Materials and Methods

Data Collection: There were standard databases of leaves from Ayurvedic medicinal plants accessible for the experiments. A collection of pictures of medicinal plant leaves was gathered from a personal botanical garden. 16 leaves were randomly selected from 30 different plant species that are used in Ayurvedic, herbal, and traditional medicine. The plants and leav es were chosen at random, and the leaves were taken from their natural habitat. The leaves were placed flat on a regular document scanner, and the highest quality setting was used to scan them. The photographs were saved in jpeg format for future manipulation. The individual leaf photos were then divided into different sets, converted to a constant resolution of 1600 X1200, and organized accordingly shown in figure1.

Fig. 1 Images of leaves

Local names and botanical names of leaves in the database define in table1.

Table 1: Plants name local and Botanical

S. No	Local Name	Botanical Name
1	Aval	Holopteleia intgegrifolia
2	Arya	Vep Azadiracta indica
3	Asokam	Saraca asoka
4	Ayyappana	Eupatorium ayyappana
5	Chittamruth	Tinospora cordifolia
7	Puthari chunda	Solanum indicum
8	Dhanthapala	Wrightia tinctoria
9	Idampirivalampiri	Helicteres isora
10	Ithi	Ficus microcarpa
12	Kanikkonna	Cassia Fistula
13	Karuvappatta	Cinnamomum verum
14	Kattu Tulasi	Ocimum americanum L
15	Kiriyath	Andrographis paniculata
16	Kudampuli	Garcinia gummigutta
17	Koovalam	Aegle marmelos
18	Kumizh	Gmelina arborea
19	Mandaram	Bauhinia acuminata
20	Manchatty	Rubia cordifolia
21	Moovila	Pseudarthira viscida
22	Murikootti	Hemigraphis colorata
23	Nagadanthi	Baliospermum montanum
24	Njezhuk	Pandanus fascicularis
25	Oorila	Desamodium ganjeticum
26	Neermaruthu	Terminalia arjuna

27	Pichakam	Jasminum grandiflorium28	Thanni	Terminalia bellirica
29	Thathiri	Woodfordia fruiticosa		
30	Thazhuthama	Boerhavia di_usa		

Model of the proposed system

There are developed datasets for training, testing, and validation. Additionally, every dataset contains every variety of medicinal leaf. The validation dataset is frequently used to verify that the output is accurate, the test dataset to test the trained model, and the training dataset to train the model. Figure 2 displays a block diagram of the proposed system. The system operates in two phases: the training phase and the testing phase. During the training phase, the system is fed individual leaf database picture inputs. All of the photographs have the same standard resolution thanks to the pre-processing stage. The color image is then changed to grayscale and finally to a binary image. Small flaws like cracks and spots are eliminated from the binary image using a morphological erosion and dilation method. To find morphological features, the binary image's biggest component is chosen.

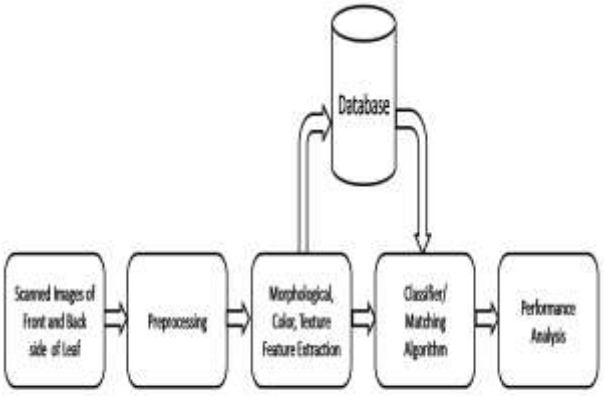

Fig. 2 Block diagram of the proposed model

Classification tools

- Support vector machines feature an algorithm that determines a function from inputted data. This approach is used to categorize the data and do a reversal analysis. Decision planes specify decision boundaries. The objects on these planes are separated into many kinds. On the basis of these planes, SVM operates. The linear classification is being done by them. Duality is the primary SVM characteristic.
- Multilayer Perceptron: A feed-forward artificial neural network is a multilayer perceptron (MLP). It directs the appropriate outputs from one group of input data to another set of outputs. a directed graph with a large number of nodes. This node appears in many levels in an MLP. These layers' unique feature is their interconnectedness. Each node in a layer can be compared to a neuron. Backpropagation is a supervised learning method that is used in the network MLP's training process. Data that cannot be separated linearly can be differentiated via MLP.
- Nearest Neighbours with Generalized Exemplars: A variation of the conventional nearest neighbours algorithm that integrates the idea of generalized exemplars is called Nearest Neighbours with Generalised Exemplars (NNGE). Exemplars are exemplary examples in NNGE that capture the traits

of a collection of related cases. By lowering the number of exemplars while keeping the algorithm's accuracy, the nearest neighbors search should become more effective and efficient.

- The classic nearest neighbors approach determines the distance between a query instance and each instance in the dataset before searching for the k-nearest neighbors of that instance. However, especially for big datasets, this can be computationally expensive. In order to represent groups of related instances, NNGE uses a reduced collection of exemplars.
- Fuzzy Lattice Reasoning: Fuzzy Lattice Reasoning (FLR) is a framework for reasoning that handles complexity and ambiguity in decision-making processes by fusing fuzzy logic and lattice theory. It offers a methodical and structured technique for representing and inferring relationships and imprecise knowledge.
- A subfield of mathematics known as lattice theory deals with partially ordered sets in which the members have a sense of hierarchy or order. Lattice structures, like those found in lattices, make it possible to analyse and manipulate various dependencies and relationships. Lattice theory is used in FLR to model and arrange fuzzy notions and connections between them.
- Naive Bayes: Based on the Bayes theorem and the premise of feature independence, the Naive Bayes algorithm is a straightforward but effective classification method. In particular, text classification, spam filtering, sentiment analysis, and document categorization are some of the machine learning tasks where it is frequently utilized.
- K-Star Instance-Based Classifier: K-Star is an instance-based classifier that fuses the ideas of decision trees and k-nearest neighbors. It is intended to solve categorization issues by making predictions locally using an instance-based method.

3. Result and Analysis

Several machine learning models, including Deep Convolutional networks, will be used to present our findings. Although ConvNet (CNN) is known to gain features when trained on larger datasets, the results obtained by training them with only original pictures will not be explored. table 2 described the comparison between the classifiers in terms of Accuracy, true positive value, false positive value, precision, recall, and F-measure.

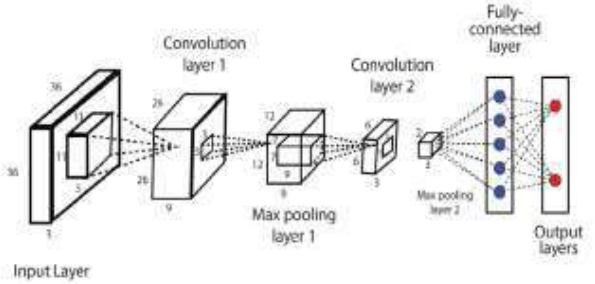

Fig. 1. Convolutional Neural Network Architecture

After the 25th training iteration, extremely accurate results were obtained with a reasonably low loss, and after the 51st iteration, a high degree of precision was used to balance inaccuracy and loss.The model was independently assessed and confirmed in each class once it had been trained. Table 2: Comparison of classifiers in the identification of green leaves using

Geometric-Colour-Texture-Zernike combination.

Classifier	Accuracy (%)	TP	FP	Precision	Recall	FMeasure
Multi Layer Perceptron	99	0.99	0	0.99	0.99	0.99
Support Vector Machine	98.7	0.98	0	0.988	0.98	0.98
Nearest Neighbor Generalized Exemplar	97.7	0.97	0.001	0.979	0.97	0.97
Fuzzy Lattice Reasoning	98.3	0.98	0	0.985	0.98	0.98
Naïve Bayes Classifier	98.1	0.98	0	0.982	0.98	0.98
K-Star Instance Based Classifier	98.3	0.98	0	0.984	0.98	0.98

Figure 4 shows the graphical representation of six classifiers multilayer perceptron gives the highest accuracy 99% and Nearest Neighbor Generalized Exemplar gives minimum accuracy 97.7%. Fig. 4 Accuracy (%) of six classifiers.

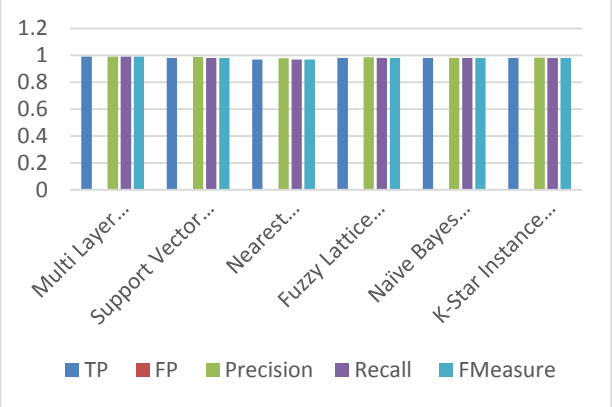

Fig.5. Graph for TP, FP, Precision, Recall, and F-Measures for six classifiers

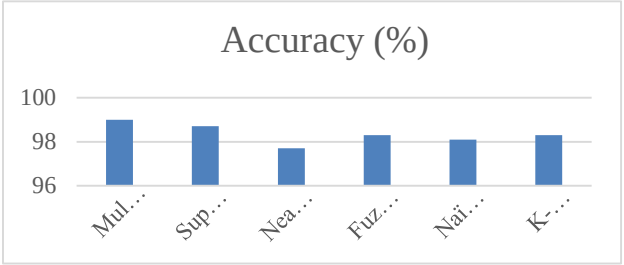

Figure 5 shows the comparison graph in terms of TP, FP, Precision, Recall, and F-Measures for the six classifiers and found that multilayer perceptron gives the highest values and Nearest Neighbor Generalized Exemplar gives minimum values.

4. Conclusion

In this study, machine learning techniques were used to evaluate a unique approach to categorizing the leaves of Ayurvedic medicinal plants. Although the final recognition rate can range from 97 to 99 percent depending on the outcomes, it is 99 percent. Although most people in our digital age have access to the internet, many do not have access to quality medical care or medications. With the help of this CNN (convolutional neural network) application, people can learn more about local Ayurvedic plants and how to use them more successfully.

Future scope

The launch of this research as a product is one of its future objectives, and we have already combined the CNN model with a web application that enables users to click on any plant photo. leaf and the app will educate users of the plant's qualities and usefulness as well as whether it is an ayurvedic plant or not. In order to continue this research and help someone who doesn't actually have access to healthcare, the researcher wanted to spread knowledge about the consequences of employing Ayurvedic botanicals.

5. References

1. 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, pp. 1-4, doi: 10.1109/ASIANCON55314.2022.9908981.
2. 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. https://doi.org/10.1007/978-981-16-2934-1_16
3. 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 Communication Networks (ICRCICN), 2017
4. Sinha, A., & Barde, S. (2022, October). Illumination invariant face recognition using MSVM. In AIP Conference Proceedings (Vol. 2455, No. 1, p. 040005). AIP Publishing LLC.. https://doi.org/10.1063/5.0100936.
5. Sinha, A., & Barde, S. (2022). Age Invariant Face Recognition Using PCA And MSVM. Journal of Pharmaceutical Negative Results, 2174-2185. 13. https://doi.org/10.47750/pnr.2022.13.S10.253.
6. Sinha, A., & Barde, S. (2022). Face recognition across age progression by using PCA. International Journal of Food and Nutritional Sciences, 11(S3), 4608-4617. ttps://doi.org/10.48047/IJFANS/S3/144
7. Sinha, A., & Barde, S. (2023). Multi Invariant Face Detection via Viola Jones Algorithm. Eur. Chem. Bull., 12(Special Issue 1), 24-32. doi: 10.31838/ecb/2023.12.s1.003
8. Barde, Snehlata. "Multimodal Biometrics: Most Appropriate for Person Identification." i-manager's Journal on Pattern Recognition 4.3 2017.
9. Ghatge, M.R. and Barde, S., 2022. Comparison of CNN-LSTM in

Sentiment Analysis for Hindi Mix Language. In: , 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS) (pp. 453-456).IEEE. https://doi.org/10.1109/ICTACS56
10. 270.2022.9988531
11. Ghatge, M.R., Barde, S. (2023). The Sentiment Analysis for Hindi Language Using Convolution Neural Network and Deep Belief Network. In:, et al. Information Systems and Management Science. ISMS 2021. Lecture Notes in Networks and Systems, vol 521. Springer, Cham. https://doi.org/10.1007/978-3-031-13150-9_1.