Ayurvedic Herb Identification and Classification

Harsh Bansal
Department of Computational Intelligence
School of Computing
Faculty of Engineering and Technology
S.R.M Institute of Science and Technology
Chennai, Tamil Nadu
hb4634@srmist.edu.in

Stuti Gambhir
Department of Computational Intelligence
School of Computing
Faculty of Engineering and Technology
S.R.M Institute of Science and Technology
Chennai, Tamil Nadu
sg0098@srmist.edu.in

Khushi Bansal
Department of Computational Intelligence
School of Computing
Faculty of Engineering and Technology
S.R.M Institute of Science and Technology
Chennai, Tamil Nadu
kb8593@srmist.edu.in

A. Alice Nithya*
Department of Computational Intelligence
School of Computing
Faculty of Engineering and Technology
S.R.M Institute of Science and Technology
Chennai, Tamil Nadu
alicenia@srmist.edu.in

***Abstract*—Identification and classification of Ayurvedic plants is an important role in the development of traditional medicine and ensures the use of medicinal plants. This project uses deep learning and image processing technology to identify plants such as neem, tulsi, aloe vera, and nagfani. Using the fine-tuned VGG16 model and custom data, the system achieves high accuracy i.e. 92% in classifying plant images. This involves several pre- processing steps including resizing, normalization, and enhancement that make it robust against the changes made in the input image. Finally it is applied using the Flask web application, allowing the identification of herbs at real time by giving a simple user interface.**

Keywords*— *Image processing, herbal classification, deep learning, VGG16.

I. Introduction

Ayurvedic medicine, based on ancient traditions, is heavily dependent on proper identification of medicinal plants and is the backbone of health care and natural remedies. Error-prone manual identification methods lead to inconsistencies in plant classification [1]. Recent advances in artificial intelligence and deep learning address these issues. This paper presents an AI-driven approach for Ayurvedic herb identification and classification using convolutional neural networks (CNNs). The proposed system leverages image processing and pre-trained models to classify plants such as Aloe Vera, Neem, Tulsi, and Nagfani, ensuring high accuracy and efficiency. This study integrates traditional knowledge with modern technology to bridge the gap between ancient practices and contemporary healthcare solutions.

II. Literature Review

Recent developments in the field of artificial intelligence, particularly deep learning, have addressed the issues of morphological similarity and environment diversity in plant species identification. Recent work by Jeon and Rhee [1] in 2017 proposed a CNN for recognizing plant leaves with high accuracy; however, it was not quite scalable to diverse datasets. Similarly, Kumar et al. [2] applied the techniques of image processing to classify Ayurvedic medicinal plants relying on handcrafted features that limited the adaptability of the approach to unseen data.

In 2020, Md Zin et al. [3] employed deep CNNs for herbal plant recognition, and they had notable success on smaller datasets but struggled with real-world noise. Nguyen et al. [4] explored CNN-based medicinal plant identification in natural settings, but the lack of data augmentation resulted in performance degradation under varying lighting conditions. Hassan and Maji [5] highlighted the possibility of deep learning for plant classification, but they underlined the need for bigger datasets and to ensure robust generalization. Muneer and Fati [6] further combined shape and texture features into a CNN-based classification model with improved accuracy and optimized computational efficiency.
In 2021, Akter and Hosen [7] applied CNNs to classify Bangladeshi medicinal plants. Though they showed promising results, it is challenging to extend to varied flora. As of 2022, Pushpa and Rani [8] proposed a lightweight CNN for Indian Ayurvedic plant classification called Ayur-PlantNet. It was designed for a resource-constrained environment. Sonia et al. [9] focused on real world applications, integrating their model for practical use, whereas Vinod et al. [10] further extended this by developing a mobile application for real-time herb classification based on machine vision. Azadnia et al. [11] presented in 2022 a GAP-based CNN model that enhanced feature extraction with reduced model complexity. Similarly, R.

S and V.M.N [12] developed a hybrid machine learning and deep learning framework for Ayurvedic plant identification, which balances accuracy with computational requirements. In 2024, Jaffino and Jose [13] proposed hardware-optimized solutions using KNN classifiers with FPGA for Ayurvedic plant classification. Moreover, Vinod et al. [14] used mobile technology for real-time herb classification using machine vision, making it more accessible. Finally, V. H. A et al.[15] applied transfer learning techniques in deep learning to enhance domain-specific classification of medicinal plants.

III. EXPERIMENTAL METHODOLOGY

A. Project Description

The Ayurvedic Herb Identification and Classification system has a unique feature. It includes deep learning techniques for classifying herbs including Neem, Tulsi, Aloe Vera, and Nagfani which includes the use of ayurveda. Such systems use modified VGG16 models as their basis and makes use of transfer learning - a method that reduces the time needed for training while keeping high classification levels.

The project helps resolve issues with the manual process of identifying herbs by making it automated thus eliminating the chances of mistakes. It provides a simple to use interface by making use of a sophisticated flask based web app that allows users to upload pictures and classify herbs in real time. This system helps in practicing most of the Ayurvedic knowledge and at the same time assists with studies of diversity and ancient medicine.

B. Data Flow

To ensure users can make easy use of the system and obtain accurate predictions in a smooth processing manner, the layout of the system is structured in a number of steps. Distribution of tasks may be done for better predictability of the model.

1. Data Collection:

Pictures of herbs are compiled within a database from public archives, field-captured images and herb-specific data sets with variety in the size, shooting position, incidence angle, light, and background. These Pictures are organized into systematized directories that connect to the herb classification.

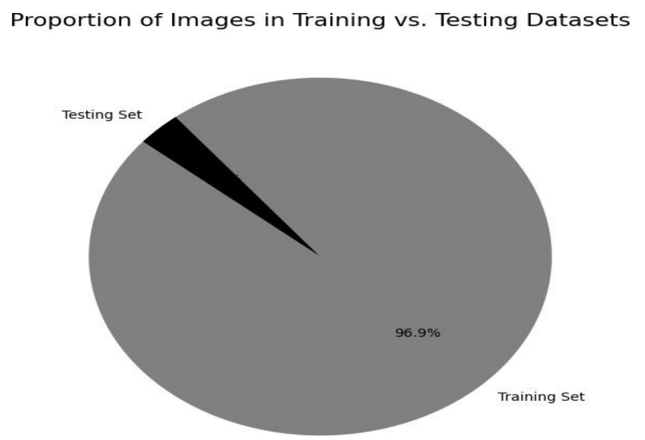

Fig.1. Training vs Testing Dataset

2. Data Preprocessing:

Determined image data is standardized in order to improve model robustness as well as ensuring uniformity. Resizing: All the images are resized to 150 by 150 pixels. This was done so that the dimensions corresponding to the input of the model VGG16 would be satisfied. Normalization: Rescaling pixel values between 0 and 1 such as data representation input is consistent.

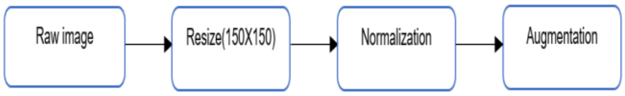

Fig. 2. Preprocessing Pipeline Visualization

Augmentation: Techniques such as Rotation, Flipping, Zooming, and changing brightness were utilized on data to increase the generalizability of the model as well as the variability of the dataset.

3. Feature Extraction:

The 'pre-processed images' are given input into the VGG16 model, the model extracts specific features of the images, these include the leaf geometric shape, vein patterns on leaves as well as the color texture of the leaves. The last layers of the models are altered to make certain that feature learning that is task specific takes place.

4. Classification:

The next layer of the Deep Neural Net is the classification layer where the features that were extracted earlier are fed into and in this layer, the softmax activation is used. From this layer, probabilities assigned to each unknown herb category is the output and class with the maximum probability is the herb that will be predicted. Whenever the confidence score is said for instance lower than a specific threshold the model considers the prediction as having a lot of uncertainty and hence should not be trusted.

Fig.3. System Architecture Diagram

5. Results and User Interface Visualization:

The anticipated output is available on a web page application developed using Flask. The interface enables users to submit images, and then the predictions and the corresponding confidence levels are displayed, together with the image that was uploaded for reference. Designing the system in this way makes it user-friendly as it gives instant feedback and is easily accessible.

By completing these steps the system is able to keep an optimal workflow from data stream into classification and interaction. Such orderly procedure combines traditional approaches toward herbal identification and new technologies that make the system sound and fit for practical application.

IV. Result and Analysis

The Ayurvedic Herb Classification and Identification system worked well because it achieved high accuracy in recognizing herbs like Neem, Tulsi, Aloe Vera, and Nagfani.

The fine-tuned VGG16 model, trained on an organized dataset, exhibited an overall accuracy of 92 percent. Among the four categories, Aloe Vera obtained the highest classification accuracy of 95 percent, while Nagfani achieved an accuracy of 94 percent. However, the system found it difficult to classify between Neem and Tulsi with accuracies of 89 percent and 90 percent, respectively, as these two are morphologically similar. All these results support the applications of transfer learning and fine-tuning in using pre-trained models for specific tasks such as herb classification.

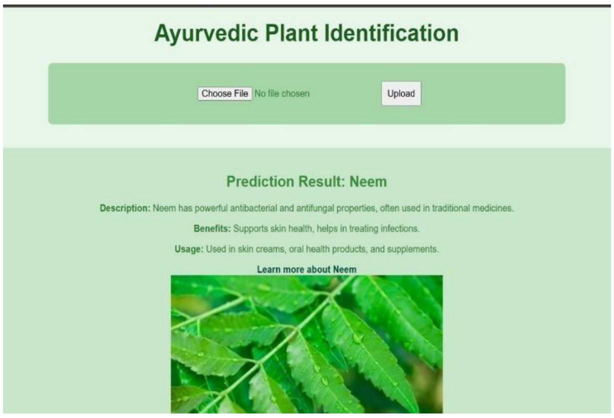

Fig. 4. Neem Identification – User Interface

This figure shows how an image is classified as neem and shows its benefits and usage for the user and also provides user a direct link to Wikipedia to gain more knowledge.

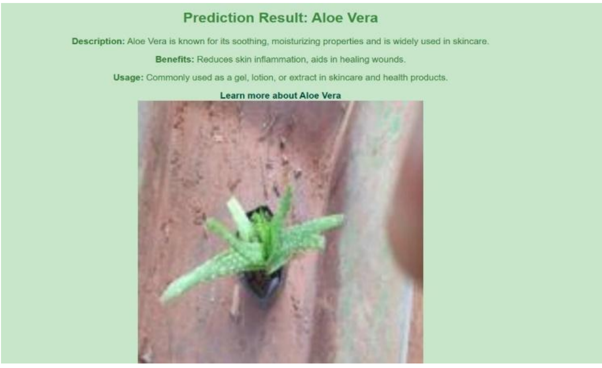

Fig.5. Aloe Vera Identification – User Interface

This figure shows how an image is classified as aloevera and shows its benefits and usage for the user and also provides user a direct link to Wikipedia to gain more knowledge.

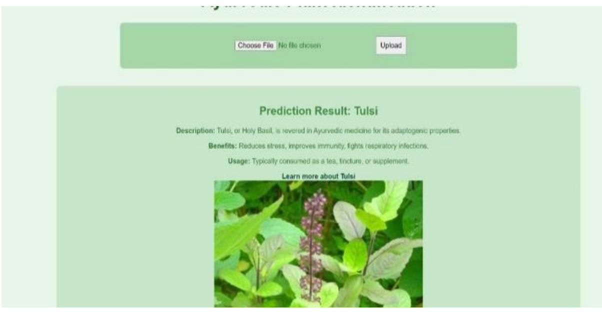

Fig.6. Tulsi Identification – User Interface

This figure shows the benefits ,usage about tulsi when user uploads an image and also gives the description about the plant.

Fig.7. Nagfani Identification – User Interface

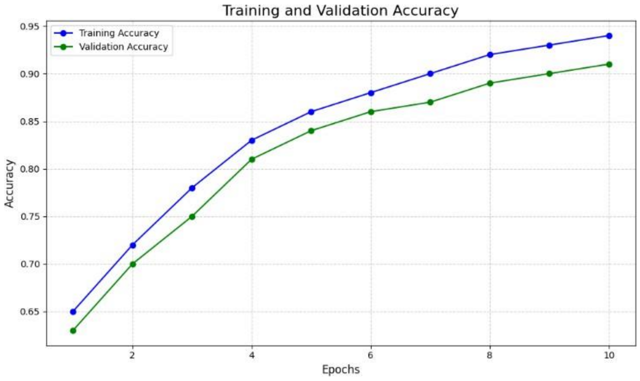

Fig.8.Training and Validation Accuracy

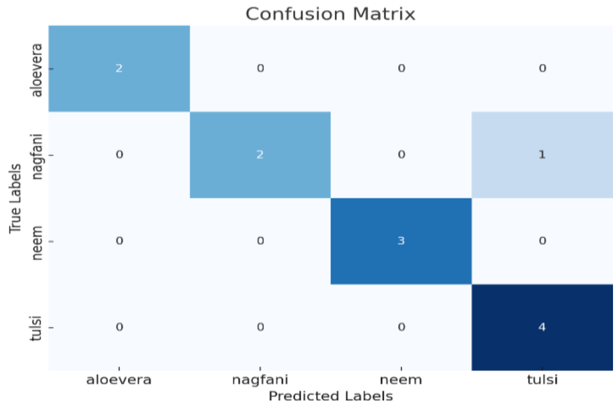

Fig.9.Confusion Matrix

According to the confusion matrix, an analysis was conducted, revealing that there are very few errors of misclassification. Thus, in this case, it proves that the model is learning meaningful patterns from the dataset and predicting in a reliable manner.

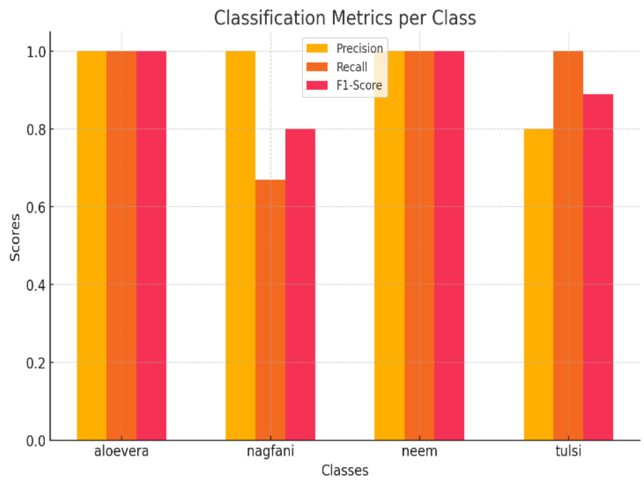

Fig.10.Classification Metrics per Class

Class-wise evaluation in terms of precision, recall, and F1-score further supports the robustness of the model. The classification metrics per class indicate that Aloe Vera and Nagfani have higher precision and recall, meaning that this model can identify these classes accurately with fewer errors. On the other hand, Tulsi and Neem performed well but had overlap issues in predictions, as indicated in the confusion matrix, which again emphasizes the need for optimization in the future work.

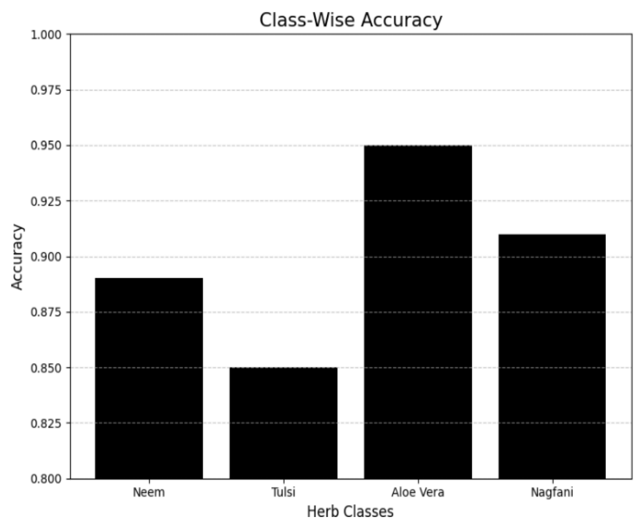

Fig.11. Class -Wise Accuracy

V. Conclusion and Future Work

The Ayurvedic Herb Classification and Identification system successfully integrates deep learning techniques with traditional knowledge to achieve high classification accuracy for medicinal herbs such as Neem, Tulsi, Aloe Vera, and Nagfani. The fine-tuned VGG16 model demonstrated its effectiveness in learning meaningful patterns, with Aloe Vera and Nagfani achieving outstanding accuracy levels. The deployment of the model in a web-based application allows real-time predictions, making it highly beneficial for healthcare practitioners, researchers, and Ayurveda enthusiasts. While the system has proven to be effective, challenges such as misclassification of morphologically similar plants like Neem and Tulsi indicate the need for a more diversified dataset and advanced feature extraction techniques. The confusion matrix analysis shows minimal misclassification, confirming the model's reliability.

References

[1] Jeon, Wang-Su Rhee, Sang-Yong. (2017). Plant Leaf Recognition Using a Convolution Neural Network. The International Journal of Fuzzy Logic and Intelligent Systems. 17.26-34. 10.5391/IJFIS.2017.17.1.26.

[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, India, 2017, pp. 231-238, doi:10.1109/ICRCICN.2017.8234512.

[3] I. A. Md Zin, Z. Ibrahim, D. Isa, N. Sabri, S. Aliman and N. N. Mangshor, ”Herbal plant recognition using a deep convolutional neural network,” Proceeding of Electrical Engineering and Informatics, vol. 9, no. 5, pp. 2198-2205, 2020.

[4] Nguyen Quoc, Trung Vn, Truong Hoang, Vinh. (2020).Medicinal Plant identification in the wild by using CNN.

[5] Hassan, Sk Maji, Arnab. (2020). Identification of Plant Species Using Deep Learning. 10.1007/978-981-15-7834-211.

[6] A. Muneer and S. M. Fati, ”Efficient and Automated Herbs Classification Approach Based on Shape and Texture Features using Deep Learning,” in IEEE Access, vol. 8, pp.196747-196764, 2020, doi: 10.1109/ACCESS.2020.3034033.

[7] Akter, Raisa Hosen, Md. (2021). CNN-based Leaf Image Classification for Bangladeshi Medicinal Plant Recognition. 10.1109/ETCCE51779.2020.9350900.

[8] B. R. Pushpa and N. S. Rani, ”Ayur-PlantNet: An unbiased lightweight deep convolutional neural network for Indian Ayurvedic plant species classification,” J. Appl. Res.Med. Aromat. Plants, vol. 34, pp. 100459, 2023.

[9] S. V. E. Sonia, D. N and D. R. S, ”Medicinal Plants Classification by Visual Characteristics of Leaves Using CNN,” 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 2023, pp. 01-05, doi: 10.1109/ICEEICT56924.2023.10157410.

[10] Marada S. Rao, S. Praveen Kumar and Konda S. Rao,” A Methodology for Identification of Ayurvedic Plant Based on Machine Learning Algorithms,” Int. J. of Computing and Digital Syst., vol. 14, no. 1, pp. 10233-10241, 2023.

[11] Azadnia, Rahim Al-Amidi, Mohammed Mohammadi, Hamed Cifci, Mehmet Akif Daryab, Avat Cavallo, Eugenio.(2022). An AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average PooAgronomy.12. 10.3390/agronomy12112723.

[12] R. S and V. M.N, ”Ayurvedic Plants Identification based on Machine Learning and Deep Learning Technologies,” 2022 Fourth InternationalConference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), Mandya, India, 2022, pp. 1-6, doi:10.1109/ICERECT56837.2022.10060533.

[13]G. Jaffino and J. Prabin Jose, ”Classification of Ayurvedic Medicinal Plant Leaf using KNN Classifier and FPGA,” 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2024, pp. 2255-2259, doi: 10.1109/ICACCS60874.2024.10717192.

[14]D. Vinod et al., ”An Android App for Ayurvedic Herb Identification and Classification Using Machine Vision,”2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST), Kochi, India, 2024, pp.01-06, doi: 10.1109/ICTEST60614.2024.10576106.

[15]V. H. A, N. M. M. S and T. D. K, ”Leveraging Deep Learning for Identification of Medicinal Plant Species,”2024 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 2024, pp. 1-9, doi:10.1109/ICDSNS62112.2024.10690945.