MEDICATIVE PLANT CLASSIFICATION USING MACHINE LEARNING

Vaishnavi Mandwekar
M.Tech, CSE, PRMITR, Amravati

G. R. Bamnote
Department of Computer Science & Engg.
PRMITR, Badnera Amravati

S.M. Iqbal
Department of CSE
PRMITR, Badnera Amravati

ABSTRACT— In an age driven by artificial intelligence and biotechnology, the convergence of deep learning and ethnobotany offers novel avenues for environmental and medical innovation. This paper presents the development of a Medicinal Plant Forensics Tool leveraging computer vision and deep learning to identify medicinal plants from images and provide comprehensive medical insights. Utilizing the ResNet50 architecture, fine-tuned on a dataset of medicinal plants, the system enables accurate, real-time identification with confidence indicators. Integrated with Google's Gemini AI, the tool retrieves structured medicinal information including active compounds, traditional uses, modern applications, and safety guidelines. The system further provides robust evaluation metrics such as confusion matrices, class-wise accuracy charts, and classification reports, ensuring performance transparency. Through intuitive GUI, users can interact seamlessly with the model for both prediction and exploration. This paper stands as a promising step toward preserving indigenous knowledge while fostering smart, data-driven healthcare accessibility.

KEYWORDS- Deep Learning, Computer Vision, Medicinal Plants, ResNet50, Gemini AI, Image Classification, GUI, Model Evaluation, Python, Plant Identification

I. INTRODUCTION

Ayurveda, an ancient system of medicine originating in India around 5000 years ago, holds its roots in the Vedic times. Central to Ayurvedic treatments are medicinal plants, encompassing various components such as leaves, roots, and bark. Among the vast array of Indian flora, more than 8000 plants have been recognized for their medicinal value, with approximately 1500 forming a subset commonly employed in herbal medicines across different Indian systems. Notably, commercial Ayurvedic formulations selectively integrate 500 of these plants. Traditionally, Ayurvedic physicians personally gathered medicinal plants from forests and wastelands, preparing medicines for their patients. However, with evolving environmental conditions, humanity faces new challenges in the form of life-threatening viruses and illnesses. The recent global pandemic, exemplified by COVID-19, underscores the urgency of developing immediate solutions to combat emerging health threats.

In the pharmaceutical and medical field, addressing novel viruses and illnesses poses significant challenges. The need for swift and effective solutions has never been more critical. Medicinal plants emerge as a valuable resource, offering new properties that pharmacists and medical practitioners can harness in the quest for defenses against life-threatening viruses and illnesses. This paper aims to bridge traditional Ayurvedic knowledge with contemporary challenges in medicine. Through the implementation of advanced techniques, particularly focusing on the utilization of Convolutional Neural Networks (CNN) and various features like shape, size, color, and texture, this paper seeks to automate the identification process of medicinal plants.

II. LITERATURE REVIEW

1. A. Gopal et. al Implement a system using image processing with images of the plant leaves as a basis of classification. The software provides the most similar match in response to the input query. The proposed algorithm was implemented and evaluated for efficiency by testing it on 10 different plant species. The software is trained with 100 (10 numbers of each plant species) leaves and tested with 50 (tested with different plant species) leaves. The implementation of the proposed algorithm achieved an efficiency of 92%.

2. Umme Habiba et. al In this paper, for automatically classifying medicinal plants, they present a Multichannel Modified Local Gradient Pattern (MCMLGP), a new texture-based feature descriptor that uses different channels of color images for extracting more significant features to improve the performance of classification. Authors have trained their proposed approach using SVM classifier with various kernels such as linear, polynomial and HI. In addition, we used different feature descriptors for comparative experimental analysis with MCMLGP by conducting rigorous experiments on our own medicinal plants dataset. The proposed approach gains higher accuracy (96.11%) than other techniques and is significantly valuable for exploration and evolution of medicinal plants classification.

3. R. Janani et. al Have proposed a method for the extraction of shape, color and texture features from leaf

images and training an artificial neural network (ANN) classifier to identify the exact leaf class. The key issue lies in the selection of proper image input features to attend to high efficiency with less computational complexity. They tested the accuracy of the network with different combinations of input features. The test result on 63 leaf images reveals that this method gives 94.4% accuracy with a minimum of 8 input features. This approach is more prominent for leaf identification systems that have minimum input and demand less computational time.

4. Vijayashree. T et. al Has created a database with 127 herbal leaves. For creating a database 11 texture parameters are considered. The parameters considered include Sum of Variance, Inverse Difference Moment, Aspect Ratio, Correlation, Sum Entropy, Mean, and Sum Average. To extract features such as entropy, homogeneity, contrast, and energy, the Gray Level Co-occurrence Matrix (GLCM) method is employed. A test image is analyzed and compared against the database by calculating dissimilarity based on these extracted features. The leaf corresponding to the minimum dissimilarity score is identified, and the result is presented as the output.

5. Venkataraman et. al A system is developed which would provide a solution for identifying the plant and providing its medicinal values, thereby helping in the cure of many ailments in a natural way. This paper discusses the dataset collection, feature extraction using texture and HOG and thereby classifying based on the Support Vector Machine algorithm.

6. Shitala Prasad et. This paper presents a new and efficient technique for leaf acquisition. The image is transformed into a device independent lαβ color space that is further used to compute VGG-16 feature map. The feature map is projected onto a PCA subspace to enhance species recognition performance. To validate the method's robustness, two distinct plant leaf datasets are utilized in the evaluation.

7. Dileep M.R et. al This work proposes Ayur Leaf, a Deep Learning based Convolutional Neural Network (CNN) model, to classify medicinal plants using leaf features such as shape, size, color, texture etc. This study also introduces a standardized dataset of medicinal plants commonly found across different regions of Kerala, a state located on India's southwestern coast. The dataset includes leaf samples from 40 distinct medicinal plant species. A deep neural network inspired by Alex net is utilized for the efficient feature extraction from the dataset. Finally, classification is carried out using both SoftMax and SVM classifiers. Our model, upon 5-cross validation, achieved a classification accuracy of 96.76% on Ayur Leaf dataset. Ayur Leaf helps us to preserve the traditional medicinal knowledge carried out by our ancestors and provides an easy way to identify and classify medicinal plants.

8. C. Amudha Lingeswaran et.al Had built a model (Deep Neural Networks) for the identification of medicinal plants. To train the model, the author used around 8,000 images belonging to four different classes. Finally, it arrived with a good accuracy of 85% when testing with images taken from the open field land areas.

9. Manojkumar P. et. al This paper explores feature vectors from both the front and back side of a green leaf along with morphological features to arrive at a unique optimum combination of features that maximizes the identification rate. A database of medicinal plant leaves is created from scanned images of the front and back side of leaves of commonly used ayurvedic medicinal plants. Leaves are classified using distinctive combinations of features, achieving identification rates as high as 99% across a broad range of classifiers. This approach has been further extended to include dry leaf identification, where a combined set of feature vectors yielded identification accuracies exceeding 94%.

10. Amala Sabu et. al The proposed system uses a combination of SURF and HOG features extracted from leaf images and a classification using k-NN classifier. Our experiments demonstrate results that are promising enough to support the development of real-world applications.

III. METHODOLOGY

The system, *Medicinal Leaf*, presents an advanced Convolutional Neural Network (CNN)-based solution for the accurate classification of medicinal plants through deep learning. Leveraging the powerful ResNet-50 architecture, the system processes user-uploaded leaf images via a Flask-based web interface, where the images undergo preprocessing steps such as resizing and normalizing before prediction. By extracting key features like shape, size, color, and texture, the model accurately identifies the plant species from among 40 commonly found medicinal plants in India. The system is backed by a standardized and diverse dataset, enhancing model robustness and generalizability. Key components include a trained model file (best_model.pth), a GUI interface (predict_gui.py), an evaluation script (evaluate_model.py), and class-label mappings (class_mapping.txt). Altogether, this unified framework aims to modernize medicinal plant identification, offering a user-friendly, reliable, and automated tool to support research, conservation, and practical applications in herbal medicine.

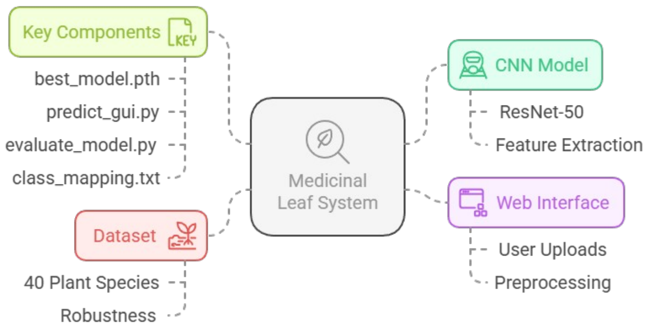

Fig 1 shows the block diagram

FLOWCHART

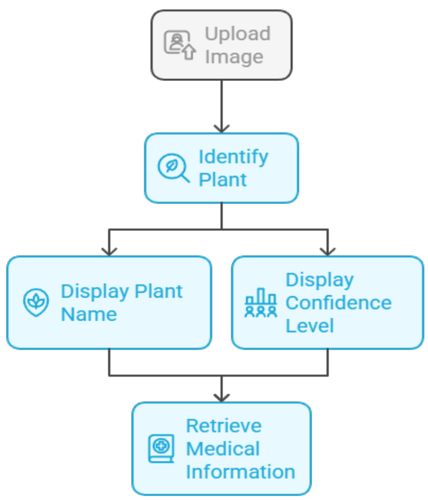

Fig 2 shows the flowchart of system

This flowchart illustrates the working of a medicinal plant identification system using image input:

1. Upload Image: The user uploads a plant image.
2. Identify Plant: The system uses a machine learning model to recognize the plant species from the image.
3. Display Plant Name: The identified plant's name is shown.
4. Display Confidence Level: The system also shows how confident it is about the identification (e.g., 95% confidence).
5. Retrieve Medical Information: Finally, it fetches and displays the medicinal uses or properties of the identified plant.

Convolutional Neural Network

A Convolutional Neural Network (CNN) is a deep learning architecture tailored for image processing and recognition. Unlike other classification methods, CNNs minimize the need for preprocessing by automatically learning hierarchical features directly from raw image data.

CNN ARCHITECTURE

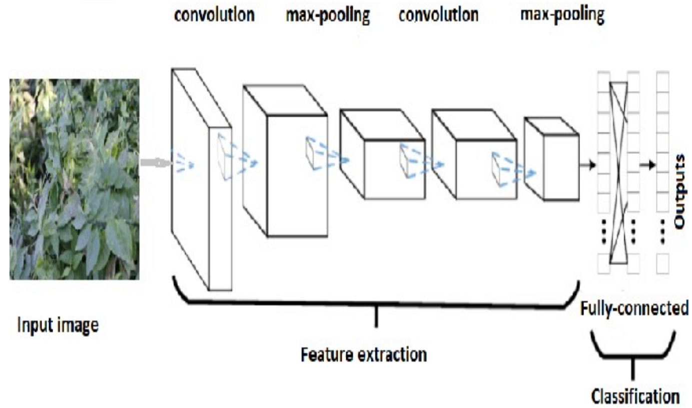

Fig 3 shows the CNN architecture

WORKING

The CNN architecture operates through an end-to-end pipeline designed for the automatic identification of medicinal plants using leaf images and deep learning. The process begins when a user uploads a leaf image via a Flask-based web interface (handled by predict_gui.py). This image is first preprocessed including steps like resizing to fit model input requirements and normalization to standardize pixel values ensuring it's compatible with the trained neural network.

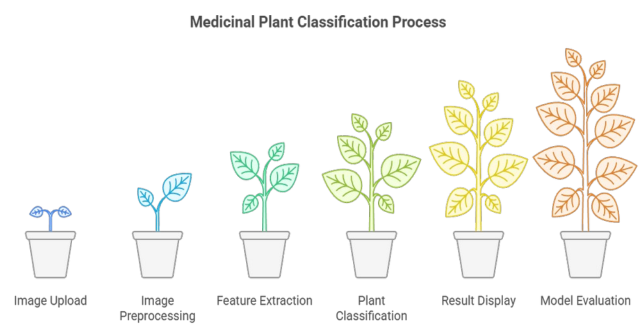

Fig 4 shows the workflow

The preprocessed image is then passed into a pretrained Convolutional Neural Network model specifically, a ResNet-50 architecture, saved in the best_model.pth file. The model extracts key features such as the leaf's shape, size, color, and texture to analyze and classify the plant. Based on the internal learned representations, the model outputs a prediction corresponding to one of 40 medicinal plant species commonly found across India.

The output prediction is then mapped to its corresponding plant name using class_mapping.txt, and the result is displayed back to the user on the GUI. wishing to analyze the model's performance, the evaluate_model.py script provides a detailed evaluation using metrics like accuracy, precision, recall, and confusion matrix visualization (using libraries such as matplotlib, seaborn, and scikit-learn).

This system is powered by deep learning libraries like PyTorch and torch vision, and relies on custom tkinter, PIL, and other supporting Python libraries for visualization and backend logic. Additionally, google-generativeai is optionally included for future enhancements using Gemini AI. In summary, the paper offers a robust, automated, and user-friendly framework for medicinal plant classification, supporting use cases in herbal research, botanical education, and conservation efforts.

IV. SYSTEM REQUIREMENT

SOFTWARE REQUIREMENT

- Frontend- HTML, CSS, JavaScript
- Backend- Python
- API-Gemini API
- IDE-Visual Studio Code (VS Code)
- PyTorch 2.0+, Torchvision, custom tkinter, PIL (Pillow), matplotlib, seaborn, scikit-learn, pandas, torch.

V. RESULT & PERFORMANCE ANALYSIS

Plant Image Analysis Output

The submitted plant images are processed using the "Medicinal Plant Tool," which utilizes computer vision and machine learning algorithms to analyze plant morphology and identify the species.

- Uploaded Image: A small potted plant placed on soil.
- System Confidence: The tool returned to a 99.7% confidence level, indicating high certainty in its classification.
- Identified Plant: Ashwagandha (*With Ania somnifera*).

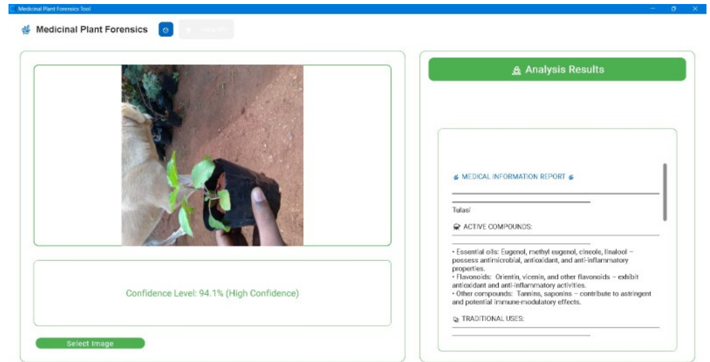

Fig 5 shows the result output of system

PERFORMANCE ANALYSIS

The Medicinal Leaf Classification paper evaluates a ResNet-50 model's ability to identify 40 medicinal plant species from leaf images. Using standard metrics (precision, recall, F1-score), the evaluation assessed performance across 1,467 samples. Implemented with PyTorch and torchvision, the model's accuracy metrics provide essential validation for real-world application and establish benchmarks for future improvements.

This confusion matrix visually presents the performance of a plant classification model in identifying 10 medicinal plants:
Bamboo, Amla, Guava, Ekka, Doddapatre, Basale, Avocado, Ashwagandha, Mint, and Tulasi.

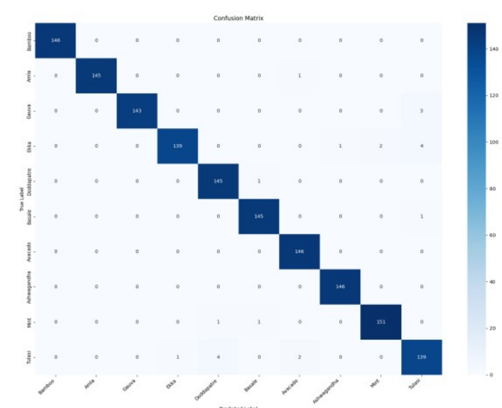

Fig 6 shows the confusion matrix

The model demonstrated exceptional classification capabilities, achieving:

- Overall Accuracy: 98.50% (0.9850034)

Fig 7 shows accuracy by plant type: Bamboo, Avocado, Ashwagandha (100%); Amla, Doddapatre, Basale (99%); Mint, Guava (98%); Ekka, Tulasi (95%). High overall performance observed, with minor improvements needed for lowest performers

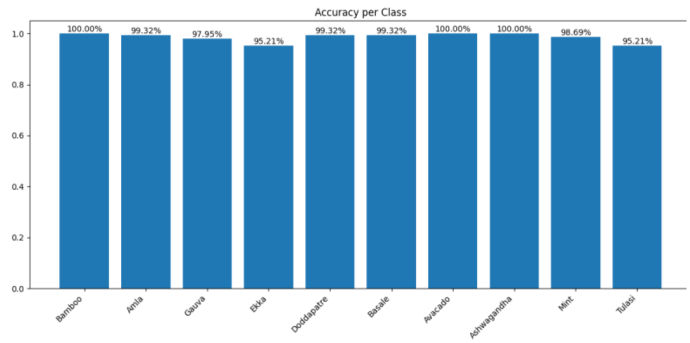

Fig 7 shows the Accuracy per class graph

Classification Metrics per Class – Summary

The heatmap titled *"Classification Metrics per Class"* showcases the model's precision, recall, and F1-score for identifying various plant leaves. Most classes like Bamboo, Amla, Ashwagandha, and Avocado achieve near-perfect scores (≈1.000), indicating excellent performance

Fig 8 shows the Classification Metrics Per Class

Basale and Mint also show strong metrics (~0.99), highlighting overall robustness. However, Tulasi (precision: 0.946, F1: 0.949) and Ekka (recall: 0.952) underperform slightly, suggesting occasional misclassifications. These weaker spots point to a need for improved data augmentation or feature extraction. Overall, the model is highly accurate, with room for fine-tuning in select classes.

This indicates that the model correctly classified 98.50% of all leaf images in the test dataset, reflecting its high reliability in identifying medicinal plant species.

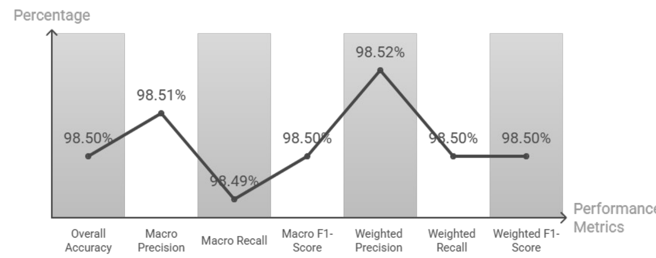

Fig 9 shows Performance Metrics of Medicinal Plant Classification Model

Additional aggregate metrics provide further insight into the model's performance:

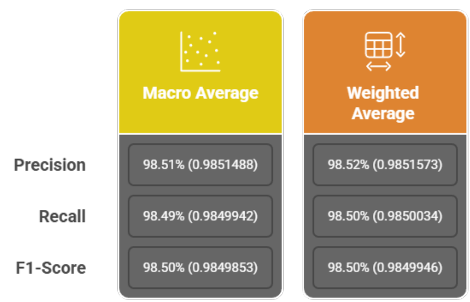

Fig 10 shows the aggregate metrics

The minimal difference between macro and weighted averages suggests consistent performance across all classes, regardless of sample size.

VI. CONCLUSION

The performance evaluation demonstrates that the Medicinal Leaf Classification paper has achieved exceptional accuracy in identifying medicinal plant species through leaf image analysis. The ResNet-50 architecture, combined with appropriate preprocessing techniques, has proven highly effective for this specialized classification task.

The remarkable consistency across different performance metrics with an overall accuracy of 98.50% and most species achieving precision and recall rates above 95% establishes the system as a reliable tool for medicinal plant identification. These results validate the technical approach and underscore the potential of deep learning for specialized botanical classification tasks.

The successful implementation of the designed model opens avenues for various applications in research, education, conservation, and medicine, demonstrating how artificial intelligence can contribute meaningfully to the field of botany and herbal medicine.

VII. REFERENCES

[1] A.Gopal, S. Prudhveeswar Reddy and V. Gayatri "Classification of Selected Medicinal Plants Leaf Using Image Processing" Published in 2021 International Conference on Machine Vision and Image Processing (MVIP)

[2] Umme Habiba, Md. Rasel Howlader, Md. Aminul Islam, Rahat Hossain Faisal, and Md. Mostafijur Rahman presented a method for automatic classification of medicinal plants using a multi-channel modified local gradient pattern combined with an SVM classifier, as detailed in the proceedings of the 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 3rd International Conference on Imaging, Vision & Pattern Recognition (IVPR).

[3] R.Janani, A.Gopal "Identification of selected medicinal plant leaves using image feature and ANN" published in 2019 international conference on advanced electronic systems (ICAES K. Elissa, "Title of paper if known," unpublished.

[4] Vijayashree.T, Dr. A. Gopal "Leaf Identification for the Extraction of Medicinal Qualities Using Image Processing Algorithm" published in 2020 International Conference on Intelligent Computing and Control (I2C2)

[5] D Venkataraman, Mangayarkarasi N "Support Vector Machine Based Classification of Medicinal Plants Using Leaf Features" published in 2019 International Conference on Advances in Computing, Communications and Informatics (ICACCI)

[6] Shitala Prasad, Pankaj P Singh "Medicinal Plant Leaf Information Extraction Using Deep Features" published in Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2019

[7] Dileep M.R. and Pournami P.N. proposed "Ayur Leaf," a deep learning-based method for medicinal plant classification, presented at the 2019 IEEE Region 10 Conference (TENCON).

[8] C. Amudha Lingeswaran, Mr. Sivakumar, and Dr. P. Renuga presented a study on the identification of medicinal plants and their applications using deep learning, published in the proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019), IEEE Xplore, ISBN: 978-1-5386-9439-8.

[9] Manojkumar P., Surya C. M., Varun P. Gopi "Identification of Ayurvedic Medicinal Plants by Image Processing of Leaf Samples" Published in 2020 Third international conference on research in Computational Intelligence and communication network (ICRCICN).

[10] Amala Sabu, Sreekumar K, Rahul R Nair "Recognition of Ayurvedic Medicinal Plants from Leaves: A Computer Vision Approach" Published in 2019 Fourth International Conference on Image Information Processing (ICIIP)