



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 11, November 2025

Indian Medicinal Plant Classification System using Deep Learning

Mr. Kambale Priti Pramod ¹, Dr. Anand Khatri ², Prof. Sachin Bhosale ³, Dr. Shubhangi Gunjal ⁴

Student, Department of Artificial Intelligence and Data Science Engineering, Jaihind College of Engineering, Pune,
Maharashtra, India¹

Professor, Department of Artificial Intelligence and Data Science Engineering, Jaihind College of Engineering, Pune,
Maharashtra, India²

Professor, Department of Artificial Intelligence and Data Science Engineering, Jaihind College of Engineering, Pune,
Maharashtra, India³

Professor & HOD, Department of Artificial Intelligence and Data Science Engineering, Jaihind College of Engineering,
Pune, Maharashtra, India⁴

ABSTRACT: The accurate identification of medicinal plants is crucial for their effective use in traditional medicine, research, and conservation. Manual identification methods are time-consuming and require expert knowledge, making them less accessible to the general public. This project presents an **Indian Medicinal Plant Classification System** using deep learning to automate plant identification based on leaf images. Our system is trained on a dataset containing **80 different classes of Indian medicinal plant leaves**, sourced from Kaggle. The dataset undergoes **preprocessing and augmentation techniques**, including feature-wise centering, standardization, rotation, width and height shifts, and horizontal flips, to enhance model performance. We implemented **EfficientNet, ResNet50, and a hybrid model** to classify the plant species accurately. The best-performing model is saved and deployed for real-time predictions. The system is integrated into a **web-based application**, developed using **HTML, CSS, JavaScript, and Flask**. Users can upload an image of a medicinal plant leaf, and the system classifies the plant while providing additional information, including **chemical components, medicinal properties, diseases it can cure, age and gender restrictions, and intake guidelines for pregnant women**. This AI-powered classification system offers an efficient, accurate, and user-friendly solution for plant identification, benefiting researchers, botanists, students, and healthcare professionals. Future enhancements include expanding the dataset, improving model accuracy, and integrating mobile applications for real-time plant identification. This project represents a significant step toward merging **artificial intelligence with traditional herbal medicine**, making plant-based healthcare more accessible and reliable.

KEYWORDS: Indian medicinal plants, deep learning, plant identification, leaf image classification, EfficientNet, ResNet50, hybrid CNN model, image preprocessing, data augmentation, Flask web application, artificial intelligence.

I. INTRODUCTION

Indian medicinal plants have played a vital role in traditional healthcare practices for centuries, offering natural remedies for a wide range of diseases and health conditions. With increasing interest in herbal medicine, Ayurveda, and plant-based treatments, there is a growing need for accurate identification and classification of medicinal plants. However, manual identification by botanists can be time-consuming and prone to human error. To address this challenge, we have developed an **Indian Medicinal Plant Classification System using Deep Learning**, which automates plant identification and provides essential medicinal information.

Our system is trained on a dataset consisting of **80 different classes of Indian medicinal plant leaves**, sourced from Kaggle. The dataset undergoes rigorous **preprocessing and augmentation**, including feature-wise centering, standardization, rotation, width and height shifts, and horizontal flips. These techniques ensure that our model can generalize well to real-world images, improving its robustness and accuracy in classification.

To achieve high performance, we implemented **EfficientNet** to classify medicinal plant leaves. Each model was trained separately to analyze their individual performance, and the best-performing model was selected for deployment. The trained model is capable of recognizing plant species with high precision and is saved for real-time inference on new images uploaded by users.

Our system is deployed as a **web-based application**, built using **HTML, CSS, JavaScript, and Flask**. The application features a simple and intuitive interface, allowing users to **upload an image of a medicinal plant leaf** for classification. Once the image is processed, the system not only identifies the plant species but also provides crucial information, such as **the chemical components present in the plant, the diseases it can cure, any age or gender restrictions, and intake guidelines for pregnant women**.

This project is highly beneficial for researchers, botanists, students, and healthcare professionals who need a quick and reliable method for identifying medicinal plants. By leveraging deep learning, our system reduces dependency on manual classification and enhances accessibility to valuable medicinal plant knowledge. In the future, this system can be expanded by incorporating more plant species, improving model accuracy, and integrating it with mobile applications for on-the-go plant identification.

Our **Indian Medicinal Plant Classification System** stands as a significant step towards combining AI with traditional medicine, offering a powerful and efficient tool for the study and application of herbal remedies. With continued advancements in machine learning and data collection, this project has the potential to revolutionize medicinal plant research and make plant-based treatments more accessible to the world.

II. METHODOLOGY

The **Indian Medicinal Plant Classification System** follows a structured approach that involves **data collection, preprocessing, model training, and web application development**. The following sections describe each stage in detail.

1. Data Collection

We used a **Kaggle dataset** containing **80 different classes of Indian medicinal plant leaves**. The dataset consists of images of plant leaves captured under various lighting conditions, angles, and backgrounds. Each class represents a unique medicinal plant species commonly used in Ayurveda and herbal medicine.

2. Data Preprocessing and Augmentation

To ensure high model performance and generalization, we applied **image preprocessing and augmentation techniques**. These steps help normalize the data and introduce variations that improve model robustness.

2.1 Preprocessing

- **Resizing:** All images were resized to a fixed dimension (e.g., **224×224 pixels**) to maintain uniformity across the dataset.
- **Normalization:** Pixel values were scaled between 0 and 1 to speed up model convergence.
- **Noise Reduction:** Gaussian blurring and median filtering were applied to remove unwanted noise and improve feature clarity.

2.2 Data Augmentation

To prevent overfitting and improve the model's ability to recognize plants under different conditions, we applied various augmentation techniques:

- **Feature-wise centering:** Centers image pixel values around zero.
- **Feature-wise standardization:** Normalizes images by scaling based on the dataset's mean and standard deviation.
- **Rotation:** Images were randomly rotated up to **20 degrees**.
- **Width and height shifts:** Random shifts of up to **20%** of the image size were applied.
- **Horizontal flip:** Images were randomly flipped to introduce more variations.

3. Model Development

For classification, we employed deep learning models based on **CNN architectures**. We trained three models separately and then analyzed their performance.

3.1 Individual Models

1. **EfficientNet**: Known for its high accuracy and efficiency, it uses **compound scaling** to optimize depth, width, and resolution.
2. **ResNet50**: A widely used deep CNN model with **residual learning**, allowing deeper networks to avoid vanishing gradient problems.
3. **Hybrid Model (EfficientNet + ResNet50)**: Combines the strengths of both models by using **feature fusion techniques**.

3.2 Model Training

- **Loss Function**: **Categorical Crossentropy**, since this is a multi-class classification problem.
- **Optimizer**: **Adam optimizer** was used for faster and efficient convergence.
- **Batch Size**: Set to **32** for balancing training speed and performance.
- **Epochs**: Trained for **50–100 epochs**, depending on convergence.
- **Evaluation Metrics**: Accuracy, Precision, Recall, and F1-score were used to assess model performance.
- **Early Stopping**: Implemented to stop training when validation loss stopped improving.
- **Best Model Saving**: The model with the highest validation accuracy was saved for deployment.

4. Model Deployment

After training, we deployed the best model using **Flask**, a lightweight Python web framework.

- **Model Serialization**: The trained model was saved using **TensorFlow's SavedModel format (.h5 or .pb)**.
- **Flask API Development**: Created an API for **loading the model, processing user-uploaded images, and returning predictions**.
- **Inference Pipeline**:
 1. User uploads an image.
 2. The image is preprocessed to match the model's input format.
 3. The model predicts the plant species.
 4. The system retrieves additional information about the plant, such as:
 - **Chemical components**
 - **Diseases it cures**
 - **Age and gender restrictions**
 - **Pregnancy intake conditions**
 5. The results are displayed on the web interface.

5. Web Application Development

We built a **user-friendly web interface** for seamless interaction.

5.1 Frontend (User Interface)

Developed using **HTML, CSS, and JavaScript**, the website provides:

- **Homepage** with an introduction to the system.
- **Image Upload Page** where users can upload medicinal plant images.
- **Results Page** displaying classification results along with medicinal details.

5.2 Backend (Flask API)

- **Handles HTTP requests** from the frontend.
- **Processes uploaded images** and feeds them into the model.
- **Fetches medicinal plant details** based on the classification result.
- **Returns predictions** to be displayed in the user interface.

6. Performance Evaluation

To assess model effectiveness, we conducted **quantitative evaluations** using multiple metrics:

- **Accuracy:** The percentage of correctly classified images.
- **Precision & Recall:** Measures model's ability to correctly classify plant species.
- **Confusion Matrix:** Visualizes classification results for each class.
- **F1-score:** Evaluates the balance between precision and recall.

We also tested the model on **new unseen images** to ensure real-world applicability.

7. Future Enhancements

To further improve the system, future versions will include:

- **A larger dataset** with more medicinal plant species.
- **Mobile App Integration** for on-the-go plant identification.
- **OCR-based plant search** allowing users to input plant names and retrieve medicinal information.
- **Geolocation-based suggestions** for finding medicinal plants in specific regions.

III. CONCLUSION

This project successfully automates the identification of Indian medicinal plants using **deep learning and a web-based interface**. By integrating **EfficientNet, ResNet50, and a hybrid model**, the system ensures high accuracy in classification while providing **medicinal information** relevant to healthcare and research. The web application makes it accessible to researchers, herbalists, students, and healthcare professionals, promoting AI-driven solutions in traditional medicine.

REFERENCES

1. Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F., & Xiang, Q.-L. (2007). A leaf recognition algorithm for plant classification using probabilistic neural network. *Proceedings of the IEEE 7th International Symposium on Signal Processing and Information Technology*, Cairo, Egypt.
2. Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F., & Shiang, C.-L. (2007). Leafsnap dataset. Retrieved from <http://leafsnap.com/dataset/>
3. Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I. (2011). Leaf classification using shape, color, and texture features. *International Journal of Computer Trends and Technology (IJCTT)*, 1(3), 306–311.
4. Ehsanirad, A., & Sharath Kumar, Y. H. (2010). Leaf recognition for plant classification using GLCM and PCA methods. *Computer Science Journal*, 3(1).
5. Šulc, M., & Matas, J. (2017). Fine-grained recognition of plants from images. *Plant Methods*, 13, Article 115. <https://doi.org/10.1186/s13007-017-0265-4>
6. Çuğu, İ., Şener, E., Erciyes, Ç., Balcı, B., Akin, E., Önal, I., & Akyüz, A. O. (2017). Treelogy: A novel tree classifier utilising deep and hand-crafted representations. *arXiv*.
7. Lee, S.-H., & Chan, C. S. (2015). DEEP-PLANT: Plant identification with convolutional neural networks. *2015 IEEE International Conference on Image Processing (ICIP)*.
8. Chaki, J., Parekh, R., & Bhattacharya, S. (2018). Plant leaf classification using multiple descriptors: A hierarchical approach. *Journal of King Saud University – Computer and Information Sciences*, (2020), 1158–1172.
9. Zhang, Y., Cui, J., Wang, Z., Kang, J., & Min, Y. (2020). Leaf image recognition based on bag of features. *Applied Sciences*, 10(15), 5177.
10. Khan, F., You, J. (2021). Leaf image classification using deep-learning network. *Academic Journal of Computing & Information Science*, 4(3), 109–115.
11. Pandey, A., (2022). A robust deep attention dense convolutional neural network for medicinal-plant leaf classification. *[Journal name]*.
12. Survey paper: Jethva, J. B., & Singla, C. R. (2018). Survey paper on plant leaf identification system. *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, 4(5), 519-528.
13. Yili, Z., Gangliang, W., Qiang, W., Yue, Z., & Yong, Z. (2017). Method of leaf identification based on multi-feature dimension reduction. *Transactions of the Chinese Society of Agricultural Machinery*, 48(3).
14. Barhate, D., (2024). A systematic review of machine-learning and deep-learning algorithms for plant species recognition. *[Journal]*.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details