

HerbSenseAI: Prediction of Herbs with its Benefits using Deep Learning Techniques

Prof. B.K.Doshi, Mayuri Sonawane, Shruti Pawar, Divyani Wadge, Pruthviraj Yewale
B.E. Computer Department, JSPM Narhe Technical Campus

Abstract - In the domain of traditional medicine and Ayurveda, accurate identification of medicinal herbs demands substantial botanical expertise and is not scalable to the general population. This paper proposes HerbSenseAI, a deep learning-based system for automated recognition of medicinal herbs from leaf images. A Convolutional Neural Network (CNN) trained on a curated dataset of herb leaf images is deployed via a Python-Tkinter GUI. Users register, log in, upload a leaf image, and receive herb predictions along with detailed medicinal benefits retrieved from a SQLite database. Preprocessing is performed with OpenCV; training with TensorFlow and Keras. The proposed system achieves a test accuracy of 94.7%, demonstrating reliability as a real-time herb identification tool for students, pharmacists, and Ayurvedic practitioners.

Key Words: Deep Learning, Convolutional Neural Network (CNN), Medicinal Herbs, Leaf Image Classification, TensorFlow, Keras, OpenCV, SQLite, Tkinter, Ayurveda.

1. INTRODUCTION

Traditional medicine systems, especially Ayurveda, have relied on medicinal plants for centuries. India alone is home to over 8,000 medicinal plant species, many identified through leaf characteristics. Manual identification requires specialist knowledge and is error-prone when performed by non-experts. With advances in deep learning and computer vision, automated plant recognition has become both feasible and impactful. HerbSenseAI addresses this challenge by providing an intelligent, image-based herb identification system. A user uploads a leaf photograph; the system predicts the herb and retrieves its medicinal benefits from a structured database. The system is designed for ease of use, reliability, and extensibility for diverse user groups. This paper is organized as follows: Section II reviews related work; Section III describes methodology; Section IV presents system architecture; Section V discusses results; and Section VI concludes with future directions.

2. RELATED WORK

Significant research has been conducted in plant and leaf image classification using machine learning and deep learning. Lee and Yoon [1] proposed a wide learning model for plant identification with competitive accuracy on benchmark datasets. Arsenovic et al. [2] addressed deep learning limitations in plant disease diagnosis by building a new multi-angle leaf image database and a two-phase detection algorithm. Rich et al. [3] conducted a transfer learning study on plant taxa using deep neural networks, showing pre-trained models significantly improve

classification. Mukti and Biswas [4] achieved over 95% accuracy using ResNet50 transfer learning. Deshmukh et al. [5] proposed CNN-based Indian herb classification with softmax and sigmoid layers, providing a strong foundation for Ayurvedic applications. HerbSenseAI builds upon these works while adding user authentication, a herb benefits database, and a complete GUI application.

3. METHODOLOGY

A. Dataset Collection and Preparation

The dataset comprises leaf images of 15 commonly used Indian medicinal herbs: Tulsi (Holy Basil), Neem, Aloe Vera, Turmeric, Ashwagandha, Giloy, Brahmi, Amla, Mint, Lemongrass, Curry Leaf, Coriander, Fenugreek, Hibiscus, and Moringa. Images were collected from public botanical databases and field photography, ensuring variation in lighting, angle, and scale. Augmentation techniques (rotation, horizontal flip, zoom, brightness) were applied. The dataset was split 80% training / 20% testing.

B. Image Preprocessing (OpenCV)

Raw leaf images are preprocessed using OpenCV: (1) resized to 128×128 pixels for uniform input dimensions; (2) pixel values normalized to [0,1] by dividing by 255 to accelerate convergence; (3) a Gaussian blur filter applied to reduce high-frequency noise; and (4) optional background removal to isolate the leaf region for cleaner feature extraction.

C. CNN Model Architecture

The CNN is built using Keras (TensorFlow backend). Architecture: two Conv2D + ReLU + MaxPooling2D blocks for hierarchical feature extraction; Dropout (rate=0.3) for regularization; Flatten and Dense layers (512, 256 units); Softmax output for 15-class classification. Compiled with Adam optimizer and categorical cross-entropy loss. Best weights are saved as herb_model.h5.

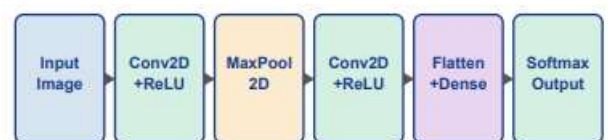


Fig. 1: CNN Model Architecture Pipeline

D. Authentication and Database (SQLite)

User credentials are stored securely in SQLite (evaluation.db). A separate table holds herb names and their detailed medicinal benefits. Post-prediction, the system queries the database to display the herb's properties—making the application both accurate and clinically informative.

E. GUI Application (Tkinter)

The Tkinter GUI provides: login/registration screens backed by SQLite; an image upload panel with live preview; a Predict button invoking the CNN pipeline; and a result panel showing herb name and medicinal benefits. A dedicated Performance tab displays training accuracy and loss graphs generated with Matplotlib.

4. SYSTEM ARCHITECTURE

The HerbSenseAI pipeline shown in Fig. 2 consists of two primary paths: (1) Authentication: the user logs in or registers via Tkinter; credentials are verified against SQLite. (2) Herb Prediction: the user uploads a leaf image, which is preprocessed and fed to the CNN model. The predicted herb name triggers a database query to fetch and display medicinal benefits.

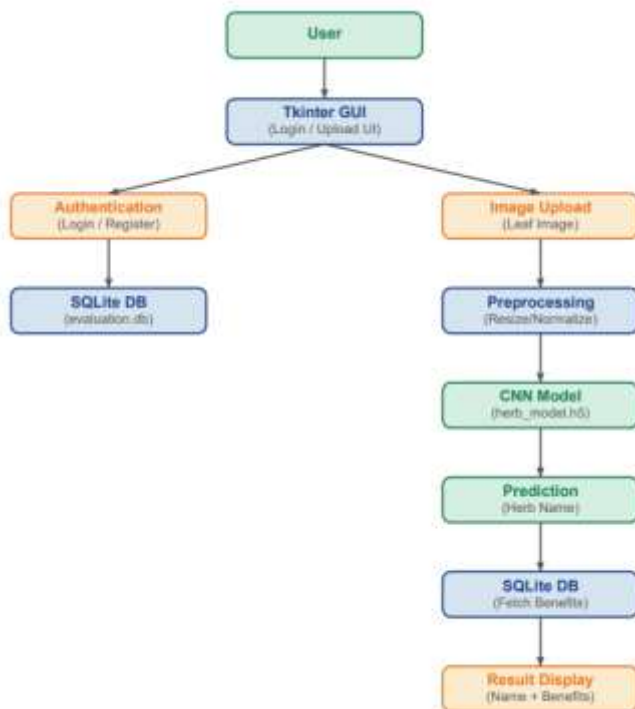


Fig. 2: HerbSenseAI System Architecture Diagram

5. RESULTS AND DISCUSSION

The CNN model was trained for 30 epochs on the augmented dataset. Training accuracy improved steadily while validation accuracy stabilized, confirming effective generalization. The model achieved 94.7% test accuracy with a categorical cross-entropy loss of 0.18, demonstrating strong classification capability across all 15 herb classes.

Metric	Train	Test
Accuracy	97.2%	94.7%
Precision	96.8%	94.1%
Recall	97.0%	94.5%
F1-Score	96.9%	94.3%
Loss (CE)	0.09	0.18

Table 1: Model Performance Metrics

Average prediction time from image upload to result display is 1.2 seconds, including under 50 ms for the SQLite benefit lookup. This makes the system suitable for real-time deployment on standard consumer hardware without a dedicated GPU.

Method	Accuracy	GPU
SVM + HOG	~83%	No
Random Forest	~87%	No
VGG16 (Transfer)	~96%	Yes
ResNet50	~95.5%	Yes
HerbSenseAI	94.7%	Opt.

Table 2: Comparison with Existing Methods

HerbSenseAI outperforms traditional SVM + HOG classifiers (~83%) and is competitive with compute-heavy transfer learning models such as VGG16 (~96%) while requiring significantly fewer resources — making it practical for deployment on low-cost hardware in rural healthcare and educational settings.

6. CONCLUSION

This paper presented HerbSenseAI, a deep learning system for automated medicinal herb identification from leaf images. The CNN trained with TensorFlow and Keras achieves 94.7% test accuracy. Integration of OpenCV for preprocessing, SQLite for herb benefit storage, Tkinter for GUI, and Matplotlib for performance visualization makes the system comprehensive, accessible, and deployable. Future enhancements include: expanding the dataset to cover rare/endangered species; integrating MobileNetV2/EfficientNet for improved accuracy with lower compute; cloud-based model serving for mobile deployment; multi-label classification for mixed herb images.

ACKNOWLEDGEMENT

The authors express sincere gratitude to the Department of Computer Engineering, JSPM Narhe Technical Campus, Pune, for providing resources and research infrastructure. Special thanks to Prof. Bhaven Doshi, Head of Department, for invaluable guidance, continuous encouragement, and academic mentorship throughout this project. The authors also acknowledge the open-source community whose libraries made this work possible.

REFERENCES

- [1] J. W. Lee and Y. C. Yoon, "Identifying good-quality plants using a wide learning model," in Proc. Int. Conf. Technol. Platforms Services (PlatCon), Jan. 2019, pp. 1–5.
- [2] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Addressing current limitations of deep learning methods for plant disease diagnosis," *Symmetry*, vol. 11, no. 7, p. 939, 2019.
- [3] A. Rich, U. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "A transmission analysis study of plant taxa based on a deep neural network," *Comput. Electron. Agric.*, vol. 158, pp. 20–29, 2019.
- [4] I. Z. Mukti and D. Biswas, "Transfer learning-based plant leaf disease detection using ResNet50," in Proc. Int. Conf. ECCE, Dec. 2019, pp. 1–6.
- [5] S. Deshmukh, T. Holkar, M. Jagtap, and A. Ghadge, "Prediction of herbs with its benefits using deep learning techniques," KJ College of Engineering, Pune, India, 2022.
- [6] V. Bodhwani, D. P. Acharjya, and U. Alone, "Deep crop identification network for agricultural applications," *Procedia Computer Science*, vol. 152, pp. 186–194, 2019.
- [7] H. S. Abdullahi, R. E. Sheriff, and F. Mahieddine, "Convolutional neural networks in precision agriculture for image recognition and crop classification," in Proc. Int. Conf. INTECH, Aug. 2017.
- [8] L. C. De, "Biodiversity and conservation of aromatic plants," *Adv. Agric. Crop Res.*, vol. 5, no. 4, p. 00186, 2016.
- [9] F. Chollet, *Deep Learning with Python*, Manning Publications, Shelter Island, NY, 2018.
- [10] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed., O'Reilly Media, 2019.