AI-Based Medicinal Plant Identification Using Deep Learning for Mobile Applications

Anumol V S
Computer science and engineering
St. Joseph's College of Engineering and Technology
Palai, India
anumolvs2026@cs.sjcetpalai.ac.in

Elna S Bijo
Computer science and engineering
St. Joseph's College of Engineering and Technology
Palai, India
elnasbijo2026@cs.sjcetpalai.ac.in

Neha Maria Joji
Computer science and engineering
St. Joseph's College of Engineering and Technology
Palai, India
nehamariajoji2026@cs.sjcetpalai.ac.i

Siya Varghese
Computer science and engineering
St. Joseph's College of Engineering and Technology
Palai, India
siyavarghese2026@cs.sjcetpalai.ac.in

Prof. Teena George
Assistant Professor
Computer science and engineering
St. Joseph's College of Engineering and Technology
Palai, India
teena@cs.sjcetpalai.ac.in

***Abstract*—Traditional rural and Ayurvedic healing systems represent a valuable component of cultural heritage, yet much of this knowledge is gradually diminishing due to modernization and limited digital documentation. This paper presents an AI-based mobile application designed to support medicinal plant identification and contribute to the digital preservation of traditional herbal knowledge. The proposed system integrates a Convolutional Neural Network (CNN) along with MobileNet V2 for image-based plant recognition with a structured medicinal knowledge base curated through expert consultation and verified data sources.**

The application enables users to identify medicinal plants using smartphone image capture and provides detailed plant profiles containing scientific and local names, medicinal uses, preparation methods, and safety precautions. The backend knowledge repository is organized to ensure efficient storage and retrieval of plant information, supporting structured and scalable data management.

By combining artificial intelligence with an intuitive mobile interface, the system aims to improve accessibility to reliable medicinal information while promoting awareness of traditional practices. The implementation demonstrates a practical approach toward sustainable digital preservation and technology-assisted community engagement in herbal healthcare knowledge.

***Index Terms*—Medicinal plant identification, Deep learning, CNN, TensorFlow Lite, Mobile application, Ayurveda.**

I. INTRODUCTION

A. Background and Motivation

Medicinal plants play an important role in traditional healthcare systems, especially in Kerala where biodiversity and Ayurvedic practices are widely preserved. However, traditional medicinal knowledge is often transmitted orally and risks being lost due to modernization and changing lifestyles. This creates a need for digital systems capable of preserving and providing accessible information about medicinal plants.

B. Problem Overview

Identifying medicinal plants correctly is difficult for non-experts. Most existing digital plant identification systems require continuous internet access and often provide limited medicinal information. Incorrect usage of medicinal plants can also lead to unsafe practices. Therefore, there is a need for an intelligent plant identification system that works offline while providing reliable medicinal knowledge.

C. Contribution of the Project

The proposed system, Herbaura, integrates an AI-based image recognition model with a structured medicinal knowledge database. The mobile application allows users to identify medicinal plants from images and search remedies based on symptoms. Using on-device deep learning with a SQLite database converted to JSON format, the system provides offline accessibility while supporting digital preservation of Ayurvedic medicinal knowledge.

II. PROBLEM STATEMENT AND MOTIVATION

A. Problem Statement

Medicinal plants are important in traditional healthcare systems like Ayurveda. However, identifying these plants can

DOI: 10.5281/zenodo.19826217

be hard for non-experts because many species look alike and there are not enough reliable reference materials available. Most traditional knowledge is passed down orally, which makes it at risk of being lost over time.

Current digital plant identification systems usually need a steady internet connection. They mainly focus on identifying species but do not provide detailed medicinal information. Many applications also lack knowledge specific to different regions or safety precautions. This can result in misuse or misinformation.

There is a need for a system that can identify medicinal plants, provide reliable information about them, and function effectively in areas with poor connectivity.

B. Motivation

The Herbaura project is motivated by the growing interest in natural and traditional medicine and the fast development of artificial intelligence technologies. Kerala has a wealth of biodiversity and a strong Ayurvedic heritage, making it an ideal place to create a system for identifying medicinal plants.

By using deep learning and mobile computing, Herbaura aims to connect traditional medicinal knowledge with modern digital platforms. The project seeks to preserve indigenous knowledge, promote safe use of medicinal plants, and give users a simple and trustworthy tool for plant identification and medicinal references.

C. Need for the Proposed System

The proposed Herbaura system addresses the limitations of existing solutions by:

- Enabling AI-based image recognition of medicinal plants
- Supporting offline usability through on-device inference
- Providing structured medicinal information including usage and safety details
- Preserving traditional Ayurvedic knowledge in a digital format

This makes Herbaura a practical, educational, and scalable solution for medicinal plant identification and knowledge preservation.

III. LITERATURE SURVEY

Recent studies have demonstrated the effectiveness of deep learning techniques in medicinal plant identification. Acasamoso Jr. et al. proposed a CNN-based approach for classifying medicinal plants using optimized deep learning architectures, highlighting the applicability of such systems in rural and remote areas for herbal identification [1]. Their results emphasize the potential of CNNs as a foundation for mobile-based plant recognition systems.

Wang and Zhao presented a plant intelligent recognition system based on transfer learning, demonstrating improved classification performance by leveraging pre-trained deep learning models [2]. Their work shows that transfer learning significantly reduces training complexity while maintaining high recognition accuracy, making it suitable for practical deployment.

Sehaba et al. introduced an embedded plant recognition benchmark focusing on low-footprint deep neural networks for resource-constrained environments [3]. This study is particularly relevant for mobile and edge-based applications, as it highlights the feasibility of deploying deep learning models on devices with limited computational resources.

Hussain et al. proposed an enhanced medicinal plant identification framework combining image processing techniques with machine learning models to improve raw material recognition accuracy [4]. Their work underlines the importance of accurate preprocessing and feature extraction in plant classification tasks.

Yuan et al. explored leaf-based image recognition using dynamic threshold segmentation combined with deep learning techniques [5]. Their findings demonstrate that effective segmentation plays a critical role in improving classification accuracy for medicinal plant leaves.

Custodio proposed an improved medicinal plant leaf classification approach using transfer learning and ensemble deep convolutional neural networks [6]. The study highlights the benefits of combining multiple deep learning models to enhance robustness and classification performance.

Although these works focus on plant identification accuracy, most existing systems mainly emphasize classification and lack integrated knowledge about medicinal uses, symptom-based searches, and offline usability. Herbaura builds on existing research by combining on-device CNN-based identification, a symptom-based search for medicines, and a structured SQLite-based knowledge database serialized in JSON format. It aims to preserve and promote the practical use of Ayurvedic medicinal knowledge.

IV. METHODOLOGY

The Herbaura methodology describes the system architecture, data handling, and deep learning approach used for medicinal plant identification and knowledge retrieval. The system is designed to operate efficiently on mobile devices while supporting offline access to medicinal information.

A. System Overview

Herbaura is a mobile app that combines an AI-powered image recognition feature with a structured database of medicinal knowledge. The system has two main functions: identifying medicinal plants through images and searching for remedies based on symptoms. The entire process starts with user interaction via a mobile interface and ends with the display of confirmed medicinal information.

A Convolutional Neural Network (CNN) optimized with TensorFlow Lite performs image classification on the device. This setup allows for quick response times and works offline. Medicinal plant data is stored in SQLite and converted to JSON format for smooth data transfer between system parts.

B. Image-Based Plant Identification

In the image-based identification process, the user takes a picture of a medicinal plant with the mobile device camera.

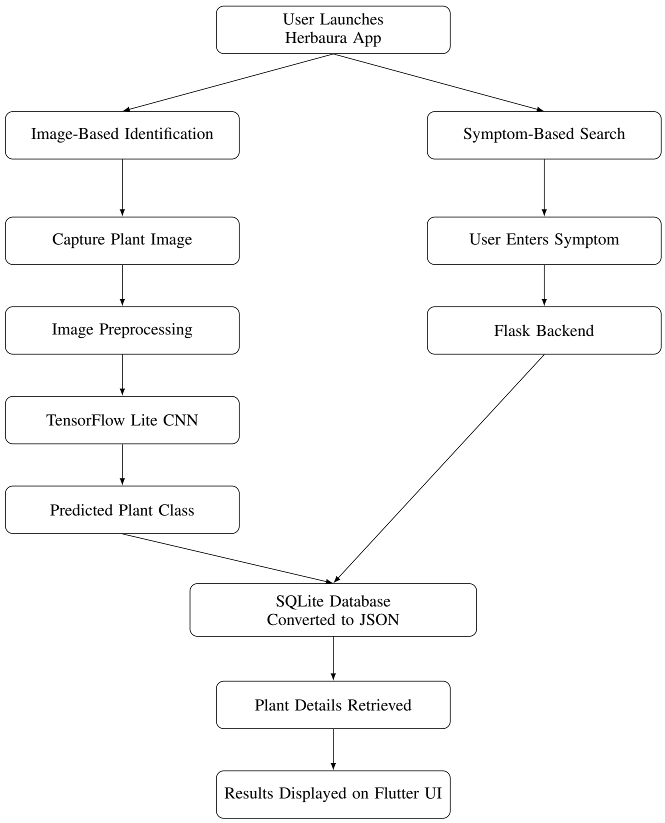

Fig. 1. Workflow of the Herbaura system illustrating image-based medicinal plant identification and symptom-based search using a SQLite knowledge base serialized in JSON format.

The picture goes through preprocessing steps like resizing and normalization to meet the model's input requirements.

The preprocessed image is then sent to the CNN model that runs on TensorFlow Lite. The model predicts the plant species based on learned visual features. After a successful classification, the predicted plant label is used to fetch related medicinal details from the knowledge base.

C. Symptom-Based Remedy Search

In addition to image-based plant identification, Herbaura provides a symptom-based remedy search feature. Users can enter symptoms such as cough, fever, or headache through the mobile application. The backend service processes the input and retrieves relevant medicinal remedies from the database associated with those symptoms.

This functionality allows users to explore suitable herbal remedies even when they are unable to identify a specific plant through images, thereby improving accessibility and practical usability of the system.

D. Knowledge Base and Data Handling

Medicinal plant information is stored locally with SQLite, a lightweight relational database that works well for mobile applications. Each plant entry includes scientific names, local names, medicinal uses, and safety precautions.

For smooth data exchange and frontend compatibility, the SQLite database is converted into JSON format. This method allows for efficient querying, organized storage, and simple integration with both the frontend and backend parts of the system.

E. Model Training Configuration

The plant classification model was developed using a transfer learning approach based on the MobileNetV2 architecture pretrained on ImageNet. The last 20 layers of the network were fine-tuned to adapt the model to the medicinal plant dataset.

The model was trained using the Adam optimizer with two learning phases. In the first phase, a learning rate of 1×10^{-3} was used, followed by fine-tuning with a learning rate of 1×10^{-4}.

The key training hyperparameters used in the experiment are shown in Table I.

TABLE I
TRAINING HYPERPARAMETERS

Parameter	Value
Architecture	MobileNetV2
Input Image Size	224×224
Optimizer	Adam
Learning Rate	10^{-3} (Phase 1), 10^{-4} (Phase 2)
Batch Size	32
Epochs	10 (Phase 1), 8 (Phase 2)
Loss Function	Categorical Cross Entropy

F. Implementation Tools and Technologies

The Herbaura system is implemented using the following technologies:

- Frontend: Flutter (Dart) for cross-platform mobile development
- AI Model: Convolutional Neural Network deployed using TensorFlow Lite
- Backend: Python Flask for handling requests and data retrieval
- Database: SQLite with JSON-based data representation

This combination ensures a lightweight, scalable, and efficient system suitable for real-world usage.

V. RESULTS AND DISCUSSION

The Herbaura system was evaluated using a dataset of 2927 images covering 14 medicinal plant species and an additional "Unknown" class, resulting in 15 classes. An 80–20 train-validation split was used, with 582 images reserved for validation.

Dataset Preparation: The dataset was primarily constructed through controlled image collection conducted by the project team. The majority of the images were captured directly by the authors to ensure consistency in lighting conditions, background variation, and plant representation.

To enhance class diversity and improve generalization performance, additional images for specific classes such as Neem, Tulsi, and Aloe Vera were integrated from publicly available Kaggle datasets. For these categories, both self-collected and publicly sourced images were combined to improve representation and variability.

All images were carefully reviewed for quality and relevance. The dataset was resized to a uniform resolution of 224×224 pixels and normalized prior to training the Convolutional Neural Network (CNN).

Image-Based Identification and App Results: The trained CNN model ran on TensorFlow Lite and successfully classified the plant species under controlled testing conditions. On-device inference allowed for faster response times and supported offline use. The Flutter-based mobile app showed smooth overall functionality. This included image capture and cropping, plant prediction, and displaying medicinal details, as shown in Fig. 2.

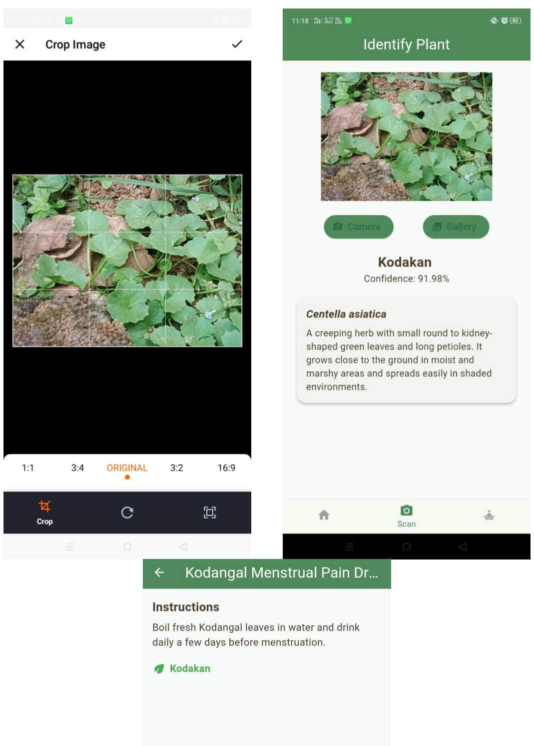

Fig. 2. Herbaura mobile application workflow showing (left) image cropping, (center) plant identification result with confidence score, and (right) medicinal instructions displayed to the user.

Database Evaluation and Discussion: Medicinal plant information was stored in a SQLite database and converted into JSON format for efficient retrieval within the mobile application. The system consistently displayed accurate medicinal details corresponding to each predicted plant class.

In addition to functional validation, quantitative evaluation was conducted to assess model performance. A comparative analysis between a baseline Simple CNN and a transfer learning-based MobileNetV2 architecture demonstrated significant improvement in validation accuracy and deployment efficiency. The MobileNetV2 model achieved substantially higher classification performance while maintaining a smaller TensorFlow Lite model size, making it more suitable for real-time mobile deployment.

The modular system design enables future expansion of the dataset and further retraining to improve generalization across additional medicinal plant classes.

Confusion Matrix and Class-wise Performance The confusion matrix shown in Fig. 3 illustrates the classification performance across individual plant classes using the validation dataset. The MobileNetV2 model achieved an overall classification accuracy of 88.83%. The macro-average precision, recall, and F1-score were 0.896, 0.902, and 0.891 respectively. Several plant classes such as Kayyunyam, Neem, Panikoorkka, and Poovamkurunila achieved perfect recall (100%), indicating highly accurate identification for these species.

However, some visually similar plant species exhibited confusion during classification. For example, Muyalcheviyan was frequently misclassified as Kizharnelli due to similarities in leaf structure. Similarly, Cherula occasionally showed confusion with Poovamkurunila.

This analysis helps identify challenging plant pairs and indicates directions for improving dataset diversity and model robustness in future work.

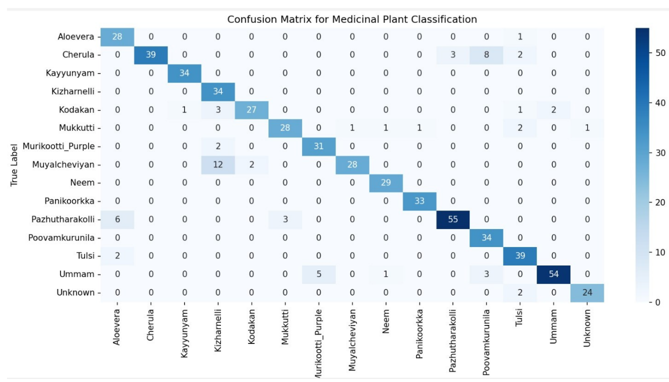

Fig. 3. Confusion matrix showing classification performance of the MobileNetV2 model across 15 classes (14 medicinal plants and one Unknown category).

Model Performance Comparison: To evaluate the effectiveness of the proposed approach, a comparative analysis was conducted between a baseline Simple CNN architecture and a transfer learning-based MobileNetV2 model. Both models were trained and evaluated using the same dataset, preprocessing steps, and an 80–20 train–validation split to ensure a fair comparison.

TABLE II
PERFORMANCE COMPARISON BETWEEN SIMPLE CNN AND MOBILENETV2

Metric	Simple CNN	MobileNetV2
Validation Accuracy (%)	51.51	88.83
Precision (Macro)	0.4749	0.896
Recall (Macro)	0.5047	0.902
F1-score (Macro)	0.4637	0.891
TFLite Model Size (MB)	10.66	2.55

Macro-average precision, recall, and F1-score were used to ensure equal importance to all plant classes during performance evaluation.

The Simple CNN achieved a validation accuracy of 51.51%, indicating limited generalization capability when trained from scratch on a moderate-sized dataset containing 14 medicinal plant species along with an additional "Unknown" class. The macro-average precision, recall, and F1-score were 0.4749, 0.5047, and 0.4637 respectively, reflecting inconsistent performance across classes.

In contrast, the MobileNetV2 model achieved a significantly higher validation accuracy of 88.83%. The macro-average precision, recall, and F1-score improved to 0.896, 0.902, and 0.891 respectively, demonstrating superior classification performance and better generalization across all plant categories. Furthermore, the MobileNetV2 TensorFlow Lite model size was considerably smaller, making it more suitable for real-time mobile deployment. These results justify the selection of MobileNetV2 as the final model for the Herbaura application.

VI. CONCLUSION AND FUTURE WORK

This project introduced Herbaura, an AI mobile app designed for identifying medicinal plants and preserving traditional herbal knowledge digitally. The system combines a Convolutional Neural Network (CNN) that uses TensorFlow Lite for image recognition on devices with a structured medicinal knowledge base stored in SQLite and serialized in JSON format. The app allows users to identify medicinal plants through images and search for remedies based on symptoms. This makes it easy for users to access reliable medicinal information.

The current implementation acts as a proof of concept.It was evaluated using a dataset containing 14 medicinal plant species along with an additional "Unknown" class used to handle non-medicinal or unrecognized plant inputs. Functional testing showed that the system integrates well, can be used offline, and effectively retrieves medicinal information. This work focused on validating the system workflow, handling data, and the interaction between mobile, AI, and backend components instead of highlighting large-scale performance metrics.

Future work will focus on expanding the dataset to include a wider range of medicinal plant species. This will help improve model generalization and real-world use. Additional improvements involve increasing classification accuracy through better data enhancement methods. We will introduce multilingual support to boost accessibility and add a chatbot for guided user interaction. We will also optimize the model for low-end devices. Community-driven knowledge validation can increase the system's scalability and reliability.

Overall, Herbaura shows how combining artificial intelligence and mobile technology can help preserve and promote traditional medicinal knowledge. This creates a solid base for future development and use.

REFERENCES

[1] D. M. Acasamoso Jr., "Classification of medicinal plant using an optimized deep learning method," in *Proc. 15th Int. Conf. Information and Communication Technology Convergence (ICTC)*, IEEE, 2024.

[2] Y. Wang and Z. Zhao, "Design of plant intelligent recognition system based on transfer learning," in *Proc. 8th Int. Conf. Electrical, Mechanical and Computer Engineering (ICEMCE)*, IEEE, 2024.

[3] M. E. A. Sehaba, C. Crispim-Junior, and L. T. Rodet, "Embedded plant recognition: A benchmark for low footprint deep neural networks," in *Proc. IEEE/CVF Int. Conf. Computer Vision Workshops (ICCVW)*, 2023, pp. 1–8.

[4] G. F. Hussain, D. M. Ismail, A. S. Almarhoon, and N. T. B. H. Ahmed, "Enhanced medicinal plant identification: Leveraging machine learning and image processing for accurate raw material recognition," in *Proc. 4th Int. Conf. Intelligent Technologies (CONIT)*, IEEE, 2024.

[5] Z. Yuan *et al.*, "Image recognition study on leaf characteristics of medicinal plants based on Otsu dynamic threshold segmentation algorithm," in *Proc. IEEE 2nd Int. Conf. Mobile Networks and Wireless Communications (ICMNWC)*, 2022, pp. 1–8.

[6] E. F. Custodio, "Improved medicinal plant leaf classification using transfer learning and model averaging ensemble of deep convolutional neural networks," in *Proc. Int. Conf. Information Technology Research and Innovation (ICITRI)*, IEEE, 2024, pp. 347–352.

[7] Y. A. Viknesh G. S., "Leaf classification for plant recognition using EfficientNet architecture," in *Proc. IEEE 4th Int. Conf. Advances in Electronics, Computers and Communications (ICAECC)*, 2022, pp. 1–7.

[8] Q. Wang, "Plant species recognition and classification algorithm based on deep learning," in *Proc. 3rd Int. Conf. Integrated Circuits and Communication Systems (ICICACS)*, Changchun, China, 2025.

[9] R. Sumathy, R. J. Shalom, A. N. Antony, and G. Ragul, "Real-time prediction of medicinal herb plants using deep learning," in *Proc. 4th Int. Conf. Sentiment Analysis and Deep Learning (ICSADL)*, 2025.

[10] L. Zhou, H. Gao, D. Gao, and Q. Zhao, "Recognition of Tibetan medicinal plants based on local feature localization," in *Proc. IEEE 5th Int. Conf. Pattern Recognition and Machine Learning (PRML)*, 2024.

[11] C. Yang, M. Yang, and S. Guo, "Research on plant feature extraction and intelligent recognition system based on computer vision," in *Proc. Asia-Europe Conf. Cybersecurity, Internet of Things and Soft Computing (CITSC)*, 2025.

[12] D. Vinod, P. M. Jacob, S. Simon, A. Johnson, and J. M. Mathews, "An Android app for Ayurvedic herb identification and classification using machine vision," in *Proc. IEEE Int. Conf. Computational Intelligence and Sustainable Engineering Solutions (CISES)*, Greater Noida, India, 2024, pp. 264–270.

[13] P. S. Bora, A. Bhatt, A. Paul, and R. K. Singh, "Identification and classification of rare medicinal plants," in *Proc. IEEE Int. Conf. Computing, Communication and Intelligent Systems (ICCCIS)*, Greater Noida, India, 2024, pp. 1–6.

[14] P. Kambli *et al.*, "Identification of Ayurvedic plants using deep learning," in *Proc. IEEE Int. Conf. Computing, Semiconductor, Mechatronics, Intelligent Systems and Communications (COSMIC)*, Bangalore, India, 2024, pp. 72–76.

[15] V. E. S. Vani, M. G. R. Rao, and K. Rai, "Deep learning techniques for herbal plant detection," in *Proc. 1st Int. Conf. Software, Systems and Information Technology (SSITCON)*, 2024, pp. 1–8.

[16] P. Indoriya and S. Barde, "Prediction of Ayurvedic herbs for specific diseases by classification techniques in machine learning," in *Proc. 2nd Asian Conf. Innovation in Technology (ASIANCON)*, Ravet, India, 2022, pp. 1–4.

[17] V. Serrao, A. K. M., and Varsha, "AyurMate: Blending Ayurveda and artificial intelligence for personalized wellness," *Int. J. Advanced Computer Science and Applications (IJACSA)*, vol. 16, no. 4, 2025.

[18] S. Ghosh and M. Thenmozhi, "Prediction of Ayurveda herbs and its benefits using CNN and MobileNet algorithms in deep learning," in *Proc. 4th Int. Conf. Intelligent Technologies (CONIT)*, Karnataka, India, 2024.

[19] B. M. Moreno and P. E. Cruvinel, "Computer vision system for identifying farming weed species," in *Proc. IEEE Int. Conf. Semantic Computing (ICSC)*, São Carlos, Brazil, 2022, pp. 287–292.

[20] Y.-K. Chen, K.-H. Liu, and C. Lin, "Plant species recognition based on hyperspectral plant canopy images with deep learning," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Pasadena, USA, 2023, pp. 7567–7570.