Deep Learning-based Automated Identification of Indian Medicinal Plants Using Leaf Imagery

Dr. Keerthi Kumar HM
Dept. of Computer Science and Engg.
Malnad College of Engineering Hassan, India

Guide
Chirag MR
Dept. of Computer Science and Engg.
Malnad College of Engineering Hassan, India

Gnanesh HJ
Dept. of Computer Science and Engg.
Malnad College of Engineering Hassan, India

Rakshith KV
Dept. of Computer Science and Engg.
Malnad College of Engineering Hassan, India

Bhoomika J
Dept. of Computer Science and Engg.
Malnad College of Engineering Hassan, India

***Abstract*—The accurate identification of medicinal plants is crucial for ethnobotanical research, pharmaceutical development, and biodiversity conservation. Traditional identification methods rely on expert knowledge, which is time-consuming and not always accessible. This paper proposes a deep learning-based framework using Convolutional Neural Networks (CNNs) for the automated classification of Indian medicinal plants through leaf imagery. We systematically compare the performance of five state-of-the-art CNN architectures: MobileNetV2, ResNet50, InceptionV3, Xception, and DenseNet121. InceptionV3 achieves the best classification accuracy of 95%, demonstrating robust performance in extracting spatial hierarchies of leaf features. Furthermore, we introduce *Tinkerit*, a lightweight desktop appli- cation with a graphical user interface (GUI), enabling real-time plant recognition. Our solution paves the way for modernizing herbal medicine practices through AI-driven tools, contributing to both healthcare innovation and ecological preservation.**

I. INTRODUCTION

Plants have always played a vital role in sustaining life, particularly in the domain of traditional medicine. India, with its rich biodiversity, hosts thousands of medicinal plant species used in Ayurveda and Siddha medicine systems. However, identifying medicinal plants based solely on morphology—especially leaves—poses a significant challenge for non-experts due to visual similarity among species.

The growing capabilities of computer vision, particularly with CNNs, offer promising avenues for plant classification tasks. Leaf images carry distinguishable features such as venation patterns, shape, and texture, which can be leveraged for species identification using deep learning techniques.

This paper aims to:

- Evaluate and benchmark multiple CNN models for Indian medicinal plant classification.
- Identify the most accurate and efficient architecture for deployment.
- Develop Tinkerit, a user-friendly application integrating the trained model.

II. RELATED WORK

Numerous attempts have been made in recent years to leverage machine learning and image processing for plant identification:

- De Luna et al. [1] implemented a simple ANN model for classifying 10 types of Philippine herbal plants using 600 leaf images.
- Nguyen and Hoang [2] used CNN architectures such as VGG16, ResNet50, and InceptionV3 for wild medicinal plant recognition from the VNPlant-200 dataset.
- The MedLeaf app [3] adopted Local Binary Pattern features with Probabilistic Neural Networks (PNN) for Indonesian plants, achieving only 56.33% accuracy, highlighting the need for deeper feature extraction methods.
- MLBP (Modified Local Binary Patterns) [4] enhanced texture-based plant leaf classification using histogram features.
- The SDAMPI algorithm [5] used chain-code and contour descriptors for shape-based matching of medicinal plants.

Despite these advancements, limitations persist due to shallow models, limited datasets, and lack of robust end-user deployment systems.

III. METHODOLOGY

A. Dataset Collection and Preprocessing

A custom dataset was curated consisting of high-resolution leaf images of various Indian medicinal plants. Images were captured under controlled lighting and also sourced from online databases and field surveys.

Each image was:

- Resized to 224x224 pixels to match CNN input requirements.
- Denoised using Gaussian filters to reduce background artifacts.
- Normalized to [0,1] pixel intensity range.
- Augmented using:

- Rotation (±30°)
- Horizontal/vertical flipping
- Random zoom and crop
- Brightness variation

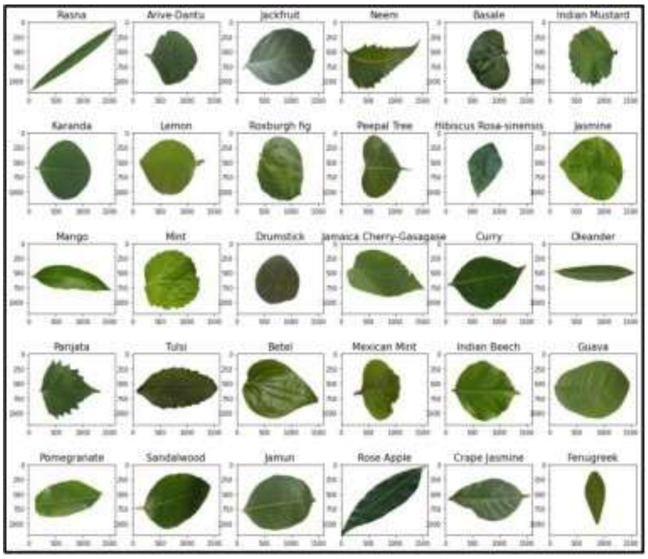

Fig. 1. Sample images of Indian medicinal plant leaves used in the dataset.

B. *CNN Architectures Evaluated*

The following pre-trained models were fine-tuned on our dataset using transfer learning:

- **MobileNetV2:** Optimized for mobile devices; uses inverted residual blocks and depthwise separable convolu- tions.
- **ResNet50:** Implements skip connections to prevent vanishing gradient issues in deep networks.
- **DenseNet121:** Enhances gradient flow by connecting each layer to every other layer.
- **Xception:** Combines depthwise separable convolutions with residual architecture for efficient computation.
- **InceptionV3:** Leverages factorized convolutions and multi-scale feature extraction within inception modules.

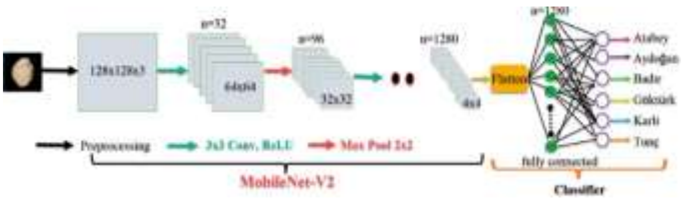

Fig. 2. MobileNetV2 architecture: Inverted residual blocks and depthwise separable convolutions.

C. *Training Configuration*

D. *Tinkerit Application*

Tinkerit is a cross-platform GUI application developed using Python's Tkinter library. It:

TABLE I
TRAINING PARAMETERS

Parameter	Value
Epochs	50
Batch Size	32
Optimizer	SGD (Stochastic Gradient Descent)
Learning Rate	0.001
Loss Function	Categorical Crossentropy
Validation Split	20%
Framework	TensorFlow / Keras

- Accepts image input via file upload or webcam.
- Preprocesses the image using OpenCV pipelines.
- Passes the image to the trained model for inference.
- Displays predicted species name and confidence level.

IV. RESULTS

TABLE II
MODEL ACCURACY COMPARISON

Model	Accuracy (%)
MobileNetV2	89.5
ResNet50	91.2
DenseNet121	92.8
Xception	93.4
InceptionV3	95.0

V. DISCUSSION

The superior performance of InceptionV3 can be attributed to its ability to capture multi-scale spatial features efficiently. Its modular architecture enables a balance between depth and computational feasibility. While MobileNetV2 offers near- real-time inference speeds ideal for edge devices, it compro- mises slightly on accuracy.

The augmentation strategies applied during preprocessing significantly improved the model's generalization on unseen test data. However, overfitting was observed in earlier trials with deeper networks on smaller datasets, which was mitigated through dropout layers and learning rate scheduling.

VI. LIMITATIONS

- Reliance on visual features alone may not suffice for species with visually similar leaves.
- Environmental variables such as lighting, shadow, and background clutter can affect model performance.
- Dataset size and class imbalance limited the learning scope for rare species.
- The system does not currently incorporate multi-modal inputs such as plant bark or flower images.

VII. OPPORTUNITIES AND FUTURE WORK

Future directions to enhance the model and system include:

- Expanding the dataset to include multi-regional medicinal flora and seasonal variations.
- Incorporating chemical, aromatic, or infrared spectral data as supplementary inputs.

- Developing mobile and web-based versions of Tinkerit with offline support.
- Integrating continual learning for adapting to new species.
- Utilizing object detection techniques (like YOLOv8 or EfficientDet) for real-time multi-leaf identification.

VIII. CONCLUSION

This research demonstrates the efficacy of deep learn- ing, particularly InceptionV3, in classifying Indian medicinal plants based on leaf images. The integration of a user-friendly GUI tool makes the system practical for field deployment by botanists, researchers, and rural health workers. With further refinements and data integration, such AI-powered tools can bridge the knowledge gap in medicinal plant usage and con- tribute to preserving India's botanical heritage.

REFERENCES

[1] R. G. De Luna, A. A. Mananghaya, J. P. Cabatuan, "Identification of Philippine herbal medicine plant leaf using artificial neural network," *IEEE Region 10 Conference (TENCON)*, 2017.
[2] Q. T. Nguyen and V. T. Hoang, "Classification of wild medicinal plants with VNPlant-200 dataset using CNN," *International Conference on Information and Communication Technology Convergence (ICICTA)*, 2020.
[3] D. A. T. Putra, A. Maulana, A. F. Sani, "MedLeaf: Herbal plant recognition app using LBP and PNN," *Proc. IES*, 2020.
[4] M. Andayani, R. Wardoyo, A. Harjoko, "Leaf image-based plant clas- sification using Modified Local Binary Pattern and histogram," *Journal of Theoretical and Applied Information Technology*, 2019.
[5] A. B. S. Kumar, V. S. K. R. Kolluru, R. V. R. Achanta, "Shape descriptor-based matching of medicinal plant images (SDAMPI)," *IEEE INDICON*, 2021.