Medical Plant Detection

Shreyash Mukund Vhanyalkar
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India
shreyashvhanyalkar@gmail.com

Abhuday Sudhir Nalawade
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India
abhyuday1012@gmail.com

Shreyash Shrikant Patil
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India
shreyashpatil218@gmail.com

Manik Shankar Kute
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India
manikkute97@gmail.com

Akshay Chinchanikar
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India

Uma Gurav
KIT's College of Engineering (Autonomous)
Kolhapur, Maharashtra, India
gurav.uma@kitcoek.in

***Abstract*—This research paper presents a deep learning approach for the classification of medicinal plants using a convolutional neural network based on the ResNet50 architecture. The study utilizes a dataset comprising 40 classes of medicinal plants, applying extensive data augmentation techniques to enhance model generalization. The pretrained ResNet50 model is fine-tuned by freezing early layers and retraining the later layers and fully connected layers to adapt to the specific classification task. The model is trained on a GPU-enabled environment to accelerate computation. Experimental results demonstrate progressive improvements in training accuracy and validation accuracy, achieving high performance in recognizing diverse medicinal plant species. The proposed method highlights the effectiveness of transfer learning combined with data augmentation for plant species classification, which can aid in botanical research and applications in traditional medicine identification.**

***Index Terms*—**

I. INTRODUCTION

Medicinal plants have long been an essential resource for traditional medicine and pharmaceutical research due to their therapeutic properties. Accurate identification and classification of these plants are crucial for ensuring their proper use and conservation. However, manual identification by experts is time-consuming and prone to errors, especially given the vast diversity of plant species.Recent advances in deep learning, particularly convolutional neural networks (CNNs), have revolutionized image classification tasks by enabling automatic feature extraction and robust pattern recognition. Transfer learning, which leverages pretrained models such as ResNet50, allows for efficient training on specific datasets with limited samples by fine-tuning existing networks. This research focuses on developing a reliable and efficient method for medicinal plant classification using a ResNet50-based CNN model. By employing data augmentation techniques and fine-tuning the pretrained network, the study aims to achieve high classification accuracy across 40 different medicinal plant species. The proposed approach not only facilitates rapid identification but also contributes to the broader application of deep learning in botanical sciences and traditional medicine.

II. RELATED WORKS

Deep learning, especially convolutional neural networks (CNNs), has become the leading approach for medicinal plant classification, often achieving high accuracy in leaf-based identification tasks. Transfer learning using pretrained models like ResNet50 is commonly employed to leverage general features and adapt to specific plant datasets. Data augmentation techniques help overcome challenges of limited and imbalanced datasets.Most studies focus on private datasets, primarily using leaf images, with reported accuracies frequently above 90%. However, challenges remain due to the scarcity of large, diverse public datasets and the complexity of plant morphology. Recent research explores hybrid models, optimization algorithms, and feature fusion to improve performance.Mobile and real-time applications are emerging but often lack continuous learning capabilities. Overall, the literature confirms the effectiveness of CNN-based methods while highlighting the need for larger datasets, better interpretability, and practical deployment strategies.

III. Related Work

Recent years have witnessed a surge in research applying deep learning techniques to the classification and recognition of medicinal plant species, motivated by the need for accurate identification to support biodiversity conservation and traditional medicine. Manual identification by botanists is labor-intensive and often impractical at scale, prompting the adoption of automated, vision-based approaches.

Trends and Approaches: Systematic reviews of the field reveal that the majority of studies focus on supervised deep learning methods, with Convolutional Neural Networks (CNNs) being the most widely used architecture for medicinal plant identification. Transfer learning with pretrained models is prevalent, allowing researchers to leverage existing feature extraction capabilities and adapt them to specific plant datasets. Augmentation and preprocessing techniques are commonly applied to address dataset limitations and improve model robustness. In addition to CNNs, other deep learning models such as Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), and Multilayer Perceptrons (MLPs) have also been explored, with CNNs and MLPs typically achieving the highest accuracy.

Datasets and Features: A significant portion of studies utilize private datasets, with 67.7% of reviewed papers relying on non-public image collections. Most research focuses on leaf images, exploiting features like shape, texture, and color for classification, with over 96% of studies using leaf organs and 74% specifically leveraging leaf shape. The lack of large, publicly available, and geographically diverse datasets remains a notable limitation, hindering the development of universally robust models.

Performance and Innovations: Deep learning models, especially those with optimized CNN architectures and feature fusion techniques, have demonstrated superior performance compared to traditional methods, achieving classification accuracies often exceeding 98%. Innovations such as residual and inverted residual block configurations, Binary Chimp Optimization, and serial feature fusion have further enhanced accuracy and speed. Some studies have also explored mobile and web-based deployment for real-time plant identification, broadening accessibility.

Challenges and Gaps: Despite these advances, key challenges persist, including dataset scarcity, the need for continuous model updating to maintain accuracy, and the complexity of plant morphology across different ages and health statuses. There is also a recognized need for more collaborative efforts to create standardized, open datasets and for research into the interpretability and trustworthiness of deep learning models in botanical applications.

Summary: The literature underscores the transformative potential of deep learning for medicinal plant identification, with CNN-based models leading the field. However, progress is constrained by data limitations and the need for ongoing methodological innovations. Future research should prioritize dataset expansion, model interpretability, and practical deployment strategies to fully realize the benefits of AI-driven plant classification.Table I.

TABLE I: Literature Review

Name	Family	Part used	Preparation/Uses
Aloevera	Asphodelaceae	Leaf gel	Gel extracted and applied or consumed
Amla	Phyllanthaceae	Fruit	Eaten raw, juice, pickles, dried, in Chyawanprash
Amruta Balli	Menispermaceae	Root, stem	Decoction, powder
Arali	Apocynaceae	Leaves, latex	Juice, paste, latex for wounds
Ashoka	Fabaceae	Bark, flowers	Bark decoction for gynecological disorders
Ashwagandha	Solanaceae	Root	Powder, decoction, capsules
Avocado	Lauraceae	Fruit, seed, leaves	Fruit eaten raw, oil, leaf tea
Bamboo	Poaceae	Young shoots, leaves	Shoots cooked, leaves in traditional medicine
Basale	Basellaceae	Leaves, stems	Leaves cooked as vegetable, paste for wounds
Betel	Piperaceae	Leaf	Chewed raw, leaf juice
Betel Nut	Arecaceae	Seed (nut)	Chewed, powder
Brahmi	Plantaginaceae	Whole plant	Juice, powder, decoction
Castor	Euphorbiaceae	Seeds, oil	Oil extracted, used externally and internally
Curry Leaf	Rutaceae	Leaves	Leaves used in cooking, decoction

Doddapatre	Lamiaceae	Leaves	Leaf juice, paste
Ekka	Apocynaceae	Leaves, roots	Paste, decoction
Ganike	Rubiaceae	Leaves, roots	Decoction, paste
Guava	Myrtaceae	Fruit, leaves	Fruit eaten raw, leaf decoction for diarrhea
Geranium	Geraniaceae	Leaves, oil	Oil extracted, leaves in tea
Henna	Lythraceae	Leaves	Paste for dyeing, decoction
Hibiscus	Malvaceae	Flowers, leaves	Flower tea, paste, hair oil
Honge	Fabaceae	Seeds, oil	Oil for lamps, medicinal use
Insulin Plant	Costaceae	Leaves	Leaves chewed or decoction for diabetes
Jasmine	Oleaceae	Flowers, leaves	Flower oil, tea, paste
Lemon	Rutaceae	Fruit, leaves	Juice, rind, leaf tea
Lemon Grass	Poaceae	Leaves	Tea, essential oil
Mango	Anacardiaceae	Fruit, leaves, bark	Fruit eaten raw/ripe, bark/leaf decoction
Mint	Lamiaceae	Leaves	Leaves in tea, paste, essential oil
Nagadali	Apocynaceae	Leaves, latex	Latex for wounds, leaf paste
Neem	Meliaceae	Leaves, bark, seeds	Paste, decoction, oil
Nithyapushpa	Apocynaceae	Leaves, flowers	Paste, decoction
Nooni (Noni)	Rubiaceae	Fruit, leaves	Juice, paste
Pappaya	Caricaceae	Fruit, leaves, latex	Fruit eaten raw/ripe, leaf juice, latex for medicinal use
Pepper	Piperaceae	Fruit (peppercorn)	Dried and ground as spice, decoction
Pomegranate	Lythraceae	Fruit, rind, seeds	Eaten raw, juice, rind decoction
Raktachandini	Fabaceae	Heartwood	Decoction, powder
Rose	Rosaceae	Petals, oil	Rose water, oil, petals in sweets
Sapota	Sapotaceae	Fruit	Eaten raw, latex used medicinally
Tulasi	Lamiaceae	Leaves	Leaves in tea, decoction, paste
Wood Sorel	Oxalidaceae	Leaves	Leaves chewed, juice

IV. Proposed Method

The proposed system leverages the ResNet50 deep convolutional neural network architecture combined with transfer learning for accurate classification of medicinal plants based on leaf images. The methodology includes the following key steps:

Data Collection: The "Indian Medicinal Leaves Image Dataset," containing diverse leaf images from 40 to 80 medicinal plant species, is used to train and evaluate the model. The dataset includes images captured in various environmental conditions and backgrounds to ensure robustness.

Data Preprocessing: To enhance model generalization and handle variability, extensive data augmentation techniques such as random flipping, rotation, scaling, and color jittering are applied. This enriches the dataset with diverse leaf appearances.

Model Architecture: A pretrained ResNet50 model initialized with ImageNet weights is employed as a feature extractor. The early layers are frozen to retain general visual features, while the later layers and fully connected layers are fine-tuned to learn species-specific characteristics. The final fully connected layer is replaced with custom dense layers including dropout and ReLU activation to improve classification performance.

Training: The model is trained using cross-entropy loss and optimized with the Adam optimizer. Learning rate scheduling and early stopping strategies are applied to enhance convergence and prevent overfitting. Training is accelerated using GPU hardware.

Deployment: A web-based interface is developed using Flask, allowing users to upload images of medicinal plants and receive real-time identification results based on the trained ResNet50 model.

This approach automates medicinal plant identification, reducing reliance on manual expertise, and achieves high classification accuracy (reported up to 99.84%), demonstrating its effectiveness and scalability for botanical research and traditional medicine applications.

V. Results and Discussions

The ResNet50-based deep learning model demonstrated outstanding performance in the classification of medicinal plant species. Using the Indian Medicinal Leaves Image Dataset, which contains images from 80 different species, the model achieved a remarkable test accuracy of 99.84%. This result signifies a substantial improvement over traditional and baseline classification techniques, highlighting the model's ability to extract intricate and discriminative features from leaf images.

The high accuracy achieved by the ResNet50 model can be attributed to several factors:

Deep Feature Extraction: The architecture's deep convolutional layers effectively capture complex patterns and subtle differences between plant species.

Transfer Learning: Leveraging pretrained weights and fine-tuning on the medicinal plant dataset enabled the model to adapt quickly and efficiently to the specific classification task.

Robust Data Augmentation: Applying extensive augmentation techniques improved the model's ability to generalize across varied environmental conditions and image qualities.

The results are consistent with other recent studies employing similar architectures. For example, a ResNet50v2-based model achieved 99 percent accuracy in identifying 30 medicinal plant species, further validating the effectiveness of deep learning approaches in this domain. Comparative studies also indicate that ResNet50 outperforms other CNN variants such as MobileNet and Inception v3, with reported accuracies of 91 percent and above on similar datasets.

The discussions emphasize that deep learning models, particularly those based on the ResNet50 family, are highly competent for medicinal plant identification. These models not only provide accurate and reliable results but also offer scalability for real-time applications, such as web-based identification tools, thus bridging the gap between traditional botanical knowledge and modern AI-driven solutions.

VI. Conclusion

This research demonstrates that deep learning techniques, specifically the ResNet50 model, offer a highly effective and scalable solution for medicinal plant classification. By leveraging deep convolutional layers and transfer learning, the model achieves outstanding accuracy in identifying medicinal plant species from leaf images, surpassing traditional methods and baseline models. The approach not only provides precise and real-time identification but also bridges the gap between traditional herbal knowledge and modern technological advancements. This work paves the way for enhanced applications in pharmacology, healthcare, and biodiversity conservation, supporting the accurate, efficient, and accessible utilization of botanical resources. Continued development and integration of such AI-driven systems can further advance the understanding and application of medicinal plants across various domains.

1. Chanyal H, Yadav RK, Saini DKJ. Classifcation of medicinal plants leaves using deep learning technique: a review. Int J Intell Syst Appl Eng. 2022;10(4):78–87.

2. Javid A, Haghirosadat BF. A review of medicinal plants efective in the treatment or apoptosis of cancer cells. Cancer Press J. 2017;3(1):22–6.

3. Barimah KB, Akotia CS. The promotion of traditional medicine as enactment of community psychology in Ghana. J Community Psychol. 2015;43(1):99–106.

4. Rao RU, Lahari MS, Sri KP, Srujana KY, Yaswanth D. Identifcation of medicinal plants using deep learning. Int J Res Appl Sci Eng Technol. 2022;10:306–22.

5. Singh V, Misra AK. Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf Process Agric. 2017;4(1):41–9.

6. Malik OA, Ismail N, Hussein BR, Yahya U. Automated realtime identifcation of medicinal plants species in natural environment using deep learning models—a case study from Borneo Region. Plants. 2022;11(15):1952.

7. Valdez DB, Aliac CJG, Feliscuzo LS. Medicinal plant classifcation using convolutional neural network and transfer learning. In: 2022 IEEE International Conference on Artifcial Intelligence in Engineering and Technology (IICAIET). IEEE; 2022. p. 1–6.

8. Abdollahi J. Identifcation of medicinal plants in ardabil using deep learning: identifcation of medicinal plants using deep learning. In: 2022 27th International Computer Conference, Computer Society of Iran (CSICC). IEEE; 2022. p. 1–6.

9. Sivaranjani C, Kalinathan L, Amutha R, Kathavarayan RS, Kumar KJJ. Real-time identifcation of medicinal plants using machine learning techniques. In: 2019 International Conference on Computational Intelligence in Data Science (ICCIDS). IEEE; 2019. p. 1–4.

10. Zin IAMd, Ibrahim Z, Isa D, Aliman S, Sabri N, Mangshor NNA. Herbal plant recognition using deep convolutional neural network. Bull Electr Eng Inform. 2020;9(5):2198–205.

11. Saikia AP, Hmangaihzuala PVL, Datta S, Gope S, Deb S, Singh KR. Medicinal plant species classifcation using neural network classifer. In: 2021 6th International Conference on Communication and Electronics Systems (ICCES). IEEE; 2021. p.1805–11.

12. Sachar S, Kumar A. Deep ensemble learning for automatic medicinal leaf identifcation. Int J Inf Technol. 2022;14(6):3089–97.

13. Manoharan JS. Flawless detection of herbal plant leaf by machine learning classifer through two stage authentication procedure. J Artif Intell Capsule Netw. 2021;3(2):125–39.

14. https://www.kaggle.com/datasets/vishnuoum/medicinal-plantdataset-augmented?select=data.

15. Saponara S, Elhanashi A. Impact of image resizing on deep learning detectors for training time and model performance. In: Applications in Electronics Pervading Industry, Environment and Society: APPLEPIES 2021. Cham: Springer International Publishing; 2022. p. 10–17.

16. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6(1):1–48.

17. Sae-Lim W, Wettayaprasit W, Aiyarak P. Convolutional neural networks using MobileNet for skin lesion classifcation. In: 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE; 2019 . p. 242–7.

18. Wang W, Hu Y, Zou T, Liu H, Wang J, Wang X. A new image classifcation approach via improved MobileNet models with local receptive feld expansion in shallow layers. Comput Intell Neurosci. 2020.