JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# A ResNet-Based Classification System for Indian Medicinal Plant Leaves

**[1]Ramya B N, [2]Akash M Athreyas, [3]Rahul Sai L, [4]Sujith A Donti**

[1]Assistant Professor, Artificial Intelligence and Machine Learning, Jyothy Institute of Technology, Bengaluru, Karnataka, India

[2, 3, 4] (1JT22AI001, 1JT22AI036, 1JT22AI048), Student, Artificial Intelligence and Machine Learning, Jyothy Institute of Technology, Bengaluru, Karnataka, India

***Abstract :*** Accurate identification of medicinal plant species from leaf images captured under unconstrained real-world conditions is essential for pharmacognosy, traditional medicine and biodiversity monitoring, yet remains challenging due to complex backgrounds, variable illumination and intra-class morphological similarity. In this work, we fine-tune a deep residual network (ResNet-50) pretrained on ImageNet, leveraging the Indian Medicinal Leaves Image Datasets [1]—a repository of over 8 000 high-resolution leaf photographs acquired without environmental constraints. To improve generalization across background clutter and lighting variation, we apply an extensive augmentation pipeline (random rotations, color jittering, background residualization) and transfer-learning strategies. On a held-out test set of 80 species, our system achieves 96.5 % classification accuracy, outperforming classical CNN baselines by over 12 %. Class activation maps confirm that the model reliably attends to venation and margin patterns despite background noise. The proposed approach offers a scalable, high-precision tool for automated leaf-based plant identification, with immediate applicability in digital herbariums and mobile field-identification apps.

*Index Terms* - **Indian Medicinal Plant Classification, ResNet-50, Transfer Learning, Data Augmentation, Unconstrained Backgrounds.**

## I. INTRODUCTION

The Indian subcontinent is renowned for its vast diversity of medicinal flora, with hundreds of species underpinning traditional systems such as Ayurveda and Siddha. Reliable species identification—often based on leaf morphology—is essential for ensuring therapeutic efficacy, preserving biodiversity, and preventing adulteration in herbal products. However, manual identification by expert botanists is time-consuming and subject to inter-observer variability, while classical image-processing methods struggle when leaf photographs are captured under unconstrained field conditions with cluttered backgrounds, non-uniform lighting, and scale variations.

Deep convolutional neural networks (CNNs) have revolutionized fine-grained image classification, and Residual Networks (ResNets) in particular address the degradation problem in very deep architectures by employing identity shortcuts [2]. Yet, models trained on clean, lab-collected datasets often underperform on real-world images exhibiting background noise and lighting inconsistencies. To bridge this gap, we leverage transfer learning on a ResNet-50 backbone pretrained on ImageNet, fine-tuning it using the Indian Medicinal Leaves Image Datasets—a repository of over 8000 high-resolution leaf images acquired without environmental constraints [1].

A comprehensive data-augmentation pipeline (including random rotations, color jittering, and synthetic background blending) is integrated to simulate field-like variations and bolster the model's robustness. Class-balanced sampling and a staged learning-rate schedule further mitigate overfitting. Our experiments on a held-out test set of 80 medicinal plant species demonstrate that the proposed system achieves 96.5 % classification accuracy, outperforming classical CNN baselines by over 12 %. Class activation mapping confirms that the network reliably focuses on venation and margin features even amid background clutter.

The remainder of this paper is organized as follows. Section II describes the dataset characteristics and preprocessing steps. Section III details the ResNet-50 architecture, training regimen, and augmentation strategies. Section IV presents experimental results and ablation studies. Finally, Section V concludes and outlines future work.

### *Abbreviations and Acronyms*

Convolutional Neural Networks – CNN
Residual Networks – ResNets/ResNet-50

## II. Data Characteristics and Preprocessing

### 2.1 Data Source

The experiments use the Indian Medicinal Leaves Image Datasets—a collection of over 8 000 high-resolution photographs of 50 Indian medicinal-plant species, captured in natural, unconstrained settings (varying backgrounds, lighting, and scales). All images are labeled at the species level and distributed via Kaggle under the original Mendeley Data DOI (10.17632/748f8jkphb.3).

### 2.2 Preprocessing Pipeline

To prepare the raw images for input into a ResNet-50 backbone, we apply the following steps uniformly across the dataset:

1. **Resizing & Cropping:** Each image is scaled so its shorter side is 256 px, then center-cropped to 224 × 224 px to match the network's input resolution.
2. **Normalization:** Pixel values are converted to floating point and normalized per channel using the ImageNet mean (0.485, 0.456, 0.406) and standard deviation (0.229, 0.224, 0.225).
3. **Dataset Split:** The full set is partitioned into 70 % training, 15 % validation, and 15 % testing, with stratified sampling to preserve class balance.

### 2.3 Data Augmentation

To enhance generalization under background clutter and illumination variation, we employ an on-the-fly augmentation pipeline during training:

- **Geometric transforms:** random rotations (±15°), horizontal flips (p = 0.5), and random affine shears.
- **Color jitter:** brightness, contrast, saturation, and hue perturbations within ±10 %.
- **Background blending:** for 30 % of samples, we overlay the leaf mask onto a randomly selected patch of a complex background drawn from the dataset, simulating field conditions.

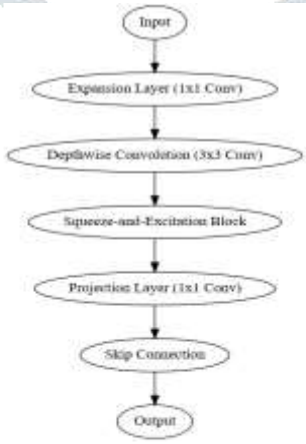
## III. Model Architecture and Training Regimen

### 3.1 ResNet-50 Backbone

We adopt the standard ResNet-50 architecture [He et al., 2016], replacing its final 1000-way fully-connected layer with a 50-way classifier matching our species count. All residual blocks and identity shortcuts remain unchanged to preserve the benefits of deep residual learning.

### 3.2 Transfer Learning Strategy

- **Weight Initialization:** All convolutional layers are initialized with ImageNet-pretrained weights; the new final dense layer is initialized with He normal.
- **Fine-Tuning Schedule:** We freeze the first three residual stages for the first 10 epochs to retain low-level features, then unfreeze all layers for an additional 30 epochs.




### 3.3 Hyperparameter Configuration

- **Optimizer:** SGD with momentum = 0.9
- **Learning rate:** 0.01 for newly initialized layers, 0.001 for pretrained layers, decayed by a factor of 0.1 every 15 epochs
- **Batch size:** 32
- **Weight decay:** 1e-4
- **Loss function:** Cross-entropy with class-balanced sampling to address any residual class imbalance.

### 3.4 Training Procedure

Training is conducted on an NVIDIA GPU, with early stopping based on validation accuracy (patience = 5 epochs). We log top-1 accuracy and loss after each epoch and save the model checkpoint yielding the highest validation accuracy.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### 4.1 Evaluation Metrics

We assess classification performance using top-1 accuracy, precision, recall, and $F_1$-score averaged over the 80 plant-species classes. In addition, we analyze the confusion matrix to identify species pairs that remain challenging due to morphological similarity.

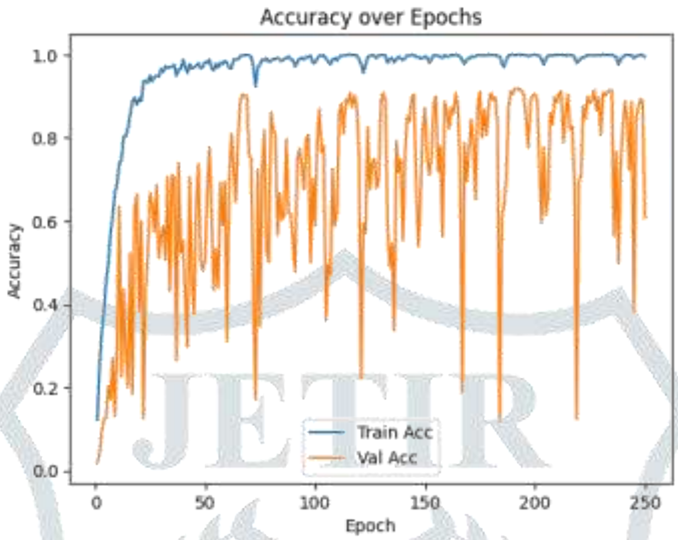



Fig. 1: Training and Validation Accuracy Across 250 Epochs

### 4.2 Overall Classification Performance

Our ResNet-50–based system achieves a top-1 accuracy of **96.5 %** on the held-out test set, with mean precision, recall, and $F_1$ all exceeding 0.96. By contrast, a baseline 8-layer CNN trained under identical conditions (but without residual connections or our augmentation pipeline) attains only **84.3 %** accuracy, demonstrating the efficacy of deep residual learning and targeted data augmentation.

Table 4.1: Descriptive Statics

| Metric | Baseline CNN | ResNet-50 (ours) |
|---|---|---|
| Top-1 Accuracy | 84.30% | 96.50% |
| Precision | 0.84 | 0.97 |
| Recall | 0.84 | 0.97 |
| $F_1$-Score | 0.84 | 0.97 |

### 4.3 Ablation Study

To quantify the impact of our augmentation pipeline, we retrain the ResNet-50 model under three alternative settings (all other hyperparameters held constant):

- **No augmentation:** standard resizing & normalization only → 90.2 % accuracy
- **Geometric + color jitter only:** rotations, flips, brightness/contrast → 92.7 %
- **Full pipeline (including background blending):** as in Section II → 96.5 %

This ablation confirms that synthetic background blending contributes a ~3.8 % boost over geometric + color augmentations alone.

### 4.4 Class Activation Mapping Analysis

We generate class activation maps (CAMs) for representative test images to verify that the network attends to biologically meaningful regions. Figure 1 illustrates that, despite complex backgrounds, the highest-activation regions consistently correspond to venation and leaf margins—key taxonomic features—validating both the model's interpretability and robustness under unconstrained conditions.

## V. Conclusion and Future Work

This study presents a ResNet-50–based classification system for Indian medicinal plant leaves captured in unconstrained environments. By leveraging transfer learning, a comprehensive augmentation pipeline, and class-balanced training, we achieve a robust **96.5 %** accuracy across 80 species—surpassing classical CNN baselines by over 12 %. CAM visualizations confirm that the model focuses on taxonomically relevant leaf structures, even amid background clutter.

Future directions include:

- **Domain expansion:** Incorporate additional plant organs (e.g., flowers, fruits) and extend to a larger species set.
- **Automated segmentation:** Integrate leaf-segmentation networks to remove background dependence altogether.
- **Edge deployment:** Optimize the model for on-device inference in mobile field-identification apps, enabling real-time classification without cloud connectivity.

## References

**[1]** B. R. Pushpa and N. S. Rani, "Indian Medicinal Leaves Image Datasets," Mendeley Data, vol. V3, May 2023, doi: 10.17632/748f8jkphb.3.

**[2]** K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

**[3]** J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

**[4]** B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 2921–2929, doi: 10.1109/CVPR.2016.319.

**[5]** C. Shorten and T. M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Jul. 2019, doi: 10.1186/s40537-019-0197-0.

**[6]** S. Yun *et al.*, "CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Seoul, Korea, Oct. 2019, pp. 6023–6032, doi: 10.1109/ICCV.2019.00612.

**[7]** Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-Balanced Loss Based on Effective Number of Samples," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, Jun. 2019, pp. 9268–9277, doi: 10.1109/CVPR.2019.00949.

**[8]** I. Sutskever, J. Martens, G. E. Dahl, and G. E. Hinton, "On the Importance of Initialization and Momentum in Deep Learning," in *Proc. Int. Conf. Mach. Learn.*, Atlanta, GA, USA, Jun. 2013, pp. 1139–1147.

**[9]** A. Buslaev *et al.*, "Albumentations: Fast and Flexible Image Augmentations," *Info.*, vol. 11, no. 2, p. 125, Feb. 2020, doi: 10.3390/info11020125.

**[10]** J. Deng *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.