International Journal of Science, Strategic Management and Technology
Volume **02** Issue **08** August-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

AI Deep Driven Framework for Automated Identification and Classification of Plant Species from Leaf Images Using Transfer Learning Techniques

Medha V

Department of Computer Science and Engineering, Bangalore, India

https://doi.org/10.55041/ijsmt.v2i8.047

Cite this Article: V, M. (2026). AI Deep Driven Framework for Automated Identification and Classification of Plant Species from Leaf Images Using Transfer Learning Techniques. International Journal of Science, Strategic Management and Technology, *02*(8), 1-9. https://doi.org/10.55041/ijsmt.v2i8.047

Abstract

Accurate plant species identification is important for biodiversity documentation, agriculture, forestry, ecological monitoring, and educational applications. Conventional identification commonly depends on expert knowledge and manual comparison of leaf morphology, which can be slow and difficult when species have similar visual characteristics or when images are captured under uncontrolled field conditions. This paper proposes an AI Deep Driven framework for automated identification and classification of plant species from leaf images using transfer learning techniques. The framework combines image acquisition, preprocessing, leaf-region enhancement, data augmentation, transfer learning, fine-tuning, confidence estimation, and class-wise evaluation in a single pipeline. Pre-trained convolutional neural network architectures such as ResNet50, EfficientNet, and MobileNet are considered as feature-learning backbones, with a task-specific classification head adapted to the target plant-species classes. The use of transfer learning reduces the need to train a deep network entirely from scratch and enables reuse of general visual representations learned from large image collections. The proposed methodology is designed to address variations in illumination, scale, orientation, background, blur, and intra-species leaf appearance. LeafSnap is considered as a suitable benchmark because it contains laboratory and field images covering 185 tree species, including 23,147 laboratory images and 7,719 field images. The framework evaluates performance using accuracy, precision, recall, macro F1-score, confusion matrix, and inference time. This paper presents the architecture, methodology, evaluation protocol, limitations, and future scope of the proposed system. Numerical experimental results are intentionally not fabricated; they should be populated after the proposed models are trained and tested on the selected dataset.

Keywords: Plant Species Identification, Leaf Image Classification, Deep Learning, Transfer Learning, Convolutional Neural Network, ResNet50, EfficientNet, MobileNet, Computer Vision, Biodiversity.

Introduction

Plant identification is a fundamental task in agriculture, forestry, ecological surveys, biodiversity conservation, and botanical education. A plant species can often be recognized from characteristics such as leaf shape, venation, margin, texture, color, and arrangement. However, manual identification requires botanical expertise and may become difficult when multiple species exhibit similar leaf structures. Field conditions introduce additional challenges because images may contain complex backgrounds, shadows, blur, occlusion, different camera viewpoints, and changes in illumination.

Computer vision provides an opportunity to automate this task by converting a leaf photograph into a set of discriminative visual features and mapping those features to a species label. Earlier recognition systems relied heavily on handcrafted descriptors and shape measurements. LeafSnap demonstrated an important practical direction by using visual recognition to identify tree species from leaf photographs. Its released dataset contains 23,147 laboratory images and 7,719 field images covering 185 tree species, making it a useful benchmark for research on leaf-based species recognition.

Deep convolutional neural networks (CNNs) have substantially improved image classification by learning hierarchical visual representations directly from images. Nevertheless, training a deep model from random initialization generally requires large amounts of labeled target-domain data and substantial computational resources. Transfer learning addresses this problem by starting from a network pre-trained on a large image dataset and adapting its learned representation to a new classification task. Research on feature transfer has shown that lower-level visual features can be reused across related image tasks, although transferability becomes more task-dependent in deeper layers.

The proposed work therefore focuses on an integrated transfer-learning framework rather than a single classifier. Multiple modern CNN backbones are considered so that accuracy, computational cost, and deployment suitability can be compared. The framework is intended to support both laboratory-style images and more challenging field photographs, with explicit preprocessing and augmentation stages designed to reduce sensitivity to acquisition conditions.

Literature Review

LeafSnap established an early computer-vision pipeline for automatic plant species identification and showed that leaf shape and contour information can support species-level recognition. The work also highlighted an important characteristic of the problem: inter-species differences may be small while intra-species variation can be substantial. This observation motivates robust feature learning rather than reliance on a single handcrafted shape descriptor.

Mohanty, Hughes, and Salathé demonstrated the feasibility of deep learning for plant-leaf image analysis using the PlantVillage dataset. Their study classified crop species and disease conditions and reported very high performance on controlled held-out data, while also showing a substantial reduction in accuracy on images collected under different conditions. This finding is especially relevant to the proposed framework because a species classifier must generalize beyond clean laboratory images.

Yosinski et al. investigated the transferability of learned CNN features and showed that early layers tend to learn more general visual patterns while higher layers become increasingly task-specific. This supports a staged transfer-learning strategy in which a pre-trained backbone is initially frozen and later selected layers are fine-tuned for the target plant dataset.

ResNet introduced residual learning to make deeper networks easier to optimize, while EfficientNet proposed balanced scaling of depth, width, and image resolution. MobileNet introduced depth-wise separable convolutions to reduce computational cost for mobile and embedded vision applications. These architectures provide complementary choices for the proposed framework: ResNet emphasizes strong residual feature learning, EfficientNet emphasizes accuracy-efficiency trade-offs, and MobileNet emphasizes lightweight inference.

The literature indicates that the major challenge is not simply achieving high accuracy on a controlled dataset, but maintaining reliable recognition under field variability. Consequently, the proposed framework includes separate evaluation of laboratory and field-like images, class-wise metrics, confidence analysis, and error inspection.

Methodology

The proposed framework follows a sequence of stages: image acquisition, dataset organization, preprocessing, augmentation, transfer-learning model construction, training, fine-tuning, validation, prediction, and performance analysis. The architecture is designed so that the backbone can be replaced without changing the rest of the pipeline.

Dataset Selection and Organization: LeafSnap is selected as a primary benchmark because its official dataset contains 30,866 images from 185 tree species, comprising 23,147 laboratory images and 7,719 field images. For a reproducible experiment, class labels should be mapped to scientific species names, corrupted or unusable images should be removed according to a documented rule, and train/validation/test partitions should be created without allowing near-duplicate images from the same capture sequence to leak across partitions.

Image Preprocessing: Each input image is resized to the input resolution required by the selected backbone. Pixel values are normalized using the preprocessing convention of the corresponding pre-trained model. Where appropriate, background reduction or leaf segmentation may be applied to emphasize the leaf region. Preprocessing should not remove biologically meaningful characteristics such as serration, venation, texture, or margin shape.

Data Augmentation: To improve robustness, training images can be augmented using controlled rotation, horizontal or vertical translation, zoom, mild brightness variation, contrast adjustment, and horizontal flipping where botanically acceptable. Excessive transformations should be avoided because they can create unrealistic leaf morphology. Validation and test images should not receive random augmentation.

Transfer Learning Models: Three complementary backbones are proposed. ResNet50 provides residual connections that support deep feature learning. EfficientNet provides a compound scaling strategy that balances depth, width, and resolution. MobileNet is included as a lightweight alternative for mobile or edge deployment. In the initial stage, the convolutional backbone is frozen and only the new classification head is trained. In the second stage, selected upper layers are unfrozen and fine-tuned using a smaller learning rate.

Classification Head: The extracted feature tensor is passed through global average pooling followed by dropout and a dense layer. A softmax output produces a probability distribution over the target species. The predicted class is the class with the highest probability, while the probability value is retained as a confidence indicator. An optional rejection threshold can be used so that low-confidence predictions are flagged for human verification instead of being presented as certain identifications.

Training Strategy: Categorical cross-entropy is used for multi-class classification. Adam or another adaptive optimizer may be used for the initial transfer stage, followed by a lower learning rate during fine-tuning. Early stopping and learning-rate reduction can be used to limit overfitting. Model checkpoints should be selected using validation macro F1-score rather than accuracy alone when the class distribution is imbalanced.

Evaluation Metrics: Accuracy measures the overall proportion of correctly classified images. Precision, recall, and F1-score should be computed per class and summarized using macro averages. A confusion matrix is used to identify species pairs that are systematically confused. Top-k accuracy is useful when the application presents several candidate species. Inference time, model size, and memory use should also be recorded when deployment on mobile or edge hardware is a goal.

System Architecture

The framework contains five logical layers. The acquisition layer receives a leaf image from a camera, mobile device, or dataset. The preprocessing layer performs resizing, normalization, quality checks, and optional segmentation. The representation layer uses a pre-trained CNN backbone to extract hierarchical visual features. The classification layer maps the features to plant-species probabilities. Finally, the decision layer returns the top prediction, confidence score, and optional top-k alternatives.

A modular architecture makes it possible to compare different transfer-learning backbones using the same data split and evaluation procedure. This reduces experimental bias and allows the final system to select a model based not only on accuracy but also on inference cost and robustness.

Results and Discussion

Because no trained model, experimental log, or measured test-set output was supplied with the source material, this paper does not claim fabricated numerical results. Instead, the results section defines the evaluation procedure and the observations that should be reported after execution. This approach preserves

International Journal of Science, Strategic Management and Technology
Volume 02 Issue 08 August-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

scientific reproducibility and prevents unsupported accuracy values from being presented as experimental evidence.

The first comparison should evaluate the three transfer-learning backbones under identical preprocessing, data splits, augmentation policy, optimizer family, and stopping criteria. The resulting table should contain test accuracy, macro precision, macro recall, macro F1-score, top-5 accuracy, parameter count, and average inference time. The best model should be selected using a multi-objective view rather than accuracy alone.

The second analysis should compare laboratory and field images. A model that performs strongly on controlled images but degrades substantially on field images may have learned background or acquisition-specific cues rather than species-specific characteristics. This issue is consistent with earlier plant-image studies showing that domain shift can significantly affect generalization. Therefore, field-oriented testing is a central component of the proposed framework.

Error analysis should focus on species pairs with similar morphology. Confusion matrices can reveal whether errors arise from similar leaf margins, venation, shape, color, or texture. Misclassified images should be reviewed together with confidence scores. High-confidence errors are particularly important because they indicate cases in which the model may be confidently relying on non-discriminative visual cues.

For deployment, MobileNet may be preferable when latency, memory, or mobile execution is more important than the last increment of classification accuracy. EfficientNet can be evaluated when a balanced accuracy-efficiency trade-off is required, while ResNet50 provides a strong conventional residual-learning baseline. These are design considerations; the final choice must be determined by measured results on the selected dataset.

Advantages and Limitations

Advantages of the proposed framework include reuse of pre-trained visual representations, a modular backbone design, explicit handling of image variability, class-wise evaluation, and the possibility of lightweight deployment. The framework can also be extended to include additional plant attributes such as disease status, growth stage, or geographic information.

Limitations include dependence on the coverage and quality of the training dataset, possible class imbalance, sensitivity to species that have visually similar leaves, and domain shift between controlled and field images. Leaf-only identification also has biological limitations because some species are difficult to distinguish from a single leaf photograph. A reliable real-world application should therefore communicate uncertainty and, where necessary, request additional evidence such as flowers, fruits, bark, or multiple leaf images.

Conclusion and Future Scope

This paper presented an AI Deep Driven framework for automated plant species identification from leaf images using transfer learning. The proposed pipeline combines preprocessing, augmentation, pre-trained CNN feature extraction, fine-tuning, confidence estimation, and comprehensive evaluation. ResNet50, EfficientNet, and MobileNet are proposed as complementary backbone choices, enabling comparison of recognition performance with computational requirements.

The framework is designed to move beyond a controlled-image demonstration by explicitly considering field variability and class-wise errors. The use of a benchmark such as LeafSnap provides a reproducible basis for experimentation, while transfer learning reduces the need for training a deep network entirely from scratch. The final numerical performance should be obtained by executing the training and evaluation protocol rather than being assumed from published results on different tasks.

Future work can include field-specific augmentation, self-supervised pretraining on unlabeled plant images, vision-transformer backbones, explainable AI techniques such as class-activation maps, calibration of prediction confidence, continual learning for newly observed species, and deployment as a mobile application. A broader regional dataset containing Indian plant species would also improve the practical relevance of the system for local biodiversity and agricultural applications.

References

[1] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. B. Soares, "Leafsnap: A Computer Vision System for Automatic Plant Species Identification," in Proc. 12th European Conference on Computer Vision (ECCV), 2012.

[2] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016, doi: 10.3389/fpls.2016.01419.

[3] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in Advances in Neural Information Processing Systems 27, 2014, pp. 3320–3328.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[5] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. 36th International Conference on Machine Learning (ICML), 2019, pp. 6105–6114.

[6] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv:1704.04861, 2017.

[7] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556, 2014.

[8] LeafSnap, "Leafsnap Dataset," Columbia University, University of Maryland, and Smithsonian Institution. Dataset documentation and image collection.

[9] M. S. S. S. et al., "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review," systematic review of leaf-image recognition datasets and methods, 2018.

[10] S. Zhang, C. Zhang, and Z. Wang, "Plant Leaf Recognition Based on Deep Learning and Image Processing," representative work illustrating CNN-based plant recognition approaches.