LeafIQ: A Hybrid Deep Learning Framework for Medicinal Leaf Classification Using EfficientNet-B1 and GLCM Texture Feature

Dr. R. Jayamala[1], Jemuna U[2], Aishwarya K S[3]

Abstract

Traditional health care systems such as Ayurveda, Siddha, and Unani need the identification of medicinal leaves. However, manual recognition is challenging due to similarity between species. Most existing work depends on traditional Convolution Neural Network (CNNs) and handcrafted features which perform poor on noised image and similar leaves. In this paper, a hybrid deep learning model called LeafIQ (Leaf Intelligence Quotient) using EfficientNet-B1 and advanced Gray Level Co-occurrence Matrix (GLCM). EfficientNet-B1 model is used for image classification and feature extraction. GLCM is the texture feature extraction technique for accurate extraction of texture. The deep feature vector and the texture feature vector is fused for accurate classification. Gradient-weighted Class Activation Mapping (Grad-CAM) is an Explainable Artificial Intelligence (XAI) used for visually represent the important part of the leaf image. By analysing the images of 30 different plant species, the results show that combined deep features and texture features helps to identify the leaves better. As a result, this framework gives accuracy of 99.82% and reliable classification. LeafIQ is deployed as a Flask web application with a professional chat-style interface supporting Gemini AI integration, YouTube plant search, Tamil–English bilingual output, and Grad-CAM visual explanations.

[1]Department of Computer Science and Engineering, [2,3]Department of Information Technology
[1,2,3]University College of Engineering, BIT Campus, Anna University, Trichy, India
[1]jayamala@aubit.edu.in, [2]jemunau@gmail.com, aishwaryaselvaraju14@gmail.com[3]
*Corresponding author
DOI: 10.70102/IJARES/V6I2/6-2-05

Keywords: Medicinal leaf classification, Deep learning, EfficientNet-B1, GLCM (Gray Level Co-occurrence Matrix), Grad-CAM (Gradient-weighted Class Activation Mapping), Attention fusion, Transfer learning and Data Augmentation

Introduction

Over centuries, medicinal plants act as the backbone of natural medicines across the world. Traditional medicine includes system like Ayurveda, Siddha, and Unani still rely on plant-based preparations for treating a broad range of illness. Given their long-lasting clinical references, accurately identifying the medicinal plants species remains a struggle for researchers, botanist and pharmaceutical scientists alike. These plants are foundational to herbal medicines, beauty products, household remedies and modern drug discovery. Yet classification between species is far from straightforward, many plants have similar leaf shapes, surface textures, vein structures, and color profiles making manual identification of species is both time consuming and prone to error [1]-[3].

The rapid growth of the computer vision and Artificial Intelligence (AI) sector makes the classification of medicinal leaves an easy process. Researchers are focusing on the Deep Learning (DL) and Convolutional Neural Network (CNN) for the automatic recognition of the Medicinal Plants in a fast and reliable manner [4]-[6]. These data driven techniques allow the models to learn complex visual patterns from a large set of medicinal leaf images and generalize it across specific leaf boundaries. Along with the complete learned features, handcrafted descriptors especially Gray Level Co-occurrence Matrix (GLCM) texture features have proven values in seizing fine surface information that complements the feature learned by deep networks [7].

Multiple studies on the comparative study covers a broad range of CNN architectures for leaf classification, including alexnet [17], vggnet [16], resnet [15], googlenet [18], densenet [19], and various optimized variants [8]-[11]. Most of the recent investigation have extended further, incorporating ensemble strategies, knowledge distillation, and model compression to maintain the accuracy along with computational efficiency [12], [13]. Among present-day architectures, efficientnet is specialized for compound scaling across width, depth, and resolution, its design philosophy yields strong performance without unnecessary code bloat [14]. Supplementary enhancement from squeeze and excitation mechanisms [20] have further improved the ability of deep networks to model channel-wise feature interconnectedness.

Beyond learned depictions, classical texture analysis through Haralick GLCM features [21] offers computationally inexpensive plan to capture the repeating surface pattern in leaf image. AI transparency has also become a growing issue in applied deep learning, and gradient-weighted class activation mapping (Grad-CAM) [22] has appeared as a practical tool for visualizing which spatial regions of the input is responsible for the output. When trust and transparency are particularly valuable Grad-CAM is required. Across a wide range of plant related task from disease detection [23] to species identification [24]-[26], deep learning has proven to be effective. Large, standardized datasets such as imagenet [27] enables this progress by playing significant role.

Moreover, advanced feature extraction efficiency has been increased with Depth wise separable convolution-based architectures like Xception [28]. Data augmentation techniques [29] are now usual procedure for improving generalization from limited samples. Meanwhile residual connections [30], dense connectivity [19], and normalization strategies [34] have together improved the training stability and final accuracy. Optimization methods like Adam [36] are used for training the deep networks effectively. Leaf based plant identification systems ranging from earlier approach like shape descriptor [31], [32] to systematic surveys [33] on deep learning to illustrate the curriculum sequence of this field toward more practical and ready-to-go solutions.

Driven by these developments, this work introduces leafiq - an intelligent framework for medicinal leaf classification using feature-enhanced Deep Learning that fuses efficientnet-B1 deep features with the Multiscale feature-enhanced textural features to improve the recognition of the medicinal plant leaves. This system achieves 99.82% accuracy by integrating data augmentation to stabilize inference, Grad-CAM visualization, youtube integration for reference videos and chat with Gemini AI section for querying the doubts.

Related work and comparitive study

Recent years deep learning is applied for solving the problem of medicinal leaf recognition. Gautam et al. [1] proposed a lightweight feature fusion approach that uniting learned and handcrafted representations, illustrating this hybrid method improves accuracy over other approaches. Bouakkaz et al. [2] recognized optimized CNN architectures and high accuracy are achieved at high training overhead cost. Further scope was increased by Dalvi et al. [3] with multi attribute with carefully designed deep CNN from plant classification to skin disease applications, which can serve dual purpose across relevant tasks.

Other notable contributions stretch synthetic data generation [4], architectural comparison studies [8], Xception based feature extraction [10], and CNN based classification for regional plant datasets [11]. Integrating domain specific preprocessing including log Gabor filter analysis [13] by researchers to sharpen inter species discrimination. Together, these works finalize that medicinal plant classification enhances when strong feature extractions are mapped with domain aware preprocessing. However, limitations in computational cost, dataset diversity and real-world robustness remains constant across many approaches [15], [16]. The present works target these gaps directly by mapping efficientnet backbone with the handcrafted textural features from GLCM and augmentation strategy designed to increase resilience under real time implementation conditions.

Table. 1. Comparison of Existing Systems

Systems	Methods	Limitations
Plantnet [31]	Multi organ images in CNN	~ 4% accuracy on leaf only
Leafsnap [32]	Shape descriptors and Support Vector Machine (SVM)	Fails on real world conditions
Plantvillage CNN [23]	Alexnet for disease classification	Accuracy occurs falls outside the controlled conditions
Inception-v3 [18]	Fine-tuning using imagenet	No textural features and prone to inter species confusion
Resnet-50 and GLCM [15], [21]	36-dimension GLCM vector with resnet	No attention mechanism and not designed for deployment
Efficientnet-B0 [14]	Compound Scaling of width, depth and resolution	Overconfident predictions on unknown species
Vision Transformer (vit) with leaf morphology [27]	Compute Optimal Scaling	High computational costs and there is no real time web interface

Table 1. Presents a comparative analysis of existing medicinal leaf classification approaches, highlighting the methods employed and their key limitations. It indicates that most approaches suffer from issues such as poor real-world performance, inadequate texture representation, high computational cost, and limited deployment capability.

A deeper examination of earlier approaches directs to the three recurring weakness that restricts their practical application for medicinal leaf classification. First, CNN based approaches such as plantnet, Inception-v3 and efficientnet-B0 depends entirely on the appearance-based features and faces challenges in distinguishing texture patterns which are important for separating visually alike medicinal species. Second, hybrid architectures such as resnet-50 augmented with GLCM descriptors address this partially through the feature concatenation, which basically lack attention mechanism which is capable of weighting the deep visual features versus statistical descriptors in an organised way. Third, there is no existing system reviewed above incorporates meaningful uncertainty estimation, efficientnet-B0 is a proper example producing softmax outputs that provides no algorithm to flag out-of-distribution inputs.

Data collection and augumentation

The proposed system uses medicinal plant leaves dataset containing 3670 images from 30 different medicinal plant species shown in Fig. 1, the images are collected from Mendeley dataset [39]. The images are taken in various light condition, background, leaf position, orientation, which makes the classification accurate. Leaf images are taken by the Samsung S9+ mobile camera, and the result of the leaves is visualized in Canon Inkjet Printer. In the context of dataset size is small for performing classification the data augmentation technique such as rotating, scaling, flipping is preformed to increase the size of the dataset for better classification. By using the diverse dataset, the overfitting issued is reduced. The dataset is split into 70:15:15 ratio for training, testing, and validation.

Fig. 1. Sample images from Dataset

System architecture

The proposed leafiq framework is a hybrid feature enhanced medicinal leaf classification system. The system combines deep semantic representation with handcrafted textural descriptors for the accurate and robust and accurate classification of medicinal leaves. Initially, the medicinal leaf image is resized into 240 x 240 pixels with the Eq. (1)

$$I_r = \text{Resize}(I_e, 240 \times 240) \tag{1}$$

Where, I_e Represents the enhanced image and I_r Denotes the resized image.
Then Normalized before passing into the feature extraction pipeline with the Eq. (2)

$$I_n = \frac{I_r - \mu}{\sigma} \tag{2}$$

Where, I_r Is the resized image used for normalization, μ represents the average pixel value, and σ denotes the variation of pixel values (standard deviation).
The system processes the input in two parallel approaches: a deep learning branch and texture analysis branch.

In the deep learning branch, the leaf image is fed as input to the efficientnet-B1 model for extracting high level deep semantic representations. The efficientnet utilizes compound scaling to balance the depth, width and input resolution of the network and maintains lower computational cost. Efficientnet-B1 also uses the Swish activation function, which is a non-linear function defined in Eq. (3)

$$\text{Swish}(x) = x \cdot \sigma(x) = \frac{x}{1+e^{-x}} \tag{3}$$

where, x is the input value and $\sigma(x)$is the sigmoid function

Unlike traditional activation functions like relu, Swish does not completely block negative values; instead, it allows small negative values to pass through. This helps the network learn smoother gradients and more complex patterns, improving overall performance.

The proposed architecture as shown in Fig. 2. Utilizes multiple Mobile Inverted Bottleneck Convolution (mbconv) blocks to capture hierarchical visual information such as leaf edges, venation structures, contours, color distribution, and morphological patterns. Global Average Pooling (GAP) is applied to generate a compact 1280-dimensional deep feature vector following the convolution layers.

Similarly, the texture analysis branch extracts statistical texture descriptors using Gray Level Co-occurrence (GLCM) analysis. The input medicinal leaf image is converted into grayscale, and to capture the spatial relationships between the neighbouring pixel intensities, GLCM matrices are computed across multiple orientations and distances. From these generated co-occurrence matrices several textural properties are obtained, including contrast, dissimilarity, homogeneity, energy, correlation, and entropy. These descriptors provide extra information that complements the deep features extracted by the CNN model.

Contrast measures the intensity difference between neighbouring pixels in the image. It indicates how much variation or local contrast exists in the texture. Eq. (4)

$$\text{Contrast } (f_1) = \sum_{i,j}(i-j)^2\, P(i,j) \qquad (4)$$

Where, i and j are pixel intensity values and $P(i,j)$ is the probability of their co-occurrence. Higher contrast values indicate greater variation in intensity.

Dissimilarity measures the absolute difference between pixel pairs and represents how different the neighboring pixels are Eq. (5)

$$\text{Dissimilarity}(f_2) = \sum_{i,j} \mid i-j \mid \; P(i,j) \qquad (5)$$

A higher value indicates more variation and less similarity between pixels.

Homogeneity measures how close the distribution of elements in the GLCM is to its diagonal. It indicates the smoothness of the image. Eq. (6)

$$\text{Homogeneity}(f_3) = \sum_{i,j} \frac{P(i,j)}{1+(i-j)^2} \qquad (6)$$

Energy measures the uniformity or repetition of pixel pairs in the image. A higher energy value indicates more uniform and consistent texture patterns. Eq. (7)

$$\text{Energy}(f_4) = \sum_{i,j} P(i,j)^2 \qquad (7)$$

Correlation measures how correlated a pixel is to its neighbour across the image. It reflects the linear relationship between pixel intensities. Eq. (8)

$$\text{Correlation}(f_5) = \sum_{i,j} \frac{(i-\mu)(j-\mu)}{\sigma^2}\, P(i,j) \qquad (8)$$

Where, μ is the mean intensity and σ is the standard deviation. Higher values indicate strong relationships between neighbouring pixels. Higher homogeneity values indicate smoother textures with less variation.

Entropy measures the randomness or complexity of the texture in the image. Higher entropy values indicate more complex and less predictable textures. Eq. (9)

$$\text{Entropy}(f_6) = -\sum_{i,j} P(i,j)\log\left(P(i,j)\right) \qquad (9)$$

After extracting the features, normalization (F_n) is applied to scale them into a common range as given in Eq. (10)

$$F_n = \frac{F-F_{min}}{F_{max}-F_{min}} \qquad (10)$$

Where, F is the original feature value, F_{min} Is the minimum value, and F_{max} Is the maximum value. This step ensures that all features are on the same scale, improving the performance of the classification model.
Finally, the texture features are combined into a single vector as represented in Eq. (11):

$$F_{texture} = [f_1, f_2, f_3, f_4, f_5, f_6] \qquad (11)$$

This 6-dimensional vector represents the texture characteristics of the leaf image and is passed to the feature fusion stage. This aims to extract finer details pertaining to texture. Vein structures, surface roughness, among others, can be obtained through this process. Texture features help in understanding more information about the microstructures of the leaves.

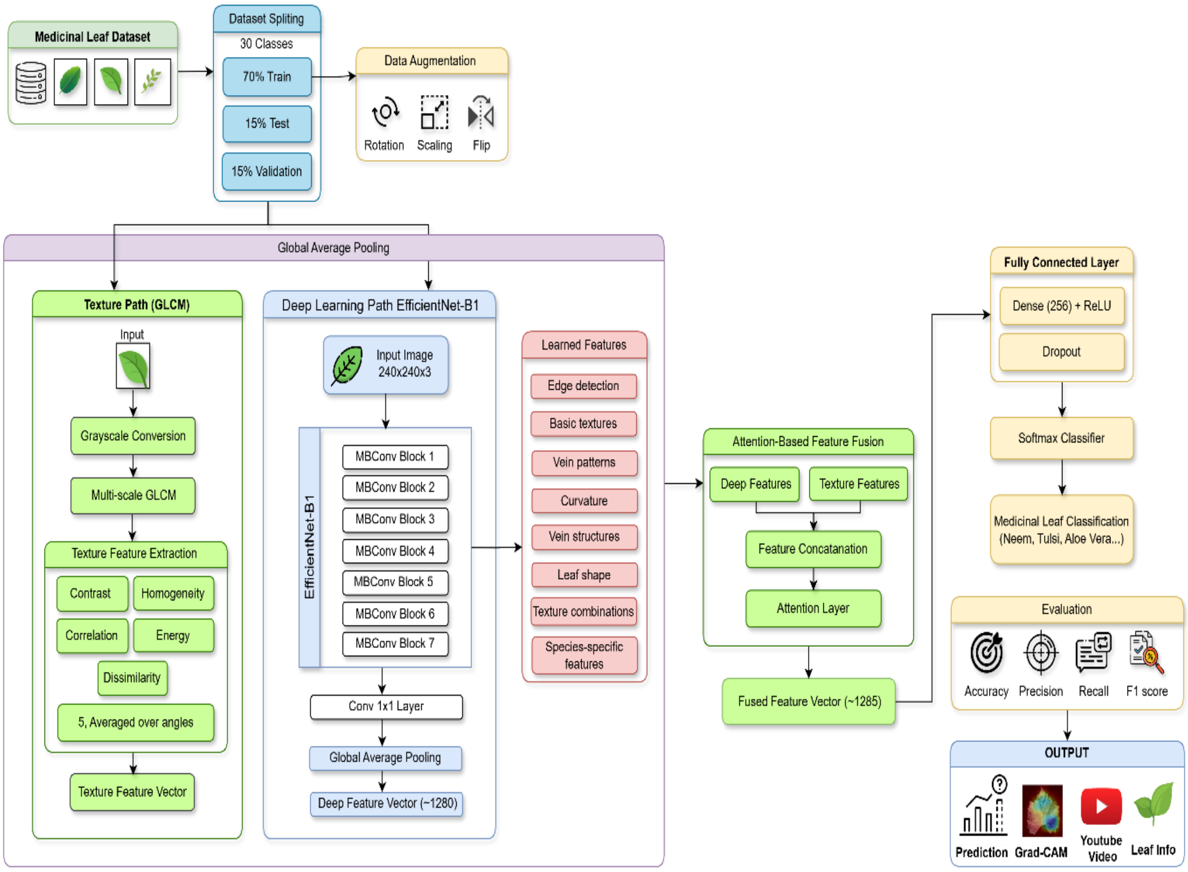

Fig. 2. System Architecture of the leafiq

The 1280-dimensional deep features vector and the 6 GLCM texture features vector is fused, and the 1285-dimensional vector is fed into the fully connected layer to perform classification. The softmax classifier is used in final layers, the class which has the highest weight is the output of the classification.

Test Time Augmentation (TTA) is incorporated during inferenced to enhance the prediction reliability. Grad-CAM visualization is integrated into the framework for highlighting the important image regions making the decision process easier to understand and improve the user confidence in the system. Bilingual language support, Chatbot integration and youtube recommendations to enhance the user experience.

The complete leafiq architecture combines semantic, structural and texture-based information to achieve more reliable medicinal leaf classification in practical real-world environments.

Methodology

The proposed methodology as described in Algorithm 1 follows a multi-stage hybrid learning process designed to improve medicinal leaf classification by combining deep and textural features to improve the medicinal leaf classification. The workflow includes preprocessing, deep feature extraction, texture feature computation, feature fusion, classification and visualization.

Algorithm 1: Proposed Hybrid Medicinal Leaf Classification Framework

Input: Medicinal leaf image (I)
Output: Predicted class (Y_{pred})

BEGIN
1. Preprocess image:
    Iout ← Resize (I, 240×240)

```
     Iout ← Normalize (Iout)
     Apply data augmentation
2. Extract deep features:
     Fdeep ← efficientnet-B1(Iout)
3. Extract texture features:
     Convert Iout to grayscale
     Compute GLCM
     FGLCM ← {Contrast, Dissimilarity,
            Homogeneity, Energy,
            Correlation, Entropy}
4. Perform feature fusion:
     Ffusion ← Fdeep ⊕ FGLCM
5. Perform Classification:
     P ← Softmax (Ffusion)
     Ypred ← argmax(P)
6. Confidence verification:
     IF max(P) < 0.65 THEN
         Reject Prediction
     END IF
7. Test-Time Augmentation:
     Generate transformed samples
     Average prediction score
8. Generate Grad-CAM explanation
9. Return Ypred
END
```

Initially, the medicinal leaf image undergoes processing steps such as resizing, normalization, and data augmentation to enhance the image quality and increase the dataset variability. Augmentation techniques such as rotation, flipping, scaling, and brightness adjustment to reduce overfitting and improves the model's ability to generalize under different climatic conditions.

The processed images are then provided as input to the efficientnet-B1 which extracts the deep visual features and related to leaf shape, structural patterns, edges and venation details.

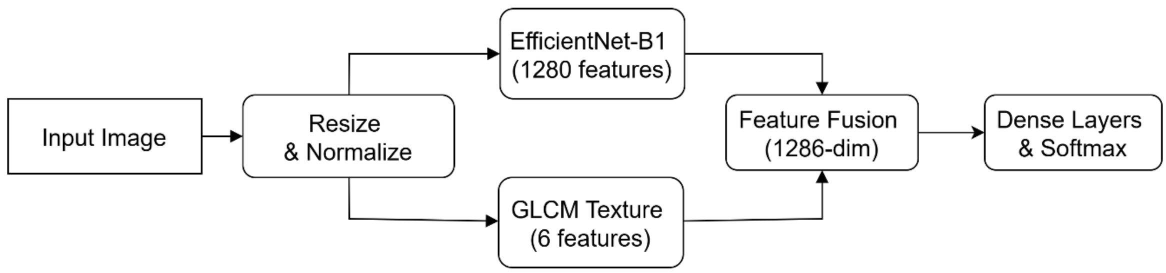

Fig. 3. Feature fusion Model

The processed medicinal leaf images undergo Gray scale conversion. Gray Level Co-occurrence Matrix (GLCM) of the processed medicinal leaf images are used to obtain statistical texture information. These textural descriptors capture subtle surface characteristics that are useful for identifying the medicinal plant species with similar visual appearances.

$$F_{fusion} = [F_{deep} \oplus F_{GLCM}] \tag{12}$$

Where, F_{fusion} Represents the final hybrid fused feature representation used for classification. F_{deep}Represents the deep feature vector extracted using efficientnet-B1, F_{GLCM}Denotes the handcrafted texture descriptor vector and $\oplus$ indicates the concatenation operation.

The extracted deep and texture features are concatenated as shown in Eq. (12). To form a fused feature vector. This fusion enables the model to capture both global semantic information and fine-grained texture details.
The refined feature vector is passed through fully connected layers with relu activation, followed by dropout regularization to prevent overfitting. Finally, a Softmax classifier is used to compute class probabilities as formulated in Eq. (13). And the class with the highest probability is selected as the predicted label and confidence score less than 0.65 is rejected. The proposed model is designed to classify medicinal leaf images into multiple categories.

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \tag{13}$$

Where, $P(y_i)$denotes the probability of the input medicinal leaf image belonging to class i, e denotes Euler's exponential constant, z_iRepresents the output score (logit) corresponding to class i, N represents the total number of medicinal leaf classes, j denotes the class index used in the normalization process and $\sum_{j=1}^{N} e^{z_j}$Represents the sum of exponential scores of all classes.

The Test-Time Augmentation as formulated in Eq. (14) method is used here to generate multiple images for better classification.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t\,(x) \tag{14}$$

Where, $\hat{y}$ Represents the final prediction after Test-Time Augmentation, T denotes the total number of augmented samples, $f_t(x)$ represents the prediction obtained from the t^{th} Augmented image, t denotes the augmentation index and x represents the input image.

To create the system transparent, the Grad-CAM visualization technique is used to make classification better. Youtube video recommendation for the identified medicinal leaf to enhance the user experience. The chatbot module was integrated using the Gemini API to provide the natural language interaction with users [40]. Bilingual Language support to make it accessible for Tamil users as well.

Experiment results at evaluation metrics

Evaluation metrics is the measure of model's performance and there are formulated. Accuracy means the overall correctness of the model, its show the proportion of correctly classified leaf image of all prediction. Precision (P_k) shows the actual correctness of positive prediction. Recall (R_k) means the ability to identify all the actual positive prediction of the model. F1 Score ($f1_k$) is the harmonic mean provides the balance between the precision and recall. Macro F1 Score is the average of F1 scores of all k classes (30 plant species), which give equal importance to all the classes.

The experimental analysis shows the comparison of Traditional and Proposed leafiq as shown in Table. 2.

Table. 2. Comparison table of Traditional and Proposed leafiq value

Model	**Train Accuracy (%)**	**Test Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Alexnet	42.00	75.59	81.53	75.49	74.42
VGG16	35.00	49.50	46.08	49.50	41.34
Resnet50	55.00	74.05	83.84	74.05	72.73
Googlenet	38.00	48.09	64.01	48.09	41.65
Densenet121	36.00	61.52	71.92	61.52	58.51
Efficientnet-B0	50.00	77.59	81.68	77.59	75.85

Proposed leafiq	**98.00**	**99.64**	**100.00**	**99.64**	**99.81**

The comparative analysis of deep learning models as shown in Figs. 4-7 highlights a notable performance improvement by the proposed leafiq system. Unlike other traditional models such as alexnet, VGG16, resnet50, googlenet, densenet121, which shows divergence between training and testing accuracy, the leafiq system shows stable and reliable performance.

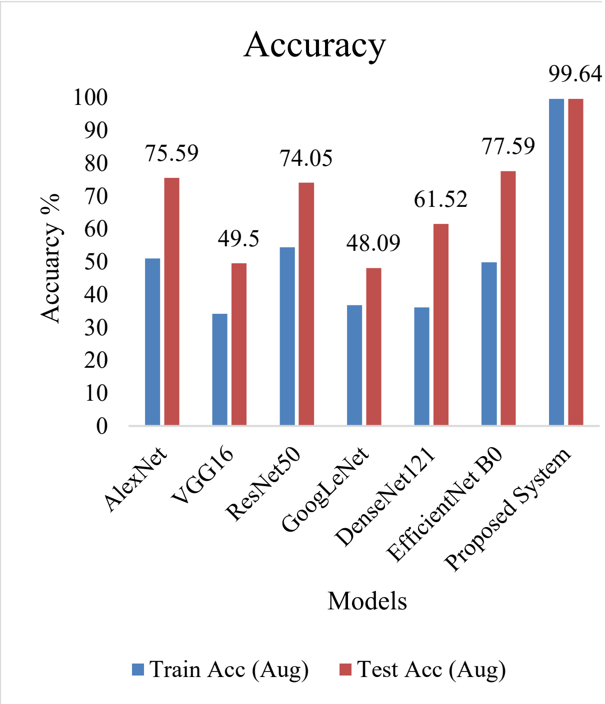

Fig. 4. Accuracy

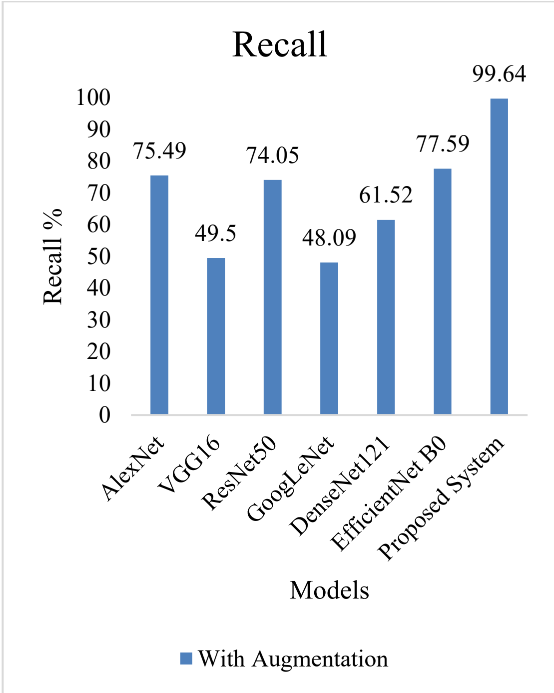

Fig. 6. Recall

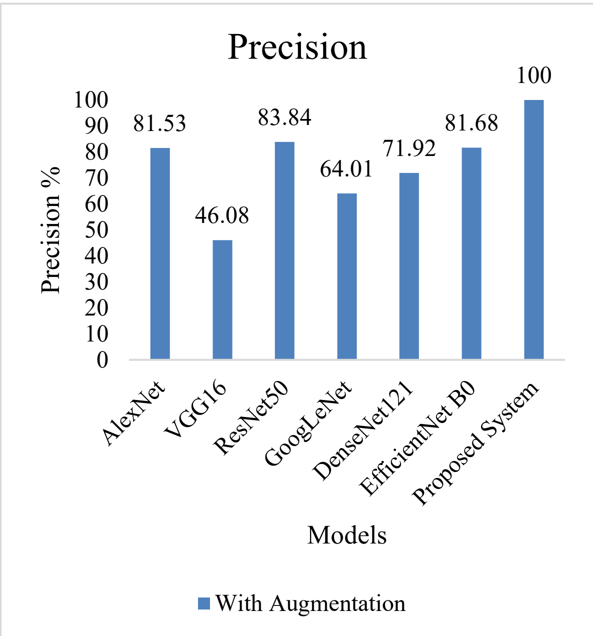

Fig. 5. Precision

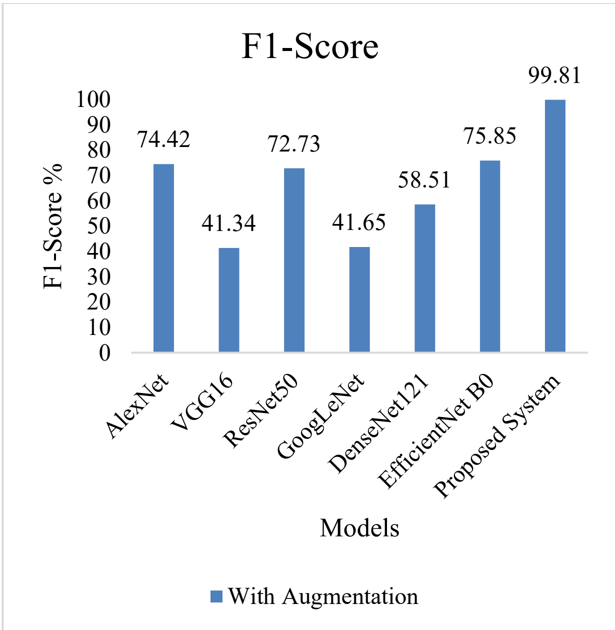

Fig. 7. F1 Score

The output visualization results illustrate the effectiveness and practical usability of the proposed leafiq framework for real-time medicinal leaf classification. The proposed system gives an interactive web-based interface that allows users to upload medicinal leaf images and obtain classification predictions along with confidence scores and medicinal information.

Figs. 8–11 together illustrates the intelligent web-based medicinal leaf classification platform developed using the proposed hybrid deep learning framework.

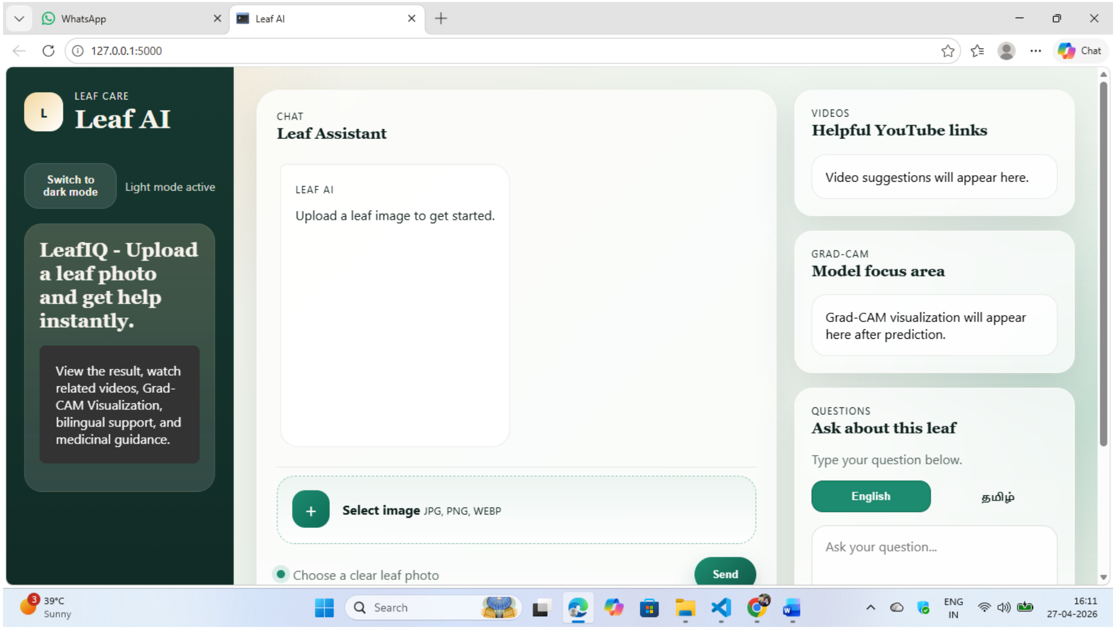

Fig. 8. User Interface of leafiq

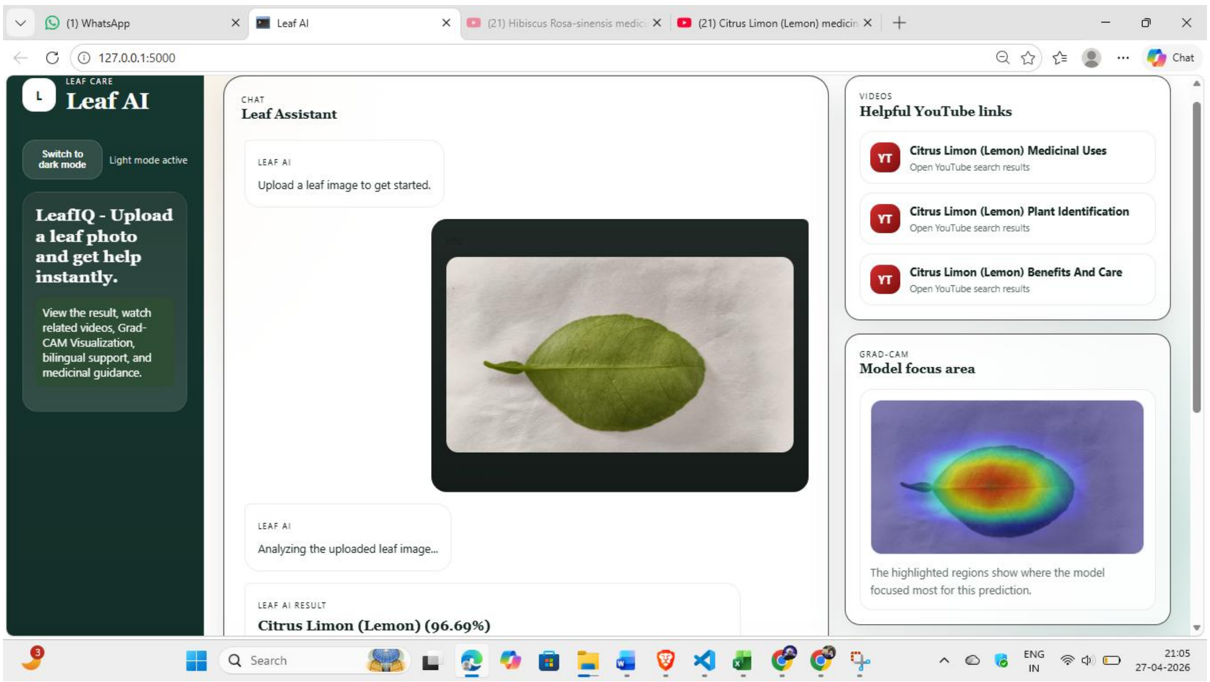

Fig. 9. Predicted Output with You Tube Recommendations, Grad-CAM Visualization Of Citrus Limon and Bilingual Support

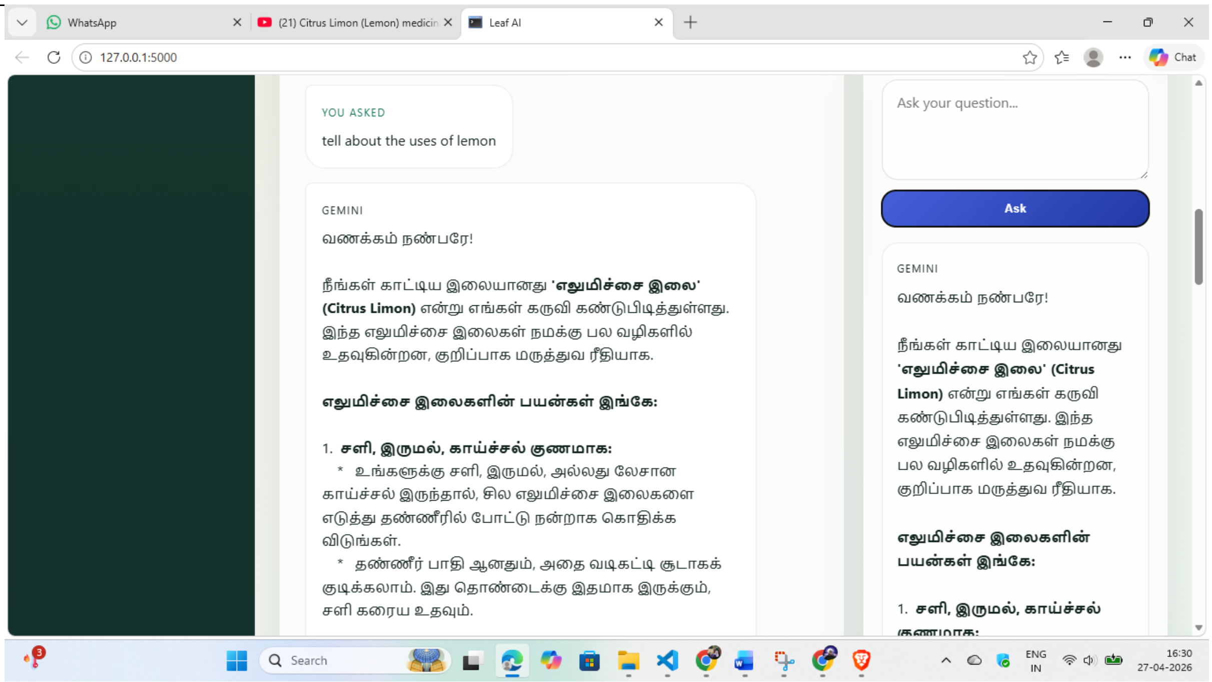

Fig. 10. Bilingual Chat Response of leafiq

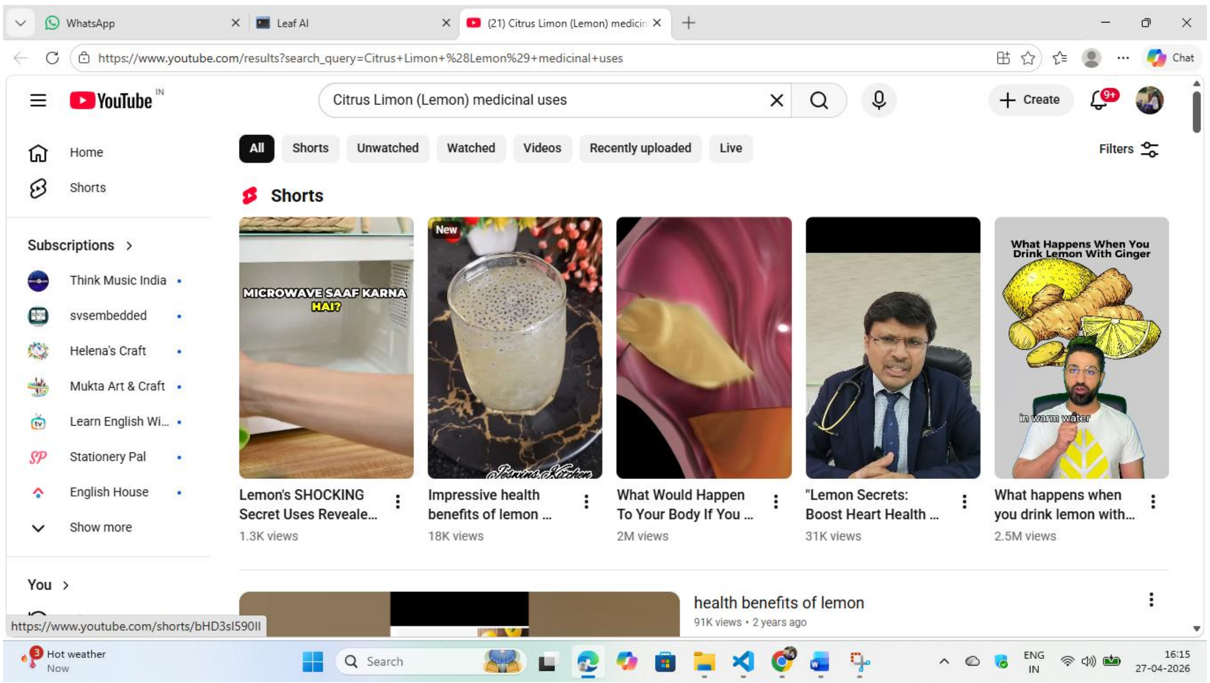

Fig. 11. You Tube API Integration to improve the user experience

The visualization results demonstrate the proposed leafiq framework achieves high classification performance and provides an explainable, interactive, and user-friendly platform for intelligent medicinal leaf identification.

Conclusion

The paper presents the system leafiq, a hybrid deep learning framework for medicinal leaf classification that fuses efficientnet-B1 for deep features and the GLCM for the texture features. The proposed system works well on similar leaves, which makes the system more reliable. The system uses the Data Augmentation and Grad-CAM for better performance of the system. Future work may focus on improving the dataset size with more species, which is accurate for real world scenarios, and developing a mobile application for the farmer and the normal people to know about the detail for the medicinal plants.

Future work

There are multiple techniques that can be used to advance the system development. For example, expanding the plant database with additional species is possible as the current system is capable of accommodating up to 100 plant species or more which necessitates the creation of the high-quality dataset with citizens' science data included. Additionally, deployment of the model in the mobile application in an offline fashion is possible via quantization and knowledge distillation in order to use the models created not by Flask but tensorflow Lite and ONNX (Open Neural Network Exchange). Another improvement to the system involves developing its ability to identify different plant organs in addition to leaves which means adding flowers, bark, and fruits recognition. Severity-aware plant disease recognition can contribute to the transition from plant identification to disease diagnostics of the plants. Self-supervised training of the model rather than the imagenet will contribute to improved feature extraction in the fine-grained classification. Active learning could become a new component of the system to train the model on incorrect identification results.

References

1. V. Gautam, G. Kaur, G. S. P. Ghantasala, P. Vidyullatha, S. Allabun, M. Othman, and A. Zheleznyak, “A lightweight deep learning method for medicinal leaf image classification using feature fusion,” Scientific Reports, vol. 15, Art. No. 34417, Jan. 2025, doi: 10.1038/s41598-025-17436-w.
2. H. Bouakkaz, M. Bouakkaz, C. A. Kerrache, and S. Dhelim, “Enhanced classification of medicinal plants using deep learning and optimized CNN architectures,” Heliyon, vol. 11, no. 3, Art. No. E42385, Feb. 2025, doi: 10.1016/j.heliyon.2025.e42385.
3. P. Dalvi, D. R. Kalbande, S. S. Rathod, H. Dalvi, and A. Agarwal, “Multiattribute deep CNN-based approach for detecting medicinal plants and their use for skin diseases,” IEEE Transactions on Artificial Intelligence, vol. 6, no. 3, pp. 710–724, Mar. 2025, doi: 10.1109/TAI.2024.3491938.
4. D. Chetia, S. Kr. Kalita, P. P. Baruah, D. Dutta, and T. Akhter, “Identification of traditional medicinal plant leaves using an effective deep learning model and self-curated dataset,” arxiv preprint, arxiv:2501.09363v1, Jan. 2025.
5. T. M. Garang Ok, M. Sibiya, Z. Wang, and M. Nkandla, “Investigation of deep learning techniques used in medicinal plants identification and classification,” IEEE Access, Aug. 2025, doi: 10.1109/ACCESS.2025.3598636.

A. S. Karnik, N. Nair, Y. Sagili, and P. B. Shanthi, “Multi-scale venation pattern analysis for medicinal plant species recognition,” IEEE Access, vol. 13, pp. 125526–125536, Jul. 2025, doi: 10.1109/ACCESS.2025.3589278.

6. M. T. Islam, M. W. Rahman, M. S. Hossain, K. Roksana, I. D. Azpiroz, R. M. Diaz, I. Ashraf, and M. A. Samad, “Medicinal plant classification using particle swarm optimized cascaded network,” IEEE Access, vol. 12, pp. 42465–42478, Mar. 2024, doi: 10.1109/ACCESS.2024.3378262.
7. S. Salsabila, A. Suharso, and P. Purwantoro, “Comparison of deep learning architectures in identifying types of medicinal plant leaf images,” Journal of Artificial Intelligence and Control, vol. 8, no. 1, pp. 39–46, Jul. 2024, doi: 10.30871/jaic.v8i1.6289.
8. R. Swarnalatha and S. Salunkhe, “EEG-centered Parkinson’s disease prediction system using deep learning techniques,” Medical Research Archives, vol. 11, no. 5, May 2023, doi: 10.18103/mra.v11i5.3699.
9. S. Roopashree and J. Anitha, “deepherb: A vision-based system for medicinal plants using Xception features,” IEEE Access, vol. 9, pp. 135927–135941, Sep. 2021, doi: 10.1109/ACCESS.2021.3116207.
10. R. Akter and M. I. Hosen, “CNN-based leaf image classification for Bangladeshi medicinal plant recognition,” in Proc. ETCCE, Dhaka, Bangladesh, 2020, pp. 1–6, doi: 10.1109/ETCCE51779.2020.9350900.
11. M. J. Hasan, S. Mazumdar, and S. Momen, “Deployable deep learning for cross-domain plant leaf disease detection via ensemble learning, knowledge distillation, and quantization,” IEEE Access, vol. 13, Aug. 2025, doi: 10.1109/ACCESS.2025.3595390.
12. S. O. Oppong, F. Twum, J. Ben Hayfron-Acquah, and Y. M. Missah, “A novel computer vision model for medicinal plant identification using log-Gabor filters and deep learning algorithms,” Computational Intelligence and Neuroscience, vol. 2022, Article ID 1189509, Apr. 2022, doi: 10.1155/2022/1189509.
13. M. Tan and Q. V. Le, “efficientnet: Rethinking model scaling for convolutional neural networks,” in Proc. ICML, 2020, arxiv:1905.11946v5.
14. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. CVPR, 2015, pp. 1–12, arxiv:1512.03385v1.
15. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in ICLR, 2015, arxiv:1409.1556v6.

A. Krizhevsky, I. Sutskever, and G. E. Hinton, “imagenet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.

B. Szegedy, W. Liu, Y. Jia, S. Reed, D. Erhan, and A. Rabinovich, “Going deeper with convolutions,” in Proc. CVPR, 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

16. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proc. CVPR, 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
17. J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proc. CVPR, 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.
18. R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.
19. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in Proc. ICCV, 2017, pp. 618–626, arxiv:1610.02391v4.
20. S. P. Mohanty, D. P. Hughes, and M. Salathe, “Using deep learning for image-based plant disease detection,” Frontiers in Plant Science, vol. 7, Art. 1419, 2016, doi: 10.3389/fpls.2016.01419.
21. J. Zhang, Z. Lin, J. Brandt, X. Shen, and S. Sclaroff, “Top-down neural attention by excitation backprop,” International Journal of Computer Vision, vol. 126, no. 10, pp. 1084–1102, 2016, arxiv:1608.00507v1.
22. P. Barré, B. C. Stöver, K. F. Müller, and V. Steinhage, “leafnet: A computer vision system for automatic plant species

identification," Ecological Informatics, vol. 40, pp. 50–56, 2017, doi: 10.1016/j.ecoinf.2017.05.005.

23. D. P. Hughes and M. Salathe, "An open access repository of images on plant health," arxiv, 2015, arxiv:1511.08060.
24. J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "imagenet: A large-scale hierarchical image database," in Proc. CVPR, 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.
25. F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proc. CVPR, 2017, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.
26. C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, p. 60, 2019, doi: 10.1186/s40537-019-0197-0.
27. B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in Proc. CVPR Workshops, 2017, pp. 136–144, doi: 10.1109/CVPRW.2017.151.
28. 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. ECCV, 2012, pp. 502–516, doi: 10.1007/978-3-642-33709-3_36.
29. T. Munisami, M. Ramsurn, S. Kishnah, and S. Pudaruth, "Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers," Procedia Computer Science, vol. 58, pp. 740–747, 2015, doi: 10.1016/j.procs.2015.08.095.
30. J. Waldchen and P. Mader, "Plant species identification using computer vision techniques: A systematic literature review," Archives of Computational Methods in Engineering, vol. 25, no. 2, pp. 507–543, 2018, doi: 10.1007/s11831-016-9206-z.
31. S. Ioffe and C. Szegedy, "Batch normalization," in Proc. ICML, 2015, pp. 448–456, arxiv:1502.03167v3.
32. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929–1958, 2014.
33. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," ICLR, 2017, arxiv:1412.6980v9.

A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, and G. Chanan, "pytorch: An imperative style, high-performance deep learning library," in neurips, 2019, arxiv:1912.01703v1.

34. G. Van Rossum and F. L. Drake, Python 3 Reference Manual, createspace, 2009.
35. S. Roopashree and J. Anitha, "Medicinal Leaf Dataset," Mendeley Data, vol. 1, 2020,
36. Doi: 10.17632/nnytj2v3n5.1.
37. Google, "Gemini API Documentation," Google AI for Developers. [Online]. Available:
38. Https://ai.google.dev/docs. [Accessed: 06-Jun-2026].