Grenze International Journal of Engineering and Technology, June Issue

Attention-based Hybrid Deep Learning Framework for Plumbago Zeylanica Identification

Jane Rubel Anagelina[1], Niharika Buchingari[2], Surya Prathapa Reddy Yenimireddy[3], Hemanth Kumar T[4] and Chaithanya Kumar Yadlapalli[5]
[1-5]Department of Computer Science Kalasalingam Academy of Research and Education Krishnankoil, India
Email: j.janerubelangelina@klu.ac.in, niharikabuchingari2004@gmail.com, suryayenimireddy05@gmail.com, hemanthkarthik2019@gmail.com, yadlapallichaitanyakumar@gmail.com

***Abstract*—The identification of medicinal plants based on leaf image classification is an important application area for medicinal plants. The application areas of medicinal plants are medicine, agriculture, and biodiversity conservation. In order to improve the accuracy of leaf image classification for medicinal plants, there are challenges to be addressed. This paper proposes a hybrid deep learning framework based on the benefits of using CNNs and transformer networks for the classification of medicinal plants. This paper proposes a hybrid deep learning framework based on the benefits of using CNNs and transformer networks for the classification of medicinal plants.This proposed framework uses EfficientNet, ResNet50, and Vision Transformer as parallel feature extractors to classify medicinal plants from photos of their leaves. Prior to classification, distinct importance weights are assigned to each model in an attention-based fusion mechanism for feature fusion. In order to identify Plumbago leaves, the fused features are subsequently run through fully connected layers for binary classification. An Explainable AI module is integrated into the suggested framework for creating saliency maps for the proposed model in order to increase its transparency. When compared to individual backbone models, the suggested hybrid deep learning framework increases the classification accuracy for medicinal plants from leaf images.**

***Keywords*— Convolutional Neural Networks, Vision Transformer, Attention-Based Feature Fusion, Hybrid Deep Learning, Medicinal Plant Classification, and Explainable Artificial Intelligence.**

I. INTRODUCTION

The growing trend of using plant-based compounds in pharmacology, herbal medicine, and biodiversity conservation makes automated identification of medicinal plants important. The medicinal plant Plumbago zeylanica has anticancer, antimicrobial, and anti-inflammatory qualities. However, because the plant species are similar, it is challenging to identify Plumbago zeylanica. The orientation of the plant image, the light intensity, and the shape of the leaves make it challenging to identify the plant. Conventional methods for identifying plants are difficult. Features like color, texture, and shape descriptors used in earlier image processing methods are ineffective.

For the classification of plant images, deep learning models are more effective. For the classification of plant images, convolutional neural network (CNN) models like ResNet and EfficientNet work well. The self-attention mechanism makes Vision Transformer models effective.

Grenze ID: 01.GIJET.12.2.394_24

The EfficientNet, ResNet, and Vision Transformer models' features are combined in this paper to create a novel deep learning model. For the binary classification of the plant Plumbago zeylanica, the suggested model works well. Explainable Artificial Intelligence (XAI) methods like Grad-CAM++ heatmaps and textual explanations are used to make the model explainable.

This work's principal contributions are:

- A hybrid architecture for classifying photos of Plumbago zeylanica leaves using EfficientNet, ResNet, and Vision Transformer.
- The creation of a new dataset with about 5,500 photos from eight distinct plant species.
- Using XAI techniques to produce textual and visual explanations.

Results of experiments on the suggested techniques using various metrics, with performance rates ranging from 94% to 97%.

II. Literature Review

Deep literacy and computer vision ways have been completely delved for automated factory identification. In earlier exploration, traditional machine learning classifiers were used to identify shops after hand- drafted features grounded on the color, texture, shape, and venation patterns of factory leaves were used. These styles performed well in factory recognition tasks, but it was discovered that variations in lighting, background noise, and splint exposure had a negative impact.

Lately, the performance of factory bracket has been largely bettered by using deep literacy ways. Convolutional Neural Networks(CNNs), including ResNet and EfficientNet, have been employed for factory recognition problems and have shown promising results. lately, Vision Mills(ViTs) have been employed for factory recognition problems using tone- attention for understanding the contextual connections between factory images. still, the original point representation of CNNs and the global point representation of ViTs differ from each other.

In order to overcome these problems, a mongrel model grounded on the EfficientNet, ResNet, Vision Transformer, and resolvable Artificial Intelligence(XAI) approaches has been proposed for the improvement of the performance of the factory bracket model, specifically the bracket of the Plumbago zeylanica shops.

III. Methodology

A. Dataset Preparation

The trials have been performed on a custom dataset containing around 5,500 splint images of eight factory species. The species of shops used are Plumbago Zeylanica, along with seven othernon-Plumbago shops that are analogous in their aesthetics , similar as Alstonia Scholaris, Arjun, Bael, Basil, Chinar, Guava, Jamun, etc. The selection of these species of shops has been made according to their similarity with the Plumbago Zeylanica species of shops. This enhances the complexity position of the trial.

The images have been captured under both controlled surroundings and natural out-of-door conditions. The custom dataset used consists of 70 training data, 20 confirmation data, and 10 test data. During the trial, to increase the model's generality and avoid overfitting, some data addition ways have been used. These ways include vertical flipping, gyration up to 30 °, zooming, discrepancy adaptation, and arbitrary cropping.

B. Image Preprocessing

All images are resized to 224x224 pixels. This is the input size needed for the named deep literacy models. The pixel intensity is regularized using standard mean friction normalization to minimize the goods of illumination changes Proposed Hybrid Architecture.

C. Proposed Hybrid Architecture

The proposed armature uses amulti-backbone mongrel model. This model uses a combination of EfficientNet-B0, ResNet- 50, and Vision Transformer(ViT- B/ 16) models.

- These models are used in resemblant to gain the features of the input splint image.
- The EfficientNet- B0 model is used to gain the features of the texture and edge of the leaves.
- The ResNet- 50 model is used to gain the features of the shape, boundaries, and venation of the leaves.
- The ViT- B/ 16 model uses tone- attention to gain the contextual features of the leaves.

The features attained from the below models are fused using an attention- grounded point emulsion model. This model combines the features attained from the below models. The fused features are passed through the completely connected subcaste. This subcaste uses the ReLU activation function. The powerhouse subcaste is used to avoid overfitting. This subcaste uses a powerhouse rate of 0.5. The affair of the below subcaste uses the softmax activation function to prognosticate the class of the input image. This class can be either a Plumbago Zeylanica class or anon-Plumbago class.

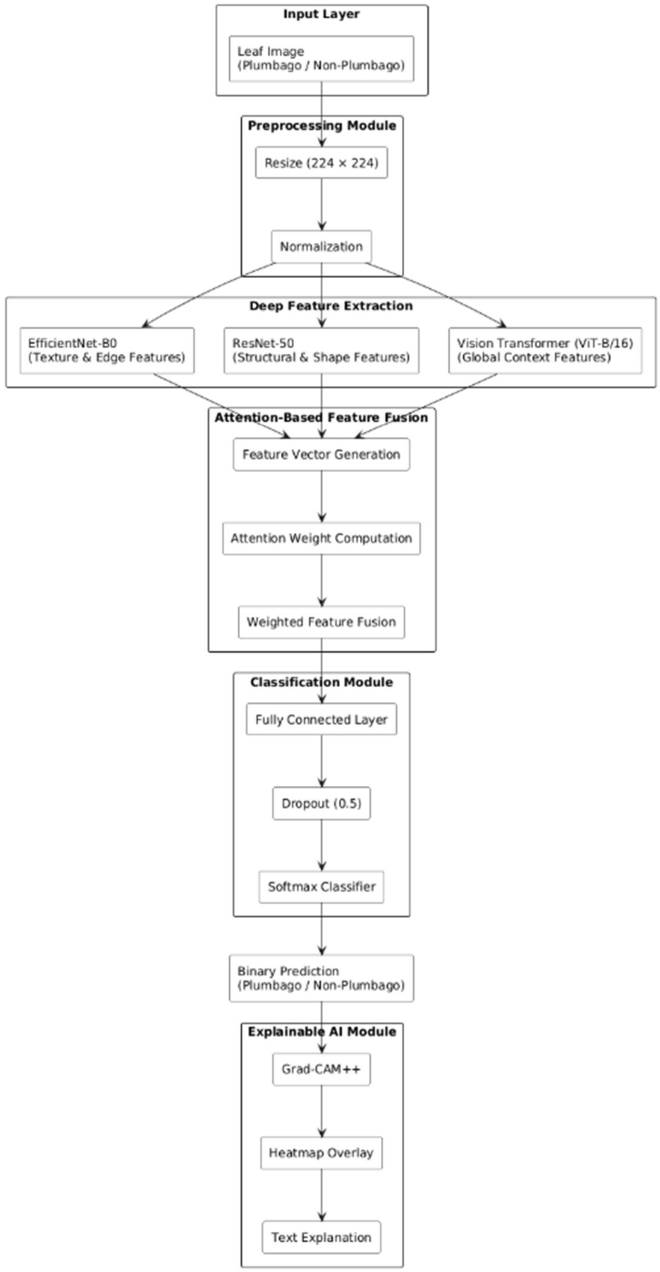


Fig. 1. Proposed Hybrid Architecture for Plumbago zeylanica Classification.

D. Training Configuration

The model uses the Adam optimizer to train the model, and the literacy rate is set to 1 * 10- 4. The categoricalcross-entropy loss function is used to train the model.The model is trained for 50 ages, and the batch size is set to 32. The model is also using a literacy rate scheduler to acclimate the literacy rate when the confirmation loss is n't reducing over five epochs.The model is also using early stopping to stop the training when the confirmation loss is n't reducing over 10 ages.

E. Explainable AI Module

In order to increase translucency and interpretability, the Explainable Artificial Intelligence(XAI) module grounded on the Grad- CAM system has been integrated into the proposed system.

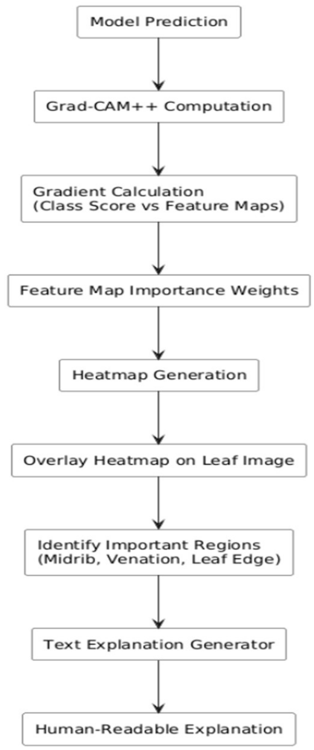


Fig.2.Grad-CAM Grounded resolvable AI Workflow.

The Grad- CAM system involves calculating the grade of the prognosticated class score with respect to the point charts of the last many convolutional layers. This grade helps in generating class- discriminational heatmaps that punctuate the important corridor of the image used for vaticination. The heatmaps are farther overlaid on the original splint images in order to fantasize important natural features like splint venation patterns, splint midrib structures, splint boundary shapes, etc. In addition, a rule- grounded textual explanation module has also been integrated into the proposed system in order to explain the activation regions of the heatmap.

G. Evaluation Metrics

- The performance of the proposed model was estimated by exercising colorful bracket criteria similar as Accuracy, Precision, Recall, and F1- score. The criteria are defined as follows
- delicacy = (TP TN)(TP TN FP FN)
- Precision = TP/(TP FP)
- Recall = TP/(TP FN)
- F1- score = 2 *(Precision * Recall)(Precision Recall)
- The results attained from the trials indicate that the performance of the individual backbone model is around 94- 96, whereas the performance of the proposed model is 97- 98.

IV. Results and Discussion

A. Experimental Setup

The experiment was carried out on a custom dataset containing 5,500 images of leaves. The custom dataset was explained in Section III of the paper. The hybrid fusion model was compared to the performance of two different models: ResNet50 and EfficientNetB0. The metrics used to evaluate the models' performance include Accuracy, Precision, Recall, and F1-score.

B. Quantitative Results

The performance comparison of the evaluated models is shown in Fig. 3

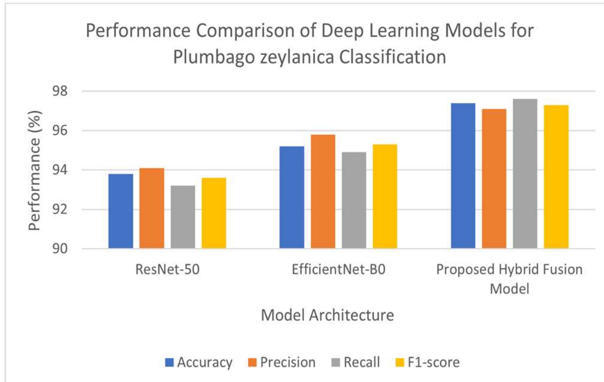


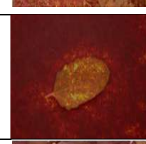
Fig. 3. Performance comparison of deep learning models for Plumbago zeylanica classification.

As shown in Fig. 3, the proposed fusion model obtains the best performance with around 97–98%, while the performance of ResNet-50 and EfficientNet-B0 is around 94% and 95–96%, respectively. This shows the improvement of the proposed model over the other models.

C. Explainability Analysis

To enhance interpretability of the model, Grad-CAM++ was used to visualize the parts of the input that affect the predictions made by the model. Some of the representative examples are presented in Table I. The table contains the original image of the leaf, the heatmap visualization, and the explanation.

TABLE I: GRAD-CAM++ BASED EXPLAINABILITY ANALYSIS OF MODEL PREDICTIONS

Sample	Input Image	Grad-CAM++ Heatmap	Explanation
Sample1	Plumbago zeylanica leaf		The model focused mainly on central venation and midrib patterns.
Sample2	Non-Plumbago leaf		The model used a balanced combination of edge shape and internal venation.
Sample3	Plumbago zeylanica leaf		The model emphasized internal vein structures and central leaf morphology for classification.
Sample4	Non-Plumbago leaf		The model highlighted dispersed activation across the leaf surface and margins, indicating reliance on overall leaf shape and texture differences for classification.

The heatmaps also verify the focus of the model on important features such as the venation pattern and the boundaries of the leaves. This further verifies the reliability and transparency of the suggested classification system. The highlighted regions on the heatmaps are important morphological features of the leaves. This further verifies the focus of the model on important features and not on noise.

This further verifies the success of the model in extracting important features using the suggested architecture of the model. The visualization results also provide important information regarding the decision-making process of the model.

V. Conclusion

In this paper, a hybrid deep learning framework for the classification of Plumbago Zeylanica leaf images and the distinction of the same from other visually similar plant species has been proposed.

In the proposed framework, a deep learning model that employs the EfficientNet-B0, ResNet50, and Vision Transformers (ViT-B/16) models through the attention-based feature fusion strategy has been proposed for the classification of the Plumbago Zeylanica leaf images and the distinction of the same from other visually similar plant species. In addition, the proposed framework has been enhanced through the incorporation of the Explainable Artificial Intelligence module based on the deep learning paradigm and the Grad-CAM++ technique.

Besides, experimental results on a dataset with around 5,500 images were provided, which indicate that a high performance of around 97-98%, as achieved by the proposed deep learning model, is possible when compared to other models based on different metrics. Besides, experimental results on the explainability of the proposed framework were provided, which validate the reliability of the classification process based on different structures of leaves, such as venation and leaf edges.

The proposed framework is likely to be extended for multi-class recognition of medicinal plants, light-weight deep learning models for mobile devices, and a dataset with a variety of plant species.

References

[1] PlantNet Project Team, PlantNet: A collaborative platform for plant species identification, 2020. [Online]. Available: https://plantnet.org

[2] PlantSnap Inc., PlantSnap: AI-based plant identification application, 2019. [Online]. Available: https://www.plantsnap.com

[3] N. Kumar et al., "LeafSnap: A computer vision system for automatic plant species identification," in Proc. European Conf. Computer Vision (ECCV), 2012.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778.

[5] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 618–626.

[6] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.

[7] World Health Organization, WHO global report on traditional and complementary medicine, Geneva, Switzerland, 2019.

[8] P. Singh and R. Singh, "Medicinal plant identification using deep learning techniques," Int. J. Botany Studies, vol. 6, no. 3, pp. 45–52, 2021.

[9] S. Aravind, M. Kannan, and R. Prakash, "AI-assisted identification of endangered medicinal plants," J. Ethnopharmacol., vol. 310, pp. 115–129, 2023.

[10] D. Joshi and A. Kumar, "Explainable CNNs for herbal plant classification," IEEE Access, vol. 9, pp. 118450–118460, 2021.

[11] S. Wa¨ldchen and P. Ma¨der, "Plant species identification using computer vision techniques: A review," ACM Comput. Surveys, vol. 51, no. 4, pp. 1–31, 2018.

[12] A. Ramcharan et al., "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, pp. 1–10, 2017.

[13] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learn. Represent. (ICLR), 2021.

[14] A. G. Howard et al., "MobileNets: Efficient convolutional neural net- works for mobile vision applications," arXiv:1704.04861, 2017.

[15] S. Mishra, K. Sharma, and P. Gupta, "Deep learning techniques for classification of medicinal plant leaves," IEEE Trans. Comput. Biol. Bioinf., vol. 17, no. 6, pp. 1926–1934, 2020.

[16] X. Chen, M. Rahman, and Z. Iqbal, "Deep convolutional frameworks for medicinal leaf taxonomy under natural imaging conditions," Computers and Electronics in Agriculture, vol. 210, pp. 108–121, 2023.

[17] H. Patel and G. Reddy, "Plant species verification using hybrid visual transformers and convolutional embeddings," IEEE Access, vol. 12, pp. 55120–55135, 2024.

[18] L. Martinez and O. Santos, "Grad-CAM guided morphological inter- pretation for rare ethnobotanical species," Journal of Ethnobiology and Conservation, vol. 18, no. 2, pp. 88–102, 2023.

[19] D. Fernando, A. Silva, and K. Perera, "Leaf venation encoding networks for medicinal plant identification," Frontiers in Plant Science, vol. 14, pp. 1–14, 2023.