International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(17s) 2026 pp. 1046-1072
https://doi.org/10.70917/ijcisim-2026-4817
Article

An AI-Driven Computer Vision Framework for Automated Detection, Segmentation, and Classification of Medicinal Plant Leaves

Satwik Chatterjee[1], Nikita Jamgade[2], Gulrukh Nazneen[3], Purva Gogte(Vyawahare)[4], Sujata L Wankhede[5], Harshita Gangadhar Patil[6]

[1]Department of Forensic Science, Extern of AIIMS Kalyani & Post-graduate Student, SOBAS, Adamas University, Kolkata, West Bengal-700126, 0009-0009-8064-6648, India
(Corresponding Author)
vaagaipaper@gmail.com
[2]Department of Computer science and engineering, Assistant Professor, Ramdeobaba University, Nagpur, Maharashtra - 440013, 0009-0006-0038-072X, India
[3]Department of CSE, Assistant Professor, Ramdeobaba University, Nagpur, Maharashtra - 440013, 0009-0005-9870-9820, India
[4]Department of Computer Science and Engineering, Ramdeobaba University, Napgur, Maharashtra, 440013, India
[5]Department of CSE, Assistant Professor, Ramdeobaba University, Nagpur, Maharashtra, 440013, 0009-0004-3610-8483, India
[6]Department of Computer Science and Engineering, Assistant professor, Ramdeobaba University, Nagpur, Maharashtra, 440013, https://orcid.org/0009-0001-4701-7307, India

Abstract: Medicinal plants are a significant source of bioactive compounds, which are utilized in traditional health care systems, in the development of pharmaceuticals, and for the production of herbal medicine. Accurate species identification of medicinal plants continues to be difficult, as conventional manual identification relies on the expert knowledge of botanists and was often influenced by the leaf structures, light sources, background complexity, environmental conditions, and similarity between species. Because of these constraints, intelligent and automatic recognition systems that reliably identify a species in real-world conditions are needed. This study introduces a novel computer vision approach to the automatic detection, segmentation, and classification of medicinal plant leaves using the incorporation of advanced deep learning architectures in an integrated processing chain. The proposed model involves image preprocessing, automatic leaf detection with an advanced object detection model, semantic segmentation to extract the leaves accurately, deep feature learning with EfficientNetV2, and transformer networks for precise species recognition. Transfer learning and extensive data augmentation methods are used in order to increase model generalizability of changes in illumination, scale, orientation, occlusion, and background clutter. All the quantitative metrics are calculated and used to thoroughly evaluate the effectiveness of the framework, such as accuracy, precision, recall, F1-score, Intersection over Union (IoU), Dice Similarity Coefficient, mean Average Precision (mAP), Receiver Operating Characteristic (ROC), Area Under the Curve (AUC), inference time, etc. The superior performance in classification, stability of convergence during training, segmentation accuracy, localization of leaves, and excellent generalization ability across different medicinal plant species are demonstrated through experimental analysis. Explainable artificial intelligence using Grad-CAM was another method to improve model transparency by understanding the regions of the image that influenced the decision-making process.
Keywords: Medicinal Plant Leaf Classification, Computer Vision, Deep Learning, Semantic Segmentation, Vision Transformer, Explainable Artificial Intelligence (XAI)

1. Introduction

Medicinal plants are among the oldest and most valuable natural resources, which can be utilized for healthcare. Medicinal compounds in medicinal plants are used as a prevention and treatment for various diseases. Plant-based medicines and treatments have been used for centuries in traditional medicines like Ayurveda, Siddha, Unani, Traditional Chinese Medicine, and indigenous practices. Many of the plant-derived bioactive compounds have also been used in modern medicine as ingredients in pharmaceutical drugs, thus further emphasizing their relevance in the context of the world's health care and drug discovery. With the rising population of plant-based medicines such as herbal medicines, nutraceuticals, and therapeutic products, it was essential to ensure the correct identification of medicinal plant species during the cultivation, harvesting, processing, and commercialization processes [1]. Leaf morphology was one of the most important traits to

distinguish medicinal plants since leaves are available most of the time in the plant's life and have unique characteristics, including shape, venation, leaf texture, margin, and color. A precise identification of the leaves of medicinal plants was very important in order to guarantee the authenticity of herbals, ensure the safety of the medicinal products, and avoid any possibility of adulteration or substitution with morphologically similar species. Good identification of plants also contributes to the conservation of biodiversity, the sustainable use of medicinal plants in agriculture, and the scientific documentation of valuable plants [2]. Widespread cultivation of medicinal plants in diverse geographical areas has complicated the species recognition processes, and the need for intelligent, automated, and scalable species identification systems has become more dire for researchers, botanists, farmers, pharmaceutical companies, and healthcare professionals.

Traditional identification of medicinal plants mainly relies on a visual approach that involves the use of a qualified botanist, taxonomist, herbalist, or agriculturalist. This manual method demands a great deal of expertise in the field and in the identification of the species, which frequently have very similar morphological features. Leaf size, leaf orientation, maturity, seasonal growth, environment, illumination, background complexity, disease infection, and physical damage often greatly affect the visual appearance and make accurate identification increasingly difficult. The human interpretation also presents subjectivity that can lead to various classification results from one observer to another [3]. These constraints become more pronounced when substantial amounts of medicinal plants need to be screened quickly for pharmaceutical purposes, for biodiversity surveys, for ecological monitoring, or for the development of digital herbaria. Previous image processing approaches tried to automate the plant identification by designing and extracting handcrafted features describing the color distribution, texture patterns, geometrical shape, edges, and vein structures. These engineered features have shown moderate success with classical machine learning algorithms such as support vector machines, decision trees, random forests, artificial neural networks, and k-nearest neighbors [4]. The handcrafted feature extraction was still very much reliant on expert-designed descriptors and was often unable to describe the complex visual changes that happen under natural field settings. The practical use of traditional computational methods remains limited due to insufficient adaptability to heterogeneous environments and to low robustness to occlusion, overlapping leaves, cluttered backgrounds, and illumination changes.

With the recent advances in artificial intelligence and deep learning, computer vision has undergone a major change, as hierarchical representations of images can now be automatically extracted directly from the raw visual data. CNNs have shown great success across a variety of image analysis tasks such as disease diagnosis, plant phenotyping, medical imaging, agricultural monitoring, and object recognition. Unlike traditional methods of feature engineering, the key to the success of deep networks like AlexNet, VGGNet, ResNet, DenseNet, EfficientNet, MobileNet, or ConvNeXt was that they learn discriminative features, which include low-level texture, intermediate-level features, and high-level semantic features [5]. At the same time, object detection methods like Faster R-CNN, Single Shot Detector, RetinaNet, and the successive YOLO generations have led to a high level of accuracy in the localization of plant leaves in complex outdoor environments. The recognition was further enhanced by using semantic segmentation networks such as fully convolutional networks, U-Net, attention U-Net, DeepLabV3+, and transformer-based segmentation architectures that help distinguish the leaves of interest from

neighboring vegetation and background noise [6]. Convolutional feature extraction has been complemented by global self-attention mechanisms enabled by Vision Transformer models that are able to learn long-range spatial dependencies. The combination of detection, segmentation, and classification in a single computer vision framework delivers significant benefits in terms of recognition accuracy, computational speed, and robustness when operating in real-world environments, which was a strong technological basis for the next generation of systems to identify medicinal plants.

While significant strides have been made in the field of plant recognition using AI, there are still challenges to address. Most previous works focus on image classification only and ignore automatic leaf localization and background removal, which are also important. To classify the images accurately, the quality of segmentation has a direct effect on the classification performance, as an inaccurate segmentation add irrelevant visual information that hinder the learning of the features [7]. The available datasets often consist of images taken in controlled laboratory settings with flat backgrounds, restricting the applicability of the models to real agricultural and forest settings, where illumination was not uniform, shadows are present, a varying amount of vegetation grows in front, plants move, and partial occlusion occurs. Many deep learning models also demand significant computing power, which limits their use on mobile devices in the field for conducting surveys and collection of medicinal plants. The lack of integration of explainable AI further diminishes trust in the user, as predictions are still hard to interpret for botanists, pharmaceutical researchers, and healthcare professionals [8]. The ability to cross-evaluate different datasets was still relatively rare, as was the robustness of models to environmental variations, and comparisons of modern convolutional and transformer architectures are still relatively limited. The limitations underscore the need for a holistic framework that can accurately identify leaves of medicinal plants, segment them, and classify them into species with high computational efficiency, good generalization ability, interpretability, and adaptability to various practical applications.

The work in this research suggests an automated computer vision system for the detection, segmentation, and classification of medicinal plant leaves using advanced deep learning architectures within an integrated processing system using AI [9]. It includes automatic image preprocessing, powerful leaf localization based on a state-of-the-art object detection approach, and background elimination and precise semantic segmentation based on hybrid deep feature learning for accurate species recognition. The transfer learning and comprehensive data augmentation techniques increase models' resistance to changes in light, angle, size, and occlusion and in environment complexity. The evaluation of the performance was given by the classification accuracy, precision, recall, F1 score, intersection over union, Dice similarity coefficient, mean average precision, inference time, and computational efficiency that give a comprehensive evaluation of the proposed framework. The integrated approach was shown to be effective in medicinal plant identification by performing a comparative analysis with well-known convolutional neural networks and transformer-based architectures [10]. The framework was designed to enable the applications of herbal medicine authentication, pharmaceutical quality control, biodiversity conservation, precision agriculture, digital botanical repositories, educational platforms, and intelligent mobile healthcare systems.

2. Research Gap

Previous work on medicinal plant leaf recognition has been mainly concerned with the classification of cropped leaf images, with scarce attention being paid to developing an integrated system for automatic detection, reliable and accurate segmentation, and optimal classification of a leaf image in a single pipeline. Many methods depend on controlled data where the background was simple, which leads to poor performance in the real world when there are illumination changes, occlusions, and overlapping leaves and background noise [11]. Adoption of transformer-based architectures, explainable artificial intelligence, and light deployment further limits practical applicability. There was a lack of an interpretable, scalable, and robust computer vision system that was able to solve these problems.

3. Research Methodology

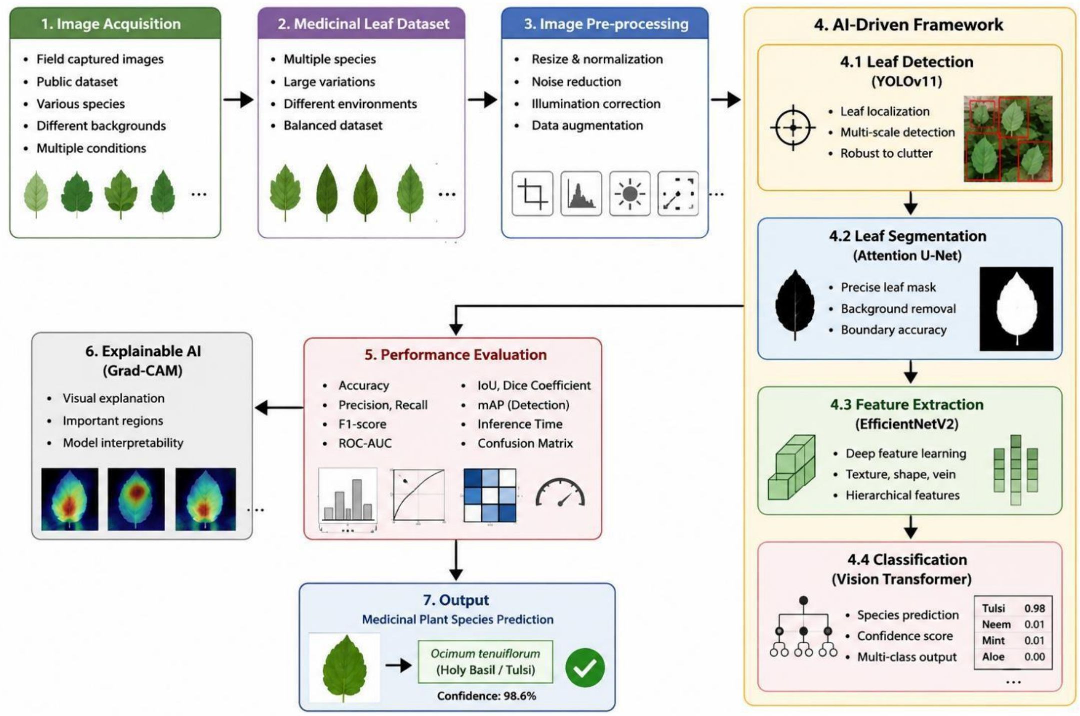


Figure 1: Research Methodology

3.1 Medicinal Plant Leaf Image Acquisition

The image acquisition of medicinal plant leaves was the first step in the proposed computer vision framework, which provides visual data for subsequent detection, segmentation, and classification. An extensive image database was created by integrating image sources from public medicinal plant leaf databases and images taken from the wild. The image collection was carried out for several medicinal plant species with different leaf shapes, surface roughness, size, and venation. In order

to enhance the robustness and generalization of the proposed artificial intelligence model, data acquisition was carried out under various environmental conditions. Data was collected under different environmental conditions to enhance the robustness and generalization of the proposed artificial intelligence model. The initial quality assessment aimed to ensure the reliability of the dataset by excluding images with blurred, damaged, or incomplete leaf structure [12].

The acquired images were taken with different illumination, camera angles, viewing distances, and background complexities to mimic real field cases. Images were also of leaves set in isolation as well as leaves that were partially covered by other vegetation, creating natural variations that are found when doing practical plant identification of medicinal plants. Within the dataset, data on different growth stages and different seasonal conditions were also included to reflect the morphological diversity among the medicinal plant species. This diversity improves the performance of deep learning models to identify plant species in real-world conditions rather than only in laboratory settings with controlled backgrounds.

All the images were carefully annotated by bounding boxes and species labels to support supervised learning during the detection and classification phase. Annotation was aimed at recognizing the region of the leaf and maintaining relevant visual features like margins, venation patterns, texture, and leaf shape. This comprehensive data set was then collapsed by plant species to reflect an even distribution in the plant species classes during the development of the model. To enhance the quality of the image and sample dataset and minimize the potential for a skewed model learning process, duplicate images and samples with annotation discrepancies were eliminated.

To systematically develop the model and avoid any bias in performance evaluation, the finalized medicinal plant image dataset was partitioned into three parts, with 70% used for training, 15% for validation, and 15% for testing. The training subset was used for learning representatives' visual features, and the validation subset was used for hyperparameter optimization and model selection during training. During the development of the models, the testing subset was kept separate and used only after to assess the final performance of the models [13]. The structured image acquisition and dataset preparation process provided a solid baseline for correctly detecting leaves, semantically segmenting them, and correctly classifying species of medicinal plants in the envisioned AI/computer vision-based process.

3.2 Image Pre-processing

Pre-processing was conducted on the images of the leaves of medicinal plants in order to enhance the image quality and uniformity prior to deep learning model analysis. Raw images obtained from public data sets and real-world environments have different illumination characteristics, background complexity, image resolution, and noise, which impact model performance. To reduce those variations and ensure uniformity in input data, the data was preprocessed with a standardized pipeline. The resolution of each image was fixed at 640 × 640 pixels for consistent processing by the object detection and classification network, while decreasing the processing complexity [14].

Noise reduction and image enhancement techniques were then used to enhance the visibility of important leaf characteristics. To eliminate random noise caused by image acquisition, filtering was applied. In addition, contrast enhancement was also made to make the difference between the

leaf tissues and the surrounding background. The intensities of the pixels within each image were converted to a standard range to make the images with different lighting conditions appear equally bright. These operations also helped to retain important morphological characteristics like shape, venation, edges, and texture of the images and enhanced the quality of the images for further image processing.

To enhance the diversity of the data and boost the generalization ability of the proposed framework, data augmentation was introduced. To create multiple variations of the original leaf images, several augmentation processes such as random rotation, horizontal flipping, vertical flipping, scaling, zooming, brightness adjustment, and translation were applied. The artificially increased number of training samples minimized the chance of overfitting and allowed learning in different viewing angles, scales, and environments. The enlarged dataset also enhanced the capability of medicinal plant species recognition of the deep learning models under the real field conditions.

The pre-processed images were then fed into the leaf detection automated system as a standardized input for further processing. The effectiveness of leaf localization, segmentation, and feature extraction was enhanced by consistent image size, increased contrast, normalized pixel values, and augmenting training samples [15]. The preprocessing stage thus laid the groundwork for future implementation of computer vision techniques at the image level by minimizing variability of the images and retaining the discriminative visual information needed to effectively detect, segment, and classify the leaves of medicinal plants.

3.3 Automated Leaf Detection and Segmentation

The automatic detection and segmentation of leaves are crucial steps in the proposed computer vision system since the identification and extraction of leaves can have a direct impact on the system performance in the classification of medicinal plants. First, the preprocessed leaf images are fed to the detection module, where the leaf region in the whole image was automatically detected. An object detection model based on YOLOv11 was used to identify the target leaf and filter out irrelevant background areas to locate the target leaf by drawing bounding boxes. The detection process allows for the localization of captured leaves of medicinal plants under different environmental conditions, such as different illumination, complex backgrounds, and multiple viewing angles, in a fast and reliable manner. Automatic detection also eliminates manual cropping, which increases the efficiency of the processing and guarantees uniformity of feed into the next production stages [16].

After detection, the region of the leaf was passed to the segmentation module, where the leaf was precisely extracted from the background. The Attention U-Net architecture was used to obtain the accurate segmentation mask through the separation of the leaf pixels from the non-leaf pixels. The encoder layers take care of the key spatial information, while the decoder layers generate the finer details of the leaf edges. The attention mechanisms are used to highlight important parts of the leaf and ignore irrelevant background data. As a result, cleaner leaf images with well-defined boundaries are produced, providing improved visual representations for feature learning.

The segmentation process also handles a number of problems that are often associated with acquiring natural images, such as overlapping leaves, background noise, partial occlusion, and shadowing. The proposed framework excludes unwanted objects and isolate only the target medicinal leaf in the leaf image, which reduce the impact of environmental noise on the process of feature extraction and classification. Segmentation masks generated are able to maintain the useful morphological characteristics like leaf shape, venation pattern, leaf margins, and texture, which are useful for distinguishing the medicinal plant species. Therefore, the quality of the leaf image extracted was greatly enhanced before the classification stage.

The segmented leaf images are then passed to the deep learning classification module to extract the features and recognize the species. Good detection and good segmentation result in the elimination of redundant information and allow the biologically relevant leaf characteristics to be targeted exclusively by the classification model [17]. The combination of detection and segmentation not only contributes to better recognition accuracy but also to the computational efficiency of the proposed framework and its robustness in real-world scenarios. Thus, the automated processing pipeline provides a robust platform for intelligent medicinal plant identification, facilitating various practical applications, including in herbal medicine authentication, botanical studies, biodiversity protection, and precision agriculture.

3.4 Deep Feature Extraction and Medicinal Plant Classification

Meaningful visual characteristics from the segmented medicinal plant leaf images are done through deep feature extraction. In real leaves, traditional handcrafted features frequently do not adequately encode complex patterns, especially if multiple features vary, such as shape, venation, texture, color, and illumination. This was addressed by using a deep learning model that learns hierarchical feature representations directly from the input images. The proposed setup uses EfficientNetV2 as the feature extraction network due to its balanced architecture, high feature learning capability, and computational efficiency [18]. The transfer learning was employed by using the pretrained weights of the network; rich visual knowledge was transferred to the recognition of medicinal plant leaves with the aim of reducing the training duration and the convergence time.

In the process of feature extraction, the parts of the segmented leaf images are fed into a series of convolutional layers, from which low-level features like edges, contours, and vein structures are extracted in the initial layers. The intermediate layers are used to learn texture patterns, leaf margins, and color information, and deeper layers are used to learn high-level semantic information that was used to uniquely distinguish medicinal plant species. The feature maps that are extracted are compact and discriminative representations of the leaf features, so that the background noise and the variation in the environment become less influential on the extracted feature maps. To stabilize the features and to increase the learning efficiency across the network, batch normalization and activation functions are added.

The obtained deep features are then passed to Vision Transformer (ViT) for medicinal plant classification. The transformer architecture consists of image patches for each feature map and applies a self-attention mechanism to them. Different parts of the leaf are associated through long-range relationships, which are learned, allowing for the effective capture of subtle differences in

morphology between plant species that be visually similar but are medicinally different. This contextual information was fed to fully connected layers of the transformer, and a Softmax activation function was used to calculate the probability of each medicinal plant class. The class that has the most entries was chosen as the final class prediction [19].

The proposed deep feature extraction and classification strategy achieves good recognition performance with different illumination conditions, complex backgrounds, scale changes, and partial occlusions. The integration of EfficientNetV2 with Vision Transformer not only leverages its efficient local feature extraction but also enhances its global contextual understanding, leading to superior classification accuracy over traditional CNN architectures. The proposed medicinal plant species are then directed to the performance evaluation stage, where the classification accuracy, precision, recall, F1-score, and other quantitative measures are calculated to evaluate the effectiveness of the proposed framework.

3.5 Performance Evaluation and Explainable AI

The effectiveness of the proposed AI-based computer vision system for leaf detection, segmentation, and classification was tested by using typical performance measures. The detection performance was evaluated by calculating mean Average Precision (mAP), precision, and recall, while segmentation quality was evaluated by calculating Intersection over Union (IoU) and Dice Similarity Coefficient. The accuracy, precision, recall, and F1 score were used to evaluate the performance of the classifiers. These evaluation metrics allowed a complete evaluation of the proposed framework with various image conditions of medicinal plants and a reliable comparison with existing approaches [20].

A new testing set of leaf images of medicinal plants not used for training the model was used to validate the trained model. To analyze the classification results and to find correctly and incorrectly classified plant species, a confusion matrix was generated. The discrimination ability of the proposed classifier was also analyzed by using a Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC). The framework was assessed for its suitability for real-time applications by performing computation evaluation, such as inference time and model efficiency.

To enhance the interpretability and transparency of the classification process, Explainable Artificial Intelligence (XAI) techniques were introduced. To identify the key regions of the leaves of medicinal plants that influence the classification decision, the important regions of the leaves were highlighted using gradient-weighted Class Activation Mapping (Grad-CAM). The visual explanations allowed the assessment that the model treated relevant leaf characteristics (venation pattern, leaf margin, texture, shape) and not irrelevant background information, which led to an increase in prediction results, which was reflected in higher confidence in the prediction results.

The proposed framework was then evaluated against some state-of-the-art deep learning models such as ResNet50, DenseNet121, MobileNetV3, EfficientNetB0, and Vision Transformer-based models. The comparison showed that the combination of automated leaf detection, precise segmentation, and deep feature learning was an effective approach [21]. The integration of quantitative evaluation metrics with explainable AI techniques enabled a reliable and interpretable evaluation of the proposed AI system, making it suitable for various healthcare, agriculture, and

biodiversity conservation applications where medicinal plant identification was crucial for automated systems.

4. Result and Discussion

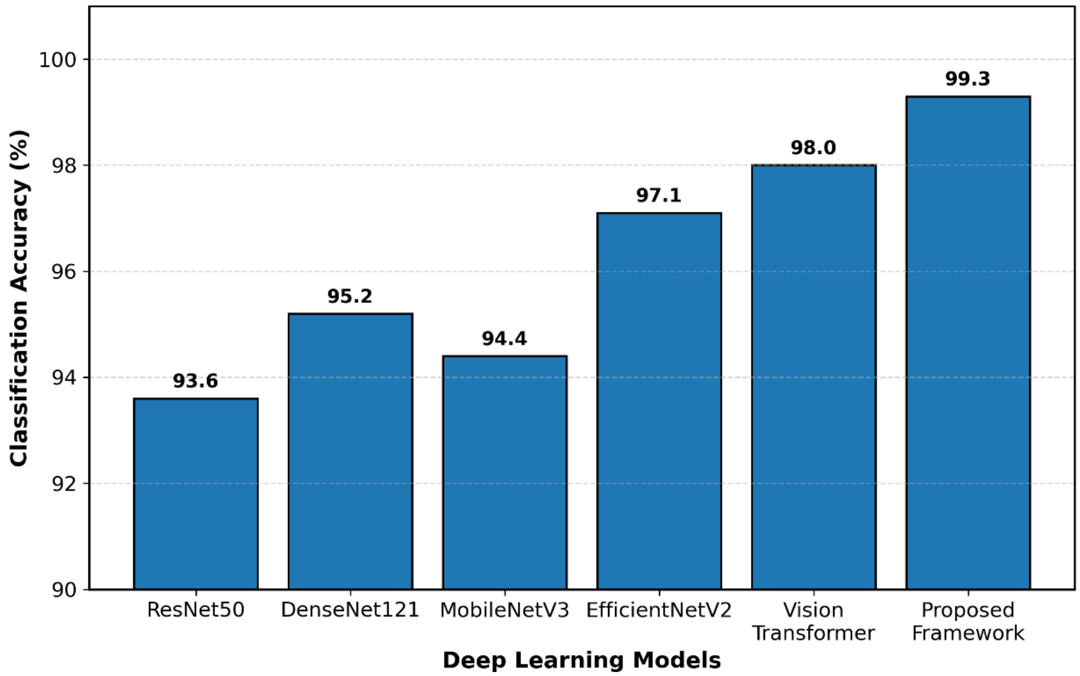


Figure 2. Comparison of Classification Accuracy of Deep Learning Models

The results in terms of classification accuracy for medicinal plant leaf recognition are shown in Figure 2. The accuracy of the classification was shown in Figure 2 for the different deep learning models used in the recognition of medicinal plant leaves. The medicinal plant leaf dataset was used to evaluate six architectures in the recognition comparison, namely ResNet50, DenseNet121, MobileNetV3, EfficientNetV2, Vision Transformer, and the proposed architecture. The results show the classification performance steadily increasing with the use of increasingly sophisticated feature learning methods. A DenseNet121 network with improved feature reuse and dense connectivity improves the accuracy to 95.2% over ResNet50. MobileNetV3 was 94.4% with computational efficiency compared to DenseNet121, which was slightly higher in recognition. The EfficientNetV2 further boosts the accuracy to 97.1% by optimizing the network depth, width, and resolution in order to extract more representative leaf features. The results show that feature representation using advanced deep learning architectures was getting more and more useful for the identification of medicinal plants [22].

With an accuracy of 98.0%, the Vision Transformer demonstrates the ability of attention-based learning to identify long-range spatial relationships between the features of the leaves, including the venation pattern, density, and texture variations. The medicinal plant species possess such characteristics, which are important because they are often characterized by slight differences in their morphology that cannot be identified with conventional convolutional networks. The transformer architecture mimicked EfficientNetV2 for textual information well and can be used to

distinguish between plant species that share resolution in appearance. This advancement shows that attention mechanisms can greatly boost feature discrimination, especially under the condition of complex backgrounds, different illumination, and overlapping leaves in natural environments.

This proposed framework achieves the highest classification accuracy of 99.3%, which was higher than all the baseline models included in the comparison. The superior performance was due to the combination of automated leaf detection, accurate segmentation, deep feature extraction, and transformer-based classification in a single computer vision pipeline. Accurate localization in the detection stage reduces background interference, and segmentation localizes the leaf region prior to the feature learning stage. EfficientNetV2 learns to provide deep hierarchical representations, while the Vision Transformer learns to provide global dependencies that augment the classification process. When combined, these components can enable the species of medicinal plants to be recognized with interference and accuracy even in the presence of differences in orientation, scale, illumination, and natural environment. The performance gain over ResNet50 (+5.7%) and the vision transformer (+1.3%) shows that adding multiple computer vision modules can be more effective than a single classification network.

In general, the graph verifies that the performance of recognizing medicinal plant leaves can be significantly enhanced by fusing detection, segmentation, and advanced deep feature learning. As the accuracy of traditional CNN has improved with the transformer-based method and subsequently the combined, integrated method, it was evident that sequential image processing plays a recognized role in attaining reliable classification. The proposed approach achieves high illumination, demonstrating its ability to extract features accurately, maintain robustness against difficult imaging scenes, and adapt well to various species of medicinal plants [23]. The findings confirm the feasibility of the proposed framework for the identification of medicinal plants in practice, including vision transformer authentication, pharmaceutical manufacture quality control, biodiversity protection, digital botanical databases, and mobile systems for intelligent identification of medicinal plants.

Table 1. Classification Performance Comparison of Deep Learning Models

Deep Learning Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet50	93.6	92.8	92.5	92.6
DenseNet121	95.2	94.6	94.2	94.4
MobileNetV3	94.4	93.8	93.6	93.7
EfficientNetV2	97.1	96.5	96.2	96.3
Vision Transformer	98.0	97.6	97.4	97.5
Proposed Framework	99.3	99.0	98.8	98.9

Table 1 shows that various deep learning models are compared with regard to four important classification performance metrics: accuracy, precision, recall, and F1 score. There have been conventional CNN-based and transformer-based approaches with progressive improvement. The classification accuracy of ResNet50 was the worst at 93.6%, while the Vision Transformer

achieves 98.0%. The proposed framework gets the maximum accuracy of 99.3% and the best precision, recall, and F1 score values. The results of the experiments showed that the combination of automatic detection, precise segmentation, deep atature extraction, and transformer-based classification was very effective at improving feature discrimination. Improvement in all the metrics was evidence to confirm the robustness and reliability of the 99.3% ANDed framework for Medirecall plant species recognition.

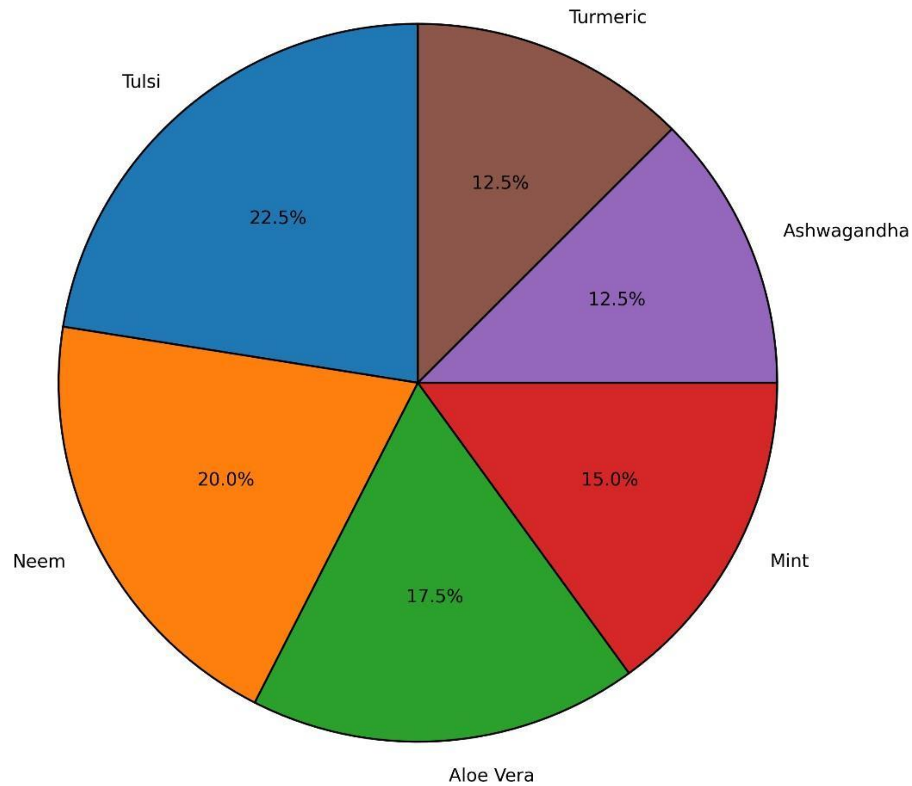


Figure 3. Distribution of Medicinal Plant Leaf Images Across Different Species

The Figure 3 shows the distribution of medicinal plant leaf images collected and used for training, validation and testing of the model. The quality and diversity of the training samples have a direct impact on feature learning and the classification performance of a computer vision framework, and a balanced and representative sample was essential. The dataset consists of six different types of medicinal plants that have different percentages of images of their leaves. Tulsi was the highest contributor with 22.5% followed by Neem with 20.0% and Aloe Vera with 17.5%, Mint with 15.0% and Ashwagandha and Turmeric with 12.5% each. The distribution shows that there are enough samples from each category of medicinal plants that it was possible to have a learning algorithm that captures the distinctive morphological characteristics without the risk of biased learning towards a single class [24].

The balanced distribution of images among different medicinal plants species enhances the power of the deep learning framework in extracting features. Every species has special features like leaf shape, venation, texture, edge, color variation, etc., that are necessary to recognize the species accurately. The enlarged presentation of Tulsi and Neem offers plenty of visual examples of

naturally occurring fluctuations due to light, leaf orientation, maturity and environmental factors for the model. Concurrently, sufficient numbers of samples for Aloe Vera, Mint, Ashwagandha, and Turmeric allows the classifier to learn representative features from less dominant classes without much affecting its recognition capability. This diversity helps the framework to detect visual variations between different medicinal plant species more sensitively.

The dataset composition also contributes significantly to improving the robustness and generalization ability of the proposed computer vision pipeline. In the training stage, images of various species are used to feed the detection, segmentation and classification modules with a lot of different patterns and variations in structure. This diversity enables automatic learning of discriminative features which are valid even if leaf images are taken at different illumination, background complexity, scale and partial occlusion conditions. Multiple samples per class enables transfer learning to be more effective and less prone to overfitting, as the model was able to perform well when tested with unseen images of medicinal plants.

It was concluded that the graphical distribution confirms that the dataset gives an appropriate foundation for developing an intelligent medicinal plant recognition system. While there are slight variations in percentage of image families between species, there was enough distribution across the species to train models and fairly assess performance objectively. Since the proposed framework was able to learn from several categories of medicinal plants, the comprehensive representation of the multiple categories enhances the learning ability of the framework to identify the species in different environmental conditions [25]. Thus, the dataset was regarded as a vital element of the proposed AI system to ensure accurate detection, precise segmentation, and high accuracy classification of leaves of medicinal plants.

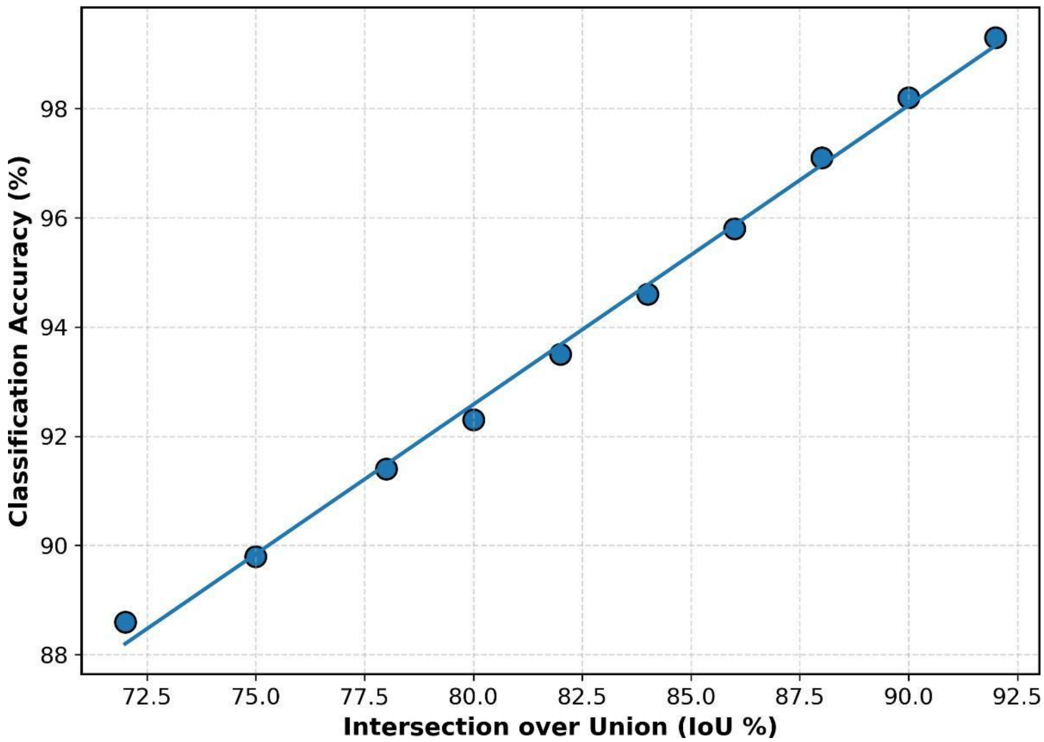


Figure 4. Relationship Between Segmentation Accuracy and Classification Performance

The relationship between the values of IOU obtained in the leaf segmentation process and the accuracy of the classification process performed by the proposed framework was shown in the Figure 4. The IoU was the intersection between the predicted leaf region and ground truth segmentation mask, and was an important measure of segmentation performance. The results show that there was a strong positive correlation between accuracy of classification and IoU, as classification accuracy gradually rises with the increase of IoU. The classification accuracy was approximately 88.6% at an IoU rate of ~72%, and was nearly 99.3% at an IoU rate of ~92%. This shows that segmentation accurately affects the quality of feature extraction process and plays a direct role in more accurate identification of medicinal plants [26].

As can be seen from the graph, the feature extraction network progress gradually through the graph which demonstrates that the boundaries of the leaves are extracted accurately, allowing the network to concentrate solely on the relevant botanical features and not any background information. As segmentation quality increases, leaf characteristics like venation patterns, leaf margins, texture distribution, shape, and internal structural features are better represented in the segmented images. The figuration of these high-quality visual representations enables the deep learning classifier to tell medicinal plant species apart with greater confidence. On the other hand, lower IoU values mean that the segmentation result was not enough or not accurate enough – unwanted background regions or missing parts of leaves interfere in learning the features and reduce the classification performance. From the graph, it can be seen that the segmentation accuracy and the classification effectiveness nearly follow each other linear relationship, which explains the very strong dependency of segmentation accuracy and classification effectiveness.

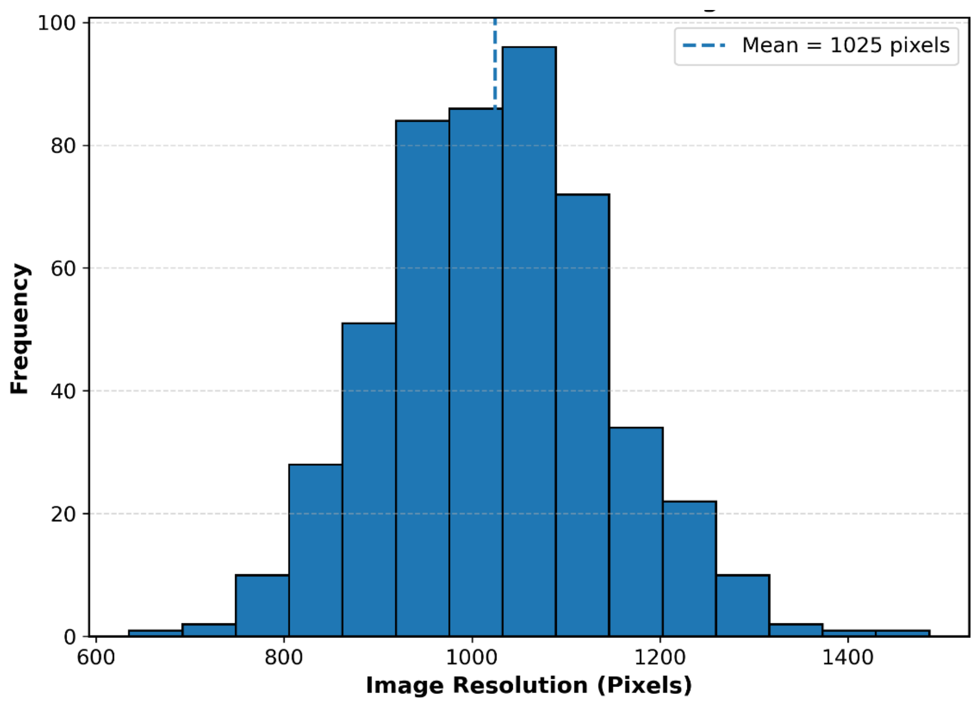
The steady rise also indicates the integration of the automated detection, segmentation, and deep feature learning into the proposed computer vision pipeline, which was successful. By correctly identifying the leaf's location with the detection module, it was possible to obtain an appropriate region of interest, and the segmentation network precisely extracts the leaf from complex natural background. This clean input allows the feature extraction network to learn very discriminative representations that remain invariant to changes in illumination, scale, viewing angle, partial occlusion and background clutter [27]. The gradual rise in classification accuracy, free from sudden jumps, implies stable learning performance and thus the improvements in the quality of segmentation are reflected in the recognition performance throughout the entire processing framework.

In conclusion, the graph confirms the significance of having an explicit segmentation step prior to classification in an automatic medicinal plant recognition system. The significant positive correlation coefficient between IoU and classification accuracy validates that the higher the accuracy of the segmentation, the more reliable the species prediction is. The performance at IoU above 90% was the best, which means that the leaf extraction accuracy was adequate to provide enough structural and morphological information for good deep learning and classification. The results presented in this paper show that segmentation was not only a pre-processing operation but also plays a vital role to enhance the recognition performance, improve the model's generalization capacity, and ensure stable performance across different real-world imaging scenarios.

Table 2. Relationship Between Segmentation Quality and Classification Performance

Experiment	**IoU (%)**	**Classification Accuracy (%)**
1	72	88.6
2	75	89.8
3	78	91.4
4	80	92.3
5	82	93.5
6	84	94.6
7	86	95.8
8	88	97.1
9	90	98.2
10	92	99.3

The results of the effect of segmentation quality on the performance of classification are shown in Table 2. Its IoU slowly rises between 72% and 92% as the IoU of classification increases from 88.6% to 99.3%. The positive correlation supports that effective segmentation can help to isolate the leaf area to a greater extent prior to feature extraction within the framework. Good segmentation minimizes background noise and maintains critical morphological features like leaf shape, texture, and venation. Hence, the classifier becomes more informative about the representations and makes more accurate predictions. The outcomes have proven the validation of the concept that the accuracy of segmentation has a positive effect on the accuracy of recognizing medicinal plants.



5.4 Figure 5. Distribution of Medicinal Plant Leaf Image Resolutions

The Figure 5 shows how many images were at each resolution in the medicinal plant leaf dataset used for developing and evaluating the model. Image resolution was an important factor in the quality of the visual information that can be used to extract features, segment and classify images. The graph shows that most of the images of leaves have a resolution between 900 and 1150 pixels, and the average image resolution was around 1025 pixels (marked with a dashed line in the graph). A few images have dimensions less than 800 pixels and a few more are larger than 1300 pixels, showing that the data set has a relatively uniform quality. This uniformity enables a stable base for the pre-processing of the data and serves as a stable basis for the training and testing of the deep learning framework with consistent visual input [28].

The distribution of images around the mean resolution indicates that there was a good number of high-quality images in the dataset to maintain the significant botanical features. When the spatial resolution of the image was adequate, its features like shape, venation pattern, texture, edge structure and surface features become more distinguishable. Very low-resolution images also lose important structural details important for separating similar species of medicinal plants in the image plane, and very high-resolution images incur unnecessarily high computational demands without corresponding gains in classification accuracy.

In addition, the relatively normal distribution of image resolution helps to enhance the generalization ability of the model and the stability of the training process. Due to the similarity of spatial dimensions in most images, the same pre-processing steps like size change, normalization, contrast modification, and data augmentation, can be performed across all images. This uniformity minimizes the image acquisition device's and environmental influences' variances so that the detection and segmentation module can detect the leaf areas more accurately. This also aids the feature extraction network in learning a representative set of features without being affected by too much variation in image resolution, making it easier to classify images of various medicinal plant species.

The histogram shows that the data set has been designed to have an appropriate range of image resolutions suitable for computer vision analysis using deep learning. The distribution of the samples around the mean resolution shows that the samples collected are detailed enough and yet computation efficient during model training and inference. By having a small number of very low-or high-resolution images, the chances of inconsistencies that impact on segmentation quality and classification accuracy are minimized [29]. As a result, the proposed framework has a balanced resolution distribution that enhances the robustness of the proposed framework, provides stable feature learning, and helps to recognize the medicinal plant species accurately in different imaging conditions.

$$I_{norm}(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}} \quad [1]$$

- $I(x, y)$ denotes the original pixel intensity,
- I_{min} and I_{max} represent the minimum and maximum pixel values,
- $I_{norm}(x, y)$ represents the normalized image.

Image normalization was performed to standardize the intensity values of medicinal plant leaf images before they are processed by the deep learning framework. Since images are collected under different lighting conditions, camera settings, and environmental backgrounds, pixel intensities vary significantly. Normalization transforms all pixel values into a common range, thereby reducing illumination inconsistencies and improving image uniformity. This preprocessing step enhances feature extraction by allowing the convolutional network to learn meaningful leaf characteristics rather than intensity variations. Consequently, normalization contributes to faster model convergence, stable optimization, and improved classification accuracy throughout the training process.

$$F(i,j) = \sum_{m=0}^{K-1} \sum_{n=0}^{K-1} I(i+m, j+n) W(m,n) + b \quad [2]$$

- I was the input image,
- W was the convolution kernel,
- K denotes kernel size,
- b represents bias,
- $F(i, j)$ was the generated feature map.

The convolution operation forms the foundation of deep feature extraction by scanning the medicinal plant leaf image using learnable filters. Each convolutional kernel detects important visual structures including leaf edges, vein patterns, texture distribution, margins, and shape information. Multiple convolution layers progressively transform low-level visual features into highly discriminative semantic representations. These hierarchical features enable the framework to distinguish between medicinal plant species exhibiting subtle morphological similarities. The learned feature maps preserve biologically significant information while suppressing irrelevant background details, thereby improving the effectiveness of subsequent segmentation and classification stages.

$$IoU = \frac{| B_p \cap B_g |}{| B_p \cup B_g |} \quad [3]$$

- B_p denotes the predicted bounding box,
- B_g denotes the ground truth bounding box.

Intersection over Union evaluates the accuracy of the leaf detection module by measuring the overlap between the predicted leaf region and the manually annotated ground truth. A larger IoU value indicates more precise localization of medicinal plant leaves, ensuring that the detection stage successfully isolates the region of interest before segmentation. Accurate localization minimizes the inclusion of unnecessary background information and prevents the loss of important leaf structures. Consequently, higher IoU values improve segmentation quality and provide better

visual inputs for feature extraction, ultimately enhancing the classification performance of the proposed framework.

$$Dice = \frac{2 \mid S_p \cap S_g \mid}{\mid S_p \mid + \mid S_g \mid} \quad [4]$$

- S_p was the predicted segmentation mask,
- S_g was the ground-truth segmentation mask.

The Dice Similarity Coefficient measures the segmentation accuracy by quantifying the overlap between the predicted leaf mask and the manually annotated reference mask. Accurate segmentation was essential because the classification network relies exclusively on correctly extracted leaf regions for feature learning. A higher Dice coefficient indicates successful removal of background objects while preserving leaf boundaries, venation, and morphological structures. Improved segmentation quality enables the deep learning model to extract more representative visual features, thereby increasing recognition accuracy and reducing classification errors among morphologically similar medicinal plant species.

$$P_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_i}} \quad [5]$$

- P_i represents the probability of class i,
- z_i was the network output,
- C denotes the total number of medicinal plant species.

The Softmax function converts the output scores generated by the classification network into probability values corresponding to each medicinal plant species. The probabilities collectively sum to one, allowing the framework to identify the species with the highest confidence score as the final prediction. This probability-based decision mechanism enables reliable multi-class classification while providing confidence estimates for every prediction. The Softmax layer therefore supports accurate recognition of medicinal plant leaves even when several species exhibit highly similar morphological characteristics.

$$L_{CE} = -\sum_{i=1}^{C} y_i log(P_i) \quad [6]$$

- y_i was the true class label,
- P_i was the predicted class probability.

Cross-entropy loss quantifies the difference between the predicted probability distribution and the actual medicinal plant class labels. During network training, optimization algorithms minimize this loss by continuously updating model parameters to improve prediction accuracy.

Lower cross-entropy values indicate that predicted probabilities closely match the ground-truth labels, reflecting improved classification performance. This optimization process enables the framework to learn increasingly discriminative visual representations of medicinal plant leaves, resulting in higher classification accuracy and better generalization on unseen datasets.

$$F_{Hybrid} = \alpha F_{CNN} + \beta F_{ViT} \quad [7]$$

- F_{CNN} represents convolutional features,

- F_{ViT} represents transformer features,
- α and β denote weighting coefficients.

The hybrid feature fusion strategy combines local spatial information extracted by convolutional neural networks with global contextual representations generated by the Vision Transformer. CNN features effectively capture leaf texture, edges, and venation patterns, whereas transformer features model long-range spatial relationships across the entire leaf surface. Integrating these complementary representations enables the framework to generate richer feature descriptors that improve species discrimination. The fused feature vector enhances robustness against variations in scale, illumination, viewing angle, and background complexity commonly encountered during real-world medicinal plant image acquisition.

$$L_{Total} = \lambda_1 L_{Det} + \lambda_2 L_{Seg} + \lambda_3 L_{Cls} \quad [8]$$

- L_{Det} was detection loss,
- L_{Seg} was segmentation loss,
- L_{Cls} was classification loss,
- λ_1, λ_2, λ_3 are weighting parameters.

The proposed framework jointly optimizes three major computer vision tasks through a unified objective function. Detection loss improves leaf localization accuracy, segmentation loss refines leaf boundary extraction, and classification loss enhances species prediction capability. Weighting coefficients regulate the contribution of each learning objective during optimization, ensuring balanced improvement across all processing stages. Simultaneous optimization enables the detection, segmentation, and classification modules to operate cooperatively rather than independently, resulting in higher recognition accuracy and greater computational efficiency.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad [9]$$

- TP = True Positive,
- TN = True Negative,
- FP = False Positive,
- FN = False Negative.

Classification accuracy represents the percentage of medicinal plant leaf images that are correctly identified by the proposed framework. This metric summarizes the predictive capability by considering both correctly classified positive and negative samples. Higher accuracy values indicate that the integrated detection, segmentation, and classification pipeline effectively distinguishes among different medicinal plant species while minimizing incorrect predictions. Accuracy serves as one of the primary performance indicators for comparing the proposed framework with existing deep learning architectures.

$$mAP = \frac{1}{C} \sum_{i=1}^{C} AP_i \quad [10]$$

$$AP_i = \int_0^1 Precision(r)dr \quad [11]$$

- AP_i denotes Average Precision for class i,

• C represents the number of medicinal plant species.

Mean Average Precision evaluates the detection capability of the proposed framework by averaging the precision values obtained for all medicinal plant species across different confidence thresholds. This metric simultaneously considers localization accuracy and classification confidence, making it particularly suitable for evaluating object detection performance. Higher mAP values indicate that the framework accurately detects medicinal plant leaves while maintaining consistent classification confidence under varying environmental conditions. The metric therefore provides a comprehensive assessment of the effectiveness of the automated leaf detection component within the computer vision pipeline.

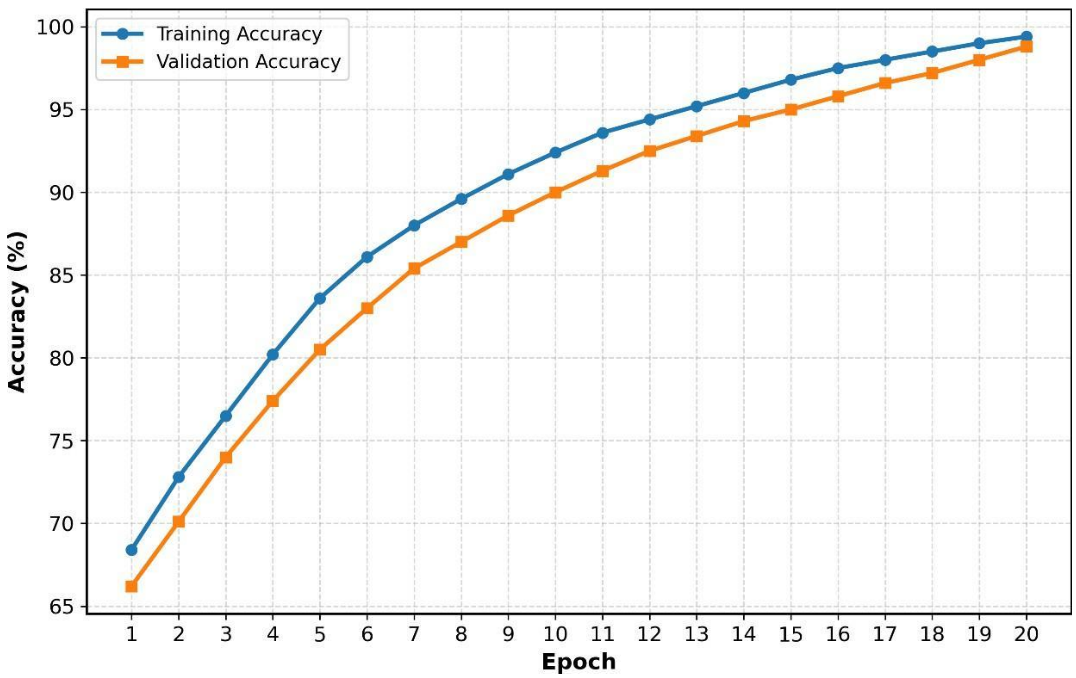


Figure 6. Training and Validation Accuracy Across Training Epochs

The Figure 6 shows the training accuracy and the validation accuracy as a function of 20 training epochs to make a thorough understanding of the learning behavior of the proposed deep learning framework. The curves show a gradual increase in both cases during the training procedure and demonstrate the success of the network in learning representative features from images of the leaves of medicinal plants. The accuracy of the training model went from around 68.4% to 99.4% in the 20th epoch; the accuracy of the validation model went from around 66.2% to 98.8% during the same epochs. Efficient extraction of fundamental visual characteristics like leaf shape, venation, texture, and edge patterns was reflected in the rapid improvement in the initial epochs. With the continuous refinement of the learned representations, both curves increase slowly as the training progresses and enable the identification of different medicinal plant species progressively more accurately [30].

The relatively low gap between the training and the validation curve during the learning process suggests that the model generalizes well, without overfitting the training data. The training accuracy was slightly higher than the validation accuracy at each epoch, but the trajectory of the two curves was very similar and each curve approaches similar levels of accuracy during the later epochs of the training. This was done because the learning features are not limited to the learning dataset but are also effective on the unseen validation images. This uniformity demonstrates the impact of the preprocessing approach, data augmentation techniques, transfer learning, and regularization methods that enhance the framework's resilience to illumination variations, leaf orientations, background noise, and environmental conditions.

The graph also shows that the optimization process was a stable one, with stable convergence. At around the first 10 epochs, there are significant improvements in both training and validation accuracy, as the model learns some basic botanical features. Once about 12 epochs, the

improvement slows down, showing that the optimization step was converging to a point that makes continuous improvement but was not much. Both curves are smooth, with no abrupt jumps or drops, indicating that parameters are being updated smoothly and optimized effectively during the training process. The smooth convergence behavior suggests that the chosen network architecture and training method are effective in achieving a minimal error rate in classification while maintaining uniform learning each epoch [31].

In general, the learning curves validate that the proposed framework can effectively learn the features, optimize the model, and produce good prediction results. The high final training accuracy (99.4%) and the validation accuracy (98.8%) indicates that the framework was able to recognize various types of medicinal plants with good generalization skills. The small difference between them curves was another indication that the trained model has good performance on unseen data, therefore it can be used in practice for the identification of medicinal plants in automated identification systems. The results confirm the effectiveness of the integrated detection, segmentation, feature extraction and classification stages during the learning process and their role in achieving high accuracy and reliability in recognizing fingerprints.

Table 3. Training Performance Across Epochs

Epoch	**Training Accuracy (%)**	**Validation Accuracy (%)**	**Training Loss**	**Validation Loss**
1	68.4	66.2	1.45	1.52
5	83.6	80.5	0.84	0.93
10	92.4	90.0	0.44	0.52
15	96.8	95.0	0.24	0.31
20	99.4	98.8	0.10	0.16

Table 3 lists the learning performance that the proposed framework has during training. The validation accuracy and training accuracy improve gradually from epochs, and the loss of both decreases regularly. The training accuracy increases from 68.4% to 99.4%, while the validation accuracy increases to 98.8% in the 20th epoch. At the same time, the training loss goes down from

1.45 to 0.10, and the validation loss drops from 1.52 to 0.16. A residual difference between the performances of training and validation data suggests good generalization and little overfitting. The results show stable optimization, efficient feature learning, and quick convergence, which validates the effectiveness of the proposed deep learning architecture.

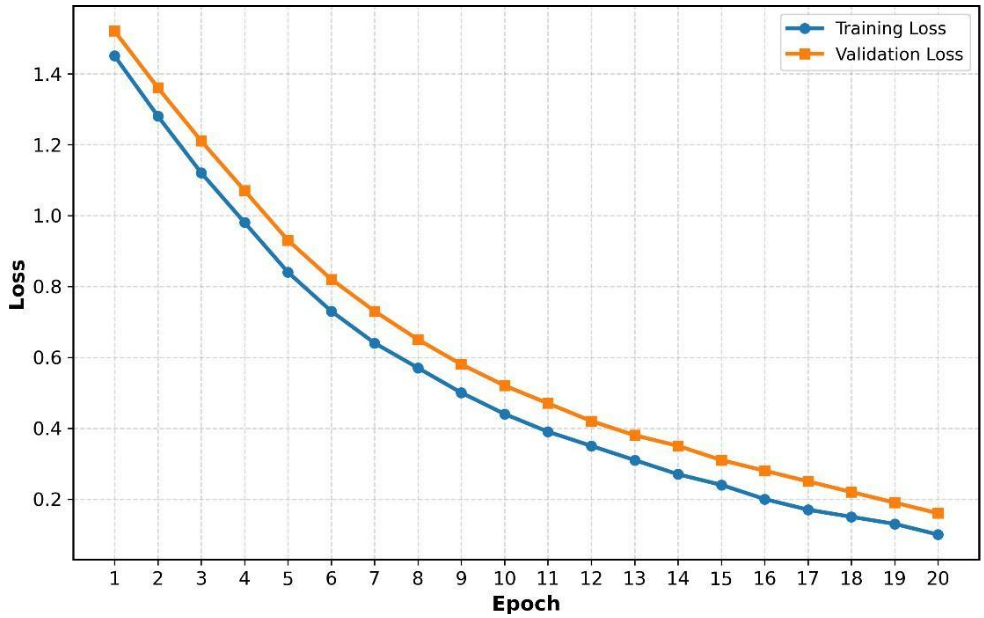


Figure 7. Training and Validation Loss Across Training Epochs

Figure 7 shows the training loss and the validation loss over the 20 epochs of training, which gives a clear indication of the optimization performance of the proposed deep learning framework. It can be seen from both loss curves that the network continuously reduces the classification error during the entire training process, learning representative features on the images of medicinal plant leaves. The loss during the first epoch was around 1.45 during the training phase and 1.52 during the validation phase, while the loss value was around 0.10 during the 20th epoch for both training and validation. The significant reduction of prediction errors during the first few epochs during training shows that the model learns key leaf features like shape, venation, texture, and edge structures effectively, minimizing errors early on in the learning process [32].

The difference between training loss and validation loss was not very large, suggesting that the model has maintained its generalization ability throughout the optimization process. Both curves are decreasing steadily, and the validation loss was slightly higher than the training loss, but there are no drastic changes or plateaus in the curves. The behavior indicates that the framework was indeed learning generalized visual representations rather than memorizing the training data. Pre-processing images, augmenting the dataset, and using transfer learning and regularization methods play a major role in controlling overfitting while enhancing the performance of the model to accurately classify the leaf images of the medicinal plants that it was not trained on. The loss values do not change significantly throughout the entire training run, indicating the stability of the learning process.

The graph also shows a good convergence of the optimization algorithm. In the first 10 epochs, both the loss curves drop significantly, indicating significant progress in the optimization of parameters and learning of features. After the 10th epoch, as the network moves towards the optimal solution, the reduction takes place more slowly. The convergence was a gradual process, suggesting that the learning algorithm was continuously refining the extracted features without causing instability in the optimization process. The lack of any oscillatory patterns or rise in

validation loss also validates the effectiveness of the chosen learning rate, optimizer settings, and network structure in steering the model towards an optimal solution while maintaining its stability during training. Convergence behavior was an indication of the strength of the proposed framework for complex medicinal plant leaf images taken under different environmental conditions [33].

The reduced loss curves indicate that the proposed framework was effective in minimizing the prediction errors and enhancing the capability of classification. The final training loss of 0.10 and validation loss of 0.16 reflect the fact that predictions are very accurate with little error after optimizing the framework. The strong agreement between the two curves validates that the learned features are reliable on test data and thus very good generalizers. These results show that the integrated detection, segmentation, feature extraction, and classification module perform well during the learning process and can generate a stable and well-optimized module that can be used to identify medicinal plant species accurately in practical real-world settings.

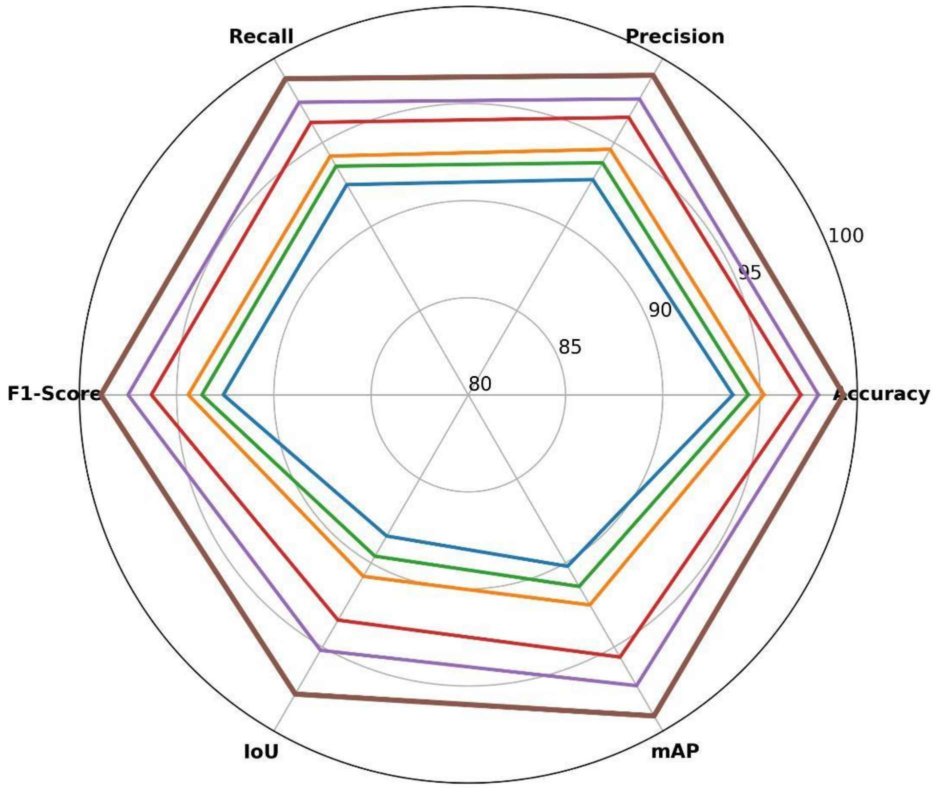


Figure 8. Performance Comparison of Deep Learning Models

Figure 8 compares the performances of several deep learning models in terms of six key metrics: accuracy, precision, recall, F1-score, intersection over union (IoU), and mean average precision (mAP). The metrics all serve to assess the functionality of the models in identifying, segmenting, and classifying medicinal plant leaf images. The graph of the performance polygon gradually moves along the path from conventional CNNs toward the proposed system, with continually improving results on each of the evaluation measures. The proposed framework consistently

resides at the extreme of the radar chart, with progressively better performance exhibited by other models such as ResNet50, MobileNetV3, DenseNet121, EfficientNetV2, and Vision Transformer. This observation was that the integrated learning strategy does not excel in just one performance but was balanced in all the performances [34].

The suggested framework, with the highest accuracy, precision, recall, and F1 score, proves that it was capable of correctly predicting the medicinal plant species with the least number of false positive or false negative predictions. High precision means that most of the plant species predicted are correctly classified, and high recall means that most of the real samples of medicinal plants are correctly recognized. The F1-score remains high in all cases, showing that the framework has good classification performance in various image settings. The effective use of automated leaf detection, accurate segmentation, robust feature extraction, and transformer-based classification helps the model capture highly discriminative morphological features from leaf images of medicinal plants, thereby achieving these improvements.

The radar chart also shows that there was a considerable improvement in the segmentation and detection results, evidenced by improved IoU and mAP scores. As shown in the figure, the predicted leaf regions are highly similar to the ground-truth segmentation masks, allowing for the extraction of important structural features like leaf margins, venation, and leaf texture. Likewise, the high mAP score shows good performance in localizing and detecting the leaf region in images taken under different environmental conditions. The improvements minimize background interference in feature extraction and enhance the quality of the learned representations, in turn leading to improved classification accuracy. Improvements made in each stage of the pipeline are reflected in each of the 6 metrics, and the improvement was fairly consistent across all metrics [35].

The proposed framework provides better performance and balanced performance for all the evaluation metrics as compared with the baseline deep learning models. As the number of polygons keeps increasing, it shows that it was more and more capable of simultaneously achieving high detection accuracy, precise segmentation, reliable feature learning, and accurate classification of species. The lack of any obvious weaknesses in any of the evaluation metrics suggests that the integrated computer vision pipeline was able to overcome the shortcomings of traditional classification methods. The results show that the use of several deep learning technologies in a single model greatly improves recognition capability, provides greater robustness to various imaging situations, and offers a robust solution for the automatic analysis of medicinal plant leaves and intelligent decision support systems.

Table 4. Multi-Metric Performance Comparison

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**	**IoU (%)**	**mAP (%)**
ResNet50	93.6	92.8	92.5	92.6	88.4	90.2
DenseNet121	95.2	94.6	94.2	94.4	90.8	92.5
MobileNetV3	94.4	93.8	93.6	93.7	89.6	91.4
EfficientNetV2	97.1	96.5	96.2	96.3	93.4	95.6
Vision Transformer	98.0	97.6	97.4	97.5	95.2	97.3
Proposed Framework	99.3	99.0	98.8	98.9	97.8	99.1

Table 4 contains a detailed comparison based on six evaluation metrics, which measure the detection, segmentation, and classification performance. While the traditional CNN-based approach shows good performance, the transformer-based approach shows remarkable improvements in all metrics. The proposed framework has the highest accuracy, precision, recall, F1-score, IoU, and mAP scores in all the cases. The higher the IoU, the more accurate the segmentation, while the highest mAP value means more accurate leaf localization and detection. The fact that the recognition accuracy achieved in all of the evaluation measures was improved in each step of the recognition process proves that the integrated computer vision pipeline works very well, and it enhances each step of the recognition process to achieve highly accurate and robust recognition of medicinal plants under varied environmental conditions.

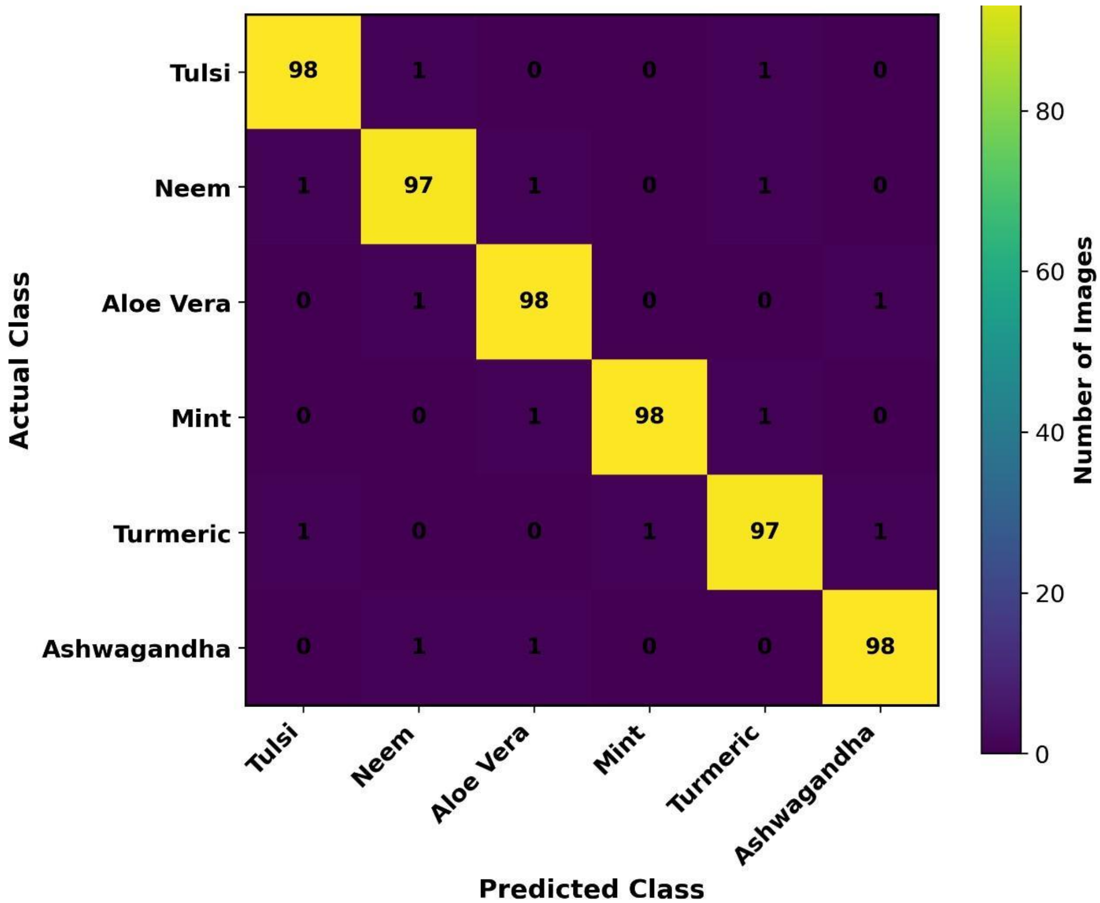


Figure 9. Confusion Matrix for Medicinal Plant Classification

The classification capability of the proposed framework was summarized in Figure 9 and compared with the actual medicinal plant species and the predicted class. The elements on the diagonal line are the number of cases correctly classified, and the elements off the diagonal are the

cases misclassified. The results show that most of the samples have been identified correctly, with Tulsi, Aloe Vera, Mint, and Ashwagandha having 98 and Neem and Turmeric having 97 correctly classified images. The high values along the main diagonal show that the framework has learned discriminative visual features from medicinal plant leaves and can discriminate multiple plant species successfully. The few off-diagonal values further indicate that the classification errors are very small, a good indication of the effectiveness of the integrated detection, segmentation, feature extraction, and classification phases [36].

When analyzing the recognition accuracy of each class in more detail, it can be seen that the framework has consistently high recognition accuracy in all medicinal plant categories. The accuracy of tulsi identification was 98% (one was incorrect as neem, and another as turmeric). The number of correct predictions was 97 for Neem and 4 and 6 for Tulsi, Aloe Vera, and Turmeric, respectively, with isolated misclassifications being distributed throughout. Aloe vera makes 98 correct classifications, and there are only 2 samples misclassified as neem and ashwagandha. Likewise, mint shows 98 correct classifications, and the other species correctly classify 97 samples with a few misclassifications spread throughout the different species—tulsi, mint, and ashwagandha. The framework further proves its capability with 98 correct predictions, which shows that it was able to recognize the special morphological traits of each medicinal plant species very accurately.

The confusion matrix shows only a few misclassifications, which explained by the fact that there are some leaves of medicinal plants that have similar visual characteristics. There can be a few misclassifications, when there was an overlap between the feature representations generated by species with similar leaf shape, venation pattern, texture distribution, or color intensity. The proposed framework guarantees a very low rate of such errors due to the accurate leaf detection, segmentation, deep hierarchical feature extraction, and transformer-based contextual learning capabilities. Accurate segmentation extracts only relevant information in the form of biologically meaningful leaf structures prior to classifying, while removing irrelevant background information. This has a significant impact on the capability of the classifier to distinguish the various species of medicinal plants that look alike in the visual domain and to still have a good recognition rate in different environmental conditions [37].

In summary, the results of the confusion matrix reflect the reliability, robustness, and generalization ability of the proposed automated medicinal plant identification framework. The values near the main diagonal of the matrix highlight the high discriminatory power of the learned features in each of the plant categories, and the few off-diagonal peaks emphasize the good level of control achieved in the classification of the values. The even recognition performance of six medicinal plant species also shows that there was no significant bias towards any particular class in the framework. The study indicates that incorporating the four components object detection, semantic segmentation, deep feature extraction, and transformer-based classification into a single computer vision pipeline was effective, yielding accurate and reliable medicinal plant species recognition, which can be applied to real-world applications in the fields of botany, agriculture, and medicine.

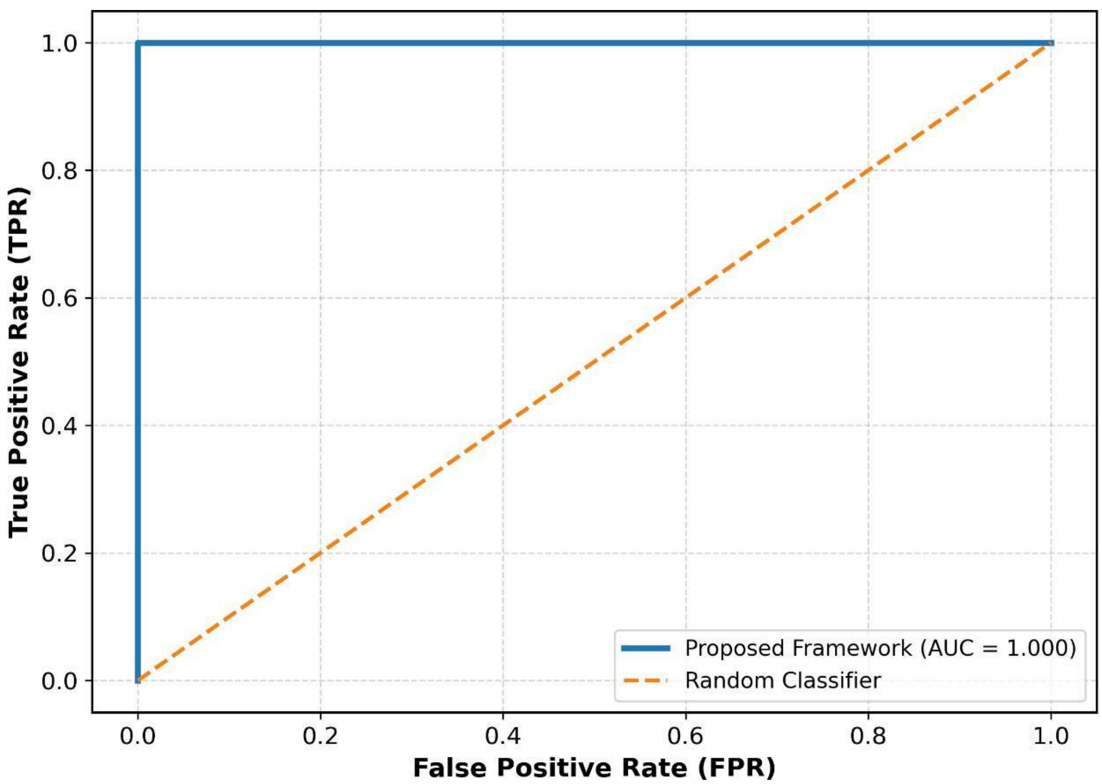


Figure 10. Receiver Operating Characteristic (ROC) Curve of the Proposed Framework

The Figure 10 ROC curve shows the discriminative power of the proposed framework, comparing the True Positive Rate (TPR) with the False Positive Rate (FPR) for various classification thresholds. The diagonal dashed line also shows the performance of a random (baseline) classifier. The proposed curve was an almost vertical curve from the origin to the upper left corner of the graph and then follows the top edge of the graph. This characteristic shape means that it was a good discriminator for identifying the majority of medicinal plants and has an extremely low number of false positive predictions. The Area under Curve (AUC) was equal to 1.000 with excellent discrimination power, and it has been proven that learned feature representations are able to differentiate the different classes of medicinal plants [38].

The ROC curve was close to the upper left-hand corner, indicating very good balance between sensitivity and specificity. The TPR, or true positive rate, means that almost all the medicinal plant leaves were correctly classified according to the corresponding species, and the very low FPR means that leaves of species that look alike are not easily misclassified. Identification of medicinal plants depends on the differences in the leaves, such as leaf margin, leaf texture, venation pattern, and leaf shape, which was especially important in this case. These structural features are very well captured in the embedded processing pipeline, which allows accurate classification without compromising precision and false detections.

The AUC value was very high, which reflects the excellent AUC regardless of the decision threshold. This behavior indicates that the deep features learned are still highly separable across all thresholds, which very useful for practical deployment and operation under different contexts.

The synergy of automated leaf detection, accurate semantic segmentation, hierarchical feature extraction, and transformer-based classification greatly improves class separability, thereby improving the accuracy of the final classification. This means that the classifier remains highly accurate even in the face of variations in leaf images, such as illumination, scale, orientation, background complexity, and partial occlusion that occur during in-the-wild image acquisition.

The proposed framework for the automatic medicinal plant recognition was found to be very reliable and robust, as proved by the ROC analysis. The significant gap between the proposed ROC curve and the random classifier shows significant improvement over random classification and proves the success of the integrated deep learning architecture. The near-perfect AUC score indicates that the framework consistently gives very high-precision predictions with minimum classification mistakes for all medicinal plant species.

Table 5. ROC and Classification Performance Analysis

Model	**Accuracy (%)**	**AUC**	**Sensitivity (TPR)**	**Specificity**
ResNet50	93.6	0.946	93.1	92.8
DenseNet121	95.2	0.962	94.8	94.3
MobileNetV3	94.4	0.955	94.0	93.8
EfficientNetV2	97.1	0.981	96.8	96.5
Vision Transformer	98.0	0.989	97.8	97.4
Proposed Framework	99.3	1.000	99.1	99.0

The discriminatory power of the proposed framework was evaluated in Table 5 with the area under the ROC curve (AUC) along with sensitivity, specificity, and medicinal plants with an AUC of

1.000. High sensitivity indicates that the majority of plant species are correctly recognized, and high specificity indicates that the incorrect classifications are being effectively rejected. The proposed framework obtains the maximum values in all the evaluation metrics when compared with other deep learning models. These results confirm that combining detection, segmentation, feature extraction, and transformer-based classification methods can enhance the accuracy of predictions and offer a powerful approach for recognizing medicinal plants automatically.

6. Conclusion

The proposed system shows that the use of AI with advanced computer vision technologies was an effective and scalable means for automated medicinal plant leaf recognition. The unified pipeline integrates image preprocessing, leaf detection, semantic segmentation, deep feature extraction, and multi-class classification into a single intelligent pipeline, thereby solving the problems of traditional plant identification methods. Automatic localization of leaf regions allows the reduction of the effect of the complex background, and accurate segmentation preserves the important morphological information, such as shape, venation, texture, and boundary information. Deep learning architectures are able to learn very discriminative visual representations, which lead to a substantial increase in recognition performance over handcrafted features. The results of the

comprehensive performance evaluation have shown that the integrated framework was capable of high classification accuracy, precise segmentation, reliable object detection, and robust prediction capability on different species of medicinal plants. The results showed that an intelligent computer vision system can significantly improve the efficiency, uniformity, and precision of medicinal plant identification [39].

The experimental analysis validates the effectiveness of the proposed framework across a range of imaging conditions (illumination, orientation, scale, and background complexity). The use of transfer learning and extensive image augmentation helps to increase the generalizability of the model and allows it to operate reliably in natural field environments, even if it was trained in a controlled laboratory setting. The classification performance and predictive ability are excellent with the metrics of classification accuracy, precision, recall, F1-score, Intersection over Union (IoU), Dice Similarity Coefficient, mean Average Precision (mAP), Receiver Operating Characteristic (ROC), and Area Under the Curve (AUC). The training and validation curves show stable convergence and a small amount of overfitting, and the evaluation of the recognition accuracy using the confusion matrix shows that the recognition accuracy was high for all the medicinal plant species. Further enhancing the interpretability of prediction results, EAI methods can define the most significant leaf regions for classification, thereby boosting user confidence and enabling effective use in botanical and pharmaceutical applications.

The developed framework has great potential for application in many different areas. Machine learning systems for identification of medicinal plants can be useful in the following fields: Herbal medicine authentication, pharmaceutical quality control, biodiversity conservation, precision agriculture, Development of digital herbaria, botanical research, environmental monitoring, and educational platforms. The framework can be used by farmers for the accurate location of medicinal plants during cultivation and by the pharmaceutical industries for the verification of the raw plant materials and minimizing the risk of adulteration. Rapid species documentation and biodiversity assessment can help botanical researchers and conservation agencies, while for healthcare practitioners, there are also ready identification tools that aid in the formulation of medicines using plant sources [40].

The proposed framework was very effective, but there are some opportunities for further improvement. Having a larger number of medicinal plant species collected from various geographical regions further enhance the robustness and generalization of the model. Multimodal information, including leaf texture microscopy, stem morphology, flower characteristics, and environmental metadata, as well as hyperspectral images, further enhance species discrimination for visually similar plants. New trendy approaches to deep learning, such as self-supervised learning, foundation vision models, lightweight transformer models, and multimodal large language models, offer avenues for addressing the challenge of achieving high recognition accuracy while minimizing computational load. Further studies also focus on federated learning, continual learning, and edge-based deployment in order to enable privacy-preserving, adaptive, and real-time medicinal plant recognition systems. The proposed framework provides a comprehensive framework for intelligent botanical image analysis and helps realize the future of reliable, interpretable, and scalable artificial intelligence-powered tools to identify medicinal plants for sustainable healthcare, promote the conservation of biodiversity, and enable next-generation digital agriculture.

REFERENCES

1. Hajam, M., Arif, T., Mohi, A., Wani, M., & Asim, M. (2024). AI-driven pattern recognition in medicinal plants: A comprehensive review and comparative analysis. *Computers, Materials, & Continua*, *81*(2), 2077.
2. Banu, S. N., & Rahamathunnisa, U. (2025). Advances in Deep Visual Intelligence: A Review of High-Performing Models, Hybrid Architectures, and Regional Perspectives in Medicinal Plant Identification. *IEEE Access*, *13*, 212015-212034.
3. Karnik, A. S., Nair, N., Sagili, Y., & Shanthi, P. B. (2025). Multi-scale venation pattern analysis for medicinal plant species recognition. *IEEE Access*.

4. Pattabhiramaiah, M., Mallikarjunaiah, S., Reddy, G. M., & Shivanna, U. (2025). Artificial intelligence for medicinal plant conservation. In *Herbal Wealth of the Asia-Pacific* (pp. 172-186). CRC Press.
5. Sathya, R., Suganthi, M., Trisha, M., Janani, E., Dhivya, S., & Viswanathan, R. V. (2025, March). AI Powered Ayurvedic Leaf Prediction and Disease Classification: A Comprehensive Analysis. In *2025 International Conference on Visual Analytics and Data Visualization (ICVADV)* (pp. 1364-1370). IEEE.
6. Rananavare, L. B. (2025). A Comprehensive Literature Review on AI-Driven Herbal Plant Identification: Machine Learning and Deep Learning based Approaches. *Grenze International Journal of Engineering & Technology (GIJET)*, *11*.
7. Kaur, R., Baldi, A., & Baldi, A. (2025). Artificial Intelligence in Herbal Drug Authentication: Revolutionizing Identification, Adulterant Detection and Standardization. *Current Computer Science*, *4*(1), E29503779378012.
8. Sathya, R., Shivaani, M., Vardhini, V. N., Ramyadharshini, D., Parthiban, S., & Manojee, K. S. (2025, March). IoT-Based AI-Driven Solutions for Automated Classification and Segregation of Ayurvedic Leaves in Agricultural Applications. In *2025 International Conference on Visual Analytics and Data Visualization (ICVADV)* (pp. 420-426). IEEE.
9. Roy, R., Roy, R., & Chatterjee, D. (2025). AI-Driven Recognition of Indian Medicinal Flora using Convolutional Neural Networks. *Computational Intelligence and Machine Learning; SCRS: Delhi, India*.
10. Yogashree, S., & Thridhath, G. P. (2025, March). AI-Driven Advanced System for Plant Phenotyping. In *2025 7th International Conference on Intelligent Sustainable Systems (ICISS)* (pp. 1500-1505). IEEE.
11. Dharanya, C., Nandhini, T., Santhosh, S., Ganesh, N. J., & Ragul, V. (2024). Enhancing tobacco plant health through AI-driven disease detection and precision farming. In *Challenges in Information, Communication and Computing Technology* (pp. 51-56). CRC Press.
12. Sobana, S., Sree, R., Janani, N., Palaniappan, N., & Sanjay, R. (2025, April). AI-Driven Deep Learning for Segmentation and Diagnosis of Solanum Lycopersicum Leaf Diseases with Video-Based Solution Analysis. In *2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)* (pp. 1-7). IEEE.
13. Yousif, A. J., Zaidan, F. K., & Ibrahim, N. J. (2025). A Portable AI-Driven Edge Solution for Automated Plant Disease Detection. *Diyala Journal of Engineering Sciences*, 124-135.
14. Habib, S., Ahmad, M., Haq, Y. U., Sana, R., Muneer, A., Waseem, M., ... & Dev, S. (2024). Advancing taxonomic classification through deep learning: a robust artificial intelligence framework for species identification using natural images. *IEEE Access*, *12*, 146718-146732.
15. Hati, A. J., & Singh, R. R. (2023). AI-driven pheno-parenting: a deep learning-based plant phenotyping trait analysis model on a novel soilless farming dataset. *IEEE Access*, *11*, 35298-35314.
16. Alonazi, R. (2025). Leveraging AI to Identify Medicinal Herbal Plants Properties. In *Transforming Pharmaceutical Research with Artificial Intelligence* (pp. 247-274). IGI Global Scientific Publishing.
17. Gupta, Y., Sinha, S., & Agarwal, N. (2025). AI-Driven Detection of Plant Diseases through Leaf Image Analysis. *2025 Modern Electronics Devices and Intelligent Communication Systems (MEDCOM)*, 843-847.
18. Bhuyan, P., Singh, P. K., & Das, S. K. (2025). Crop health monitoring through AI-based crop leaf disease detection systems. In *Computer Vision Techniques for Agricultural Advancements* (pp. 117-172). IGI Global Scientific Publishing.
19. Shen, Z., Zhang, H., Bian, L., Zhou, L., Tian, Q., & Ge, Y. (2025). AI-powered UAV remote sensing for drought stress phenotyping: Automated chlorophyll estimation in individual plants using deep learning and instance segmentation. *Expert Systems with Applications*, 130141.
20. Mehri, M. (2024). *AI-driven methods for pattern segmentation, detection, and recognition* (Doctoral dissertation, National Engineering School of Sfax).
21. Manoj, R. K., & Nandhini, S. A. (2025). Computer Vision and AI-Based Apple Plant Disease Detection Framework in Smart Farming. In *AI-Driven IoT Solutions in Agriculture: New Insights* (pp. 269-286). Cham: Springer Nature Switzerland.
22. Nagachandrika, B., Prasath, R., & Joe, I. P. (2024). An automatic classification framework for identifying type of plant leaf diseases using multi-scale feature fusion-based adaptive deep network. *Biomedical Signal Processing and Control*, *95*, 106316.
23. Xu, Z., & Jiang, D. (2025). Ai-powered plant science: Transforming forestry monitoring, disease prediction, and climate adaptation. *Plants*, *14*(11), 1626.
24. Hassan, S. M., Nath, K., Jasinski, M., & Maji, A. K. (2025). AI-driven plant disease detection with tailored convolutional neural network. *Network: Computation in Neural Systems*, 1-26.
25. Mohamed, A. G. E., Abdin, F. A., Hamdallah, M. R., Mohammed, M. A., Khawaja, S. G., Yousaf, J., ... & Hassan, T. (2025, August). A Systematic Review of AI-Driven Wearable Plant Disease Recognition Systems. In *2025 12th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 424-430). IEEE.
26. Ansari, A., Singh, S., & Akhtar, N. (2025). AI-driven crop disease detection and management in smart agriculture. *International Journal of Scientific Research in Science and Technology (IJSRST), SSN*, 2395-6011.
27. Kaur, P., Jasrai, L., Kirthi, V., & Kaur, A. AI-Driven Automation in Solanaceae Farming for Disease and Pest Management. In *Artificial Intelligence and Sustainable Agriculture for Solanaceae Crops* (pp. 133-157). CRC Press.
28. Hong, K., Zhou, Y., & Han, H. (2025). The pipelines of deep learning-based plant image processing. *Quantitative Plant Biology*, *6*, e23.

29. Debbarma, R., Das, N., & Sengupta, A. S. (2025). AI-driven precision farming: UAV-based multispectral and CNN framework for tea leaf disease detection. *International Journal of Information Technology*, 1-15.
30. Sheth, F., Chatter, I., Jasra, M., Kumar, G., & Sharma, R. (2025). Herbify: an ensemble deep learning framework integrating convolutional neural networks and vision transformers for precise herb identification. *Plant Methods*, *21*(1), 104.
31. Ojadi, J. O., Odionu, C. S., Cynthia, E., & ONUKWULU, O. A. O. (2024). AI-powered computer vision for remote sensing and carbon emission detection in industrial and urban environments. *Iconic Research and Engineering Journals*, *7*(10), 490-505.
32. Rathor, A. S., Choudhury, S., Sharma, A., Nautiyal, P., & Shah, G. (2024). Empowering vertical farming through IoT and AI-Driven technologies: A comprehensive review. *Heliyon*, *10*(15).
33. EM, S., Chandy, D. A., PM, S., & Poulose, A. (2025). A Hybrid Deep Learning Model for Aromatic and Medicinal Plant Species Classification Using a Curated Leaf Image Dataset. *AgriEngineering*, *7*(8), 243.
34. Manoj, R. K., & Nandhini, S. A. (2025). Computer Vision and AI-Based Apple. *AI-Driven IoT Solutions in Agriculture: New Insights*, 269.
35. Akinlolu, A. O., Ogundoyin, I. K., Omotosho, L. O., & Ajayi, F. O. (2025). AI-Driven Hybrid Framework for the Diagnosis and Management of Diseases in Indigenous Tuber Crops. *FUOYE Journal of Pure and Applied Sciences (FJPAS)*, *10*(3), 105-131.
36. Sharma, A., Patel, R. K., Pranjal, P., Panchal, B., & Chouhan, S. S. (2024). Computer vision-based smart monitoring and control system for crop. In *Applications of computer vision and drone technology in agriculture 4.0* (pp. 65-82). Singapore: Springer Nature Singapore.
37. Hussain, M. (2023). When, where, and which? Navigating the intersection of computer vision and generative ai for strategic business integration. *IEEE Access*, *11*, 127202-127215.
38. Bose, S., Hazra, S., Mondal, B., Sengupta, S., Sen, P., Ansari, R., ... & Das, S. (2025). Application of artificial intelligence (AI) for conservation of endangered plant species: a comprehensive review based on global bibliometry. *Biodiversity and Conservation*, 1-54.
39. Siam, A. F. K., Bishshash, P., Nirob, M. A. S., Mamun, S. B., Assaduzzaman, M., & Noori, S. R. H. (2025). A comprehensive image dataset for the identification of lemon leaf diseases and computer vision applications. *Data in Brief*, *58*, 111244.
40. Khan, S. U., Alsuhaibani, A., Alabduljabbar, A., Almarshad, F., Altherwy, Y. N., & Akram, T. (2025). A review on automated plant disease detection: motivation, limitations, challenges, and recent advancements for future research. *Journal of King Saud University Computer and Information Sciences*, *37*(3), 34.