

Article

Not peer-reviewed version

An Effective Ensemble Convolutional Learning Model with Fine-tuning for Medicinal Plant Leaf Identification

[Mohd Asif Hajam](#) , Tasleem Arif , [Akib Mohi Ud Din Khanday](#) , [Mehdi Neshat](#) *

Posted Date: 20 September 2023

doi: 10.20944/preprints202309.1285.v1

Keywords: Ensemble Convolutional Learning; Transfer learning; Fine-tuning; Multiclass Classification; Medicinal plant identification



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

An Effective Ensemble Convolutional Learning Model with Fine-Tuning for Medicinal Plant Leaf Identification

Mohd Asif Hajam^{1,†,‡}, Tasleem Arif^{1,‡}, Akib Mohi Ud Din Khanday^{2,‡} and Mehdi Neshat^{3,*}

¹ Department of Information Technology, Baba Ghulam Shah Badshah University, Rajouri-Jammu and Kashmir-India ; aasifahmad0001@gmail.com(M.A.H), t.arif@bgsbu.ac.in(T.A)

² Department of Computer Science and Software Engineering-CIT, United Arab Emirates University, Al Ain, UAE; akibkhanday@uaeu.ac.ae(A.M.U.D.K)

³ Center for Artificial Intelligence Research and Optimisation, Torrens University Australia, Brisbane, QLD 4006, Australia;mehdi.neshat@torrens.edu.au(M.N)

* Correspondence: mehdi.neshat@torrens.edu.au;

‡ These authors contributed equally to this work.

Abstract: Accurate and efficient medicinal plant image classification is of utmost importance as these plants produce a wide variety of bioactive compounds that offer therapeutic benefits. With a long history of medicinal plant usage, different parts of plants, such as flowers, leaves, and roots, have been recognized for their medicinal properties and are used for plant identification. However, leaf images are extensively used due to their convenient accessibility and are a major source of information. In recent years, transfer learning and fine-tuning, which use pre-trained deep convolutional networks to extract pertinent features, has emerged as an extremely effective approach for image identification problems. This study is leveraging the power of three component deep convolutional neural networks, namely VGG16, VGG19 and DenseNet201, to derive features from the input images of the medicinal plant dataset, containing leaf images of 30 classes. The models were compared and ensembled to make four hybrid models to enhance the predictive performance by utilizing the averaging and weighted averaging strategies. Quantitative experiments were carried out to evaluate the models on the Mendeley medicinal leaf dataset. The resultant ensemble of VGG19+DensNet201 with fine-tuning showcases enhanced capability in identifying medicinal plant images with an improvement of 7.43% and 5.8% compared with VGG19 and VGG16. Furthermore, VGG19+DensNet201 can outperform its standalone counterparts by achieving an accuracy of 99.12% on the test set. A thorough assessment with metrics such as accuracy, recall, precision, and f1-score firmly establishes the effectiveness of the ensemble strategy.

Keywords: ensemble convolutional learning; transfer learning; fine-tuning; multiclass classification; medicinal plant identification

1. Introduction

Medicinal plants are gaining popularity in the pharmaceutical industry as they are less likely to have adverse effects and are less expensive than modern pharmaceuticals. According to the World Health Organization, there are over 21,000 plant species that can potentially be utilized for medicinal purposes. It is also reported that 80% of people around the world use medicinal plants for the treatment of their primary health ailments [1].

Systematic identification and naming of plants is often carried out by professional botanists (Taxonomists) who have deep knowledge of plant taxonomy [2]. Manually identifying plant species is a challenging and time-consuming process. Furthermore, the process is prone to errors, as every aspect of the identification is entirely based on human perception [3]. There is also the dearth of these plant identification subject matter experts, which gives rise to a situation of "Taxonomic impediment"

[4]. It is, therefore, important to develop an effective and reliable method for the accurate identification of these valuable plants.

Machine learning (ML) is a sub-field of artificial intelligence that solves various intricate problems in different application domains with minimal human intervention. Deep learning (DL), inspired by the structure and functionality of biological neurons, is a sub-field of machine learning (ML). DL involves the training of algorithms and demands a large amount of data to get better identification results. Due to the advancements in hardware technology and extensive availability of the data, DL has gained popularity in various tasks like natural language processing, game playing and image processing, and outstanding performance is being achieved, which could be otherwise impossible for humans to discern [5].

Various research studies reveal that researchers are showing a great interest in the automatic identification and classification of plant species using plant image specimens and by employing ML and DL techniques [6–10]. This automatic identification is carried out in different stages viz a) Image acquisition, b) Image preprocessing, c) Feature extraction d) Classification. Plant species identification can be performed using the different parts of plants like flowers, bark, fruits, and leaves or using the entire plant image. Researchers prefer to use leaf images for the identification process as leaves are easily identifiable parts of the plant. Leaf images are usually available for a long time during the year, while flowers and fruits are specific to a particular season only. There are more than one hundred studies that have used plant leaf images for the automatic identification process [3].

In this research, transfer learning and ensemble learning approaches were employed to classify medicinal plant leaf images into thirty classes. The transfer learning approach uses the existing knowledge gained from one task to solve problems of a related nature [11]. This approach is extensively used in image classification problems, particularly with Convolutional Neural Network models pre-trained on advanced GPUs capable of categorizing objects across a wide range of 1000 classes. Transfer learning can be used in medicinal plant image classification as the approach facilitates the migration of acquired features and parameters, so it reduces the need for extensive training from scratch.

The main aim of ensemble learning is to improve the overall performance of classifiers by combining the predictions of individual neural network models. Ensemble learning has recently gained popularity in image classification using deep learning [12–14]. We trained VGG16, VGG19 and DenseNet201 on the Mendeleev medicinal leaf dataset and evaluated the efficiency of these component models. After individual evaluation of the models, an ensemble learning approach using averaging and weighted averaging strategy was employed to make the final prediction.

In this section, we provide an overview of medicinal plants, highlighting their numerous benefits and the challenges associated with their identification. We also discuss existing methods used for identifying medicinal plants.

The upcoming sections will explore various related research studies shedding light on the advancements made in utilizing deep learning techniques for plant identification using their image data. Following the literature review, we describe the methodology employed and present the results obtained from our experiments. Furthermore, we discuss the significance and relevance of our proposed approach in addressing the existing challenges in medicinal plant identification.

2. Related Work

Numerous efforts have been undertaken to tackle the challenge of identifying medicinal plants from their different parts [15], employing different machine-learning approaches. However, researchers and experts often rely on leaf images to accurately classify different plant species as leaf images are widely recognized as the most accessible and trustworthy source of information for plant species identification.

Various authors relied on low-level features such as leaf shape, colour, and texture to differentiate between species. For example, the mobile application is known as Leafsnap by Kumar et al. [16],

identifies plant species based on plant leaf images. The feature extraction method used by the system is extracting features by capturing the curvature of leaf margins at multiple scales. The authors of [17] emphasized the incorporation of all these features, including shape, colour, texture, and venation, into a comprehensive feature vector for the Probabilistic Neural Network (PNN). By utilizing this feature vector, they achieved an impressive average accuracy of 93.75% on the publicly accessible Flavia dataset [18]. Amala Sabu et al. [19] proposed an ayurvedic plant identification system and extracted features with SURF and HOG-based techniques. In this study, a total of 200 images of 20 different plant leaves were collected, and K-NN was chosen as the classifier. It is evident that the aforementioned studies have primarily concentrated on recognition using hand-engineered image features. However, this approach has several limitations that need to be considered. The expressiveness and ability to discern intricate patterns and subtle changes in the data may be lacking in handcrafted features. Manual feature engineering is laborious, time-consuming, and subject to bias. Handcrafted features might not be able to manage noise and variations effectively. For huge datasets, the approach can also be computationally costly and ineffective. These techniques are difficult to apply for practical applications and necessitate the need to design a method that is less affected by the environment and is well suited for the recognition of real-world plant images.

In recent years, there has been a notable shift towards the adoption of Convolutional Neural Networks (CNNs), a powerful deep learning algorithm. CNNs have gained prominence due to their ability to automatically learn and extract meaningful features from input images, reducing the need for extensive human intervention in the feature extraction process. These algorithms have received a lot of attention in the literature. Sobitha Raj et al. [20] proposed a deep learning architecture by extracting features using MobileNet and DenseNet-121. In the study, different classifiers viz K-Nearest Neighbor, multinomial Logistic Regression, and linear discriminant analysis were used. Barre et al. [21] devised a deep-learning approach for extracting distinctive features from leaf images and subsequently classifying plant species. The authors revealed that the learned features obtained from a Convolutional Neural Network (CNN) outperformed handcrafted features. In order to automatically detect plant diseases, researchers in [22] have successfully created a robust and deep CNN model with nine layers. To enhance the learning capability of the CNN, the input data was augmented, resulting in an increased number of samples. The model was subjected to rigorous testing using various classifiers, and it achieved an impressive accuracy rate of 96%.

In computer vision tasks, several widely recognized and open-source models have gained popularity, including VGG-16, VGG-19 [23], Inception V3 [24], DenseNet [25] and ResNet-50 [26]. These state-of-the-art architectures have been extensively trained, and their parameters and weights can be utilized by other researchers to address problems in various domains. This allows for efficient knowledge transfer and facilitates the application of pre-trained models in different computer vision applications. Roopashree et al. [9] employed transfer learning to extract features from the pre-trained networks and used Artificial Neural Networks (ANN), Support Vector Machines (SVM), and SVM with Batch Optimization (SVM with BO) as classification algorithms. The results of their study demonstrated that the model trained with the Xception network and ANN achieved an impressive accuracy of 97.5% for real-time leaf images. Transfer learning was also employed by [27] to create different combinations of pre-trained MobileNet CNN for the classification of medicinal plant leaves.

3. Methods and Material

3.1. Convolutional Neural Network (CNN)

In the domain of artificial intelligence, significant progress has been made in recent years, particularly in the field of computer vision. Researchers have dedicated their efforts to unlocking the potential of machines in understanding and interpreting visual information, such as images and videos. The application of deep learning in computer vision has propelled the field forward, enabling machines to analyze, interpret, and make intelligent decisions based on visual input. This advancement has been

made possible through the development of convolutional neural networks (CNNs) and deep learning techniques. A schematic of a CNN can be seen in Figure 1.

The CNN structure is inspired by the human visual cortex and is widely used for automatic feature extraction from large datasets [28]. It consists of a series of convolutional layers, followed by sampling layers and a fully connected layer.

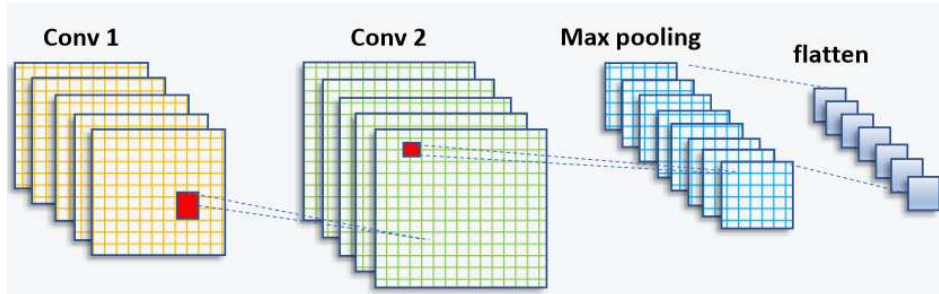


Figure 1. CNN architecture (feature extraction block)

Following the convolution layer, the ReLU activation function is applied to extract nonlinear features. The purpose of incorporating the ReLU layer is to introduce non-linearity into the network. The mathematical representation of the ReLU function is defined as follows, as shown in Equation 1.

$$\text{relu}(v) = \max(v, 0) \quad (1)$$

The pooling layer plays a crucial role in reorganizing the feature maps to effectively decrease factors [29], allocation of memory, and computational budget within the framework. Each feature map undergoes pooling, with the two commonly used approaches being max pooling and average pooling, as denoted by can be seen in Equation 2 and Equation 3, respectively. These pooling functions help in summarizing the information within the feature maps while reducing their spatial dimensions.

$$a_k = \max_{j \in R_k} (M_j) \quad (2)$$

$$a_k = \frac{1}{|R_k|} \sum_{j \in R_k} M_j \quad (3)$$

The pooling layer is denoted as M , and R_k represents all parameters within this layer. In the fully connected layers, the confidential criteria are computed and represented as a three-dimensional array, like $1 \times 1 \times c$. Each component within this volume corresponds to structural scores, with c representing the categories. In a fully connected (FC) layer, all neurons are connected to neurons in the preceding layers; that is a type of interconnection. This connectivity is a characteristic feature of a typical convolutional neural network (CNN) [29], where all layers are sequentially interconnected, as depicted in Equation 4. This connectivity pattern allows for information flow and feature extraction across the network.

$$m_r = F_r(m_{r-1}) \quad (4)$$

In the following, Equation 5 represents a deeper architecture of CNN by a greater number of fully connected layers, which may lead to vanishing or exploding the gradient values. There are various techniques to deal with this issue, such as the application of shortcut connections among the layers that can be seen in ResNet [30].

$$m_r = F_r[(m_{r-1}) + m_{r-1}] \quad (5)$$

As another example of the architecture of CNN, there is a direct connection between whole layers in DenseNet where the input in the current layer comes from the previous one as a standard model [31]. The formulation can be seen below.

$$m_r = F_r [(m_0, m_1, m_2, \dots, m_{r-1})] \quad (6)$$

In this study, three prominent convolutional neural network architectures, including VGG16 [23], VGG19 [23], and DenseNet201 [25] are employed for the identification of medicinal plant images.

3.2. VGG-16

VGG-16 is a very popular convolutional neural network (CNN) model that was proposed in [23]. It achieved remarkable performance in image detection and classification tasks, and this network is easy to use with transfer learning. The model consists of 13 convolutional layers, five max-pooling layers, and three fully connected layers. The final layer employs the softmax activation function for classification. The architecture of VGG-16 is relatively straightforward, which takes an input tensor of size 224 X 224 with three RGB channels.

3.3. VGG-19

VGG-19 [23] is an extended version of VGG-16, designed to enhance image recognition performance. It achieved remarkable success in the ImageNet Challenge 2014, where the Visual Geometry Group (VGG) team secured top rankings. VGG-19 architecture consists of 16 convolutional and three fully connected layers. Like the VGG16 network, this model also takes a 224 x 224 input image size with 3 RGB channels. The convolutional blocks serve as feature extractors, generating bottleneck features. VGG-19 exemplifies the VGG team's commitment to advancing image recognition capabilities.

3.4. DenseNet-201

DenseNet-201 [25], a deep convolutional neural network (CNN) model introduced in 2017, is known for its dense connectivity pattern. With 201 layers, it employs dense blocks where each layer is directly connected to every other layer, promoting feature reuse and gradient flow. Transition layers are inserted to control parameters and reduce spatial dimensions. Batch normalization and ReLU activation enhance training. DenseNet201 achieves high accuracy in image recognition tasks by capturing intricate patterns and hierarchical representations. Pretrained on ImageNet, it serves as a feature extractor or can be fine-tuned for transfer learning, making it a powerful CNN model for image classification and related applications.

DenseNet201 has various advantages due to its 201 convolutional layers. These advantages include addressing enabling feature adaptation to be used again, vanishing-gradient problems, achieving optimal feature distribution, and reducing the number of parameters [29]. Let's consider an image H_0 being fed into a neural network comprising S layers with non-linear transformation denoted as $F_s(.)$. In this case, DenseNet incorporates conventional skipping connections within the feed-forward network (See Figure 2). These connections allow bypassing the non-linear alteration using an identity function, as represented as follows.

$$H_s = F_s (H_{s-1}) + H_{s-1} \quad (7)$$

Further, DenseNet offers a distinct benefit in that the gradient can flow directly through the identity function from the primary layers to the last layer. On the other hand, dense networks utilize direct end-to-end connections to maximize the details within each layer. In this case, the $s - th$ layer receives all the data from the preceding layer, as follows.

$$H_s = F_s [(H_0, H_1, \dots, H_{s-1})] \quad (8)$$

DenseNet's architectural design incorporates a particular process for downsampled data, which occurs within Dense Blocks. These blocks are divided into Transition layers, each including a 1×1 convolutional layer (CONV), an average pooling layer, and batch normalization (BN). The elements from the transition layer progressively disseminate to the dense layers, making the network less intricate. In an effort to enhance network utility, the average pooling layer was wholly converted into a 2×2 max pooling layer. Each convolution layer is preceded by batch normalization. The network's growth rate, denoted by the hyperparameter k , equips DenseNet to yield state-of-the-art results. The conventional pooling layers have been removed, and the proposed detection layers have been fully amalgamated and connected to the classification layers for detection purposes. DenseNet-264 encompasses even more complex network designs than the 201-layer network. However, due to its narrower network footprint, the 201-layer structure was deemed suitable for plant leaf detection tasks. DenseNet201, despite a smaller growth rate, still performs excellently because its design employs feature maps as a network-wide mechanism. The architecture of DenseNet201 is depicted in Figure 2.

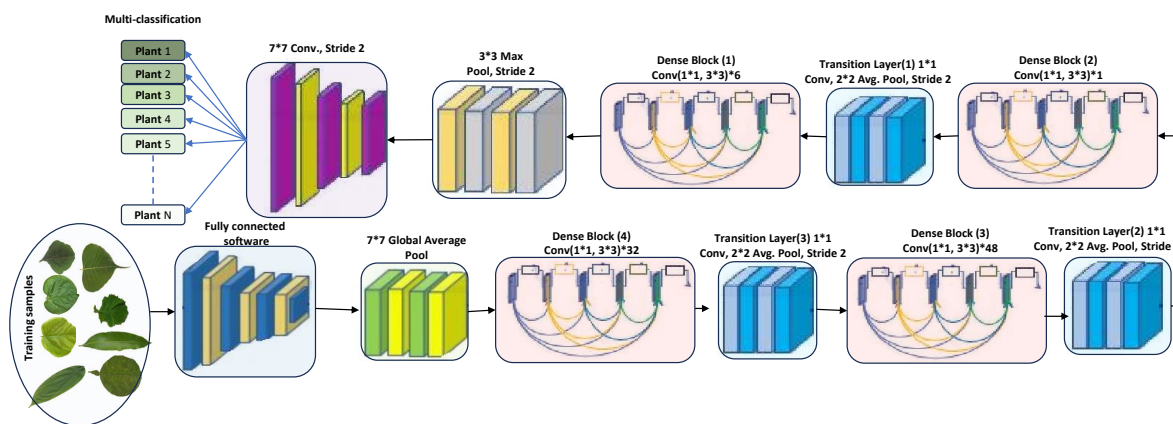


Figure 2. DenseNet201 architecture (feature extraction block)

DenseNet-201, as a deep learning architecture, possesses distinct advantages and disadvantages. The benefits of DenseNet-201 are as follows. Firstly, it effectively addresses the vanishing gradient problem by establishing direct connections between layers. This facilitates smoother gradient flow during training, leading to improved optimization based on gradients. Secondly, DenseNet-201 excels in distributing features evenly across layers due to its dense connections. This enhances the model's overall representational capabilities. Additionally, DenseNet-201 enables the reuse of feature maps at various depths, promoting efficient information flow and potentially enhancing the network's learning capacity. Lastly, it reduces the number of parameters compared to conventional deep learning architectures by eliminating the need for redundant feature learning in individual layers. This results in a more efficient utilization of parameters in the model.

3.5. Ensemble Learning Approach

Ensemble learning offers a compelling approach to elevate the performance of classifiers by encompassing diverse techniques, prominently including bagging, boosting, and stacking [32]. Ensemble learning considers either a homogeneity or heterogeneity approach. In the case of homogeneity, a single base classifier is trained on different datasets, whereas in heterogeneity, different classifiers are trained on a shared dataset. The resultant ensemble, by harnessing the collective wisdom of its components, generates predictions through a synthesis of approaches such as averaging, weighted averaging, and voting. These mechanisms draw from the individual outputs furnished by the base classifiers. In our pursuit of automating medicinal leaf detection, we have embraced the heterogeneous ensemble paradigm, where weighted average and average play a pivotal role in amalgamating the contributions of diverse classifiers to culminate in the ultimate and refined outcome.

3.6. Proposed Ensemble learning for Medicinal Plant Leaf Identification

The proposed ensemble learning approach illustrated in Figure 3 was tailored for the precise and reliable identification of medicinal plant leaves. The approach capitalizes on the collective intelligence of multiple Convolutional Neural Network (CNN) models, offering a holistic and robust strategy for accurate leaf classification. The framework initiation encompasses the individual training of three prominent CNN models: VGG16, VGG19, and DenseNet201. Central to our strategy is the utilization of transfer learning, a technique where the final layers of the aforementioned CNN models are replaced with custom pooling and dense layers. This adaptation facilitates the alignment of models with the unique attributes of our medicinal leaf dataset, enabling them to discriminate between various leaf species effectively. Each model undergoes extensive training for 100 epochs to ensure thorough convergence and optimal performance.

Through averaging and weighted averaging, the ensemble model combines the predictions of individual CNN models. The ensemble's predictive synergy amplifies the accuracy, resilience, and overall efficacy of our medicinal plant leaf identification system, surpassing standalone model performance.

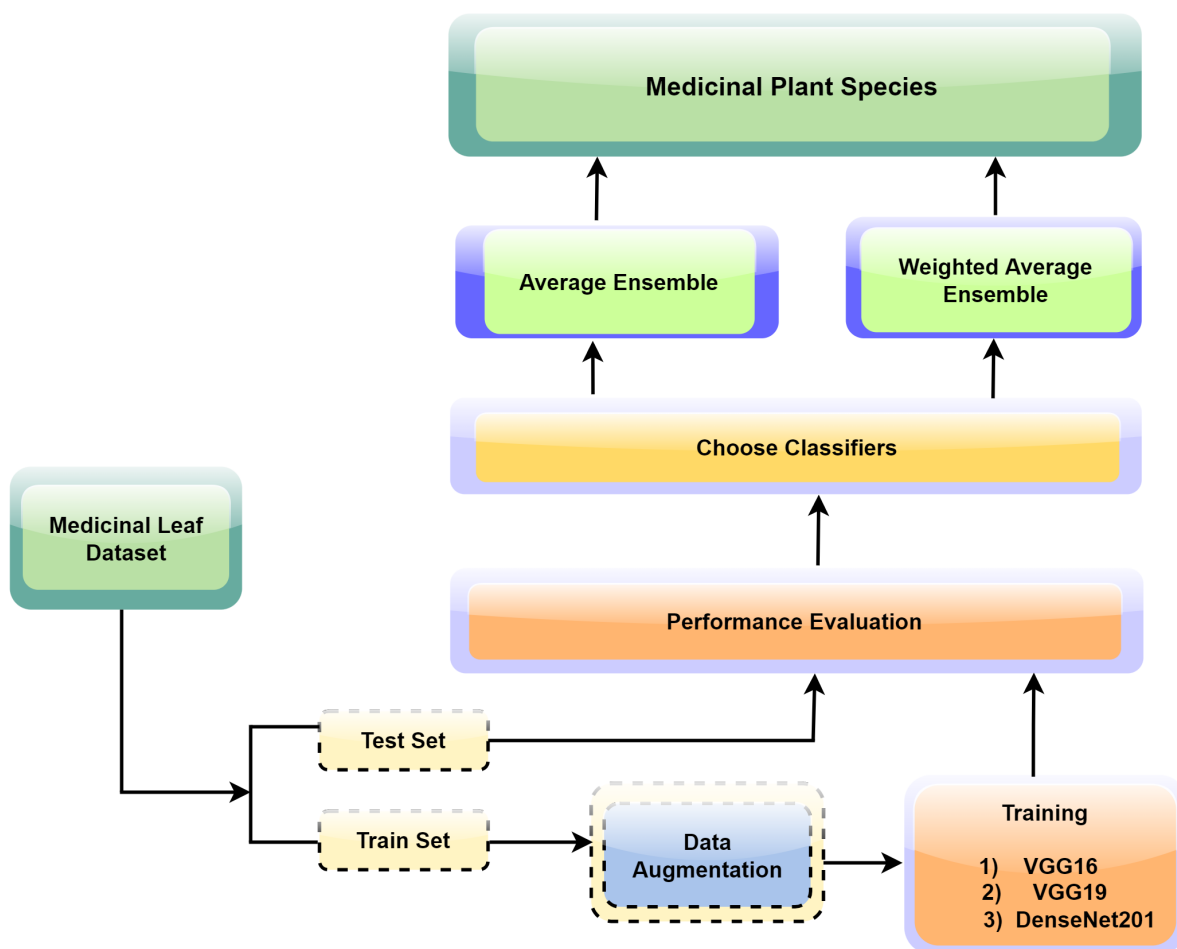


Figure 3. Outlines the process flow of the proposed research study.

Subsequent sections discuss component models, ensemble learning approach, experimental results, and discussions, shedding light on its efficacy and contribution to the field. The applied steps of the proposed ensemble model can be listed as follows.

- (1) Data loading and splitting: Collect "Mendeley Data–Medicinal Leaf Dataset" (1835 images, 30 species) and split into training (70%) and testing (30%).

- (2) Model selection: Choose VGG16, VGG19, and DenseNet201 as base models.
- (3) Image Standardization: Resize images to 224 x 224 pixels that are compatible with the input size expected by the CNN models.
- (4) Data augmentation: enhance model learning and diversity by applying random rotations, flips, translations, and adjustments to brightness or contrast. This exposure to varied image variations during training improves the model's generalization.
- (5) Batch generation: dividing the dataset into smaller subsets of images, which are then fed into the CNN model during training. This approach enhances computational efficiency by processing a portion of the dataset at a time rather than the entire dataset at once.
- (6) Training and transfer learning: The models were trained individually employing transfer learning with softmax activation for classification using Adam optimizer and categorical cross-entropy.
- (7) Validation models: for each trained model, prediction was performed by calculating class probabilities on the test set.
- (8) Hybridization models: Ensemble models were created by combining the results generated by individual classifiers, using averaging and weighted averaging strategies. Using the three components, VGG16, VGG19, and DenseNet201, the feasible ensemble models were created, validated and compared to select the best-performed model. The proposed learning framework can be seen in Figure 4.

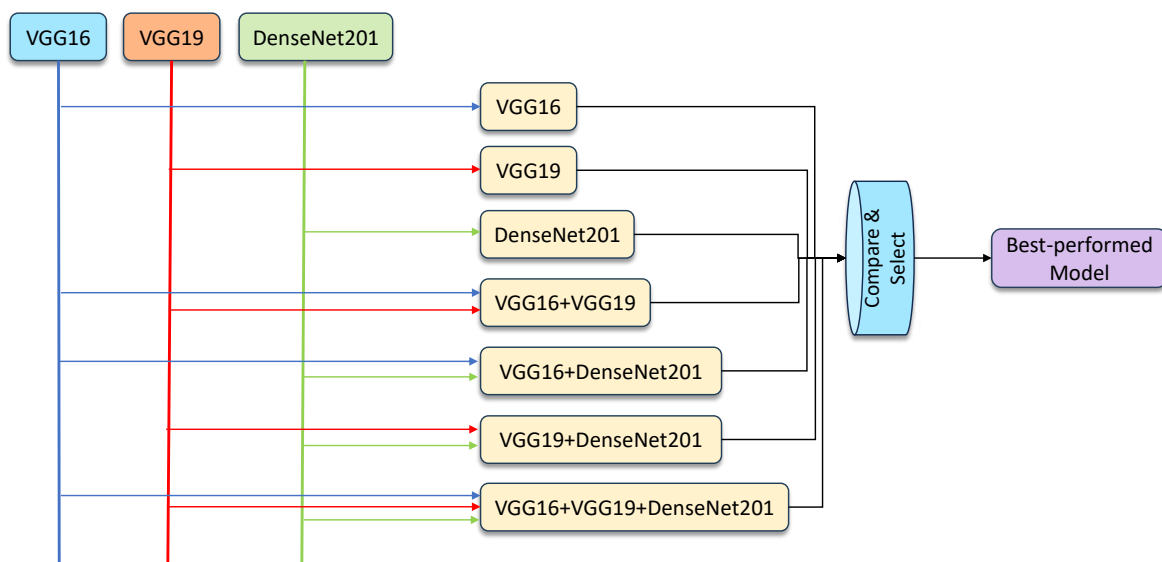


Figure 4. An overview of the proposed Ensemble model framework

3.7. Dataset description

To start the process of plant species identification, the first step involves collecting the dataset. A qualitative study was conducted to identify suitable medicinal leaf dataset sources and determine the dataset format. Finally, a dataset published in Mendeley [33], representing images from 30 different medicinal plants, was selected for this study. The selected dataset contained a total of 1835 leaf images. Figure 2 depicts some of the leaf samples of the Mendeley medicinal leaf dataset.

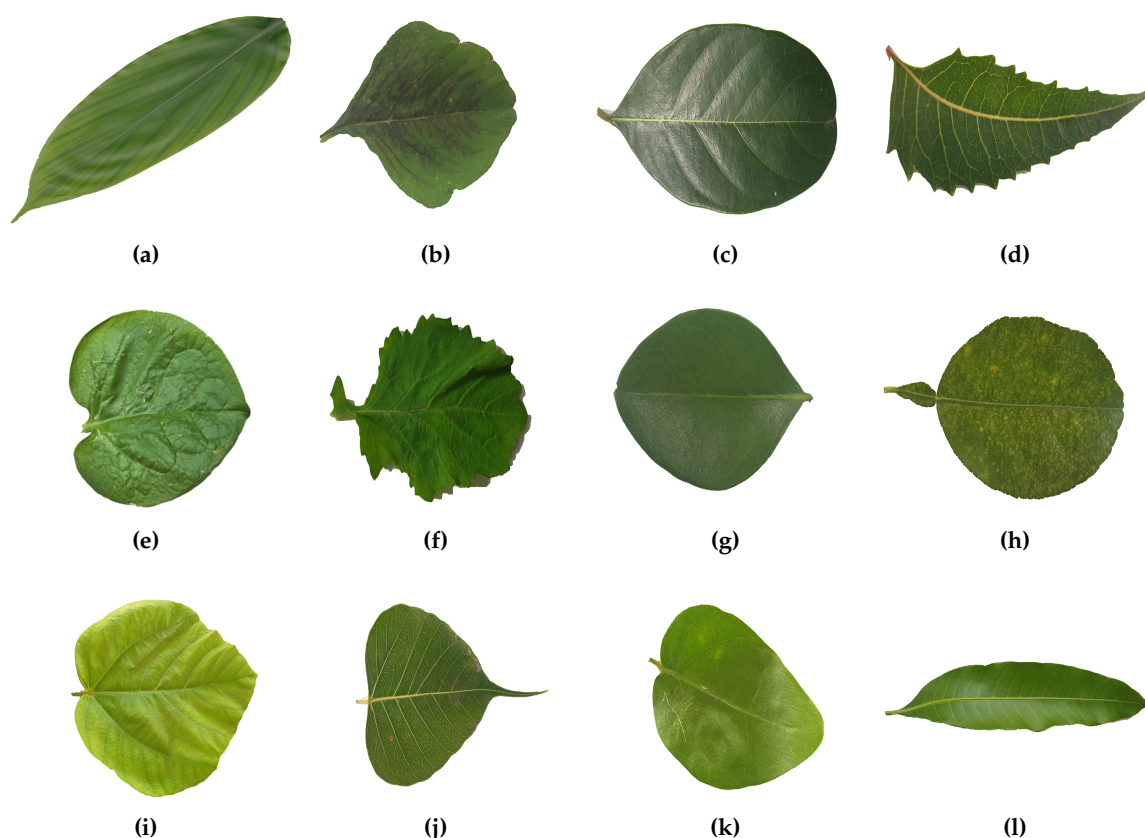


Figure 5. Showcases a sample of Medicinal Leaf dataset images, illustrating the diverse range of plant images incorporated for classification purposes. a) *Alpinia Galanga* (Rasna), b) *Amaranthus Viridis* (Arive-Dantu), c) *Artocarpus Heterophyllus* (Jackfruit), d) *Azadirachta Indica* (Neem), e) *Basella Alba* (Basale), f) *Brassica Juncea* (Indian Mustard), g) *Carissa Carandas* (Karanda), h) *Citrus Limon* (Lemon), i) *Ficus Auriculata* (Roxburgh fig), j) *Ficus Religiosa* (Peepal Tree), k) *Jasminum* (Jasmine), l) *Mangifera Indica* (Mango).

3.8. Data preprocessing

Data preprocessing plays an important role in preparing raw data for machine learning models, ensuring that the data is in a suitable format for effective training and analysis. The dataset collected was meticulously divided into distinct training and testing subsets, adhering to the widely adopted 70-30 ratio. This partitioning ensured an optimal allocation of data for both model development and rigorous evaluation. During the component model training phase, an additional step of further subdivision was introduced. Within the training subset, 20% of the data was strategically reserved for validation. This segregation endowed us with a dedicated validation set, essential for gauging the convergence and generalization capabilities of our evolving models. During the training phase of component models (VGG16, VGG19, and DenseNet201), and ensemble models, images were resized to a uniform size of 224 x 224 pixels. This preprocessing step ensured that all images were of the same dimensions, allowing them to be fed into the neural networks consistently. The images were then normalized to ensure pixel values fell within a specific range, which aids in stable and efficient training. To enhance the robustness and prevent overfitting, data augmentation techniques were applied to the training dataset. These techniques involve generating new instances by applying transformations to the existing images. Various transformations, such as random rotation by up to 20 degrees, zooming by 5%, horizontal and vertical shifting by 5%, shearing at a 5% angle, and horizontal flipping. The "nearest" fill mode handled new pixels resulting from transformations. These transformations helped

increase the diversity of the training data, enabling the model to learn a wider range of features and generalize better to unseen data.

4. Evaluation Metrics

Evaluation metrics play a vital role in optimizing classifiers for the accurate detection and classification of medicinal plant images. In this study, we have evaluated the trained models on the basis of commonly used performance metrics in this domain, which include Accuracy, Sensitivity (Recall), Precision and F1-Score. Accuracy measures the proximity between predicted and target values, while Sensitivity focuses on the ratio of correctly identified positive instances. True Negatives (TN) and True Positives (TP) represent correctly classified negative and positive instances, respectively, contributing to successful classification and detection. False Negatives (FN) and False Positives (FP) denote misclassifications. These metrics guide the fine-tuning of classifiers to achieve optimal performance in medicinal plant image classification. The formulae for calculation of the metrics used in this study are shown in equations (9), (10), (11), and (12)

4.1. Accuracy

Accuracy is a metric that calculates the proportion of correctly predicted values out of the total number of instances evaluated.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

4.2. Recall

Recall, also known as sensitivity, quantifies the proportion of positive values that are accurately classified.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

4.3. Precision

Precision assesses the accuracy of positive predictions within the predicted values of the positive class.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

4.4. F1 Score

The F1-Score quantifies the balanced performance of a classifier by taking into account both the recall and precision rates through their harmonic average.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

The classification reports and confusion matrices were analyzed to evaluate the performance of the classifiers. Similar procedures were carried out for the other feature sets, allowing a comparison of the classifiers' performance.

5. Experimental Results

5.1. Classification outcomes of component deep neural networks

Employing transfer learning played a pivotal role where the pre-trained VGG16, VGG19, and DenseNet201 deep neural network architectures were fine-tuned to classify medicinal plants. It is

evident from Table 1 and Figure 6 that DenseNet201 displayed an impressive training accuracy of 100%. And validation accuracy of 94.64%. Notably, DenseNet201 exhibited remarkable generalization, achieving an exceptional test accuracy of 98.93%. In Figure 7, the training and validation accuracy and loss for the DensNet201 are graphically illustrated, and Figure 8 displays the distribution of predicted classes against the actual classes. These results underscore the capacity of transfer learning to adapt models to the specific characteristics of the medicinal plant dataset, yielding robust and accurate classifiers for medicinal leaf image classification.

Table 1. Training, Test and Validation accuracy of VGG16, VGG19 and DenseNet201

Deep Neural Network	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
VGG16	96.19	89.7	93.67
VGG19	95.41	87.94	92.26
DenseNet201	100	94.64	98.93

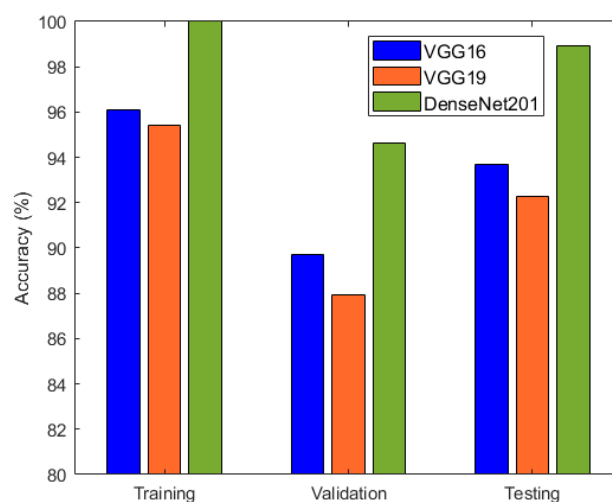


Figure 6. Accuracy comparison of VGG16, VGG19 and DenseNet201.

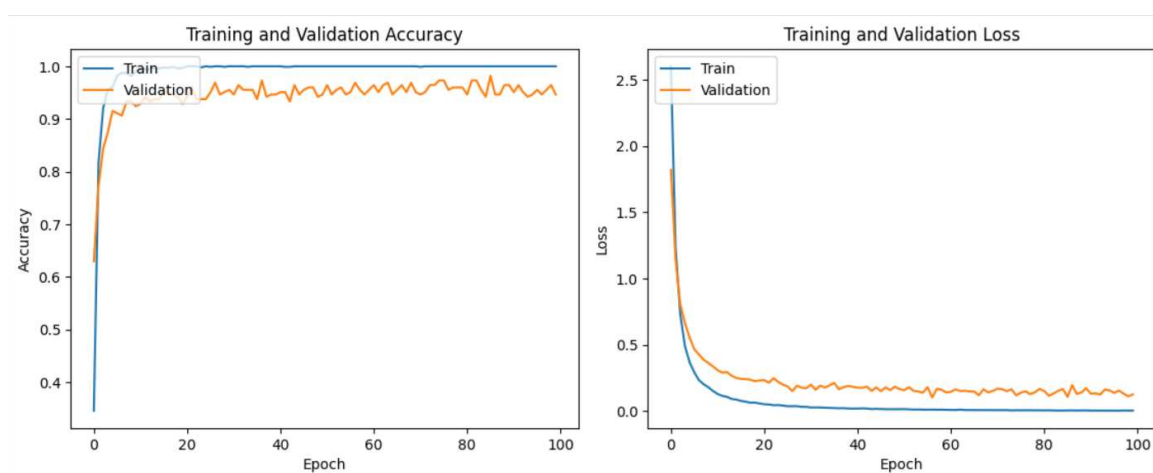


Figure 7. Depicts the training and validation loss and accuracy of DenseNet201.

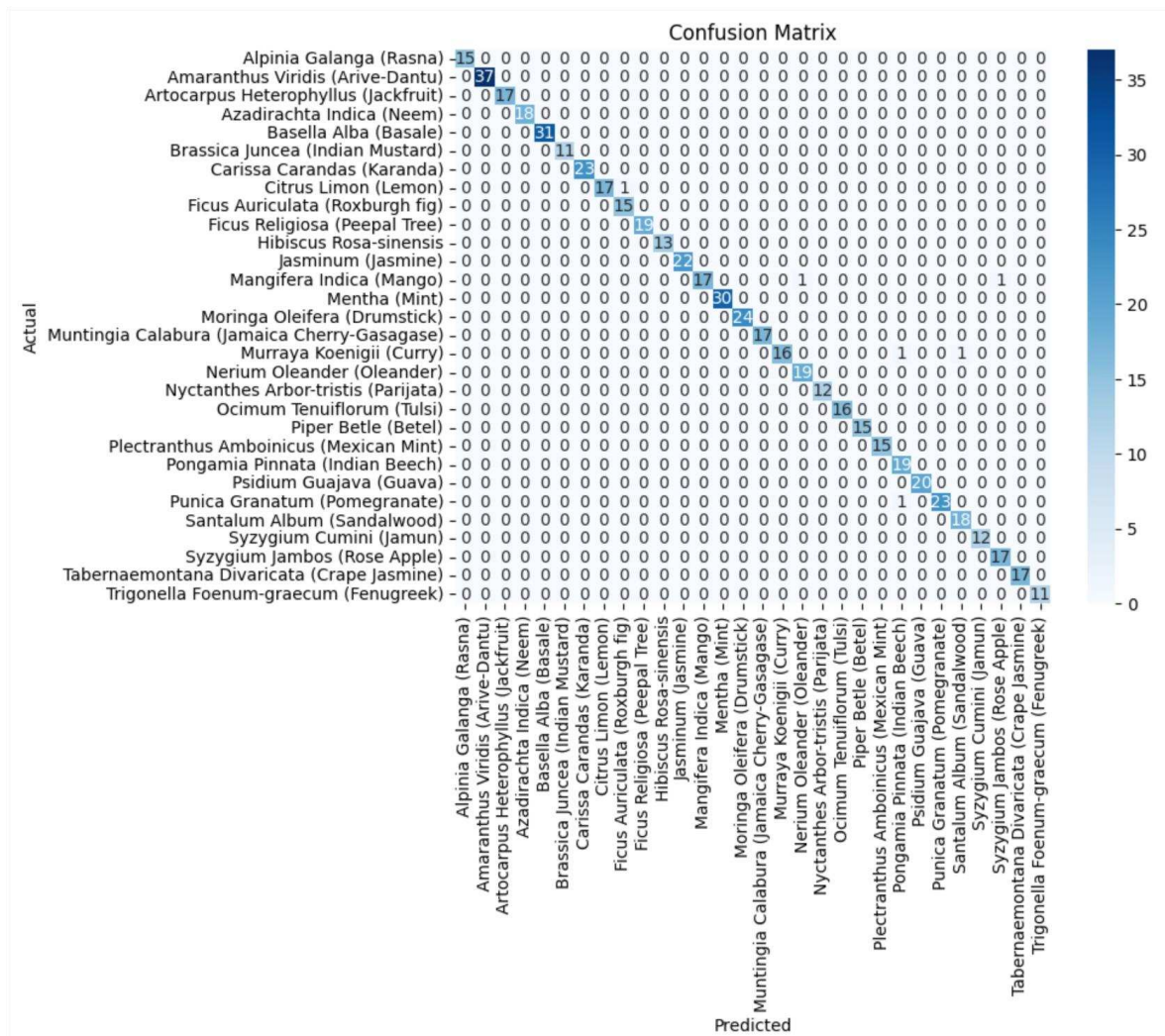


Figure 8. Shows the confusion matrix of the DenseNet201 component model

The precision, recall, and F1-score results of the individual deep neural network models are summarized in Table 2 and visualized in Figure 9. These metrics provide insights into the models' performance in terms of correctly identifying positive cases, capturing actual positive instances, and achieving a balance between precision and recall. Among the models, DenseNet201 exhibited the highest precision, recall, and F1 score, indicating its excellence in classification accuracy.

Table 2. Precision, recall and F1 score obtained by component deep neural networks.

Deep Neural Network	Precision (%)	Recall (%)	F1-Score (%)
VGG16	94.02	93.67	93.62
VGG19	92.67	92.26	92.17
DenseNet201	99.01	98.94	98.93

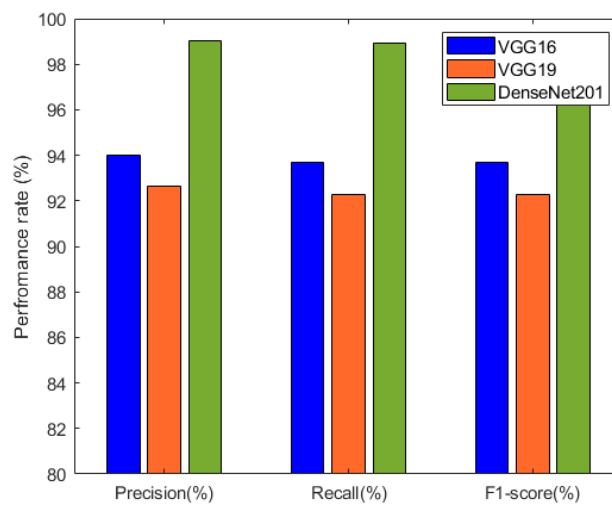


Figure 9. Performance comparison between VGG16, VGG19 and DenseNet201 based on Precision, Recall and F1-score.

5.2. Ensemble Approaches for Improved Classification Performance

After individually evaluating component deep neural network models, the exploration turned towards ensemble techniques for further refinement. Specifically, two ensemble approaches, namely averaging and weighted averaging, were employed to combine the strengths of individual models. This culminated in the creation of diverse ensemble models that revealed enhanced classification capabilities. Table 3 and Figure 10 present the accuracy outcomes of four developed ensemble models, including VGG19 + DenseNet201, VGG16 + VGG19, VGG16 + DenseNet201, and VGG16 + VGG19 + DenseNet201 based on the average performance strategy.

Table 3. Reporting accuracy outcomes of different ensembled deep neural networks using averaging ensemble approach.

Average Ensemble			
Ensemble Deep Neural Networks	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
VGG19 + DenseNet201	100	95.52	99.12
VGG16 + VGG19	99.04	90.65	96.66
VGG16 + DenseNet201	100	95.52	98.76
VGG16 + VGG19 + DenseNet201	99.90	93.90	98.41

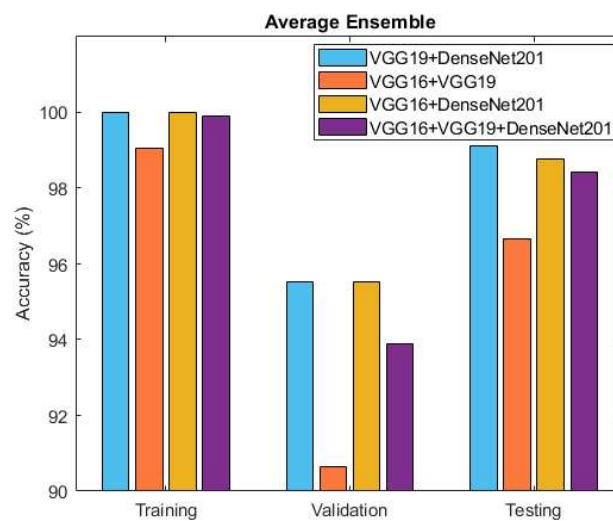


Figure 10. Accuracy comparison between Ensemble models: VGG19 + DenseNet201, VGG16 + VGG19, VGG16 + DenseNet201 and VGG16 + VGG19 + DenseNet201.

We can see the statistical analysis of the weighted average ensemble model for four developed approaches based on the training, validation and testing classification accuracy in Table 4.

Table 4. Reporting accuracy outcomes of different ensembled deep neural networks using weighted average ensemble.

Weighted Average Ensemble					
Ensemble Networks	Deep	Neural	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
VGG16 + DenseNet201	+	VGG19 +	99.61	92.68	97.89
VGG19 + DenseNet201			99.23	91.86	96.83
VGG19 + VGG16			98.75	89.02	96.66
VGG16 + DenseNet201			99.80	91.05	98.06

In Figure 11, the confusion matrix for the ensemble of VGG19+DenseNet201 is presented, which highlights the effectiveness of the ensembled approach using the average ensemble strategy. The diagonal values of a confusion matrix represent the number of data points where the predicted label matches the true label, indicating correct predictions. They signify instances where the classifier identifies the positive class correctly. In contrast, the off-diagonal elements of the confusion matrix denote mis-classifications made by the classifier. These errors can be false positives or false negatives, depending on whether the predicted label is incorrect for the positive class or the negative class, respectively. In other word, a higher value on the diagonal of the confusion matrix indicates better performance of the classifier, as it suggests a larger number of correct predictions. The ultimate goal is to maximize the values on the diagonal while minimizing the off-diagonal elements.

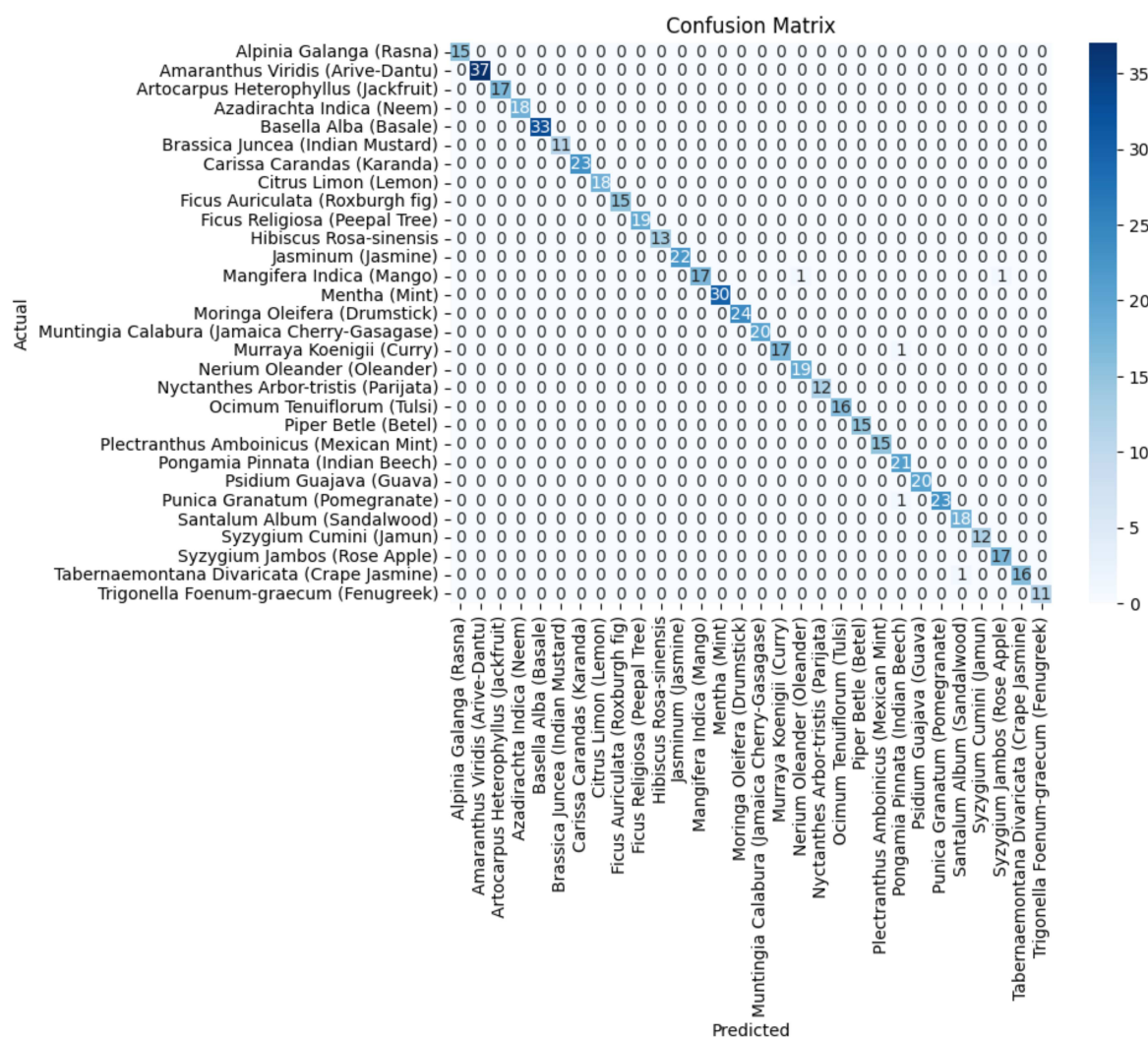


Figure 11. Shows the confusion matrix of the average ensemble of VGG19+densenet201

6. Discussion

In our experimental analysis, we examined the effectiveness of various pre-trained models such as VGG16, VGG19, and DenseNet201. For each of these models, we excluded the top layers to utilize their pre-trained weights and the acquired feature representations. We created new models by incorporating a flattening layer to transform the extracted features into a high-dimensional vector, followed by the addition of a dense layer. These models were then compiled using the Adam optimizer and categorical cross-entropy loss function. Table 1 encapsulates the accuracy outcomes achieved by three distinct deep neural network architectures: VGG16, VGG19, and DenseNet201. Notably, among these architectures, DenseNet201 stands out as a model of exceptional performance. It effortlessly attains a perfect training accuracy of 100%, underscoring its capability to internalize intricate nuances within the training dataset. This mastery in learning translates cohesively to the validation phase, where DenseNet201 achieves an impressive accuracy of 94.64%. Moreover, during the rigorous test phase, the model showcases a striking test accuracy of 98.93%, reinforcing its prowess in accurately classifying diverse and previously unseen medicinal plant samples.

The provided metrics in Table 2 highlight the precision, recall, and F1-Score of three prominent deep neural network architectures: VGG16, VGG19, and DenseNet201. Among these models, DenseNet201 emerges as a standout performer. Impressively, it achieves a remarkable precision of 99.01%, reflecting its capability to accurately identify positive cases while minimizing false positives.

The strength of DenseNet201 becomes evident in recall as well, with a value of 98.94%. This signifies its effectiveness in capturing a substantial portion of actual positive cases. The F1-Score of 98.93% for DenseNet201 affirms its harmonious balance between precision and recall, showcasing its overall excellence in classification accuracy.

The accuracy outcomes attained through distinct ensemble deep-learning approaches are summarized in Table 3 and Table 4. Among the strategies applied to the individual deep neural network models, it is evident that the most effective solution arises from combining VGG19 and DenseNet201 utilizing the averaging technique. With a perfect training accuracy of 100%, impressive validation accuracy of 95.52%, and exceptional test accuracy of 99.12%, this ensemble demonstrates the ability to master intricate patterns, generalize well to new data, and achieve precise classification. A comparative analysis of this study with the previously state-of-the-art approaches for the identification of medicinal plants using the same dataset was conducted, and it can be observed from Table 5 that leveraging the complementary features of VGG19 and DenseNet201, the ensemble approach improved robustness, balanced performance, making it a potent solution for medicinal plant identification. Its success underscores the significance of ensemble techniques in enhancing complex classification tasks and holds promising applications in diverse fields.

Table 5. Comparative analysis of the proposed study with the previously state-of-the-art approaches for the identification of medicinal plant images.

Reference	Technique	Medicinal Leaf Dataset	Accuracy
[34]	MobileNetV1	Mendeley Medicinal Leaf Dataset	98%
[35]	MobileNetV2	Mendeley Medicinal Leaf Dataset	81.82%
[36]	Mask RCNN	Mendeley Medicinal Leaf Dataset	95.7%
Proposed Approach	Ensemble Learning	Mendeley Medicinal Leaf Dataset	99.12%

7. Conclusion

In the contemporary world, traditional medicines have gained importance due to the high costs and potential negative impacts of allopathic medicines. Medicinal plant classification is usually carried out by plant taxonomists; however, the human-centric nature of this approach can introduce inaccuracies or errors in judgment. Artificial intelligence methodologies have found extensive application in automating plant recognition processes. This research paper investigates the application of ensemble learning to elevate the accuracy of medicinal plant identification. The dataset consisted of medicinal leaf images sourced from a published collection on Mendeley. To construct ensemble models based on Convolutional Neural Networks (CNNs), we compared three leading CNN architectures: VGG16, VGG19, and Densenet201. By employing Transfer Learning, a trio of three-component classifiers were harnessed without their upper layers. This adaptation allowed them to discern crucial attributes within medicinal leaf images, and these attributes were then integrated into Dense Layers. Subsequently, these classifiers underwent training on a dataset comprising 30 diverse classes of medicinal leaves using a softmax classifier.

Following the individual assessment of component deep neural network models, the medicinal plant identification approach was enhanced through ensemble techniques employing averaging and weighted averaging strategies. Upon assessing the ensemble models, it became evident that the most potent approach emerged from the fusion of VGG19 and DenseNet201, employing the averaging method. This ensemble displayed remarkable attributes, including an impeccable training accuracy of 100%, a notable validation accuracy of 95.52%, and an outstanding test accuracy of 99.12%. These results collectively underscore the capacity of the ensemble learning approach to proficiently capture intricate patterns, generalize effectively to novel data, and achieve accurate classification.

Our future plans involve collaboration with domain experts to ensure the high-quality collection and annotation of our dataset. We also aim to develop intuitive user interfaces and adapt the model for real-time applications.

Author Contributions: Conceptualization, M.A.H, T.A and A.M.U.D.K.; Methodology, A.M.U.D.K. and T.A; Validation, M.A.H; Formal analysis, M.A.H; Investigation, M.A.H, T.A. and M.N; Data curation, M.A.H and T.A; Writing—original draft, A.M.U.D.K. and M.A.H.; Writing—review & editing, A.M.U.D.K., M.A.W., S.T.R., M.N. and Q.R.K.; Visualization, A.M.U.D.K., M.A.H., T.A and M.N.; Supervision, T.A, and A.M.U.D.K.; Project administration, T.A; Funding acquisition, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge Baba Ghulam Shah Badshah University for their valuable support. Also, the authors would like to thank the Center for Artificial Intelligence Research and Optimisation, Torrens University Australia, for supporting APC in this publication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. www.nhp.gov.in. https://www.nhp.gov.in/introduction-and-importance-of-medicinal-plants-and-herbs_mtl. (Accessed on 09/12/2023).
2. Azlah, M.A.F.; Chua, L.S.; Rahmad, F.R.; Abdullah, F.I.; Wan Alwi, S.R. Review on techniques for plant leaf classification and recognition. *Computers* **2019**, *8*, 77.
3. Wäldchen, J.; Mäder, P. Plant species identification using computer vision techniques: A systematic literature review. *Archives of computational methods in engineering* **2018**, *25*, 507–543.
4. de Carvalho, M.R.; Bockmann, F.A.; Amorim, D.S.; Brandão, C.R.F.; de Vivo, M.; de Figueiredo, J.L.; Britski, H.A.; de Pinna, M.C.; Menezes, N.A.; Marques, F.P.; et al. Taxonomic impediment or impediment to taxonomy? A commentary on systematics and the cybertaxonomic-automation paradigm. *Evolutionary Biology* **2007**, *34*, 140–143.
5. Sarker, I.H. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science* **2021**, *2*, 420.
6. Husin, Z.; Shakaff, A.; Aziz, A.; Farook, R.; Jaafar, M.; Hashim, U.; Harun, A. Embedded portable device for herb leaves recognition using image processing techniques and neural network algorithm. *Computers and Electronics in agriculture* **2012**, *89*, 18–29.
7. Gokhale, A.; Babar, S.; Gawade, S.; Jadhav, S. Identification of medicinal plant using image processing and machine learning. In Proceedings of the Applied Computer Vision and Image Processing: Proceedings of ICCET 2020, Volume 1. Springer, 2020, pp. 272–282.
8. Puri, D.; Kumar, A.; Virmani, J.; Kriti. Classification of leaves of medicinal plants using laws texture features. *International Journal of Information Technology* **2019**, pp. 1–12.
9. Roopashree, S.; Anitha, J. DeepHerb: a vision based system for medicinal plants using xception features. *IEEE Access* **2021**, *9*, 135927–135941.
10. Sachar, S.; Kumar, A. Deep ensemble learning for automatic medicinal leaf identification. *International Journal of Information Technology* **2022**, *14*, 3089–3097.
11. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* **2009**, *22*, 1345–1359.
12. Nakata, N.; Siina, T. Ensemble Learning of Multiple Models Using Deep Learning for Multiclass Classification of Ultrasound Images of Hepatic Masses. *Bioengineering* **2023**, *10*, 69.
13. Kuzinkovas, D.; Clement, S. The detection of covid-19 in chest x-rays using ensemble cnn techniques. *Information* **2023**, *14*, 370.
14. D'Angelo, M.; Nanni, L. Deep Learning-Based Human Chromosome Classification: Data Augmentation and Ensemble **2023**.

15. Nazarenko, D.; Kharyuk, P.; Oseledets, I.; Rodin, I.; Shpigun, O. Machine learning for LC–MS medicinal plants identification. *Chemometrics and Intelligent Laboratory Systems* **2016**, *156*, 174–180.
16. Kumar, N.; Belhumeur, P.N.; Biswas, A.; Jacobs, D.W.; Kress, W.J.; Lopez, I.C.; Soares, J.V. Leafsnap: A computer vision system for automatic plant species identification. In Proceedings of the Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7–13, 2012, Proceedings, Part II 12. Springer, 2012, pp. 502–516.
17. Kadir, A.; Nugroho, L.E.; Susanto, A.; Santosa, P.I. Leaf classification using shape, color, and texture features. *arXiv preprint arXiv:1401.4447* **2013**.
18. Wu, S.G.; Bao, F.S.; Xu, E.Y.; Wang, Y.X.; Chang, Y.F.; Xiang, Q.L. A leaf recognition algorithm for plant classification using probabilistic neural network. In Proceedings of the 2007 IEEE international symposium on signal processing and information technology. IEEE, 2007, pp. 11–16.
19. Sabu, A.; Sreekumar, K.; Nair, R.R. Recognition of ayurvedic medicinal plants from leaves: A computer vision approach. In Proceedings of the 2017 Fourth International Conference on Image Information Processing (ICIIP). IEEE, 2017, pp. 1–5.
20. Sundara Sobitha Raj, A.P.; Vajravelu, S.K. DDLA: dual deep learning architecture for classification of plant species. *IET Image Processing* **2019**, *13*, 2176–2182.
21. Barré, P.; Stöver, B.C.; Müller, K.F.; Steinhage, V. LeafNet: A computer vision system for automatic plant species identification. *Ecological Informatics* **2017**, *40*, 50–56.
22. Geetharamani, G.; Pandian, A. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers & Electrical Engineering* **2019**, *76*, 323–338.
23. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* **2014**.
24. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
25. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
27. Duong-Trung, N.; Quach, L.D.; Nguyen, M.H.; Nguyen, C.N. A combination of transfer learning and deep learning for medicinal plant classification. In Proceedings of the Proceedings of the 2019 4th International Conference on Intelligent Information Technology, 2019, pp. 83–90.
28. Prashar, N.; Sangal, A. Plant disease detection using deep learning (convolutional neural networks). In Proceedings of the Second International Conference on Image Processing and Capsule Networks: ICIPCN 2021 2. Springer, 2022, pp. 635–649.
29. Akhtar, M.J.; Mahum, R.; Butt, F.S.; Amin, R.; El-Sherbeeny, A.M.; Lee, S.M.; Shaikh, S. A Robust Framework for Object Detection in a Traffic Surveillance System. *Electronics* **2022**, *11*, 3425.
30. Leonidas, L.A.; Jie, Y. Ship classification based on improved convolutional neural network architecture for intelligent transport systems. *Information* **2021**, *12*, 302.
31. Ahsan, M.; Naz, S.; Ahmad, R.; Ehsan, H.; Sikandar, A. A deep learning approach for diabetic foot ulcer classification and recognition. *Information* **2023**, *14*, 36.
32. Al-Kaltakchi, M.T.; Mohammad, A.S.; Woo, W.L. Ensemble System of Deep Neural Networks for Single-Channel Audio Separation. *Information* **2023**, *14*, 352.
33. Medicinal Leaf Dataset - Mendeley Data. <https://data.mendeley.com/datasets/nnytj2v3n5/1>. (Accessed on 09/12/2023).
34. Patil, S.S.; Patil, S.H.; Pawar, A.M.; Patil, N.S.; Rao, G.R. Automatic Classification of Medicinal Plants Using State-Of-The-Art Pre-Trained Neural Networks. *Journal of Advanced Zoology* **2022**, *43*, 80–88.

35. Ayumi, V.; Ermatita, E.; Abdiansah, A.; Noprisson, H.; Jumaryadi, Y.; Purba, M.; Utami, M.; Putra, E.D. Transfer Learning for Medicinal Plant Leaves Recognition: A Comparison with and without a Fine-Tuning Strategy. *International Journal of Advanced Computer Science and Applications* **2022**, *13*.
36. Almazaydeh, L.; Alsalameen, R.; Elleithy, K. Herbal leaf recognition using mask-region convolutional neural network (mask R-CNN). *Journal of Theoretical and Applied Information Technology* **2022**, *100*.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.