



Assessing deep convolutional neural network models and their comparative performance for automated medicinal plant identification from leaf images

Biplob Dey^{a,c}, Jannatul Ferdous^a, Romel Ahmed^{a,c,*}, Juel Hossain^b

^a Department of Forestry and Environmental Science, Shahjalal University of Science and Technology, Sylhet, 3114, Bangladesh

^b Department of Soil Science, University of Chittagong, Chattogram, 4331, Bangladesh

^c Center for Research in Environment, iGen and Livelihoods (CREGL), Sylhet 3114, Bangladesh

ARTICLE INFO

Keywords:

Artificial intelligence
Species identification
Deep learning
Medicinal plants
CNN

ABSTRACT

Medicinal plants have got notable attention in recent years in the field of pharmaceutical and drug research. The high demand of herbal medicine in the rural areas of developing countries and drug industries necessitates correct identification of the medicinal plant species which is challenging in absence of expert taxonomic knowledge. Against this backdrop, we attempted to assess the performance of seven advanced deep learning algorithms in the automated identification of the plants from their leaf images and to suggest the best model from a comparative study of the models. We meticulously trained VGG16, VGG19, DenseNet201, ResNet50V2, Xception, InceptionResNetV2, and InceptionV3 deep neural network models. This training utilized a dataset comprising 5878 images encompassing 30 medicinal species distributed among 20 families. Our approach involved two avenues: the utilization of public data (P_I) and a blend of public and field data (PF_I), the latter featuring intricate backgrounds. Our study elucidates the robustness of these models in accurately identifying and classifying both interfamily and interspecies variations. Despite variations in accuracy across diverse families and species, the models demonstrated adeptness in these classifications. Comparing the models, we unearthed a crucial insight: the Normalized leverage factor (γ_w) for DenseNet201 stands at 0.19, elevating it to the pinnacle position for P_I with a remarkable 99.64 % accuracy and 98.31 % precision. In the PF_I scenario, the same model achieves a γ_w of 0.15 with a commendable 97 % accuracy. These findings serve as a guiding beacon for shaping future application tools designed to automate medicinal plant identification at the user level.

1. Introduction

Almost 70 % people in developing countries directly depend on traditional medicine for primary health care and treating common ailments [1]. Due to limited government medical facilities and the high cost of allopathic medicine in developing countries, many people rely on herbal medicine [2]. Besides, the emerging drug industries in developed countries depend on medicinal plants to some extent for the pharmaceutical products [3]. In the world, 18 % of the top 150 prescription medications, along with 25 % of modern pharmacopoeia, are plant-based. Asia is one of the major hubs for the world's bioresources [4], and around 50 % of the world's

* Corresponding author.

E-mail address: romel-fes@sust.edu (R. Ahmed).

<https://doi.org/10.1016/j.heliyon.2023.e23655>

Received 28 May 2023; Received in revised form 30 November 2023; Accepted 8 December 2023

Available online 13 December 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

traditional medicine exports are made up of Asian medicinal plants [5]. The recognition of plant medicines among the public improved as a result of the positive and reassuring results of several clinical trials. There are studies that shed light on and spur the therapeutic use of plant-based medicines by examining the many active ingredients of herbs along with their clinical functions [6].

In most of the cases, people who are older than 45 years seem to be better knowledgeable about medicinal plants, likely due to their greater experiences in dealing with these plants grown in their surrounding habitats [7]. In fact, ethnobotanical knowledge is gradually eroding among young people, importantly among the educated youth, that poses a threat to transfer of the knowledge to the future generations. The paucity of standardized preparation techniques and scientific evidences about their effectiveness and possible toxicity would result in inefficient use of the plants or misidentification of the species, consequently impacting the potential use of herbal medicine in the future generation [8]. Poisoning from medicinal plants is commonly documented as a result of either incorrect identification of the plant when sold or inadequate preparation and administration by untrained individuals. Without any prior botanical knowledge, finding information about medicinal plants from books or internet can be challenging and time-consuming, especially when dealing with diverse local names for the same species. According to Refs. [9,10], there is a significant disparity of herbal knowledge acquired by people in cities or researchers in compared to the knowledge of tribes or villages. Low-cost, efficient, and accurate identification of medicinal plants can drive a revolution in the field of medical research as well as the conservation of these precious natural resources. Leaves of the plants could perform a significant role in plant identification because of their uniqueness and abundance throughout the seasons.

Adoption of cutting-edge technology reduces the labor involved in expert inspection for detecting signs of disease, nutrient deficiency, and plant identification [11,12]. Significant progress has been seen in recent decades with a variety of well-known architectures, including InceptionV3, GoogleNet, VGG16, AlexNet, and ResNet for handling a wide range of image classification tasks with the introduction of deep learning techniques. Numerous studies [11–15] have used these pre-trained models for identification of plants and diseases owing to the outstanding performance of these models. Applications of deep learning pre-trained models are more accurate due to the fact that these architectures were designed to identify the 1000 classes in the Imagenet dataset [16].

A number of studies [17–24] have been carried out in recent years to provide tools for the identifications of medicinal plants. Le et al. [25] used modified kernel descriptor and support vector machine for visual identification of Vietnamese medicinal plants. Although, overall performance was satisfactory in their study, the highest accuracy of identifying individual species was around 80 %. Authors in other studies [20,26,27] used ML algorithms like k-nearest neighbour (KNN), naïve Bayes, Support vector machine (SVM), Decision tree, neural networks for the automated medicinal plant identification, the number of species used in their studies varies from 10 to 24 and trained maximum 97 % accuracy. Moreover, a study [18]) obtained 85 % accuracy employing Deep Neural Networks (DNN). Using transfer learning with convolutional neural network for identifying medicinal plants using 10 species obtained 98.7 % accuracy [28]. Pushpa and Rani [29] utilized the Ayur-PlantNet deep convolutional neural network to conduct a comparative analysis involving MobileNetV3Large, Densenet121, Resnet34, VGG16, and Resnet50 for ayurvedic plant identification. Notably, the highest accuracy of 92.27 % was achieved by Ayur-PlantNet. Muneer and Fati [30] leveraged shape and texture features of leaves for the classification of Malaysian leaf types. Their study evaluated the performance of both DNN and ML classifiers on a comprehensive dataset. Additionally, Pushpanathan [31] delved into the identification of twelve locally found perennial herbs in Malaysia, renowned for their medicinal significance. Kan [32] focused on twelve Chinese medicinal plant species, extracting distinctive shape and texture attributes. Their exploration also highlighted the distinctive characteristics of plant photographs and medicinal variations. Amuthalingeswaran [18] employed a DNN to classify medicinal plant species across four distinct plant types, achieving a notable average accuracy. VijayaLakshmi and Mohan [33] introduced a method that harnessed texture, shape, and color attributes for plant leaf type identification, employing a fuzzy relevance vector machine for classification. The authors encountered challenges in distinguishing leaves with similar shapes, handling images with shadows, addressing cluttered backgrounds, and effectively recognizing flawed leaves. Russel and Selvaraj [34] introduced a multiscale parallel DNN for disease detection based on leaf images. Their experimental endeavors were conducted using existing datasets captured in controlled environments with plain backgrounds.

Previous studies [26,35–37] have predominantly relied on public datasets or focused on specific backgrounds, leaving a gap in research concerning the identification of medicinal plants from raw data with complex backgrounds that collected from different geographic distribution. In real-world scenarios, variations in leaf size and shape can significantly impact accurate identification. While a few studies have utilized real field datasets [19,38] with complex backgrounds, there is a need for comprehensive and rigorous investigations to validate and establish models for rapid and precise identification of medicinal plants under different geographic field conditions. Furthermore, a thorough comparative study utilizing advanced deep learning algorithms to assess the performance and predictive capabilities of these models across different species and families has not been conducted yet. This research gap needs to be addressed in order to gain a better understanding of the strengths and limitations of different algorithms in identifying medicinal plants.

To bridge this research gap, our study aims to analyze the performance of seven advanced deep learning algorithms (VGG16, VGG19, DenseNet201, InceptionV3, ResNet50V2, Xception, InceptionResnetV2) in a family-wise manner using various public data sources and our own real field images of medicinal plants. By utilizing data sources that vary in terms of resolution and contrast, reflecting the challenges encountered in real-world scenarios. These seven DNN algorithms were selected based on their proven performance and success in various image recognition tasks. We also performed their comparative assessment to find out best model for the advancement of drug development whether it be non-pharmacoepial, pharmacoepial, or synthetic drugs or benefiting the local communities.

2. Materials and methods

2.1. Data acquisition and pre-processing

Medicinal plant leaf images from 30 species belonging to 20 families (Table-1) were gathered from the data archive of Kaggle [39] with a plain background (P_1) and from local field images (F_1) with complex real backgrounds (Fig. 1 (d, f, g, h, k, l, m, p)). The images were saved in JPG format with corresponding scientific names. The Pillow library (Version 8.4.0) was used for resizing the images to 224×224 pixels before using a pretrained deep learning model. Resizing a large image reduces computational load on the GPU and potentially speeds up the processing of the model. Fig. 1(a–p) shows some representative images that we used to train the DNN models.

2.2. Deep convolutional neural networks

The study evaluated seven distinct Deep Convolutional Neural Networks (DCNN) (Fig. 2). DCNN models use convolutional layers to extract features from medicinal plant leaf images in a spatial hierarchy way, where the lower layers can learn simple features such as edges and textures and higher layer learn complex features. Leaf input images (224×224) comprised of three matrices or color (I_C) channels (RGB). The Convolution Layer plays a crucial role in our Deep Convolutional Neural Network models. It is responsible for extracting and learning features from input images, producing a feature map as its output.

After the convolution operation, a non-linear activation function was applied to allow the network to learn more complex representations of the input image. For input image height (I_H), width (I_W) this CL can be represented according to Eq. (1).

$$\dim(\text{image}) = (I_H, I_W, I_C) = (224, 224, 3) \quad (1)$$

The kernel or filter (k) must have a same i_C of the image. Then, the filter dimension can be calculated using Eq. (2).

$$\dim(\text{filter}) = (k, k, I_C) \quad (2)$$

For the input image of medicinal plant leaf (I), padding (p) and stride (s) the tensor dimension calculated based on Eq. (3).

$$(I_H, I_W, I_C) \times (k, k, I_C) = \left(\left[\frac{I_H + 2p - k}{s} + 1 \right], \left[\frac{I_W + 2p - k}{s} + 1 \right] \right), \text{ for } s > 0 \quad (3)$$

The next step in the process involved a 2D Convolution Layer (CL) with a max pooling function and a Rectified Linear Unit (ReLU)

Table 1

Medicinal plant species list that used for species identifications with their common name in respect of Bangladesh and India.

Scientific name	Family	Local name	Image number	
			Public (P_1)	Field image (F_1)
<i>Amaranthus viridis</i> L.	Amaranthaceae	Data shak, Marissag	122	157
<i>Artocarpus heterophyllus</i> Lam.	Anacardiaceae	Jackfruit, Kanthal	92	133
<i>Brassica juncea</i> (L.) Czern	Basellaceae	Shorisha, Indian mustard	85	102
<i>Azadirachta indica</i> (L.)	Apocynaceae	Neem	95	112
<i>Basella alba</i> (L.)	Apocynaceae	Puishaak	103	98
<i>Carissa carandas</i> (L.)	Brassicaceae	Koromcha, Kilakkai	74	107
<i>Citrus limon</i> (L.) Burm. f. (pro. sp.)	Fabaceae	Lebu, Goranebu	77	115
<i>Ficus auriculata</i> Lour.	Fabaceae	Trimmel, Puroi khak	80	90
<i>Ficus religiosa</i> L.	Lamiaceae	Ashvattha, Peepal	73	93
<i>Hibiscus rosa sinensis</i> L.	Lamiaceae	Joba, China rose	84	100
<i>Jasminum officinale</i> L.	Lamiaceae	Jasmine	71	145
<i>Mangifera indica</i> L.	Lythraceae	Aam, Mango	92	135
<i>Mentha piperita</i> L. (pro. sp.)	Malvaceae	Mentha, Pudina	97	97
<i>Moringa oleifera</i> Lam.	Meliaceae	Shojne, Moringa	77	120
<i>Alpinia galanga</i> (L.) Willd	Moraceae	Kulanjan, Blue ginger	80	87
<i>Muntingia calabura</i> L.	Moraceae	Jamaica cherry, Calabura	86	105
<i>Murraya koenigii</i> (L.) Spreng	Moraceae	Curry leaf	60	113
<i>Nyctanthes arbor tristis</i> Linn.	Rutaceae	Har singar, Shiuli	79	100
<i>Santalum album</i> L.	Rutaceae	Sandalwood	88	95
<i>Syzygium jambos</i> (L.) Alston	Rutaceae	Golapjaam, Mountain apple	78	139
<i>Trigonella foenum-graecum</i> L.	Santalaceae	Fenugreek, Methi	66	92
<i>Syzygium cumini</i> (L.) Skeels	Zingiberaceae	Jamun	78	133
<i>Tabernaemontana divaricate</i> (L.) R. Br. ex Roem. & Schult.	Apocynaceae	Thoka tagar, Pinwheel flower	66	89
<i>Plectranthus amboinicus</i> (Lour.) Spreng	Piperaceae	Indian borage, Mexican mint	85	94
<i>Punica granatum</i> L.	Moringaceae	Dalim	82	141
<i>Psidium guajava</i> L.	Muntingiaceae	Guava	80	167
<i>Nerium oleander</i> L.	Myrtaceae	Kaner, Rokto korobi	70	143
<i>Piper betle</i> L.	Myrtaceae	Betel, Paan	55	165
<i>Ocimum tenuiflorum</i> L.	Oleaceae	Kalotulsi, Tulshi	58	119
<i>Pongamia pinnata</i> (L.) Pierre	Oleaceae	Karanja	61	98

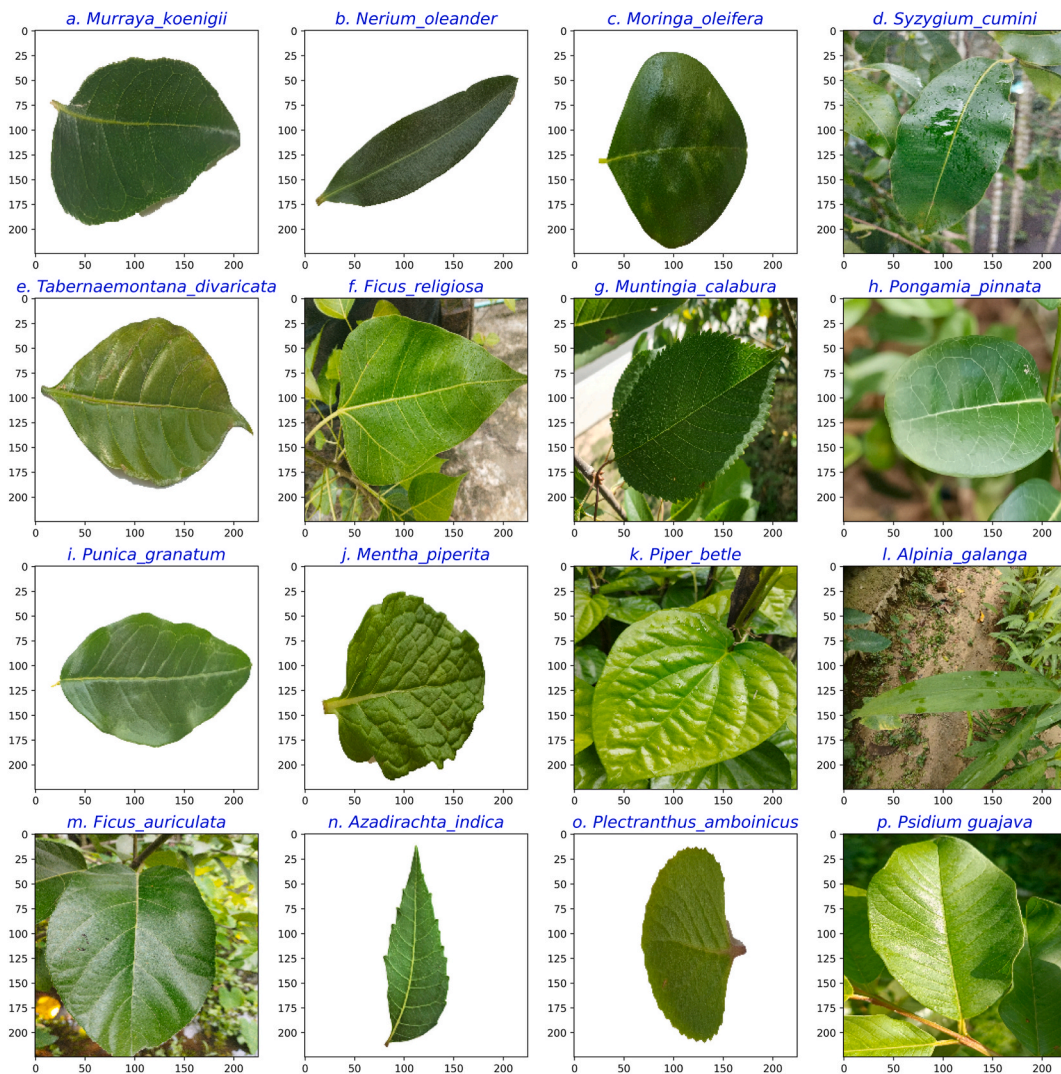


Fig. 1. Representative leaf sample images from 30 medicinal plant species used in the study for automated identification. The images denoted by labels d, f, g, h, k, l, m, p represent real field data, while the remaining images represent public data from Kaggle.

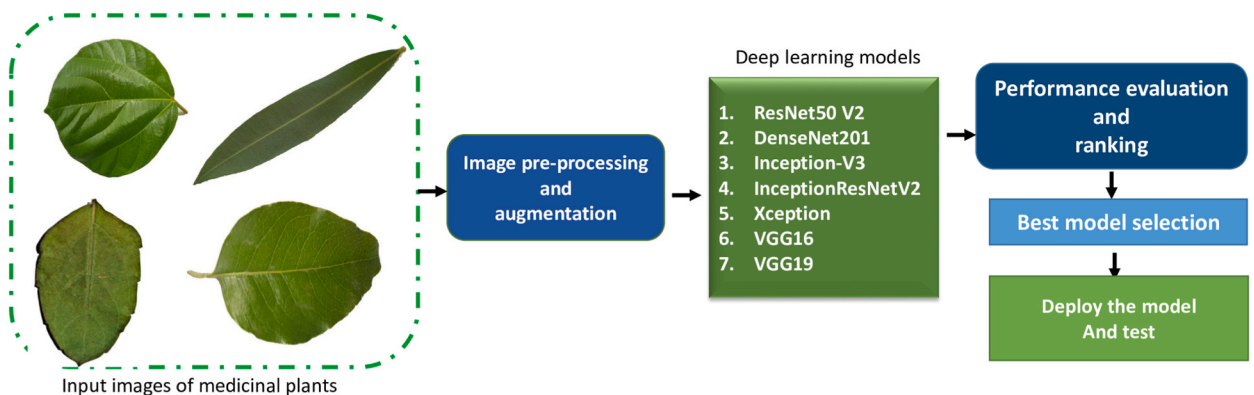


Fig. 2. The overall workflow that employed in this study for medicinal plant species identification using leaf.

activation. The ReLU activation was used to handle nonlinearities and ensure efficient activation, as it does not activate all neurons simultaneously. The CL utilized convolutional products with filters and an activation function (ψ) as inputs. The l th layer can be expressed as Eq. (4).

$$\forall n \in [1, 2, 3, \dots, I_C^{[l]}],$$

$$\text{con}(d^{[l-1]}, f^{(n)})_{x,y} = \psi^{[l]} \left\{ \sum_{i=1}^{[l-1]} I_H \sum_{j=1}^{[l-1]} I_W \sum_{k=1}^{[l-1]} I_C \cdot K_{i,j,k}^{(n)} d_{x+i,y+j-1,k}^{[l-1]} + b_n^{[l]} \right\} \quad (4)$$

Then, in each layer 3×3 max pooling was included to prevent overfitting issue and to make efficient and robust performance. For the pooling function $\phi^{[l]}$ the pooling layer can be expressed by Eq. (5).

$$\text{Pool}(d^{[l-1]})_{x,y,z} = \phi^{[l]} \left(d_{x+i,y+j-1,z}^{[l-1]} \right)_{(i,j) \in \{(1,2,3,4,\dots,i^{[l]})^2\}} \quad (5)$$

The output is finally processed through a flattened layer, followed by a fully connected layer where each neuron is connected to an activation unit and a 40 % dropout is applied. This fully connected layers received input a $a^{[i-1]}$ vector and give back a $a^{[i]}$ vector, for the i th layer of j th node, then we can express it by Eq. (6),

$$Z_j^{[i]} = \sum_{l=1}^{n_{i-1}} W_{j,l}^{[i]} d_l^{[i-1]} + b_j^{[i]} \quad (6)$$

The input features were then processed through the ReLU function, which classifies the images into specific labels. The softmax activation function makes the final decision on classification based on the output of the neurons' classification labels.

Besides three fully connected layers (FCL) of VGG 16 (Fig. 3a) and VGG 19 (Fig. 3b), they have 16 layers with 13 convolution layers and 19 layers with 16 convolution layers, respectively. Both models use max-pooling in feature maps for reducing the spatial dimensions and enhance their translation invariance. To reduce the volume size the models modified with pre-trained weights by incorporating max-pool function, as well as a Softmax classifier for the output from the previous fully connected layer. These models differ in their deep layer architecture and the use of small convolutional filters, which enable them to learn detailed features from images.

DenseNet201 is a CNN architecture that is employed for this leaf image classification tasks. Which is a variation of DenseNet architecture, and dense connections between layers (Fig. 3c). Which contains 201 total layers and a total of 19,628,095 parameters, with both parameters trainable and non-trainable. Additionally, the layers in each block are connected to all preceding layers, facilitating feature learning from all previous layers [12], and enhancing the flow of information and gradients through the network. The model also employs a global pooling and full-connected layer, also implements "transition layers" to decrease the number of channels and spatial resolution between dense blocks.

An Inception model comprises of multiple parallel branches, each with a different convolutional filter size ($1 \times 1, 3 \times 3, 5 \times 5$) with a max-pool layer. The 36 convolutional layers of model act as the foundation for constructing the network architecture (Fig. 4a). In order to enable the model to learn features at various scales and capture various levels of abstraction, these branches are then concatenated along the channel dimension. This also employs a method known as "factorization," which lowers the computational cost

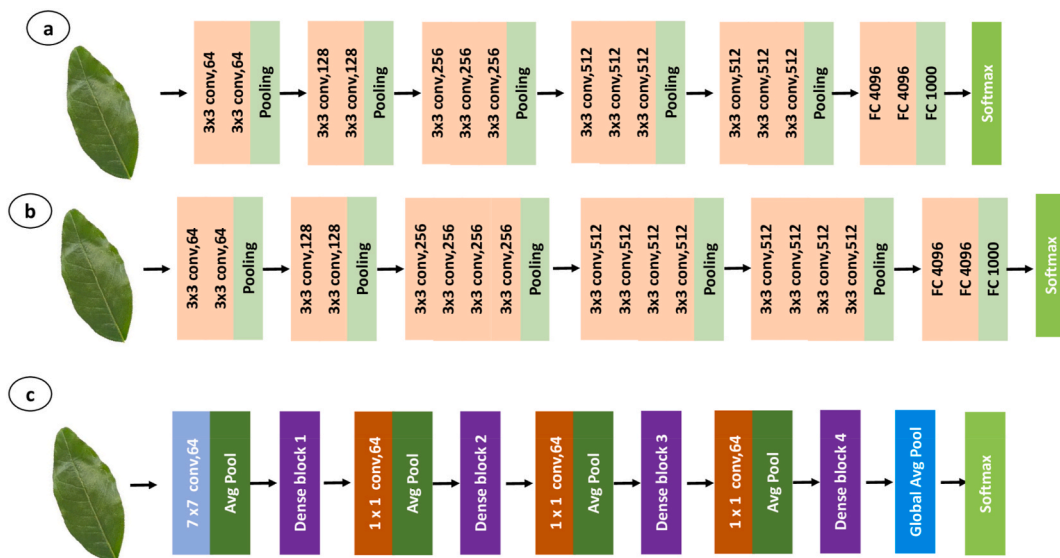


Fig. 3. Basic deep neural network architecture of (a) VGG16, (b)VGG19, (c) DenseNet201.

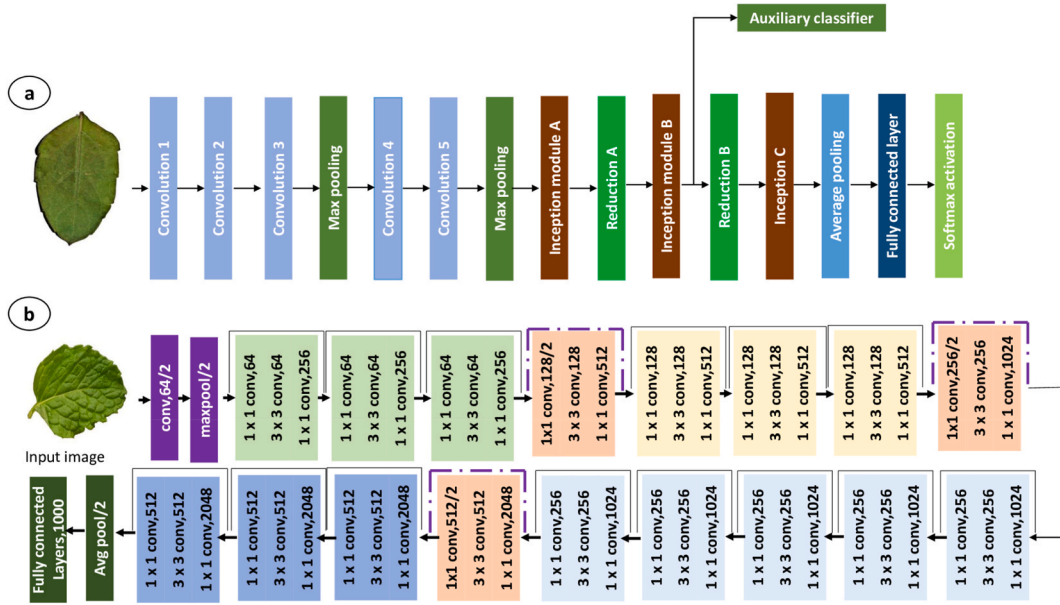


Fig. 4. Basic deep neural network architecture of (a) Inception, (b) ResNet50.

of the 3×3 and 5×5 convolutional filters by utilizing a 1×1 convolutional filter to cut down on the number of input channels before the 3×3 or 5×5 convolutional filter is applied.

ResNet50V2 is a reformed variant of ResNet50, incorporating a new residual-unit and featuring 49 convolutional layers, 1 fully connected layer for classification, one average pooling and one max-pool layer (Fig. 4b). ResNet152V2 is similar to ResNet152 and is a modified version of the latter, which comprises 151 convolutional layers with one max-pool, one average pooling, and one full-connected layer for the classification purpose.

Inception is an effective architecture for classifying images. It is distinguished by its use of residual connections and depth wise separable convolutions, both of which increase computing efficiency. Which enhances the network’s gradients and information flow. This also uses “residual connections”, where the input of a layer is added to its output before passing it through the next layer. This architecture is especially appropriate for embedded and mobile devices because they have constrained processing resources.

2.3. Hyperparameter setting

We employed two kinds of data to train our DCNN models. The first kind was using public data (P_1) exclusively, while the second kind (PF_1), combined public data with field data ($P_1 + F_1$) that had complex backgrounds. For the training process, we utilized 85 % of the dataset in both types. Additionally, 5 % of the data was allocated for validation during each epoch to monitor the model performance and prevent overfitting. Furthermore, we ensured an unbiased evaluation of the model’s performance by using a completely separate dataset for testing, 10 % from both P_1 and F_1 dataset. This test dataset contained data that the model had never encountered before, replicating its performance on new and real-world data scenarios. Through this approach, we could meter how well the model generalized to previously unseen samples, which is the ultimate objective of a DNN model.

During the training phase, we used batch size 32 which refers to the number of input images the model processes at one time before updating the model’s parameters. While a bigger batch size can lead to more precise gradient estimations, it also necessitates more memory [40]. Moreover, average pooling is used in DCNN to reduce the spatial dimensions of feature maps that work by taking the average values in the feature map region, resulting in a smaller output feature map. “softmax” function is also employed in the output layer for mapping the output of the final layer for improving the interpretability of the model’s predictions. To ensure the reproducible results of the models, the random generator’s seed is set to random state 42. The parameters of the model are modified during the training phase based on input data and the optimization algorithm. This makes it possible to train and evaluate the model consistently, which is crucial for comparing its performance across several runs or implementations.

For turning both rows and columns horizontally, the “horizontal flip” option is set to “True” in the ImageDataGenerator constructor. This technique is used to increase the amount of data available for training from the existing ones. This allows the model to be more resistant to translations in the test data by enabling the model to learn features that are invariant to horizontal translations. For multiclass classification, the “categorical” class mode was employed for labelling the image data with one-hot encoded categorical labels. When there are more than two classes, this mode is helpful as it enables the models to understand the relationships between classes and forecast numerous classes. Additionally, each image was randomly rotated between $[32, +32]$ degrees with 20 % shifting along the X-axis and Y-axis. We used shearing and zoom range 0.2–20 % zoom in and 20 % zoom out of the image, respectively. Additionally, pixel intensity was arbitrarily scaled between $[0.5, 1.5]$. We employed an “Adam” optimizer function, which is an

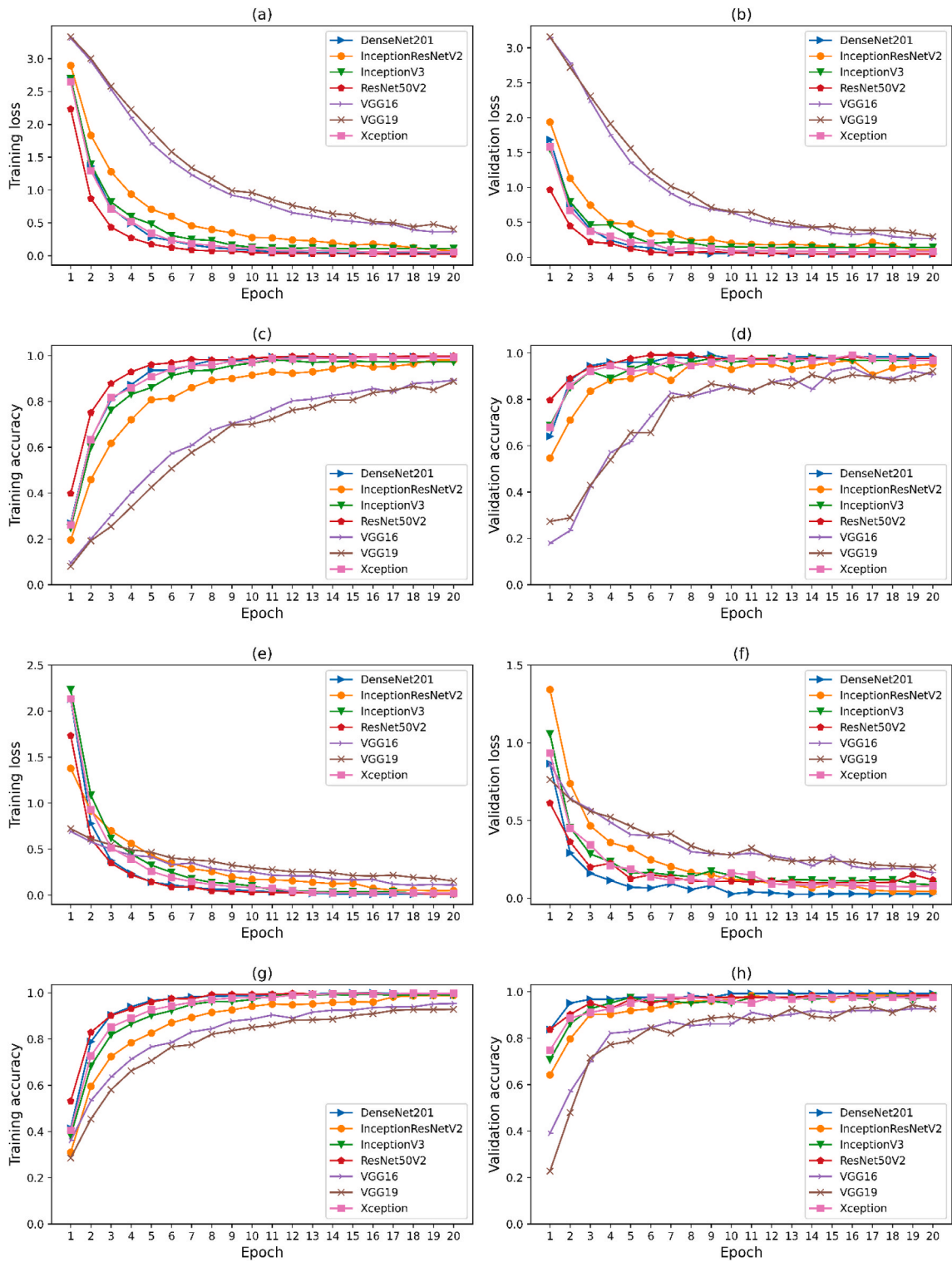


Fig. 5. Plots of loss and accuracy of training and validation according to epoch of seven neural network models. Where, a,b,c,d for mix dataset (PF₁) and e,f,g,h for public dataset (P₁).

extension of the stochastic gradient descent optimization algorithm. This function adapts the learning rate for each parameter individually [41], that improves the stability and performance of the model during training, also corrects the estimator bias of the first and second moments of the gradient. In this study, we evaluate the performance of DCNN models using the sensitivity and accuracy based on [11]. Due to the interdependency of various evaluation parameters, ranking AI models based solely on accuracy becomes challenging. To address this issue, we implemented the “Dey & Ahmed method” [12,42] (supplementary file S3).

3. Results

3.1. Training and validation of DCNN models

A comparison experiment was conducted using seven different models, namely VGG16, VGG19, ResNet50V2, Xception, InceptionResnetV2, DenseNet201, and InceptionV3, based on P_1 and PF_1 . Fig. 5 presents the accuracy and loss curves for all seven models. Initially, during the early epochs, the losses of the PF_1 dataset were higher (Fig. 5a) than those of the P_1 dataset (Fig. 5e), such differences can be attributed to the complexity of the backgrounds in the field data. Among the models, VGG19 showed the highest training losses for the PF_1 dataset and InceptionV3 for the P_1 . As the number of epochs increased, the gap between the training and validation accuracy decreased, indicating that the models were becoming more robust (Fig. 5 c-d,g-h). Initially, during the first four epochs (Fig. 5 a-b,e-f), the training loss was higher for most models. However, it gradually reduced after the 10th epoch, indicating improved model learning and eventually, all models started to converge. Fig. 5 (c,g) demonstrates that the models reached their peak performance on the training data, and further training is unlikely to significantly improve performance. However, for the validation dataset, the increasing validation accuracy over time suggests that the models were generalizing well to new, unseen data and were not overfitting to the training data (Fig. 5 d,h). Additionally, the high training accuracy and similarly high validation accuracy indicate that the models were not overfitted to the training data. This balance between training and validation accuracy shows that the models were able to generalize effectively to new data.

3.2. Evaluation of DCNN models

The comprehensive evaluation of the seven DCNN models is detailed in Table 2. The accuracy ratings for each model varied between 0.90 and 0.996 for both P_1 and PF_1 . However, it was evident that the PF_1 dataset exhibited comparatively lower performance than the P_1 dataset in terms of accuracy, precision, recall, and F1-score (Table 2). In the PF_1 , the InceptionResNetV2 model achieved an impressive recall score of 0.97, signifying its ability to correctly identify 97 % of all positive instances. This was the highest recall score among all the models, and it also achieved 97 % accuracy and 94 % precision. On the other hand, in the P_1 , the ResNet50V2 model achieved a precision of 0.9732, indicating a low false positive rate and accurate identification of 97.32 % of medicinal plant species. Among all the DCNN models, DenseNet201 stood out with the highest precision and accuracy for both the P_1 and PF_1 datasets (Table 2). DenseNet201 surpassed all other models tried in the study.

It is important to consider multiple performance metrics, to evaluate the effectiveness of the adopted DCNN models thoroughly. These metrics provide different perspectives on the model’s performance and help ensure accurate predictions and an overall measure of its effectiveness. According to normalized leverage factor (NLF), DenseNet201 attained a weight of 0.19, ranking it as the best-performing model, while the weight of NLF for ResNet50V2 was 0.1752, placing it as the second-best model in the ranking (supplementary file S3). For both P_1 and PF_1 , though each model shows different performance matrices, the overall rankings were the same.

The confusion matrix (CM) describes the functional abilities of the seven implemented DCNN models used to predict each of the 30 medicinal species (Fig. 6–7). The horizontal and vertical axes of the matrix represent true and predicted labels, where the normalized diagonal values represent the correct prediction numbers for each species, ranging from 0 to 1. This also represents the true positive of the specific species. As same as performance result, the rate of misclassification is comparatively high in PF_1 . In DenseNet201 of P_1 , 28 species were identified with 100 % accuracy while identification accuracy of the other two species namely *Ficus auriculata* and *Moringa oleifera* were 93 % and 96 % respectively (Fig. 6). On the contrary, in PF_1 identification performance of DenseNet201 was not as strong as P_1 , where classification accuracy of 3 species ranged from 88 to 93 % and the rest was 100 % (Fig. 7), though both were best performing model based on their respective dataset.

Table 2

The performance scores of seven deep convolutional neural network models were evaluated for identifying medicinal plants in our dataset. In this evaluation, γ_{ω_m} was used as the normalized leverage factor.

	Public data (P_1)						Public field mix data (PF_1)					
	Accuracy	Precision	Recall	F1	γ_{ω_m}	Rank	Accuracy	Precision	Recall	F1	γ_{ω_m}	Rank
Resnet50V2	0.986	0.973	0.984	0.977	0.176	2	0.951	0.940	0.947	0.953	0.118	2
DenseNet201	0.996	0.983	0.995	0.973	0.191	1	0.974	0.968	0.970	0.962	0.152	1
VGG16	0.946	0.949	0.931	0.921	0.104	6	0.919	0.917	0.891	0.911	0.074	6
VGG19	0.911	0.918	0.894	0.901	0.071	7	0.902	0.886	0.902	0.894	0.063	7
InceptionV3	0.978	0.982	0.964	0.969	0.166	3	0.940	0.933	0.927	0.931	0.098	3
Xception	0.975	0.963	0.967	0.972	0.155	4	0.959	0.944	0.959	0.960	0.130	4
Inception ResNetV2	0.963	0.968	0.945	0.958	0.138	5	0.968	0.943	0.976	0.962	0.140	5

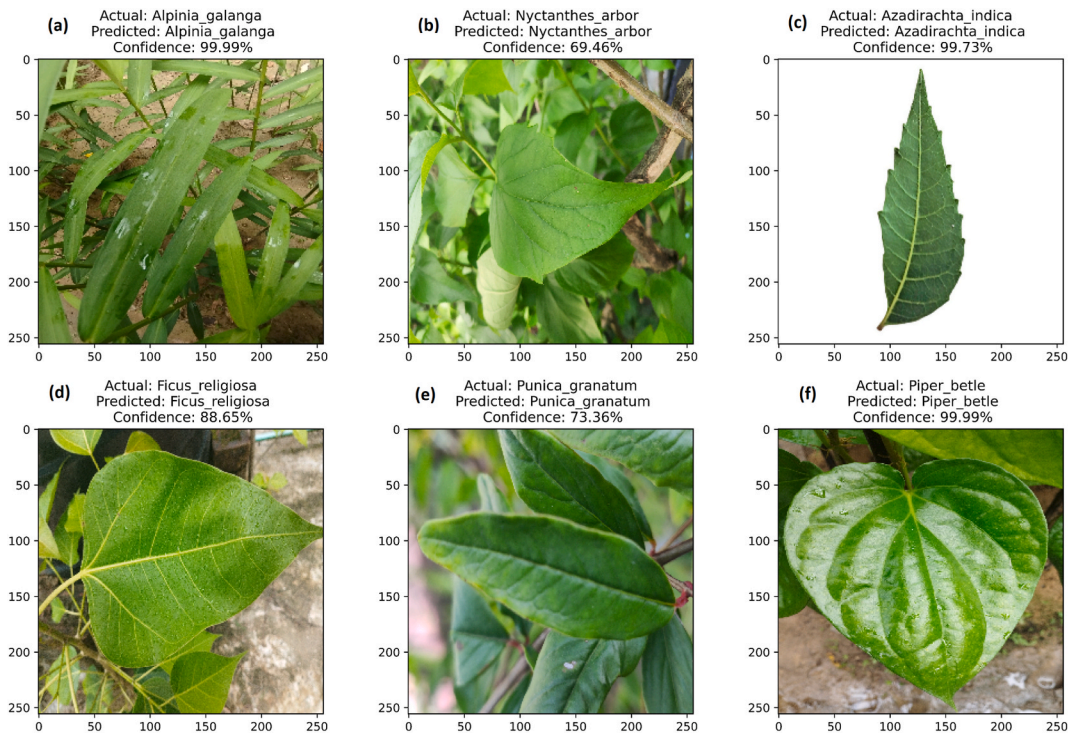


Fig. 8. Test performance of the leaf based automated medicinal plant detection using trained DenseNet201 model.

4. Discussion

While automatic identification has exhibited favorable outcomes, specific constraints persist, demanding attention and resolution in the forthcoming years. Our investigation into the plant species parts and images that influence accuracy can serve as a compass for prospective refinements. Remarkably, our findings unveiled that the proficiency of identification remained consistent for both solitary

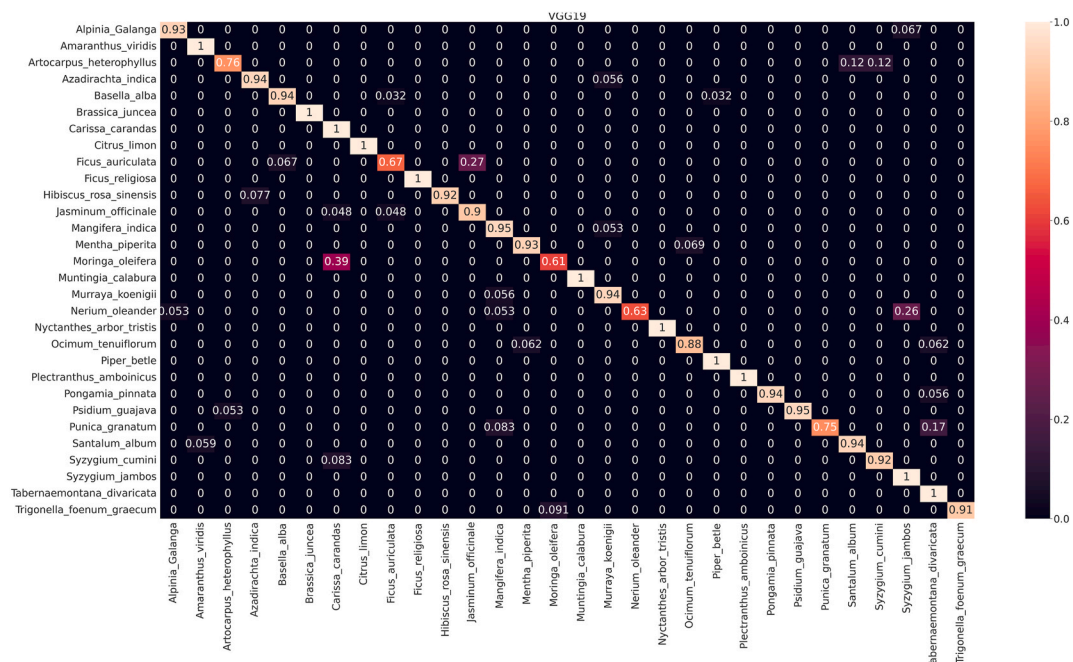


Fig. 9. Confusion matrix for automatic identifying medicinal plants using VGG19 for public dataset (P₁).

images extracted from our source and observations amassed from the field, often encompassing multiple images. It is noteworthy that the precision in identifying both uncommon and prevalent species across distinct regions remained unwavering, and the primary species habitat did not exert influence on outcomes. In contrast to our anticipations, the analogous performance observed between database images and field observations could plausibly be ascribed to the meticulous curation of the observation database, ensuring images were discernible to experts. This equitable performance accentuates the application's potential for versatile utility, encompassing tasks such as identifying endangered plants or conducting comprehensive biodiversity evaluations, particularly for frequently encountered species. The accomplished performance of this application implies its potential for extended employment. Furthermore, it harbors the prospect of aiding in the identification of the burgeoning trove of digitized plant image collections, contingent upon the database's image quality adhering to reasonable standards.

4.1. Intra-family, inter-family, and species-level performance

To reveal the model's ability to distinguish broader taxonomic categories, it is crucial to scrutinize family-level identification. This analysis complements species-level evaluation by highlighting potential challenges in classifying higher-level groups. Addressing errors at the family level can enhance the model's understanding of underlying patterns and improve its accuracy across different taxonomic levels, ensuring a more comprehensive and reliable classification system. The confusion matrix (CM) of least performing model (VGG19) shown in Fig. 9 demonstrates the analyses of inter-family and intra-family performance, while rest of the CM are reported in supplementary files S1 and S2. Among the models evaluated, DenseNet201 is the best performer. However, it recognized *Ficus auriculata* (Fabaceae family) with an accuracy of 93 % (Fig. 6) which was misidentified as *Jasminum officinale* (Lamiaceae family) in around 7 % cases. Just above 4 % *Moringa oleifera* of Meliaceae family was misclassified as *Carissa carandas* of Brassicaceae by this model.

On contrary, in the VGG19 model for P_1 (Fig. 9), predicts incorrectly around 5 % *Jasminum officinale* as *Ficus religiosa*, both of which belong to Lamiaceae family. On the other hand, 6 % *Ocimum tenuiflorum* (Oleaceae), 5.6 % *Pongamia pinnata* (Oleaceae), 1.7 % *Punica granatum* (Moringaceae) misclassified as *Tabernaemontana divaricata* (Apocynaceae). Additionally, ResNet50V2 did not show intra-family misclassification issue like DenseNet50V2. Such as, 8.7 % *Moringa oleifera* of Meliaceae family identified as *Carissa carandas* that belongs to the Brassicaceae family in the ResNet50V2 (Figure S1.5). In the case of P_{F_1} , we observed inter-family misclassification rather than intra-family misclassification. Specifically, the Fabaceae (*Citrus limon*, *Ficus auriculata*) family demonstrated a tendency to be misclassified with the Moraceae (*Muntingia calabura*) and Moringaceae (*Punica granatum*) families (Fig. 10). Conversely, the Moraceae and Moringaceae families were frequently confused with the Oleaceae family. This indicates that all models were not efficient for prediction or classification of species or families based on leaf features. Additionally, Table-2, and Fig. (6,7) clearly suggest that implementation of DenseNet201 can solve the problem of medicinal plant identification that is comparatively less biased and free from intrafamily and interspecies misidentification.

4.2. Reason of misidentification of species

Misclassification of medicinal plants may arise due to diverse factors inherent to the classification process. A comprehensive comprehension of these underlying causes is of paramount significance for the enhancement and refinement of the accuracy inherent in the identification model. Medicinal plants frequently exhibit analogous morphological attributes encompassing leaf configuration, pigmentation, and dimensions [43]. Such shared resemblances can engender perplexity during image-oriented classification endeavors, particularly when discerning between closely akin species. The intrinsic variability in growth conditions, soil composition, and climatic nuances introduces dissimilarity in plant morphology and foliar manifestation [11]. These environmental intricacies may engender incongruities in feature extraction, subsequently impeding precise identification efforts. Certain medicinal plants boast elaborate and intricate leaf structures, presenting intricate challenges for automated systems that necessitate intricate feature extraction methodologies to effectively encapsulate pertinent information [44]. Within a single species, considerable diversity in leaf characteristics may prevail. Aspects such as age, well-being, and genetic heterogeneity contribute to this multifaceted variability, thereby complicating the generalized depiction of plant attributes. The absence of contextual insights, encompassing factors such as habitat, geographical coordinates, and correlated flora, may serve as impediments to the precise discernment of species.

4.3. Significance of the study

Local guides or books on medicinal plants are common sources for identification, however, they can be difficult to use and are sometimes unreliable. Since key books are beneficial when the user is aware with the observable plant family, which the purpose accomplished most of the cases, it has the ability to direct the observer in the right direction and help bring people closer to plants. Combining the AI program and the key book could be a useful tool for fieldwork for observers with limited experience. Recent approaches for the similar identification strategy achieved more reliable results (Table-3), which is also in line with our results. This also indicates the robustness of our study as we obtained almost 98 % accuracy. Our fine-tuned DenseNet50v2 an updated variant of DenseNet50, has the ability to efficiently propagate gradients and reduce the number of parameters compared to traditional convolutional neural networks.

Our findings of automated medicinal species identifications through deep convolutional neural network approaches using images offer excellent opportunity in the fields of research, academia, pharmaceuticals industry, pharmacognosy, phytochemistry, and horticulture. However, it never eliminates the need for taxonomists; rather, it only speeds up the process of accurate species

localized data could result in optimal performance. Advancing the user interface and refining identification algorithms are pivotal for future enhancements. Leveraging deep learning algorithms holds the potential for effective real-world identification solutions.

It is important to acknowledge and elaborate on certain limitations that we encountered during our research. Sometimes it's hard to automatically identify plant species, especially in places like Bangladesh and India with many different species. One way to make it better is by using geographical information from national or international sources, like where these plants are usually found. But this introduces its own problems because species can move to new places, and the same species can look different depending on their age and where they grow. Having more pictures of the same species is important for accurate identification, and the leaves change throughout the year, which makes it even more complicated. Addressing these limitations through further study holds the potential to refine and enhance the robustness of our approach.

However, our study has demonstrated the feasibility of using leaves from a wide variety of plant species, spanning different botanical families, for medicinal species identification. This finding holds significant promise for the pharmaceutical industry, as it can facilitate the identification and utilization of plant species with medicinal properties.

Code availability

Codes are available in this repository <https://github.com/biplobforestry/Automated-medicinal-plant-identification-.git>.

Data availability statement

Data will be made available on request.

Funding

This research did not receive any funding from public, private, or not-for-profit organizations.

CRediT authorship contribution statement

Biplob Dey: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jannatul Ferdous:** Formal analysis, Visualization, Writing – original draft. **Romel Ahmed:** Conceptualization, Investigation, Project administration, Resources, Supervision, Writing – review & editing. **Juel Hossain:** Data curation, Resources, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e23655>.

References

- [1] S.M. Jeelani, G.A. Rather, A. Sharma, S.K. Lattoo, In perspective: potential medicinal plant resources of Kashmir Himalayas, their domestication and cultivation for commercial exploitation, *Journal of Applied Research on Medicinal and Aromatic Plants* 8 (2018) 10–25, <https://doi.org/10.1016/J.JARMAP.2017.11.001>.
- [2] J. Nankaya, J. Nampushi, S. Petenya, H. Balslev, Ethnomedicinal plants of the loita Maasai of Kenya, *Environ. Dev. Sustain.* 22 (2020) 2569–2589, <https://doi.org/10.1007/s10668-019-00311-w>.
- [3] K. Chapman, N. Chomchalow, Production of medicinal plants in Asia, *Acta Hort.* 679 (2005) 45–59, <https://doi.org/10.17660/ACTAHORTIC.2005.679.6>.
- [4] C.P. Kala, P.P. Dhyani, B.S. Sajwan, Developing the medicinal plants sector in northern India: challenges and opportunities, *J. Ethnobiol. Ethnomed.* 2 (2006) 32, <https://doi.org/10.1186/1746-4269-2-32>.
- [5] K. Vasisht, N. Sharma, M. Karan, Current perspective in the international trade of medicinal plants material: an update, *Curr. Pharmaceut. Des.* 22 (2016) 4288–4336, <https://doi.org/10.2174/1381612822666160607070736>.
- [6] N. Martínez-Aledo, D. Navas-Carrillo, E. Orenes-Piñero, Medicinal plants: active compounds, properties and antiproliferative effects in colorectal cancer, *Phytochemistry Rev.* 19 (2020) 123–137, <https://doi.org/10.1007/s11101-020-09660-1>.
- [7] Y.M. Mbuni, S. Wang, B.N. Mwangi, N.J. Mbari, P.M. Musili, N.O. Walter, G. Hu, Y. Zhou, Q. Wang, Medicinal plants and their traditional uses in local communities around cherangani hills, western Kenya, *Plants* 9 (2020) 331, <https://doi.org/10.3390/PLANTS9030331>, 9 (2020) 331.
- [8] W.Y. Ng, L.Y. Hung, Y.H. Lam, S.S. Chan, K.S. Pang, Y.K. Chong, C.K. Ching, T.W.L. Mak, Poisoning by toxic plants in Hong Kong: a 15-year review, *Hong Kong Med. J.* 25 (2019) 102–112, <https://doi.org/10.12809/HKMJ187745>.
- [9] S. Roopashree, J. Anitha, T.R. Mahesh, V. Vinoth Kumar, W. Viriyasitavat, A. Kaur, An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach, *Measurement* 200 (2022), 111484, <https://doi.org/10.1016/J.MEASUREMENT.2022.111484>.
- [10] Y. Herdiyeni, N.K.S. Wahyuni, *Mobile application for Indonesian medicinal plants identification using fuzzy local binary pattern and fuzzy color histogram*, in: 2012 International Conference on Advanced Computer Science and Information Systems, ICACSIS, 2012, pp. 301–306.

- [11] B. Dey, M. Masum Ul Haque, R. Khatun, R. Ahmed, Comparative performance of four CNN-based deep learning variants in detecting Hispa pest, two fungal diseases, and NPK deficiency symptoms of rice (*Oryza sativa*), *Comput. Electron. Agric.* 202 (2022), 107340, <https://doi.org/10.1016/j.compag.2022.107340>.
- [12] B. Dey, R. Ahmed, J. Ferdous, M.M.U. Haque, R. Khatun, F.E. Hasan, S.N. Uddin, Automated plant species identification from the stomata images using deep neural network: a study of selected mangrove and freshwater swamp forest tree species of Bangladesh, *Ecol. Inf.* 75 (2023), 102128, <https://doi.org/10.1016/j.ecoinf.2023.102128>.
- [13] K.P. Ferentinos, Deep learning models for plant disease detection and diagnosis, *Comput. Electron. Agric.* 145 (2018) 311–318, <https://doi.org/10.1016/j.compag.2018.01.009>.
- [14] S.P. Mohanty, D.P. Hughes, M. Salathé, Using deep learning for image-based plant disease detection, *Front. Plant Sci.* 7 (2016), <https://doi.org/10.3389/fpls.2016.01419>.
- [15] X. Zhang, Y. Qiao, F. Meng, C. Fan, M. Zhang, Identification of maize leaf diseases using improved deep convolutional neural networks, *IEEE Access* 6 (2018) 30370–30377, <https://doi.org/10.1109/ACCESS.2018.2844405>.
- [16] M. Agarwal, S.K. Gupta, K.K. Biswas, Development of Efficient CNN model for Tomato crop disease identification, *Sustainable Computing: Informatics and Systems* 28 (2020), 100407, <https://doi.org/10.1016/j.suscom.2020.100407>.
- [17] J. Abdollahi, Identification of medicinal plants in ardabil using deep learning: identification of medicinal plants using deep learning, in: *Proceedings - 2022 27th International Computer Conference, Computer Society of Iran, 2022*, <https://doi.org/10.1109/CSICC55295.2022.9780493>.
- [18] C. Amuthalingeswaran, M. Sivakumar, P. Renuga, S. Alexpandi, J. Elamathi, S.S. Hari, Identification of medicinal plant's and their usage by using deep learning, in: *Proceedings of the International Conference on Trends in Electronics and Informatics, 2019, Institute of Electrical and Electronics Engineers, 2019*, pp. 886–890, <https://doi.org/10.1109/ICOEL2019.8862765>.
- [19] L. Gao, X. Lin, Fully automatic segmentation method for medicinal plant leaf images in complex background, *Comput. Electron. Agric.* 164 (2019), 104924, <https://doi.org/10.1016/j.compag.2019.104924>.
- [20] A. Kumar, D.B. Kumar, Automatic recognition of medicinal plants using machine learning techniques, *Gedrag & Organisatie Review* 33 (2020) 166–175, <https://doi.org/10.37896/gor33.01/012>.
- [21] T. Nguyen Quoc, V. Truong Hoang, Medicinal Plant identification in the wild by using CNN, in: *2020 International Conference on Information and Communication Technology Convergence, 2020*, pp. 25–29, <https://doi.org/10.1109/ICTC49870.2020.9289480>.
- [22] S. Patil, M. Sasikala, Segmentation and Identification of Medicinal Plant through Weighted KNN, *Multimedia Tools and Applications, 2022*, <https://doi.org/10.1007/s11042-022-13201-7>.
- [23] C. Sivaranjani, L. Kalinathan, R. Amutha, R.S. Kathavarayan, K.J. Jegadish Kumar, Real-time identification of medicinal plants using machine learning techniques, in: *2019 International Conference on Computational Intelligence in Data Science, Institute of Electrical and Electronics Engineers, 2019*, pp. 1–4, <https://doi.org/10.1109/ICCIDS.2019.8862126>.
- [24] O.A. Malik, N. Ismail, B.R. Hussein, U. Yahya, Automated real-time identification of medicinal plants species in natural environment using deep learning models-A case study from borneo region, *Plants* 11 (2022), <https://doi.org/10.3390/plants11151952>.
- [25] T.-L. Le, D.-T. Tran, V.-N. Hoang, Fully automatic leaf-based plant identification, application for Vietnamese medicinal plant search, in: *Proceedings of the 5th Symposium on Information and Communication Technology, Association for Computing Machinery, New York, NY, USA, 2014*, pp. 146–154, <https://doi.org/10.1145/2676585.2676592>.
- [26] A.M. Raghukumar, G. Narayanan, Comparison of machine learning algorithms for detection of medicinal plants, in: *Proceedings of the 4th International Conference on Computing Methodologies and Communication, vol. 2020, 2020*, pp. 56–60, <https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00010>.
- [27] L.D.S. Pacifico, L.F.S. Britto, E.G. Oliveira, T.B. Ludermir, Automatic classification of medicinal plant species based on color and texture features, in: *2019 8th Brazilian Conference on Intelligent Systems, Institute of Electrical and Electronics Engineers, 2019*, pp. 741–746, <https://doi.org/10.1109/BRACIS.2019.00133>.
- [28] N. Duong-Trung, L. Da Quach, M.H. Nguyen, C.N. Nguyen, A Combination of Transfer Learning and Deep Learning for Medicinal Plant Classification, *Association for Computing Machinery International Conference Proceeding Series, 2019*, pp. 83–90, <https://doi.org/10.1145/3321454.3321464>. Part F1479.
- [29] B.R. Pushpa, N.S. Rani, Ayur-PlantNet: an unbiased light weight deep convolutional neural network for Indian Ayurvedic plant species classification, *Journal of Applied Research on Medicinal and Aromatic Plants* (2023), 100459, <https://doi.org/10.1016/j.jarmap.2023.100459>.
- [30] A. Muneer, S.M. Fati, Efficient and automated herbs classification approach based on shape and texture features using deep learning, *IEEE Access* 8 (2020) 196747–196764, <https://doi.org/10.1109/ACCESS.2020.3034033>.
- [31] K. Pushpanathan, M. Hanafi, S. Masohor, W.F.F. Ilahi, MYLPHerb-1: a dataset of Malaysian local perennial herbs for the study of plant images classification under uncontrolled environment, *Pertanika Journal of Science and Technology* 30 (2022) 413–431, <https://doi.org/10.47836/PJST.30.1.23>.
- [32] H.X. Kan, L. Jin, F.L. Zhou, Classification of medicinal plant leaf image based on multi-feature extraction, *Pattern Recogn. Image Anal.* 27 (2017) 581–587, <https://doi.org/10.1134/S105466181703018X>.
- [33] B. VijayaLakshmi, V. Mohan, Kernel-based PSO and FRVM: an automatic plant leaf type detection using texture, shape, and color features, *Comput. Electron. Agric.* 125 (2016) 99–112, <https://doi.org/10.1016/j.compag.2016.04.033>.
- [34] N.S. Russel, A. Selvaraj, Leaf species and disease classification using multiscale parallel deep CNN architecture, *Neural Comput. Appl.* 34 (2022) 19217–19237, <https://doi.org/10.1007/s00521-022-07521-w>.
- [35] S. Sachar, A. Kumar, Deep ensemble learning for automatic medicinal leaf identification, *Int. J. Inf. Technol.* 14 (2022) 3089–3097, <https://doi.org/10.1007/s41870-022-01055-z>.
- [36] M.R. Dileep, P.N. Pournami, AyurLeaf: a deep learning approach for classification of medicinal plants, *IEEE Region 10 Annual International Conference, Proceedings/TENCON. 2019-October (2019)* 321–325, <https://doi.org/10.1109/TENCON.2019.8929394>.
- [37] R. Azadnia, M.M. Al-Amidi, H. Mohammadi, M.A. Cifci, A. Daryab, E. Cavallo, An AI based approach for medicinal plant identification using deep CNN based on global average pooling, *Agronomy* 12 (2022), <https://doi.org/10.3390/agronomy12112723>.
- [38] R. Azadnia, K. Kheiralipour, Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier, *Journal of Applied Research on Medicinal and Aromatic Plants* 25 (2021), 100327, <https://doi.org/10.1016/j.jarmap.2021.100327>.
- [39] Kaggle. <https://www.kaggle.com/>, 2022. (Accessed 25 December 2022).
- [40] T. Das, Y. Zhong, I. Stoica, S. Shenker, Adaptive stream processing using dynamic batch sizing, in: *Proceedings of the Association for Computing Machinery Symposium on Cloud Computing, Association for Computing Machinery, New York, NY, USA, 2014*, pp. 1–13, <https://doi.org/10.1145/2670979.2670995>.
- [41] A.K. Jain, P.P. Rao, K. Venkatesh Sharma, Optimization techniques in deep learning scenarios: an empirical comparison, in: *Advanced Analytics and Deep Learning Models, 2022*, pp. 255–282, <https://doi.org/10.1002/9781119792437.ch11>.
- [42] B. Dey, K.A.M. Abir, R. Ahmed, M.A. Salam, M. Redowan, M.D. Miah, M.A. Iqbal, Monitoring groundwater potential dynamics of north-eastern Bengal Basin in Bangladesh using AHP-Machine learning approaches, *Ecol. Indicat.* 154 (2023), 110886, <https://doi.org/10.1016/j.ecolind.2023.110886>.
- [43] Y. Zhang, Y. Wang, Recent trends of machine learning applied to multi-source data of medicinal plants, *Journal of Pharmaceutical Analysis* (2023), <https://doi.org/10.1016/j.jpha.2023.07.012>.
- [44] S. Mihai, M. Yaqoob, D.V. Hung, W. Davis, P. Towakel, M. Raza, M. Karamanoglu, B. Barn, D. Shetve, R.V. Prasad, H. Venkataraman, R. Trestian, H.X. Nguyen, Digital twins: a survey on enabling technologies, challenges, trends and future prospects, *IEEE Communications Surveys and Tutorials* 24 (2022) 2255–2291, <https://doi.org/10.1109/COMST.2022.3208773>.
- [45] A. Bahri, Y. Bourass, I. Badi, H. Zouaki, K. El moutaouakil, K. Satori, Dynamic CNN combination for Morocco aromatic and medicinal plant classification, *International Journal of Computing and Digital Systems* 11 (2022) 239–249, <https://doi.org/10.12785/IJCDs/110120>.
- [46] R. Akter, M.I. Hosen, CNN-Based leaf image classification for Bangladeshi medicinal plant recognition, in: *ETCCE 2020 - International Conference on Emerging Technology in Computing, Communication and Electronics, 2020*, <https://doi.org/10.1109/ETCCE51779.2020.9350900>.
- [47] G.E. Austen, M. Bindemann, R.A. Griffiths, D.L. Roberts, Species identification by experts and non-experts: comparing images from field guides, *Sci. Rep.* 6 (2016) 1–7, <https://doi.org/10.1038/srep33634>.

- [48] A. Solow, W. Smith, M. Burgman, T. Rout, B. Wintle, D. Roberts, Uncertain sightings and the extinction of the ivory-billed woodpecker, *Conserv. Biol.* 26 (2012) 180–184, <https://doi.org/10.1111/J.1523-1739.2011.01743.X>.
- [49] M. Phumthum, K. Srithi, A. Inta, A. Junsongduang, K. Tangjitman, W. Pongamornkul, C. Trisonthi, H. Balslev, Ethnomedicinal plant diversity in Thailand, *J. Ethnopharmacol.* 214 (2018) 90–98, <https://doi.org/10.1016/J.JEP.2017.12.003>.
- [50] J. Pärtel, M. Pärtel, J. Wäldchen, Plant image identification application demonstrates high accuracy in Northern Europe, *Annals of Botany (AoB) Plants* 13 (2021), <https://doi.org/10.1093/AOBPLA/PLAB050>.
- [51] J.C.D. Terry, H.E. Roy, T.A. August, Thinking like a naturalist: enhancing computer vision of citizen science images by harnessing contextual data, *Methods Ecol. Evol.* 11 (2020) 303–315, <https://doi.org/10.1111/2041-210X.13335>.