Received: 13 September 2025 | Revised: 8 December 2025 | Accepted: 5 March 2026 | Published online: 20 May 2026

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2026, Vol. 00(00) 1–17
DOI: 10.47852/bonviewJCCE62027652

BON VIEW PUBLISHING

Deep Learning for Medicinal Plant Classification: A Comprehensive Review of Recent Advances and Challenges

Monir Hossain[1, 2], **Fahmid Al Farid**[3], **Momotaz Begum**[2,*], **Jia Uddin**[4] **and Hezerul Bin Abdul Karim**[3,*]

[1] Department of Computer Science and Engineering, International Islamic University Chittagong, Bangladesh

[2] Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Bangladesh

[3] Centre for Image and Vision Computing, Multimedia University, Malaysia

[4] AI and Big Data Department, Woosong University, Republic of Korea

Abstract: Deep learning classifies medicinal plants, driven by the need to preserve traditional knowledge and automate identification for practical uses. This review extensively summarizes 30 recent studies (2021–June 2025) on applying deep learning, primarily using image data, to classify medicinal plants. This review analyzes research distribution, dataset preparation, image preprocessing, augmentation, and deep learning architectures like convolutional neural networks, Vision Transformers, and hybrid models. Our analysis reveals a strong geographic focus, with 50% of the selected studies originating from India and Bangladesh. The focus is overwhelmingly on leaf imagery, with 29 out of the 30 studies relying on this approach. The field is also characterized by its dependence on existing data, as 56.6% of studies utilized public datasets and another 26.6% employed a hybrid of public and private data, with dataset sizes ranging from a minimum of 637 to a maximum of 13,500 images. Methodologically, the vast majority of studies rely on a transfer learning approach (36.7%), achieving robust accuracy rates between 74% and 99.9%. Furthermore, we recognize significant limitations, such as the absence of standardized and diverse datasets, insufficient inclusion of uncommon or endangered species, and inadequate representation of whole-plant imaging. The research underscores the necessity for collaborative, multidisciplinary initiatives to develop centralized, high-quality, and geographically comprehensive datasets. We delineate prospective avenues, including multimodal feature integration, the development of real-world applications, and optimization for privacy-preserving frameworks such as federated learning. This study guides academics advancing deep learning for medicinal plant classification and biodiversity conservation.

Keywords: medicinal plant classification, plant image dataset, deep learning, feature extraction, image preprocessing

1. Introduction

Medicine plays an important role in preventing, diagnosing, and treating diseases, improving health, relieving pain, and extending life expectancy, using natural remedies, medical knowledge, and modern or traditional methods. Plants were the primary source of healing prior to the development of modern medicine. The ancient Egyptians used honey and garlic to treat wounds. Herbs like neem and turmeric were used in Indian Ayurveda to promote healing and balance. Ginseng, ginger, and licorice were used in traditional Chinese medicine to promote health and vitality. Native Americans made use of willow bark and echinacea. Mint and chamomile are among the numerous herbal remedies that the Greeks and Romans recorded. Through new discoveries and translations, Islamic scholars increased our understanding of herbs. Long before the invention of synthetic drugs, these traditions believed that plants were effective means of curing disease [1].

Modern medicine has achieved remarkable progress, but it also faces several significant problems. Medicinal plants can help solve a number of issues facing modern medicine. Drug side effects are a major problem. When used appropriately, plant-based remedies frequently have fewer negative effects than synthetic drugs, which frequently cause harmful reactions [2]. For example, compared to non-steroidal anti-inflammatory drugs, ginger and turmeric have anti-inflammatory properties while causing less stomach discomfort. Antibiotic resistance, which is caused by the overuse of synthetic antibiotics, is another significant problem. With their broad-spectrum antimicrobial qualities, medicinal plants such as tea tree, neem, and garlic may be useful against resistant strains without encouraging resistance [3]. Medicinal

***Corresponding authors:** Momotaz Begum, Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Bangladesh. Email: drmomotaz@duet.ac.bd and Hezerul Bin Abdul Karim, Centre for Image and Vision Computing, Multimedia University, Malaysia. Email: hezerul@mmu.edu.my

plants can also be used to treat chronic conditions like diabetes and hypertension. Synthetic drug use over a lifetime frequently results in cumulative side effects. Some studies have shown that herbs like hibiscus, cinnamon, and bitter melon can lower blood pressure and blood sugar [4]. Due to the high expense and restricted availability of modern medicine, people in many low-income areas, such as parts of Africa, South Asia, and Latin America, rely on medicinal plants [3, 5]. For primary healthcare in underserved and rural areas, traditional methods are still essential [6]. In these situations, locally accessible medicinal plants provide an affordable, socially acceptable, and efficient substitute [2].

For natural medicine to be used safely and effectively, medicinal plant classification is crucial. It guarantees precise plant identification and use, to start. While many plants have similar appearances, only a few are safe or beneficial. As biodiversity decreases [7], classifying medicinal plants is like creating a detailed inventory for a burning library. It helps us know exactly what plants we have, which are most at risk, and which are most useful. This allows us to focus our conservation efforts on saving the most important species before they go extinct. It also helps scientists find new medicines by studying related plants, ensures that local communities can correctly identify and safely use traditional remedies, and guides the sustainable harvesting of these valuable resources so they can regrow and be available for future generations. According to Saggar et al. [8], proper classification aids in preventing the use of hazardous or incorrect plants. Medicinal plant classification promotes the advancement of medicine and research. Research is easier to disseminate and replicate when scientists name and group plants according to a standard system. This guarantees that results are based on the same plant species globally and aids in the discovery of new medications derived from plants [9]. Classification also aids in plant conservation and sustainable use. We can save rare or endangered plants if we know the exact species. It also enables people to cultivate and gather plants in environmentally friendly ways. Many communities still rely on these plants for medical care, so this is significant [10].

Deep learning is a subset of artificial intelligence (AI) that helps accurately identify medicinal plants from images by learning features like shape and color automatically [11, 12]. Deep learning enables computers to recognize patterns in data, much like people do. Its ability to accurately identify plant parts (such as leaves, flowers, or seeds) has led to its widespread use in the classification of medicinal plants. One explanation is the high level of image analysis capabilities of deep learning models, particularly convolutional neural networks (CNNs). By learning characteristics like shape, color, and texture, these models are able to distinguish between plants that look alike [13]. Human error is decreased by deep learning. Even professionals can make mistakes, and many medicinal plants have similar appearances. According to Archana and Jeevaraj [14], deep learning models are capable of processing thousands of images rapidly and producing accurate predictions. As more data are collected, these models continue to get better. The model learns more effectively as we add more plant photos or data. This aids in the development of mobile applications or tools that enable real-time plant identification for farmers, students, or researchers [12]. Deep learning aids in accelerating research and preserving plant biodiversity, particularly in isolated locations where specialists might not be accessible.

There are many machine learning algorithms that have been previously studied for the recognition of medicinal plants. But in this paper, our aim is to focus only in the effective and reliable deep learning algorithm and technique such as CNN [15–17], Fast RCNN [18], Faster RCNN [19], VGGNet, ResNet (Residual Networks) [20], Inception (GoogLeNet) [21], MobileNet [22], DenseNet (Densely Connected Networks) [23], EfficientNet [24], Alexnet, UNet [25], Mask-RCNN [26], Xception [27], transfer learning [28], ensemble learning, and transformer-based architectures [29] like Vision Transformer (ViT).

The key contributions of this paper are as follows:

1) This paper gives a clear and thorough explanation of the methodical way that studies on classifying medicinal plants using deep learning are chosen.
2) This paper includes the number of research publications in this field over time and across different regions.
3) This paper carefully looks into the datasets and the most important features of the datasets that are needed to train deep learning models to classify and recognize medicinal plant species.
4) This paper shows how different parts of medicinal plants (such as leaves, flowers, and seeds) are used and the methods used to extract features from.
5) This paper gives a full review of the newest deep learning methods used for classifying medicinal plants.
6) This paper analyzes the learnable parameter needs of different deep learning architectures to see how easy they are to compute and install.
7) This paper lists current problems and suggests interesting areas for future research and real-world use.

The remainder of this paper is organized as follows: Methodology sections describing how relevant papers and studies were collected for this review. In the Results section, we present the results, summarizing key findings regarding publication demographics, dataset preparation, preprocessing, augmentation strategies, and the methods and tasks employed in the selected studies. The Discussion section delivers a detailed discussion, highlighting research gaps and proposing prospective solutions. Finally, in the Conclusion section, we draw conclusions and provide directions for future work.

2. Methodology

2.1. Paper selection criteria

We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. To concentrate on the latest research trends in the classification of medicinal plants utilizing deep learning techniques, we restricted our search from January 2021 to June 2025. We assessed records for inclusion in our evaluation by determining whether they met our established inclusion and exclusion criteria based on title and abstract.

The following inclusion criteria were adhered to:

1) Only investigations that were published in peer-reviewed academic journals and the proceedings of prestigious conferences were considered. This ensured that the included research had undergone expert evaluation for methodological soundness and validity.
2) The review exclusively included studies whose focus was the classification of medicinal plants using deep learning methodologies like CNNs to identify and categorize plant species from images.
3) The included investigations were required to have been published between January 2021 and June 2025. This time frame was selected to capture the most recent and relevant advancements in the field.

4) All selected documents were documented in English. This criterion ensured the research was accessible to a wide international audience and prevented potential misinterpretation from translation.

The following exclusion criteria were adhered to:

1) Studies were excluded if they concentrated on the categorization of medicinal plants but did not employ deep learning methodologies.
2) Review studies, abstracts, commentaries, book chapters, protocols, brief papers, and editorials were not considered for inclusion.
3) Any articles identified as duplicates were removed from the selection.

2.2. Paper database source and searched keyword

To guarantee the selection of pertinent and high-caliber publications, five distinguished databases and libraries were referenced: ScienceDirect, Google Scholar, PubMed, IEEE Explore, and Springer Link. The investigation concentrated on titles, keywords, and abstracts to discern relevant material within the field. It was exceedingly challenging to identify the basic query that encompasses all results pertinent to our study. We noted that the predominant terms referenced in the majority of publications are medicinal plants and classification. Additionally, as we aimed to identify publications specifically linked to deep learning, we incorporated this keyword as well. We utilize these keywords and their related synonyms with AND and OR to create our query. Following numerous trial-and-error rounds, the final search query is as follows: ["convolutional neural network" OR "deep learning" OR "CNN"] AND ["medicinal plant" OR "ayurvedic plant"] AND ["identification" OR "classification"] were utilized to acquire published publications.

2.3. Paper selection and collection process

Upon examining all five databases, we gathered the following data for each entry: Item Title, Publication Title, DOI, Authors, Publication Year, URL (uniform resource locator), and Content Type, which were subsequently stored as individual CSV (Comma Separated Values) files for each database and have been provided. The files were subsequently uploaded into Google Sheets and consolidated for additional processing. We implemented our inclusion and exclusion criteria to filter the data, eliminated duplicates, and meticulously examined the titles to ensure that only the most pertinent publications were included in our final list. Figure 1 presents the comprehensive PRISMA 2020 flow diagram, delineating our rigorous research selection procedure according to established eligibility criteria. Our extensive database search initially yielded 737 articles that satisfied the inclusion criteria: 120 from Google Scholar, 253 from ScienceDirect, 47 from PubMed, 37 from IEEE Xplore, and 280 from Springer Link. A total of 372 records were gathered during the refined search phase: 94 from Google Scholar, 150 from ScienceDirect, 29 from PubMed, 19 from IEEE Xplore, and 80 from Springer Link. Upon examining the title, 126 out of 372 were elected. Subsequently, the 126 records were examined, resulting in the elimination of 5 duplicate articles. After abstract and full text screening, 121 papers were chosen for retrieval, whereas 74 reports were inaccessible and were found that did not focus solely on this topic. A quality assessment was performed on 47 remaining articles, and we picked 30 high-quality articles, ensuring that all sorts of models are represented in the list. In conclusion, the final review included 30 studies. The

Figure 1
Research screening the PRISMA framework

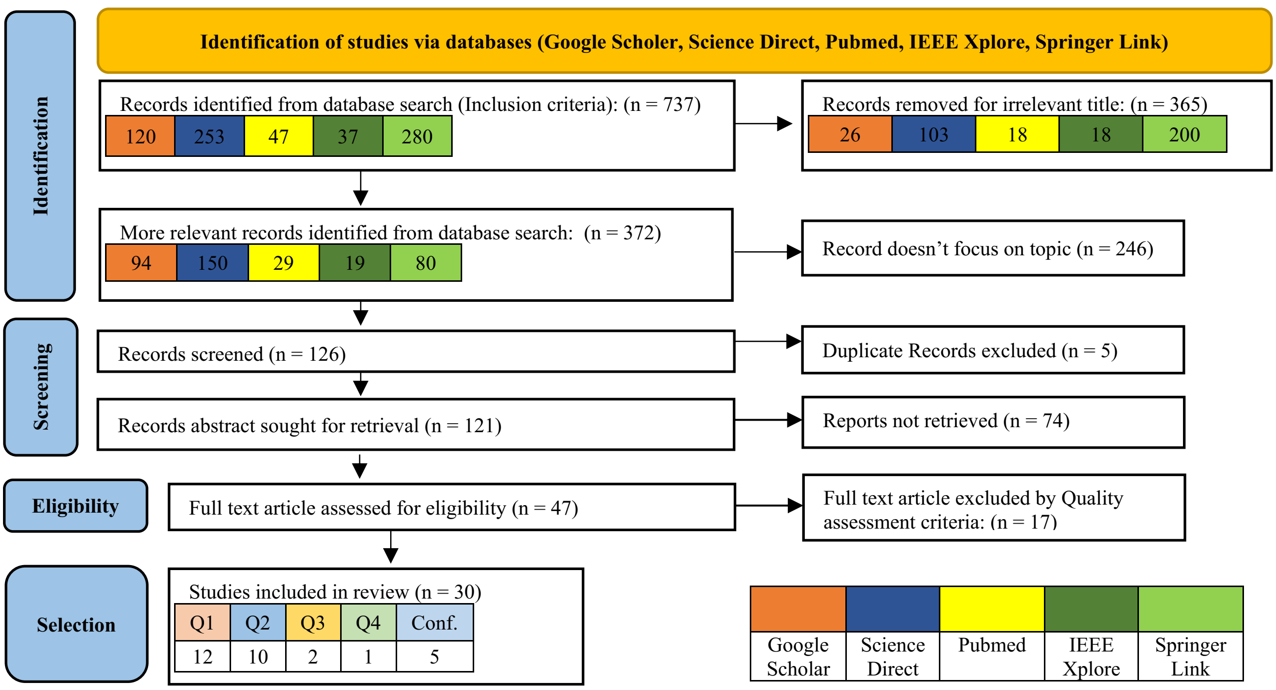
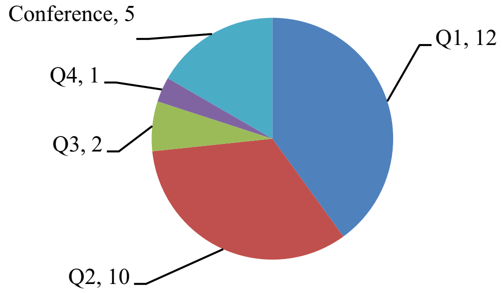


Figure 2
Selected publication distribution by quality measure



compilation comprised 12 Q1 journal articles, 10 Q2 journal articles, 2 Q3 journal articles, 1 Q4 journal article, and 5 conference papers, shown in Figure 2.

2.4. Data collection process

Our principal aim was to collect data across three core categories such as Table 1, publication demography, dataset details, and deep learning tasks and methods. We established a table to extract and document data from the chosen publications. During the initial phase, data were gathered from all selected articles. In the second and third phases, we conducted cross-validation on each entry to verify the collected data in the first phase. All differences were candidly addressed and handled cooperatively. In the context of publishing demography, we concentrated on the publishing year and the nationality of the primary author of each selected study. The dataset details section includes data on species count (number of classes), images per species (number of images per class), dataset source (public, private, or mixed), and employed preprocessing techniques. We recorded the feature extraction techniques, implemented deep learning models, and architectural specifications under deep learning tasks and methods.

Table 1
Data collection review template

Publication demography	Dataset details	Deep learning task and methods
Publication year, Country of the first author	No of species, no of images in each species, dataset source (private, public, or mixed), data preprocessing technique	Feature extraction technique, deep learning models, and architecture details

2.5. Data analysis

After collecting the data, we analyzed it across three main categories: publication demographics, dataset details, and deep learning tasks and methods. In the Publication demography section, we aimed to identify the geographic and disciplinary origins of the research, determining which regions and academic fields were contributing the most. Within the dataset details category, our focus was on understanding the characteristics of the datasets used. This included the number of classes, data size, quality, data augmentation techniques, and whether the datasets were publicly available or privately held. Finally, for deep learning tasks and methods, we focused on how researchers approached feature extraction, the most prevalent model types, and the number of parameters these models utilized. Based on this analysis, we then identified key research gaps that warrant further investigation.

2.6. Quality assessment

This literature study evaluated the validity threat in primary research by examining eligibility criteria, publication names, abstracts, and keywords. The search encompassed numerous studies that fulfilled the criterion, although the ultimate conclusions remained impartial. Data extraction was objective; however, certain research contained unreported or misconstrued data. The validity threat was ascribed to bias in research selection and data extraction. To alleviate this, both manual and automated searches were conducted across several databases, and a data extraction template was created to fulfill study objectives.

3. Results

3.1. Publication demography

To comprehend the trends in the classification of medicinal plants, we performed an analysis of publication demographics. Figure 3 demonstrates a significant surge in academic interest, with over 75 papers published in just the first half of 2025, compared to roughly 27 in the entire year of 2021. This upward trend, illustrated in our final review list of 30 papers (Figure 4), strongly indicates the growing significance of this research area. Our analysis of first-author affiliations, summarized in Figure 5 and Table 2, reveals that the selected publications originated from scholars in 13 different nations. India was the dominant contributor, accounting for 40% (n = 12) of the publications, followed by Bangladesh at 10% (n = 3). Ethiopia, Vietnam, Iran, and Indonesia each contributed 6.67% (n = 2) of the papers. Furthermore, we found that the vast majority of studies were led by researchers from deep learning-related fields, such as computer science, information technology, and computer engineering.

Figure 3
Distribution of publication per year

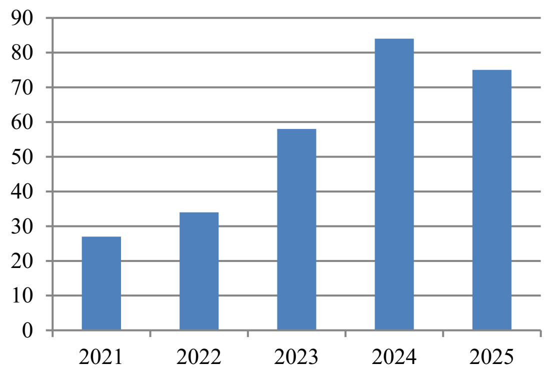


Figure 4
Distribution of selected 30 publications per year

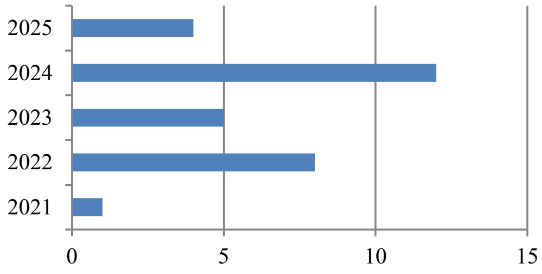
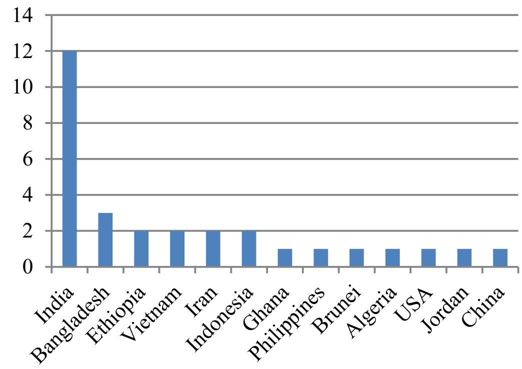


Figure 5
Distribution of selected 30 studies across different countries



3.2. Dataset preparation

This review examines datasets for medicinal plant species, with an emphasis on deep learning methods for plant classification and recognition. The discussion encompasses data acquisition methods, sources of origin, image volume, and image details. The review identifies challenges in dataset preparation, noting that primary studies frequently utilize public datasets. The review's findings indicate that 56.6% (n = 17) of primary studies

Table 2
Geographic region of origin

Refs.	Publication year	Countries of the author	Background
[30]	2021	India	Computer Science
[31]		India	Computer Science
[32]		Brunei	Computer Science
[33]		Indonesia	Computer Science
[34]	2022	Iran	Biosystems
[35]		Ghana	Computer Science
[36]		Iran	Computer Science
[37]		Philippines	Computer Science
[38]		Jordan	Computer Science
[39]		Bangladesh	Computer Science
[40]		India	Computer Science
[41]	2023	India	Computer Science
[42]		India	Computer Science
[43]		India	Computer Science
[44]		Vietnam	Computer Science
[45]		USA	Computer Science
[46]		China	Medicine
[47]		Ethiopia	Computer Science
[48]		India	Computer Science
[49]	2024	Indonesia	Computer Science
[13]		Bangladesh	Forestry and Environmental Science
[50]		Bangladesh	Computer Science
[51]		Ethiopia	Computer Science
[52]		India	Computer Science
[53]		India	Computer Science
[54]		Vietnam	Computer Science
[55]		India	Computer Science
[56]	2025	India	Computer Science
[57]		Algeria	Biochemistry
[58]		India	Computer Science

utilized public datasets. Furthermore, 26.6% (n = 8) of the studies utilized private datasets specifically designed for comparable challenges. A total of 16.6% (n = 5) utilized their private datasets in conjunction with public datasets. The private dataset contains a minimum of 1270 images and a maximum of 13,500 images. The public dataset contains a minimum of 637 images and a maximum of 38,066 images. The minimum average number of images per class in the private dataset is 19.78, while the maximum average is 2700. The minimum average number of images per class in the public dataset is 19.9, while the maximum average is 3806. The total average number of images per class in private datasets is 260.8, while in public datasets, it is 155.45. All datasets contain more than 1000 images, except for one public dataset, which contains only 637 images. Over 50% of the dataset contains 5000 or more images. Approximately 30% contains fewer than 2000 images. All datasets utilized leaf images of medicinal plants, with the exception of one that features fruit images. In addition to leaf images, a limited number of studies also include images of whole plants. The datasets contain various types of backgrounds. Eight participants utilized a white background. To achieve a white background, some individuals utilized a white backdrop during photography, while others employed preprocessing techniques. Several studies employed natural backgrounds in both controlled and uncontrolled environments. Some utilized a mixed environment. All the details about the dataset are listed in Table 3.

3.3. Preprocessing and augmentation technique

Augmentation involves generating new instances through minor modifications of existing examples. This enables systems to learn more efficiently without the necessity of gathering extensive new data. For instance, when a computer is trained to recognize apples, exposure to the same apple from various angles, sizes, or lighting conditions enhances its understanding. Augmentation enhances learning by increasing variety and better equipping the system for real-world scenarios. Table 4 presents the various augmentation techniques employed in the selected studies. The majority of the studies (n = 22), representing 73.33%, employed the standard augmentation technique. Augmentation techniques

Table 3
Dataset details

Refs.	Type	Dataset	No of images	No of species	Image details	Image background condition
[47]	Private	Private	2200	44	Leaf images	White background
[50]			6,427	7	Plant/leaf images	Natural background
[56]			4000	13	Leaf images	Complex backgrounds
[51]			1853	35	Leaf images	Standardized background
[34]			13,500	5	RGB images of leaves	White background
[36]			3000	30	Leaf images	White background
[37]			7405	10	Leaf images	Natural and complex backgrounds
[43]			1270	8	Leaf images (single object per image)	White background
[46]	Private and Public	Private	5118	14	RGB images of medicinal fruits and derived products	Controlled environment with reflective gray coating and uniform lighting
		Public (Chinese medicinal blossom	12538	12		
		and medicinal leaf dataset)	1500	30		
[30]		Private: DeepHerb	2515	40	Leaf images	White background through segmentation
		Public: Flavia	1907	32		
[35]		Private: MyDataset	2450	49	Leaf images	Closed environment with constant illumination
		Public: (Flavia, Swedish Leaf, MD2020, Folio)	–	–		
[32]		Private: UBD Botanical Garden Dataset	2097	106	Leaf and whole plant images	Mixed backgrounds (real-world conditions)
		Public: PlantCLEF 2015	23708	1000		
[49]		Private: Indonesia medicinal plants	10000	100	Images of whole plants, leaves, and other parts from Google Images	Natural environments
		Public: Vietnam Medicinal Plant	20000	200		

(*Continued*)

Table 3
(*Continued*)

Refs.	Type	Dataset	No of images	No of species	Image details	Image background condition
[39]	Public	Bangladeshi medicinal plants	5000	10	RGB leaf images	Controlled conditions with background removal
[45]		BDMediLeaves	38066	10	Leaf images	Controlled (augmented)
[42]		DIMPSAR dataset	6000	40	Whole plant images	Uncontrolled environments
[53]		Folio and Mendeley medicinal leaf dataset	637 1835	32 30	Leaf images	White and natural backgrounds
[52]		Indian Medicinal Leaves and Indian Plants Dataset	6900 5945	80 40	Plant/leaf Images	Natural background
[48]		Indian Medicinal Plants Database	2000	40	Leaf images	Segmented background
[31]		Medicinal leaf dataset	1841	30	RGB leaf images	Natural backgrounds
[40]		Medicinal leaf dataset	1835	30	Leaf images	Natural background
[33]		Medicinal leaf dataset	1500	30	Leaf images	Segmented (background removed)
[38]		Medicinal leaf dataset	1800	30	Leaf images	White background
[55]		Medicinal leaf dataset and medicinal plant dataset from Mendeley	19745	120	Plant/leaf images	Mixed (natural + controlled)
[57]		Medicinal plant	1000	30	Leaf images	Natural background
[58]		Medicinal plant datasets	1907 6000	33 40	RGB images of leaves and plants	Mixed backgrounds
[13]		Public from Kaggle (not mentioned)	5878	30	Leaf images	Plain + complex background
[41]		Swedish and Flavia	1932 3090	15 32	Leaf images	White background
[44]		VNPlant-200	20000	200	RGB images of whole plants	Natural background
[54]		VNPlant-200	20000	200	Leaf images	Natural background

Table 4
Augmentation, preprocessing, studied organs, and feature extraction techniques details

Refs.	Augmentation techniques	Preprocessing techniques	Studied organ	Feature extraction techniques
[44]	RandAugment (rotation, zoom, flip, brightness, shear)	Resizing (224 × 224, 384 × 384), normalization	Whole plant	MBConv (EfficientNet), ViT, BEiT
[45]	Rotation (60°), zoom (10%), flip, brightness (0.2–0.3), shear (15%)	Hybrid denoising (Wavelet + Gaussian blur), resizing (512 × 512)	Leaves	Custom CNN, ConvMixer, CCT (ViT hybrid)
[55]	CLAHE for contrast enhancement	CLAHE (LAB color space), resizing (224 × 224)	Leaves + whole plants	MBConv + Residual Spatial Attention (RSA)/Residual Channel Attention (RCA) (Attention)
[46]	Random cropping, shadowing, rotation, flip	Random local enhancement, unsharp masking, CLAHE, morphological gradient	Fruits and derived products	Hybrid MAE (ViT + MBConv)

(*Continued*)

Table 4
(*Continued*)

Refs.	Augmentation techniques	Preprocessing techniques	Studied organ	Feature extraction techniques
[31]	Not specified	Resizing, standardization	Leaves	MobileNetV2, InceptionV3, ResNet50
[39]	Not specified	Background removal, unsharp masking, CLAHE, morphological gradient	Leaves	VGG16, ResNet50, DenseNet201, InceptionV3, Xception
[40]	Random rotation (20°), zoom (5%), shift (5%), shear (5°), flip	Resizing (224 × 224), normalization, cropping	Leaves	VGG16, VGG19, DenseNet201
[47]	Rotation (30°), shift (20%), shear (20%), zoom (20%), flip	Resizing (224 × 224, 260 × 260, 380 × 380), normalization, cropping	Leaves	EfficientNetB0-B4
[48]	Mirroring, twisting, shearing, cutting, color adjustment	Resizing (64–256px), background replacement, segmentation (GrabCut), noise reduction	Leaves	Modified ResNet50 (VCP, VAP, VMP)
[49]	Flips, rotations (balanced classes)	Resizing (128 × 128), normalization (0–1)	Leaves, flowers, whole plants	Transfer learning (ResNet34, DenseNet121, VGG11)
[13]	Rotation (±32°), flip, shift (20%), shear, zoom (0.2–20%), intensity scaling	Resizing (224 × 224), color normalization	Leaves	VGG16, VGG19, DenseNet201, ResNet50V2, Xception
[50]	Likely standard augmentations	Resizing (224 × 224), color normalization	Leaves and whole plants	ResNet50 + PSO
[56]	Rotation (90°, 45°, –45°), flip, brightness/hue changes	Gaussian/median blur, scaling, translation	Leaves	Hybrid models (VGG16 + MobileNet, ResNet50 + MobileNet)
[51]	Rotation (0–90°), shift (20%), shear (20%), zoom (20%), flip	Resizing (224 × 224), normalization, cropping	Leaves	Pretrained CNNs (VGG16, VGG19, InceptionV3)
[32]	AutoAugment (Reinforcement Learning (RL)-based)	Resizing (224 × 224), class balancing	Leaves and whole plants	EfficientNet-B1 (transfer learning)
[33]	Flip, shift, rotation, zoom, brightness, shear	Normalization (×1/255), augmentation	Leaves	VGG16, VGG19, MobileNetV2
[57]	Flip, rotate 90°, synthetic data	Feature fusion, optimization	Leaves	Residual blocks, BCO, feature fusion
[30]	Flip, rotation, edge noise	Segmentation (Gaussian blur, edge detection), resizing (1600 × 1200)	Leaves	Xception, VGG16, VGG19, InceptionV3
[41]	Resizing, rotation, scaling, flip, histogram equalization	Background removal, grayscale conversion	Leaves	CNN, VGG-16, ESPCA
[52]	Assumed standard augmentations	Vein morphometrics, Sobel edge detection, resizing (224 × 224)	Leaves and whole plants	CNN-based feature fusion
[34]	Rotation (45°–180°), color manipulation	Background removal, resizing (64 × 64 – 256 × 256), dilation	Leaves	CNN blocks, GAP layer
[58]	Not specified	Contour detection, edge segmentation, color histograms, GLCM, LBP	Leaves and whole plants	Multi-level CNN, self-attention, BERT
[35]	Not specified	Log-Gabor filters for texture	Leaves	Log-Gabor + DenseNet201
[42]	Rotation (180°), flip, contrast adjustments	Segmentation (color thresholding, morphology)	Leaves, stems, flowers	CNN-based self-learned features
[36]	Not specified	Resizing (256 × 256), mean subtraction	Leaves	MobileNetV2

(*Continued*)

Table 4
(*Continued*)

Refs.	Augmentation techniques	Preprocessing techniques	Studied organ	Feature extraction techniques
[37]	Random crop, flip, rotation	Resizing (224 × 224), normalization	Leaves	MobileNetV3
[53]	Resizing, scaling, flip	Median filtering, noise removal	Leaves	Hybrid CNN-RNN (channel-wise feature extraction (CWFE))
[38]	Resizing (800 × 1024), flip, rotation	Mask annotation, noise filtering	Leaves	Mask RCNN (ResNet101)
[54]	Not specified	Resizing, normalization	Leaves	VGG16, ResNet50, ConvNext, MaxVit
[43]	Not specified	Grayscale, binarization, edge segmentation (Canny, Sobel)	Herbal leaves	Shape/texture/color features; RCNN variants

encompass geometric transformations such as rotation, shifting, zooming, shearing, cropping, shadowing, mirroring, twisting, and scaling, as well as various color adjustment methods including intensity, brightness, contrast, and hue. One study found that CLAHE (Contrast Limited Adaptive Histogram Equalization) is also utilized in augmentation. In addition, 23.33% of studies (n = 7) did not specify the use of augmentation techniques. Preprocessing refers to the preparation of raw data to enhance its comprehensibility and usability by a computer. Similar to the process of cleaning and organizing ingredients prior to cooking, in technology, data are cleaned, formatted, or adjusted—such as eliminating errors, resizing images, or converting text—to prepare it for analysis or training. This step enhances the accuracy and efficiency of systems utilizing the data. Resizing emerged as the predominant preprocessing technique across the studies, utilized in 24 instances, with 224 × 224 identified as the most commonly employed dimension. Normalization was present in 11 studies, frequently alongside resizing or cropping techniques. Cropping was employed in three studies, whereas color normalization was explicitly referenced in two. CLAHE and unsharp masking were utilized in three studies, frequently in conjunction. Background removal was addressed in four studies, while segmentation techniques, including GrabCut, morphology, and edge-based methods, were utilized in five studies. Five studies utilized noise reduction or denoising methods, such as Gaussian or median filtering. Edge detection methods, including Sobel, Canny, and morphological gradients, were utilized in four studies, while grayscale conversion was employed in two studies. Additional techniques included class balancing, data augmentation, feature fusion and optimization, mask annotation, log-Gabor filtering, contour detection, color histograms, GLCM (Gray Level Co-occurrence Matrix) and LBP (Local Binary Patterns) texture features, and mean subtraction. Several studies employed hybrid approaches that integrate multiple techniques. Most studies employed a combination of resizing, normalization, denoising, and segmentation, highlighting their significance in standardizing and improving image quality for subsequent analysis or model training.

3.4. Studied organs and feature extraction technique

Feature extraction involves identifying the most significant components of data that facilitate pattern recognition by a computer. It resembles emphasizing essential concepts in a text rather than engaging with the entire material. In images, this may refer to shapes or edges; in text, it may pertain to significant words. This step enhances the data's simplicity and utility for tasks such as recognition or prediction. Of the 30 studies, 24 concentrated exclusively on leaves as the target organ. Four studies examined both leaves and whole plants, whereas one study focused solely on the whole plant. One study examined fruits and their derived products, while another focused on leaves, stems, and flowers. Leaves emerged as the predominant organ of study, underscoring their diagnostic importance in plant research. The majority, approximately 23, employed deep learning models such as CNN variants (e.g., VGG16, ResNet50, DenseNet201) or their hybrids. VGG16 was referenced in 10 studies, whereas ResNet50 and its variants were included in 9. Five studies utilized MobileNetV2 and V3. ViTs and their variants, including BEiT (Bidirectional Encoder representation from Image Transformers), CCT (Compact Convolutional Transformer), and MAE (Masked Autoencoders), were utilized in five studies, reflecting an increasing interest in this area. Eight studies employed hybrid techniques that integrate CNNs with attention mechanisms, transformers, or handcrafted features. Only one study utilized traditional handcrafted methods such as shape, texture, or color features, indicating a distinct transition toward deep feature extraction.

3.5. Deep learning task and method

A deep learning classifier is a system that learns to make decisions by identifying patterns in data through layers of artificial neurons, analogous to the functioning of the human brain. In contrast to conventional approaches that typically require human intervention to identify significant features, deep learning autonomously extracts these features from raw data. This capability is particularly effective for intricate tasks such as facial recognition, speech comprehension, and language translation, where manual feature selection poses significant challenges. Table 5 provides an overview of selected studies utilizing deep learning techniques for the classification of medicinal plants. The analysis of these studies identifies distinct patterns in model selection and performance. CNNs represent the predominant methodology, featured in around 12 studies, with notable architectures such as VGG (16/19), ResNet variants, and MobileNet versions. Transfer learning strategies, utilized in approximately nine studies, predominantly employed the same CNN architectures as base models, specifically VGG16 and EfficientNet variants, illustrating their adaptability across various datasets.

Table 5
The employed DL (deep learning) task, method, and performance of selected studies

Refs.	DL task	Deep learning method	Performance (%)
[44]	Classification	EfficientNetB0, EfficientNetV2-S, ViT, BEiT	BEiT: 99.14% (accuracy)
[45]	Classification	VGG16, CCT, ConvMixer, CNN, Hybrid CNN-ViT	CNN: 83% AUC, Hybrid: 74% AUC
[55]	Classification	AELGNet (LocalNet + GlobalNet)	99.71% accuracy, 99.80% precision
[46]	Classification	Hybrid MAE with parallel classification branch	98.73% (Top-1 accuracy)
[31]	Classification	Ensemble of MobileNetV2, InceptionV3, ResNet50 (weighted average)	99.66% (test set), 99.9% (cross-validation)
[39]	Classification	Ensemble of DRD, DRCD, IRCD (hard and soft voting)	99% (soft ensemble), 98% (hard ensemble), 97% (DRCD)
[40]	Classification	Ensemble of VGG16, VGG19, DenseNet201 (averaging and weighted averaging)	VGG16: 93.67%, VGG19: 92.26%, DenseNet201: 98.93%, Ensemble (VGG19 + DenseNet201): 99.12%
[47]	Classification	Majority voting-based ensemble of EfficientNetB0, B2, B4	EfficientNetB0: 99.74%, EfficientNetB2: 99.91%, EfficientNetB4: 99.93%, Ensemble: 99.96%
[48]	Classification	Enhanced CNN with modified ResNet50 + Progressive Transfer Learning; optimized SVM (OSVM) for classification	Modified ResNet50 + OSVM: 98.5% (training), 96.8% (testing); ResNet50 + SVM: 82% (training), 79% (testing); VGG16 + SVM: 78% (training), 73% (testing)
[49]	Classification	Transfer learning with fine-tuning of pre-trained models (ResNet34, DenseNet121, VGG11, ConvNeXt, Swin Transformer)	ConvNeXt: 92.5%; DenseNet121: 89.1%; ResNet34: 86.5%; VGG11: 87.0%; Swin t: 76.6%; Scratch model: 53.9%
[13]	Classification	VGG16, VGG19, DenseNet201, ResNet50V2, InceptionResNetV2, InceptionV3	DenseNet201: 99.64% (Pt), 97% (PFt); ResNet50V2: 98.6%; VGG16: 94.6%; VGG19: 91.1%; InceptionV3: 97.8%; Xception: 97.5%; InceptionResNetV2: 96.3%
[50]	Classification	Cascaded ResNet50-PSO-SVM	ResNet50-PSO-SVM: 99.6%; VGG19-PSO-SVM: 98.93%; VGG16-PSO-KNN: 98.61%
[56]	Classification	Hybrid models (feature fusion + KNN/DL classifiers)	Hybrid Model 1: 85.85%, Hybrid Model 2: 88%, Hybrid Model 3: 94.24%
[51]	Classification	Transfer learning with fine-tuning	VGG19: 94%, VGG16: 92%, Inception-V3: 91%, Xception: 87%
[32]	Classification	Transfer learning with focal loss and class weighting	Offline: 87% (Top-1), real time: 78.5% (Top-1)
[33]	Classification	Transfer learning (VGG16, VGG19, MobileNetV2)	MobileNetV2 (fine-tuned): 96.02% validation, 81.82% testing
[57]	Classification	Hybrid CNN (residual + inverted residual blocks), BCO, feature fusion	WNN classifier: 99.9% accuracy (inverted residual block)
[30]	Classification	Transfer learning with fine-tuning of pre-trained models (Xception, VGG16, VGG19, InceptionV3), classified using ANN and SVM (with Bayesian optimization for hyperparameter tuning)	- Xception + ANN: 97.5% accuracy (best model)—Xception + SVM + Bayesian optimization: 95.20% accuracy—InceptionV3 + ANN: 96.16% accuracy—VGG19 + ANN: 95.97% accuracy
[41]	Classification	Hybrid CNN and VGG-16 with HP-BSGD classifier	97% (Flavia), 98.85% (Swedish)
[52]	Classification	Multi-task joint learning network (MTJNet)	99.71% accuracy, 99.60% precision, 99.62% recall, 99.58% F1-score
[34]	Classification	Custom CNN with GAP layer	99.66% (64 × 64), 99.32% (128 × 128), 99.45% (256 × 256)
[58]	Classification	HCVINet (Hierarchical CNN + NLP integration)	98.3% (Flavia), 97.36% (Indian Medicinal Dataset)
[35]	Classification	OTAMNet (DenseNet201 + Log-Gabor fusion)	MyDataset: 98%, Flavia: 99%, Swedish Leaf: 100%, MD2020: 99%, Folio: 97%
[42]	Classification	Ayur-PlantNet (lightweight CNN with feature concatenation blocks)	92.27% (with cross-validation), 93.10% (without cross-validation)

(*Continued*)

Table 5
(*Continued*)

Refs.	DL task	Deep learning method	Performance (%)
[36]	Classification	MobileNetV2 (transfer learning)	98.05%
[37]	Classification	MobileNetV3 (proposed model with feature fusion)	97.43%
[53]	Classification	DeepHybrid-OptNet	95.7% (Folio), 94.51% (Mendeley)
[38]	Instance segmentation	Mask RCNN	95.7% (classification), 100% (segmentation)
[54]	Classification	FedAvg, FedProx (federated learning)	94.51% (IID), 82.65% (non-IID)
[43]	Classification and object detection	CNN, Recurrent Neural Networks (RNN), LSTM, BILSTM, RCNN, Fast RCNN, Faster RCNN	Faster RCNN: 96.53%, Fast RCNN: 94.6%, RCNN: 93.63%, CNN: 92.56%

Ensemble methods were implemented in five studies, often combining multiple CNN architectures through techniques like weighted averaging or majority voting, consistently achieving some of the highest reported accuracies (up to 99.96%). Hybrid approaches, featured in six studies, typically merged CNN architectures with other techniques such as ViTs or traditional machine learning classifiers, yielding competitive performance. Notably, four studies developed custom architectures specifically for plant classification tasks, with AELGNet achieving particularly impressive results (99.71% accuracy). Transformer-based models, while less common (appearing in three studies), demonstrated strong potential, with BEiT reaching 99.14% accuracy. Performance metrics across studies were predominantly high, with the majority of models attaining accuracy levels between 97% and 99%. OTAMNet, the highest-performing model, attained a perfect classification accuracy of 100% on the Swedish Leaf dataset. While complex models frequently yield superior results, certain studies indicate that simpler CNN architectures can attain competitive performance when optimized effectively. Table 5 presents several examples of emerging methodologies, such as federated learning (FL) and multi-task networks, indicating potential avenues for future research. The data suggest that, although researchers are investigating new architectures and combinations, established CNN-based methods continue to serve as the foundation for the majority of effective medicinal plant classification systems, especially when improved by transfer learning or ensemble techniques.

4. Discussion

4.1. Demography of publication

In recent years, classifying medicinal plants using deep learning has become a promising topic of research, as shown by a clear rise in academic interest around the world (Figure 3). This trend is probably the result of the necessity to digitize and protect botanical information, especially for healthcare and pharmaceutical uses, coming together with progress in AI. Even if things are going well, our research shows that contributions are not evenly spread out across different areas. A lower, but still important, number of works are coming out of African countries. However, there isn't much information coming from other biodiverse areas, like South America, especially the Amazon basin, which has a huge variety of medicinal plants that haven't been studied enough in terms of computer classification. This shows that more people throughout the world need to be aware of and work together to make the most of the world's medicinal plant resources. Our research also shows that there is an imbalance in the disciplines of the researchers working in this area. Furthermore, the disciplinary dominance of computer science and engineering (Table 2), while ensuring technical proficiency, may limit the biological and ecological depth of the classification models. Experts in biosystems engineering, biochemistry, forestry, or medicine are only involved in a tiny number of investigations. This lack of involvement from people with different backgrounds can make it harder to come up with more complete and useful answers. So, encouraging collaboration between different fields is important to make sure that future research is both technically sound and relevant to the field. Building connections across countries and fields could greatly speed up innovation and make it easier to use deep learning-based systems for classifying medicinal plants in the real world.

4.2. Dataset preparation

Our analysis reveals critical challenges in dataset curation that hinder progress in the field. A primary issue is the lack of standardization, leading to fragmented datasets with significant species overlap and redundancy. The common recurrence of similar images across different sources introduces bias, compromising model reliability. Furthermore, the high variation in image backgrounds—from plain white to complex natural settings—creates inconsistent conditions that impede fair comparison of deep learning architectures across studies.

We advise setting up a centralized and collaborative system for dataset curation to solve these problems. This kind of system might bring together contributions from different areas and experts, making sure that each species is represented in a unique and full way. To improve the dataset's coverage of biodiversity, it should focus on species that are not well represented and are only found in certain areas. Also, following rules for how to collect images—such as using the same lighting, backgrounds, and annotation formats—would make the dataset more consistent. The PlantVillage dataset [59] is an example of a comparable project that has already helped the plant pathology and agricultural community. It shows how a well-organized repository may speed up development in automated plant identification and disease detection. A consolidated benchmark dataset with clear documentation and contributions from a variety of ecosystems would be a big step toward making it possible to compare models fairly and perform research that can be repeated.

Another key insight is that model efficacy is not solely dependent on large image volumes per class. Some current datasets, even if they only have a small number of samples for each species, have shown good classification results. This means that even with little amounts of data, the quality of the model, the choice of architecture, and the preprocessing of the data can have a big effect on performance. For instance, CNNs that use attention processes or transfer learning from pre-trained models like ResNet or EfficientNet have been demonstrated to operate effectively on small datasets. So, when making new datasets in the future, the main goal should be to include more plant species (i.e., classes) rather than just adding more samples within each class. This change in focus can help create models that are stronger and can recognize a wider range of medicinal plants, many of which are only found in certain areas and are very important for medicine.

Beyond imagery, integrating multimodal data offers a significant opportunity for improvement. In plant classification, leaf size and texture are two important distinguishing properties. But in standard image databases, these features are often lost or not used enough. Deep learning models can be made better at telling the difference between things by adding manual measures like leaf length, width, and surface texture descriptors to picture data. This multimodal method might make it less necessary to use huge image datasets and let us employ simpler designs with fewer learnable parameters. For example, fusion models that combine scalar metadata with visual inputs have been used successfully in fields like medical imaging and remote sensing. A similar feature fusion technique can be employed to identify medicinal plants where image and tabular features (leaf length, leaf width, etc.) can be concatenated to make a final feature vector. Also, this could make it easier to deploy in real time and on the go, especially in places where resources are limited and computing efficiency is very important.

Finally, we recommend a pivot toward developing lightweight, hardware-efficient models for practical deployment. ViTs and deep ResNets are examples of huge models that have shown amazing accuracy, but they can't be used in the real world very often because they need a lot of computing power. Combining multimodal data with frameworks like MobileNet or TinyML, on the other hand, could lead to useful tools for farmers, herbalists, and researchers in remote places that can identify therapeutic plants on their phones. Also, these kinds of applications would be better off with a benchmark dataset that is approved around the world and includes species and environmental conditions from all over the world. We think this path can help make botanical knowledge more accessible to everyone and help protect biodiversity. In short, our study shows that deep learning can help with medicinal plant classification by making it more standardized, collaborative, and multimodal. A well-organized dataset that focuses on species diversity and background variation and is created by the community can be a great base for new ideas. Future models can be more accurate, accessible, and useful by combining both picture and physical aspects and making them work better.

4.3. Preprocessing and augmentation technique

The study of 30 papers on applying deep learning to classify medicinal plants shows that preprocessing and data augmentation are very important for improving model performance and resilience. All of the research we looked at included preprocessing as an important and necessary step. This shows how important it is to get picture data ready for training and inference. Resizing, normalizing, denoising, and segmentation were some of the most common techniques used. Resizing, notably to 224×224 pixels, was the most common method used. These processes all work together to lower variability, cut down on noise, and make sure that the dataset is consistent, which is necessary for getting high classification accuracy.

Meanwhile, 73.33% of the research used data augmentation, which shows that it is useful for making datasets more diverse without having to collect more data by hand. Standard methods of augmentation, such as rotation, flipping, cropping, and changing brightness, help models work better with different situations in the actual world. But you should be careful when choosing enhancement strategies. Non-rigid or non-isometric transformations, including twisting or excessive warping, can change the appearance of medicinal plant leaves or structures in a way that makes it hard to tell what they really look like. This could lead to model confusion or learning of distorted features. So, it is important to choose augmentation methods carefully so that they don't change the meaning of the object while adding useful variation.

One intriguing thing about the research we looked at is that they used both advanced enhancement methods like CLAHE, unsharp masking, and background removal and more traditional image processing approaches like edge detection and grayscale conversion. These hybrid preprocessing pipelines show a trend toward using both deep learning and standard computer vision techniques to make features clearer and easier to find. Additionally, other research used domain-specific feature engineering methods like GLCM and LBP texture analysis and morphological segmentation approaches to make the data representation even better and possibly improve the results of classification.

The fact that most research consistently use preprocessing and carefully apply augmentation implies that they all recognize how important these steps are. Preprocessing makes sure that the input is of excellent quality, and augmentation mimics real-world changes in leaf form, texture, and color, which helps make the model more robust. Future studies should keep improving these stages, focusing on biologically realistic changes and using adaptive preprocessing methodologies that are specific to the plant domain.

4.4. Studied organs and feature extraction techniques

One key thing to note from the literature that was looked at is that leaves are the main organ used to classify medicinal plants. Twenty-four of the 30 studies looked only at leaf photos. Only a handful looked at other sections of the plant, such as stems, flowers, fruits, or the whole plant. This trend shows how important leaves are for diagnosis because they frequently include unique morphological traits including form, border, venation, and texture that may accurately tell one species from another. Not all medicinal plants have big, easy-to-see leaves, but for most of them, leaves are a great and easy way to tell them apart. Their availability for most of the year, ease of collecting, and capacity to work with image-based analysis all make them even more important in automatic classification systems.

Most studies have moved away from traditional handcrafted methods and toward deep learning-based models, especially CNNs like VGG16, ResNet50, and DenseNet201. The fact that more and more studies are using ViTs and hybrid CNN-transformer architectures shows that researchers are moving toward more advanced ways of extracting features and learning based on attention. This evolution shows how the field is keeping up with bigger advances in computer vision, using deep features instead of handmade descriptors to make plant classification jobs more accurate and generalizable.

4.5. Deep learning task and method

The fact that there are so many different deep learning models used in the 30 research we looked at shows how quickly computer vision systems are changing and how hard it is to classify medicinal plants. CNNs, either in their basic forms or with unique improvements, are clearly the most popular type of neural network. Many people used popular CNN designs like VGG16, ResNet50, DenseNet201, InceptionV3, and Xception since they were known to be good at extracting hierarchical visual information. Table 6 shows a list of the learnable parameters that these models need. Most of these models have between 20 and 120 million learnable parameters. VGG16 (~138M) and DenseNet201 (~20M) are two examples of models that commonly find a good balance between model complexity and performance. We also employed lighter models like MobileNetV2 and EfficientNetB0, which have between 2 million and 7 million parameters and are good for contexts with limited resources.

It's interesting that more and more research are looking into ViTs and their offshoots, such as BEiT, CCT, MAE, and Swin Transformer. These architectures were first created for big vision problems, but now they may be used for plant picture categorization as well. Their parameter sizes usually range from 85M to 150M; however, smaller versions like CCT can make things much simpler, bringing the size down to as low as 5M–20M. The use of ViTs in five research shows a move toward models that can better capture long-range dependencies and contextual cues

Table 6
Summary of approximate learnable parameters for selected models

Model category	Subtypes	Models	Parameters (approx.)
CNN-based models	Lightweight CNNs	MobileNetV2, MobileNetV3, EfficientNetB0	~2M–7M
	Medium-sized CNNs	VGG16, VGG19, ResNet34, DenseNet121	~20M–40M
	Large CNNs	ResNet50, ResNet50V2, InceptionV3, Xception, DenseNet201	~23M–55M
	Very large CNNs	InceptionResNetV2, EfficientNetB4, ConvNeXt	~55M–120M+
	Custom/modified CNNs	Ayur-PlantNet, OTAMNet, GAP-layer CNN, HCVINet	~1M–40M (varies widely)
Transformer-based models	Base Transformers	ViT-B, BEiT, MAE, Swin Transformer	~85M–120M
	Compact Transformers	CCT (Compact Convolutional Transformer)	~5M–20M
	Larger ViT Variants	ViT-Large, Swin-Large	~150M–300M+
Hybrid and fusion models	CNN + Transformer	CNN-ViT hybrids, Hybrid MAE, AELGNet	~50M–150M+
	CNN + Attention/Other	Hybrid CNN-VGG-HP-BSGD, DeepHybrid-OptNet	~20M–100M
	Multi-branch/Fusion Nets	MTJNet, HCVINet, OTAMNet	~15M–70M
Ensemble models	Light CNNs	MobileNetV2 + ResNet50	~25M–60M
	Mixed large CNNs	VGG16 + DenseNet201 + InceptionV3	~80M–140M+
		EfficientNetB0 + B2 + B4 ensemble	~60M–100M
Transfer learning	Feature extraction	frozen backbone	~1M–10M (classifier only)
	Full fine-tuning	ResNet50, VGG	~20M–50M
	TL with transformers	Swin, ViT	~80M–120M+
Detection and segmentation models	–	Mask RCNN	~40M–140M+
	–	Faster RCNN	~40M–60M
Recurrent models	Basic RNN/LSTM	–	~1M–10M
	Stacked + Bidirectional	BiLSTM, RCNN	~10M–30M+
	Integrated with CNNs or attention	–	~20M–60M+
Federated learning	–	FedAvg, FedProx (MobileNetV2, ResNet50)	Same as base model

in images. This is especially helpful for spotting small changes in the architecture of medicinal plants.

There were also a lot of hybrid models that combined CNNs with transformers, attention mechanisms, handcrafted features, or other classical algorithms. Depending on how hard it is to integrate them, these models can be anywhere from 15 million to more than 150 million parameters. The goal of hybridization is often to use the best parts of several architectures together. For example, CNNs can extract local features while transformers can reason globally. Models like Hybrid MAE, AELGNet, DeepHybrid-OptNet, and multi-branch networks like MTJNet showed these kinds of combinations. These methods are part of a larger trend in AI research to create domain-specific architectures that adapt general-purpose networks to more complex classification tasks.

Another important group used ensemble techniques, which integrated several pre-trained models (such as VGG16, DenseNet201, and InceptionV3) using methods like majority voting or weighted averaging. These ensembles, which require a lot of computing power (with parameter counts frequently above 100M), demonstrated encouraging results in making classification more robust by lowering the bias of each model. But there are still worries about how well they can be used in the real world, especially in places with few resources.

Several research found that transfer learning had a big difference, especially when the datasets were small. A lot of people used fine-tuning on pre-trained CNN or transformer backbones. This lets them keep the benefits of high-capacity models while simply training a few of the parameters. When only the last classifier was taught, the number of learnable parameters reduced significantly to between 1M and 10M. This meant that there was a trade-off between performance and training cost. Lastly, there was a minor but growing trend toward FL (like FedAvg and FedProx) and recurrent models (like Long Short-Term Memory (LSTM) and Region Based Convolutional Neural Networks (RCNN)). These were early attempts to deal with privacy, decentralization, or temporal modeling in plant data. In the past, people sometimes utilized classical classifiers like Support Vector Machine (SVMs) or K-Nearest Neighbors (KNNs) in hybrid scenarios, but now deep architectures that can automatically extract features have mostly taken their place. The fact that it consistently performs well in many studies implies that deep learning is now a reliable tool for this type of categorization task.

4.6. Future direction

The research landscape in medicinal plant classification using deep learning has significantly evolved in recent years. But there are still a few potential paths that haven't been fully explored, and they offer chances for important future work. Most of the studies that are out there have one big problem: they only look at leaf-based classification. Leaves are usually the easiest and most recognizable portion of a plant, but they don't show the whole range of morphological variability. Many medicinal plants have stems, flowers, fruits, and bark that are different from each other, especially when there aren't many leaves or none at all. Future studies should focus on creating models that include images of full plants, even those without leaves, to make sure they are strong and can be used all year round.

Another important goal is to identify and classify uncommon and endangered medicinal plants. These species are very important for both the environment and medicine; however, they are typically not included in existing datasets. Making big databases that include rare species and teaching deep learning models to find them could help protect biodiversity and support traditional medicine. Additionally, these kinds of models could help researchers and conservationists keep an eye on and track endangered animals in real time, giving them useful information.

There is still a lack in the literature about how to use medicinal plant classification models in the actual world, beyond their use in school. Making these models available in easy-to-use mobile or web-based apps can make it easier for local populations, traditional healers, herbal product makers, and researchers to use them. This kind of deployment would make plant identification tools available to more people and increase public knowledge and awareness of medicinal plants. Some researchers are working not only to identify a leaf's species but also to assess its health and then evaluate either its medicinal value (if healthy) or mark it as unusable for medicine (if unhealthy), which will help Ayurvedic practitioners, researchers, and pharmaceutical companies [60].

Another interesting subject to look at is how to make deep learning models work better in real-world and distributed settings, such as FL. Most of the models we have now use a lot of resources and need a lot of computing power, which makes them hard to adapt to decentralized networks. Making models lighter and with fewer learnable parameters can make them operate with FL frameworks. This would allow for the safe and private collection of medicinal plant data from different areas, which would make the models more useful across different ecological zones. Federated updates let models get better all the time with data obtained in different places, so there is no need to send sensitive information. This makes the system scalable and privacy-conscious.

In short, future research should look at more plant traits, make sure that rare species are included, make sure that the models can be used in the real world, and focus on making the models work better in federated and low-resource settings. These steps not only fix problems that already exist, but they also make it possible for deep learning to be used in more useful, accessible, and inclusive ways in research on medicinal plants.

5. Conclusion

The application of deep learning to medicinal plant classification has gained substantial momentum in recent years. This is because AI has made great strides and there is an urgent need to digitize traditional botanical knowledge. We looked over 30 research publications and found a number of important insights and ongoing problems that show how far this subject has come and how far it still has to go. Most of the contributions come from Asian countries (25 out of 30 studies), which have a lot of traditional medical knowledge. However, biodiverse places like the Amazon basin are still very underrepresented. This uneven distribution of research around the world shows that we need to work together more internationally and across disciplines, especially with biological and ecological specialists, to make AI models more useful in the real world for plant science. Preparing datasets became a major issue, as most research used image datasets that were made by individuals and had problems with how species were represented, how backgrounds were set up, and how annotations were made. These differences make it harder to compare models and make them harder to reproduce. Our study strongly supports the creation of a centralized, community-driven dataset architecture that puts biodiversity, geographic variety, and consistent data gathering techniques at the top of its list of priorities. We observed that it is still possible to get accurate classifications with smaller datasets when utilizing sophisticated designs and good preprocessing. This suggests that we should focus on boosting species diversity instead of the number of samples within a class. All of the investigations used preprocessing and data augmentation techniques, with scaling, normalization, and denoising being the most common. Data augmentation, especially modifications that make sense biologically, such as rotating and changing brightness, was very important for making models broader. Combining traditional computer vision methods with contemporary deep learning made features even clearer, showing how useful hybrid methods can be.

Leaves were the most researched plant organs (29 out of 30 studies) since they are easy to get to and have a lot of different shapes. But this narrow focus makes categorization systems less reliable, especially when there aren't any leaves around. We think that future studies should incorporate other parts of plants, such as flowers, stems, fruits, and bark, to make plant identification more complete. CNNs are still the most popular architecture, although ViTs and hybrid CNN-transformer models are getting more and more attention and most of the models offer 90–99% accuracy. Transfer learning and ensemble approaches were often utilized to get the best results, even with small datasets. Lightweight models like MobileNet showed promise for application in places with low resources.

This assessment points up several future directions, such as classifying rare and endangered species, merging different types of data (such as visual data with physical measures), and using mobile or online apps in the real world. Optimizing models for privacy-preserving frameworks like FL is also a scalable and safe way to collect data in a decentralized way, especially in areas that are environmentally varied but don't have a lot of technology. In short, deep learning has shown a lot of promise for classifying medicinal plants, but to fully realize its potential, we need to fix problems with the quality of the datasets, the geographic coverage, and the need for collaboration between different fields. The discipline may make big steps toward helping healthcare, conservation, and traditional medicine by using standardized data methods, lightweight and easy-to-understand models, and doing more studies on plant traits and locations that aren't well covered.

Funding Support

This research is funded by Multimedia University, Cyberjaya, Selangor, Malaysia (Grant Number: PostDoc (MMUI/240029)).

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Monir Hossain: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Fahmid Al Farid:** Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Visualization. **Momotaz Begum:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization. **Jia Uddin:** Conceptualization, Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Visualization, Supervision, Project administration. **Hezerul Bin Abdul Karim:** Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Visualization, Funding acquisition.

References

[1] Ur Rehman, F., Kalsoom, M., Adnan, M., Fazeli-Nasab, B., Naz, N., Ilahi, H., . . . , & Ali, Z. (2021). Importance of medicinal plants in human and plant pathology: A review. *International Journal of Pharmacy & Biomedical Research*, *8*(2), 1–11. https://doi.org/10.18782/2394-3726.1110

[2] van Wyk, A. S., & Prinsloo, G. (2020). Health, safety and quality concerns of plant-based traditional medicines and herbal remedies. *South African Journal of Botany*, *133*, 54–62. https://doi.org/10.1016/j.sajb.2020.06.031

[3] El-Saadony, M. T., Saad, A. M., Mohammed, D. M., Korma, S. A., Alshahrani, M. Y., Ahmed, A. E., . . . , & Ibrahim, S. A. (2025). Medicinal plants: Bioactive compounds, biological activities, combating multidrug-resistant microorganisms, and human health benefits-A comprehensive review. *Frontiers in Immunology*, *16*, 1491777. https://doi.org/10.3389/fimmu.2025.1491777

[4] Salleh, N. H., Zulkipli, I. N., Mohd Yasin, H., Ja'afar, F., Ahmad, N., Wan Ahmad, W. A. N., & Ahmad, S. R. (2021). Systematic review of medicinal plants used for treatment of diabetes in human clinical trials: An ASEAN perspective. *Evidence-Based Complementary and Alternative Medicine*, *2021*(1) 5570939. https://doi.org/10.1155/2021/5570939

[5] Tilburt, J. C., & Kaptchuk, T. J. (2008). Herbal medicine research and global health: An ethical analysis. *Bulletin of the World Health Organization*, *86*(8), 594–599. https://doi.org/10.2471/BLT.07.042820

[6] Ssenku, J. E., Okurut, S. A., Namuli, A., Kudamba, A., Tugume, P., Matovu, P., . . . , & Walusansa, A. (2022). Medicinal plant use, conservation, and the associated traditional knowledge in rural communities in Eastern Uganda. *Tropical Medicine and Health*, *50*(1), 39. https://doi.org/10.1186/s41182-022-00428-1

[7] Raven, P. H., & Wagner, D. L. (2021). Agricultural intensification and climate change are rapidly decreasing insect biodiversity. *Proceedings of the National Academy of Sciences*, *118*(2), e2002548117. https://doi.org/10.1073/pnas.2002548117

[8] Saggar, S., Mir, P. A., Kumar, N., Chawla, A., Uppal, J., Shilpa, & Kaur, A. (2022). Traditional and herbal medicines: Opportunities and challenges. *Pharmacognosy Research*, *14*(2), 107–114. https://doi.org/10.5530/pres.14.2.15

[9] Pawar, A. T., Tagalpallewar, A. A., Mishra, M., Swami, A. G., & Baheti, A. M. (2025). Medicinal plant biotechnology. In U. E. Odoh, S. S. Gurav, & M. O. Chukwuma (Eds.), *Pharmacognosy and phytochemistry: Principles, techniques, and clinical applications* (1st ed., pp. 347–370). Wiley. https://doi.org/10.1002/9781394203680.ch18

[10] Eshete, M. A., & Molla, E. L. (2021). Cultural significance of medicinal plants in healing human ailments among Guji

semi-pastoralist people, Suro Barguda District, Ethiopia. *Journal of Ethnobiology and Ethnomedicine*, *17*(1), 61. https://doi.org/10.1186/s13002-021-00487-4
[11] Wäldchen, J., & Mäder, P. (2018). Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering*, *25*(2), 507–543. https://doi.org/10.1007/s11831-016-9206-z
[12] Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition*, *71*, 1–13. https://doi.org/10.1016/j.patcog.2017.05.015
[13] Dey, B., Ferdous, J., Ahmed, R., & Hossain, J. (2024). Assessing deep convolutional neural network models and their comparative performance for automated medicinal plant identification from leaf images. *Heliyon*, *10*(1), e23655. https://doi.org/10.1016/j.heliyon.2023.e23655
[14] Archana, R., & Jeevaraj, P. E. (2024). Deep learning models for digital image processing: A review. *Artificial Intelligence Review*, *57*(1), 11. https://doi.org/10.1007/s10462-023-10631-z
[15] Elngar, A. A., Arafa, M., Fathy, A., Moustafa, B., Mahmoud, O., Shaban, M., & Fawzy, N. (2021). Image classification based on CNN: A survey. *Journal of Cybersecurity and Information Management*, *6*(1), 18–50. https://doi.org/10.5281/zenodo.4897990
[16] Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of image classification algorithms based on convolutional neural networks. *Remote Sensing*, *13*(22), 4712. https://doi.org/10.3390/rs13224712
[17] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *arXiv Preprint: 1409.1556*
[18] Girshick, R. (2015). Fast R-CNN. In *2015 IEEE International Conference on Computer Vision*, 1440–1448. https://doi.org/10.1109/ICCV.2015.169
[19] Li, W. (2021). Analysis of object detection performance based on Faster R-CNN. *Journal of Physics: Conference Series*, *1827*(1), 012085. https://doi.org/10.1088/1742-6596/1827/1/012085
[20] Shafiq, M., & Gu, Z. (2022). Deep residual learning for image recognition: A survey. *Applied Sciences*, *12*(18), 8972. https://doi.org/10.3390/app12188972
[21] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . , & A, Rabinovich. (2015). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594
[22] Ashwinkumar, S., Rajagopal, S., Manimaran, V., & Jegajothi, B. (2022). Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks. *Materials Today: Proceedings*, *51*, 480–487. https://doi.org/10.1016/j.matpr.2021.05.584
[23] Zhou, T., Ye, X., Lu, H., Zheng, X., Qiu, S., & Liu, Y. (2022). Dense convolutional network and its application in medical image analysis. *BioMed Research International*, *2022*(1) 2384830. https://doi.org/10.1155/2022/2384830
[24] Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, *61*, 101182. https://doi.org/10.1016/j.ecoinf.2020.101182
[25] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
[26] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *2017 IEEE International Conference on Computer Vision*, 2980–2988. https://doi.org/10.1109/ICCV.2017.322
[27] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 1800–1807. https://doi.org/10.1109/CVPR.2017.195
[28] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., . . . , & He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555
[29] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . , & Polosukhin, I. (2017). Attention is all you need. *arXiv Preprint:1706.03762*.
[30] Roopashree, S., & Anitha, J. (2021). DeepHerb: A vision-based system for medicinal plants using Xception features. *IEEE Access*, *9*, 135927–135941. https://doi.org/10.1109/ACCESS.2021.3116207
[31] Sachar, S., & Kumar, A. (2022). Deep ensemble learning for automatic medicinal leaf identification. *International Journal of Information Technology*, *14*(6), 3089–3097. https://doi.org/10.1007/s41870-022-01055-z
[32] Malik, O. A., Ismail, N., Hussein, B. R., & Yahya, U. (2022). Automated real-time identification of medicinal plant species in natural environment using deep learning models—A case study from Borneo region. *Plants*, *11*(15), 1952. https://doi.org/10.3390/plants11151952
[33] Ayumi, V., Ermatita, E., Abdiansah, A., Noprisson, H., Jumaryadi, Y., Purba, M., . . . , & Putra, E. D. (2022). Transfer learning for medicinal plant leaves recognition: A comparison with and without a fine-tuning strategy. *International Journal of Advanced Computer Science and Applications*, *13*(9), 138–144. https://doi.org/10.14569/IJACSA.2022.0130916
[34] Azadnia, R., Al-Amidi, M. M., Mohammadi, H., Cifci, M. A., Daryab, A., & Cavallo, E. (2022). An AI based approach for medicinal plant identification using deep CNN based on global average pooling. *Agronomy*, *12*(11), 2723. https://doi.org/10.3390/agronomy12112723
[35] Oppong, S. O., Twum, F., Hayfron-Acquah, J. B., & Missah, Y. M. (2022). A novel computer vision model for medicinal plant identification using Log-Gabor filters and deep learning algorithms. *Computational Intelligence and Neuroscience*, *2022*(1), 1189509. https://doi.org/10.1155/2022/1189509
[36] Abdollahi, J. (2022). Identification of medicinal plants in Ardabil using deep learning: Identification of medicinal plants using deep learning. In *2022 27th Iranian Conference on Electrical Engineering*, 1–6. https://doi.org/10.1109/CSICC55295.2022.9780493
[37] Valdez, D. B., Aliac, C. J. G., & Feliscuzo, L. S. (2022). Medicinal plant classification using convolutional neural network and transfer learning. In *2022 IEEE International Conference on Artificial Intelligence Engineering and Technology*, 1–6. https://doi.org/10.1109/IICAIET55139.2022.9936868
[38] Almazaydeh, L., Alsalameen, R., & Elleithy, K. (2022). Herbal leaf recognition using mask-region convolutional neural network (Mask R-CNN). *Journal of Theoretical and Applied Information Technology*, *100*(11), 3664–3671.
[39] Uddin, A. H., Chen, Y.-L., Borkatullah, B., Khatun, M. S., Ferdous, J., Mahmud, P., . . . , & Por, L. Y. (2023).

Deep-learning-based classification of Bangladeshi medicinal plants using neural ensemble models. *Mathematics*, *11*(16), 3504. https://doi.org/10.3390/math11163504

[40] Hajam, M. A., Arif, T., Khanday, A. M. U. D., & Neshat, M. (2023). An effective ensemble convolutional learning model with fine-tuning for medicinal plant leaf identification. *Information*, *14*(11), 618. https://doi.org/10.3390/info14110618

[41] Barhate, D., Pathak, S., & Dubey, A. K. (2023). Hyperparameter-tuned batch-updated stochastic gradient descent: Plant species identification by using hybrid deep learning. *Ecological Informatics*, *75*, 102094. https://doi.org/10.1016/j.ecoinf.2023.102094

[42] Pushpa, B. R., & Rani, N. S. (2023). Ayur-PlantNet: An unbiased lightweight deep convolutional neural network for Indian Ayurvedic plant species classification. *Journal of Applied Research on Medicinal and Aromatic Plants*, *34*, 100459. https://doi.org/10.1016/j.jarmap.2023.100459

[43] Duth, P. S., Vedavathi, S., & Roshan, S. (2023). Herbal leaf classification using RCNN, Fast RCNN, Faster RCNN. In *2023 7th International Conference on Computing, Communication, Control and Automation*, 1–8. https://doi.org/10.1109/ICCUBEA58933.2023.10392240

[44] Nhut, D. T. N., Tan, T. D., Quoc, T. N., & Hoang, V. T. (2024). Medicinal plant recognition based on Vision Transformer and BEiT. *Procedia Computer Science*, *234*, 188–195. https://doi.org/10.1016/j.procs.2024.02.165

[45] Hossain, S., Hasan, R., & Uddin, J. (2025). Medicinal plant leaf classification using deep learning and vision transformers. *Baghdad Science Journal*, *22*(3), 30. https://doi.org/10.21123/bsj.2024.10844

[46] Tan, C., Tian, L., Wu, C., & Li, K. (2024). Rapid identification of medicinal plants via visual feature-based deep learning. *Plant Methods*, *20*(1), 81. https://doi.org/10.1186/s13007-024-01202-6

[47] Kiflie, M. A., Sharma, D. P., Haile, M. A., & Srinivasagan, R. (2024). EfficientNet ensemble learning: Identifying Ethiopian medicinal plant species and traditional uses by integrating modern technology with ethnobotanical wisdom. *Computers*, *13*(2), 38. https://doi.org/10.3390/computers13020038

[48] Diwedi, H. K., Misra, A., & Tiwari, A. K. (2024). CNN-based medicinal plant identification and classification using optimized SVM. *Multimedia Tools and Applications*, *83*(11), 33823–33853. https://doi.org/10.1007/s11042-023-16733-8

[49] Musyaffa, M. S. I., Yudistira, N., Rahman, M. A., Basori, A. H., Mansur, A. B. F., & Batoro, J. (2024). IndoHerb: Indonesia medicinal plants recognition using transfer learning and deep learning. *Heliyon*, *10*(23), e40606. https://doi.org/10.1016/j.heliyon.2024.e40606

[50] Islam, M. T., Rahman, W., Hossain, M. S., Roksana, K., Azpíroz, I. D., Diaz, R. M., . . . , & Samad, M. A. (2024). Medicinal plant classification using particle swarm optimized cascaded network. *IEEE Access*, *12*, 42465–42478. https://doi.org/10.1109/ACCESS.2024.3378262

[51] Kiflie, M. A., Sharma, D. P., & Haile, M. A. (2024). Deep learning for Ethiopian indigenous medicinal plant species identification and classification. *Journal of Ayurveda and Integrative Medicine*, *15*(6), 100987. https://doi.org/10.1016/j.jaim.2024.100987

[52] Sharma, S., & Vardhan, M. (2024). MTJNet: Multi-task joint learning network for advancing medicinal plant and leaf classification. *Knowledge-Based Systems*, *299*, 112147. https://doi.org/10.1016/j.knosys.2024.112147

[53] Renukaradhya, S., & Narayanappa, S. S. (2024). Deep HybridNet with hybrid optimization for enhanced medicinal plant identification and classification. *International Journal of Electrical and Computer Engineering*, *14*(5), 5626–5640. https://doi.org/10.11591/ijece.v14i5.pp5626-5640

[54] Viet, K. L. D., le, Ha., Quoc, K., N, T., & Hoang, V. T. (2024). Medicinal plants identification using federated deep learning. *Procedia Computer Science*, *234*, 247–254. https://doi.org/10.1016/j.procs.2024.02.171

[55] Sharma, S., & Vardhan, M. (2025). AELGNet: Attention-based enhanced local and global features network for medicinal leaf and plant classification. *Computers in Biology and Medicine*, *184*, 109447. https://doi.org/10.1016/j.compbiomed.2024.109447

[56] Pushpa, B. R., Jyothsna, S., & Lasya, S. (2025). HybNet: A hybrid deep models for medicinal plant species identification. *MethodsX*, *14*, 103126. https://doi.org/10.1016/j.mex.2024.103126

[57] Bouakkaz, H., Bouakkaz, M., Kerrache, C. A., & Dhelim, S. (2025). Enhanced classification of medicinal plants using deep learning and optimized CNN architectures. *Heliyon*, *11*(3), e42385. https://doi.org/10.1016/j.heliyon.2025.e42385

[58] Vani, E. S., Gururaj, H. M., & Rai, B. K. (2025). HCVINet: A multimodal deep learning approach for medicinal plant classification using visual and semantic features. *SN Computer Science*, *6*(5), 551. https://doi.org/10.1007/s42979-025-04099-7

[59] Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv Preprint:1511.08060*.

[60] Tejas, A., Reddy, K. C., Battu, S. K., & Devi, V. V. R. (2025). AI-based Ayurvedic leaf analysis and recommendations. In *Proceedings of International Conference on Computer Science and Communication Engineering*, 1018–1033. https://doi.org/10.2991/978-94-6463-858-5_85

How to Cite: Hossain, M., Farid, F. A., Begum, M., Uddin, J., & Karim, H. B. A. (2026). Deep Learning for Medicinal Plant Classification: A Comprehensive Review of Recent Advances and Challenges. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE62027652