Grenze International Journal of Engineering and Technology, January Issue

A Comprehensive Literature Review on AI-Driven Herbal Plant Identification: Machine Learning and Deep Learning based Approaches

Mahalakshmi B[1] and Dr. Laxmi B Rananavare[2]
[1-2]School of Computer Science and Engineering, Reva University, Bengaluru, Karnataka, India.
Email: mahabit@gmail.com, laxmib.rananavare@reva.edu.in

***Abstract*— Deep Learning (DL) and Machine Learning (ML) approaches have been extensively explored in recent times for their efficiency in identifying and classifying plant species and their diseases. These techniques have the potential to increase efficiency and precision in diagnosing diseases, thereby revolutionizing agriculture by improving output and yields. Many industries depend on these plants for their products, making the prospect of their disappearance a critical issue. Therefore, identifying these plants has become increasingly important. Given the vast number of plant species, manually identifying them is impractical due to the experience and time required. To expedite this process, experts suggest using DL and ML techniques. This review provides a comprehensive analysis of various DL and ML methods for identifying plant diseases, highlighting technological approaches, data sources, and metrics of efficacy. Key models discussed include Support Vector Machines (SVMs), Convolutional Neural Networks (CNNs), and hybrid models that combine multiple feature extraction techniques. While these methods have shown promising results, they still face challenges such as inconsistent datasets, high resource demands, difficulties in result interpretation, and practical implementation issues. This study offers an in-depth review of the current research landscape, identifies persistent problems, and suggests future directions for advancing the application of these technologies in the agricultural industry.**

***Index Terms*— herbal plant identification, deep learning, machine learning, literature review.**

I. INTRODUCTION

Undoubtedly, plants are essential for food, medicine, spices, clothing, shelter, and fertilizer. Above all, they play a significant role in regulating climate change. For millennia, human health has relied on medicinal plants [1]. For example, Andrographis paniculata, commonly known as Kalmegh, has been used in traditional medicine to treat skin diseases, respiratory issues, and liver disorders. Similarly, Gymnema sylvestre has been found effective in treating fever, lowering blood sugar, protecting the liver, inhibiting cancer, and reducing inflammation. Additionally, Azadirachta indica, or Neem, is believed to offer therapeutic benefits for many diseases, including those that have become resistant to multiple drugs.[4].

Due to their lower cost and fewer side effects compared to modern medicines, an increasing number of therapeutic plants are being utilized in the pharmaceutical industry. According to the World Health Organization, more than 21,000 plant species are used for medicinal purposes [5].

Grenze ID: 01.GIJET.11.1.29

Moreover, 80% of people worldwide use medicinal herbs to address their basic health conditions [6]. Additionally, medicinal plants represent a significant resource for developing novel antibiotics and biocompatible therapies. For instance, Withania somnifera has been extensively researched for its various pharmacotherapeutic uses, demonstrating a wide range of therapeutic qualities including stress and anxiety reduction, immune system modulation, anti-inflammatory effects, and improvement in sexual dysfunction [7,8].
On Earth, there are several million plant species, some of which can be used medicinally, while others are endangered or potentially dangerous. Plants are essential not only for human survival but also for maintaining the balance of the entire food web. Medicinal plants are used in herbal, Ayurvedic, and conventional medicine. These plants are defined by their natural ability to treat illnesses. Approximately 80% of people worldwide still rely on medicinal plants to address their health needs. Plants considered medicinal are those whose stems, roots, or leaves can be utilized in drug manufacturing. Many such medicinal plants are found in the wild and can provide valuable information about their therapeutic properties. One method of assessing these plants involves examining their leaves. Accurate research and classification of plants are crucial for preserving plant species and understanding their potential benefits [9].
Centuries ago, Ayurvedic physicians would personally collect plants and prepare remedies for their patients. Today, however, this traditional practice is followed by only a small segment of the population. The production and sale of Ayurvedic medicines now generate an annual turnover of about Rs. 4,000 crores, with an estimated 8,500 manufacturers of Ayurvedic preparations in India [10]. Despite this growth, the quality of the raw materials used in Ayurveda has come under scrutiny. Children and women, lacking Traditional Ecological Knowledge, have begun collecting wild plants, which has led to issues with misidentification or the use of spurious herbs.

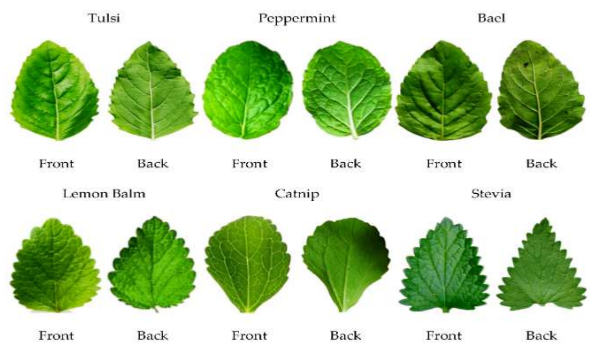

Figure 1: Description of Leaves

Many manufacturing sites receive therapeutic herbs that are not properly identified and lack the necessary quality control measures. Additionally, the use of local names for herbs varies regionally, contributing to confusion. The use of dried herbs can also complicate manual identification. Improper use of medicinal herbs can render Ayurvedic treatments ineffective and potentially cause adverse effects. Therefore, implementing robust quality control measures for Ayurvedic pharmaceuticals and raw materials is crucial to ensure the sector's growth while preserving the efficacy and authenticity of these medicines.

TABLE I. USES OF HERBS

General Name	Botanical Name	Usages
Tulsi (Holy Basil)	*Ocimum sanctum*	Boosts immunity, treats respiratory disorders, reduces stress, anti-inflammatory.
Neem	*Azadirachta indica*	Antibacterial, antifungal, anti-inflammatory, used in skin disorders, dental care, and as a blood purifier.
Kalmegh (Green Chiretta)	*Andrographis paniculata*	Treats liver disorders, respiratory issues, skin conditions, and acts as an immune booster.
Ashwagandha (Indian Ginseng)	*Withania somnifera*	Reduces stress, improves brain function, boosts immunity, and increases strength.
Giloy (Heart-leaved Moonseed)	*Tinospora cordifolia*	Boosts immunity, treats chronic fever, improves digestion, and has anti-inflammatory properties.
Amla (Indian Gooseberry)	*Phyllanthus emblica*	Rich in Vitamin C, boosts immunity, improves skin and hair health, and aids digestion.

Brahmi (Water Hyssop)	*Bacopa monnieri*	Enhances memory, reduces anxiety and stress, and improves cognitive functions.
Gymnema (Gurmar)	*Gymnema sylvestre*	Lowers blood sugar levels, protects the liver, reduces inflammation, and helps in weight management.
Shatavari (Wild Asparagus)	*Asparagus racemosus*	Supports reproductive health, boosts immunity, and acts as an adaptogen.
Turmeric	*Curcuma longa*	Anti-inflammatory, antioxidant, improves liver function, and supports joint health.

Conservative estimates suggest that the Earth loses at least one key drug candidate every two years. The current rate of plant species extinction is at least 100 to 1,000 times higher than the natural extinction rate. According to the World Wildlife Fund and the International Union for Conservation of Nature, between 50,000 and 80,000 species of flowers are used medicinally around the world. Currently, 20% of these species are already extinct in the wild, and another 15,000 species are threatened by overharvesting and habitat loss. Human activities, both direct and indirect, are the primary causes of the accelerated extinction rate. Therefore, it is essential to recognize and categorize medicinal plant species as quickly and accurately as possible for effective research and biodiversity management.

Plant species identification is typically carried out by professional botanists or taxonomists with expertise in plant taxonomy [14, 15]. Manual identification of plant species is a challenging and labour-intensive task, as each step relies on human perception, making the process prone to inaccuracies [16]. The shortage of plant identification experts has resulted in a "taxonomic impediment," hindering the accurate identification of plant species [17]. Therefore, developing a reliable and effective method for correctly identifying these valuable plants is crucial.

This includes a class of artificial intelligence known as Machine Learning (ML), which is capable of addressing complex issues across various application fields with minimal human involvement. Deep Learning (DL), a subset of ML, is inspired by the architecture and functioning of biological neurons. DL involves training algorithms to distinguish data and improve identification accuracy. Advances in hardware technology and the availability of Big Data have significantly enhanced performance, achieving notable results in areas such as natural language processing, game playing, and image processing—tasks that would otherwise be challenging for human capabilities. Recent investigations have shown considerable interest in the automatic identification of plant species using ML and DL approaches with plant image specimens. The automatic identification process typically involves several steps: (a) Image Acquisition; (b) Image Preprocessing; (c) Feature Extraction; and (d) Classification. Identification can be conducted on the entire plant or by focusing on specific parts, such as leaves, fruits, flowers, and bark. Since leaves are the most readily identifiable part of a plant and can be photographed throughout the year, researchers often use leaf images for this purpose. In contrast, photographs of flowers and fruits are generally limited to specific seasons. More than one hundred research studies have utilized leaf images for automatic plant identification. This study reviews the most advanced ML and DL techniques available today for classifying common plant leaves, plant illnesses, and therapeutic plants. The rest of the study is organized as follows: Section II briefly reviews the challenges and literature related to common and medicinal plants, while Section III presents the conclusions and future directions of the work.

II. LITERATURE SURVEY

A concise overview of the relevant literature on cutting-edge ML and DL methods for the classification and identification of plant illnesss is provided in this section. This session will discuss models for ML and DL in plant identification using their leaves.

A. Literature Review on Medicinal Plant

Roopashree et al. [19] proposed an automatic recognition model to classify medicinal plants using Machine Learning (ML) techniques, with the goal of enhancing the traditional medical system of India. Despite the acceptance of conventional medicine in many countries as a viable alternative to synthetic drugs, limitations such as insufficient public awareness and challenges in accessing source evidence have restricted its broader acceptance and usability. To address these issues, they developed an intelligent system that utilized a Raspberry Pi 3 Model B+ (RPi) and the RPi camera to capture images of Indian medicinal herbs and identify their medicinal properties. They implemented five different models for plant identification, including one known as the Herb model. This model extracted feature maps from captured medicinal leaves by combining three distinct feature extraction techniques: Scale-Invariant Feature Transform (SIFT), Oriented FAST and Rotated BRIEF (ORB), and Histogram of Oriented Gradients (HOG). Using a Support Vector Machine (SVM) as the classifier, the Herb model achieved a mean precision of 96.22% on a custom dataset of medicinal leaves, which included

40 species and 2,515 samples. Additionally, they generated a Bag of Visual Words (BoVW) by applying K-Means clustering to both SIFT and ORB descriptors to reduce dimensionality. The combined feature vector was further analysed using Random Forest and k-Nearest Neighbours (k-NN) classifiers. The effectiveness of their approach was benchmarked using the Flavia dataset with an Artificial Neural Network (ANN) as a classifier.

Naeem et al. [20] developed a machine learning-based system for classifying medicinal plant leaves. They compiled a dataset consisting of six varieties of therapeutic plant leaves from the Department of Agriculture at The Islamia University of Bahawalpur, Pakistan. The dataset was created using both multispectral and digital images captured through a computer vision laboratory setup. In the preprocessing phase, they cropped the leaf regions and converted them into grayscale format. They then employed seed intensity-based edge/line detection using the Sobel filter to delineate five regions of interest. From this process, they extracted a dataset of 65 fused features, which included a combination of texture, run-length matrix, and multispectral features. For feature optimization, they applied a chi-square feature selection method, reducing the number of features to 14. Finally, they tested five machine learning classifiers—multi-layer perceptron, logit-boost, bagging, random forest, and simple logistic—on the optimized dataset of medicinal plant leaves.

Kan et al. [21] emphasized the importance of therapeutic herbs as the primary source of Traditional Chinese Medicine (TCM), which is crucial for safeguarding human health. They highlighted that effective medicinal plant classification methodologies are vital for preserving TCM resources, authenticating TCM practices, and enhancing teaching methods for TCM identification. To address the limitations of manual classification methods, they proposed an automatic classification system based on leaf images of medicinal plants. Their approach included several key steps: first, preprocessing the leaf images; second, extracting ten shape features (SF) and five texture characteristics (TF); and finally, classifying the medicinal plant leaves using a Support Vector Machine (SVM) classifier.

Azadnia et al. [22] addressed the challenges of identifying medicinal plants, noting that the process is often time-consuming, tedious, and requires specialist expertise. To aid researchers and the general public in recognizing herb plants more quickly and accurately, they proposed an intelligent system utilizing a Convolutional Neural Network (CNN). Their deep learning model consists of two main components: a CNN block for feature extraction and a classifier block for feature classification. The classifier block includes Global Average Pooling (GAP), dense, dropout, and SoftMax layers. Their system was tested on leaf images of five different medicinal plants at three resolutions (64 × 64, 128 × 128, and 256 × 256 pixels). Using a vision-based approach, the system achieved a precision of over 99.3% across all image resolutions, demonstrating its effectiveness in plant identification.

Muneer et al.[23] noted that identifying the desired herb among thousands is an exhausting and time-consuming task. Therefore, they proposed a vision-based system for herb identification, which alleviates the need for traditional collection methods by pharmacists and botanists. Their study aimed to develop an efficient and automatic classification system for recognizing Malaysian herbs used in medical or culinary applications. Given the lack of similar studies on medicinal herbs in Malaysia, their research was pioneering in this context. In their proposed system, they explored various classifiers to create an effective model, which was then integrated into a mobile app for real-time classification. The system utilized two classifiers: Support Vector Machine (SVM) and Deep Learning Neural Network (DLNN).

Mustafa et al. [24] highlighted the challenges of manual plant identification, which is often time-consuming and requires prior knowledge of plant characteristics such as shape, Odor, and texture. To address these challenges, their research aimed to develop a computerized method for both recognizing plant species and detecting early illnesses in herbs. They created a system that integrated computer vision and an electronic nose to analyse the Odor, shape, colour, and texture of herb leaves. Their system employed a hybrid intelligent approach, combining several techniques: a fuzzy inference system, Naïve Bayes (NB), Probabilistic Neural Network (PNN), and Support Vector Machine (SVM) classifiers. These methods were applied to achieve efficient and accurate herb species recognition and early illness detection across samples of ten different herb species.

Quoc et al. [25] highlighted the increasing interest and significant efforts in the research community towards plant identification using Deep Learning (DL), which has produced many promising results and become a prominent trend in recent years. In their study, they employed Convolutional Neural Networks (CNNs) to recognize images of Vietnamese medicinal plants. They assessed various frameworks, including VGG16, ResNet50, InceptionV3, DenseNet121, Xception, and MobileNet. Among these, the Xception framework achieved the highest precision, reaching 88.26%.

Chouhan et al. [26] explored the application of Internet of Things (IoT) technology, which is crucial in global infrastructure for connecting devices to collect, contribute, analyse, and share information. They focused on leveraging this technology to enhance precision farming, specifically through automated disease detection in plant leaves. Their study introduced a novel framework called IoT_FBFN, which integrates a Fuzzy-Based Function Network (FBFN) with IoT capabilities. Their process starts with acquiring images of plant leaves, which are then pre-processed and analysed using the Scale-Invariant Feature Transform (SIFT) method to extract features. The FBFN is employed to detect galls caused by the insect Pauropsyllatuberculate. The network's training process is optimized using the Firefly algorithm, which enhances its efficiency. The IoT_FBFN framework combines the computational power of fuzzy logic with the learning adaptability of neural networks, achieving higher precision in identifying and classifying galls compared to other methods.

Mia et al. [27] emphasized the crucial role of herb leaves in treating various human illnesses worldwide and their significant impact on clinical science. However, recognizing these herbs can be challenging due to limited information, and misidentifying harmful plants as medicinal ones can pose serious health risks. This highlights the importance of accurately identifying beneficial herb leaves. To address this challenge, they proposed a neural network model using YOLO (You Only Look Once) for real-time localization and classification of herb leaves. Their study focused on five types of herbs: Mehndi, Betel, Mint, Basil, and Aloe Vera. The YOLO-based model achieved a classification precision of approximately 95%, which is promising compared to other approaches. In future work, they aim to further develop this model into a versatile application to help users learn about herbal remedies. Sharma et al. [28] developed an intelligent approach for identifying Indian medicinal plant leaves by leveraging a combination of heterogeneous features. Their method employs a dual strategy for feature extraction. The first approach involves extracting handcrafted feature descriptors, including edge histograms, oriented gradients, and binary patterns. The second approach utilizes deep features extracted through Convolutional Neural Networks (CNNs) using a transfer learning technique. To standardize the extracted features, Z-score normalization was applied. The normalized features were then combined using intermediate-level fusion techniques to create a diverse set of heterogeneous features. To tackle the challenges associated with the high dimensionality of the heterogeneous feature vector dataset, they employed Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction. Finally, the reduced feature vectors were used as inputs for various classifiers to identify and classify medicinal herbs.

TABLE II. MACHINE LEARNING MODELS

Author(s)	Methodology	Dataset	Feature Extraction	Classifier(s)	Precision
Roopashree et al. [19]	Raspberry Pi 3 Model B+ and RPi camera	Custom dataset of 40 species, 2515 samples	SIFT, ORB, HOG	SVM	96.22%
Naeem et al. [20]	ML classification using multispectral and digital image dataset	Six varieties from The Islamia University of Bahawalpur	Texture, run-length matrix, multispectral features	MLP, Logit-Boost, Bagging, Random Forest, Simple Logistic	Not specified
Kan et al. [21]	Automatic classification using SVM	Not specified	Shape and texture features	SVM	Not specified
Muneer et al. [23]	Vision system for Malaysian herbs	Not specified	Not specified	SVM, DLNN	Not specified
Mustafa et al. [24]	Hybrid intelligent system for species and illness recognition	Ten herb species	Odor, shape, colour, texture	Fuzzy Inference System, Naive Bayes, PNN, SVM	Not specified
Chouhan et al. [26]	IoT and FBFN for automated illness detection	Not specified	SIFT	FBFN	Not specified
Sharma et al. [28]	Bilateral approach using handcrafted and deep features	Indian medicinal herbs	Edge histograms, oriented gradients, binary patterns, CNN features	Various classifiers	Not specified
Borman et al. [30]	Multiclass SVM for herbal leaf identification	Not specified	First-order feature extraction (mean, skewness, variance, kurtosis, entropy)	Multiclass SVM	Not specified
Shailendra et al. [31]	IoT and ML for real-time herb identification	Custom dataset of 25 species, 1500 images	SIFT, HOG	SVM	98.98%

Hao et al. [29] addressed the challenges in classifying Chinese Herbal Medicine (CHM), a critical area of research in Intelligent Medicine. They identified two main issues: the limited size of available Chinese Herbal datasets and the limitations of traditional CHM classification models, which hindered achieving satisfactory results. To address these challenges, they developed a novel, large-scale CHM classification (CHMC) dataset, which includes 100 classes and approximately 10,000 samples. This dataset features a diverse range of medicinal materials and natural backgrounds. Building on this dataset, they proposed the EfficientNetB4 model for CHM classification. EfficientNet is designed to uniformly scale the depth, width, and resolution of the network, achieving improved precision by balancing all dimensions of the model. Borman et al. [30] investigated the use of plant leaves, including herbal leaves, as medicine, noting that many people are unfamiliar with these herbs. Their study focused on identifying different types of herbal leaves based on their shape characteristics. To achieve this, they employed first-order feature extraction combined with the Multiclass Support Vector Machine (Multiclass SVM) algorithm. The first-order feature extraction process involved capturing features using parameters such as mean, skewness, variance, kurtosis, and entropy. The Multiclass SVM algorithm was then used to classify the leaves by finding the optimal hyperplane that separates data points of different classes in a two-dimensional space. Their approach enabled effective grouping and identification of herbal leaves based on the extracted features.

Shailendra et al. [31] developed an automatic recognition model using Internet of Things (IoT) and Machine Learning (ML) techniques to classify therapeutic plants and enhance traditional medicinal systems. They proposed an intelligent system that leverages a Raspberry Pi 3 Model B+ (RPi) and its camera to capture real-time images of Indian medicinal herbs and identify their medicinal properties. In their study, they developed four ML models, including one that provides detailed information about a captured medicinal leaf via the RPi user interface. Their model achieved a top-1 precision of 98.98% on a custom dataset consisting of 1,500 leaf images from 25 different medicinal species. The model combined two feature extraction techniques—Scale Invariant Feature Transform (SIFT) and histogram of oriented gradients (HOG). Additionally, k-means clustering was applied to SIFT descriptors to generate a Bag of Visual Words for feature selection. The system's performance was evaluated using a Support Vector Machine (SVM) classifier. Pushpa et al. [32] identified limitations in current state-of-the-art methods for photo-based plant species classification, particularly in addressing the complexities of plant image recognition. To overcome these challenges, they proposed a novel, lightweight deep convolutional neural network called Ayur-PlantNet, designed specifically to classify forty Ayurvedic plant species. Ayur-PlantNet was developed from scratch and trained on 6,000 samples of segmented plant regions. In their comparative study, they evaluated Ayur-PlantNet against several pre-trained models, including ResNet34, ResNet50, VGG16, MobileNetV3_Large, EfficientNet_B4, and DenseNet121. They found that Ayur-PlantNet achieved a precision of 92.27%, while maintaining fewer trainable parameters and lower computational complexity compared to the pre-trained models. Arnandito et al. [33] conducted a study comparing the performance of EfficientNetB7 and MobileNetV2 for classifying herbal plant species using Convolutional Neural Networks (CNNs). The primary goal was to achieve high precision in automatically identifying herbal plant species. Their evaluation revealed that both EfficientNetB7 and MobileNetV2 achieved approximately 98% precision in recognizing herbal plant species. While both models excelled in precision, recall, and F1-score across most plant species, EfficientNetB7 showed a slight advantage in certain evaluation metrics. These findings provide valuable insights into the use of CNN architectures for automatic plant recognition and suggest that such models could be effectively implemented in web-based systems for herbal plant identification.

TABLE III: DEEP LEARNING MODELS

Author(s)	Methodology	Dataset	Feature Extraction	Classifier(s)	Precision
Azadnia et al. [22]	CNN for herb identification	Five medicinal plants, images at 64x64, 128x128, 256x256 pixels	CNN features	CNN	>99.3%
Quoc et al. [25]	CNN for Vietnamese medicinal plant images	Not specified	CNN features	VGG16, ResNet50, InceptionV3, DenseNet121, Xception, MobileNet	88.26% (Xception)
Mia et al. [27]	Neural network using YOLO for real-time classification	Five herb types	YOLO features	YOLO	95%
Hao et al. [29]	EfficientNetB4 for CHM classification	CHMC dataset (100 classes, ~10,000 samples)	EfficientNet features	EfficientNetB4	Not specified
Pushpa et al. [32]	Ayur-PlantNet CNN for Ayurvedic plants	40 species, 6000 samples	CNN features	Ayur-PlantNet, ResNet34, ResNet50, VGG16, MobileNetV3_Large, EfficientNet_B4, DenseNet121	92.27% (Ayur-PlantNet)
Arnandito et al. [33]	Comparison of EfficientNetB7 and MobileNetV2 for herbal plant identification	Not specified	CNN features	EfficientNetB7, MobileNetV2	~98%

B. Issues and Challenges for Medicinal Plant Classification

Several major challenges complicate medicinal plant classification. One significant issue is the lack of widespread recognition and accessibility of traditional medicine, which has hindered its applicability and acceptance despite its potential benefits. Manual identification of medicinal plants is not only time-consuming and error-prone but also involves discerning characteristic features such as shape, Odor, and texture. The task is further complicated by the diversity and high dimensionality of plant traits, including a wide range of shapes, textures, and colours. This complexity makes accurate classification difficult. Additionally, the scarcity of comprehensive datasets exacerbates these challenges, increasing the risk of overfitting and reducing model generalization.

Significant challenges arise in balancing computational efficiency with precision in deep learning (DL) models. While these models often achieve high precision, they are typically computationally intensive and thus not well-suited for real-time applications or environments with limited resources. Integrating multiple feature extraction techniques—such as texture, shape, and colour—into a single model presents its own difficulties. This integration requires meticulous optimization to fully leverage all involved features. Environmental variables, such as changes in lighting, background, and leaf orientation during image capture, can negatively impact model performance. Therefore, consistent precision demands high-level preprocessing and normalization techniques. Additionally, early diagnosis of illnesses in medicinal plants is crucial for timely intervention. However, identifying subtle symptoms that vary with conditions remains challenging. The proposed study aims to address these issues by developing and evaluating DL and state-of-the-art ML models for the automatic identification and classification of medicinal plants. The project will focus on achieving high classification precision while balancing computational needs through the use of diverse datasets and efficient storage of high-dimensional data. Automated vision-based systems will replace traditional manual methods. The inclusion of advanced preprocessing and integration of additional features, such as illness detection, will enhance model performance and facilitate early detection and prompt treatment of plant illnesses.

C. Literature Review on Ordinary Plant

Similar to the studies mentioned, numerous works have been conducted on the classification of ordinary plant leaves and plant illnesses. While traditional feature extraction and machine learning (ML) approaches are widely discussed, their performance can be limited by various factors. Consequently, deep learning (DL) systems have gained significant attention in this domain due to their ability to overcome many of these limitations.

Kaur et al. [34] investigated the application of deep learning-based Convolutional Neural Networks (CNNs) in computer vision tasks such as image classification, object detection, segmentation, and image analysis. They observed that existing models often used mixed crop approaches or single-class datasets, resulting in average precision levels. In their study, they employed a Modified InceptionResNet-V2 (MIR-V2) model, a variant of the CNN architecture, combined with a transfer learning approach to identify diseases in tomato leaf images. Their proposed model was trained on both a public dataset and a self-collected dataset, which included images of seven different tomato leaf diseases as well as healthy leaves. The performance of their model was evaluated using various parameters, including dropout rate, learning rate, batch size, number of epochs, and precision.

Sharma et al. [35] addressed the limitations of existing convolutional neural network (CNN) models, which often focus on detecting illnesses in specific crops. To overcome these limitations, they proposed a new, deeper, and lightweight CNN architecture named DLMC-Net, designed for real-time plant leaf illness detection across multiple crops. The DLMC-Net model introduced a series of collective blocks and a passage layer to enhance the extraction of deep features. This design improved feature propagation and reuse, addressing the vanishing gradient problem commonly encountered in deep networks.

Additionally, the model utilized point-wise and separable convolution blocks to reduce the number of trainable parameters, enhancing efficiency. The efficacy of the DLMC-Net model was validated using four publicly available datasets—citrus, cucumber, grapes, and tomato—demonstrating its effectiveness in illness detection across diverse crop types.

Haridasan et al. [36] proposed an automated approach for accurately detecting and classifying diseases in rice plants using photographs. Their system employs a computer vision-based methodology that integrates image processing, machine learning (ML), and deep learning (DL) techniques to improve the protection of paddy crops against common diseases such as bacterial leaf blight, false smut, brown leaf spot, rice blast, and sheath rot, which frequently affect Indian rice fields. The approach involves several stages: image preprocessing, followed by image segmentation to isolate the disease-affected sections of the rice plant. The diseases are then identified based on their visual characteristics. Their system combines a support vector machine (SVM) classifier with Convolutional Neural Networks (CNNs) to recognize and classify the specific diseases. Using ReLU and

SoftMax activation functions, the DL-based strategy achieved a validation precision of 0.9145, demonstrating its effectiveness in disease detection and classification.

Hosny et al. [37] proposed an innovative deep Convolutional Neural Network (DCNN) model that replaces traditional convolution layers with deeper separable convolutions. This modification aims to reduce the number of model parameters and iteration times. They combined deep features with Local Binary Pattern (LBP) features, which efficiently capture local texture information, thereby complementing the deep features. Their approach significantly reduced both training parameters and iteration times compared to conventional transfer learning models like AlexNet, GoogleNet, and VGG16. The integrated model, which effectively captures both deep and local spatial texture information, achieved high precision in detecting plant leaf diseases across various datasets.

Reddy et al. [38] tackled the challenges associated with manual plant disease inspections, which often suffer from low precision, time consumption, and limitations in handling multiple plant diseases. To address these issues, they developed a customized model called PDICNet for identifying and classifying plant leaf diseases. Their approach begins with using ResNet-50 to extract a range of features from plant leaf images, particularly focusing on colour and texture properties. They then employed a Modified Red Deer Optimization Algorithm (MRDOA) to select optimal features, refining them to a reduced and significant set. Subsequently, a Deep Learning Convolutional Neural Network (DLCNN) classifier was utilized to improve classification performance. The PDICNet model demonstrated exceptional results, achieving a precision of 99.73% and an F1-score of 99.78% on the Plant Village dataset, and a precision of 99.68% and an F1-score of 99.71% on the Rice Plant dataset.

Mahum et al. [39] addressed limitations in existing ML and DL techniques that typically use datasets like "The Plant Village Dataset" to classify potato leaves into only two categories. They proposed an enhanced DL algorithm designed to classify potato plants into five distinct categories: Potato Late Blight (PLB), Potato Early Blight (PEB), Potato Leaf Roll (PLR), Potato Verticillium Wilt (PVw), and Potato Healthy (PH). Their approach involved augmenting "The Plant Village" dataset, originally containing images of only two diseases (Early Blight and Late Blight) plus a Healthy class, with manually collected data for the additional categories. Dense connections with regularization further mitigated overfitting, particularly with small training sets. Their algorithm achieved a precision of 97.2% on the testing set, representing a significant advancement in identifying and categorizing multiple potato leaf diseases.

Sahu et al. [40] emphasized the critical need for early detection of plant illnesses to prevent their spread and minimize damage. They proposed a novel hybrid model combining Random Forest and Multiclass Support Vector Machine (HRF-MCSVM) for plant foliar illness detection. Their approach involves preprocessing and segmenting image features using Spatial Fuzzy C-Means to improve computational accuracy. They utilized the Plant Village dataset, which contains 54,303 images of healthy and diseased leaves. The effectiveness of their HRF-MCSVM model was evaluated through various performance metrics, including precision, F-measure, specificity, sensitivity, and recall, showcasing its potential for effective plant illness detection.

Kotwal et al. [41] highlighted the importance of automated detection for early-stage identification of plant illnesses using deep learning (DL). They proposed a hybrid model, Plant Leaf Ailment Classification (PLDC), which integrates several advanced techniques.. For feature extraction, they combined Gray Level Co-occurrence Matrix (GLCM), Scale-Invariant Feature Transform (SIFT), and a Gabor filter. The classification was performed using an Artificial Driving-EfficientNet (AD-ENet) model based on these features, leading to improved precision and effectiveness in illness detection. Diana Andrushia et al. [42] focused on detecting illnesses in grape leaves using convolutional capsule networks, an advanced deep learning architecture. Capsule networks use groups of neurons (capsules) to capture spatial relationships more effectively. Their approach involved incorporating convolutional layers before the primary capsule layer to reduce the number of capsules and speed up the dynamic routing process. Testing on both augmented and non-augmented datasets, their method achieved a high precision of 99.12% in detecting grape leaf illness.

TABLE IV: LITERATURE SURVEY

Author(s)	Methodology	Dataset	Feature Extraction	Classifier(s)	Precision
Kaur et al. [34]	Modified InceptionResNet-V2 (MIR-V2) with transfer learning to recognize tomato leaf illness	Public and self-collected dataset of tomato leaves	CNN features with dropout, learning rate, batch size, epochs	CNN	Not specified
Sharma et al. [35]	Deeper lightweight Convolutional Neural Network architecture (DLMC-Net) for plant leaf illness detection	Public datasets: citrus, cucumber, grapes, and tomato	Collective blocks and passage layer	CNN	Not specified

Haridasan et al. [36]	DL, ML, and image processing combined to identify rice plant illness	Not specified	CNN features with image pre-processing, segmentation	CNN, SVM	91.45%
Hosny et al. [37]	Local Binary Pattern (LBP) and separable convolutions in a Deep Convolutional Neural Network (DCNN)	Various plant leaf illness datasets	Deep features combined with LBP features	CNN	High precision rate
Reddy et al. [38]	Personalised PDICNet model for identifying and categorising plant leaf illness	Plant Village dataset and Rice Plant dataset	ResNet-50 features, MRDOA for feature selection	DLCNN	99.73% (PlantVillage), 99.68% (Rice Plant)
Mahum et al. [39]	Improved DL algorithm using Efficient DenseNet for potato leaf illness classification	The Plant Village dataset and manually gathered data	Efficient DenseNet with extra transition layer	CNN	97.2%
Sahu et al. [40]	For the identification of plant foliar illness, hybridised Random Forest Multiclass SVM (HRF-MCSVM)	Plant Village dataset	Preprocessing and segmentation with Spatial Fuzzy C-Means	Random Forest, SVM	Not specified
Kotwal et al. [41]	Hybrid strategy with UNet segmentation and Artificial Driving-EfficientNet (AD-ENet) for the classification of plant leaf illness	Not specified	GLCM, SIFT, Gabor filter features	EfficientNet	Not specified

D. Issues and Challenges for Ordinary Plant Classification

The intrinsic diversity in plant images, resulting from several environmental factors such as illumination, background, and leaf orientation, is a significant challenge and can negatively impact the precision of the model. Another crucial issue is the availability of comprehensive and high-quality datasets. This means that the models rely on publicly accessible datasets that may not represent the spectrum of infections or may contain insufficient representatives from all relevant classes, leading to overfitting and poor generalization. One of the more challenging tasks is balancing computing efficiency against precision. Certain DL models, like CNNs, are not suitable for real-time or resource-constrained applications as they require high levels of computational power. Additionally, the procedure for feature extraction and selection is further complicated due to the high dimensionality of plant features, including shape, texture, and colour. To tackle this complexity, elaborate methodologies are often needed. These also raise questions about model interpretability. While precision in DL models is expected to be high, interpreting their decision-making process—an important factor in gaining end-user trust and confidence, especially within agriculture—becomes difficult due to the "black box" nature of their design. Another challenge lies in combining different methods of feature extraction within one coherent model, as it requires careful optimization to fully leverage information about colour, shape, and texture. Although challenging, diagnosing illness in its early stage is very important since the symptoms are subtle and diverse. This necessitates effective preprocessing and normalization techniques to maintain model performance across different settings and datasets. Additionally, since most models currently in use are trained for specific crops or ailments, developing models capable of detecting numerous illnesses across various plant species adds another layer of complexity. Moreover, while most proposed models show very high precision on test datasets, it remains an open question whether they would be effective in real-world situations. Practical challenges to overcome include inconsistent image quality, multiple illnesses on one leaf, and immediate processing requirements in the field. Furthermore, implementing such models in real-world scenarios, especially in developing countries, must consider infrastructure and resource availability. This often precludes the use of computationally expensive solutions.

III. CONCLUSION

The evolution of DL approaches has shown promising results in improving the precision and effectiveness of medicinal plant detection and classification. This review highlights the broad range of techniques—from traditional ML methods to advanced DL models using CNNs and hybrid systems. Each approach has its own benefits in their respective fields. However, several important issues remain unresolved. The significant impact of environmental conditions and the variability in dataset quality on model performance underscore the need for more comprehensive and diverse datasets. Practical limitations also arise from the high computational requirements of complex DL models, particularly in real-time applications and low-resource settings.

The interpretability and acceptability of DL models are often challenged by their complexity and "black box" nature. These issues can only be addressed through ongoing research aimed at enhancing the resilience, interpretability, and efficiency of the models. For broader real-world applicability, models should be able to identify multiple plant species universally. Future studies should focus on developing models that balance interpretability, computational efficiency, and precision, informed by recent advancements in transfer learning and hybrid modelling. Additionally, significant efforts are needed to generate higher-quality and more comprehensive datasets that represent the full spectrum of plant species across various environmental settings.

REFERENCES

[1] Jamshidi-Kia, F., Lorigooini, Z., & Amini-Khoei, H. (2017). Medicinal plants: Past history and future perspective. Journal of herbmed pharmacology, 7(1), 1-7.

[2] Goswami, A., Malik, M., Malik, A., Kumar, A., & Chaudhary, S. (2020). Andrographis paniculata (Kalmegh): A critical study on medicinal effects. Annals of Horticulture, 13(2), 115-123.

[3] Tiwari, P., Mishra, B. N., & Sangwan, N. S. (2014). Phytochemical and pharmacological properties of Gymnema sylvestre: an important medicinal plant. BioMed research international, 2014(1), 830285.

[4] Wylie, M. R., & Merrell, D. S. (2022). The antimicrobial potential of the neem tree Azadirachta indica. Frontiers in pharmacology, 13, 891535.

[5] Anand, U., Jacobo-Herrera, N., Altemimi, A., & Lakhssassi, N. (2019). A comprehensive review on medicinal plants as antimicrobial therapeutics: potential avenues of biocompatible drug discovery. Metabolites, 9(11), 258.

[6] Süntar, I. (2020). Importance of ethnopharmacological studies in drug discovery: role of medicinal plants. Phytochemistry Reviews, 19(5), 1199-1209.

[7] D'Cruz, M., & Andrade, C. (2022). Potential clinical applications of Ashwagandha (Withania somnifera) in medicine and neuropsychiatry. Expert Review of Clinical Pharmacology, 15(9), 1067-1080.

[8] Syed, A. A., Reza, M. I., Singh, P., Thombre, G. K., & Gayen, J. R. (2021). Withania Somnifera in neurological disorders: Ethnopharmacological evidence, mechanism of action and its progress in delivery systems. Current drug metabolism, 22(7), 561-571.

[9] Sachar, S., & Kumar, A. (2021). Survey of feature extraction and classification techniques to identify plant through leaves. Expert Systems with Applications, 167, 114181.

[10] Kavitha, S., Kumar, T. S., Naresh, E., Kalmani, V. H., Bamane, K. D., & Pareek, P. K. (2023). Medicinal Plant Identification in Real-Time Using Deep Learning Model. SN Computer Science, 5(1), 73.

[11] Chen, S. L., Yu, H., Luo, H. M., Wu, Q., Li, C. F., & Steinmetz, A. (2016). Conservation and sustainable use of medicinal plants: problems, progress, and prospects. Chinese medicine, 11, 1-10.

[12] Kumar, A., Ramchiary, N., Singh, P. J. S. (2021). Role of traditional ethnobotanical knowledge and indigenous communities in achieving sustainable development goals. Sustainability 13 (6), 3062. doi: 10.3390/su13063062

[13] Murphy, G. E., & Romanuk, T. N. (2014). A meta-analysis of declines in local species richness from human disturbances. Ecology and evolution, 4(1), 91-103.

[14] Azlah, M.A.F.; Chua, L.S.; Rahmad, F.R.; Abdullah, F.I.; Wan Alwi, S.R. Review on techniques for plant leaf classification and recognition. Computers 2019, 8, 77.

[15] Tripathi, K.; Khan, F.A.; Khanday, A.M.U.D.; Nisa, K.U. The classification of medical and botanical data through majority voting using artificial neural network. Int. J. Inf. Technol. 2023, 15, 3271–3283.

[16] Wäldchen, J.; Mäder, P. Plant species identification using computer vision techniques: A systematic literature review. Arch. Comput. Methods Eng. 2018, 25, 507–543.

[17] De Carvalho, M.R.; Bockmann, F.A.; Amorim, D.S.; Brandão, C.R.F.; de Vivo, M.; de Figueiredo, J.L.; Britski, H.A.; de Pinna, M.C.; Menezes, N.A.; Marques, F.P.; et al. Taxonomic impediment or impediment to taxonomy? A commentary on systematics and the cybertaxonomic-automation paradigm. Evol. Biol. 2007, 34, 140–143.

[18] Rabani, S.T.; Khanday, A.M.U.D.; Khan, Q.R.; Hajam, U.A.; Imran, A.S.; Kastrati, Z. Detecting suicidality on social media: Machine learning at rescue. Egypt. Inform. J. 2023, 24, 291–302.

[19] Roopashree, S., Anitha, J., Mahesh, T. R., Kumar, V. V., Viriyasitavat, W., & Kaur, A. (2022). An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach. Measurement, 200, 111484.

[20] Naeem, S., Ali, A., Chesneau, C., Tahir, M. H., Jamal, F., Sherwani, R. A. K., & Ul Hassan, M. (2021). The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach. Agronomy, 11(2), 263.

[21] Kan, H. X., Jin, L., & Zhou, F. L. (2017). Classification of medicinal plant leaf image based on multi-feature extraction. Pattern recognition and image analysis, 27, 581-587.

[22] 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.

[23] Muneer, A., & Fati, S. M. (2020). Efficient and automated herbs classification approach based on shape and texture features using deep learning. IEEE Access, 8, 196747-196764.

[24] Mustafa, M. S., Husin, Z., Tan, W. K., Mavi, M. F., & Farook, R. S. M. (2020). Development of automated hybrid intelligent system for herbs plant classification and early herbs plant disease detection. Neural Computing and Applications, 32, 11419-11441.
[25] Quoc, T. N., & Hoang, V. T. (2020, October). Medicinal Plant identification in the wild by using CNN. In 2020 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 25-29). IEEE.
[26] Chouhan, S. S., Singh, U. P., & Jain, S. (2021). Automated plant leaf disease detection and classification using fuzzy based function network. Wireless Personal Communications, 121(3), 1757-1779.
[27] Mia, J., Bijoy, H. I., Uddin, S., & Raza, D. M. (2021, July). Real-time herb leaves localization and classification using YOLO. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE.
[28] Sharma, M., Kumar, N., Sharma, S., Kumar, S., Singh, S., & Mehandia, S. (2024). Medicinal plants recognition using heterogeneous leaf features: an intelligent approach. Multimedia Tools and Applications, 83(17), 51513-51540.
[29] Hao, W., Han, M., Yang, H., Hao, F., & Li, F. (2021). A novel Chinese herbal medicine classification approach based on EfficientNet. Systems Science & Control Engineering, 9(1), 304-313.
[30] Borman, R. I., Rossi, F., Jusman, Y., Rahni, A. A. A., Putra, S. D., & Herdiansah, A. (2021, October). Identification of Herbal Leaf Types Based on Their Image Using First Order Feature Extraction and Multiclass SVM Algorithm. In 2021 1st international conference on electronic and electrical engineering and intelligent system (ICE3IS) (pp. 12-17). IEEE.
[31] Shailendra, R., Jayapalan, A., Velayutham, S., Baladhandapani, A., Srivastava, A., Kumar Gupta, S., & Kumar, M. (2022). An IoT and machine learning based intelligent system for the classification of therapeutic plants. Neural Processing Letters, 54(5), 4465-4493.
[32] Pushpa, B. R., & Rani, N. S. (2023). 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, 34, 100459.
[33] Arnandito, S., & Sasongko, T. B. (2024). Comparison of EfficientNetB7 and MobileNetV2 in Herbal Plant Species Classification Using Convolutional Neural Networks. Journal of Applied Informatics and Computing, 8(1), 176-185.
[34] Kaur, P., Harnal, S., Gautam, V., Singh, M. P., & Singh, S. P. (2023). A novel transfer deep learning method for detection and classification of plant leaf disease. Journal of Ambient Intelligence and Humanized Computing, 14(9), 12407-12424.
[35] Sharma, V., Tripathi, A. K., & Mittal, H. (2023). DLMC-Net: Deeper lightweight multi-class classification model for plant leaf disease detection. Ecological informatics, 75, 102025.
[36] Haridasan, A., Thomas, J., & Raj, E. D. (2023). Deep learning system for paddy plant disease detection and classification. Environmental monitoring and assessment, 195(1), 120.
[37] Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E., & Papakostas, G. A. (2023). Multi-class classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern. IEEE Access, 11, 62307-62317.
[38] Reddy, S. R., Varma, G. S., & Davuluri, R. L. (2023). Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification. Computers and Electrical Engineering, 105, 108492.
[39] Mahum, R., Munir, H., Mughal, Z. U. N., Awais, M., Sher Khan, F., Saqlain, M., ... & Tlili, I. (2023). A novel framework for potato leaf disease detection using an efficient deep learning model. Human and Ecological Risk Assessment: An International Journal, 29(2), 303-326.
[40] Sahu, S. K., & Pandey, M. (2023). An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model. Expert systems with applications, 214, 118989.
[41] Kotwal, J. G., Kashyap, R., & Shafi, P. M. (2024). Artificial driving based Efficient Net for automatic plant leaf disease classification. Multimedia Tools and Applications, 83(13), 38209-38240.
[42] Diana Andrushia, A., Mary Neebha, T., Trephena Patricia, A., Umadevi, S., Anand, N., & Varshney, A. (2023). Image-based disease classification in grape leaves using convolutional capsule network. Soft Computing, 27(3), 1457-1470.