FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 9, September, 2025, pp 190 – 193
DOI: https://doi.org/10.33003/fjs-2025-0909-4029

AI TECHNIQUES FOR IDENTIFICATION AND STUDY OF MEDICINAL PLANTS: A REVIEW

***[1]Bashiru Aliyu Sani, [2]Mansur Lawal, [1]Nasiru Aliyu Jeka, [1]Asmau Usman and [1]Bashir Ahmed**

[1]Department of Computer Science, Abdu Gusau Polytechnic, Talata Mafara – Zamfara, Nigeria.
[2]Department of Pharmacognosy and Ethno-pharmacy, Usmanu Danfodiyo University, Sokoto, Nigeria.

*Corresponding authors' email: baliyusani@gmail.com Phone: +2348062477796

ABSTRACT

Medicinal plants have been integral to human health for centuries, offering a wealth of bioactive compounds with therapeutic potential. However, their identification and study pose significant challenges due to the vast number of species, morphological similarities, and the need for expert knowledge. Traditional methods are time-consuming and often require specialized skills. With the advent of artificial intelligence (AI), particularly machine learning, there is a growing opportunity to streamline and enhance the processes involved in medicinal plant research. This review explores the application of AI techniques, focusing on machine learning and deep learning, in the identification and study of medicinal plants. By synthesizing recent research, this paper highlights how AI can address key challenges in this field, the combination of machine learning algorithms and multi-source data analysis facilitates a comprehensive analysis and aids in the effective evaluation of the quality of medicinal plants. Deep learning and Convolutional Neural Networks (CNNs); Product Decision Rule (PDR); EfficientNet-B1-based deep learning model; Direct Ensemble Classifier for Imbalanced Multiclass Learning (DECIML) are among the prominent machine learning tools used to identify medicinal plants based on their leaf textural features in an ensemble manner are used to compare their performance accuracies over this data. Also, Artificial Neural Networks, Deep Neural Networks, Neuro-fuzzy Logic have eased the time required in classical experimental strategy and paved a pathway for understanding from accurate species recognition to predicting bioactive compound biosynthesis, thereby paving the way for more efficient drug discovery and conservation efforts.

Keywords: Artificial Intelligence (AI), Medicinal Plants, Identification

INTRODUCTION

Medicinal plants have been a cornerstone of human healthcare for centuries, providing a rich source of bioactive compounds for therapeutic use (Singh *et al.,* 2021). The global attention on traditional remedies during the COVID-19 pandemic has further emphasized their value (Paulson & Ravishankar, 2020). However, identifying medicinal plants and studying their pharmacological properties present significant challenges. Traditional identification methods rely heavily on expert knowledge and are labor-intensive, while pharmacological screening is costly and time-consuming (Singh *et al*., 2021). Additionally, the increasing demand for medicinal plants has led to fraudulent practices like adulteration, which threatens consumer safety (Paulson & Ravishankar, 2020).

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a transformative tool to address these challenges. ML algorithms can process vast datasets, including images and chemical profiles, to create comprehensive resource databases for medicinal plants (Zhang & Wang, 2023). DL, especially convolutional neural networks (CNNs), excels in image-based identification tasks, enabling rapid and accurate species recognition (Azadnia *et al.,* 2022). This review synthesizes recent research to explore AI applications in medicinal plant identification and pharmacological studies, highlighting their potential to revolutionize the field.

.

MATERIALS AND METHODS

Data Acquisition and Query Criteria

Published peer-reviewed articles on phytochemistry were retrieved from the Dimensions™ database (PubMed and AJOL inclusive) on Mar 24, 2024. The database houses millions of research works associated with over 1.6 billion citations, grants, clinical trials, datasets, policy documents and patents. It is the most comprehensive research grants database which links grants to millions of resulting publications, clinical trials and patents (Hook *et al*., 2018). We used the search term "Artificial intelligence AND Medicinal Plants Identification" to identify primary research articles published between 2013 and 2023. All available information was retrieved. To obtain subject-specific results and for the sake of accuracy, article titles and abstracts were searched using full counting method.

The below diagram is a procedures that achieve the aim of developing an AI-based techniques for plant identification using machine algorithm. The methodology is a steps taken from data collection to continuous improvement, providing a structured framework for the whole work

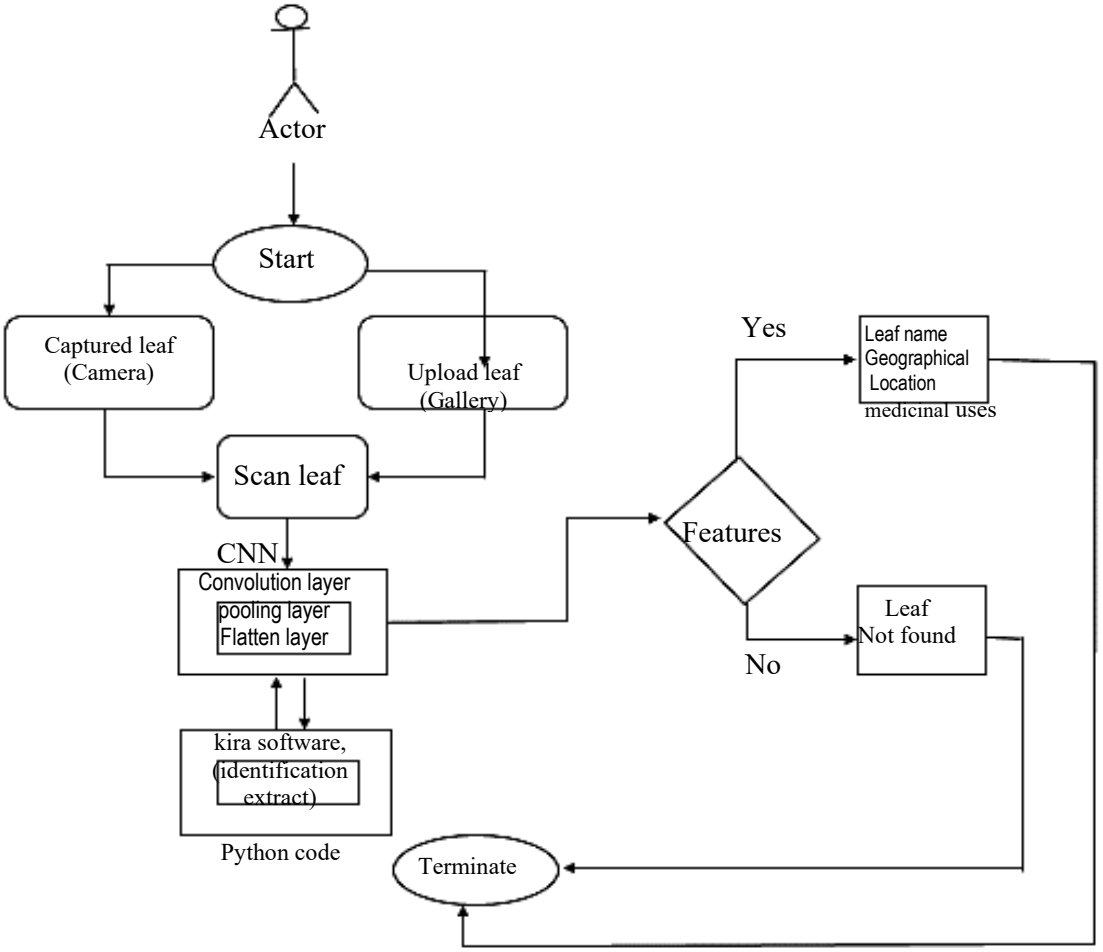

Figure 1: Architecture of AI Medicinal Plant

Data Processing and Analysis

The review focuses on addressing the limitations of traditional methods, such as the time-intensive nature of plant identification and the complexity of pharmacological screening. Key aspects analyzed include:

i. AI Models: Types of algorithms used, such as CNNs, VGG16, ResNet50, and neuro-fuzzy logic.
ii. Datasets: Characteristics of datasets, including image-based, chemical, and genetic data, and their regional specificity.
iii. Methodologies: Approaches to data preprocessing, feature extraction, and model training.
iv. Outcomes: Accuracy, efficiency, and practical applicability of AI models in real-world settings.

The synthesis of these studies provides a comprehensive overview of AI's role in medicinal plant research, identifying effective methodologies and areas for improvement.

RESULTS AND DISCUSSION

Deep Learning for Medicinal Plant Identification

Deep learning, particularly convolutional neural networks (CNNs), has shown remarkable success in identifying medicinal plants based on visual features, such as leaf images. Paulson and Ravishankar (2020) developed a system using pre-trained CNN models (VGG16, VGG19, and a custom CNN) to identify 64 types of indigenous ayurvedic medicinal plants. Their dataset comprised 64,000 leaf images, split into 80% training, 10% validation, and 10% testing sets. The VGG16 model achieved the highest accuracy at 97.8%, followed by VGG19 at 97.6% and the custom CNN at 95.7% (Paulson & Ravishankar, 2020).

Azadnia *et al.* (2022) proposed a CNN model with global average pooling for identifying five medicinal plant species. The model was tested across different image resolutions (64x64, 128x128, and 256x256 pixels), achieving up to 99% accuracy. This robustness across varying image qualities highlights the adaptability of CNNs in diverse settings.

Oppong *et al*. (2022) introduced a novel approach combining Log-Gabor filters with deep learning algorithms to enhance feature extraction from leaf images. This method improved classification accuracy by addressing morphological similarities among plant species, which is a common challenge in plant identification.

Malik *et al*. (2022) applied deep learning models for real-time identification of medicinal plants in the natural environment of the Borneo region. Their study demonstrated the feasibility of AI in field settings, achieving high accuracy and supporting conservation efforts through rapid species identification.

Herdiyeni *et al*. (2013) developed a computer-aided system for tropical leaf medicinal plant identification, likely using image processing and early machine learning techniques. While not explicitly deep learning-based, their work laid the groundwork for subsequent AI-driven approaches.

Machine Learning for Pharmacological Studies

Machine learning has significantly advanced pharmacological research on medicinal plants by predicting bioactive compound biosynthesis and facilitating drug discovery. Singh et al. (2021) reviewed computational approaches, including molecular docking, molecular dynamics (MD) simulations, pharmacophore modeling, and quantitative structure-activity relationship (QSAR) analysis. These methods have reduced the time required for traditional experimental strategies, enabling researchers to identify potential drug candidates and understand their mechanisms of action.

García-Pérez *et al.* (2021) combined untargeted metabolomics with machine learning to predict the biosynthesis of phenolic compounds in Bryophyllum plants (genus Kalanchoe). Their approach integrated complex metabolic data, providing insights into plant metabolism and identifying novel bioactive compounds with therapeutic potential.

Data Sources and Datasets

The success of AI models in medicinal plant research relies heavily on the quality and diversity of datasets. Zhang and Wang (2023) emphasized the importance of multi-source data integration, combining images, chemical profiles, and genetic information to enhance analysis. This approach allows for a more comprehensive understanding of plant characteristics and their therapeutic potential.

Sainin and Alfred (2014) and Sainin *et al.* (2014) focused on feature selection for Malaysian medicinal plant leaf shape identification, highlighting the need for region-specific datasets that capture local biodiversity. Their work underscores the importance of tailoring datasets to specific ecological and cultural contexts.

Table 1 summarizes dataset characteristics from a systematic review by Mulugeta *et al*. (2024), which provides insights into dataset usage in medicinal plant research

Table 1: Datasets Characteristics

Aspect	Details	Exact Numbers
Dataset Usage	67.7% used private datasets, 16.1% used public datasets, 16.1% combined both	67.7% (n=21), 16.1% (n=5), 16.1% (n=5)
Image Quantity	54.8% used >5000 images, 45.1% used <5000 images	54.8% (n=17), 45.1% (n=14)
Species Analyzed	77.4% analyzed >10 species	77.4% (n=24)
Preprocessing and Augmentation	74.1% used image preprocessing, 38.7% used data augmentation	74.1% (n=23), 38.7% (n=12)

This table highlights the reliance on private datasets and the need for large, diverse image sets to achieve high accuracy in AI models

Applications and Case Studies

Several case studies illustrate the practical applications of AI in medicinal plant research:

i. Borneo Region Field Identification: Malik *et al*. (2022) demonstrated the use of deep learning for real-time identification of medicinal plants in Borneo's natural environment, supporting conservation and sustainable harvesting.
ii. Biosynthesis Prediction: García-Pérez *et al*. (2021) used machine learning to predict phenolic compound biosynthesis in Bryophyllum plants, advancing the discovery of new bioactive compounds.
iii. Malaysian Plant Identification: Sainin and Alfred (2014) and Sainin *et al*. (2014) applied feature selection techniques to identify Malaysian medicinal plants, emphasizing the importance of region-specific datasets.

Discussion

The reviewed studies demonstrate the transformative potential of AI in medicinal plant research. Deep learning techniques, particularly CNNs, have achieved high accuracy in identifying medicinal plants, with models like VGG16 and custom CNNs reaching up to 99% accuracy (Paulson & Ravishankar, 2020; Azadnia *et al*., 2022). These advancements are critical for overcoming the limitations of traditional identification methods, which are often slow and require expert knowledge. The ability of AI to process large volumes of image data and learn complex patterns makes it particularly suited for distinguishing morphologically similar species (Oppong *et al*., 2022).

In pharmacological studies, machine learning has accelerated drug discovery by predicting bioactive compound biosynthesis and elucidating plant metabolism (Singh *et al*., 2021; García-Pérez *et al*., 2021). These approaches reduce the time and cost associated with traditional screening methods, offering a pathway for developing new therapeutic agents from medicinal plants.

However, several challenges remain. The heavy reliance on private datasets, as shown in Table 1, limits the generalizability and reproducibility of AI models (Mulugeta *et al*., 2024). Public, standardized datasets are needed to facilitate broader research and validation. Additionally, the black-box nature of many deep learning models raises concerns about interpretability, particularly in critical applications like drug discovery (Anchitaalagammai *et al*., 2021). This lack of transparency can hinder trust among researchers and practitioners.

CONCLUSION

AI techniques, particularly machine learning and deep learning, are revolutionizing the identification and study of medicinal plants. By enabling rapid and accurate species recognition and facilitating pharmacological analysis, AI addresses key challenges in traditional methods. However, issues such as data scarcity and model interpretability must be addressed to fully realize AI's potential. By leveraging standardized datasets and developing interpretable models, researchers can unlock new opportunities for drug discovery and the sustainable utilization of medicinal plants, ultimately benefiting global health and biodiversity conservation.

Future research should focus on:

i. Public Datasets: Developing globally accessible, high-quality datasets to support AI model development and validation.
ii. Interpretable Models: Designing AI models that provide transparent decision-making processes to enhance trust and usability.
iii. Technological Integration: Combining AI with technologies like drones for field data collection to improve data acquisition in remote areas (Paulson & Ravishankar, 2020).

iv. Cross-Regional Studies: Conducting studies that account for regional variations in plant characteristics to improve model generalizability.

REFERENCES

Anchitaalagammai, J. V., Shantha Lakshmi Revathy J. S., Kavitha S. and Murali S. (2021). Factors influencing the use of deep learning for medicinal plants recognition. *Journal of Physics*: Conference Series, 2089(1), 012055. https://doi.org/10.1088/1742-6596/2089/1/012055

Azadnia, R., Al-Amidi M. M., Mohammadi H. Cicif M. A., Daryab A. and 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

García-Pérez, P., Zhang L., Miras-Moreno B., Lozano-Milo E., Mariana L., Lucini L., Gallego P. P. (2021). The combination of untargeted metabolomics and machine learning predicts the biosynthesis of phenolic compounds in Bryophyllum medicinal plants *(Genus Kalanchoe)*. *Plants*, 10(11), 2430. https://doi.org/10.3390/plants10112430

Herdiyeni, Y., Nurfadhilah E., Zuhud A. E. M., Damayanti E., Arai K., Okumura H. (2013). A computer aided system for tropical leaf medicinal plant identification. *International Journal on Advanced Science, Engineering and Information Technology*, 3(1), 23-27. http://dx.doi.org/10.18517/ijaseit.3.1.23

Hook D. W., Porter S. J., Herzog C. (2018) Dimensions: building context for search and evaluation. *Front Res Metr Anal*.;3:23. doi: 10.3389/frma.2018.00023.

Malik, O. A., Ismail N., Hussein B. R., Yahya U. (2022). Automated real-time identification of medicinal plants species in natural environment using deep learning models—A case study from Borneo region. *Plants,* 11(15), 1952. https://doi.org/10.3390/plants11151952

Mulugeta, A. K., Sharma, D. P., & Mesfin, A. H. (2024). Deep learning for medicinal plant species classification and recognition: a systematic review. *Frontiers in Plant Science*, 14, 1286088. https://doi.org/10.3389/fpls.2023.1286088

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, 1-21. https://doi.org/10.1155/2022/1189509

Paulson, A., & Ravishankar, S. (2020). *AI based indigenous medicinal plant identification*. In 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA) (pp. 57-63). IEEE. https://doi.org/10.1109/ACCTHPA49271.2020.9213224

Sainin, M., & Alfred, R. (2014). Feature selection for Malaysian medicinal plant leaf shape identification and classification. In *International Conference on Computational Science and Technology* (ICCST 2014).

Sainin, M. S., Ghazali T. K., and Alfred R. (2014). *Malaysian medicinal plant leaf shape identification and classification*.

Singh, H., Bharadjava N. (2021). Treasuring the computational approach in medicinal plant research. *Progress in Biophysics and Molecular Biology,* 164, 19-32. https://doi.org/10.1016/j.pbiomolbio.2021.07.004

Zhang, Y., & Wang, Y. (2023). Recent trends of machine learning applied to multi-source data of medicinal plants. *Journal of Pharmaceutical Analysis*. https://doi.org/10.1016/j.jpha.2023.03.001