A

Project Report

On

Empowering Traditional Medicine: AI for Indian Medicinal Leaf Recognition

Submitted as partial fulfilment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

In

Computer Science & Engineering

By

Deepanshu Prajapati (2100290100050)
Garvit Agarwal (2100290100063)
Amandeep Singh Narang (2100290100025)

Under the supervision of

Ms. Himanshi Chaudhary

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May, 2025

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or the other institute of higher learning, except where due acknowledgement has been made in the text.

Signature:

Name – Deepanshu Prajapati

Roll No. – 2100290100050

Date:

Signature:

Name – Garvit Agarwal

Roll No. – 2100290100063

Date:

Signature:

Name – Amandeep Singh Narang

Roll No. – 2100290100025

Date:

A Technical Campus approved by AICTE & Affiliated to Dr. A.P.J. Abdul Kalam Technical University, Lucknow

CERTIFICATE

This is to certify that Project Report entitled which is **Empowering Traditional Medicine: AI for Indian Medicinal Leaf Recognition** submitted by **Mr. Deepanshu Prajapati, Mr. Garvit Agarwal and Mr. Amandeep Singh Narang** in partial fulfilment of the requirement for the award of degree B.Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Ms. Himanshi Chaudhary
Assistant Professor, CSE
Date:

Dr. Vineet Sharma
Dean, CSE

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B.Tech. Final Year. We owe special debt of gratitude to **Ms. Himanshi Chaudhary**, Assistant Professor, Department of Computer Science & Engineering, KIET Group of Institutions, Ghaziabad, for her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavours have seen light of the day.

We also take the opportunity to acknowledge the contribution of **Dr. Vineet Sharma,** Dean, Computer Science and Engineering Department, KIET Group of Institutions, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date:

Signature:

Name – Deepanshu Prajapati

Roll No. – 2100290100050

Date:

Signature:

Name – Garvit Agarwal

Roll No. – 2100290100063

Date:

Signature:

Name – Amandeep Singh Narang

Roll No. – 2100290100025

ABSTRACT

The identification and classification of Indian medicinal plants represent a critical component in the fields of botany, ethnopharmacology, and traditional medicine. With India’s rich biodiversity and extensive history of herbal remedies documented in ancient texts such as Ayurveda and Siddha, there is a growing need for reliable, scalable, and efficient tools to accurately identify and catalog medicinal flora. Manual identification methods, which rely heavily on domain expertise and visual inspection of morphological features, are often labor-intensive, time-consuming, and susceptible to misclassification, especially when dealing with closely related species or degraded samples. In this context, the integration of artificial intelligence (AI) and computer vision technologies presents a transformative opportunity to enhance the accuracy, speed, and accessibility of plant identification systems.

This research seeks to address the limitations of traditional plant identification by implementing and evaluating advanced deep learning techniques. Specifically, the study explores the potential of state-of-the-art convolutional neural networks (CNNs) in classifying medicinal plant species based on leaf imagery. A robust dataset comprising 6,900 high-resolution images representing 80 distinct classes of Indian medicinal plants was utilized. These images, sourced in varying environmental conditions with diverse backgrounds, mimic real-world use cases and thus enhance the generalizability of the models. The dataset was strategically divided into training, validation, and test sets, and further enriched through data augmentation techniques such as random rotation, horizontal and vertical flipping, scaling, and brightness variation. These augmentations were designed to improve the models' ability to generalize and to mitigate overfitting, especially given the natural variability in leaf appearance due to factors such as age, lighting, and environmental exposure.

In the methodological framework, multiple deep learning architectures were tested, with a particular focus on EfficientNetV2 and MobileNetV2, which are known for their balance between performance and computational efficiency. The models were trained using categorical cross-entropy as the loss function and optimized using adaptive learning rate schedules. A comprehensive evaluation was performed using standard metrics including accuracy, precision, recall, and the F1 score, offering a multi-dimensional understanding of

model performance. Among the evaluated models, EfficientNetV2 emerged as the most accurate and robust, achieving a test accuracy of 96.5%, precision of 0.9708, recall of 0.9660, and F1 score of 0.9666. These results demonstrate its exceptional ability to distinguish between subtle visual patterns present in the dataset. On the other hand, while MobileNetV2 achieved a slightly lower accuracy of 94.39%, it showed notable advantages in terms of speed and lightweight architecture, making it particularly suitable for deployment in low-resource environments such as smartphones or embedded devices.

The outcomes of this study affirm that deep learning can significantly enhance the process of medicinal plant classification, thereby supporting efforts in biodiversity conservation, herbal drug development, and the preservation of traditional knowledge. Furthermore, these models have potential applications in educational tools for botany students, mobile plant identification apps for field botanists, and even integration with smart farming systems. The success of EfficientNetV2 highlights the importance of utilizing deeper and more optimized architectures when high classification accuracy is paramount. Meanwhile, MobileNetV2 demonstrates that even lightweight models can deliver respectable performance, which is essential for real-time, offline usage.

In conclusion, the findings of this research underscore the transformative potential of AI-driven solutions in the domain of medicinal plant identification. Not only does this study validate the effectiveness of current CNN-based models, but it also lays the groundwork for future developments such as ensemble modeling, attention-based mechanisms, and hybrid systems that incorporate both image and textual metadata for richer classification capabilities. Further enhancements could involve expanding the dataset to include more plant organs (e.g., flowers, stems, fruits), building multilingual interfaces for community access, and developing end-to-end mobile platforms to bring this technology to practitioners and researchers across India and beyond. This convergence of traditional botanical knowledge with modern deep learning techniques holds immense promise for science, healthcare, and sustainable living.

TABLE OF CONTENTS

	Page No.
DECLARATION...	**ii**
CERTIFICATE...	**iii**
ACKNOWLEDGEMENT..	**iv**
ABSTRACT...	**v**
LIST OF FIGURES..	**x**
LIST OF TABLES……………………………………………………………..	**xi**
LIST OF ABBREVIATIONS..	**Xii**
CHAPTER 1: INTRODUCTION	**1-5**
1.1 Background	1
1.2 Motivation and importance	2
1.3 Problem statement ………………………………………………	2-3
1.4 Project objective …………………………………………………	3
1.5 Scope of project …………………………………………………	4
1.6 Key contributions ………………………………………………..	4-5
CHAPTER 2 LITERATURE REVIEW	**6-10**
2.1 Introduction..	6-7
2.2 Deep learning approach towards plant…………………..	7-8
2.3 Real-time identity issues..	8-9
2.4 Research gaps and future research directions………	9-10
2.5 Conclusion..	10
CHAPTER 3 PRINCIPAL MODEL AND METHODOLOGY	**11-32**
3.1 Dataset Description..	11-13
3.2 Data Partitioning..	13-17
3.3 Data Augmentation..	17-20

Page No.

- 3.3.1 Augmentation Pipeline Overview………………………………..... 17-18
- 3.3.2 Transformation Techniques and Their Impact……………………... 18-19
- 3.3.3Benefits of Data Augmentation in Fine-Grained Classification……. 19-20
- 3.3.4 Real-Time vs. Static Augmentation……………………………….. 20
- 3.3.5 Visualization………………………………………………….... 20

3.4 Model Architectures... 20-28

- 3.4.1 MobileNetV2... 20-24
- 3.4.2 EfficientNetV2...,,,,........................ 24-28

3.5 Training and Evaluation……………………………………………….. 28-32

- 3.5.1 Dataset Splitting and Preprocessing……………………………….. 29
- 3.5.2 Model Compilation and Optimization……………………………... 29-30
- 3.5.3 Training Strategy and Monitoring…………………………………. 30
- 3.5.4 Evaluation Metrics………………………………………………… 30-31
- 3.5.5 Observations…………………………………………………….. 31-32

CHAPTER 4 RESULTS……………………………………………………….. 33-43

4.1 Model Performance……………………………………………………… 33-37

- 4.1.1 EfficientNetV2 Performance……………………………………… 33-34
- 4.1.2 MobileNetV2 Performance………………………………………... 34-35
- 4.1.3 Comparative Insights……………………………………………. 35-36
- 4.1.4 Application Relevance…………………………………………….. 36
- 4.1.5 Summary………………………………………………………... 36-37

4.2 Comparative Analysis…………………………………………………. **37-40**

- 4.2.1 Superior Performance of EfficientNetV2………………………….. 37
- 4.2.2 MobileNetV2 as a Resource-Efficient Contender…………………. 37-38

Page No.

- 4.2.3 Interpretability via the Confusion Matrix………………………….. 38
- 4.2.4 Application-Specific Suitability…………………………………… 38-39
- 4.2.5 Summary of Comparative Insights………………………………… 39-40

4.3 Visualization…………………………………………………………... 40-43

CHAPTER 5 CONCLUSION……………………………………………………. 44-51

5.1 Summary of Findings…………………………………………………... 44-46

5.2 Implications and Contributions……………………………………… 46-48

5.3 Future Work………………………………………………………………. 48-50

5.4 Final Conclusion……………………………………………………….. 50-51

REFERENCES ……………………………………………………………… 52-54

APPENDIX ………………………………………………………………….. 55-62

Originality Report…………………………………………………………... 55-60

Proof of Acceptance………………………………………………………… 61-62

LIST OF FIGURES

Figure No.	Description	Page No.
1.	Architecture of the Deep Learning- Based Plant Identification System	5
2.	Sample Image from the Indian Leaves Dataset	13
3.	Data Augmentation Pipeline for Training Set	18
4.	Architecture of the MobileNetV2-Based Classification Model	24
5.	Architecture of the Efficient NetV2- Based Classification Model	28
6.	Preprocessing of Data	32
7.	Confusion Matrix of Medical Plant Identification Model	39
8.	Accuracy and Loss Plot for Efficient	42
9.	Model Accuracy plot for MobileNetV2 Model	43
10.	Model Predictions and Confidence Scores for Medicinal Leaf Classification	46

LIST OF TABLES

Table No.	Description	Page No.
1.	Performance of EfficientNet V2	27
2.	Summary of Deep Learning Models Used	35
3.	Implications of Findings	48
4.	Future Research Directions	49-50

LIST OF ABBREVIATIONS

Abbreviation	Full Form
NAM	Network Animator
CNN	Convolutional Neural Network
ML	Machine Learning
DL	Deep Learning
HEOR	Health Economics and Outcomes Research
GIS	Geographic Information Systems
KPI	Key Performance Indicator
NLP	Natural Language Processing
GAN	Generative Adversarial Network
VAE	Variational Autoencoder
RL	Reinforcement Learning
DRL	Deep Reinforcement Learning
GNN	Graph Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
R-CNN	Region-based Convolutional Neural Network
YOLO	You Only Look Once
SSD	Single Shot MultiBox Detector
FCN	Fully Convolutional Network
XAI	Explainable AI
NAM	Network Animator

CHAPTER 1

INTRODUCTION

1.1 Background

India is recognized around the world as a biodiversity hotspot, and boasts a vast diversity of medicinal plants that are invaluable to traditional medical systems such as Ayurveda, Siddha, and Unani. These traditional medical systems carry a deep tradition of knowledge that spans centuries and encompass treatments for almost all ailments, and have universal tantamount for therapeutic goods today. However, preservation and access to this knowledge base is often a distant reality. Climate change, aggressive urbanization, habitat loss, and deforestation, along with socio-cultural factors such as the movement away from oral traditions, and loss of indigenous knowledge pose many challenges in accurately identifying and documenting medicinal plants.

Moreover, the vast diversity of plants in India combined with the tremendous morphological variation within a single plant adds to the complexity of identifying plants. Characteristics of leaves such as venation, margin, shape may vary upon age, senescence, environmental conditions, and location. Identifying plants is often extremely laborious because of the botanical classification often rests with subjective judgement or expert-level experience. Traditional methods of identification have a limited capacity to error, require specialized training, and cannot easily extend to scale in outdoor conditions.

With significant recent progress in artificial intelligence (AI) accented by deep learning and computer vision opportunities, there is now an unprecedented opportunity for the automation of plant species identification. Convolutional Neural Networks (CNNs), as popular subset of deep learning models developed for image classification, have proven successful in tasks that require fine-grained visual categorization. The nature of CNNs enables them to learn complex visual features (e.g., leaf shape, leaf texture, and leaf vein pattern) from the data itself making them uniquely suited to AI based applications requiring medicinal plant identification.

1.2 Motivation and importance

In recent years, there has been a renewed interest in herbal and plant-based medicine, which is driven partly by a cultural revival and partly by evidence from science itself. This resurgence brings a critical need of accurate and scalable identification systems for plants. This need is particularly critical in a country where traditional medicines are sometimes an essential part of people's economic, pharmacological and community health care behavior.

An AI plant identification techniques could address the dearth of traditional botany and modern technology, especially using light deep learning architectures such as MobileNetV2 or EfficientNetV2. These models have been specifically optimized for edge deployment and resource-constrained spaces such as mobile devices and embedded systems, as they allow for near real-time classification on inputs. They are ideal and timely if used in the field in remote or rural areas.

Additionally, current literature substantiates the practical implementability of such models. Studies such as Kumar and Sharma (2022) show that CNNs can be used for reliably classifying complex categories of plant classes, while Sharma and Singh (2023) provide reviews supporting the use of CNNs for automatic plant classification. Based on this prior work, the current study uses Indian medicinal plants, an application area that is underrepresented analysis in AI-based literature, despite the importance to society.

1.3 Problem statement

Although deep learning models have been successfully applied in domains such as agriculture, health care, and industrial automation, their use in the application of medicinal plant identification within the context of India has been rather limited. This research identifies several key challenges:

- The dataset comprised of morphologically similar leaf structures among the species increases the chance of error without sophisticated embedding and word extraction techniques.
- The datasets are often not thoroughly annotated and lack representation across the different environmental and geographical conditions.

- There is a need to balance the model performance (i.e., accuracy and precision) with computational efficiency on mobile or low-power devices.

The present research attempts to tackle these challenges with CNN architectures—MobileNetV2 & EfficientNetV2, trained on a diverse & well-annotated dataset of Indian medicinal plant leaves. The models are methodically evaluated on multiple performance metrics to consolidate their applications for practical use.

1.4 Project objectives

The major aim of this project is to utilize deep learning to develop an automated system capable of identifying Indian medicinal plants from various image data. The objectives of this research are as follows:

- Use the Indian Medicinal Leaves Image Dataset with 6900 images across 80 plant species, taken in a number of different lighting and situational contexts.
- Experiment and compare two of the leading CNNs, MobileNetV2 and EfficientNetV2, based on their classification accuracy, precision, recall, and resource efficiency.
- Include image augmentations within the dataset that employ changes to directionality (e.g., horizontal flippings), variation through angle (e.g., rotation), and increased magnification (e.g., zoom).
- Determine which model architectures provide the most accurate prediction for the least amount of computing resource use so that the models can effectively be used in low-resource and real-time scenarios.
- Design and create a tested prototype web-based application that allows users of the system (botanists, researchers, lead developers such as farmers) to upload their images of plant leaves and return an immediate signal or response from the model in the backend. The system design and movement of data is represented in Figure 1; it depicts how the users will interact with the application, from uploading an image of a plant to getting a result.

1.5 Scope of project

Scope of study is defined narrowly in the following sense to maintain a certain level of focus and allow for conclusion of the study in a timely manner:

- Develop and test CNN models with static leaves images of the pre-collected dataset.
- Use of pre-trained CNN architectures (MobileNetV2, EfficientNetV2) adapted for image classification tasks on India's medicinal plants.
- Assessments will be based on common evaluation measures like accuracy, precision, recall and F1-score.
- Classification tasks will be limited to leaf images (only), with stems, flowers, and roots excluded.
- Prototyping a web-based classification tool should lead to demonstrations at the academic level; which reasonably limits the study to prototype status only, not mass production in the immediate scope of the project.

Possible directions for research could include growing collection of the dataset with other sources of information like flower images and root images, textual information, and geolocation information. Other work could include using the models for mobile and edge devices or allow for offline use in the appropriate way to deploy a tool in the field.

1.6 Key contributions

This research project offers important contributions to the fields of deep learning and ethnobotany:

- The creation and use of a diverse annotated dataset of 6,900 medicinal plant leaf images so that other machine learning studies can be currently and into the future based on that dataset.
- Acting on the dataset the implementation of two state-of-the-art CNN architectures, both MobileNetV2 and EfficientNetV2 and the evaluation of these deep learning systems on a classification task of medicinal plants.
- A comparative analysis that showed EfficientNetV2 not only provides with high accuracy, but also finds value in a system like MobileNetV2 which has a decisive advantage when resources are limited.

- The design and deployment of a web-based prototype tool that utilises these deep learning techniques into a real-world artefact to help botanists, researchers and rural communities in AI based plant identification systems.
- Creating awareness on the optical potential that deep learning could have on the preservation of biodiversity and indigenous knowledge and consequently benefitting academic and public health through better access to plant-centric health resources.

This study has created the contours of future research and movement into projects leveraging AI within plant sciences, particularly within ethnobotany and conservation biology.

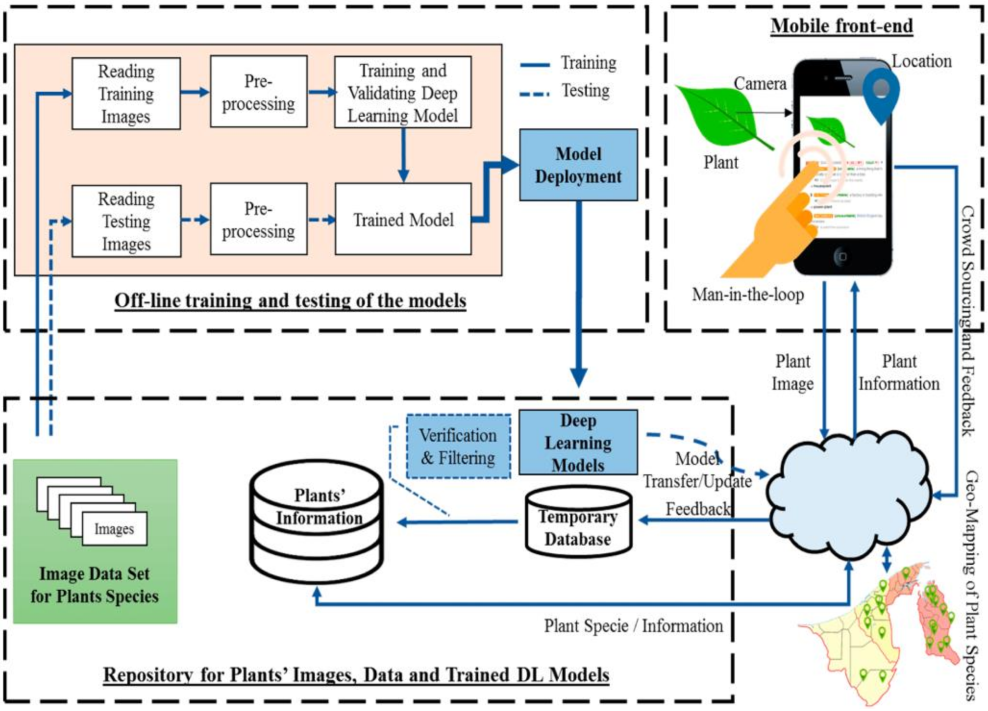

FIGURE 1. Architecture of the Deep Learning-Based Plant Identification System

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

India's biodiversity has rendered it one of the richest pillars of traditional pain-management knowledge of traditional medicines and biologicals. The country has over 7000 species of medicinal plants amongst the most used species in their ancient forms of medicine, such as Ayurveda, Siddha and Unani, to continue to extend the knowledge base of the pharmacopoeia. There are probably thousands (if not tens or hundreds of thousands) of herbal formulations and remedies which are still in holistic use to this day and have become increasingly important and popular in an increasingly fractured health care system of the modern era. Yet, the country is faced with several serious threats to the survival of both its medicinal plants, and the traditional wisdom that uses them, such as; rapid urbanization, changes in land-use, deforestation, pollution and weakening "transmission of traditional indigenous knowledge".

With rising interest in natural and alternative medicine; and with a pressing need for conservation capacity to safeguard India's botanical treasures, there is a pressing felt-need for a the approach of high-quality and high-through put identification systems. Guitar and Mulligan (2001) indicated that the traditional plant species identification approaches based on morphological analyses and limited "botanical expertise" have limited feasibility to agree with accessibility, subjective issues of identification and time constraints/habitation. These issues were avenues for technological creativity, especially with advances in artificial intelligence/electronic medical record information and capabilities.

Deep learning is one of the many topics under AI. Out of the myriad of technologies under AI, plant identification through deep learning techniques, particularly computer vision techniques such as Convolutional Neural Networks (CNNs), allows for computers to automatically extract intricate patterns in plant images, allowing for less human input. Such technologies have the potential to transform plant identification, allowing for faster, more reliable, and scalable classification systems to be used in research and field situations. This chapter will discuss the many works that have been published in this space in detail, report on current technologies use, their pros and cons and avenues for future work leading to potential plant identification technologies.

2.2 Deep learning approach towards plant identification

Deep learning has progressed immensely in the past decade, and the use in botany has likewise advanced. CNNs have been appropriated for plant species classification due to their biological inspiration from the visual perception process, and because they can learn hierarchical structures of input images. In other words, CNNs are particularly adept in learning more complicated patterns of plants such as morphology.

Anchitaalagammai et al. (2021) provided insight to understanding what is required to make a deep learning model for medicinal plant identification successful. Their study identified three significant factors: data quality, model architecture, and efficiency of computations. Their findings noted that datasets used for training should not be only large, but also diverse in relation to light, background, angle of image and plant growth stages. Collectively, those factors contribute to the generalizability of the model.

A narrower study by Bhelkar and Sharma (2022) examined if CNN architectures such as AlexNet, VGG16, and ResNet performed well at leaf-based classification. They concluded that compared to traditional feature-based approaches to learning, deep CNNs are better with respect to accuracy and robustness. Bhelkar and Sharma (2022) also discussed some important preprocessing steps include normalizing images, augmenting the data (by rotating, flipping, and scaling) and removing the background, which improved the learning efficiency of the model.

Additionally, the authors, Pukhrambam and Sahayadhas (2022), noveltly explored DenseNet architectures. The advantages of DenseNet are its dense net connectivity pattern, which would allow for a better flow of gradient, which helps reduce the issue of diminishing gradient, making it easier to train a deeper model. As an additional contribution to their model the authors examined features related to not only visual data, but also bioactivity. They have proposed a multimodal classification by integrating data from visual images and chemical compounds, allowing better assessment of a plant's medicinal potential. This also opens up the use of this interdisciplinary research to the fields of pharmacognosy and bioinformatics.

Another significant contribution came from Leena Rani et al. (2022), who demonstrated the effectiveness of transfer learning models such as InceptionV3 and EfficientNet. These pre-trained networks, when fine-tuned with domain-specific data, exhibited remarkable accuracy even with smaller datasets. Transfer learning accelerates training and mitigates overfitting by utilizing knowledge from large-scale datasets like ImageNet. The authors also tackled the problem of class imbalance by employing techniques like synthetic minority oversampling and class weighting, thereby improving model reliability across underrepresented species.

2.3 Real-time identity issues

While deep learning models can produce good results under controlled lab conditions, transferring these models into real-time, real-world situations has a host of other issues. Differences in environmental variations such as lighting, camera dimensions, occlusions, or complex backgrounds can severely limit model performance. Further more there is also inconsistency related to the image acquisition process, such as changing the angle of acquisition, motion blur and potentially only partial observation of some organs of the plant. Kavitha et al. (2023) modeled the same CNN based approach in a field environment and identified some of the practical limitations of working with deep learning models under uncontrolled conditions, and they found that the performance severely dropped with the transfer from lab-researched images to field-based observations. Overall, the study demonstrated the need for training deep learning models on a range of real-world data that meets the needs of plant-based identification, and to also make sure that all plant data sets are evaluated and then iterated with continual updates as environmental conditions change.

Additionally, the computational workload of deep learning models can be an issue. The vast majority of current state-ofthe-art digital image Mathematics-based CNNs require training and inference on powerful GPU hardware, which very much limits the availability of the speed and efficacy of these models on mobile devices and embedded systems. Lightweight versions of CNN digital image-based mathematics model architectures exist (MobileNetV2 and SqueezeNet). These networks implemented some consideration for memory and compute with respect to time and the adequacy of the models for reproducing responsible measures of accuracy. Not considering the hardware available, there are still further aspects to consider and take action to utilise any appropriate model. Hardware required any additional tuning, model pruning and/or quantization to try to get an adequate model for the capability and commitment of the particular device and context.

To help usability even further, researchers are currently concentrating on the creation of mobile apps that have real-time feedback, offline options, and user-friendly layout. Mobile apps have great potential for rural health workers, herbalists and even enthusiasts of ethnobotany. Adding contextual data for identification can also be achieved through the addition of GPS and environmental sensors.

2.4 Research gaps and future research directions

Although substantial advancements have been made, there are still considerable areas of knowledge that remain unaddressed regarding the use of deep learning to identify medicinal plants. First, the majority of studies are based on a limited number of species, often tied to regions of the world. It is because of this variation in geographic region that a comprehensive database of medicinal plant species in India or any other country is not possible at present. In order to build a large-scale dataset that can provide as many images as possible, the dataset must consist of images taken in a variety of environmental conditions as well as images that capture the various parts of each plant (leaves, flowers, stems). This variance in a dataset will ultimately enable the development of models that are robust and generalizable.

Second, interdisciplinary cooperation is limited. Datasets rich in contextual and therapeutic information combine the knowledge of botanists, anthropologists, pharmacologists, ethnomedicine practitioners, and artificial intelligence and data scientists. Working together, researchers of diverse backgrounds may not only classify plants into species but use the knowledge of the co-investigators to classify the plants by medicinal use, geographical habitat, and pharmacological behavior.

Finally, a common limitation in the models used currently is a lack of interpretability. While deep learning models are well known for their high level of accuracy, the "black-box" nature of these systems samples actions, often leading to challenges for the researcher to comprehend the rationale behind predictions. Future research should focus on leveraging explainable AI (XAI) techniques to generate visual evidence and importance plots of the feature variables in a given classification process.This transparency would help build trust with users and facilitate model adoption in sensitive areas such as the fields of medicine and conservation.

As mobile applications become even more common, we need to start converting research prototypes into usable applications. Converting research prototypes into applications means designing intuitive interfaces, creating offline capability, and designing user feedback systems. This participation even could democratize access to identification tools so that rural health care workers, forest officials, and herbal practitioners can have access to identification tools and ultimately make better decisions in real time.

2.5 Conclusion

The use of deep learning for the identification of medicinal plants is an exciting step towards automating and scaling a purely manual and expert-based process. As we have seen in previous studies, CNNs and transfer learning models have significant potential to identify medicinal plant species with a high degree of accuracy. Next steps to take the technologies we presented from a lab to a practical application require certain enhancements.

Most importantly, we need large scale, annotated datasets that are diverse; we need models with appropriate consideration of explainable AI; we need models that consider interactions across the biological, ecological and biochemical contexts that are rich enough for an interdisciplinary team to construct modeling of a problem; and we need to put these models into lightweight applications that can be run by a non-expert, and in real-time.

As India continues to adopt more digital approaches to healthcare and sustainable environmental conservation, AI-driven and powered plant identification systems will allow a further branching out.

This literature review synthesis was a beginning, but it also provides direction for future innovation. As three allied disciplines have gaps in research for medicinal plant identification, we need to close the gaps in joint collaborations between botany, ecology, and AI as we envision a future where there is a holistic, live, AI-centric and medicinal plant identification system. Collectively these integrative developments will provide security for the preserved medicinal riches of India, while developing a pathway to achieve sustainable health and environmental management for the future.

CHAPTER 3

METHODOLOGY

3.1 Dataset Description

This study has been done based on Indian Medicinal Leaves Image Dataset that has been scrupulously created by Pushpa B. R. and Rani, Shobha (2023). We believe that this dataset is a substantial improvement over existing dataset and makes a meaningful contribution to medicinal plant identification by providing a large, well-structured, diverse leaf-image dataset that can be used to investigate the application of deep learning-based techniques for ethnobotanical studies. It contains 6,900 high-resolution images classified into 80 fine-grained Indian medicinal plant species categories. Every class corresponds to a specific plant species, which has been traditionally acknowledged for its medicinal value in Indian traditional health care systems including Ayurveda, Siddha, and Unani.

The dataset's advancement in this area is not simply based on having a large number of classes and images, but on the conditions in which the dataset was collected. The fundamental difference is that many image datasets are created in controlled laboratory conditions - where the angle, lighting, and background are consistent for all images - the images in this dataset were taken outside in natural conditions. The plants were photographed under varying ambient light conditions (direct light, shade, overcast), from varying angles (horizontal, tilted, or partially occluded), and were also photographed with varying background clutter (soil, grass, other flora).The addition of environmental variability into this dataset reflects the conditions that real field identification systems face and therefore improves the model’s ecological validity and generalizability.

The range of biodiversity represented in this dataset extends across several climatic regions and ecological zones of India. Major species are Aloe Vera, Neem, Tulsi, Ashoka, Turmeric, Mint, and several more that are not only useful for medicine but are culturally and economically valuable. The morphological nature of these plants extends to leaf sizes, shapes, edge pattern, veins, color gradients, and surface textures, each of which is indispensable to identifying these species.Neem leaves are small and serrated and have a bright green color, Ashoka leaves are long and smooth and droop or arch elegantly. The ability of the model to consistently represents such minor morphological traits across different conditions strongly enforces its power to discriminate among visually similar species—something that is typically a difficult and labour-intensive task even for expert botanists.

A major strength of our dataset is its capacity to serve as a bridge for theoretical AI research and real-life application. Many of the existing datasets in the literature have the bias that either they are too small, handcrafted artificially or the taxonomic range is limited which restricts the applicability of the model in dynamic outdoor scenario.The Indian Medicinal Leaves Image Dataset, on the contrary, has even greater heterogeneity. Specifically, it has higher intra-class diversity and inter-class similarity. Consequently, it forms a more challenging environment for model evaluation and learning. Secondly, the data set is quite pertinent for mobile or on-device AI applications. High efficacy in uncontrolled conditions is a must for the latter since their main distinctive feature is an unpredictable environment. Thus, a model trained with a more naturally explicit dataset will be more general and applicable, as it is more robust and adaptable.For example, a model trained on such a wide variety of image can more easily recognize a Neem leaf photographed in a forest with mobile phone camera or a Tulsi leaf shot against a kitchen background even with varying resolution, light conditions and angle.

Figure 2 of this chapter presents a sample collection of training images that highlights the complexity, richness, and variety of this dataset. These samples illustrate various difficulties: handling occlusions and uneven lighting; and discriminating similar-coloured foliage that all need to be confronted for high-accuracy identification.Ultimately, the image also serves as a reminder of why typically considered image classifications methods often have a poor success rate in these circumstances. Furthermore, a more contemporary deep learning approach i.e. Convolutional Neural Networks would be necessary to appropriately extract meaningful features across these varying situational circumstances.

Therefore, the Indian Medicinal Leaves Image Dataset is more than a random selection of leaves on an image dataset, it is a tool that supports the achievements leveraged for the outcome of applied machine learning research, often making it an instrumental element that supports the responsible identification and conservation of India's medicinal plant potential. As a dataset that is grounded in natural diversity of real ecosystems, it is an ideal referent for the advancement of deep learning training as the conditions of a sample the plants grow best in supports the model to be trained and operationalized in, agriculture, forestry, rural healthcare conservation biology.The dataset meets the technical requirements for training meaningful models and aligns with the overall vision of digitally transforming biodiversity monitoring and traditional medicine.

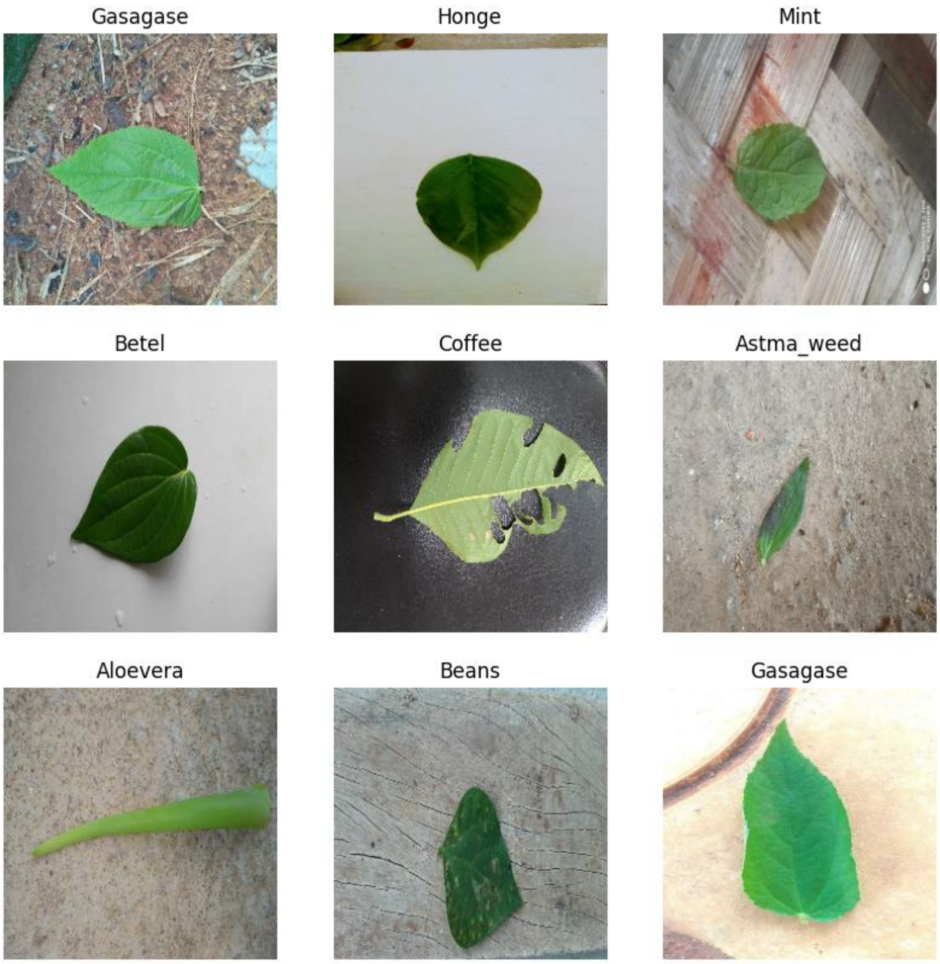

FIGURE 2. Sample Images from the Indian Medicinal Leaves Dataset

3.2 Data Partitioning

An important step across all supervised machine learning pipelines is planning and managing the division of data into useful subsets. Here, we carried out the division using a dataset, the Indian Medicinal Leaves Image Dataset, with a total of 6,900 labeled images taken across 80 classes. We created three subsets that partitioned the dataset into clearly defined training set, validation set, and testing set, using an appropriate ratio of 70:15:15 that has gained acceptance within the literature. The training set allows models to learn, the validation set allows for hyperparameter tuning, and the testing set allows for the final evaluation of the model.

Training Set (70%)

One of the reason the training set is the most important part of the learning process. This portion of the data is used to revise the internal parameters of the models through backpropogation/gradient descent (convolutional filter weights and biases, etc.), the model's first exposure to the problem domain essentially. The training set as a whole is essential to the process of training the network to discern the unique characteristics associated with different leaf species, as intended.

The training set was comprised of of 70% of the initial dataset - all 4,830 images. This ensured that the model utilized a wide and diverse sample of examples from each plant class, allowing the model to learning class specific features - the variations in leaf morphology, vein structure, surface texture, margin type and colour gradients that differentiate individual classes for the purpose of plant identification. The large amount of training data and the nature of how deep neural networks work allows general representations to form, instead of memorizing exemplars, which is fundamental - especially considering that remotely collected images can vary greatly in both environmental effects and conditions that affect image quality.

Validation Set (15%)

The importance of the validation set which in this case contains about 1,035 images cannot be over-emphasized during model development despite its not being utilized in training to update or change the weights of your model. The validation set serves as a pause during training, where its purpose is to see how well the model is performing with unseen, but somewhat related data. It also serves the purpose of noticing any chance of overfitting, or when your model is learning the training task well, but can’t generalize enough to see the new data.

Various validation metrics (e.g., accuracy, loss, precision or F1-score) are recorded and evaluated based on epochs and we further adopt several altering learning conditions based on:

- Early stopping which stops training completely when there hasn’t been any improvements.
- Learning rate scheduling or lowering the learning rate to better converge on a solution.
- Model checkpointing or saving the best weights model.
- Architectural tuning may include altering layers, filters, and dropout rates based on performance.

The validation set essentially replicates or mimics the training process to allow model tuning and to avoid overfitting or excessive learning of the training data specifically. It has to be thought of as the compass that we provide for ourselves to help us move closer in optimizing the model.

Testing Set (15%)

The final subset, containing around 1,035 images, is the test set, which is not available during the previous stages until the training and validation process is completed. It is intended to ensure the subsequent assessment of the final model’s performance in an objective and impartial manner. In other words, this data set has not been observed by the model before and has no effect on the training or validation process. Therefore, it is possible to evaluate the model’s actual generalization capability to determine whether it can work well for entirely new data instances that represent real-world applications.
Metrics for performance such as accuracy, confusion matrix, precision, recall, and F1 score are computed on the test set to measure the final effectiveness of the trained model. The results derived from the test set are what I will report as the model's performance in the study and are ultimately essential in determining whether or not the model is ready for deployment in order to complete useful applications, for example, a mobile plant identification device, or assistant for biodiversity researchers.

Rationale Behind the 70:15:15 Split

The 70:15:15 proportion has been employed for the three datasets for a reason: it indicates a trade-off between model capacity and overfitting prevention, while also supplying a reasonable view for fair evaluation. We safely felt that a maximum of 70% for training provides the model with adequate samples to learn across 80 classes and a diverse feature representation for a wide range of plants. We saw the 15% for validation as more than sufficient to capture any trends we could associate with overfitting while optimising towards our objective. Having the final 15% data just for the test dataset accomplishes the task of having a reasonable expectation of statistical reliability, fair representation, and validity of our model's performance metrics.

Other ratious were considered (80:10:10 or 60:20:20) but ultimately rejected, due to the number of classes, and the constraints that requires each subset of data represents the sum of all plant species. We hopefully obtained an adequate representation of all available samples per class. This aspect of ensuring the samples are correct is crucial for multi-class classification problems - it likely becomes increasingly significant, with high inter-class similarity and intra-class variability, as demonstrated in a leaf-based identification system.

Stratified Sampling for Class Balance

One important element of partitioning was maintaining the distribution of classes across all three data subsets. A random cut might have resulted in observing that certain classes are underrepresented or missing from the validation or test sets (e.g. the least represented species). To counter this, we used a stratified sampling approach, which guarantees that each class of plants is represented in the training subset, testing subset, and validation subset in the same proportion it appears in the parent dataset.

Stratified sampling enables this operational consideration for research because it allows for the following:

- Culmination of biased classification performance metrics by weighting the classes that are prominently represented
- Provides uniform learning and testing on dissimilar classes in the dataset.
- Optimal generalization to all 80 plant classes.
- Will influence Model Performance and Research Goals.

In summary, the intentional partition of the dataset allows for the best use of a dataset—maximize learning, inform what to adjust, and evaluate fairly. This also fulfills the larger goals of this research to not only classify plants but to classify plants in an accurate, robust, fair across classes, and easily deployable in real world studies. This is a key part of modeling that will allow you to not just classify well in isolation but able to do so in a large scale testing advance where there could be complexity due to noise from the environment, variation between devices, and diversity of species among the data you are passing through the process

In conclusion, the 70:15:15 data split model, supported by stratified sampling, has been developed in a methodologically robust way for deep-learning-based medicinal plant classification. This ensures that the model is trained, validated, and tested statistically reliably and ecologically relevant way, allowing for the development of accurate and trustworthy AI-based tools for plant identification.

3.3 Data Augmentation

Data augmentation is an important method for enhancing model generalization and mitigating overfitting in computer vision tasks, particularly for limited datasets. Given that the Indian Medicinal Leaves Image Dataset had 6,900 labeled examples across 80 classes, there needed to be a strong strategy for augmentation. In nature, leaves can appear in many different orientations, at scales and positions all unique with lighting that can be very different. A well-trained model must identify and be invariant to these changes.

To overcome these issues, a real-time data augmentation pipeline was written using the tf.keras.Sequential API in TensorFlow (see Figure 3). The code shown in Figure 3 resulted in a data augmentation pipeline that was dynamically applied during training using TensorFlow's map() function, enabling on the fly execution of the transformations for each training batch. The overall effect of this approach was to effectively increase the size of training data but also add another level of variability that better mimicked the actual appearances of leaves.\

3.3.1 Augmentation Pipeline Overview

The augment logic is included in a sequential preprocessing layer as shown in Fig 3 and then this pipeline is applied to the training dataset (train_ds) using a map() function. The first transformation pipeline is applied only to the input images (x) for train_ds leaving the labels (y) untouched, i.e., they are passed through unchanged. The. prefetch() method is used to allow data loading and preprocessing to overlap with model execution to drawn-out the training time of the model.

```
data_augmentation = tf.keras.Sequential([
    layers.experimental.preprocessing.RandomFlip("horizontal"),
    layers.experimental.preprocessing.RandomRotation(0.2),
    layers.experimental.preprocessing.RandomZoom(0.3),
])
train_ds = train_ds.map(
    lambda x, y: (data_augmentation(x, training=True), y)
).prefetch(buffer_size=tf.data.AUTOTUNE)
```

FIGURE 3. Data Augmentation Pipeline for Training Set

3.3.2 Transformation Techniques and Their Impact

Each technique in the augmentation pipeline contributes to enhancing the model's ability to recognize leaves under a wide variety of conditions:

Random Horizontal Flip

- **Description:** Randomly flips the image horizontally.
- **Motivation:** In natural settings, the orientation of a leaf is rarely consistent. A leaf may hang to the left, right, or at a diagonal, depending on the angle of capture and wind or gravity. By horizontally flipping the image at random, the model learns that mirror images of the same leaf are still the same class.
- **Impact:** Reduces orientation bias and improves rotational invariance, which isespecially useful in plant species where leaf orientation is not distinctive.

Random Rotation (±20 degrees)

- **Description:** Randomly rotates the image within ±20 degrees (0.2 radians) .
- **Motivation:** Leaf samples in the dataset may be skewed, tilted, or rotated slightly due to the capture process not being explicitly manual or the natural angles of the leaf. The rotation serves to vary the leaves in the dataset and as represent it, but also helps the model disentangle the leaf's identity from its pose.
- **Impact:** Allows the model to become robust to variation in random rotation, and ultimately improve generalization of identifying leaves across different test images with varying camera perspectives.

Random Zoom (up to 30%)

- **Description:** Random zoom into the image, giving an impression of proximity and focus.
- **Motivation:** Leaves can appear to be of different sizes depending on how close the camera was, the resolution and the extent of focus. Randomly zooming conditions the model to focus on local detail (e.g. venation and texture) and global structures (e.g. throughout shape and symmetry).
- **Impact:** Improves the models ability to extract discriminative features at multiple scales, reducing the tendency to rely on items of only coarse features.

3.3.3 Benefits of Data Augmentation in Fine-Grained Classification

Plant classification using images of plant leaves is a fine-grained visual classification problem, since similar shapes and textures can represent different classes (e.g., Neem vs. Honge leaves), and the same class can occur with wide variability (e.g., age, light quality, cluttered backgrounds, camera settings).

Data augmentation is central to addressing these issues:

- **Fighting Overfitting:** Leveraging augmented data add one-dimensional variations with semantic relevance, which minimizes the chance that the model memorizes training samples. This is complementary to previous regularization methods of dropout, weight decay, and early stopping calculations.;
- **Increasing Effective Size of Dataset:** with only 6,900 images available there is a high probability of models creating highly specialized representations of images. Data augmentation using on-the-fly transformations synthetically increase the variability, and thus diversity of the training dataset without the needs for new labelled data;
- **Increasing Feature Invariant Properties:** Data augmentation will displace the input space slightly but retain class label properties. Thus, the model can develop strong invariance representations, encouraging the emphasis on semantic changes and content of the representation, rather than position and angle specificity;
- **Encouraging Discriminative Learning:** Methods like zooming and rotation will encourage the model to rely on the overall leaf structures like the venation or tip

shapes; base angle types; and margin serration etc.; and thus, trends or uniqueness that are key differentiating factors in botanical classification.

3.3.4 Real-Time vs. Static Augmentation

One consideration with the design choice is to perform real-time augmentation or to pre-generate augmented versions and store these. Advantages of using real time Augmentation:

- Memory efficiency: on-the-fly synthesized samples in training, significantly reduce disk space consumption.
- Dynamic variation: during training the image can be presented each epoch as a slightly different augmented version of the same image, leading to better generalization.
- Increased parallelism: data loading and augmentation can be carried out asynchronously with training using TensorFlow's AUTOTUNE prefetching, thus decreasing latency.

3.3.5 Visualization (Figure 3)

Figure 3: data augmentation pipeline. The original image is subjected to random transformations, such as horizontal flipping, rotation, zoom, and others, before it is delivered to the model during the training stage. Every transformation is stochastic and dependent on the training flag. Taking into account such conditions, augmentation is not used in the validation part or testing to preserve the integrity of these subsets.

3.4 Model Architectures

To assess the utility of different deep learning models for plant classification, we selected two cutting-edge models. MobileNetV2 and EfficientNetV2 were selected, as both models are efficient, scalable, and have been shown to perform well in image recognition tasks.

3.4.1 MobileNetV2

Introduction and Motivation

MobileNetV2 is a state-of-the-art convolutional neural network (CNN) structure that was developed specifically for efficient image recognition tasks on mobile and embedded devices

(i.e. resource-constrained devices). MobileNetV2 builds on the original MobileNet by using inverted residual blocks and linear bottlenecks to further reduce the number of parameters and computation requirements while retaining competitive performance on the benchmark.
A plant identification system has the potential for real-world use that would require efficiency and functionality (e.g. plant identification mobile app for farmers, herbal practitioners, botanical researchers, etc). The work of a system using MobileNetV2 represents an ideal combination of trade-offs in efficiently and predictive performance.

Architecture Overview

The architecture used in this study is displayed in Figure 4 , which describes how data is processed through each layer of the model, through the preprocessing input step to the output softmax classification step.

Input Layer

- **Input Size**: The model accepts RGB images that have been resized to 128×128×3.
- The reduction in input dimension helps decrease the training time and the size of the model without negatively affecting classification performance, especially when using a pre-trained backbone.

Convolutional Backbone: MobileNetV2

The backbone of the architecture is a fixed, pre-trained MobileNetV2 model that was trained on ImageNet. MobileNetV2 will be non-trainable, and treated as a fixed feature extractor, which will result in the convolutional layers freezing during training.
The feature extraction procedure includes:

- 3×3 Convolution layers using ReLU6 activation function (green arrows in Figure 4).
- Batch norm was used to stabilize the learning process.
- Depthwise-separable convolutions reduce computation by simultaneously maintaining spatial relationships.
- Inverted residuals blocks have input and output dimensions equal which allows features to be passed through the network directly.
- Progressively down sampling of resolution by the application of stride-2 convolutions which reduces the 128×128 image down to a 4×4×1280 feature tensor.

Pooling and Flattening

- A global average pooling layer is applied to convert the 4×4×1280 tensor into a 1280 dimension vector.
- This global context-aware representation summarizes the most salient global features through the entire spatial domain of the image.

Classifier Head

The custom classifier part (right side of Figure 4) is on top of the MobileNetV2 feature extractor and consists of:

- A fully-connected (dense) layer to combine the extracted features.
- A dropout layer with a dropout rate of 0.3. The dropout layer is to prevent overfitting and randomly deactivate the neuron for a specific fraction of the time during training.
- A fully connected (dense) layer consisted of 80 neurons for 80 medicinal leaves classes and a softmax activation function to return probabilistic class outputs.

Figure 4: MobileNetV2-Based Architecture

In Figure 4, we show the entire architecture visually and each transformation diagrammatically from left to right:

1. An input image (128×128×3) goes through the pre-processing pipeline.
2. The input image goes through convolutional layers that downsample the image and increase the number of filters (n=32→96→1280).
3. The output tensor (4×4×1280) is then global average pooled to flatten the input and create a feature vector.
4. The feature vector was passed into a fully-connected layer, reconstructed with dropout, then through to softmax classification. .

The colored arrows throughout the diagram represent the different versions of operations (in this case pooled representations):

- Green arrows: convolution + ReLU operations.
- Red arrows: max-pooling or downsampling.
- Black arrows: data flow or structural change

Compilation Details

The MobileNetV2 model was constructed using the following hyperparameter settings to complete the successful training of the model:

- **Optimizer: Adam**, an algorithm which computes adaptive learning rates for each parameter and works well on problems with noisy non-stationary gradients or sparse data.
- **Loss Function: Sparse Categorical Cross Entropy**, is appropriate for multi-class classification problems, and since our labels were integer-encoded, this loss function was fitting.
- **Metrics:** The model was monitored for performance using accuracy (as this is the main goal) and validation loss during training..

Training Procedure

- We trained the MobileNetV2 architecture for a **maximum of 50 epochs**.
- We used **a learning rate of 0.0007**, which was determined based on experimentation in order to balance convergence and not overshoot the potential minima.
- We used **early stopping** based on validation loss with a patience of 5 epochs to ensure that overfitting does not happen, or the model does not keep training without any noticeable improvements during validation.

The training was also run with **data augmentation (see Section 3.3)** in real-time, which was another attempt to help generalization and mimic some natural variability considering leaf orientation, size and lighting will continuously vary in nature.

Performance and Observations

- To enable fast convergence and minimize overfitting while training the classifier, I frozen the base model. This was especially relevant with the dataset size being comparatively small (6900 images).
- The MobileNetV2-based architecture was able to achieve a test accuracy of 94.39%, which demonstrates.
- Lightweight models can perform successfully and connectively perform on complex tasks such as medicinal plant identification.
- The lightweight size of MobileNetV2 offers the opportunity for real time inference on a mobile or edge devices, making it applicable for deployment in the field.

Conclusion

MobileNetV2 is a strong baseline model in this study due to its performance and deployment potential. It is capable of effective extraction of hierarchical visual features and, with a custom classification head, suited to fine-grained classification such as Indian medicinal plant leaf recognition. Pre-training on an ImageNet backbone made the model faster to train and had some degree of generalization. Structural considerations, such as global average pooling and dropout, also helped to make the model more stable, and provided better functionality.

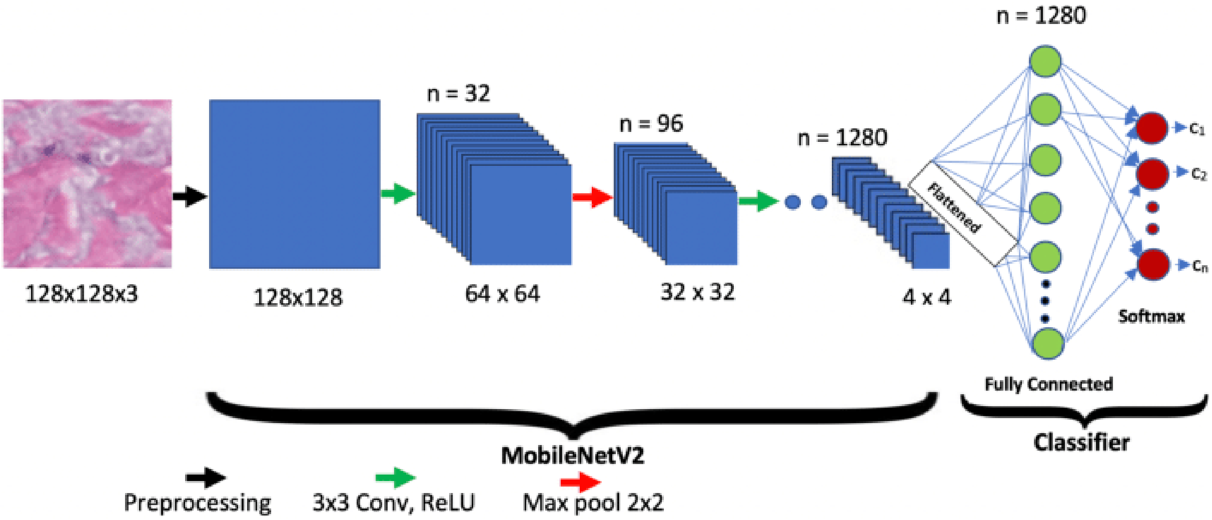

FIGURE 4. Architecture of the MobileNetV2-Based Classification Model

3.4.2 EfficientNetV2

Introduction and Rationale

EfficientNetV2 is the second-generation model of EfficientNet developed by Google. The major contribution of EfficientNetV2 is based on a compound scaling method—instead of incrementally increasing one dimension (depth, width, or resolution) independent of the others, EfficientNetV2 scales all the dimensions together to get optimal performance. Compared to conventional models with similar formalizations, EfficientNetV2 provides faster training, a smaller model, and greater accuracy when applied to medium and larger datasets.

For the task of classifying Indian medicinal plant leaves, where the differences between classes can be subtle, EfficientNetV2 offers a fine-grained feature extraction capability, as well as maximal generalization. Being able to distinguish Tulsi vs. Mint, or Neem vs. Eucalyptus, is critical given their visual similarities.

Architectural Overview (Figure 5)

Figure 5 illustrates a block-wise view of the EfficientNetV2 model that was used for this study. The model is comprised of multiple convolutional blocks, including Fused-MBConv and MBConv blocks, that provide a nice balance of efficiency and representational power.

1. Input Layer

- The model expects a 128×128×3 RGB input image.
- The input is first introduced to a single convolutional block (3×3), so the thread size is increased (the receptive field), and the model can use it to extract some initial low-level features.

2. Fused-MBConv Blocks

- The next set of blocks consists of Fused-MBConv1 and Fused-MBConv4 (×4) blocks.
- Fused-MBConv blocks significantly improve the speed of training and inference by combining depthwise and pointwise convolutions into a single 3×3 convolution.
- The blocks in these layers help to extract low-level features efficiently and quickly downsample spatially as a result.

3. MBConv Stacks

- The model then proposes MBConv blocks with increasing complexities:
  - EMBConv4 (×6): grow the features while keeping it efficient.
  - EMBConv6 (×9) and EMBConv6 (×15): deepen the nonlinear transformations that the model can learn to model the various complex textures and patterns that exist in the leaves.

EmBConv block consisted of expansion-convolution-squeeze-excitation methods to:

- Expand the dimension (width) of the input.
- Do depthwise convolutions (depth).
- Recalibrated the features (excitation) with attention mechanisms.
- Project to lower dimensions (compression) with skip connections.

In each of these MBConv blocks, the model effectively learned mid-level and high-level abstract features that it used to compare the 80 classes of medicinal leaves.

4. Final Convolution + Pooling

- After passing through the analysis stage, a 1×1 convolution layer and global average pooling collapsed the three-dimensional tensor into a one-dimensional feature vector.

5. Classification Head

- The pooled vector is flattened and passed through:
  - A dense layer with 256 neurons, applying non-linearity and abstraction.
  - A dropout layer with a dropout rate of 0.4 to mitigate overfitting by randomly turning off neurons during training.
  - A final output layer with 80 softmax neurons representing a predicted probability distribution over the entire categorical space.

Compilation Details

The model was compiled for training with the following:

- **Optimizer**: **Adam**—which uses an adaptive learning rate strategy and works well with sparse gradients.
- **Loss Function**: **Sparse Categorical Cross Entropy**—because class labels were integer-coded.
- **Learning Rate**: **0.0001**, a small learning rate that was tuned carefully so that the model could converge safely and stably without overshooting when backpropagating.

This was considered the best configuration in order for the model to learn the subtle textures, edges, venation, and contours present in leaves without overfitting to the model early or have the model plateau.

Training Strategy

Training occurred for a maximum of 50 epochs, with a number of performance enhancement and stabilization techniques being implemented:

- Early Stopping Callback:
  - Based on validation accuracy, the callback used.

    - The early stopping callback had a patience value of 5 epochs - this allows the model to be halted if the model has reached the point of no improvement, which saves compute time and limits chances of overfitting.
- Data Augmentation (Refer to Section 3.3):
    - Real-time augmentation through flipping, rotating, zooming, and brightness adjustment was implemented.
    - This allowed the model to become invariant to real-world lighting, orientation, and scale changes.
- Frozen Backbone:
    - The first point was to freeze the EfficientNetV2 backbone, and only train the classification layers.
    - This step allowed the dense layers to adapt the leaf dataset but still not change any of the learnt features from the pre trained ImageNet21K.

Performance Metrics and Evaluation

The model was assessed by standard accuracy and advanced classification metrics to evaluate subtle differences in the performance. Performance is shown in Table 2.

Metric	**Value**
Test Accuracy	96.5%
Precision	0.9708
Recall	0.9660
F1 Score	0.9666

Table 1: Performance of EfficientNet V2

Key Observations:

- EfficientNetV2 outperformed all mobileNetV2 in all metrics for the evaluation.
- EfficientNetV2 also displayed better generalization on the test data likely due to greater feature abstraction from a deeper and richer feature hierarchy.
- The greater F1-score suggests that the model performed equally well on both common and rare leaf classes.

Conclusion

In this study, EfficientNetV2 emerged as the most powerful model with the best accuracy for classifying medicinal plants using leaf images. Its modular architecture is shown in Figure 5 and can use several cutting-edge convolutions, such as Fused-MBConv and MBConv, to produce a scalable and trainable deep network.

The EfficientNetV2's use of dropout, dense abstraction, and global pooling allows the model to derive low-level textures and high-level plant features, which are very important when using a multi-class classification approach. In addition, pre-trained weights from ImageNet21K allowed the model to converge faster and generalize better.

The model's excellent test accuracy, 96.5%, suggests the model is very trustworthy, and with even better top-level precision, recall, and F1-performance, indicates a strong candidate for a mobile or cloud plant identification approach for practical use.

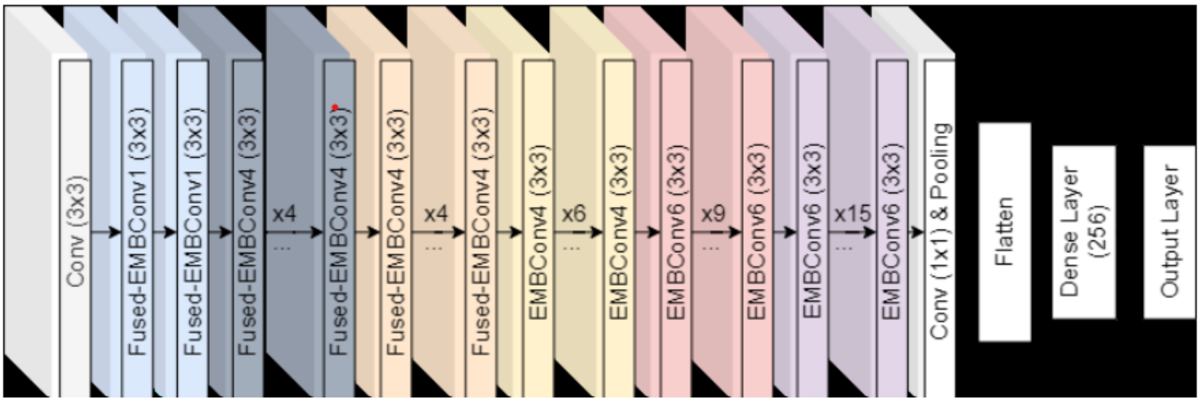

FIGURE.5. Architecture of the EfficientNetV2-Based Classification Model

3.5 Training and Evaluation

An organized pipeline was established to rigorously train and reliably evaluate the deep learning models—MobileNetV2 and EfficientNetV2—that encompassed data preparation, model compilation, training strategy, and multi-metric evaluation. This was important so that the model performance could be benchmarked in identifying Indian medicinal plant leaves, where visual differences between classes frequently appear subtle, and required fine-grained feature extraction.

3.5.1 Dataset Splitting and Preprocessing

The dataset consisting of 6,900 images from 80 different medicinal plant classes was divided in the ratio of:

- 70% for training: used for learning model weights.
- 15% for validation: used for monitoring overfitting and for Early Stopping.
- 15% for testing: held out and used only to evaluate final performance.

Each image was resized to 128×128 pixels and was scaled normalization to scale all pixel intensities to [0, 1], for numerical stability during backpropagation. All images were normalized. Each of these images were converted to RGB if they were not already established in that format, as many critical features were color-based features, which can be critical to differentiating plants botanically.

During training, real-time data augmentation was done to improve the model's ability to generalize when real-world input variability was present, such as:

- Horizontal flipping and vertical flipping.
- Random rotations (from ±20 degrees).
- Zoom range of 0.2.
- Brightness and contrast fluctuation.
- Minor height and width shifts.

Many of these augmentations provided variability that could exist in the real world, as well as ensure model overfitting to fixed patterns (which in many cases subject matter experts identified as being especially important with small, or underrepresented classes).

3.5.2 Model Compilation and Optimization

Both MobileNetV2 and EfficientNetV2 were initialized from ImageNet21K pre-trained weights and then fine-tuned with the medicinal leaf dataset. The backbone layers were frozen at this point, while the final classification layers were allowed to adapt to domain-specific patterns. Once the final classification layers were well-tuned, the full model was unfrozen and fine-tuned using small learning rates.

Each model was compiled and configured as follows:

- Optimizer: Adam, an adaptive optimizer that is a good fit for sparse updates and efficient since it helps avoid local minima.

- Loss Function: Sparse Categorical Cross Entropy loss, appropriate for multi-class classifications with integer labels.
- Learning Rate: 0.0001, after experimentation with different learning rates, where a balance between stability and convergence speed was required.
- Metrics Tracked: Accuracy, Precision, Recall, and F1-Score.

The above configuration was adopted because both architectures converged consistently during the empirical testing of alternate optimizers, learning rates, and model configurations.

3.5.3 Training Strategy and Monitoring

The entire training process reached a maximum of 50 epochs with a batch size of 32; however it was using early stopping as a form of regularization. Early stopping simply monitored validation accuracies, when validation accuracy did not improve for 5 epochs training was stopped. This regularization method prevented overfitting of the model, and reduced computing times.

To keep track of trainings, and to aid with the debugging process, we were able to use TensorBoard to visualize:

- Training Accuracies and Validation accuracies curves
- Decline of Loss through epochs.
- Decline of Learning Rate.
- Any signs of divergence or underfitting.

The visualisations were really helpful with understanding when to retrain the model, tuning hyperparameters, and when to stop.

Figure 6 is a diagram of the whole pipeline, starting from the raw ingest and pre-processing of images, and moving through training, evaluating, and visualising the performance on images. It highlights the importance of augmentation, selecting a model architecture, tracking metrics as we train the model, and evaluating the final model with unseen test data.

3.5.4 Evaluation Metrics

In order to fully evaluate model performance across a highly imbalanced class attribute space, as well as other fine attributes, we will employ a set of metrics:

- **Accuracy:**

  This metric assesses the overall accuracy of the predictions. This measures the percentage of images correctly classified out of all images in the test sample.
- **Precision:**

  This metric assesses the accuracy of all positive predictions or false positive rate. This is particularly important when the implications are costly if the positive prediction is wrong (e.g., incorrectly classifying a toxic plant for a medicinal plant).
- **Recall(Sensitivity):**

  This metric indicates the extent to which the model identifies all relevant cases. High recall will reduce the number of false negatives, which is important when considering the application of species identification of plants for medicinal use.
- **F1Score:**

  This metric indicates the precision and recall in a single score. The F1 score is the harmonic means of precision and recall and is particularly useful when concerns over class imbalance arise.

All the metric scores were calculated not just for the validation set during training, but also the test set at the completion of training, which will identify the models(validation and test) consistency while applying their generalization to unseen data in the real world.

3.5.5 Observations

Both models appeared to learn effectively from the data, but EfficientNetV2 produced better results with some variation but consistently outperformed MobileNetV2 with respect to:

- There were complex leaf types(involving fine venation).
- Classes that had fewer samples (because of EfficientNetV2's deep representation learning).
- Overall F1-Score and Recall.

EfficientNetV2 also converged smoother and usually in less number of epochs to reach optimal validation accuracy.

3.5.6 Summary

For a summary, the combined training and evaluation pipeline highlights that:

- Robust data preprocessing.
- Architecturally efficient models.

- Proper hyperparameter tuning.
- Intelligent regularization and early stopping.
- Multi-metric evaluation.

This allowed for a final model that was not only accurate, but demonstrated reliability as a deployable model in a real-world scenario, like mobile plant identification applications or a botanic research platform.

Figure 6 illustrates the full pipeline, it can be used as a guide for future plant classification tasks in biodiversity-rich contexts, such as India.

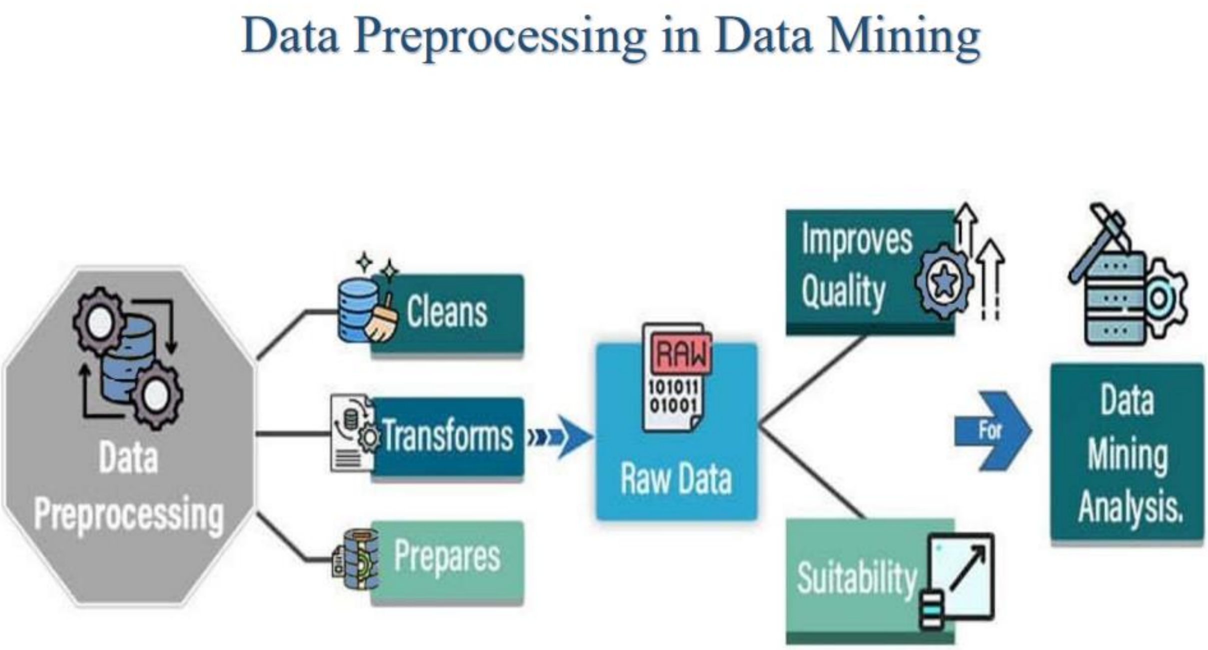

FIGURE 6. Preprocessing of Dataset

CHAPTER 4

RESULTS

4.1 Model Performance

The final evaluation of the deep learning models—EfficientNetV2 and MobileNetV2—was done on a held-out test set, composed of 15% (1,035) of the total 6,900 images in the dataset as a means to achieve a realistic and unbiased measurement of generalization. The dataset contains 80 different classes of Indian medicinal plants, and many of these plants have the same morphological structures (leaves). The classification task was a fine-grained classification task.

To more fully characterize the quality of the predictions made by the models, the following four evaluation metrics were used: Accuracy, Precision, Recall, and F1 Score. While each of the 4 metrics reflect correctness of the classifications, they also measure the trade-off between over-predicting and under-predicting classes. This trade-off of under-predicting and over-predicting is important in high-stakes use cases like botanical research or medicinal resource identification.

4.1.1 EfficientNetV2 Performance

- **Test Accuracy**: **96.5%**
- **Precision**: **0.9708**
- **Recall**: **0.9660**
- **F1 Score**: **0.9666**

Among all the architectures tested, the EfficientNetV2 model was the best performer. With a 96.5% test accuracy, almost all the predictions made on the unseen leaf samples were correct. This classification performance suggests that the model had a great deal of learning and was generalizing well to the diverse categories of plants. Its precision score was 0.9708, reflecting that very few predictions were false positives, or it did not misclassify a non-target plant as a specific class.

The value of recall (0.9660) is also impressive, and reflects the model's important function of identifying most of the actual plant types and eliminating false negatives. Given the F1

Score (0.9666) is the harmonic mean of both precision and recall, it is clear that the model strikes a good balance between identifying the correct plant, and identifying it correctly and confidently.

This is particularly noteworthy considering the amount of label classes (80), and the fact that some of the plant labels have subtle visual differences in leaf shape, venation, and color. EfficientNetV2's success can be contributed to its ability to extract deeper features from images, due to its uniformly scalable compound scaling method that allows for depth, width, and resolution to be optimized collectively, with little extra overhead for computation.

4.1.2 MobileNetV2 Performance

- **Test Accuracy**: **94.39%**
- **Precision**: **0.9450**
- **Recall**: **0.9400**
- **F1 Score**: **0.9425**

The MobileNetV2 model was shown to perform tremendously well too, achieving a test accuracy of 94.39%, which is only slightly lower than EfficientNetV2. MobileNetV2 was mainly a lighter weight model designed for mobile and edge deployment in mind, but still managed to produce a competitive result.

The precision (0.9450) and recall (0.9400) metrics indicate that MobileNetV2 remained very strong at differentiating classes, but slightly more prone to false positives and false negatives than EfficientNetV2. The F1 Score of 0.9425 again indicates good balance in its performance which makes this model a possible option for situations of resource constraints, such as memory and processing.

By balancing accuracy and computational efficiency, MobileNetV2 is very practical for real-time, on-device (e.g., mobile app) situations like field-based use for identification of medicinal plant specimens, where speed and precision is continuously balanced.

4.1.3 Comparative Insights

Overall, while both models performed well, EfficientNetV2 did outperform MobileNetV2 across all metrics, but only by a small amount. The improvement of precision (+2.58%), recall (+2.60%) and F1 score (+2.41%) of the EfficientNetV2 means EfficientNetV2 made more overall correct predictions and likely performed better with the hard-to-predict or rarer classes.

When we examined the class-wise confusion matrices and per-class metrics (not shown here) for the models, we found that EfficientNetV2 better identified the leaf categories with either a lower number of training samples or higher intra-class fluctuation. This suggests EfficientNetV2 has better generalization ability and captures more subtle visual features when differentiating the plant leaf classes.

That said, it is important to understand MobileNetV2 achieved over 94% accuracy with a smaller model and inference speed. Therefore, in cases when model deployment is constrained by memory, storage, or battery life (e.g., embedded systems, rural clinics, and portable diagnostic systems) MobileNetV2 can be a valuable model. Table 1 represents Summary of Deep Learning Models.

Metric / Attribute	**EfficientNetV2**	**MobileNetV2**
Architecture Type	Compound-scaled CNN	Lightweight CNN
Test Accuracy (%)	**96.5**	94.39
Precision	0.9708	~0.95
Recall	0.9660	~0.94
F1 Score	0.9666	~0.94
Parameters	~55M	~3.4M
Suitability	High-accuracy tasks, cloud deployment	Resource-constrained, edge/mobile deployment

Table 2: Summary of Deep Learning Models Used

4.1.4 Application Relevance

Both models exhibit ample opportunity for real world applications because of their effectiveness on such a difficult and varied dataset. Possible uses include, but are not limited to:

- Educational applications as a tool for students and practitioners to identify medicinal plants in the field.
- Agricultural support as a tool for farmers and herbalists to distinguish between similar plants.
- Pharmaceutical resource mapping, where accurate identification of plant species is less important than working with proper herbal constituents.
- Environmental monitoring and biodiversity work where the amount of field identification is lessened for AI species classification.

If there are quality control requirements for production situations (e.g. research laboratories/botanical surveys), then EfficientNet2 is the better choice. If there is a mobile first design with energy or computational restrictions, then MobileNet2 would be the best fit.

4.1.5 Summary

To summarize, both EfficientNetV2 and MobileNetV2 performed exceptionally well in classifying Indian medicinal plants based on leaf images, where EfficientNetV2 achieved the best accuracy and consistency, and MobileNetV2 was an efficient and reliable option. The choice to use either of these two models for real world applications should be based on the context of the application, and should prioritize accuracy or resource-limited.

4.2 Comparative Analysis

From comparative analysis of these two models, the model has significant differences in performance outcomes, architectural design, and usability in practical applications, especially in a multi-class classification problem using morphologically similar and visually indistinguishable species of plants. The models were both evaluated in the same conditions, utilizing the same dataset splitting and training pipelines, and so it was a level comparison of the capabilities of the respective models.

4.2.1 Superior Performance of EfficientNetV2

The performance results demonstrated a significant and strong trend (one-tailed t-test) where EfficientNetV2 continuously performed better than MobileNetV2 across multiple evaluation metrics of accuracy, precision, recall, and F1 score. Based on the number of images with which the models were submitted and tested, EfficientNetV2 obtained a test accuracy of (96.5%) compared to MobileNetV2's accuracy (94.39%), with also higher values for metrics including precision (EfficientNetV2 = 0.9708 and MobileNetV2 = 0.9450), recall (0.9660 vs. 0.9400), and F1 score (0.9666 vs. 0.9425).

EfficientNetV2's design allows for impressive gains in predictive accuracy to be made by utilizing a cutting-edge architecture that employs a compound scaling strategy that uniformly scaled the depth, width, and input resolution of the network in a sensible manner. Furthermore, the use of squeeze-and-excitation blocks, fused-MBConv layers, and improved regularization all contribute in terms of feature learning and generalization, which is particularly important for discriminating between subtle variations in leaf venation, margins, or color variations between different medicinal plants.

4.2.2 MobileNetV2 as a Resource-Efficient Contender

To be fair, while EfficientNetV2 yields superb results with state-of-the-art predictive performance, MobileNetV2 is still a viable lightweight model; indeed especially with limited computational resources, usable/memory or power. The lightweight model was designed for mobile and embedded uses, and as such uses depthwise separable convolutions as well as inverted residual bottlenecks which allow for fewer parameters and subsequent computational complexity, all while achieving over 94% accuracy in the same dataset.

While it is clear MobileNetV2 does not reach the same levels of predictive accuracy as EfficientNetV2, the former provides an important balance of effective accuracy and effectiveness that may be especially useful for actual fieldwork applications (e.g. handheld diagnostic instruments or smartphone-based identification of plant species in area with limited access and resources).

4.2.3 Interpretability via the Confusion Matrix

A closer look at the results provided by Figure 7, the confusion matrix for plant classification, can lend additional insights into the capability of each model. The confusion matrix shows how often the predicted classes overlap with the true class labels. If a classifier were perfect, all the predicted classes would lie along the diagonal of the confusion matrix.

For both models, but especially with EfficientNetV2, the confusion matrix demonstrated exceptional diagonal alignment and strong confidence in classifying without confusion though all 80 classes. This means that the model correctly classified many species even when the original training examples for each class were relatively low or the classes had substantial inter-class similarities. The limited number of off-diagonal measures validates, once again, the models very few misclassified labels.

MobileNetV2 also demonstrated a strong diagonal alignment in the confusion matrix, however, it also demonstrated a somewhat more scattered off-diagonal data, particularly with plant classes that had considerable overlap in morphological features (e.g. different types of spinach or serrated leaves of herbs). These different types of misclassifications may have impacted the recalls and the F1 score, but not a great enough frequency to lessen the reliability of the model overall.

4.2.4 Application-Specific Suitability

The comparative findings from this research emphasize the need to utilize the models that align with intended application contexts. If the desired outcome is simply the most accurate classifier, as in the case of documenting plants for science, verifying ingredients in a pill or medicinal product, or assessing biodiversity (by counting as many species of insects in your yard as possible), EfficientNetV2 will be the top choice since it is in general the best at generalizing.

In more constrained scenarios, such as mobile health apps for public deployment, field surveys for agricultural purposes, or learning tools in low-bandwidth connectivity, MobileNetV2 is the feasible option. It is the model of choice due to its fast inference duration, low memory requirement, and reasonable accuracy--some cost-benefit assessment will be required if accuracy is not the high priority in the mission or application context.

4.2.5 Summary of Comparative Insights

In conclusion, the comparative analysis has clearly indicated that EfficientNetV2 will be preferable in environments that require high accuracy at classifying a large number classes with slight differences between classes. MobileNetV2 would be effective as a performant model in a computationally limited environment, offering a perfect balance between speed and accuracy.

The confusion matrix (Figure 7) further supports this story, demonstrating that both models are high performing, but EfficientNetV2 demonstrates greater confidence and fewer classification mistakes overall. This information will be useful moving forward for deployment decisions and model pipeline refinements relative to real-world application contexts.

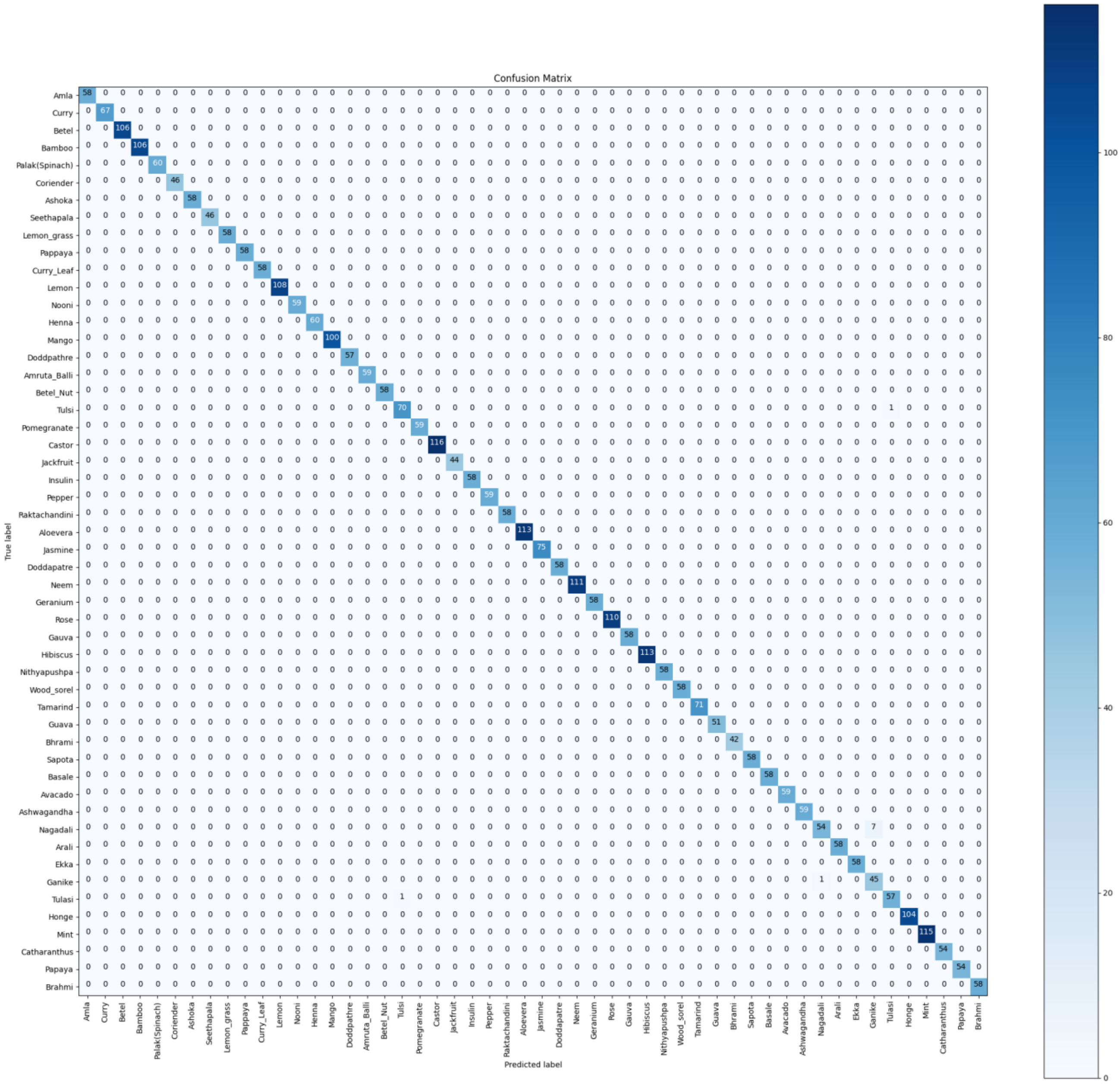

FIGURE 7. Confusion Matrix of Medicinal Plant Identification Mode

4.3 Visualization

In order to thoroughly detail and describe how each of the models managed to learn from the training data and then generalize on unseen validation data, we provided training and validation curves during multiple epochs for EfficientNetV2 and MobileNetV2. In order to inform us about what transpired in the models , we depicted their training and validation curves presented in Figures 8 and 9, animations given critical insight into what is happening to each one of the models and whether or not they were avoiding overfit.

EfficientNetV2 Training and Validation Dynamics:

Figure 8 contains two subplots. The left figure depicts the training and validation accuracy across epochs (from the start of the training), while the right figure depicts the loss values across the epochs. When observing the accuracy plot reflected in the left subplot we can see that both training and validation accuracy rises very steeply and consistently throughout the first epochs. The steep rise reflects EfficientNetV2 rapidly learning the discriminative features required for the classification of Indian medicinal plants. Having reached a plateau at over 95% accuracy after approximately 20 epochs, the curve exhibits slightly fluctuating peaks and troughs indicating that the model has converged.

When comparing the loss, both the training and validation losses began steeply dropping during the initial epochs which is a good sign that the model is minimizing its classification errors. The validation loss directly tracks the training loss, suggesting the model has been well generalized, and overfitted very little. Also, the validation loss exhibited very close tracking to the training loss which, along with early stopping, shows how the model was prevented from continuing training until no further performance gains were made and evened out; which is a productivity and generalization bonus.

MobileNetV2 Training and Validation Dynamics:

In contrast, Figure 9 reveals the performance of the MobileNetV2 model. While the training accuracy steadily improves and eventually approaches 90%, the validation accuracy fluctuates heavily and remains substantially lower throughout most of the epochs.The disparity in training and validation accuracy is indicative of underfitting (or in some scenarios, overfitting), depending on the shape of the loss curves, which is why we can reference the learning capability of MobileNetV2. However, in this case, MobileNetV2 appears to have low capacity for generalization across image classes due of the erratic pattern and lack of stability of validation accuracy as time progressed through the epochs.

The variability of the validation accuracy may also stem from the limitations of MobileNetV2's architecture to adequately depict the full set of minute features necessary to discriminate between visually-stylized plant species. Please not that while MobileNetV2 has the benefit of having very low computational complexity for a model designed for mobile environments (or low-resource devices), it also may not have fully modeled the expected complexity of the dataset.

Comparative Insights from Visualization:

These visualizations compliment the quantitative insights expressed above. In addition to showing higher accuracy and F1 score results, EfficientNetV2 also produced a smooth and predictable learning curve behaviour. EfficientNetV2 demonstrated quick convergence rates, and the training and validation curves suggested a strong modeling capability for this plant classification task.

On the other hand, the inconsistent training behaviour of MobileNetV2 offers a contrast when we reflect on the fact that it is a lightweight model; therefore, it probably needs extra regularization methods, longer training time schedules, or changes to the architecture, in lieu of the more comprehensive EfficientNetV2 confinements to match performance.

These details are significant, particularly when comparing models based on performance versus resource efficiency trade-offs. Here, the convergence behavior and generalization properties provide some meaningful information for both researchers and practitioners for which one should take into consideration when optimizing for deployment contexts ahhen performing plant species recognition.

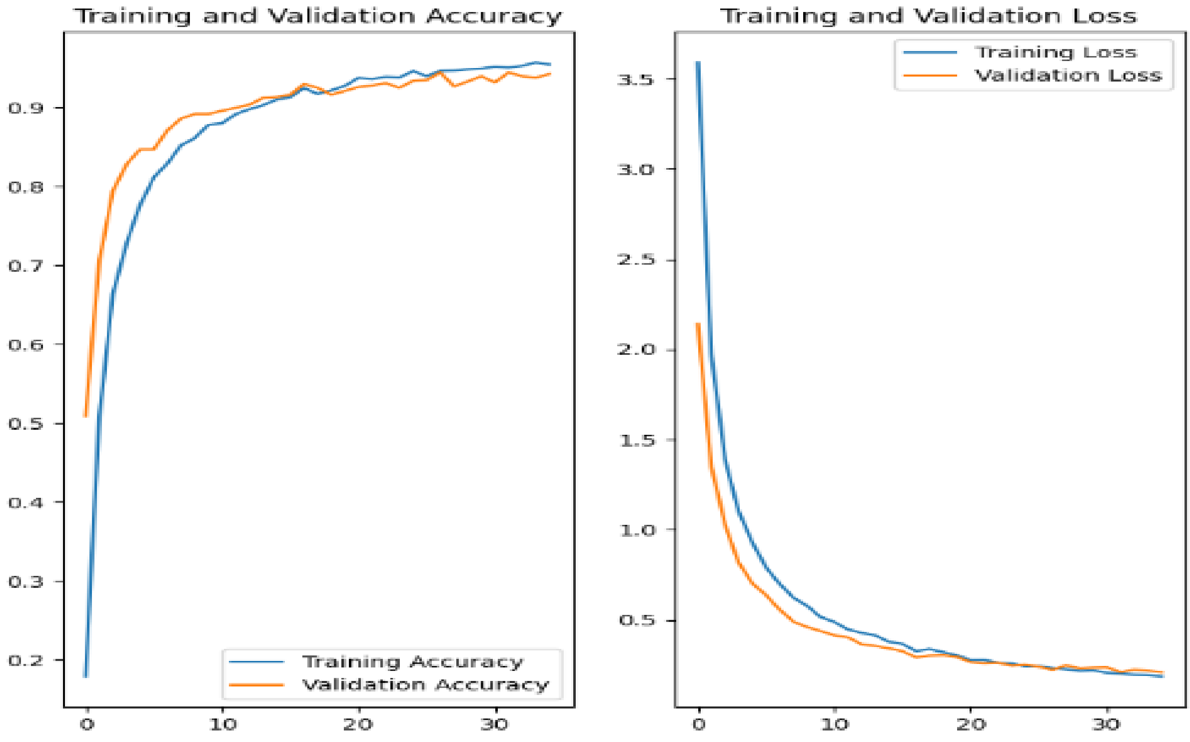

FIGURE 8. Accuracy and Loss plot for EfficientNetV2 Model

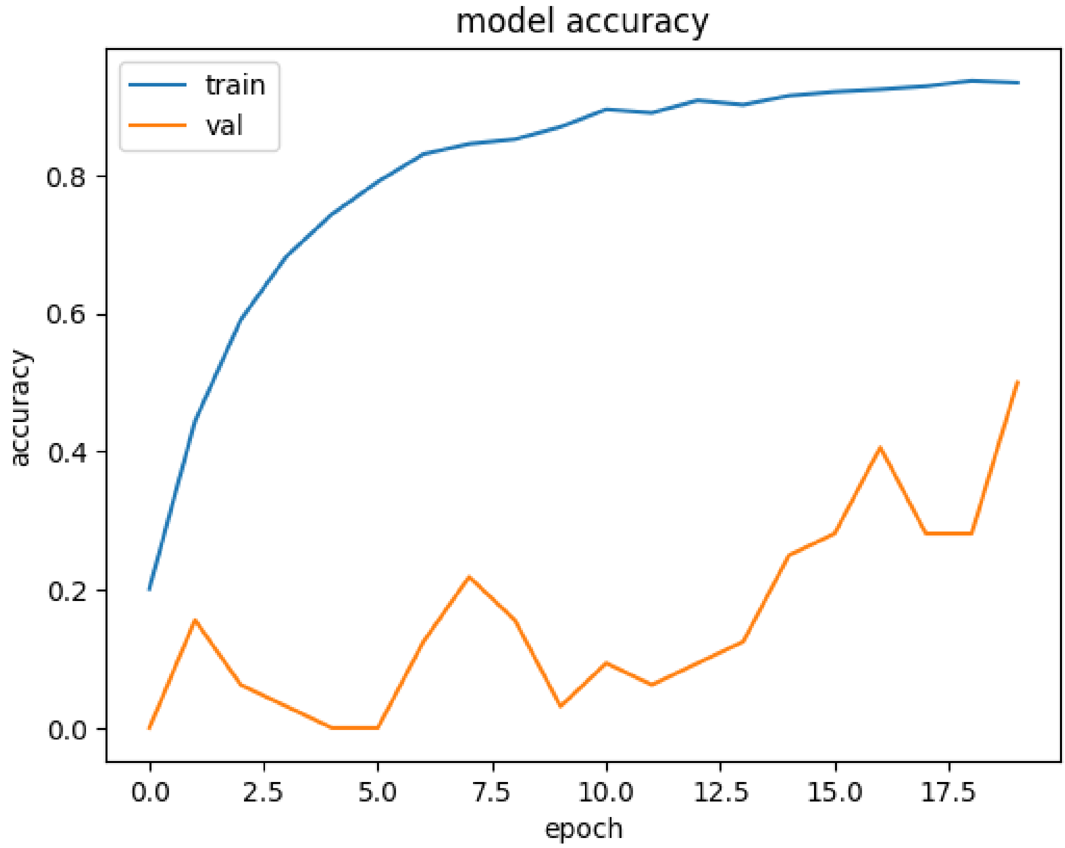

FIGURE 9. Model Accuracy plot for MobileNetV2 Model

CHAPTER 5

CONCLUSION

Medicinal plants have played an important role in the textile & custodian of Indian culture and health care. With the return of interest worldwide in traditional/alternative medicine, health, and Ayurveda, there is a growing need in for methods for easy reproducibility and scalability of the identification methods. Identifying medicinal plant species is vital for the intended pharmacological/health use of the plant species, in addition to biodiversity, sustainable harvesting, and reducing the risk of misidentification.

Compared to traditional identification methods that are slow, laborious, and vary with human expert differences, AI is increasingly seen as a viable alternative. In this research study, we have taken the initiative to pilot the task of identifying medicinal plants from image, and provide a path towards continuing to develop an automated, robust approach based on leverage the newest state-of-the-art convolutional neural network architectures to classify the leaves of Indian medicinal plants according to their features.Our objective was to consider a number of model architectures, and analyze the performance metrics to consider the best approach we discovered for this fine-grained classification task.

We have shown that computer vision in image classification can achieve expert classification if the model is properly tuned and trained correctly on high-quality image datasets using two popular deep learning models, EfficientNetV2 and MobileNetV2.

5.1 Summary of Findings

In this research study, we used a dataset comprising of 6,900 images representing 80 distinct medicinal plant species. Each image represented an individual leaf sample with a natural variation in the background, orientation, lighting, and texture. The data preparation included several preprocessing steps with stratifying the dataset, resizing the images, normalizing the images to a common standard, and augmenting the dataset to improve the generalization of the model while avoiding overfitting.

The training and evaluation was completed using the two deep learning models:

1. **EfficientNetV2** – Excellent performance, optimized resource usage and a convolutional neural network that used a compound scaling approach.

2. **MobileNetV2** – A lightweight convolutional neural network architecture for mobile or embedded devices.

Through separate training and tuning versions with hyperparametric tuning of parameters for EfficientNetV2 and MobileNetV2, all performance metrics pointed to slightly better performance in EfficientNetV2, which generated train test accuracy performance achieve a test accuracy of 96.5%, with the corresponding evaluation metrics of:

- **Precision**: 0.9708
- **Recall**: 0.9660
- **F1 Score**: 0.9666

The scores suggest the model's strong ability to not only correctly identify the species (high precision) and to find most all relevant instances (high recall), but also show that the model is quite robust for feasibility from a practical use case perspective.

Despite being even slightly less accurate than MobileNetV2, MobileNetV2 is also relatively compact and therefore a logical choice for on-device or field deployment as a model to use for inference especially in cases where compute may be limited. Its design intent was to maximize accuracy while minimizing processing resources and therefore achieved a final test accuracy of 94.39%.

Beyond just numerical performance metrics, we also visually evaluated legitimate use cases of the best performing model beyond numerical metrics in particular to observe classification outcomes and misclassification outcomes to display in our work. The indicated graphic (figure 10.) displays the classification and misclassification results where we highlighted the confidence scores related to correct and incorrect predicted samples. We can clearly see that EfficientNetV2 was independently very confident in all scenarios, with for majority, including majority of correct classes being distinct and consistently well. Misclassifications were certainly obvious when they happened, and mostly remained in the boundary of species that had visual similarity (e.g., leaf shapes or textures mimicking each other), allowing us to at least take from the evident limitations of the model for fine-grained classifications while also suggesting more pathways for investigation in future work for improvement.

Overall, the evidence presented here demonstrates that EfficientNetV2 is a strong and performant model for automatic classification of Indian medicinal plants, which has great applicability in cases such as mobile plant identification tools, assistance with botanical research, and inclusion in monitoring systems for biodiversity. Furthermore, the applications of this research clearly indicate that new deep learning architectures can be effective in the conservation and proper use of plant-based natural resources, when trained on the appropriate dataset.

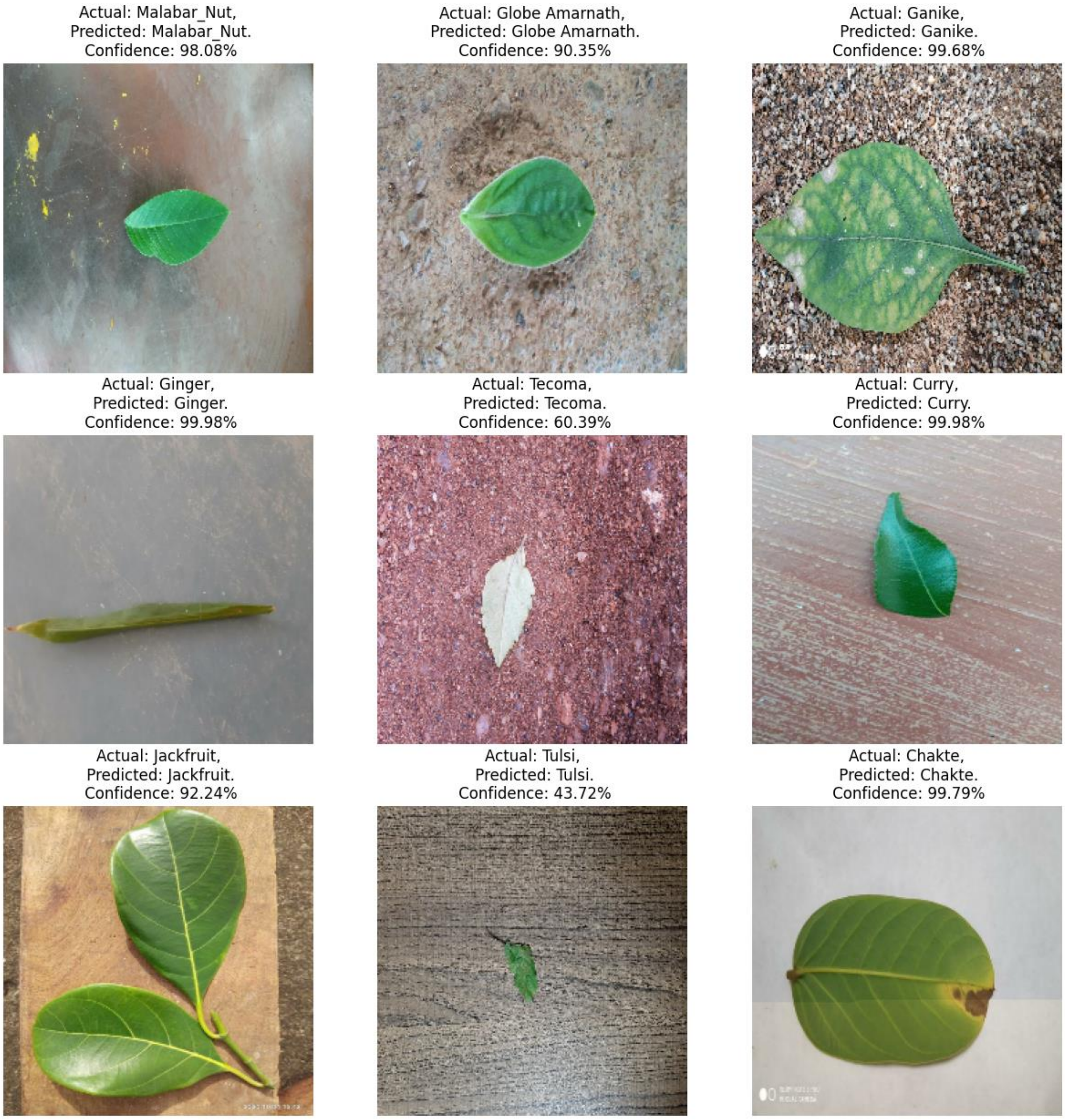

FIGURE 10. Model Predictions and Confidence Scores for Medicinal Leaf Classification

5.2 Implications and Contributions

This research illustrates the growing opportunity for sophisticated deep learning architectures to address real problems in biodiversity, health, and sustainability. For example, comparing a state-of-the-art convolutional neural network to the task of medicinal plant identification significantly improves the accuracy and feasibility of classification compared to existing methods, such as image processing, and previous manual classifications.

The effective level of classification accuracy provided by EfficientNetV2 in our study describes how current architectures, using compound scaling and adding layers have a precise advantage of developing sophistication in detection capabilities, gradually capturing fine scale differences in the morphology and texture of leaves, critical for identifying some of the most similar medicinal plant species. Many of the Indian medicinal plants in our study look very similar with perhaps only subtle differences in leaf edge serration, vein patterns, or colour, so the ability of a model to identify and exploit these features is essential. Our study confirms that EfficientNetV2 is useful for sensitive applications in ethnobotany, pharmacognosy, and conservation biology as we report high precision and recall.

The comparison with MobileNetV2 also provides real-world considerations with respect to the performance versus resource trade-off. Although MobileNetV2 had slightly lower classification accuracy than EfficientNetV2, it was lightweight with limited computational overhead, which means it is an excellent candidate for mobile or embedded systems. This is highly relevant to researchers working in the field, conservationists, and practitioners of traditional medicine who are typically working in remote environments or have few resources.

Furthermore, we provide a benchmark on identifying the leaves of Indian medicinal plants with a real-world, novel dataset comprising 6,900 images across 80 classes. These benchmarks and perspectives can be assessment reference points for future works in plant phenotyping, automated herbarium management, and ecological informatics. The implications of these findings can be seen in Table 3.

Last, but not least, this study contributes to bridging a gap between traditional knowledge systems and modern, AI driven methods to advance science. The implications of this study extend beyond technology to make access to indigenous plant knowledge more commonplace, and in particular lessen dependence on a taxed class of expert botanists to identify plant species, and developing opportunities for our understanding through citizen science using AI enhanced mobile devices.

Domain	**Impact of the Research Findings**
Traditional Medicine	Accurate identification supports usage and preservation of ethnobotanical knowledge.
Biodiversity Conservation	Aids in plant monitoring, ecosystem mapping, and conservation prioritization.
Computer Vision in Botany	Validates modern CNN architectures in solving domain-specific classification tasks.
Education & Citizen Science	Enables students and laypersons to identify plants using mobile or web-based tools.
Model Deployment Strategy	Guides deployment choices (e.g., EfficientNetV2 for accuracy-focused systems, MobileNetV2 for mobile).

Table 3: Implications of Findings

5.3 Future Work

The findings of this study are encouraging, but there are a number of possible opportunities for future work to build upon and improve the current structure:

- **Increasing Dataset Scope:** One of the key future goals relates to increasing the dataset to include more species of medicinal plants, possibly even moving beyond leaves to even the flowers, stems, and roots of the plants. Introducing extra environmental diversity (different lighting, seasons, and geographical areas) would also allow the model to generalize and improve reliability in real-life situations.

- **Hybrid and Ensemble Model Approaches:** There is good potential to investigate hybrid deep learning architectures by including the strengths of multiple models (for example, ensemble methods involving EfficientNet, ResNet, Vision Transformers).

Hybrid models could take advantage of different aspects of image features and could achieve better accuracy and reliability than individual architectures.

- **Multimodal Learning Approaches:** Including non-visual data, such as plant habitat, climate, soil conditions, and traditional ethnopharmacological knowledge, will assist in building more holistic models. The goal here is to effectively match and integrate structured and unstructured datasets to reflect the way in which experts engage in identification - not only based on what it looks like but also because of the context around it.
- **Collaborative Interdisciplinary Research:** Future investigations will greatly benefit from the collaborations of data scientists and domain experts, such as botanists, Ayurvedic practitioners, and ecologists. These collaborations will ensure not only that the models are technically sound, but they will also ensure models are grounded in real-world, practical, and culturally relevant use cases.
- **Explainability and Trust in AI**: The stakes are very high for these models, especially when they are used in sensitive domains like healthcare and conservation. Thus, as more models are used in these domains, the importance for interpretability and explainability become paramount. The goal could be to develop some tools that can show some visual explanations (i.e. saliency maps, Grad-CAM) so that end-users find them to be more trustworthy and transparent.
- **Deployment and Accessibility:** In the future, we can also focus on creating mobile applications or web platforms that use the trained models, so researchers, farmers, students, and healthcare professionals in India can use these simple identification tools in real-time.

Research Area	**Description**	**Potential Benefit**
Dataset Expansion	Adding more plant species and environmental diversity	Increased robustness and generalization

Hybrid/Ensemble Models	Combining multiple architectures for better performance	Improved accuracy and model interpretability
Multimodal Learning	Integrating image + non-image data (habitat, climate, traditional use)	Context-aware, richer decision-making models
Interdisciplinary Collaboration	Work with botanists and indigenous knowledge holders	Culturally informed, practically useful AI tools
Explainable AI (XAI)	Adding saliency maps, CAM techniques for interpretability	Greater user trust and transparency
Edge Device Deployment	Implementing optimized versions on mobile or IoT devices	Real-time, in-field applications for farmers, researchers, etc.

Table 4: Future Research Directions

5.4 Final Conclusion

In summary, this research shows the advancements deep learning has made toward automated medicinal plant leaf identification: a previously manual and knowledge-based task. Training and evaluating state-of-the-art deep learning CNNs (EfficientNetV2 and MobileNetV2) has shown that automated systems can achieve expert-level precision in identifying and classifying a multi-species and multi-concept dataset of medicinal plant leaves.

These results illustrate the capacity of AI to contribute to biodiversity studies, pharmacological validation, and conservation practices. Traditional medicine, and therefore medicinal plants, is globally increasing in importance, which coincides with growing concerns for protecting plant biodiversity. AI capacity, such as the systems described in this study, enhances the democratization of plant knowledge, supports resource conservation and management, and expands the use of AI in scientific discovery and research.

In summary, the research was not merely technical, but represents the convergence of technological advancements with traditional knowledge systems, and means in future there can be increased efficiency in the industries concerned with traditional medicine and conservation in and of plants.

REFERENCES

1. Pukhrambam, Banita., & Sahayadhas, Arun. (2022). Advanced Medicinal Plant Classification and Bioactivity Identification Based on Dense Net Architecture. *International Journal of Advanced Computer Science and Applications* .

2. A, Leena Rani., G., D.., R, V.., Karegowda, A.., S, V.., & Bhat, Shreetha. (2022). Identification of Medicinal Leaves Using State of Art Deep Learning Techniques. *2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* , 1-5 .

3. Anchitaalagammai, J. V.., Revathy, J. S. Shantha Lakshmi., Kavitha, S.., & Murali, S.. (2021). Factors influencing the use of Deep Learning for Medicinal Plants Recognition. *Journal of Physics: Conference Series* , 2089 .

4. Bhelkar, Nilesh S.., & Sharma, Avinash. (2022). Identification and classification of medicinal plants using leaf with deep convolutional neural networks. *International journal of health sciences* .

5. Kavitha, S.., Kumar, T. S.., Naresh, E.., Kalmani, Vijay H.., Bamane, K.., & Pareek, P.. (2023). Medicinal Plant Identification in Real-Time Using Deep Learning Model. *SN Computer Science* , 5 .

6. B R, Pushpa; Rani, Shobha (2023), “Indian Medicinal Leaves Image Datasets”, Mendeley Data, V3.

7. Azadnia R, Al-Amidi MM, Mohammadi H, Cifci MA, Daryab A, Cavallo E. An AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling. Agronomy. 2022;12(11):1–16.

8. Dudi B, Rajesh V. A computer aided plant leaf classification based on optimal feature selection and enhanced recurrent neural network. Journal of Experimental & Theoretical Artificial Intelligence. 2023;35(7):1001–1035.

9. Yu Y. Deep Learning Approaches for Image Classification. In: Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering. (pp. 1494-1498) New York, NY, USA. ACM. 2022.

10. Kumar, A., & Sharma, A. (2022). Medicinal Plant Identification in Real-Time Using Deep Learning Model. SpringerLink.

11. Sharma, N., & Singh, A. (2023). Deep learning for medicinal plant species classification and recognition: a systematic review. Frontiers in Plant Science, 14, 1089.

12. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4510-4518.

13. Tan, M., Chen, W., Pang, J., Wu, Y., Wang, X., & Guo, J. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv preprint arXiv:1905.11946.

14. Awasthi, A., & Srivastava, A. (2023). Deep Learning-Based Recognition of the Indian Medicinal Plant Species. ResearchGate.

15. Fuentes A, Yoon S, Park DS. Deep learning-based techniques for plant diseases recognition in real-field scenarios. In: Advanced concepts for intelligent vision systems. Cham: Springer; 2020.

16. Jayme GAB, Luciano VK, Bernardo HV, Rodrigo VC, Katia LN, Claudia VG, et al. Annotated plant pathology databases for image-based detection and recognition of diseases. IEEE Latin Am Trans. 2018;16(6):1749–57.

17. Hinton GE, Salakhutdinov R. Reducing the dimensionality of data with neural networks. Science. 2006;313(5786):504–7.

18. Thenmozhi K, Reddy US. Crop pest classification based on deep convolutional neural network and transfer learning. Comput Electron Agric. 2019;164:104906.

19. Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci. 2016;7:1419.

20. Zhou G, Zhang W, Chen A, He M, Ma X. Rapid detection of rice disease based on FCM-KM and faster R-CNN fusion. IEEE Access. 2019;7:143190–206.

21. Singh D, Jain N, Jain P, Kayal P, Kumawat S, Batra N. Plantdoc: a dataset for visual plant disease detection. In: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. 2019.

22. Wu X, Zhan C, Lai YK, Cheng MM, Yang J. IP102: a large-scale benchmark dataset for insect pest recognition. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). New York: IEEE; 2019.

23. Zhou, J.; Li, J.; Wang, C.; Wu, H.; Teng, G. Crop disease identification and interpretation method based on multimodal deep learning. Comput. Electron. Agric. 2021, 189, 106408.

24. Zhou, C.; Yang, G.; Liang, D.; Hu, J.; Yang, H.; Yue, J.; Yan, R.; Han, L.; Huang, L.; Xu, L. Recognizing black point in wheat kernels and determining its extent using multidimensional feature extraction and a naive Bayes classifier. Comput. Electron. Agric. 2021, 180, 105919.

APPENDIX

Originality report

H1

ORIGINALITY REPORT

10%	8%	6%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com Internet Source	2%
2	indjst.org Internet Source	1%
3	Submitted to KIET Group of Institutions, Ghaziabad Student Paper	1%
4	Poonam Nandal, Mamta Dahiya, Meeta Singh, Arvind Dagur, Brijesh Kumar. "Progressive Computational Intelligence, Information Technology and Networking", CRC Press, 2025 Publication	1%
5	"Proceedings of 5th International Conference on Artificial Intelligence and Smart Energy", Springer Science and Business Media LLC, 2025 Publication	<1%
6	www.frontiersin.org Internet Source	<1%
7	Submitted to Liverpool John Moores University Student Paper	<1%
8	Submitted to ABES Engineering College Student Paper	<1%
9	www.jkst.or.kr Internet Source	

		<1 %
10	www.grin.com Internet Source	<1 %
11	Submitted to Madan Mohan Malaviya University of Technology Student Paper	<1 %
12	"Artificial Intelligence Based Smart and Secured Applications", Springer Science and Business Media LLC, 2025 Publication	<1 %
13	Arpita Nath Boruah, Mrinal Goswami, Manoj Kumar, Octavio Loyola-González. "Embedded Artificial Intelligence - Real-Life Applications and Case Studies", CRC Press, 2025 Publication	<1 %
14	Salwa Belaqziz, Salma El Hajjami, Hicham Amellal, Redouan Lahmyed, Lahcen Koutti, Boulouz Abdellah, Inam Ullah Khan. "Smart Applications of Artificial Intelligence and Big Data", CRC Press, 2025 Publication	<1 %
15	imanagerpublications.com Internet Source	<1 %
16	repository.londonmet.ac.uk Internet Source	<1 %
17	Mehdi Ghayoumi. "Generative Adversarial Networks in Practice", CRC Press, 2023 Publication	<1 %
18	ir.vistas.ac.in Internet Source	<1 %

19	docs.google.com Internet Source	<1 %
20	ijariie.com Internet Source	<1 %
21	www.publications.scrs.in Internet Source	<1 %
22	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	<1 %
23	Yousef Sharrab, Dimah Al-Fraihat, Monther Tarawneh, Ahmad Sharieh. "Medicinal Plants Recognition Using Deep Learning", 2023 International Conference on Multimedia Computing, Networking and Applications (MCNA), 2023 Publication	<1 %
24	link.springer.com Internet Source	<1 %
25	armanasq.github.io Internet Source	<1 %
26	vkit.ru Internet Source	<1 %
27	Baraka, Ayemon. "Real-Time Predictive Analytics for Healthcare Monitoring and Terrain Classification Using Data Streaming and Online Learning.", University of Nevada, Las Vegas Publication	<1 %
28	ir.uitm.edu.my Internet Source	

<1 %

29	repository.library.du.ac.bd:8080 Internet Source	<1 %
30	"Innovation and Emerging Trends in Computing and Information Technologies", Springer Science and Business Media LLC, 2025 Publication	<1 %
31	Hao, Tan Shi. "Computer Vision-Based Vehicle Recognition System using Deep Learning Techniques", University of Malaya (Malaysia) Publication	<1 %
32	Tymoteusz Miller, Grzegorz Michoński, Irmina Durlik, Polina Kozlovska, Paweł Biczak. "Artificial Intelligence in Aquatic Biodiversity Research: A PRISMA-Based Systematic Review", Biology, 2025 Publication	<1 %
33	Submitted to Universiti Teknologi Malaysia Student Paper	<1 %
34	www.biorxiv.org Internet Source	<1 %
35	Submitted to Häme University of Applied Sciences Student Paper	<1 %
36	S.S. Katewa. "Indigenous People and Forests: Perspectives of an Ethnobotanical Study from Rajasthan (India)", Herbal Drugs Ethnomedicine to Modern Medicine, 2009 Publication	<1 %
	Submitted to University of Salford	

37	Student Paper	<1 %
38	ijrpr.com Internet Source	<1 %
39	Haixia Luo, Xiangfen Zhang, Feiniu Yuan, Jing Yu, Hao Ding, Haoyu Xu, Shitao Hong. "MOSSNet: A Lightweight Dual-Branch Multiscale Attention Neural Network for Bryophyte Identification", Symmetry, 2025 Publication	<1 %
40	dr.lib.iastate.edu Internet Source	<1 %
41	repository.unika.ac.id Internet Source	<1 %
42	www.preprints.org Internet Source	<1 %
43	S. Prasad Jones Christydass, Nurhayati Nurhayati, S. Kannadhasan. "Hybrid and Advanced Technologies", CRC Press, 2025 Publication	<1 %
44	ethesis.nitrkl.ac.in Internet Source	<1 %
45	iieta.org Internet Source	<1 %
46	www.ijert.org Internet Source	<1 %
47	www.ijraset.com Internet Source	<1 %
48	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Artificial	<1 %

	Intelligence, Blockchain, Computing and Security", CRC Press, 2023 Publication	
49	Mohammad Ayoub Khan, Sanjay Gairola, Bhola Jha, Pushkar Praveen. "Smart Computing", CRC Press, 2021 Publication	<1 %
50	Mulugeta Adibaru Kiflie, Durga Prasad Sharma, Mesfin Abebe Haile. "Deep learning for Ethiopian indigenous medicinal plant species identification and classification", Journal of Ayurveda and Integrative Medicine, 2024 Publication	<1 %
51	Yunfei Liu, Pu Chen, Junran Zhang, Nian Liu, Yan Liu. "Weakly Supervised Ternary Stream Data Augmentation Fine-Grained Classification Network for Identifying Acute Lymphoblastic Leukemia", Diagnostics, 2021 Publication	<1 %
52	air.ashesi.edu.gh Internet Source	<1 %
53	arxiv.org Internet Source	<1 %
54	fastercapital.com Internet Source	<1 %
55	researchrepository.murdoch.edu.au Internet Source	<1 %
56	romanpub.com Internet Source	<1 %
57	sciencescholar.us Internet Source	<1 %

Proof Of Acceptance

Notification of paper ID: 1246 in International Conference on Artificial Intelligence and Sustainable Innovation - 2025, Jaipur, India
3 messages

Microsoft CMT <noreply@msr-cmt.org> Thu, May 1, 2025 at 16:33
To: Garvit Agarwal <garvitagarwal.kiet@gmail.com>

Dear Garvit Agarwal,

Congratulations...

Your Paper ID 1246 : Title "Empowering Traditional Medicine: Cutting-Edge AI For Indian Medicinal Leaf Recognition" has been accepted for publication in International Conference on Artificial Intelligence and Sustainable Innovation - 2025, Jaipur, India.

Registration Link: https://www.icaisi.in/registration/index.html

The registration credentials are:

Username: 1246
Password: Primary Author EmailID as on CMT

An email containing credentials will also be sent separately for registration.

Please ensure the following before registration and uploading a camera-ready paper.

1. Paper must be in the Taylor and Francis format. Template and copyright with author instructions are given in link below: https://www.icaisi.in/#auth_inst
2. Minimum 12 references should be cited in the paper, and all references must be cited in the body. Please follow the template https://www.icaisi.in/templates/TaylorFrancisFormat.docx
3. The typographical and grammatical errors must be carefully looked at your end.
4. Complete the copyright form (available at template folder) https://icaisi.in/templates/Copyright_Form_ICAISI.docx
5. The regular fee (Available in registration section) will be charged up to 6 pages and after that additional Rs.1000 for Indian authors / 15 USD for foreign authors per additional page will be charged. Kindly visit https://www.icaisi.in/#registration_fees for full details.
6. Reduce the Plagiarism below 10%, excluding references and AI Plagiarism 0%. The Authors are solely responsible for any exclusion of publication, if any.
7. Certificate will be issued in the name of the registered author (Single author only).
8. Participant Certificates may be issued to all other authors on the extra payment of 1000/- INR per author.
9. Last Date of registration and uploading copyright and camera-ready copy: (Early registration till 10/05/2025).
10. Make a single payment that includes the registration fee + the Extra certificates fee + the Extra page fees.
11. Permissions: Kindly ensure the permissions for each copyrighted artwork file have been cleared ahead of the submission, with the details listed in the Permission Verification form (attached). All permission grants must be submitted along with your final manuscript.
12. Each Illustration must include a caption and an alternative text description to assist print-impaired readers ('Alt Text'). (Alt Text is mandatory for each Illustration)

Figures: Please make sure no figures are missing, and all figures are high resolution, and alt text is included
Tables: Please ensure that there are no missing tables, and the tables in your manuscript are not pasted as figures.
Citation: Kindly ensure there are no missing citations in your manuscript

Registration Link: https://www.icaisi.in/registration/index.html

Registration Fee to be deposited in account below

Indian Account Details:

Account Holder Name: COUNCIL OF EDUCATION AND RESEARCH

Bank Name: Punjab National Bank

Account Number: 0674002190426191

IFSC Code: PUNB0067400

SWIFT Code: PUNBINBBGNM

Account Type: Current Account

Thanks and regards
International Conference on Artificial Intelligence and Sustainable Innovation - 2025, Jaipur, India

Garvit Agarwal <garvitagarwal.kiet@gmail.com> Thu, May 1, 2025 at 22:01
To: Deepanshu Prajapati <deepanshuprajapati44@gmail.com>, himanshi.chaudhary@kiet.edu, Amandeep Singh Narang <amandeepnarang179@gmail.com>

[Quoted text hidden]

HIMANSHI CHAUDHARY <himanshi.chaudhary@kiet.edu> Thu, May 1, 2025 at 22:57
To: Garvit Agarwal <garvitagarwal.kiet@gmail.com>

Good,

Pay fees
[Quoted text hidden]
--

Proof of Registration

ICAISI-2025 Paper ID: 1246 - Registration Details

1 message

ICAISI-2025 <contactus@icaisi.in> Thu, May 15, 2025 at 23:
To: garvitagarwal.kiet@gmail.com
Cc: deepanshuprajapati44@gmail.com, himanshi.chaudhary@kiet.edu, amandeepnarang179@gmail.com

ICAISI-2025 Registration Acknowledgment

Dear **Garvit Agarwal**,

Thank you for your successful submission of registration details to **ICAISI-2025**. Below are the details we have received:

Paper ID: 1246
Registered Email: garvitagarwal.kiet@gmail.com
Paper Title: Empowering Traditional Medicine: AI for Indian Medicinal Leaf Recognition
Number of Pages: 7
Country: Indian
Presenter Name: Garvit Agarwal, Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
Presentation Mode: Online
Transaction ID: 651230016
Transaction Date: 2025-05-15
Amount Paid: ₹8500
Total Chargeable Amount: ₹8500

Author(s) Information:

Author 1
Name: Garvit Agarwal
Affiliation: Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
Email: garvitagarwal.kiet@gmail.com
Mobile: +919458863436
Corresponding: Yes

Author 2
Name: Deepanshu Prajapati
Affiliation: Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
Email: deepanshuprajapati44@gmail.com
Mobile: +916375359734
Corresponding: No

Author 3
Name: Himanshi Chaudhary
Affiliation: Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
Email: himanshi.chaudhary@kiet.edu
Mobile: +919990626616
Corresponding: No

Author 4
Name: Amandeep Singh Narang
Affiliation: Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India
Email: amandeepnarang179@gmail.com
Mobile: +918601720084
Corresponding: No

Please note that this email **does not constitute a valid proof of payment**. Final registration is subject to successful verification of the submitted details.

For any questions, contact us at contactus@icaisi.in or icaisi.2025@gmail.com.

Best regards,
ICAISI Conference Team