IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

Medicinal Plant Identification using Machine Learning

Dr T. Vigneswari[1], K. Sibi[2], B. Sudhir[3], A. Vanthiyath Thevan[4], K. Yogeshwaran[5]
Assistant Professor, Department of Computer Science and Engineering[1]
Student, Department of Computer Science and Engineering[2,3,4,5]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamil Nadu, India

Abstract: *Medicinal plants play a crucial role in traditional and modern medicine due to their properties. In this study, we propose the identification of medicinal plants using machine learning. The proposed system comprises several key components: data collection, feature extraction, model training, and prediction. Initially, a comprehensive dataset of medicinal plant images is gathered from databases. Next, relevant features such as leaf shape, texture, and colour are extracted from images using image processing techniques. Subsequently, ML algorithms such as Convolutional Neural Networks (CNNs) are employed to train a classification model on the extracted features. The trained model is then capable of accurately identifying medicinal plants from input images. To facilitate user interaction, a user-friendly interface may be developed, allowing users to upload images and receive instant identification results. Furthermore, the system's performance is evaluated in real-world scenarios to assess its practical utility and reliability. Overall, the proposed automated medicinal plant identification system represents a significant advancement in leveraging ML technology to streamline the identification process.*

Keywords: Machine learning, Convolutional Neural Networks

I. INTRODUCTION

The identification of medicinal plants has long been a laborious and error-prone task, often relying on manual observation and expert knowledge. However, this traditional approach is fraught with challenges such as subjectivity, time consumption, and limited scalability. As a result, there is a pressing need for automated systems that can streamline the identification process while maintaining accuracy and reliability. Enter machine learning, a branch of artificial intelligence that empowers computers to learn from data and make predictions without explicit programming. By leveraging ML algorithms, we can analyze vast datasets of botanical images, extract intricate patterns, and develop models capable of identifying medicinal plants with remarkable precision. Identifying medicinal plants accurately is crucial for various domains, from traditional medicine preservation to drug discovery and biodiversity conservation.

Conventional methods reliant on expert knowledge are often time-consuming and subjective. In response, the integration of machine learning techniques presents a promising solution, automating identification processes through computational algorithms trained on extensive datasets of plant images and associated metadata.Furthermore, machine learning can enhance quality control measures in the herbal medicine industry, ensuring the authenticity and safety of plant-based products .

Despite challenges such as limited data availability and species variation, the integration of machine learning holds immense potential for advancing medicinal plant identification, with implications across healthcare, pharmacology, and conservation biology. This paper aims to explore the current state-of-the-art in machine learning-based medicinal plant identification, shedding light on recent advancements, challenges, and future research directions. Identifying medicinal plants accurately is a critical task with wide-ranging implications across various fields, including traditional medicine, pharmacology, and conservation biology.

Traditional methods of identification often rely on the expertise of botanists and ethnobotanists, leading to inconsistencies and subjectivity. Moreover, the rapid loss of traditional knowledge and the threat of species extinction underscore the urgency of developing efficient and reliable identification techniques. In recent years, the emergence of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

machine learning has revolutionized the field of plant identification. envelopment of automated, data-driven approaches to plant identification.

II. HISTORY OF MEDICINAL PLANT IDENTIFICATION

The history of medicinal plant identification using machine learning is a convergence of ancient herbal wisdom and modern computational techniques. Initially, traditional healers and botanists relied on manual observation and classification methods to identify medicinal plants based on their physical characteristics and healing properties. However, with the advent of machine learning algorithms, particularly in the latter half of the 20th century, a paradigm shift occurred. Researchers began exploring the application of computational models to automate the identification process.

Early efforts involved the development of image recognition systems trained on extensive botanical datasets. These systems utilized techniques such as feature extraction and pattern recognition to distinguish between different plant species. Over time, advancements in machine learning algorithms, coupled with the availability of large-scale botanical databases and high-resolution imaging technologies, greatly enhanced the accuracy and efficiency of medicinal plant identification. Today, machine learning-powered applications and tools are increasingly utilized by botanists, pharmacologists, and traditional medicine practitioners to expedite the identification and authentication of medicinal plants, thus facilitating drug discovery, conservation efforts, and the preservation of indigenous knowledge

Early efforts focused on digitizing botanical data and developing algorithms capable of recognizing plant species based on visual cues. As machine learning algorithms advanced, especially with the rise of deep learning models like convolutional neural networks, the accuracy and efficiency of plant identification improved significantly. This progress facilitated interdisciplinary collaborations between botanists, computer scientists, and traditional medicine practitioners, leading to the development of practical applications such as mobile apps for field identification and quality control systems for herbal medicine production. Today, machine learning continues to play a crucial role in preserving traditional knowledge, conserving biodiversity, and unlocking the therapeutic potential of medicinal plants for global health and well-being

III. RELATED WORKS

Abhishek Gokhale, Sayali Babar, Srushti Gawade(B) , and Shubhankar Jadhav(2020)Medicinal plants have long been utilized in traditional medicine. Identification of medicinal plants is a very challenging task without external resources or assistance. Identification of the right medicinal plants that are used for the preparation of medicines isimportant in the medicinal industry. [6].

Adams Begue, Venitha Kowlessur ,Upasanasingh, Fawzi Mahomoodally (2017) Systems developed so far use varying number of steps to automate the process of automatic classification, though the processes are quite similar. Essentially, these steps involve preparing the leaves collected, undertaking some pre-processing to identify their specific attributes, classification of the leaves, populating the database, training for recognition and finally evaluating the results [2].

R.Kiruthika, M.Kousalya, Mr. C. Rathnakumar(2022) A prepared Botanist searches for every one of the accessible elements of the plants, for example, leaves, blossoms, seeds, roots and stem distinguish plants. Aside from the leaf, all others are 3D items and increment the intricacy of examination by PC. Notwithstanding, plant passes on are 2D articles and convey adequate data to distinguish the plant. Leaves can be gathered effectively and picture obtaining might be done utilizing economically advanced cameras, cell phones, or record scanners [3].

Anil kumar, Dr. Brinder kumar(2020) In this paper we are concerned with different types of medicinal plants, it is motivating and encouraging for us to handle the classification of medicinal plants in this paper medicinal plants can be identified based on their leaves are considered based on their leaves, flowers, fruits are planted as a whole. Among all the modalities, leaves are considered as a major and promising modality for effective classification of medicinal plants. [4].

Nayana G. Gavhala, Dr. A. P. Thakare (2020) In the ancient past, the Ayurvedic physicians themselves picked the medicinal plants and prepared the medicines for their patients. Most of the plants are identified using their leafs the common steps to classify the leaf of a plants are Capturing image, noise removal and resizing , Extracting features, use proposed methodology and finally identify or recognized the plant. [1].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

G. KAYHAN, E. ERGÜN (2020) In this study, different machine learning (ML)methods were used to classify medicinal and aromatic plants(MAP) namely St. John's wort (Hypericum perforatum L.),Melissa (Melissa officinalis L.), Echinacea (Echinacea purpureaL.), Thyme (Thymus sp.) and Mint (Mentha angustifolia L.)based on leaf shape, gray and fractal features. [5].

D.M.C.Dissanayake, W.G.C.W. Kumara (2021) Classical plant identification process is timeconsuming and complicated. On the other hand, knowledge of plants and the ability to identify the plant species are depleting through generations. This lack of knowledge and drawbacks of manual identification were the underlying causes to develop this study. [8].

Ke Dongs and Yihan Ruan (2020) Classical plant identification process is timeconsuming and complicated. On the other hand, knowledge of plants and the ability to identify the plant species are depleting through generations. This lack of knowledge and drawbacks of manual identification were the underlying causes to develop this study.[7]

Ms. Ridhanya[1], Ms. Kalpana Chawla S[2], Ms. Shahin Nisha S[3], Mr. R. Rajkumar[4] **(2024)** There are thousands of plant species in the globe, and many of them have therapeutic uses. In medical customs,medicinal plants are often used. One of the earliest systems of medicine that is still in use today is Ayurveda.Therefore, accurate identification of medicinal plants has significant advantages for the government, the public,physicians, life scientists, forest department personnel, pharmaceutical laboratories, and the manufacturing ofmedications. [9].

IV. SYSTEM ARCHITECTURE

The proposed system architecture for medicinal plant identification using a Convolutional Neural Network (CNN) based on the Mobile Net model involves several components as in figure1. Obtain a large dataset of images containing various medicinal plants. Ensure that the dataset is diverse and representative of different species, variations, and conditions. Split the dataset into training, validation, and test sets. Evaluate the performance of the trained model using metrics such as accuracy, precision, recall. Once the model is trained and evaluated satisfactorily, deploy it to a production environment. Develop a user-friendly interface that allows users to upload images of medicinal plants for identification. Display the results of the identification process the user

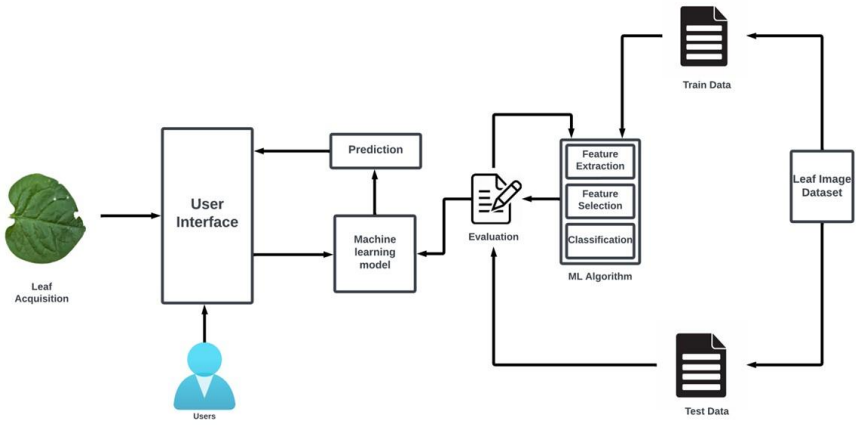
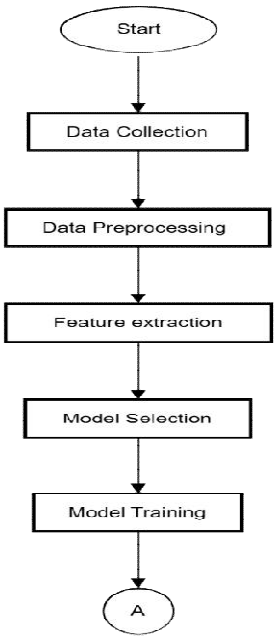


Fig. 1. System Architecture for Medicinal Plant Identification

The steps involved in the machine learning algorithm are depicted in figure 2. In first step, data related to plants is collected. Preprocessing steps may include resizing images to enhance the quality and quantity of the data. Convolutional Neural Network (CNN) MobileNetV2 model is a deep learning architecture commonly used for image classification tasks. In this step, the extracted features are fed into the MobileNetV2 model for training. The CNN MobileNetV2 model is trained using the extracted features and corresponding labels from the training dataset. Once the model is trained, it needs to be evaluated to assess its performance and generalization ability. This step involves testing the trained model on a separate validation dataset to measure metrics such as accuracy, precision, recall, etc. After successful evaluation, the trained model is deployed into a production environment where it can be used for making

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

predictions on new, unseen data. Users input images of plants through the user interface. Depending on the prediction made by the model, the user interface displays information about the predicted plant species if it's recognized. If the plant is not recognized, an appropriate message indicating "Plant not found" is displayed..



if model
the p

Figure 2:Steps involved in plant identification using Machine learning algorithm

VI. IMPLEMENTATION AND PERFORMANCE EVALUATION

The implementation of the above-described process involves the following steps

Data Collection & Preprocessing. It involves gathering a diverse dataset of images showcasing different medicinal plants. This dataset should cover various angles, lighting conditions, and backgrounds to ensure robust model training. After user gives the input image, the images are preprocessed by resizing them to a standardized size (typically 224x224 pixels) and normalizing the pixel values to a common scale (usually between 0 and 1).

Model Training. the focus is on training a robust model for medicinal plant identification. First, the dataset is divided into three subsets: training, validation, and testing, typically using an 80-10-10 split ratio. Then, MobileNetV2, a pre-trained convolutional neural network (CNN) architecture known for its efficiency, is fine-tuned on the training dataset. This involves adjusting the network's weights to learn features specific to medicinal plants. Throughout training, the model's performance is regularly evaluated on the validation set.

Performance Evaluation. the trained model's performance is assessed using a separate testing dataset to measure its accuracy, precision, recall, and F1-score. These metrics quantify how well the model identifies medicinal plants. Visualizing confusion matrices helps identify which plant species are often confused, guiding improvements. This evaluation ensures the model meets quality standards before deployment

Response Generation and Presentation completes the process. Using the retrieved information from the database, the chatbot generates tailored responses to the user's queries. These responses are formulated to address the specific query category and are presented to the user in a user-friendly format.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

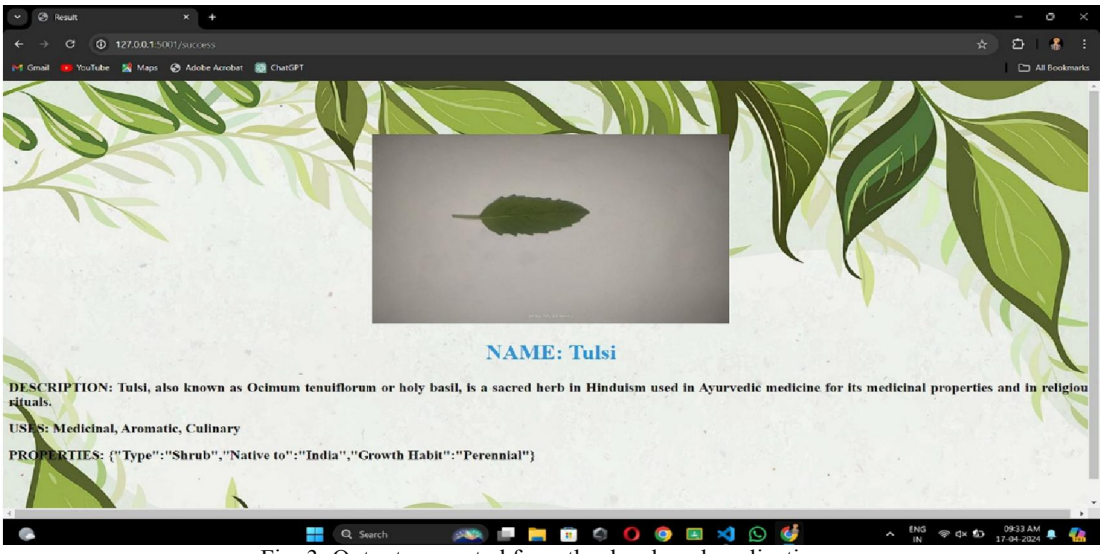


Fig. 3: Output generated from the developed application

The application developed receives the image of an unknown medicinal plant and provides the accurate details of the same after appropriate identification by using CNN algorithm. It is shown in figure 3

Performance Evaluation:

The performance of the proposed algorithm is compared with PNN and RandomForest. The results are as depicted in figure 4.The proposed algorithm works efficiently compared to the other two algorithms.

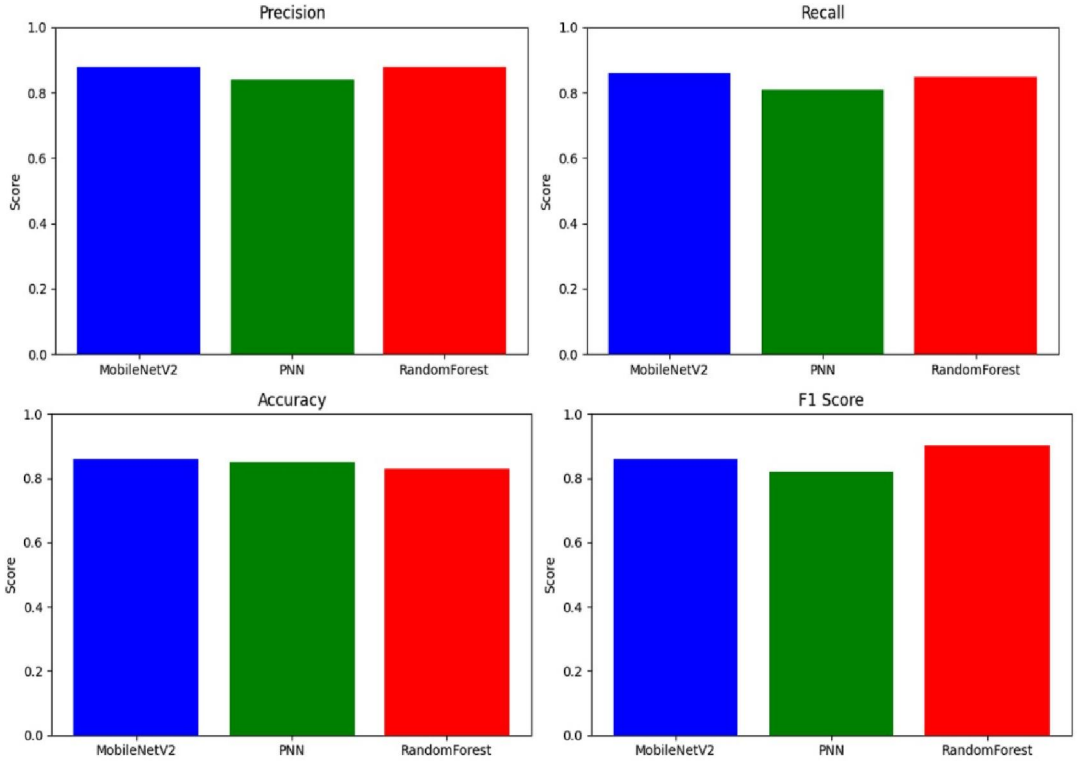


Fig.4. Performance Evaluation of MobileNetV2,PNN,RandomForest

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

VII. CONCLUSION

The integration of machine learning into medicinal plant identification represents a significant leap forward in the fields of healthcare, conservation, and cultural heritage preservation. Its ability to process vast amounts of data with accuracy and efficiency not only streamlines identification processes but also opens up new avenues for research and application. As we look to the future, further enhancements in machine learning models and technologies will continue to refine the accuracy and scope of medicinal plant identification systems. This includes improving algorithms, expanding datasets, integrating with remote sensing technologies, and developing user-friendly applications. Additionally, addressing ethical considerations surrounding data usage and ownership will be crucial for ensuring responsible deployment of these technologies.

The future of medicinal plants identification using machine learning presents an exciting frontier in both scientific research and practical applications. As machine learning algorithms become increasingly sophisticated and data collection methods more comprehensive, the potential for accurately identifying medicinal plants grows exponentially. This technology enables the development of robust models that can analyze vast amounts of botanical data, including images, chemical compositions, genomic sequences, and geographical information. By leveraging these diverse datasets, researchers can train machine learning models to recognize intricate patterns and associations, leading to more precise identification of medicinal plants and their bioactive compounds.

Moreover, the integration of machine learning with other emerging technologies such as computer vision, natural language processing, and deep learning offers new avenues for advancing medicinal plant research. For instance, computer vision techniques can aid in the automatic recognition of plant features from images, while natural language processing algorithms can extract valuable information from traditional medicine texts and ethnobotanical literature. By combining these capabilities, researchers can create comprehensive databases and knowledge repositories that streamline the identification process and enhance our understanding of the medicinal properties of plants.

Furthermore, the application of machine learning in medicinal plant identification has significant implications for drug discovery, personalized medicine, and conservation efforts. By accurately identifying and characterizing medicinal plants, researchers can expedite the discovery of new drugs and therapies, potentially addressing unmet medical needs and improving patient outcomes. Additionally, machine learning-driven approaches can facilitate the sustainable utilization of plant resources by guiding conservation efforts and promoting responsible harvesting practices.

In essence, the future scope for medicinal plants identification using machine learning is multifaceted and promising, offering transformative opportunities for scientific innovation, healthcare advancement, and environmental stewardship. As research in this field continues to evolve, we can anticipate groundbreaking discoveries that harness the power of technology to unlock the full potential of nature's pharmacy for the benefit of humanity.

VIII. AUTHOR CONTRIBUTIONS

All Authors equally contributed in this work.

REFERENCES

[1]. M. G. M. Perera, A. N. Karunarathna, and K. P. Wijewardana Year: 2023 Title: "Machine learning-based identification of medicinal plants: A review

[2]. S. Gupta, R. Singh, and A. SharmaYear: 2024Title: "Enhanced Medicinal Plant Identification Using Deep Learning Architectures"

[3]. N. S. Srinivasan, K. S. Babu, and P. Kumar Year: 2021 Title: "Medicinal Plant Recognition using Convolutional Neural Networks and Genetic Algorithms"

[4]. R. Sharma, S. Choudhary, and S. Kumar Year: 2022 Title: "Recent Advances in Medicinal Plant Identification: A Machine Learning Perspective"

[5]. H. Liu, Y. Wang, and L. Zhang Year: 2023 Title: "Medicinal Plant Identification Using Hybrid Machine Learning Models: A Comprehensive Review"

[6]. M. S. Costa, F. A. Dória, T. G. Alves, F. M. Queiroz, and R. L. H. Nascimento Year: 2022 Title: "Automated Identification of Medicinal Plants using Machine Learning Techniques"

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

[7]. A. Gupta, S. Yadav, and R. Ranjan Year: 2021 Title: "Medicinal plant identification using convolutional neural networks and transfer learning"

[8]. G. Lahanas, P. Karamolegkos, A. Gasteratos, and N. Kalatzis Year: 2020 Title: "Deep learning for the identification of medicinal plants"

[9]. S. A. Raghu, P. Kavitha, and M. C. Suma Year: 2018 Title: "Medicinal plant leaf identification using texture and shape features with machine learning algorithms"