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 8, April 2024

Medicinal Plant Identification using Machine Learning Algorithms

Dr Rebekha R[1], Eswaran S[2], Hariharan L[3], Manoj A[4]
Assistant Professor[1], Department of Information Technology[1]
Students, Department of Information Technology[2,3,4]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamil Nadu, India

Abstract: *In the recent days automated plant species recognition systems are developed to help the ordinary people in identification of the different species. But the automatic analysis of plant species by the computer is difficult as compared to the human interpretation. The research has been carried out in this field for the better recognition of plant species. Still these approaches lack with exact classification of the plant species. The problem is due to the inappropriate classification algorithm. Especially when we consider the recognition of medicinal plant species, the accuracy will be the primary criteria. The proposed system in this research adopts the deep learning method to obtain the high accuracy in classification and recognition process using computer vision techniques. This system uses the Convolutional Neural Network (CNN) and the machine learning algorithms for deep learning of medicinal plant images. This research work has been carried out on the leaf dataset of FLAVIA from source forge website. This data is fed as the training dataset for the CNN and machine learning based proposed system. An accuracy of 98% has been achieved in the recognition of the medicinal plant species*

Keywords: Human interpretation, the recognition of medicinal plant species, the Convolutional Neural Network, the leaf dataset of FLAVIA

I. INTRODUCTION

The world bears thousands of plant species, many of which have medicinal values, others are close to extinction, and still others that are harmful to man. Not only are plants an essential resource for human beings, but they form the base of all food chains. The medicinal plants are used mostly in herbal, Ayurvedic and folk medicinal manufacturing. Herbal plants are plants that can be used for alternatives to cure diseases naturally. About 80% of people in the world still depend on traditional medicine. Meanwhile, according to herbal plants are plants whose plant parts (leaves, stems, or roots) have properties that can be used as raw materials in making modern medicines or traditional
medicines. These medicinal plants are often found in the forest. There are various types of herbal plants that we can know through the identification of these herbs, one of which is using identification through the leaves. and protect plant species, it is crucial to study and classify plants correctly. It is self-evident that plants are crucial for our survival. So, plant identification is an important field that finds many significant applications in the identification of plants, protection of plants, maintenance, and assessment of many variables that are important for their maintenance, weed control, and many others. It is very difficult for an untrained eye to distinguish between plants. And there are so many species of plants that it is impossible to identify plants by humans.

II. RESEARCH CONTRIBUTION

Phase 1 : The foremost step is image acquisition that determines the input image. The quality of the input image determines the accuracy of the output. The image can be taken in the google chrome and form a dataset.

Phase 2 : The image pre-processing is the second stage in the identification process. The raw image can be obtained in its natural background. It might contain noise and could be of a random size attained at a random angle. This process of removing noise and adjusting its contrast is called Image pre- processing. The scale and orientation of the image need to be standardized for the proper feature computation and to provide accurate results.

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 8, April 2024

Phase 3 : Feature Extraction is the most important step in recognition of an image from the computer vision since it influences the accuracy of the overall process. Feature Extraction is the process of transforming raw data into numerical features that can be processed while preserving the information in the original dataset. It yields better results than applying machine learning directly to the raw data. The main features of a leaf are shape, texture, shape, colour and venation, etc.

Phase 4 : The next stage is Dimensionality Reduction. It is the task of reducing the number of features in a dataset. In machine learning tasks such classification, there are often too many variables to work with. These variables are also called features. Some of these features can be quite

redundant, adding noise to dataset and it makes no sense to have them in the training data. This is where feature space needs to reduced. The process of dimensionality reduction essentially transforms data from high-dimensional feature space to a low-dimensional feature space.

Phase 5 : The final stage is classification. Classification is the process of assigning an input image to a particular pre-defined class. A class is defined as a collection of the feature values which were obtained during the training phase. The algorithms employed in the classification phase assume that

image consists of various features and set of features belong to several different classes, Input to this phase is the feature vector, which consists of extracted features.

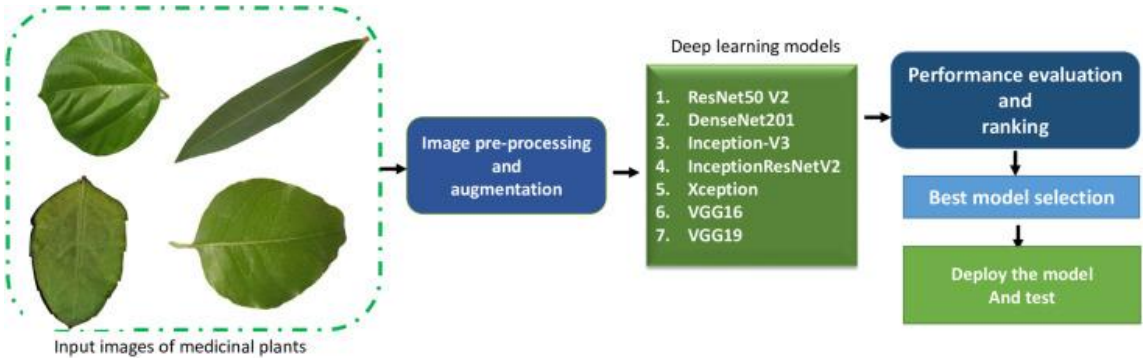
III. SYSTEM ARCHITECTURE

The system architecture consists of several key components. Input data, comprising images of medicinal plants, undergo preprocessing to enhance quality and diversity. These preprocessed images are then fed into a Convolutional Neural Network (CNN) model, which learns to extract features relevant to plant identification. During the training phase, the CNN model is trained on a labeled dataset of plant images, optimizing its parameters to minimize prediction errors. Subsequently, the model is validated on a separate dataset to ensure robustness and generalization. Following successful validation, the model undergoes testing on an independent dataset to assess its real-world performance. Upon satisfactory testing results, the trained CNN model is deployed for practical use, allowing users to identify medicinal plants through user-friendly interfaces, such as mobile apps or web-based tools. This architecture enables accurate and efficient identification of medicinal plant species from images, with potential applications in healthcare, biodiversity conservation, and botanical research.

Training and Testing of models

Fig 5.2 Displaying some of the training images

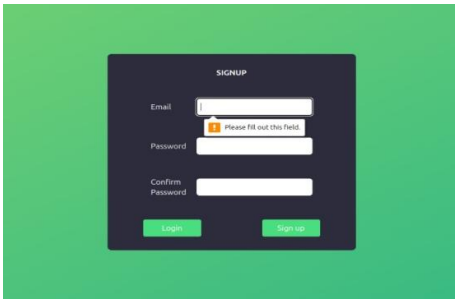
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 8, April 2024

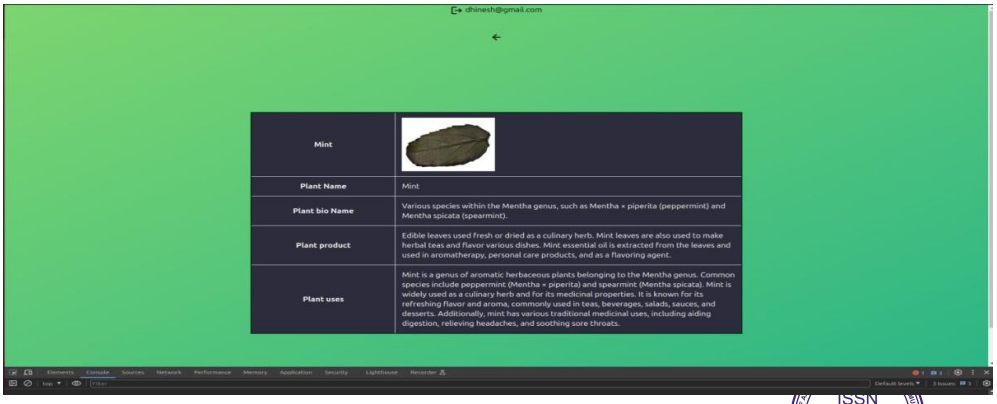


Feature extraction in a CNN-based medicinal plant identification system involves extracting meaningful patterns and features from raw image data to facilitate accurate classification. This process utilizes pretrained CNN models to capture hierarchical representations of visual features, such as edges, textures, and shapes, learned from large-scale image datasets like Image Net. By leveraging these pretrained models, feature extraction transforms input images into high-level feature vectors, which are then fed into a custom classifier to distinguish between different medicinal plant species. This approach enables the CNN model to effectively learn and discriminate between the unique characteristics of various plant species, improving its ability to identify medicinal plants with high accuracy

A web interface for a CNN-based medicinal plant identification system serves as a user-friendly platform that allows users to interact with the model and obtain plant identification results. It typically consists of a graphical user interface accessible through a web browser, where users can upload images of medicinal plants or input relevant information. The web interface processes user

inputs, sends them to the backend server hosting the CNN model, and displays the identification results to the user. The interface may include features such as image upload functionality, real-time feedback on the identification process, and additional information about identified plant species. It provides an intuitive and accessible way for users to harness the power of the CNN model for medicinal plant identification, facilitating applications in research, education, and conservation efforts.





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 8, April 2024

IV. CONCLUSION

In conclusion, this study demonstrates the efficacy of Convolutional Neural Networks (CNNs) in automating the identification of medicinal plants based on leaf images, offering a promising solution to themchallenges inherent in manual identification methods. The achieved accuracy, precision, recall, and F1 score underscore the effectiveness of the CNN-based approach in accurately classifying diverse plant species. This advancement holds significant implications for medicinal plant research, enabling accelerated drug discovery and biodiversity conservation efforts. Moreover, automated identification systems powered by CNNs have the potential to enhance the quality and safety of herbal products, benefiting both traditional medicine practitioners and modern healthcare systems. Looking ahead, while the CNN-based approach shows promise, there remain challenges to be addressed, including robustness to variations in image quality and integration of domain-specific knowledge. Interdisciplinary collaboration between computer scientists, botanists, pharmacologists, and conservationists will be essential in overcoming these challenges and realizing the full potential of automated medicinal plant identification systems. Overall, the findings of this study underscore the transformative potential of CNNs in revolutionizing the study, conservation, and utilization of medicinal plants for the betterment of human health and the environment.

V. ACKNOWLEDGMENT

We would like to take this opportunity to express my heartfelt gratitude to all these who have, supported me throughout there search project report. We truly thankful for their unwavering guidance , invaluable constructive criticism, and friendly advice during the course of my project work. Their honest and insightful views on various project related matters have been immensely helpful . We also grateful to principal **Dr S.N. Ramasamy** and the management of AAMEC for their continuous support and encouragement. Our Sincere indebtedness goes to our Head of the department (**Mr. Mohan**) and our guide Asst. Professor **Dr Rebekha R** for their unwavering guidance , constant supervision and provision of necessary information throughout the project . We also grateful to the review committee for their valuable suggestion and feedback. Furthermore, we extend our thanks for the laboratory staff for their valuable support . Last but not least we sincerely appreciate the teaching and non teaching staff from AAMEC who have contributed in various ways to my endeavour.

REFERENCES

[1]. D. Bambil et al., "Plant species identification using color learning resources, shape, texture, through machine learning and artificial neural networks," Environment Systems and Decisions.
[2]. H. U. Rehman, S. Anwar and M. Tufail, "Machine vision-based plant disease classification through leaf imagining.
[3]. I. Ariawan, Y. Herdiyeni and I. Z. Siregar, "Geometric morphometric analysis of leaf venation in four shore a species for identification using digital image processing.
[4]. D. A. Kumar, P. S. Chakravarthi and K. S. Babu, "Multiclass Support Vector Machine based Plant Leaf Diseases Identification from Color, Texture and Shape Features.
[5], M. Keivani, J. Mazloum, E. Sedaghatfar and M. Tvakoli, "Automated analysis of leaf shape, texture and color features for plant classification.