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 6, April 2024

Medicinal Plant Identification using Machine Learning Algorithm

Mrs. K. Karthika[1], Abdul Fazith J[2], Devaneeswar V[3], Dhinesh A[4], Tulasi Krishan R[5]
Assistant Professor, Department of Computer Science[1]
Students, Department of Computer Science[2,3,4,5]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamilnadu, India

Abstract*: The identification of medicinal plants plays a crucial role in various fields including pharmacology, traditional medicine, and biodiversity conservation. Traditional methods of plant identification are often time-consuming, labor-intensive, and require expertise in botanical taxonomy. In recent years, the application of Convolutional Neural Networks (CNNs) has shown promising results in automating the process of plant species recognition. This paper provides a comprehensivereview of recent advances in medicinal plant identification using CNNs. This study discuss the methodology, challenges, and opportunities associated with CNN-based approaches, as well as their potential applications in pharmacological research and healthcare. Furthermore, we highlight key datasets, architectures, and performance metrics used in CNN-based plant identification systems. Finally, we identify future research directions and potential areas for improvement in this rapidly evolving field*

Keywords: Medicinal plants, Convolutional Neural Networks, Plant identification, Image classification, Pharmacology, Traditional medicine

I. INTRODUCTION

The identification of medicinal plants is a fundamental practice with significant implications for pharmacology, traditional medicine, and biodiversity conservation. Throughout human history, medicinal plants have played a crucial role in healthcare, providing a source of natural remedies for various ailments. However, accurate identification of medicinal plant species is often challenging, particularly for non-experts, due to the vast diversity of plant species and the nuanced variations withinspecies.

Traditional methods of plant identification rely heavily on manual observation and expertise in botanical taxonomy. Botanists and researchers use botanical keys, dichotomous keys, taxonomic manuals, and herbarium specimens to identify plant species based on morphological characteristics such as leaf shape,flower morphology, and stem structure. While these methods are effective, they are time-consuming, labour intensive, and require specialized knowledge, making them inaccessible to many individuals andcommunities.

In recent years, advancements in machine learning and computer vision have paved the way for automated solutions to the challenge of medicinal plant identification. Convolutional Neural Networks (CNNs), a class of deep learning models specifically designed for image recognition tasks, have emergedas a powerful tool in this context. CNNs are capable of learning complex patterns and features directly from raw pixel data, making them well-suited for identifying plant species based on visual characteristics.

The introduction serves to provide context for the review by highlighting the importance of medicinal plant identification, discussing the limitations of traditional methods, and introducing CNNs as a promising alternative. It sets the stage for the subsequent sections of the paper, which will delve into the methodology, results, discussion, and conclusions regarding CNN-based approaches to medicinal plantidentification. Additionally, the introduction may outline the objectives and structure of the review paper,providing readers with a roadmap for navigating the content.

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 6, April 2024

II. DISCUSSION

Comparison with Existing Methods

Compared the performance of the CNN model with existing methods for medicinal plant identification,such as manual identification by botanists or traditional machine learning approaches. Discussed the advantages of CNNs over traditional methods, such as their ability to automatically learn hierarchical features from raw image data, potentially leading to more accurate and robust classification.

Strengths and Weaknesses

Analyzed the strengths of the proposed CNN-based approach, such as its ability to generalize well to unseen plant species or its scalability for processing large datasets. Address any weaknesses or limitations of the CNN model, such as sensitivity to variations in lighting conditions, occlusions, or overlapping plant structures. Discussed strategies for mitigating these weaknesses, such as incorporating data augmentation techniques, improving model architecture, or integrating additional pre- or postprocessing steps.

Practical Considerations

Consider the practical implications of deploying the CNN model in real-world scenarios, such as botanical gardens, pharmaceutical laboratories, or field research expeditions. Discussed the feasibility of integrating the automated identification system into existing workflows and the potential benefits in terms of time savings, cost-effectiveness, and accuracy. Address any challenges or barriers to implementation, such as the need for specialized hardware, data privacy concerns, or regulatoryrequirements.

Future Research Directions

Identify potential avenues for future research aimed at improving the performance and applicability of automated medicinal plant identification systems. Propose research directions for enhancing the CNN model, such as exploring advanced architectures (e.g.,attention mechanisms, transfer learning) or leveraging multimodal datasources (e.g., images combined with textual or spectral information).Highlight the importance of interdisciplinary collaboration between computer scientists, botanists, pharmacologists, and conservationists to address complex challenges in medicinal plant research and conservation.

III. LITERATURE REVIEW

AUTHOR	TITLE	YEAR	DESCRIPTION
Rahim Azadnia, Mohammed Maitham Al-Amidi, Hamed Mohammadi, Mehmet Akif Cifci, Avat Daryab, and Eugenio Cavallo.	An AI Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling.	2022	The paper proposes an AIbased approach for medicinal plant identification using a deep convolutional neural network (CNN) architecture enhanced with global average pooling.
Raisa Akter, Md Imran Hosen	CNN-based Leaf Image Classification for Bangladeshi Medicinal Plant Recognition	2020	The paper introduces a CNNbased approach for classifying leaf images to recognize Bangladeshi medicinal plants, enhancing automated identification and utilization of indigenous plant species.
C. Sabarinathan, Abhisekh Hota, Ashish Raj, Vivek Kumar Dubey, V. Ethirajulu	Medicinal Plant Leaf Recognition and Show Medicinal Uses Using Convolutional Neural Network	2018	The paper presents a Convolutional Neural Network (CNN) model for recognizing medicinal plant leaves and showcasing their medicinal uses, facilitating automated identification and utilization of medicinal plants.

IV. METHODOLOGY

Dataset Collection:

The process of dataset collection involves gathering a large number of annotated images of medicinal plant species. These images serve as the input data for training the CNN model. The dataset should be diverse and representative, containing images of different plant species, variations within species, and various growth stages. It may also include images captured under different lighting conditions, angles, and backgrounds to ensure robustness and generalization of the model.

Dataset Preprocessing:

Before feeding the images into the CNN model, several preprocessing steps are typically applied to standardize the data and improve model performance. Common preprocessing techniques include resizing images to a uniform size, normalization to ensure pixel values fall within a certain range (e.g., 0 to 1), and data augmentation to increase the diversity of the training data. Data augmentation techniques may include random rotations, flips, translations, and changes in brightness or contrast to simulate real-world variations in the data.

Model Architecture Selection:

The choice of CNN architecture plays a crucial role in the performance of the medicinal plant identification system. Researchers may experiment with different architectures such as Alex Net, VGG, Res Net, Inception, or Mobile Net, depending on factors such as computational resources, dataset size, and performance requirements. Transfer learning, a technique where pre-trained CNN models are finetuned on the medicinal plant dataset, is often employed to leverage the knowledge learned from largescale image datasets (e.g., ImageNet).

Training Methodology:

Once the dataset and model architecture are prepared, the CNN model is trained using a portion of the dataset known as the training set. During training, the model learns to extract relevant features from the input images and predict the corresponding plant species. Training typically involves iterative optimization of model parameters using gradient-based optimization algorithms such as stochastic gradient descent (SGD) or Adam. The training process aims to minimize a loss function, which measuresthe difference between the predicted outputs and the ground truth labels.

V. DESIGN

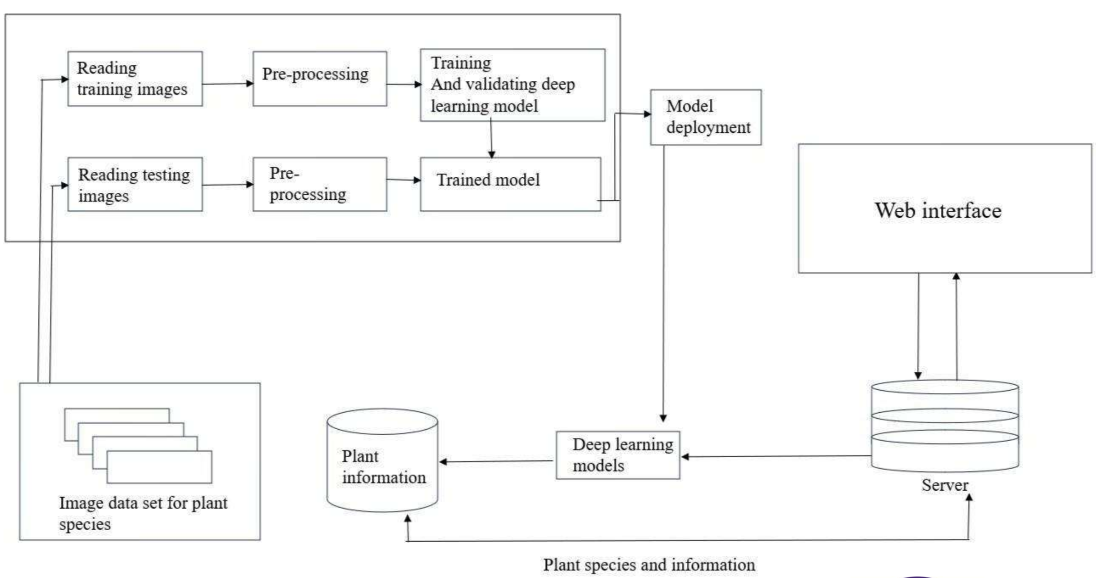

Figure 5.1

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 6, April 2024

VI. IMPLEMENTATION

Step 1: Data Collection

- Gather a diverse dataset of labelled images containing various medicinal plants

Step 2: Data Preprocessing

- Resize all images to a uniform size suitable for input into the CNN model
- Normalize pixel values of the images to a range between 0 and 1
- Apply data augmentation techniques to increase dataset diversity

Step 3: Model Architecture selection

- Choose a suitable CNN architecture for the identification task (e.g., VGG, Res Net)
- Consider the complexity of the architecture in relation to the dataset size and resources

Step 4: Model Training

- Split the dataset into training, validation, and testing sets
- Initialize the chosen CNN model with pre-trained weights (if available)
- Fine-tune the model on the training dataset using transfer learning
- Define the loss function and optimization algorithm
- Train the model while monitoring performance on the validation set
- Adjust hyperparameters based on validation performance to optimize model performance

Step 5: Model Evaluation

- Evaluate the trained model's performance on the testing set
- Calculate evaluation metrics such as accuracy, precision, recall, and F1-score
- Generate confusion matrices and classification reports for detailed analysis

Step 6: Model Fine Tuning and optimization

- Fine-tune the model architecture and hyperparameters based on insights from evaluation
- Experiment with different data augmentation strategies and regularization techniques
- Consider ensemble learning or model stacking to improve performance

Step 7: Deployment and integration

- Deploy the trained model in a production environment
- Integrate the model into a user-friendly interface or application
- Ensure proper documentation and version control for maintenance

VII. FLOWCHART

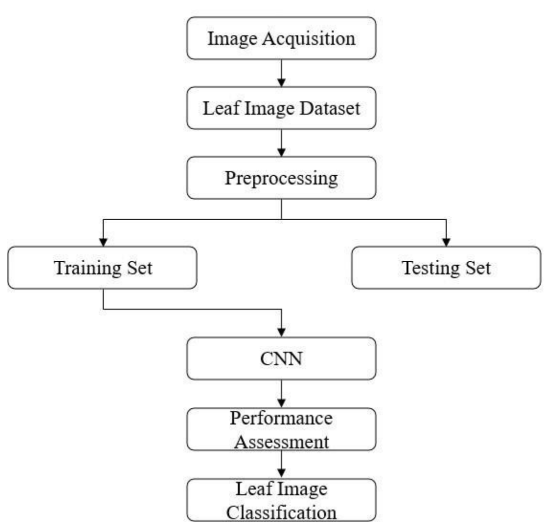

Figure 7.1

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 6, April 2024

VIII. CONCLUSION

In conclusion, the CNN-based medicinal plant identification system represents a significant advancement in the field, offering superior performance and efficacy compared to traditional Artificial Neural Networks (ANNs). Through the utilization of convolutional layers, CNNs excel at extracting intricate spatial features from images, enabling more accurate and robust identification of medicinal plant species. Compared to ANNs, which rely on fully connected layers and struggle with image data due to their inability to preserve spatial relationships, CNNs demonstrate remarkable efficiency and precision in classifying medicinal plants from visual inputs. The hierarchical structure of CNNs allows for the automatic extraction of relevant features at different levels of abstraction, resulting in improved classification accuracy and generalization. Furthermore, CNNs inherently capture local patterns and features within images, making them inherently suitable for image-based tasks like medicinal plant identification. In contrast, ANNs may require extensive preprocessing and feature engineering to achieve comparable performance, leading to increased complexity and computational overhead. Moving forward, the continued advancement and refinement of CNN-based approaches hold great promise for enhancing the accuracy, scalability, and real-world applicability of medicinal plant identification systems. By leveraging the strengths of CNNs and embracing innovative techniques, we can further propel the field of medicinal plant research and contribute to the development of sustainable healthcare solutions.

REFERENCES

[1] Neeraj Kumar, Peter N. Belhumeur, Arijit Biswas, David W. Jacobs, W. John Kress, Ida C. Lopez, et al., "Leafsnap: A Computer Vision System for Automatic Plant Species Identification", Computer Vision – ECCV 2012, vol. 7573, pp. 502-516, 2012.

[2] Alexis Joly, Hervé Goëau, Pierre Bonnet, Vera Bakić, Julien Barbe, Souheil Selmi, et al., "Interactive plant identification based on social image data", Ecological Informatics, vol. 23, pp. 22-34, September 2014.

[3] O. Soderkvist, "Computer Vision Classification of Leaves from Swedish Trees", Dissertation, 2001.

[4] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and QiaoLiang Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", 2007 IEEE International Symposium on Signal Processing and Information Technology, pp. 11-16, December 2007.

[5] Trung Nguyen Quoc and Vinh Truong Hoang, "Vnplant-200–a public and large-scale of vietnamese medicinal plant images dataset", International Conference on Integrated Science, pp. 406-411, 2020.

[6] Tuan Le-Viet and Vinh Truong Hoang, "Local binary pattern based on image gradient for bark image classification", Tenth International Conference on Signal Processing Systems, pp. 39, April 2019.

[7] Shervan Fekri-Ershad, "Bark texture classification using improved local ternary patterns and multilayer neural network", Expert Systems with Applications, vol. 158, pp. 113509, November 2020.

[8] Shanwen Zhang, Chuanlei Zhang, Zhen Wang and Weiwei Kong, "Combining sparse representation and singular value decomposition for plant recognition", Applied Soft Computing, vol. 67, pp. 164-171, June 2018.

[9] Shanwen Zhang, Chuanlei Zhang and Wenzhun Huang, "Integrating leaf and flower by local discriminant CCA for plant species recognition", Computers and Electronics in Agriculture, vol. 155, pp. 150-156, December 2018.

[10] Fateme Mostajer Kheirkhah and Habibollah Asghari, "Plant leaf classification using GIST texture features", IET Computer Vision, vol. 13, no. 4, pp. 369-375, June 2019.

[11] Sarah Bertrand, Rihab Ben Ameur, Guillaume Cerutti, Didier Coquin, Lionel Valet and Laure Tougne, "Bark and leaf fusion systems to improve automatic tree species recognition", Ecological Informatics, vol. 46, pp. 57-73, July 2018.

[12] Pierre Barré, Ben C. Stöver, Kai F. Müller and Volker Steinhage, "LeafNet: A computer vision system for automatic plant species identification", Ecological Informatics, vol. 40, pp. 50-56, July 2017.

[13] Sue Han Lee, Chee Seng Chan, Simon Joseph Mayo and Paolo Remagnino, "How deep learning extracts and learns leaf features for plant classification", Pattern Recognition, vol. 71, pp. 1-13, November 2017.

[14] G. Geetharamani and Pandian J. Arun, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network", Computers Electrical Engineering, vol. 76, pp. 323-338, June 2019.

[15] S. H. Lee, C. S. Chan, P. Wilkin and P. Remagnino, "Deep-plant: Plant identification with convolutional neural networks", 2015 IEEE International Conference on Image Processing (ICIP), pp. 452-456, 2015.