Classification and Analysis of Medicinal Plants Using Deep Learning – A Review

Pulkit Ahuja[1], Manish Bhardwaj[1], Sumit Kumar[1], Sanju Choudhary[2]
[1]Department of Computer Science and Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur-302017(India)
[2]Department of Information Technology, Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur-302017(India)
Email: **b210648@skit.ac.in, manish.bhardwaj@skit.ac.in, sumit.kumar@skit.ac.in, sanju@skit.ac.in**
Received 18.03.2025 **received in revised form** 05.05.2025, **accepted** 05.05.2025
DOI: 10.47904/IJSKIT.15.1.2025.1-4

Abstract- Correct drugs should be administered in dealing with the disease appropriately. Since allopathic medicine accelerates relief, it is not meant to deal radically with the root causes of disease manifestation. Ayurvedic drugs take steps directed at the root cause by establishing long-term solutions to restore health holistically. For accurate identification and classification of these plants represent a key factor for improving Ayurvedic medicinal products in modern health care delivery. It deals with the concern of classification and identification of Ayurvedic medicinal plants using advanced deep learning techniques. In this case, we utilized the methodology of Convolutional Neural Networks (CNNs), which analyze image data about medicinal plants, thereby allowing accurate classification to take place. This paper also integrates several methods-Fuzzy C-means clustering, and H-Gabor filters-for enhanced feature extraction and segmentation. By taking a comprehensive Medicinal Plant Data set, the approach facilitates identification and promotes the better usage of Ayurvedic resources. The final finding drawn from the conducted research on plant-based medicine, incorporating traditional knowledge, open prospects for future innovations for improvement.

Keywords– Image Classification, Convolutional neural network (CNN), H-Gabor filter, Medicinal Plant Data, Fuzzy C means (FCM), Deep Learning.

1. INTRODUCTION

Medicinal plants play a very crucial role in many fields such as medicine, pharmacology, and conservation, identification processes must adopt new methodologies. This study aims to change the tide of identification using state-of-the-art technologies, CNNs, and ML algorithms on the interface between traditional botanical knowledge and technological progress. The medicinal plants due to their numerous applications are very important and are essential both for the classic curative treatments and the new medicines because they are of paramount sources of bioactive chemicals. Yet, the identification of such plants still remains a hard nut to crack. This project shall help in making an automatic yet efficient way for the recognition of many medicinal plants through the development of a system based on the use of the capabilities of the image classification algorithm so as to overcome this problem.

The proposed approach is Convolutional Neural Networks--a family of deep learning algorithms more specifically designed for applications of images. Since CNN can learn hierarchical feature representations from raw pixel data, a significant amount of success can be seen in image classifications. This makes CNNs the best candidates for the challenging task of identification between various categories of medicinal plants considering hierarchical structure, allowing them to identify complex patterns and relationships that exist in images. CNN possesses many advantages over other methods in image classification. Because these networks learn proper features needed from given data on which CNN does not depend on human feature engineering like traditional approaches to machine learning, hence feature extraction becomes less cumbersome and the model remains more general for a diversity of plant features. The other reason CNNs can differentiate between high-level features that describe the overall shape of a plant and low-level features like texture and colors is due to hierarchical feature learning. In this process, identification of medicinal plants becomes on its head with the integration of CNNs into the whole ML platform. Hence, the system can process images of High resolution that may comprise multicomponent of the plant including leaf, flower, and stems.

2. LITERATURE REVIEW

As per C Dileep M.R. [1] AyurLeaf which is a Convolutional Neural Network Deep Learning

model, used in the classification of medicinal plants based on their leaf characters like color, shape, texture size and so on. AyurLeaf is the CNN-based classification model tested and trained on its own dataset as well as tested on DLeaf dataset also. SVM classifier gave a better classification accuracy on the AyurLeaf dataset also. SVM classifiers and Softmax were used by AyurLeaf for the purpose of classification achieving the accuracy of 96.76%. An automatic identification system was proposed by Gopal A et al. [2] for specific medicinal plant leaves. This approach suggests the identification of various categories of leaves using features as boundary-based, moment and color. In data a bunch of 100 and 50 leaves were trained and tested respectively and classification efficiency was demonstrated as 92% by the software. A CNN model known as DLeaf, venation-based, was proposed by Jing Wei Tan et al. [3] the DLeaf for the classification of plant leaves. DLeaf used ANN and CNN for classification and feature extraction, respectively. The Sobel edge detection algorithm is applied on the resized leaf images to get the venation segmentation. The classification accuracy of the D-Leaf model was 94.88%.

Mostafa Mehdipour Ghazi et al. [8] have used various deep CNNs for plant species classification. The comparison networks which are very popular namely GoogLeNet, VGGNet, AlexNet have been evaluated for formulating the features involved in these performance networks. It has also been used in training towards various components of the plant like the stem, leaf, flower, fruit, branch and even the whole plant.

In which, the optimized VGGNet gave the highest accuracy value, which is 78.44%. Shitala Prasad et al. [9] presents methods for capturing leaves and techniques to convert the captured image into a device independent color space. Those are then used in determining the VGG-16 feature map. The principal component analysis is applied on that feature map in order to reduce or optimize this feature map PCA. The dataset used is ImageNet. The Fully Connected layer has the output of a feature vector of size 3x4096. In applying the PCA approach, SVM was applied to the feature set and achieved accuracy of 97.6% for lVGG-16 and 98.2% for l-VGG-16. Bella Dwi Mardiana [10] works are devoted to gaining even better performance for Convolutional Neural Networks (CNNs) on the herbal leaf classification task. In this research study, in particular, she has used a VGG16 model together with a Transfer Learning approach. This project rationale is based on the fact that people know less about herbal leaves, which may bring in risks towards health issues because one can misidentify some plant species. With an effort to curb such risks, this research is attempting to give an accurate and exact model of classification. There are ten classes of herbal leaves in the dataset, and the technique proposed here reaches 97 % accuracy, which is very much better than the accuracy reached in previous studies, which was 82 %. To reduce overfitting and further increase the quality of the images, the classification accuracy becomes more perfect. It applies data augmentation on the dataset through Image Data Generator. This work has made a great contribution in the area by using the advanced architecture of CNN.

In this review paper we are focusing on some deep learning techniques that emphasize analyzing healthcare data in recent times. The analysis and understanding of healthcare data has become a complex task and requires high computational solutions. Artificial Intelligence and Deep Learning technologies have emerged in a prominent way to classify the characteristics of healthcare data. A deep learning architecture is simply defined as an Artificial Neural Network (ANN) consisting of two or more hidden layers which aim at enhancing the accuracy of the prediction [6]. In DNN a weighted and biased input value is passed on to a non-linear activation function to produce output. The connection helps to optimize the weight of the network [7].

3. METHODOLOGY

Accurate identification of most medicinal plants challenging work of great importance to traditional medicine, pharmacy, and conservation where traditional medicinal knowledge and newest technologies are brought together. Substructure of modern image classification techniques: more precise, Convolutional Neural Networks (CNNs), and Machine Learning algorithms are used to modernize the process itself. Holistic approach the rationale behind application of traditional medicine and bioactive discovery in it. Figure1 shows the images of a healthy Plant Leaf. Identification of Medicinal Plants Depends Primarily on the Identification of Medicinal Plants in Pharmaceuticals and Protection of Threatened Plant Species Identification of medicinal plants mostly depends on traditional approaches. Traditional approaches, however, are very time-consuming and mostly dependent on people's skills. This research will focus on developing an efficient and automated process that ensures an accurate identification of various species of medicinal plants. The current stress on simplifying the process of identification of

medicinal plants has made for this study:

Connecting the old botanical knowledge with the present technological advancement. Surprisingly, by the presence of CNNs, it benefits the field of image classification as the task which was erstwhile labor-intensive is replaced by it. This paper aims to accelerate the identification process and enrich its accuracy, efficiency, and accessibility through CNNs and ML. In addition to the boundaries of academia, this study has some social implications. The developed system can most probably be applied by both.

This identification system serves the purpose of identifying medicinal plants for amateurs and professionals alike, thereby supporting the cause of biodiversity conservation as well as environmentally friendly collection. Here figure 1 shows images of a Ayurvedic medicinal healthy plant leaf One can envision placing this system online or in mobile applications, whereby using a smartphone one can identify medicinal plants anywhere, at any time and, hence, establish a human link with rich botanical material used in. The very idea of applying ML and CNNs to classify images for identification purposes of many medicinal plants marks an innovative application with tremendous implications. Apart from solving the challenges facing traditional medicine, drugs, and conservation, it fits into more general goals - health procedures, biodiversity conservation, and sustainable relationships with nature.

CNNs are a class of deep neural networks designed specifically for the processing of structured grid data, such as images. Indeed, the architecture of the networks [1] is inspired by how the human brain processes visual information through successive layers of learning abstract features. Indeed, the constituents are the convolutional, pooling, and fully connected layers. This combination of constituents might capture feature representations at various hierarchical levels such that CNNs can classify complex patterns in images. Different approach is the use of ensemble learning by various CNN models. Ensemble approaches betterment the performance by making predictions from several models. Ensemble learning, using multiple Convolutional Neural Network (CNN) models, can greatly enhance medicinal plant identification. Ensemble methods combine several models' predictions to result in better efficiency. Here, a few CNNs, with differing initializations or architectures, perform better to enhance the Hence, the model will be more accurate and robust. The strengths of the individual models in the ensemble will provide much stronger reliable systems for the automated medicinal plants to identify processes for this model of image classification, and this system makes it more viable concerning the subtle differences of these various plant characteristics and environments. Using data augmentation techniques can increase the strength and ability to generalize CNNs, because we can increase the size of the dataset which helps in better model training and for generation of more accurate results. Data augmentation is a procedure that, on the fly, takes existing images, rotates them, flips, or zooms them, to make new training samples. Less overfitting would happen, and surely, there will be recognition of the medicinal plants by the model in varied scenarios with diversified image quality.

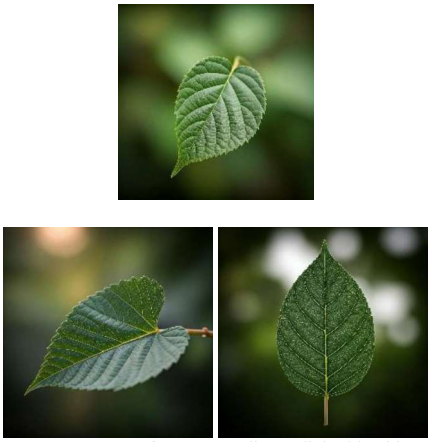

Figure 1: Images of an Ayurvedic medicinal healthy Plant Leaf

4. RESULTS AND DISCUSSION

The results of the project have proved that the proposed methodology was so efficient in automatically identifying medicinal plants. Better results were obtained from the CNN classifier with an amazing accuracy of 93.4%, CNN is so efficient and generates correct outputs and reliable results. That high accuracy gives a reflection of the strong and dependable system developed because it can identify and put many species of medicinal plants right into their categories. This, therefore, is a good testimony of the good performance by Feature extraction techniques, namely, Gabor filters and fuzzy C-means, in extracting and exploiting relevant visual patterns to achieve good CNN classifier accuracy. The results further consolidate the fact that not only does this project boast good integration between advanced image processing technology with traditional knowledge in botany, but there also is good scope for this to affect rather significantly most of these projects in that field covering pharmacy, medicines, and conservation. That would be one great step forward in building up automated detection machinery of medicinal plants in order to

assure enhanced conservation of biodiversity as well as a more resilient relation with resources. Here, figure 2 show a comparison of various classification methods on ayurvedic plant disease classification methods with accuracy which shows deep learning model achieve highest accuracy on plant diseases complex data set using pattern in plant images.

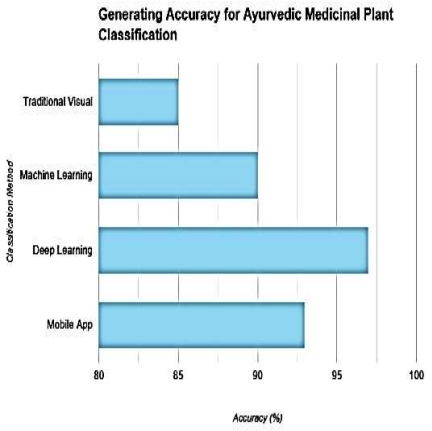

Figure 2: Comparison of Ayurvedic plant disease classification methods with accuracy

5. CONCLUSION

All this marked a paradigm shift within the domain of Ayurvedic medicinal plants identification, a fusion of traditional knowledge of botany with contemporary image processing technologies. There are 60 classes of Ayurvedic medicinal plants, so a comprehensive diverse analysis took place. Through the application of the state-of-the-art methods, specifically H-Gabor filters, fuzzy C-means, and CNNs, the successful automation of the identification process came to pass identify Ayurvedic medicinal plants. Adding Fuzzy C-means response improved the ability of the feature extraction module in being subtle to the texture features. Then these Fuzzy C-means provided a more extensive and true reproduction of Ayurvedic medicinal plant images. Therefore, this study impacts more than the intellectual dimension as it creates scope for better healthcare processes and applies aside from enabling sustainable interactions with Ayurvedic resources. This study serves as a model for future study and research because of its: proper combination of traditional ayurvedic science with contemporary technology. It fosters a harmonious coexistence of the former and the potential of the latter that highly automated systems provide. Lastly, as a symbol of what interdisciplinary work can do for the practical solving of problems in radically transformative ways.

6. ACKNOWLEDGEMENT

Acknowledgement of the work of others must be This will help to better appreciate all those who contributed in doing the research on how AI and machine learning can be used for medicinal plants classification. Special thanks to my academic advisor whose insight and guidance were invaluable to this project.

I would cherish other students and friends who have made constructive comments, encouragement, and collaboration within the work which made spice to the research. Least but not last, I will appreciate the experts and their organizations who offer information, knowledge, and contribution to such botanical research. Their level of professionalism allows for accuracy and relevance in any type of research conducted. This opportunity also enables me to give special thanks to my family and friends for supporting and being extremely patient with me during this last period. The constant trust that they have shown in me has always motivated me and encouraged me continuously. Thanks to all those for whom the success of this project depends on.

7. REFERENCES

[1] Dileep, M. R., and Pournami, P. N., Ayurleaf: A Deep Learning Approach for Classification of Medicinal Plants, 2019 IEEE Region 10 Conference, (2019), Kochi, India.

[2] A. Gopal, S. Prudhveeswar Reddy, and V. Gayatri, Classification of Selected Medicinal Plants Leaf Using Image Processing, International Conference on Machine Vision and Image Processing (MVIP), (2012), Dec.

[3] J. W. Tan, S. Chang, S. Binti Abdul Kareem, H. J. Yap, and K. Yong, Deep Learning for Plant Species Classification Using Leaf Vein Morphometri, (2018), [publication venue not specified].

[4] R. Janani and A. Gopal, Identification of Selected Medicinal Plant Leaves Using Image Features and ANN, International Conference on Advanced Electronic Systems (ICAES), (2013), Sept.

[5] A. Sabu, K. Sreekumar, and R. R. Nair, Recognition of Ayurvedic Medicinal Plants from Leaves: A Computer Vision Approach, Fourth International Conference on Image Information Processing (ICIIP), (2017), Dec.

[6] L. Deng and D. Yu, Deep Learning: Methods and Applications, Foundations and Trends® in Signal Processing, (2014), 7, pp. 197–387.

[7] J. Schmidhuber, Deep Learning in Neural Networks: An Overview, Neural Networks, (2015), 61, pp. 85–117.

[8] M. M. Ghazi, B. Yanikoglu, and E. Aptoula, Plant Identification Using Deep Neural Networks via Optimization of Transfer Learning Parameters, Neurocomputing, (2017).

[9] S. Prasad and P. P. Singh, Medicinal Plant Leaf Information Extraction Using Deep Features, TENC0N 2017 – IEEE Region 10 Conference, (2017), Nov.

[10] B. D. Mardiana, W. B. Utomo, U. N. Oktaviana, G. W. Wicaksono, and A. E. Minarno, Herbal Leaves Classification Based on Leaf Image Using CNN Architecture Model VGG16, Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), (2023), 7(1), pp. 20–26.