Leveraging Convolutional Neural Networks for Efficient and Accurate Identification of Medicinal Plants

G. Sathar[1], Husnaara[2], G . Akshaya[3], M . Bhanu Teja[4]
[1]Assistant Professor, Department of CSE - (CyS, DS) and AI & DS, VNR-VJIET, Hyderabad
[2-4]Undergraduate Students, AI & DS, VNR-VJIET, Hyderabad

Abstract*—— Identification of medicinal plants is essential for traditional health care systems, but the limited availability of experts often deters timely and accurate classification. This paper introduces a Deep Neural Network (DNN) model trained on a dataset of 117 medicinal plant species to automate the identification process. The model extracts fine-grained morphological features from leaf samples, which leads to high accuracy and scalability, avoiding the limitations of geographic locations and dependence on specialists. This technology-driven approach bridges the gap between ancestral knowledge and modern practices, ensuring the preservation of traditional medicinal heritage. Beyond efficiency, it promotes inclusivity by enabling access to accurate plant identification in resource-constrained settings. The model's ability to process large datasets enhances its adaptability for diverse applications, from research to fieldwork. The innovative blend of the power of advanced machine learning with cultural importance gives the continuity of the age-old traditional medicine along with the creation of modern pathways for herbal practice. Therefore, the solution will also maintain valuable indigenous knowledge with itself, thus sustaining future practice on medicinal plants. Medicinal plant identification

***Index Terms*—Keywords: Deep Neural Networks; Morphological feature extraction; Automated classification; Preservation of Traditional Medicine; Scalable Solutions**

I. Introduction

Future antennas will be of extreme importance in achieving faster communication and broadcasting technology. With the increase in demand for more bandwidth, the antennas should have a smaller size with increased performance. Multi-operable frequency bands and fractal geometry have highly influenced antenna design. More compact shapes of antennas are allowed without compromising on multiband and broadband applications. The fractal antenna design has particularly been quite popular in recent times because of the favorable features that can be extracted from them. Advancements in 5G and 6G, where high data rate transmissions are expected, IoT devices, radar systems, and navigation applications have accelerated interest in antenna development. On top of this, there is growing awareness and use for antennas that are sensitive enough for short-range transmissions, while also being effective with longer ranges, thereby fitting diverse communication needs.

II. Literature Survey

The earlier method for identifying plants was based on characteristics such as texture, shape, and color. However, with respect to current demands for precision and scalability, these methods are not as good. Other earlier approaches to machine learning include KNNs, SVMs, and random forests. These methods were promising but could not handle the complexity and variability of plant structures and environments. Deep learning, especially CNNs, has since revolutionized image classification by automatically extracting hierarchical features. Studies have proven its potential as it achieved 71.3% accuracy in classifying 10 Bangladeshi medicinal plants and 83.3% accuracy in the identification of three Ficus species using ANN and SVM models. These results prove the efficiency of CNNs in handling complex classification problems.

III. Proposed System

The proposed system starts by curating a dataset of special medicinal plant images, field and laboratory. The data set contains preprocessed well images to remove background noises so that the model gives proper focus. Augmentation of data techniques generated images under other environmental conditions like illumination scale and perspective, which helped in enhancing the robustness of the model. For the system, the architecture of three-layer CNN is being used. Such critical features are extracted by convolutional layers through the detection of various patterns, edges, and textures, and pooling layers reduce spatial dimensions of feature maps that conserve the most important details. Finally, fully connected layers will classify the input into clear categories, with output being the probabilities for each plant species. The system's performance is evaluated using accuracy and AUC metrics, and iterative hyperparameter tuning keeps refining the system. Initial evaluation results are competitive and have proved that this method can be competitive. The methodology provides not only an efficient alternative to manual identification but also steps toward integration of leading-edge technology into botanical and agricultural domains for greater sustainability and precision.

IV. Methodology

The methodology is based on the systematic approach, combining modern machine learning techniques with traditional approaches in plant classification to develop a CNN model that could automatically identify medicinal plants. The system will be efficient, scalable, and accessible for identification purposes and ensures its effectiveness in real-world scenarios

A. Dataset Preparation

The methodology begins with careful collection and preparation of a dataset. A diversified set of images of leaves from 117 medicinal plant species was collected. This dataset is vital because it acts as the training base for the model. Images are represented in series for every species of the plant taken at different times to enable the model to differentiate among differences like different lights, angles, and slightly morphed difference in them. These images are typically preprocessed before training because of its dependency on such improvements in model precision, mainly consisting of removal of noises, as well as of backgrounds clutters, normalizing color variation that considers light discrepancy difference, and size equalization into some common pixels in resolutions. Proper preparation of the dataset will ensure that background artifacts, which the model might use to learn, are eradicated, and it is focused on essential characteristics of the plant.

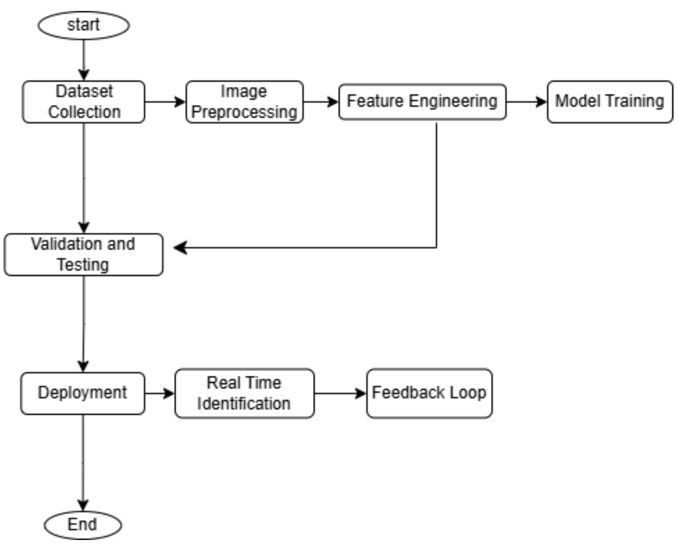

Fig. 1. Methodology Overview

B. Feature Extraction

After the dataset is ready, the next step will be to extract meaningful features from images. Feature extraction refers to a process where key characteristics of leaves that distinguish one plant from the other are identified and extracted. Features used to describe these may be based on texture, vein pattern, shapes at edges, and geometry. Deep Neural Networks are advantageous for this task since with each step of the multiplicity of layers, features found here are automatically learned to describe something more complex about the leaves and have deep layers that use extraction of high-level features. Hence, it enables classification even when the visual similarity between the two is quite subtle.

C. Training Model

The very heart of the methodology was the training of the DNN model. Training involved passing preprocessed and feature-extracted data to the network and letting it learn from this input to adjust its parameters so that prediction errors would be minimal. The dataset is divided into three parts: a training set, a validation set, and a testing set. The training set presents labeled examples of plants. As the model goes through these examples, it continues to adjust its internal weights and parameters to improve the predictions. It is utilized in checking the model's performance to avoid overfitting (when the model becomes too customized to the training data and doesn't generalize well). It's used at the final stages of training to test the performance of the model. It checks the model's accuracy, and further refinement is made.

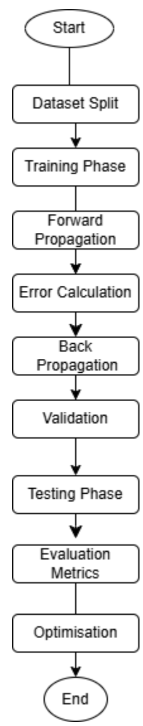

Fig. 2. Work Flow Diagram

D. Evaluation and Optimization

After training the model, testing is the most important factor in its performance improvement so that it works at its optimum level. It includes testing the model according to the performance metrics and optimization of it to have better accuracy and generalization. Evaluation Metrics, There are many important metrics through which its performance can be evaluated regarding how well the model is performing:

1) Accuracy: Calculates the overall correctness of the model by considering the correct predictions out of all made predictions.

- **TP** - True Positives
- **TN** - True Negatives
- **FP** - False Positives
- **FN** - False Negatives

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

2) Precision : This is the number of positive predictions that the model correctly identified. It gives how many of the positive plant species that were predicted were indeed correct

$$Precision = \frac{TruePositives\ (TP)}{TruePositives\ (TP) + FalsePositives\ (FP)}$$

3) Recall (Sensitivity):: Measures the extent to which the model identifies the actual positive cases; in this case, correct plant species.

$$Recall = \frac{TruePositives\ (TP)}{TruePositives\ (TP) + FalseNegatives\ (FN)}$$

4) F1-Score: : This is the harmonic mean of precision and recall. It balances both to provide a better measure.

$$F1\text{ Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

These metrics give a view of how well the model is predicting and where it needs to be improved.

E. Deployment

The final step is the deployment, in which the model is applied practically. The deployment step requires integration of the trained model into a user-friendly platform, such as a mobile or web application. Users can upload images of plants, and the model processes them in real time, producing immediate results for plant identification. The deployment interface is user-friendly. Therefore, any layperson with less technical know-how can easily upload images and receive correct results.

F. Validation and Feedback Loop

Following the deployment of the model, predictions must be constantly validated so that the results do not degenerate in the long term. This can be achieved by cross verifying the predictions from the model with expert verification and then cross-checking them against already established plant databases. It is self-updating, as it improves itself based on the users' and experts' feedback. This is every time a new dataset of images of other plant species is added or there is an error in classification. These new data are then fed into the model, further fine-tuning its predictions. This continues and makes sure the model improves with learning about the new species, changes in the environment, and evolving user requirements.

V. Results

The main objective of this project was to design a robust and reliable Convolutional Neural Network (CNN) model for the accurate classification of medicinal plants. The results clearly demonstrate the success and efficiency of the approach, with the CNN model achieving exceptional outcomes across various performance metrics and evaluation scenarios.

Fig. 3. Medicinal Leaves

A. Image Upload and Prediction

The model has practical usability because it enables the interface where users can input images of different parts of plants like flowers or leaves for instant prediction. The platform makes things easier as many predictions can be carried out effortlessly through this feature, which therefore provides ease and convenience

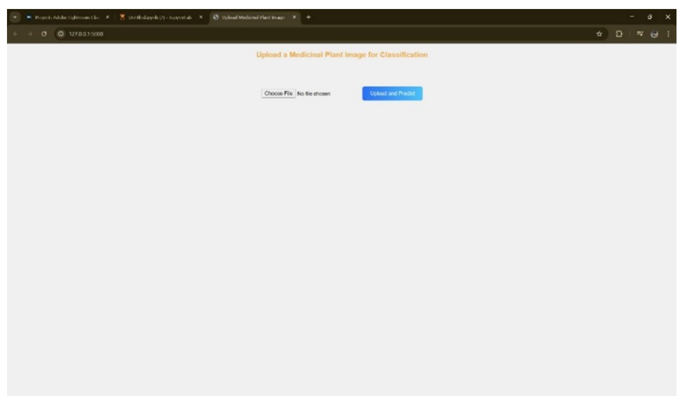

Fig. 4. Upload Page

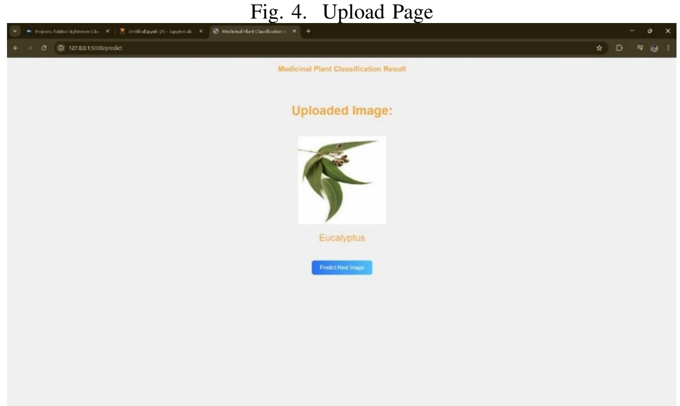

Fig. 5. Identification Page

B. Performance of the Model

CNN showed remarkable accuracy of 97% on the test set. Such accuracy appeared during different rounds of validation with this model indicating good reliability and stability of this model. These results show how the CNN models are able to generalize the things despite facing real-time problems that exist while testing: differences in scale, illumination, and other complex variations related to background. It goes way beyond all other state-of-the-art traditional models of image classification as used for these same kinds of application.

```
validation_loss, validation_accuracy = model.evaluate(validation_generator)
print(f'Validation accuracy: {validation_accuracy}')

12/12 [==============================] - 73s 6s/step - loss: 0.0599 - accuracy: 0.9763
Validation accuracy: 0.9763157963752747
```

Fig. 6. Accuracy

C. Comparison

To establish the performance of the CNN, its accuracy was compared with other well-known models:

- **MobileNet:** It was weak and performed poorly, as it had trouble dealing with background and lighting variations, and the accuracy level was low.
- **InceptionV3:** This model was better than KNN but was not as sophisticated as to handle the complexity of the visual features that were required for this work.
- **DenseNet121:** It showed robustness but failed to achieve the desired accuracy in distinguishing closely related plant species.

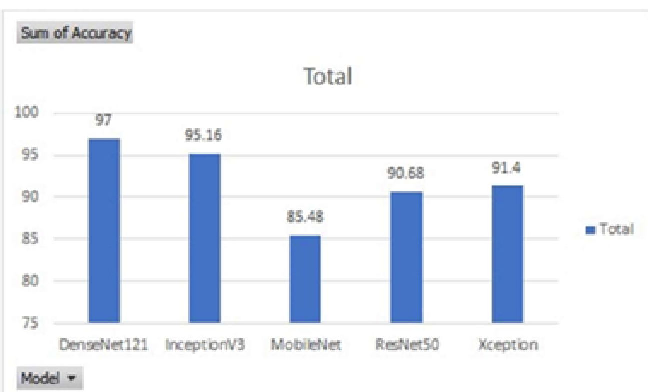

Fig. 7. Comparison

The accuracy scores for the alternative models varied from 85% to 92%, which is way less than that of the CNN model, which was 97%. This analysis reveals the better ability of CNN's in handling high intra-class variability and intricate image features.

D. Training and Validation

The accuracy curve for the training and validation datasets showed the model's ability to learn and generalization effectively. This succession of curves demonstrated the stability and robustness of the CNN across different test cases. A visual comparison of training and validation accuracy indicated minimal overfitting, showing well-balanced performance throughout

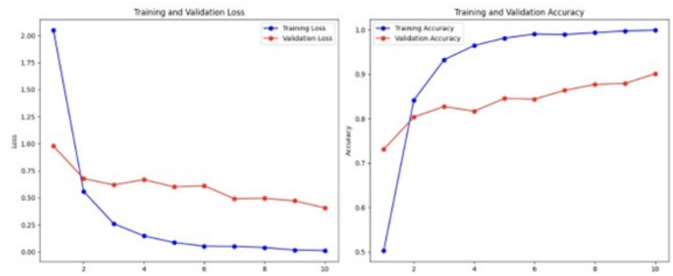

Fig. 8. Train and valid plot

Conclusion

In conclusion, the present study demonstrates the ability of deep learning techniques, such as Convolutional Neural Networks (CNNs), in the automatic classification of medicinal plants. In this work, the capabilities of CNN in extracting features from images are utilized in order to overcome some challenges associated with plant identification, including lighting, background variations, and intra-species diversity. Not only did the proposed system prove a notable performance in comparison with traditional machine learning methods, such as MobileNet, InceptionV3 , and DenseNet121 , which generally fail when dealing with plant images having complex details. Our trained system on the curated medicinal plant dataset gives a lot of promise toward faster identification of plants, which finally can be contributed to other fields like agriculture, medicine, and conservation. The success of this model opens up more doors for broader applications in the field of botanical research and sustainable practices. Future prospects include expanding the dataset with more advanced model architectures that could further improve the system's accuracy and adaptability. As the technology grows, this tool will soon become an integral part of traditional and modern practices for preserving valuable plant knowledge in the future.

References

[1] **H.K. Kan**, Classification of medicinal plant leaf image based on multi-feature extraction, 2017.
[2] **Desta Sandya Prasvita and Yeni Herdiyeni**, MedLeaf: Mobile Application For Medicinal Plant Identification Based on Leaf Image, 2013.
[3] **R. Janani**, Identification of selected medicinal plant leaves using image features and ANN, 2013.
[4] **A. Gopal**, Classification of selected medicinal plants leaf using image processing, 2012.
[5] **Tejas D. Dahigaonkar**, Identification of Ayurvedic Medicinal Plants by Image Processing of leaf samples, 2018.
[6] **P. Manoj Kumar**, Identification of Ayurvedic Medicinal Plants by Image Processing of leaf samples, 2017.
[7] **Raisa Akter and Md Imran Hosen**, CNN-based Leaf Image Classification for Bangladeshi Medicinal Plant Recognition, 2020.
[8] **C. Sabarinathan**, Medicinal Plant Leaf Recognition and Show Medicinal Uses Using Convolutional Neural Network, 2018.
[9] **Kalananthni Pushpanathan**, Machine learning in medicinal plants recognition: a review, 2014.

[10] **Rahim Azadnia**, An AI-Based Approach for Medicinal Plant Identification Using Deep CNN Based on Global Average Pooling, 2022.

[11] **Srdjan Sladojevic**, Deep Neural Networks-Based Recognition of Plant Diseases by Leaf Image Classification, 2016.

[12] **Anand Handa**, A Review and a Comparative Study of Various Plant Recognition and Classification Techniques using Leaf Images, 2015.

[13] **M.A.F. Azlah**, Review on Techniques for Plant Leaf Classification and Recognition, 2019.

[14] **Shanwen Zhang**, Plant species recognition methods using leaf image: Overview, 2020.

[15] **Anubha Pearlines**, A study on plant recognition using conventional image processing and deep learning approaches, 2019.

[16] **Zhaobin Wang**, Review of Plant Identification Based on Image Processing, 2017.

[17] **Chaitali G.**, A modern approach for plant leaf disease classification which depends on leaf image processing, 2017.

[18] **Abdolvahab Ehsanirad**, Plant Classification Based on Leaf Recognition, 2010.

[19] **Halimatu Sadiyah Abdullahi**, Convolution Neural Network in Precision Agriculture for Plant Image Recognition and Classification, 2017.

[20] **Enes Yigit**, A study on visual features of leaves in plant identification using artificial intelligence techniques, 2019.

[21] **Jana Waldchen and Patrick Mader**, Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review, 2017.

[22] **G. Saleem**, Automated analysis of visual leaf shape features for plant classification, 2019.

[23] **A. D. A. D. S. Jayalath**, Identification of Medicinal Plants by Visual Characteristics of Leaves and Flowers, 2019.

[24] **Sue Han Lee**, Deep-plant: Plant identification with convolutional neural networks, 2015.

[25] **Pierre Barré, Ben C. Stöver, and Kai F. Müller**, LeafNet: A computer vision system for automatic plant species identification, 2019.

[26] **Unknown Author**, Comparative assessment of feature selection and classification techniques for visual inspection of pot plant seedlings, 2013.

[27] **Peter N. Belhumeur**, Searching the World's Herbaria: A System for Visual Identification of Plant Species, 2008.

[28] **Shivlika Pathania**, Phytochemica: a platform to explore phytochemicals of medicinal plants, 2015.

[29] **F. Freiburghaus**, Evaluation of African medicinal plants for their in vitro trypanocidal activity, 1996.