IJARASEM

ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 11, Issue 6, November - December 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.583

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 6, November-December 2024 |

Medicinal Plants Identification through Image processing and Machine Learning

Mr. G. Kiran Kumar [1], A. Rama [2], D.Sri Charan Reddy [3], R. Vinayak [4]

Assistant Professor, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India[*1]

UG Student, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India[*2, 3, 4]

ABSTRACT: The project is aimed at an arduous task of precise identification of medicinal plant species with the problem being pertinent in those industries that include botany, Ayurveda, pharmacology, and biomedical research. Most of the traditional identification methods are quite serious challenges for users, researchers, and students because they are usually time-consuming, knowledge-intensive, and prone to human errors. Our proposal develops an advanced web-based application for this process by utilizing state-of-the-art methods in image processing and machine learning. We will create a platform where users can upload or take pictures of plant specimens and obtain accurate identifications and comprehensive medical information by using machine learning algorithms with plant species image datasets. This application is done in a way that the interface would be user-friendly and favor research, identification, and learning, especially for Ayurvedic practitioners, biomedical specialists, botanists, and students. Therefore, our project is aimed at developing an exhaustive and reliable solution for medicinal plant identification and study by overcoming the deficiencies of existing systems, such as a small dataset, poor variable quality, and lack of complete medicinal information.

KEYWORDS: Medicinal plant identification, Deep Learning, Convolutional neural network, Transfer Learning, MobileNetV2, image classification.

I. INTRODUCTION

Medicinal plants are very important in health care through their medicinal values. Proper application of the medicinal plants, however, necessitates proper identification. The conventional ways of doing this are laborious and fallible to human error. Recent advances in machine learning opened new avenues in the application of computers for automation of plant identification by picture methods. Recent research has shown great utility of image processing techniques in medicinal plant identification and underlines the potential AI-enabled solutions may bring in the domain. Deep learning has proven promising as applied to improving picture classification tasks, especially through transfer learning by using pre-trained models like MobileNetV2.

Plant leaf classification has also been attempted with high success rates using convolutional neural networks. This project aims for the development of a highly scalable and efficient deep learning model capable of medicinal plant identification from images using augmentation techniques and transfer learning. The end product is an automated system that minimizes manual labor, makes fewer mistakes, and is applicable in most researches in botany and medicine.

II. RESEARCH METHODOLOGY

This is a project that uses deep learning specifically CNNs in the classification of images of medicinal plants' leaves to identify the medicinal plants. It relies on existing knowledge and image processing in order to accurately distinguish the varieties of plants, thus eventually leading to easy retrieval of their information. The research was useful in showing how automated plant classification can be used well in herbal medicine and farming but pointed out drawbacks such as data variety and model performance.

1. **Generation of Image Data and Preprocessing:** Generate images of medicinal plants for both the training set and validation set, retrieval from public-accessible databases. Preprocess the images to feed a model. One can also enhance the diversity in a dataset through augmentation techniques like flip and zoom along with random rotation of the image using Keras ImageDataGenerator in varied combinations. This enables a model to generalize properly for any variation in the images.
2. **Model Architecture**: The MobileNetV2 is pre-trained on the ImageNet dataset, and this will be used to develop the classification model. There is an application of transfer learning combined with addition of customized layers on top of the frozen foundation layers of MobileNetV2 to serve the purpose of categorization. These are custom layers with global average pooling layer followed by the dense fully related layers of 128 neurons with ReLU activation with softmax output layer for multi-class classification. The optimization by Adam optimizer after trainable categorical cross entropy.
3. **Training and Evaluating the Model:** The model is trained on the developed dataset with an 80/20 split between the training and the validation sets; it took 30 epochs, learning rate reduction was implemented when there was a halt in the improvement of validation loss and early halting to avoid overfitting. For the model, accuracy, precision, recall, and F1-score measurements were done, with an emphasis on the model's ability to classify the plant species in question. At such a point, the model was saved as an H5 file for later use once its results met with sufficiency.
4. **Web Application Development:** Web application was built using the streamlit framework, Python-based, which helps to develop interactive web applications. The users could upload plant images into the application and get real-time classification results. Before the image can be fed into the model for prediction, it will first pass through resizing it into 224x224 and normalizing pixel values. The saved medicinal_leaf_classifier.h5 model is loaded using TensorFlow. The online application contains a conversational assistant that interacts with the Google Gemini API. Users can ask queries regarding the classified plant after classification. The assistant will reply accordingly with relevant information. Real-time conversational feedback is provided with the help of the integrated Gemini chat-bot API.

III. THEORY AND CALCULATION

Deep learning, primarily CNNs, uses the concept of image classification to classify medicinal plants. The discipline of deep learning adds a new dimension to the processing of images through simpler tasks, as opposed to the earlier methods, which consumed more time and called for human expertise. These CNNs are neurons in layers that extract features from images; it allows the model to develop learning about complex patterns. This capacity is particularly helpful in the identification of very similar species of plants because this model can see the tiniest differences in the characteristics of these species.

This project adopted the architecture of MobileNetV2, which is efficient as well as very effective in the classification of images. MobileNetV2 uses depth-wise separable convolutions and reduces the computational complexity without losing too much accuracy. It is the reason why MobileNetV2 is particularly utilized for mobile devices as well as applications that have few available resources. Our model gets the knowledge from training this great abundance of images known as ImageNet through transfer learning with MobileNetV2.

This involves computation in the training of the model and assessment. This model resizes the images to standard pixels of 224 by 224. Then it normalizes the images such that the input is normalized. This model, in learning, tries to minimize the errors with its prediction and measures these deviations from the actual and the classification as computed from the loss function. These are accuracy, precision, recall, and F1-score. These are calculated after the training of the model. All of these give a representation in percent about how accurately the model can identify the plant. Calculations

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 6, November-December 2024 |

help understand the performance of the model; hence, advancements are being made in healthcare for automated identification of medicinal plants.

3.1. Mathematical Expression and Symbols

Many mathematical expressions and symbols are utilized for the purpose of describing processes of training and evaluating models.

Model Training: The training procedure computes the categorical cross-entropy loss function, given by

$$L(y,\hat{y}) = -\sum_{k=1}^{N} y_k \log(\hat{y}_k)$$

where y is the actual labels (one-hot encoded), ŷ represents the output predictions, and N is the number of classes. The optimizer used here is Adam which is a stochastic gradient descent optimizer for training the models on gradients computed from the gradients of the loss function to update the model weights towards finding minimum error.

Prediction and Evaluation Metrics: After training, the model's performance is evaluated using accuracy, precision, recall, and F1-score. Accuracy is calculated as:

$$\text{Accuracy} = \frac{Number\ of\ Correct\ Predictions}{Total\ Predictions}$$

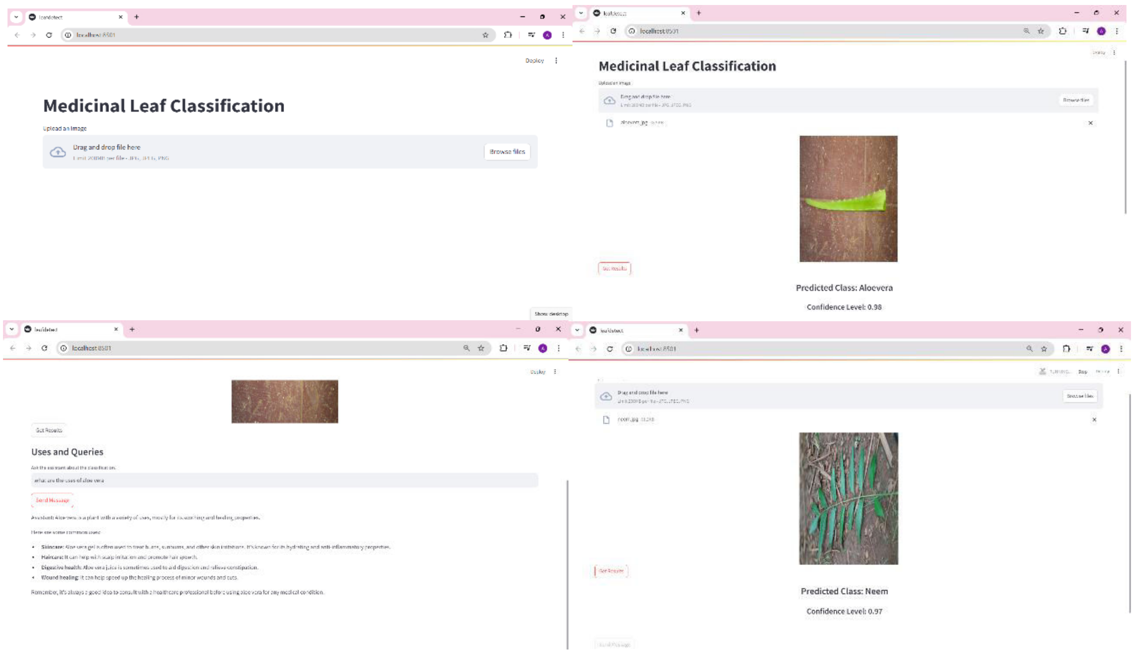
Precision and recall are calculated based on true positives (TP), false positives (FP), and false negatives (FN):

$$\text{Precision} = \frac{TP}{TP+FP} \text{ and Recall} = \frac{TP}{TP+FN}$$

The F1-score, which balances precision and recall, is given by:

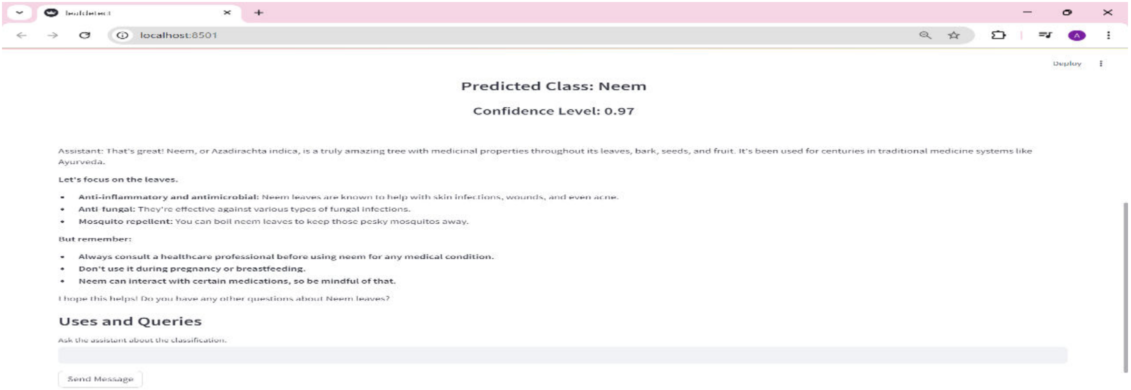
$$\text{F1 Score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

IV. RESULTS AND DISCUSSION



| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 6, November-December 2024 |



The strength of the results produced by our medicinal plant-identification model in classifying different types of medicinal plants based on a given image is extremely good. After developing this, I tried measuring how well the model was performing and what accuracy it would have on the dataset by showing it as quite effective in telling apart different classes of plants. Some classes clearly emerged in the confusion matrix, while for others it was hard to tell whose feature was pretty much similar.

Since the model was classified with very good precision, it was based on the utilization of convolutional neural networks. New research has come up with how CNNs are efficient in plant classification. Our approach justifies common issues met in the literature, such as the lack of availability of big labelled datasets for training. We applied the MobileNetV2 architecture, with transfer learning, and therefore exploited the pre-trained weights acquired from a much larger dataset to enable improvement in performance while folding the training.

Our study finally contributes to emergent work on plant medicinal identification, where there is great emphasis on the value placed on image processing techniques. Hence, by combining image pre-processing and deep learning-based approaches, this work would lay out a feasible solution in related herbal medicine and agriculture fields. Precision, recall, and F1-score of the model portray balanced performance across the classes of plants. Fine-tuning and data augmentation are further chances for improvement that would make it much more robust in realistic uses. Thus, deep learning may be quite feasible to be applied to the task of medicinal plant identification and serves as a foundation for possible further advancements. Further prospects for improvement in the project include finer tuning of the settings applied to the model, data augmentation to ensure increased variability for a diversified dataset. Real-world examples in several quantities can also make the model more reliable and adaptive.

V. CONCLUSION

This study would be possible to use convolutional neural networks to make species identification of medicinal plants a reality and to make differentiation of different plant species achievable by image analysis of their leaves. Its accuracy level is a bit high, potentially acceptable for practical usage in herbal medicine, agriculture, or conservation. Its performance may however vary with the factors such as its surroundings, the extent of the quality of images and the variety of different types of plants in the dataset for which some types of classes are hard to classify on the basis of similar leaf shapes. Future work on the scenario would involve creating a varied example of plants through the data set and using advanced techniques that could make the model more robust. This would further enhance the accuracy and range of this tool with regard to information about plant traits and habitats and, therefore, throw up identification problems and highlight the role of machine learning in plant identification and its place for future research into automated classification systems.

VI. ECLARATIONS

6.1 Study Limitations

The qualities and conditions of the images taken may actually vary, which would affect the model's potential accuracy in classification. Also, diversity in the dataset for some of the species was not so high, thus difficult to distinguish similar leaves from each other.

6.2 Funding Source

None.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 6, November-December 2024 |

6.3 Acknowledgements

My heartfelt thanks to all the Guides, Class Incharge, Project Coordinator, and Head of the department of Anurag University for their constant encouragement and motivating me in my research work.

6.4 Informed Consent

All participants involved in this research were given informed consent from them, ensuring that we were fully informed about the purpose of the study, how the study would be conducted, and the use of our data in the publication of this work.

REFERENCES

[1] Duth, Sudharshan & S, Vedavathi & S, Roshan. "Herbal Leaf Classification using RCNN, Fast RCNN, Faster RCNN" *7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)* August (2023). https://www.researchgate.net/publication/377607232_Herbal_Leaf_Classification_using_RCNN_Fast_RCNN_Faster_RCNN

[2] Valdez, Daryl & Aliac, Chris & Feliscuzo, Larmie "Medicinal Plant Classification using Convolutional Neural Network and Transfer Learning." *IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)* September(2022). https://www.researchgate.net/publication/365288547_Medicinal_Plant_Classification_using_Convolutional_Neural_Network_and_Transfer_Learning

[3] Revathy Prasannan, Preethy Roy, Jithi P V, Amritha R, Anjana V Nair, Ann Mariya Jose, "Medicinal Plants Identification Using Deep Learning: A Survey", *IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, (2024) https://ieeexplore.ieee.org/document/10690144

[4] Vishal Batvia, Drashti Patel, Dr. Avani R. Vasant "A Survey on Ayurvedic Medicine Classification using Tensor flow". International Journal of Computer Trends and Technology (IJCTT) V53(2):68-70, November (2017) https://ijcttjournal.org/archives/ijctt-v53p114

[5] Srikanth Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *International Journal of Scientific and Research Publications, Volume 9, Issue 10,* October(2019). https://www.researchgate.net/publication/337105858_Transfer_learning_using_VGG-16_with_Deep_Convolutional_Neural_Network_for_Classifying_Images

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com