Medicinal Plant Identification via ResNet50 Transfer Learning

[1]Dr. Rayapati Venkata Sudhakar, [2]Vadlapatla Bhavana, [3]Mendu Priya Spandana, [4]Manda Sowmya

[1]Associate Professor, [2,3,4] Undergraduate Students
Department of Computer Science and Engineering, Accredited by NBA
Geethanjali College of Engineering and Technology (UGC Autonomous), Affiliated to JNTUH, Approved by AICTE, Cheeryal(V)-501301, Hyderabad, Telangana, India.

***Abstract*- Traditional methods of identifying medicinal herbs present significant challenges to various stakeholders involved in herbal medicine. Herb collectors often rely on experience and expertise passed down through generations, which can be subjective and inconsistent. Researchers face hurdles in accurately cataloging and studying medicinal plant species due to manual identification processes' laborious and time-consuming nature. Medicinal Plant Identification project merges state-of-the-art technology with age-old botanical wisdom to offer a holistic herb identification and usage recommendation solution. ResNet50 model was trained for the project and an intuitive web interface was developed using Flask, users can effortlessly upload images of medicinal plants and receive precise predictions.**

Keywords: **ResNet50, medicinal plants, transfer learning, deep learning.**

I. INTRODUCTION

Identifying and utilizing medicinal herbs hold profound significance in traditional healing practices globally. India, renowned for its rich biodiversity of medicinal plants, has preserved this heritage for centuries. There's a rapid increment in the utilization of herbal-based items all over the world. Medicinal herbs play a major part in creating those drugs. However, as modern challenges intersect with traditional wisdom, innovative approaches to medicinal plant identification are imperative. Identification of herbs can be automated so this paper proposes a transfer learning technique that can help ordinary people and researchers to identify the herbs at a faster rate with great accuracy.

1.1 Motivation:

In the modern era, the manufacturing and marketing of Ayurvedic drugs have evolved into a thriving industry, generating substantial revenue. However, concerns regarding the quality of raw materials used in these medicines have emerged. With plants often collected by untrained individuals, incorrect or substituted medicinal plants are common, leading to ineffective or harmful medicines.

1.2 Problem statement:

Identifying medicinal herbs poses substantial challenges in herbal medicine. Traditional methods rely on subjective expertise and experience, leading to inconsistencies and risks in medicinal formulations. Additionally, manual identification processes hinder research efforts due to their time-consuming nature. There is a critical need for innovative solutions merging technology with botanical wisdom to provide accurate and efficient herb identification and usage recommendations.

.

II. SYSTEM ANALYSIS

2.1 Existing System:

In the existing system of medicinal plant identification, various approaches exist within the current system to address these challenges. Professionals manually identify and outline medicinal plants in images using computer software. However, this approach is slow and resource-intensive, limiting scalability and efficiency. Some systems utilize machine learning algorithms such as support vector machines or random forests to aid in plant identification. However, these approaches often necessitate manual feature engineering and parameter tuning, adding complexity to the process. Disadvantages of the existing system are:

- Manual segmentation of plants is slow and requires significant labor.
- Current methods may struggle with accurately identifying plants due to variations in their characteristics.
- The system relies heavily on the expertise of specialists or professionals, leading to delays and potential limitations in scalability due to limited access to such expertise.

2.2 Proposed System:

The proposed system for identifying medicinal plants from their leaves utilizes advanced deep learning techniques, specifically ResNet50, in conjunction with transfer learning. Through transfer learning, these pre-trained models are

fine-tuned to accurately classify and identify medicinal plant species based on leaf images. , ResNet50 is selected as the primary model for medicinal plant identification in the proposed system. Subsequently, a web application is developed using Flask, providing users with an intuitive interface to interact with the identification system. Users can upload images of medicinal plant leaves through the web interface, and the system employs ResNet50 to analyze the images and provide precise identification results. Advantages of the proposed system:

• Automated plant identification
• Enhanced accuracy with deep learning
• User-friendly web interface

III. IMPLEMENTATION

The implementation of the "Medicinal Plant Identification" project progresses through distinct phases, each focusing on specific workflow activities to construct different parts of the overall solution. Illustrated below are the key phases involved in project development:

1. **Training Phase**: The Training Phase encompasses workflow activities related to training the deep learning models for medicinal plant identification. This phase is delineated below:

• **Data Collection**: Gathering images of medicinal plants from the dataset “Indian Medicinal Leaves Dataset” from Kaggle, it is a repository that consists of medicinal plant images. The images are captured with varying backgrounds without any environmental constraints.

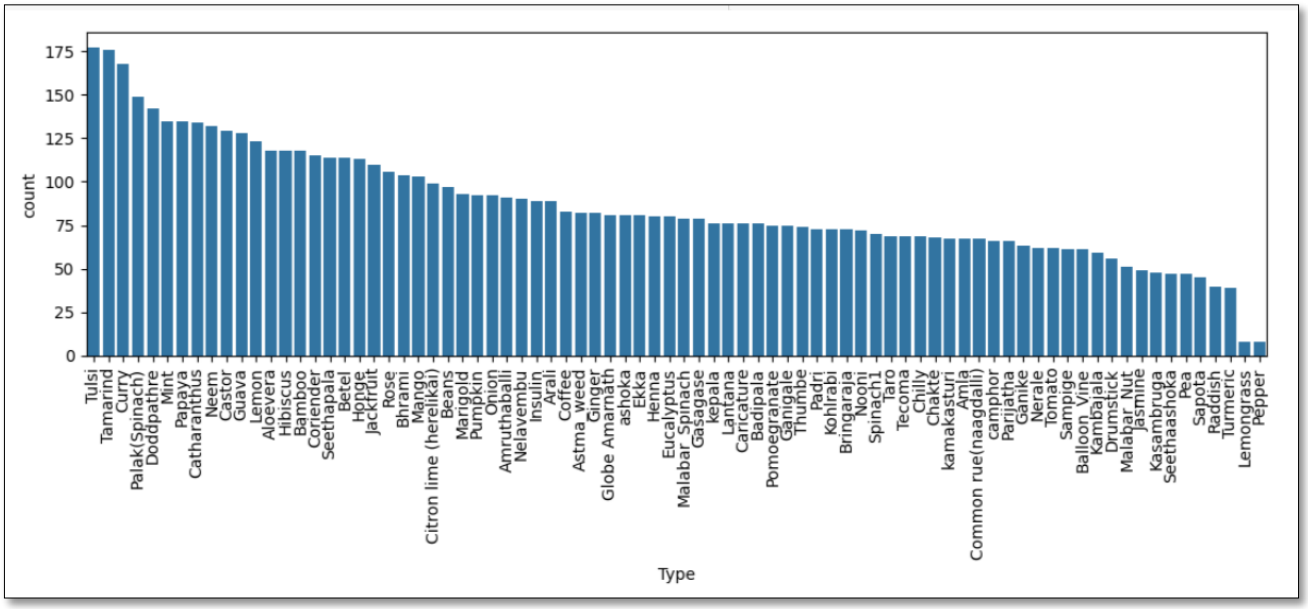

Fig 3.1 Number of images across 80 different classes

• **Data Preprocessing**: Partitioning the dataset into training and testing sets, resizing images to a standardized resolution, and dividing the data into 70% training, 15% validation, and 15% testing.

```
Found 4795 images belonging to 80 classes.
Found 1000 images belonging to 80 classes.
Found 1105 images belonging to 80 classes.
```

Fig 3.2 Number of images in train, validation, and test directories

• **Defining Model Architecture**: Designing and implementing deep learning architecture of ResNet50, for feature extraction and classification tasks.
• **Model Training**: Training the ResNet50 model using the training dataset, optimizing parameters using the Adam optimizer, and monitoring convergence.
• **Evaluating Model Performance**: Assessing the performance of the trained model using evaluation metrics such as accuracy, loss, and validation scores to ensure robustness and reliability.

```
Epoch 1/8
150/150 [==============================] - 2792s 19s/step - loss: 2.8184 - accuracy: 0.3483 - val_loss: 1.7058 - val_accuracy: 0.5690
Epoch 2/8
150/150 [==============================] - 1659s 11s/step - loss: 1.1279 - accuracy: 0.7103 - val_loss: 1.1054 - val_accuracy: 0.7010
Epoch 3/8
150/150 [==============================] - 1654s 11s/step - loss: 0.6505 - accuracy: 0.8323 - val_loss: 0.9476 - val_accuracy: 0.7210
Epoch 4/8
150/150 [==============================] - 1609s 11s/step - loss: 0.3975 - accuracy: 0.9049 - val_loss: 0.7695 - val_accuracy: 0.7790
Epoch 5/8
150/150 [==============================] - 1609s 11s/step - loss: 0.2597 - accuracy: 0.9416 - val_loss: 0.7474 - val_accuracy: 0.7880
Epoch 6/8
150/150 [==============================] - 1656s 11s/step - loss: 0.1803 - accuracy: 0.9625 - val_loss: 0.6227 - val_accuracy: 0.8140
Epoch 7/8
150/150 [==============================] - 1658s 11s/step - loss: 0.1085 - accuracy: 0.9819 - val_loss: 0.5975 - val_accuracy: 0.8240
Epoch 8/8
150/150 [==============================] - 1657s 11s/step - loss: 0.0772 - accuracy: 0.9881 - val_loss: 0.5535 - val_accuracy: 0.8300
```

Fig 3.3 Training Process

Fig 3.4 Sample images of the dataset

2. **Deployment Phase:** This phase focuses on deploying the trained model as a web application using Flask for seamless interaction. Key activities include:

• **Web Page Creation**: Creating a user-friendly web page where users can upload images of medicinal plants and receive predictions. The web page design includes intuitive elements for easy navigation and interaction.

• **Model Integration**: Integrating the trained ResNet50 model into the Flask application to perform predictions on uploaded images. This involves loading the model, preprocessing the input images, and generating predictions based on the model's output.

IV. SYSTEM CONFIGURATION

- **Software requirements**: Minimum software requirements are:
  - o Operating system: Windows 7/higher, Mac
  - o Server-Side Script: python 3.6+
  - o IDE: Jupyter Notebook
  - o Libraries Used:
    - ▪ NumPy
    - ▪ Pandas
    - ▪ TensorFlow
    - ▪ OS
    - ▪ Scikit-Learn
    - ▪ Keras
- **Hardware requirements**: Minimum hardware requirements are:
  - o RAM: 8GB
  - o Hard Disk: 128 GB

o Processor: I3/ Intel processor

V. OUTPUT

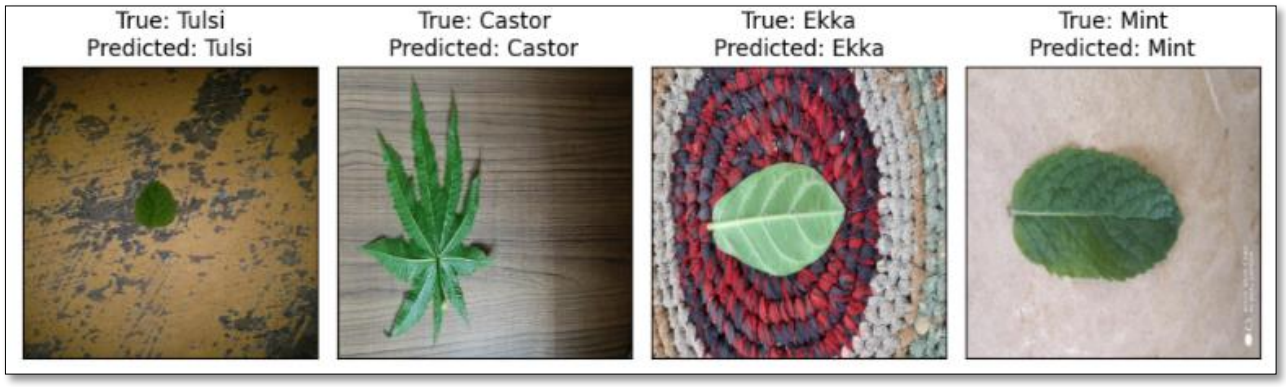

Fig 5.1 Output of model showing true label and the model predicted label

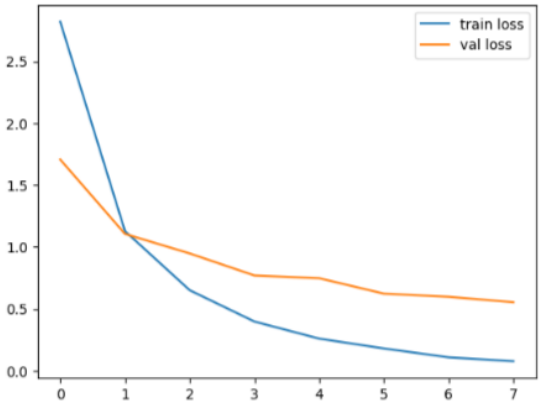

Fig 5.2 Epoch vs Loss graph

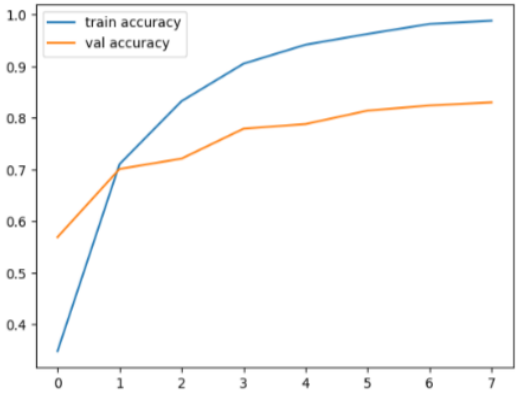

Fig 5.3 Epoch vs Accuracy graph

Fig 5.4 Home screen

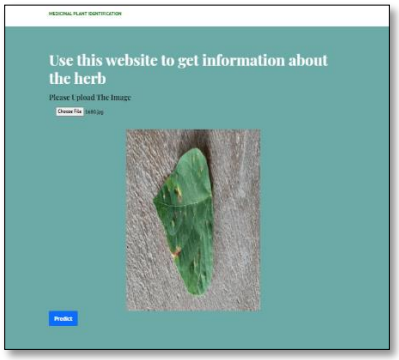

Fig 5.5 Uploading image

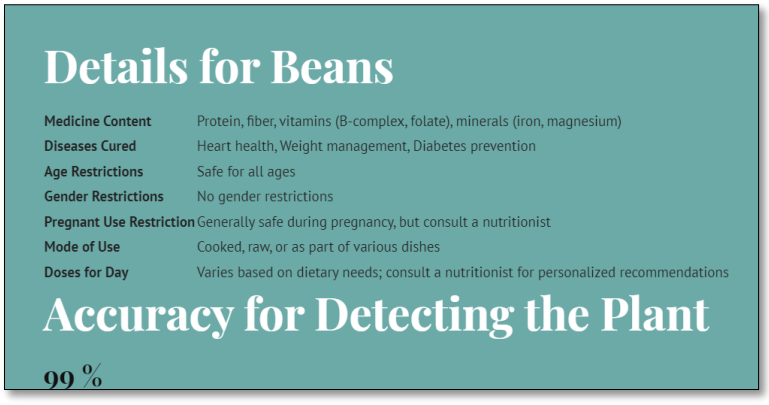

Fig 5.6 Final result

VI. CONCLUSION

The proposed system leverages deep learning's ResNet50 model for medicinal plant identification from leaf images. Transfer learning refines this pre-trained model for accurate classification. A Flask web application provides a user-friendly interface for uploading images and receiving identifications. The project implementation involves distinct phases. The training phase gathers data, preprocesses it, defines the ResNet50 architecture, trains the model, and evaluates its performance. The deployment phase creates a web interface using Flask and integrates the trained model for image analysis and prediction. This system offers automated plant identification with enhanced accuracy and a user-friendly web interface for ease of use.

VII. FUTURE SCOPE

- **Mobile Application Development**: Creating a mobile application would significantly improve accessibility and usability. Users could identify plants directly in the field, fostering on-the-go identification.
- **Dataset Expansion and Diversification**: Continuously adding new medicinal plant species and variations to the existing dataset is crucial. This will enhance the model's robustness and accuracy in identifying a wider range of plants. The dataset should include images with diverse characteristics, such as different lighting conditions, growth stages, and leaf orientations.

REFERENCES:

1. Prasvita, Desta Sandya, and Yeni Herdiyeni. "MedLeaf: mobile application for medicinal plant identification based on leaf image." *International Journal on Advanced Science, Engineering and Information Technology* 3.2 (2013): 5-8.
2. Fitzgerald, Martin, Michael Heinrich, and Anthony Booker. "Medicinal plant analysis: A historical and regional discussion of emergent complex techniques." *Frontiers in pharmacology* 10 (2020): 423244.
3. Mardiana, Bella Dwi, et al. "Herbal Leaves Classification Based on Leaf Image Using CNN Architecture Model VGG16." *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 7.1 (2023): 20-26.
4. Haq, Mohd Anul, Ahmed Ahsan, and Jayadev Gyani. "Implementation of CNN for plant identification using UAV imagery." *International Journal of Advanced Computer Science and Applications* 14.4 (2023).
5. Reddy, Satti RG, GP Saradhi Varma, and Rajya Lakshmi Davuluri. "Optimized convolutional neural network model for plant species identification from leaf images using computer vision." *International Journal of Speech Technology* 26.1 (2023): 23-50.
6. Sun, Yu, et al. "Deep learning for plant identification in natural environment." *Computational intelligence and neuroscience* 2017 (2017).
7. Lee, Sue Han, et al. "Deep-plant: Plant identification with convolutional neural networks." *2015 IEEE international conference on image processing (ICIP)*. IEEE, 2015.