



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Identification and Classification of Medicinal Plants Species Based on Leaf Structure Using Deep Learning Techniques

Dr. S B Mallikarjuna ¹, Prof. Shwetha G ², Abhinit Raj³, Harsha A Sindhe⁴, Lakshmi M S⁵, Roja A S⁶

Head of the Department, Department of Artificial Intelligence and Machine Learning, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India¹

Assistant Professor, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India²

U.G. Student, Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India^{3 4 5 6}

ABSTRACT: Identification of the correct medicinal plants that goes in to the preparation of a medicine is very important in ayurvedic, folk and herbal medicinal industry. The main features required to identify a medicinal plant is its leaf shape, colour and texture. Colour and texture from both sides of the leaf contain deterministic parameters to identify the species. In this project we explore feature vectors from both the front and back side of a green leaf along with morphological features to arrive at a unique optimum combination of features that maximizes the identification rate. A database of medicinal plant leaves is created from scanned images of front and back side of leaves of commonly used medicinal plants. The leaves are classified based on the shape and dimension combination. It is expected that for the automatic identification of medicinal plants this system will help the community people to develop their knowledge on medicinal plants, help taxonomists to develop more efficient species identification techniques and also participate significantly in the pharmaceutical drug manufacturing.

KEYWORDS: Convolutional Neural Network, Deep Learning Medicinal plant species.

I. INTRODUCTION

Medicinal plants have been a vital source of healthcare and wellness for centuries, with their diverse array of bioactive compounds holding significant therapeutic potential. However, the accurate identification and classification of these plants, as well as the diagnosis of diseases affecting them, remain challenging tasks. Traditional methods relying on manual observation and expertise are often time-consuming and prone to errors. In this context, the application of deep learning techniques offers a promising avenue for automating and enhancing the process of medicinal plant identification and disease classification. Deep learning, a subset of machine learning inspired by the structure and function of the human brain, has revolutionized various fields, including computer vision and natural language processing. By leveraging complex neural network architectures and vast amounts of data, deep learning algorithms can extract intricate patterns and features from raw input, making them particularly well-suited for tasks such as image recognition and classification. In the domain of medicinal plants, deep learning holds the potential to significantly expedite the identification of plant species based on visual characteristics.

A Convolutional Neural Network is a type of Deep Learning neural network architecture commonly used in Computer Vision. Computer vision is a field of Artificial Intelligence that enables a computer to understand and interpret the image or visual data.

When it comes to Machine Learning, Artificial Neural Networks perform really well. Neural Networks are used in various datasets like images, audio, and text. The success of deep learning relies heavily on the availability of large amounts of labelled data for training, as well as significant computational power for training complex models. Advances in hardware, such as GPUs and TPUs, have greatly accelerated the training process, making it feasible to train deep learning models on massive datasets.

Additionally, the open-source community and research institutions have contributed to the development of deep learning frameworks and algorithms, democratizing access to these powerful tools. Additionally, deep learning algorithms encompass a variety of architectural designs tailored to different types of data and tasks.

Below Figure 1 shows few leaves of medicinal plants. All of these Leaves have different shapes and structure. This helps in Identifying and Classifying the leaves. To do this we have used Deep Convolutional Neural Network, Machine Learning Architecture.









Medicinal LeafName	Front View	Back View
Tulsi		
Peppermint		
Basel		
Lemon		

Fig 1: Shows Different Medicinal Plant Leaves

II. LITERATURE SURVEY

Begue et.al. (2019): Utilized computer vision and machine learning method for automatic recognition of medicinal plants. Author collected data from tropical island of Mauritius. It consist of 720 images of distinct plant species The shape and colour features are extracted for classification and applied 10- fold cross validation procedure using random forest classifier and achieved 70.1% accuracy. [1]

Bahri et.al. (2019): Developed a Dynamic CNN (DCNN) classification model for Morocco aromatic and medicinal plants. The author created dataset of 16 plant species with 320 images. In this study, entropy impurity was used to dynamically integrate CNN classification results. Inception V3, ResNet50, and VGG16 is employed. The dynamic Resnet50 method used by CNN was 77.4% accurate. [2]

Sachar and Kumar (2020): Created a deep ensemble learning model for automatically identifying medicinal leaves, and utilizing pre-trained models like ResNet50 InceptionV3 and MobileNetV2, their accuracy on the test set was 79.66%, and their average accuracy was 89.9%. 30 classes make up the dataset for medicinal leaves. [3]

Dileep et. Al (2020): Fifty medical images from Google images were used for edge detection. When CNN's texture patch was applied for classification, the findings revealed that it was 98% accurate. His work proposes AyurLeaf, a deep learning-based CNN model to classify medicinal plants using shape and colour characteristics of the leaves. 40 therapeutic plants species leaf samples datasets used. [4]

Pearline et.al (2021): 185 plants image dataset is utilized. Alex Net, GoogLeNet, VGG19, ResNet50, and MobileNetV2 these 5 CNN based model are pretrained. 82.3% greatest accuracy achieved by MobileNetV2 with ADAM optimizer. A plant identification system was created. [5]

Karahan and Nabiyevev (2022): Using the pretrained network MobileNetV2. 5,345 flower and plant photos from 76 different species were utilized in the collection. The suggested model had accuracy on the training set of 79.71% and accuracy on the test set of 78.97% after 15 iterations. In their CNN- LSTM network design. [6]

Banzi and Abaya (2023): 20 layers were included with the SoftMax function for classification. An open collection of 100 plant species photos was used to train the models. According to experiments, the suggested CNN-LSTM works better than the convolutional CNN in classifying plant species because it achieves an accuracy of 75.06%. Table I summarized the classification methods of ML and DL of this study. [7]

Sharma et al. (2019): This survey paper reviews recent trends in plant leaf recognition using deep learning techniques, with a focus on CNN models. It discusses various methodologies for leaf image acquisition, preprocessing, feature extraction, and classification. The paper also examines the influence of factors such as dataset size, diversity, and quality on the performance of deep learning models. Furthermore, it discusses challenges related to model interpretability, scalability, and deployment in real-world applications. [8]

Nazni et al. (2019): This survey paper explores the application of deep CNN models for the classification of medicinal plant leaves. It covers a wide range of topics, including preprocessing techniques, dataset selection, model architectures, and performance evaluation metrics. The paper provides insights into the strengths and weaknesses of different CNN approaches and discusses potential avenues for enhancing classification accuracy and robustness. [9]

III. METHODOLOGY

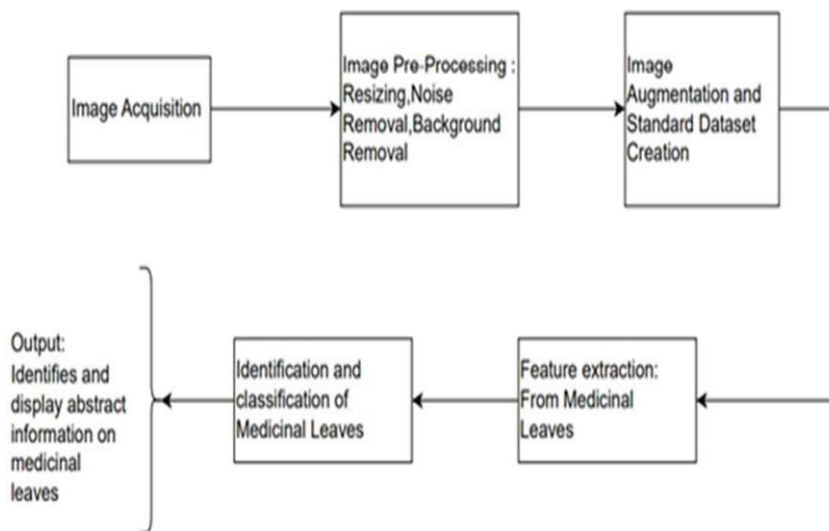


Fig 2 : Methodology of the proposed system

Description of Methodology

Data Acquisition:

The data is collected using high-resolution camera with 13 MP. These sample data was taken from university of agriculture, Bangalore. When taking images of the plant leaves, to avoid data augmentation, the different light intensity strength and direction of leaf positions were considered. Backsides of the leaves were captured to extract vein feature accurately. We made the focal length different depending on the broadness of leaves to make the prediction consistent. The dataset contains medicinal plant leaf images.

The following guidelines are taken for this data acquisition: • Medium aged and young parts of the plant leaf are selected.

The focal length of the camera (distance between leaves and camera) is taken depending on Broadness and narrowness of the plant leaves.

The white background of the image was selected to avoid confusion and get a clear leaf image structure.

Dried and diseased leaves excluded to make the prediction consistent • The leaf images are captured immediately after cut out from the plants to reduce the loss of leaf features.

The camera stands perpendicular to the leaf stand (i.e. the position taken 90° from the leaf stand).

Data Pre-Processing:

Data pre-processing in deep learning involves preparing raw data for training machine learning models. This includes cleaning data, normalizing numerical features, handling categorical data, augmenting data for diversity, and addressing preprocessing may involve tokenization and word embedding. Time series data may require resampling and feature extraction. Data splitting into training, validation, and test sets is crucial, and efficient data pipelines are designed for on-the-fly processing. Effective pre-processing ensures that the deep learning model can learn effectively and generalize well to new data.

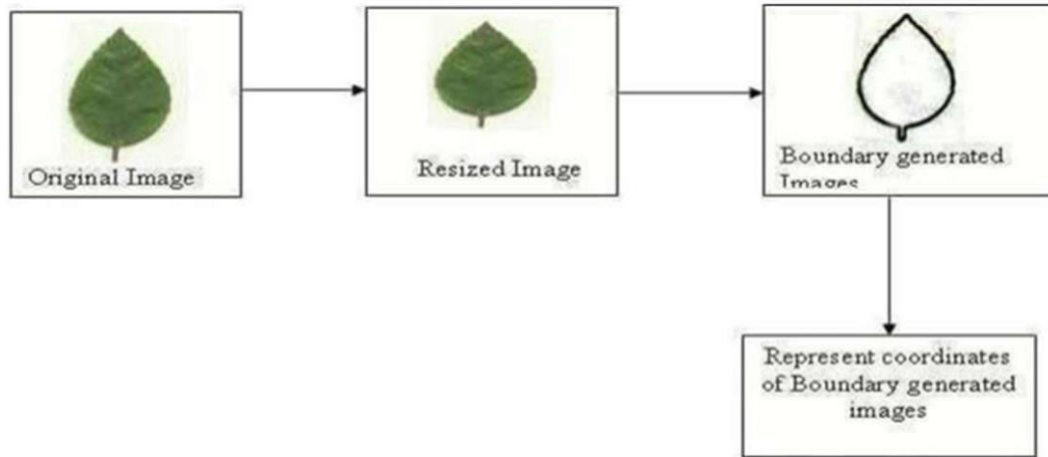


Fig 3: Data Preprocessing Technique Adopted

Data Augmentation:

Data augmentation in deep learning involves artificially expanding the size of a training dataset by applying various transformations to existing data. This technique is commonly used for image data and includes operations like rotation, flipping, zooming, and brightness adjustment. The goal is to enhance the model's ability to generalize to new, unseen data by exposing it to a more diverse set of examples during training. Data augmentation is implemented using parameters that control the degree of transformation, and it helps reduce over fitting by introducing randomness and variability. The technique is applicable to various data types, and deep learning frameworks often provide tools for its straightforward implementation.

Flow Diagram

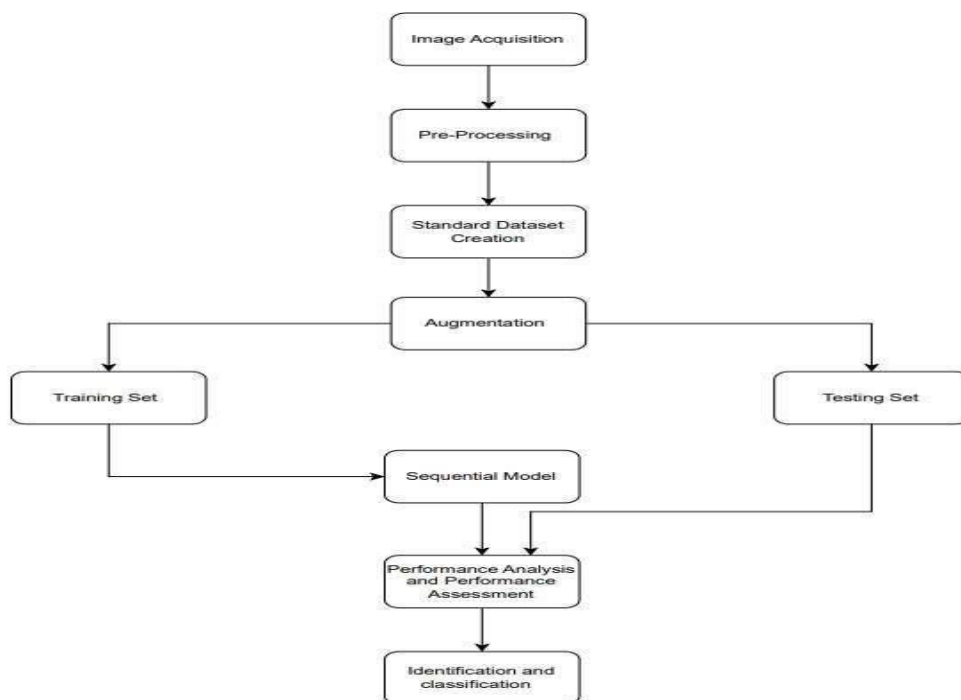


Fig 4: Flow Graph of the Proposed System



IV. EXPERIMENTAL SETUP

Step 1: Setting up Python Environment

Install Python: Download and install Python Step 2: Install anaconda prompt

Step 3: Install all the libraries pip install pillow

pip install pandas pip install seaborn

pip install django==3.0

pip install tensorflow==1.15.0 pip install keras==2.2.4

Step 4: Data Collection: We have capture images from multiple angles and under different lighting conditions to improve the model's robustness.

We have performed quality checks on the collected data to ensure accuracy. Step 5: Data Training

Import all the libraries run the command Python training.py

Step 6: Results \ output by running the server by command python manage.py runserver

V. EXPERIMENTAL RESULTS

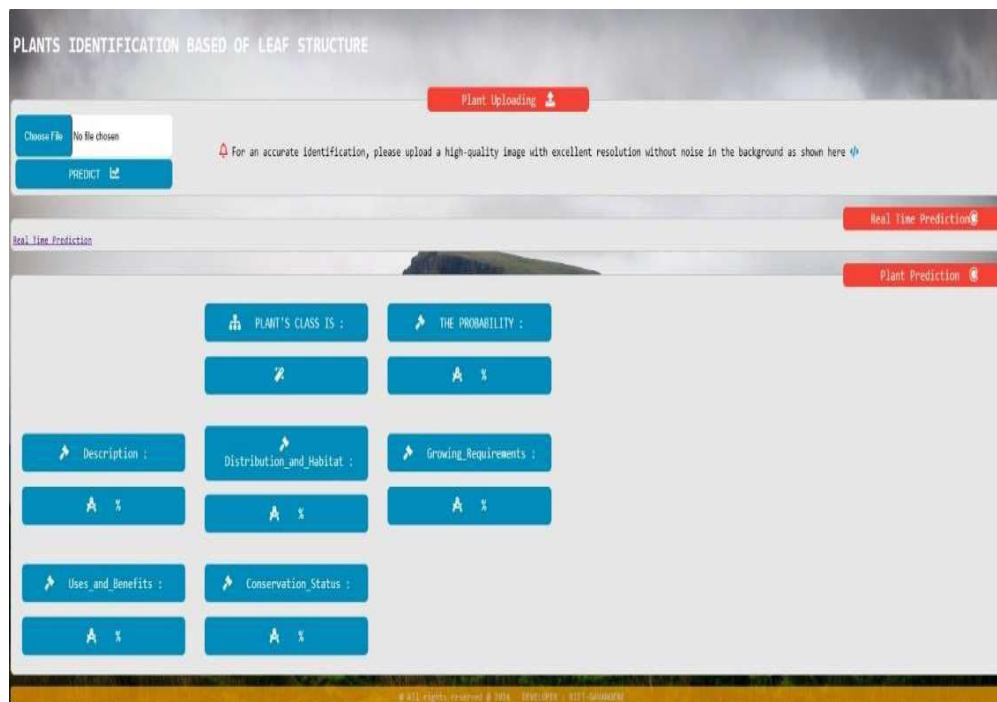


Fig 5: Home Page

It is a template of home page which consists of plant Class Name, Probability, Description, Growing Requirements, Distribution and habitat, use and benefits and conservation status.

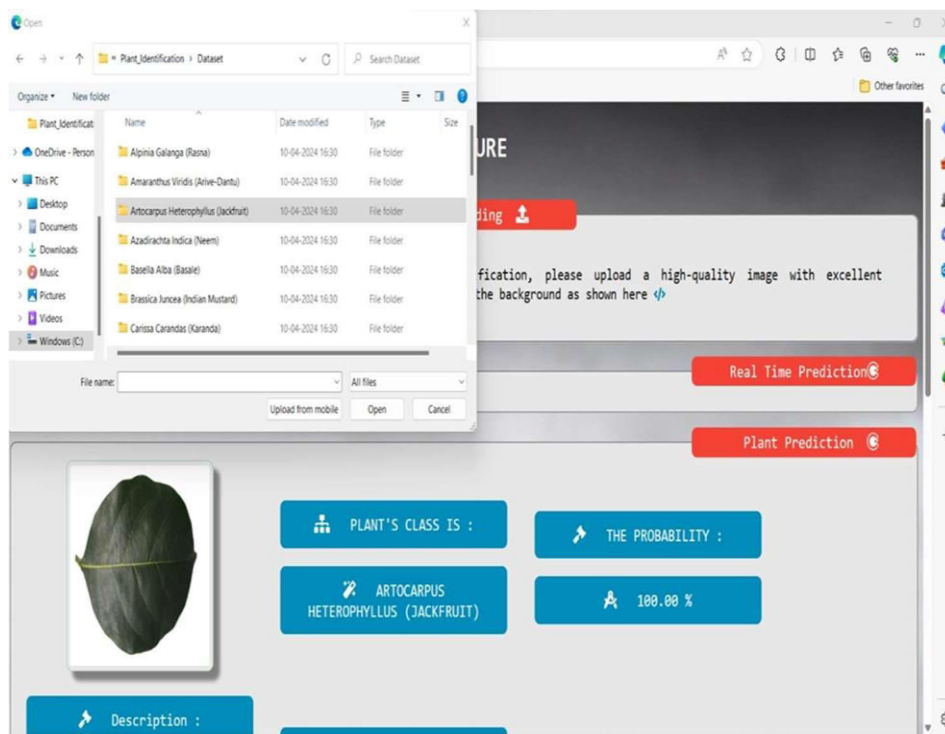


Fig 6 : Select Dataset

To predicate the medical plant leaf click on the button provided on home page and select the pretrained dataset

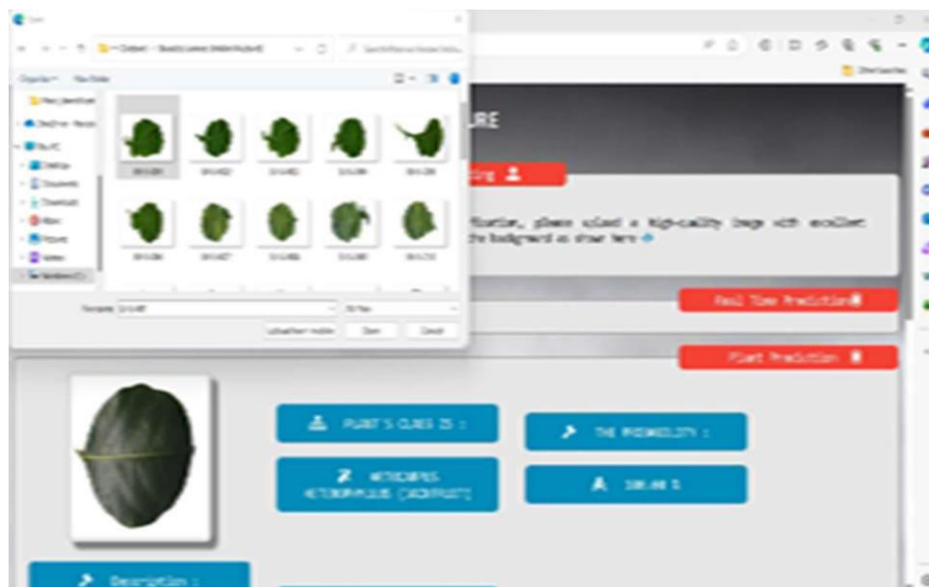


Fig 7: Select Medicinal Leaves

Once the dataset folder is selected choose any one medical plant leaf image among 45 images

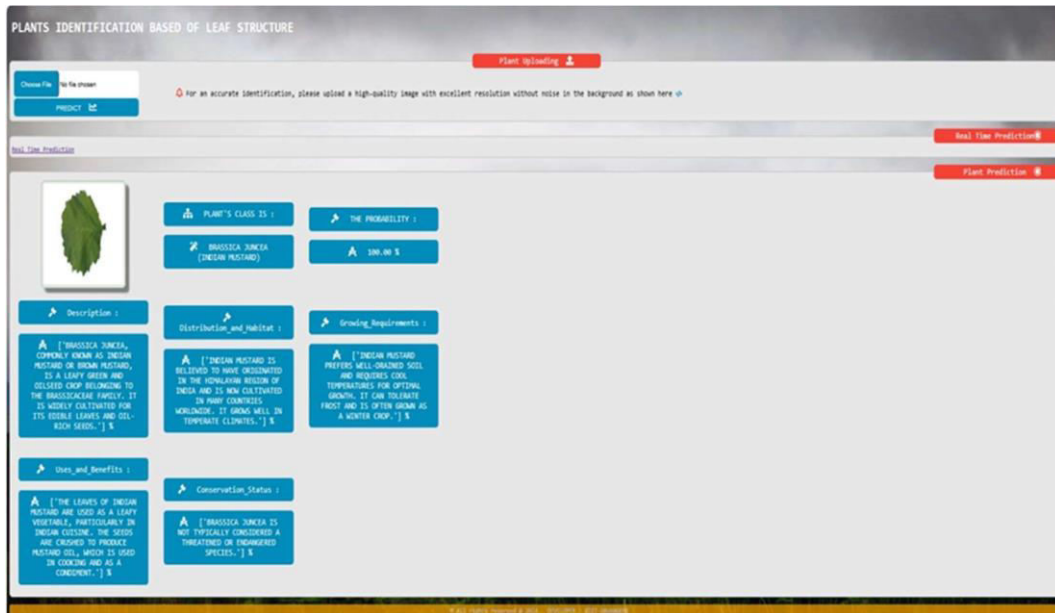


Fig 8: Description of Medicinal Leaves

once the leaf is selected click on the predict button present in the home page it will display the leafname, probability, benefits and uses.

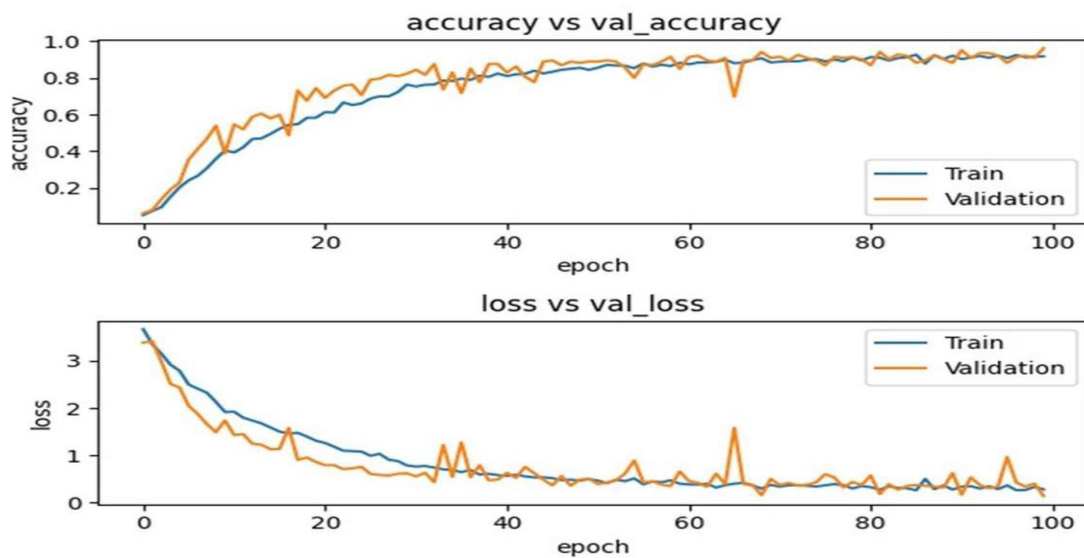


Fig 9: Accuracy Graph

Accuracy and Loss is plotted depending on epoch, here we have taken 100 epoch

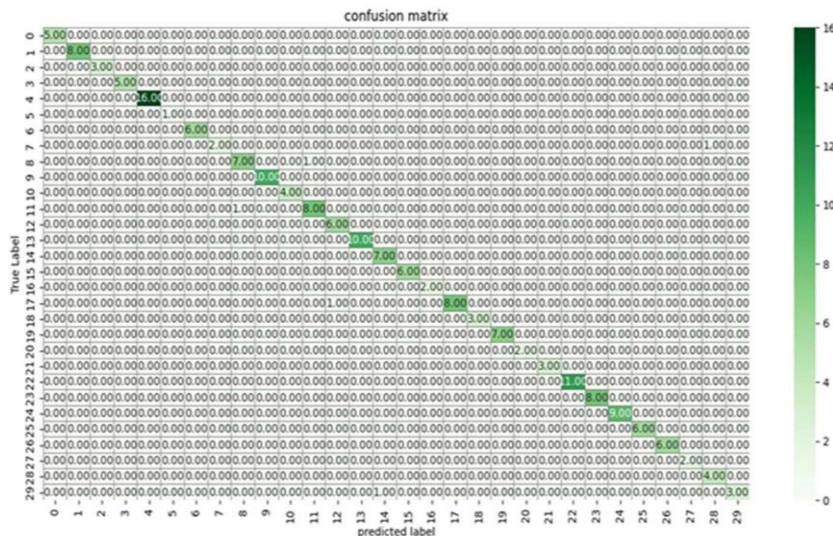


Fig 10: Confusion Matrix

The accuracy of a model is calculated by the ratio of correctly predicted instances to the total instances in the dataset. It is calculated as $(TP + TN) / (TP + TN + FP + FN)$, where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives,

VI. CONCLUSION

Deep learning offers precise and efficient identification and classification of medicinal plant species through accurate feature extraction from images. It enables high accuracy, scalability, we have collect various leaves of medicinal plants and processed the data sets for training and testing . as Data augmentation techniques enhance dataset size and model robustness. at last we are able to identify and classify the leaves of medicinal plants with high accuracy above 95 percentage

REFERENCES

1. **Begue et.al. (2019):** Utilized computer vision and machine learning method for automatic recognition of medicinal plants. Author collected data from tropical island of Mauritius. It consist of 720 images of distinct plant species The shape and color features are extracted for classification and applied 10- fold cross validation procedure using random forest classifier and achieved 70.1% accuracy.
2. **Bahri et.al. (2019):** Developed a Dynamic CNN (DCNN) classification model for Morocco aro aromatic and medicinal plants. The author created dataset of 16 plant species with 320 images. In this study, entropy impurity was used to dynamically integrate CNN classification results. Inception V3, ResNet50, and VGG16 are employed. The dynamic Resnet50 method used by CNN was 77.4% accurate.
3. **Sachar and Kumar (2020):** Created a deep ensemble learning model for automatically identifying medicinal leaves, and utilizing pre-trained models like ResNet50 InceptionV3 and MobileNetV2, their accuracy on the test set was 79.66%, and their average accuracy was 89.9%. 30 classes make up the dataset for medicinal leaves
4. **Dileep et. Al (2020):** Fifty medical images from Google images were used for edge detection. When CNN's texture patch was applied for classification, the findings revealed that it was 98% accurate. His work proposes AyurLeaf, a deep learning-based CNN model to classify medicinal plants using shape and colour characteristics of the leaves. 40 therapeutic plants species leaf samples datasets used.
5. **Pearline et.al (2021):** 185 plants image dataset is utilized. Alex Net, Google Net, VGG19, ResNet50, and MobileNetV2 these 5 CNN based model are pretrained. 82.3% greatest accuracy achieved by MobileNetV2 with ADAM optimizer. A plant identification system was created.
6. **Karahan and Nabiye (2022):** Using the pretrained network MobileNetV2. 5,345 flower and plant photos from 76 different species were utilized in the collection. The suggested model had accuracy on the training set of 79.71% and accuracy on the test set of 78.97% after 15 iterations. In their CNN- LSTM network design.



7. **Banzi and Abaya (2023):** 20 layers were included with the SoftMax function for classification. An open collection of 100 plant species photos was used to train the models. According to experiments, the suggested CNN-LSTM works better than the conventional CNN in classifying plant species because it achieves an accuracy of 75.06%. Table I summarized the classification methods of ML and DL of this study.
8. **Sharma et al. (2023):** This survey paper reviews recent trends in plant leaf recognition using deep learning techniques, with a focus on CNN models. It discusses various methodologies for leaf image acquisition, preprocessing, feature extraction, and classification. The paper also examines the influence of factors such as dataset size, diversity, and quality on the performance of deep learning models. Furthermore, it discusses challenges related to model interpretability, scalability, and deployment in real-world applications.
9. **Nazni et al. (2023):** This survey paper explores the application of deep CNN models for the classification of medicinal plant leaves. It covers a wide range of topics, including preprocessing techniques, dataset selection, model architectures, and performance evaluation metrics.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details