MEDICINAL PLANTS IDENTIFICATION BY IMAGE PROCESSING USING DEEP LEARNING TECHNIQUES AND IMPLEMENTING SUPPLY CHAIN INTEGRITY

V Karthick, Associate Professor
Department of CSE
Rajalakshmi Engineering College
Chennai, India
vkarthick86@gmail.com

Varunesh B, UG Student
Department of CSE
Rajalakshmi Engineering College
Chennai, India
210701303@rajalakshmi.edu.in

Vinoth N, UG Student
Department of CSE
Rajalakshmi Engineering College
Chennai, India
210701311@rajalakshmi.edu.in

Udhaychander R J, UG Student
Department of CSE
Rajalakshmi Engineering College
Chennai, India
210701294@rajalakshmi.edu.in

***Abstract*— This research presents a novel method for identifying medicinal plants using deep learning and image processing. The system uses a Convolutional Neural Network (CNN) trained on a large dataset of medicinal plant photos to extract and analyze essential visual aspects, classifying plants into their respective species. The study also proposes integrating supply chain integrity measurements to address concerns about the authenticity and quality of medicinal plants throughout the supply chain. Potential solutions include blockchain technology or tamper-proof labeling systems. The proposed system provides a reliable tool for plant identification, promoting responsible conservation and use of medicinal plant resources. It also ensures the quality and authenticity of therapeutic items, building market trust and encouraging ethical sourcing. The study demonstrates the potential of deep learning and supply chain integrity measures in transforming the discovery,handling,identification of medicinal plants.**

Keywords—Deep Learning,Convolutional Neural Network,Supply chain Integrity,Blockchain,Image processing

I. INTRODUCTION

Medicinal plants have played an important part in both traditional and modern healthcare.Their vast range of bioactive chemicals provides therapeutic benefits for a variety of conditions.The Convolutional Neural Network (CNN) [1]-based system can accurately categorize plants based on visual cues from a large dataset of medicinal plant photos. The system also emphasizes the importance of maintaining supply chain integrity [2] for medicinal plants. By integrating measures like blockchain technology and tamper-proof labeling, the system provides transparency and traceability, ensuring every step of a plant's journey is documented. This creates trust and confidence in the supply chain. The system also links captured plant images with relevant blockchain[3] data or information in tamper-proof labels, protecting the quality and authenticity of medicinal plants. This approach contributes to medicinal plant research by combining accurate identification with strong supply chain integrity. The proposed system encourages responsible conservation and use of valuable medicinal plant resources, while incorporating supply chain integrity measures ensures the quality and authenticity[4] of medicinal products. The research opens the path for future advances in the field by proving the power of deep learning and blockchain technology in transforming the discovery and handling of medicinal plants.

The work by Bella Dwi Mardiana, Wahyu Budi Utomo, Ulfah Nur Oktaviana, Galih Wasis Wicaksono, Agus Eko Minarno 2023 [5] was carried out to increase accuracy, this study suggests a method for classifying Medicinal plants using a pretrained VGG16 model and a convolutional neural network (CNN). In order to improve image quality and lessen overfitting, augmentation is used. Ten kinds of herbal leaves make up the dataset, which also includes testing, validation, and training sets. Comparing the results to earlier studies, the accuracy rose from 82% to 95%, with augmentation being a major contributing factor. Evaluations of precision, recall, and F1-score provide additional support for the suggested method's efficacy.It doesn't implement supply chain integrity.
The work by Lawrence C. Ngugi, Moataz Abdelwahab, Mohammed Abo-Zahhad 2023 [6] clearly explains the comparison to lesion-based image analysis, this study's semi-automated approach achieves an improvement of approximately 15% in the accuracy of leaf disease recognition. With an accuracy of 84.48%, the suggested KijaniNet model shows state-of-the-art segmentation accuracy, scoring highly on mean boundary F1 scores and mean intersection over Union (mIoU).The paper greatly enhances the accuracy of nearly 85% but doesn't contribute to the authenticity of medicinal plants and varying angles of the images.

II. MATERIALS AND METHODS

Collect a broad assortment of photos depicting numerous therapeutic plants. Ensure that the dataset represents a variety of species, growth stages, and environmental circumstances.Labeling the photos with plant species-specific information.In the next steps, photos are preprocessed [7] to improve quality and standardize them for consistency. Techniques such as scaling, normalization, and augmentation can be used to improve the model's resilience. Image processing is carried out utilizing the Convolutional Neural Networks (CNN) technique.The selected CNN model is trained on the training set and optimized to reduce classification errors.Then, install authentication procedures to assure the accuracy of identification results throughout the supply chain.

Hardware requirements of the project include

- Laptop or Personal Computer

Software requirements include

- Internet browser (Chrome/Edge/Mozilla Firefox)
- Jupyter Notebook
- Streamlit API
- TensorFlow
- MongoDB
- Firebase
- Polygon
- Metamask
- Etherium
- Solidity

III. EXISTING SYSTEM

A classification method based on leaf pictures of medicinal herbs. The technique will first preprocess medicinal plant leaf photos, then calculate 10 shape and 5 texture features, and then identify medicinal plant leaves using a support vector machine (SVM)[8] classification. The classification model was used to identify twelve different medicinal plant leaf photos, achieving an average success rate of 93.3%. The findings show that by using multi-feature extraction of leaf images in conjunction with SVM, it is possible to automatically categorize therapeutic herbs. The article provides an informative conceptual framework for medicinal herb categorization model research and development.

H Chanyal and RK Yadav [2022] [9] created an information interchange from object recognition to plant genetic research that employs deep qualities to describe the original plant leaf image. Experiments have revealed that such deep features exceed current methods for identifying plant species. The study presented an innovative and successful leaf collection method. The image is then converted to a device-independent lab color space, which is used to create the VGG-16 [10] feature map. This feature map is re-projected to the PCA subspace to improve species identification accuracy. The study employs two types of plant leaf collections to demonstrate durability.

Turkoglu and Hanbay [2019] [11] used image processing methods such as color, vein properties, Fourier Descriptors, and Gray-Level Co-occurrence Matrix (GLCM) methodologies to extract features from images. Rather than acquiring qualities for the complete leaf, researchers recommend collecting features from leaves divided into two or four parts. The Extreme Learning Machines (ELM) [12] classifier computes the individual and aggregate performances of each attribute extortion method. The Flavia leaf database was utilized to evaluate the suggested approach. The proposed technique's efficiency was assessed using 10-fold cross-validation, which was then compared and tabulated with methodologies from prior studies. On the Fluvial leaf database, the suggested method's results were rated as 94.12%. Nuril Aslina, Nursuriati Jamil et al. used the Scale Invariant Feature Transform (SIFT) [13] as a shape descriptor.

IV. PROPOSED SYSTEM

4.1. Dataset

The herbal leaf dataset is now hosted in an online repository. The dataset consists of herbal leaf pictures in.jpg format.The images have a resolution of 1600 by 1200 pixels.The herbal leaf dataset consists of 20 classes:Abelmoschus moschatus medik (Ambrette), Aloe vera (L.) Burm.F (Aloe Vera), Hibiscus rosasinensis (Red Hibiscus), Kaempferia, Galanga (Aromatic ginger), Kalanchoe Pinnata (Lam.) Pers (Miracle leaf), Lasia Spinosa (L.) Thwaites (Lesia), Lawsonia inermis L. (Henna), Leucas aspera Link (Thumba), Menthol. Benth. Ex Kurz (Serpentine root), Rotheca serrata (L.) Steane & Mabb. (Clerodendrum, Bharangi), Vanilla planifolia (Flat-leaved vanilla), Vitex negundo L. (Chinese Chaste Tree), Zanthoxylum nitidum DC. (Shiny-leaf prickly-ash), Zingiber officinale Rosc. (Ginger rhizome), Ziziphus Jujuba Mill. Figure 4.1.1 is an illustration of a sample dataset.In regard to the studies cited in [14],With reference to research referenced in the dataset's makeup is divided by the proportion.Parismita Sarma, "MED117_Medicinal Plant Leaf Dataset & Name Table" (accessed at 14:27 on September 21, 2022, from https://data.mendeley.com/datasets/dtvbwrhznz/1)80% of the data were used for training, 10% for validation, and 10% for testing.Table 1 displays the dataset classes.

Table 4.1.1 Herbal Plants classes data

SNO	CLASS
1	Abelmoschus moschatus medik(Ambrette)
2	Aloe vera (L.) Burm.F(Aloe Vera)
3	Hibiscus rosa sinensis(Red Hibiscus)
4	Kaempferia Galanga(Aromatic ginger)
5	Kalanchoe Pinnata (Lam.) Pers(Miracle leaf)
6	Lasia Spinosa (L.) Thwaites(Lesia),
7	Lawsonia inermis L.(Henna)
8	Leucas aspera Link(Thumba)
9	Mentha arvensis L(Corn Mint)
10	Mesua ferrea L.(Nagkesar)
11	Mimusops elengi L.(Spanish cherry)
12	Nyctanthes arbor-Tristis L.(Night Blooming Jasmine)
13	Psidium guajava L.(Guava Seed)
14	Rauvolfia serpentina Benth. Ex Kurz(Serpentine root)
15	Rotheca serrata (L.) Steane & Mabb.(Clerodendrum, Bharangi)
16	Vanilla planifolia(Flat-leaved vanilla)
17	Vitex negundo L.(Chinese Chaste Tree)
18	Zanthoxylum nitidum DC.(Shiny-leaf prickly-ash)
19	Zingiber officinale Rosc.(Ginger rhizome)
20	Ziziphus Jujuba Mill.(Jujube)

4.2. Model Architecture

The architecture model and the layers that make up the architecture, which are Conv2D, MaxPooling2D, Flatten, and Dense Layer, are displayed in Tables 4.2.1. A parameter value and an output shape are generated by each layer. The suggested model architecture is more sophisticated than the previous and uses a pretrained model VGG16 architecture, with higher parameter values and an output shape. The suggested model first performs the convolution process and filters the input image by resizing its dimensions to 150 x 150. Following convolution, the Pooling Layer [14] is applied, and if it is complete, the subsequent convolution is applied, changing the dimensions and allowing the Pooling Layer to continue. Moreover, the process of the Convolution and Pooling Layer will occur once more.The last Pooling Layer output will be processed by the Flatten Layer once the Convolutional and Pooling Layer has been finished at the last stage. One by one, the outcomes from the Flatten Layer will be added to the Dense Layer. Depending on the number of classes that are classified, the Dense Layer in the final layer has a value of 20.

Figure 4.2.1 CNN Equation

$$\dot{x}_{ij} = -x_{ij} + \sum_{kl \in S_{ij}(r)} a_{kl} y_{kl} + \sum_{kl \in S_{ij}(r)} b_{kl} u_{kl} + z_{ij}$$

$$y_{ij} = f(x_{ij}) = \frac{1}{2}\left(|x_{ij}+1| - |x_{ij}-1|\right)$$

$$i = 1, 2, \ldots, M, \; j = 1, 2, \ldots, N$$

Figure 4.2.2 Proposed Model Architecture

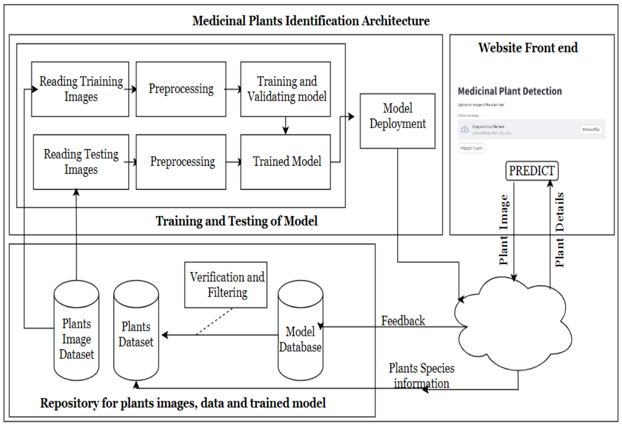


Table 4.2.2 Proposed Model Layers

Layer (type)	Output Shape	Param #
Conv2D_1 (Conv2D)	(None, 150, 150, 32)	9,248
Conv2D_2 (Conv2D)	(None, 148, 148, 32)	0
Maxpool_1 (MaxPooling2D)	(None, 74, 74, 32)	0
dropout (Dropout)	(None, 74, 74, 32)	
Conv2D_3 (Conv2D)	(None, 74, 74, 64)	36,928
Conv2D_4 (Conv2D)	(None, 72, 72, 64)	0
Maxpool_2 (MaxPooling2D)	(None, 36, 36, 64)	0
dropout_1 (Dropout)	(None, 36, 36, 64)	
Conv2D_5 (Conv2D)	(None, 36, 36, 128)	1,47,584
Conv2D_6 (Conv2D)	(None, 34, 34, 128)	0
Maxpool_3 (MaxPooling2D)	(None, 17, 17, 128)	0
flatten (Flatten)	(None, 36992)	0
Dense_1 (Dense)	(None, 512)	1,89,40,416
dropout_2 (Dropout)	(None, 512)	0
Dense_2 (Dense)	(None, 128)	65,664
Output (Dense)	(None, 6)	774

4.3.Training and Testing

A callback function is implemented in this study and used in the model training procedure. The purpose of this callback is to halt the model training procedure when val_accuracy reaches the given value. If there has been a rise in the val_accuracy matrix value, storage is carried out once each epoch. This callback is also used to retrieve the weight that the model ultimately uses, which is learned from the best epoch. Due to the potential for issues, this callback is also crucial for scheduling the learning speed. The number of epochs increasing is the factor that can lead to this issue.Tests will be conducted in a number of scenarios,

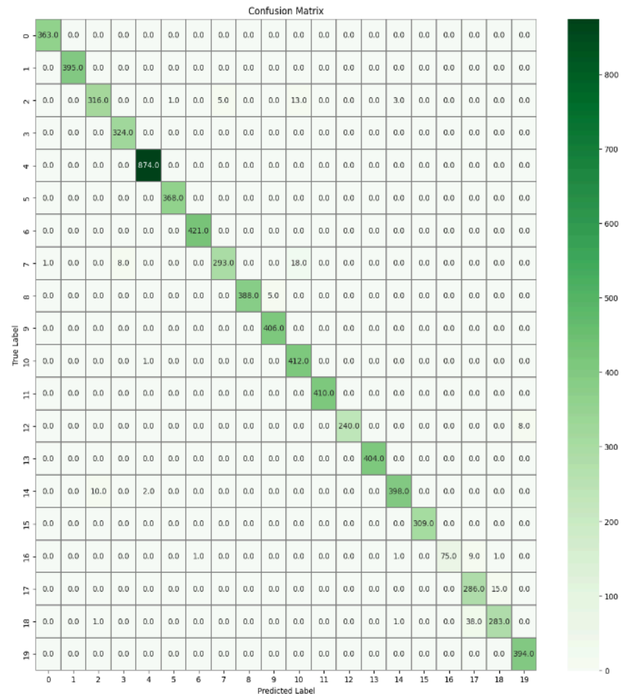
Number of training files : 9377
Number of training targets : 9377

V. RESULTS AND DISCUSSION

2.1.Confusion Matrix

The proposed model is evaluated and the confusion matrix [15] for the trained model is attached in below Figure 5.1.1

Figure 5.1.1



2.2.Training and Testing Accuracy Graph

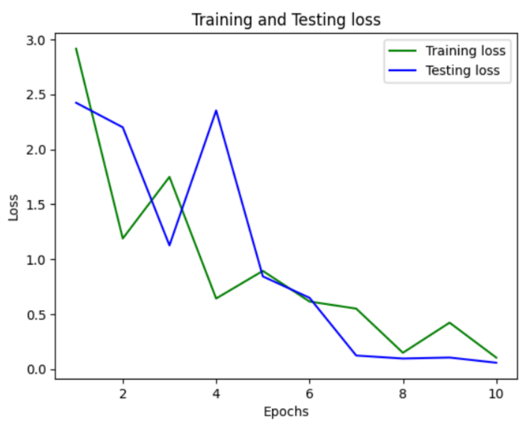
The proposed model is evaluated and the testing and training accuracy graph is obtained.The training and testing accuracy of the model is attached in the format of line graph with epochs in x-axis and accuracy in the y-axis,where one blue line indicates Testing accuracy,Green line indicates Training below figure 5.2.1

Figure 5.2.1

2.3.Training and Testing Loss Graph

The proposed model is evaluated and the testing and training loss graph is obtained.The training and testing loss rate of the model is attached in the below figure 5.3.1

Figure 5.3.1



VI. CONCLUSION

The proposed approach can accurately identify and detect the kind of herbal leaves when combined with the root mean square propagation with multiple layers of neural network model and the augmentation process on the training dataset used for medicinal leaf image classification.The collection of herbal leaf datasets that are converted into training, validation, and testing datasets makes up the complete dataset used in this study. The quality of the leaf picture obtained during testing has an impact on the medicinal leaf classification procedure.The quantity of training data is another element that influences the classification accuracy value.The model will learn a lot and perform better the more training data it uses, but it will also take longer.Training data accuracy can be increased to 97.89% and testing data accuracy to 98.17% by using the Fully Connected Layer, a Dense Layer type in the proposed model, and the results of the classification that uses the Convolutional NeuralNetwork (CNN) method for image classification using augmentation (Image Data Generator) and the addition of layers. The testing data's accuracy value may drop to 96% if the augmentation technique is not used.With the various kinds of herbal leaves utilized, it can be inferred that the application of the method used in this research is able to produce very good results with the accuracy value of the testing data higher than the primary reference journal.

REFERENCES

[1] I. Volkmer, News in the Global Sphere: A Study of CNN and Its Impact on Global Communication. Indiana University Press, 1999.

[2] C. Crossley, *Software Supply Chain Security*. "O'Reilly Media, Inc.," 2024.

[3] S. Inam, S. Kanwal, R. Firdous, and F. Hajjej, "Blockchain based medical image encryption using Arnold's cat map in a cloud environment," *Sci. Rep.*, vol. 14, no. 1, p. 5678, Mar. 2024.

[4] D. Chhabra, *Authenticity and Authentication of Heritage*. Routledge, 2021.

[5] Herbal Leaves Classification Based on LeafImage Using CNN Architecture Model VGG16

[6] A new approach to learning and recognizing leaf diseases from individual lesions using convolutional neural networks Lawrence C. Ngugi, Moataz Abdelwahab, Mohammed Abo-Zahhad 2023

[7] Comparison of data preprocessing approaches for applying deep learning to human activity recognition in the context of industry 4.0-X Zheng, M Wang, J Ordieres-Meré - Sensors, 2018

[8] A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting A Kurani, P Doshi, A Vakharia, M Shah - Annals of Data Science, 2023.

[9] Classification of medicinal plants leaves using deep learning technique: A review H Chanyal, RK Yadav, DKJ Saini - … Journal of Intelligent Systems,2022.

[10] A deep learning based model for the detection of pneumonia from chest X-ray images using VGG-16 and neural networks S Sharma, K Guleria - Procedia Computer Science, 2023

[11] Plant disease and pest detection using deep learning-based features M Türkoğlu, D Hanbay Turkish Journal of Electrical- 2019

[12] A review on extreme learning machine J Wang, S Lu, SH Wang, YD Zhang - Multimedia Tools and Applications, 2022

[13] Iris Features Extraction and Recognition based on the Scale Invariant Feature Transform (SIFT) MA Taha, HM Ahmed, SO Husain - Webology, 2022

[14] A five-layer deep convolutional neural network with stochastic pooling for chest CT-based COVID-19 diagnosis

[15] Confusion matrix-based modularity induction into pretrained CNN-S Ahmad, SU Ansari, U Haider, K Javed… - Multimedia Tools,2022.