Medical Leaf Classification using Deep Learning

[1] Shruthi M T, [2] Soundarya N
[1].Assistant professor ,Department of MCA ,BIET ,Davangere
[2].Student, Department of MCA, BIET, Davangere

Abstract

This study suggests classifying medicinal plant leaves using a machine learning-based method. The following six types of medicinal plant leaves are included in a dataset that was gathered from the Department of Agriculture: peppermint (Mentha balsamea), bael (Aegle marmelos), lemon balm (Melissa officinalis), catnip (Nepeta cataria), stevia (Stevia rebaudiana), and tulsi (Ocimum sanctum). These leaves are captured digitally and in multispectral utilizing a computer vision lab configuration.

Leaf sections are clipped and turned into gray-level pictures during the preprocessing stage. The Sobel filter is thenused for edge and line detection based on seed intensity, and five regions of interest are drawn. Extracted are 65 fused features that combine multispectral properties, run- length matrix, and texture. The dataset is reduced to 14 optimized features by the useof the chi-square feature selection approach for feature optimization. The improved dataset is then subjected to five machine learning classifiers: bagging, random forest, simple logistic, multi-layer perceptron, and logit-boost. Of the classifiers studied, the multi-layer perceptron classifier shows the most potential accuracy.

Keyword: Medicinal plant leaves, Machinelearning, Image classification, Texture features.

1. INTRODUCTION

A vital role that the classification of leaves from medicinal plants plays in botany, agriculture, and medicines. For a variety of reasons, including the creation of herbal remedies and biodiversity preservation, it is crucial to recognize and differentiate between various species of medicinal plantsbased on the features of their leaves. Recent

developments in machine learning and computer vision have completely changed the way that plant species are identified. Researchers can accurately classify leaves by capturing precise information about their spectral properties and shape by utilizing techniques like digital and multispectral imaging. Six types of leaves from medicinalplants are the subject of this study: Bael (Aegle marmelos), Lemon Balm (Melissa officinalis), Catnip (Nepeta cataria), Peppermint (Mentha balsamea), Tulsi (Ocimum sanctum), and Stevia (Stevia rebaudiana). These plants are noteworthy not only for their therapeutic qualities but also for the wide variations in color, texture,and form of their leaves.

Preprocessing methods including leafregion clipping and gray-level image conversion are part of the methodology. The Sobel filter's edge and line detection improves the ability to extract significant features from the leaf image data even more. A full set of 65 fused features, including multispectral properties, texture descriptors, and run-length matrix statistics,are extracted. A chi-square feature selection strategy is used to produce a subset of 14 optimal features and enhance feature selection and classification performance. Then, the dataset is subjected to five machine learning classifiers: multi-layer perceptron, logit-boost, bagging, randomforest, and simple logistic, to assess how well they can differentiate between the six types of medicinal plants.

This study's main goal is to evaluate how well these machine learning classifiers perform the task of reliably classifying medicinal plant leaves using attributes that have been retrieved. It is anticipated that theresults will have a major impact on automated plant species identification systems, supporting both research and real- world uses in agriculture and medicine.

Objective

1. Create an automated system that uses cutting-edge machine learning to recognize the leaves of medicinal plants.

2. Boost species identification reliability by using deep learning techniques to increase classification accuracy.

3. By expediting the processes involved in plant identification, the pharmaceutical and herbal medicine industries can promote speedier drug discovery and botanicalresearch.

2. LITERATURE REVIEW

This research's primary goal is to create an automated system that uses photos of plant leaves to reliably identify species of medicinal plants. This method attempts to eliminate human mistake in plant identification, which can have unfavorable impacts in Ayurvedic healthcare, by utilizing cutting-edge technology like computer vision and machine learning. A convolutional neural network (CNN) trained on a self-created dataset of 4,390 pictures from 35 different medicinal leaf species is used in the Android application. The system's stated accuracy is 94.10%.[1]

The primary goal of this research is to create a deep learning-based system that uses leaf photos to reliably identify different kinds of medicinal plants. With the help of a convolutional neural network (CNN) that is based on the VGG-16 model, the suggested method obtains a remarkable 98% recognition rate. By making it possibleto accurately identify and classify medicinal plants and removing the obstacles presented by their complicated appearance, this dependable technology has the potential to have a substantial impact onresearch on herbal medicine.[2]

The main objective of this study is to createa deep learning-based system that can recognize medicinal plant species with accuracy from photographs of their leaves. With the help of a convolutional neural network (CNN) that is based on the VGG- 16 model, the suggested method obtains a remarkable 98% recognition rate. By making it possible to accurately identifyand classify medicinal plants and removing the obstacles presented by their complicated appearance, this dependable technology has the potential to have a substantial impact on research on herbal medicinal[3]

Using deep neural networks and transfer learning techniques, the primary goal of this research is to develop an automated systemfor the identification of several medicinal leaf species. Using cutting-edge deep learning models such as VGG16, MobileNet, and InceptionResNetV2, the research attempts to effectively classify medicinal plants. Users can take pictures ofleaves while on the go with the resulting smartphone app, which also offers extradetails like Indian and botanical names andapplications for the detected leaves. This work is noteworthy because it applies transfer learning in a novel way to datasetsof medicinal leaves that are accessible to the public.[4]

Your system uses the enhanced vegetation index known as ExG-ExR to identify plantspecies from leaf samples. This index doesnot require a user-defined threshold andperforms effectively in difficult lighting circumstances. Color and texture cues are utilized for classification, and leaves are segregated as sub-images. Achieved accuracy is 93.3%.[5]

3.METHODOLOGY

Using the computer vision software package OpenCV, the dataset photos are analyzed to initiate the proposed methodology for categorizing leaves of medicinal plants. The Sobel filter is then used to detect edges and lines, which improves the ability to distinguish differentfeatures in the leaf images. In order to define regions of observation (ROOs) based on a pixel threshold value greater than six, this procedure uses a seed intensity approach. The pixel intensity patterns of these ROOs indicate important regions of interest.

A complete set of fused features is recovered from these regions after picture preprocessing and ROO identification. These features include a variety of variables, including multispectral attributesfrom the leaf images, statistical metrics from run-length matrices, and texture

descriptors. Together, these combined characteristics effectively convey theminute nuances and distinctive qualities of every kind of medicinal plant.

In the following stage, machine learning methods are integrated for classification using the fused features that have been extracted. To accurately classify the leaves of medicinal plants, models like logistic regression, random forest, and multi-layer perceptrons are used. A labeled dataset is used for model training, and performance measurements including as accuracy,precision, recall, and F1-score are used for performance validation to make sure the classification system is robust and reliable.

descriptors. Together, these combined characteristics effectively convey theminute nuances and distinctive qualities of every kind of medicinal plant.

In the following stage, machine learning methods are integrated for classification using the fused features that have been extracted. To accurately classify the leaves of medicinal plants, models like logistic regression, random forest, and multi-layer perceptrons are used. A labeled dataset is used for model training, and performance measurements including as accuracy,precision, recall, and F1-score are used for performance validation to make sure the classification system is robust and reliable.

Figure 1: A suggested framework for classifying leaves of medicinal plants utilizing a machine learning approach with multispectral and textural features.

This methodology's progression from image preprocessing and feature extraction to machine learning-based categorization is described in figure1.It draws attention to the ways in which every stage advances themain objective of automating and optimizing the classification of leaves frommedicinal plants, hence enabling uses in herbal medicine, botanical research, and pharmaceuticals. By improving the

efficiency and precision of plant species identification, this methodology hopes to contribute to larger initiatives forsustainable agriculture and biodiversityconservation.

This study's methodology uses machine learning techniques to develop and use an accurate classification system for medicinalplant leaves in a methodical manner.

A. Data Collection and Preparation

This study's dataset includes six different varieties of leaves from medicinal plants that are titled in Latin Tulsi (Ocimum sanctum), Peppermint (Mentha balsamea), Bael (Aegle marmelos), Lemon balm (Melissa officinalis), Catnip (Nepeta cataria), and Stevia (Stevia rebaudiana), respectively. The dataset is carefully gathered in a computer vision lab environment under carefully monitored circumstances, guaranteeing its noise-free nature. These plants are described in Figure2

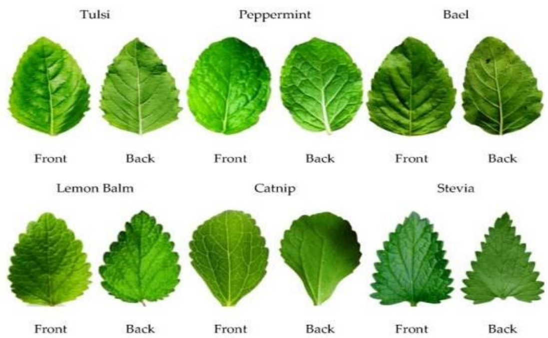


Figure 2: Sample of the medicinal plants leaves belonging to the dataset.

The popular computer vision package OpenCV is used to process the digital photographs of leaves from medicinal plants during the image preparation step. In order to identify leaf boundaries and features, the Sobel filter is used to find edges and lines within the images it describes figure 3. To make further analysis easier, every leaf image is then transformed to an 8-bit gray-level format and cropped to a standard size of 800 × 800pixels.

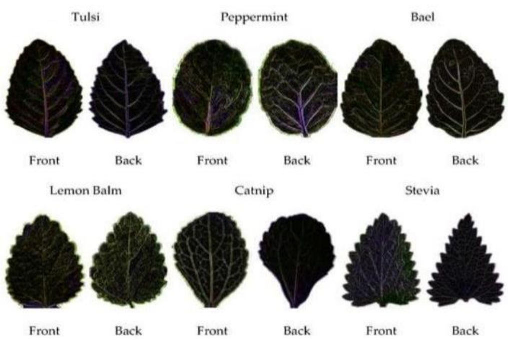


Figure 3: Sample of Transformation RGB to edge/line detection process.

On every sample image, five regions of observation (ROOs) are drawn in order to examine the leaf properties in more detail.There are two steps to this: First, ROOs with dimensions of 220 x 220 pixels are defined, and then ROOs with dimensions of 280 x 280 pixels. As a result of this process, several datasets with ROOs ofvarious sizes are produced with the intention of examining various experimental setups.it described in figure 4.

Each medical plant variety generates 1000ROOs (5 ROOs per leaf type, 200 sampleseach), for a total of 6000 ROOs created throughout the six plant kinds. This method guarantees thorough coverage of leaf characteristics, which makes it easier to conduct in-depth testing and analysis inthe study's later phases.

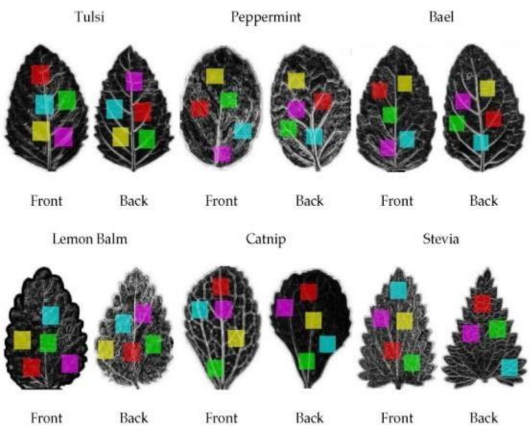


Figure 4: Sample of transformation intogray level and draw 5 colorful regions of

observation on medicinal plant leavesdataset.

B. Image Preprocessing

The system's data processing component manages the collection and processing of photographs of leaves from medicinal plants. Users can choose file locations for retrieval or use a graphical interface to addphotographs. The system makes use of OpenCV to carry out operations including noise reduction, image normalization, and edge detection using the Sobel filter. Theseprocedures enhance the quality of the images and guarantee that they are prepared for feature extraction andadditional analysis.

C. Feature Extraction:

One essential part of the system is feature extraction, which takes pre-processed leaf images and extracts pertinent properties andfeatures from them. Shape, texture, and color are among the features that are examined to identify the unique patterns and traits that distinguish various plant species. To effectively extract these features, the system uses techniques like Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG). The features that are extracted are used as input data for classification models in both machine learning and deep learning.

D. Machine Learning Model:

The training and use of conventional machine learning algorithms for leaf classification are covered by the machine learning model component of the system. For this, classification algorithms like Random Forests, Support Vector Machines(SVM), and k-Nearest Neighbors (k-NN) may be used. A labeled dataset of previously processed leaf images is utilizedto train the model. The extracted features are used as input features, and the labels relating to the plant species are used astarget labels. Based on patterns and relationships it has learned between input features and target labels, the trained modelis able to identify new photos of leaves.

E. Deep Learning Model:

The system has a deep learning model component for more sophisticated and complex feature extraction and categorization in addition to standard machine learning algorithms. A Convolutional Neural Network (CNN) architecture is utilized for leaf grouping.Transfer learning approaches are used to pre-trained CNN models (e.g., VGG16, ResNet) in order to take advantage of characteristics that have been learned from extensive picture datasets such as ImageNet.

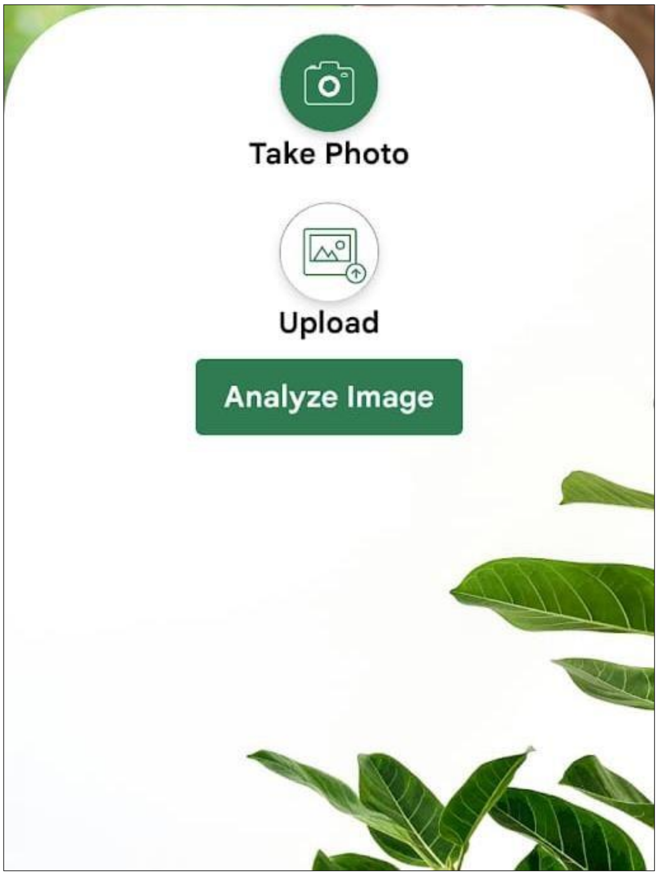
Using the dataset of leaves from medicinalplants, the pre-trained model is adjusted to make it particularly suitable for leaf classification tasks. The deep learningmodel uses the data's complex patterns and hierarchical representations to improveclassification accuracy.

F. Real-time Classification:

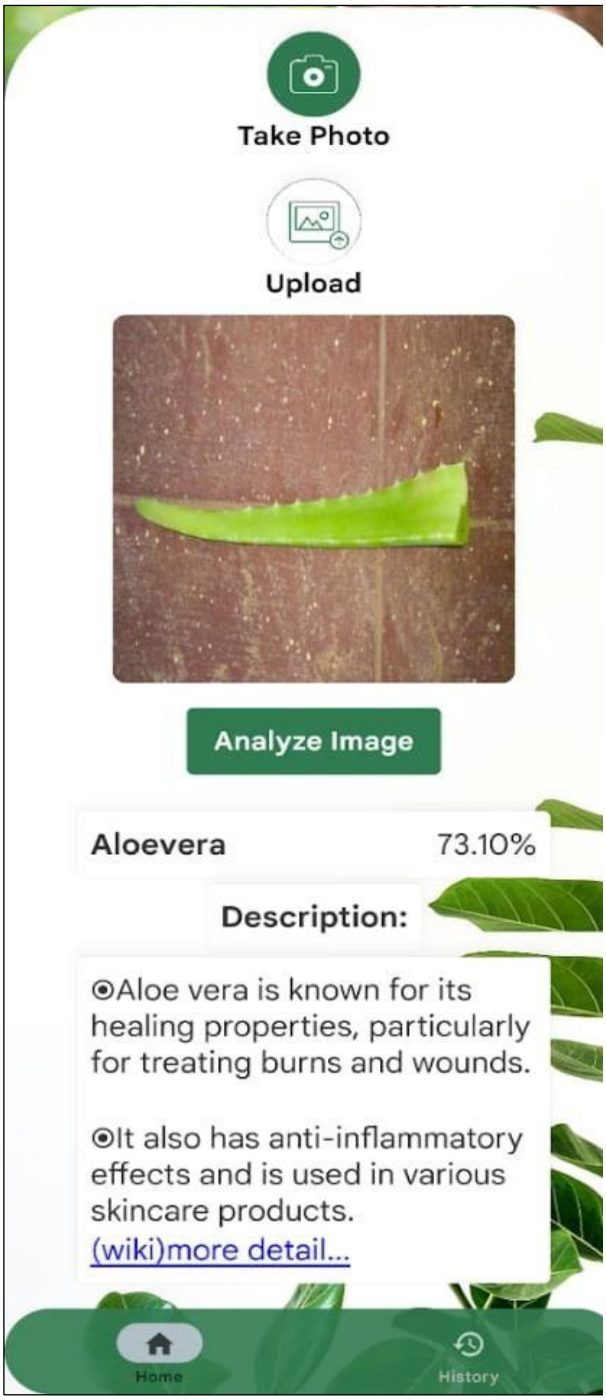
Users can obtain immediate categorization results for medicinal plant leaves thanks tothe system's real-time classificationcomponent. Through the GUI, users can interact with the system by entering photographs of leaves and starting the classification process. The system functions real-time feedback on theprojected plant species and confidence scores based on the input photos utilizing the trained machine learning and deep learning models. The GUI presents classification results to users, allowing for easy interaction and decision-making.

3. RESULT

Step 1: This result, shows the user takes a leaf image and uploads it to the page.



Step 2: This result describes an analysis of a particular leaf image, and it shows a description of the image.



4. CONCLUSION

In conclusion, this study successfully classified medicinal plant leaves using sophisticated image processing and machine learning techniques. To improve dataset quality for feature extraction and classification, we used the Sobel filter for edge detection and OpenCV for preprocessing. Creating several observational regions allowed for a detailed examination of the spectral and textural characteristics. These features were used totrain and assess machine learning models, such as random forest and multi-layer perceptron, which demonstrated encouraging results for automated plant species identification. This methodology enhances the precision of classifying medicinal plants and has potential applications in both agriculture and pharmaceutical research. These techniques could be further improved in the future for use in industry and more comprehensive botanical research.

REFERENCES

1. Singh, A., Sharma, P., & Sharma, D. (2019). "Medicinal Plant Identificationusing Deep Learning Techniques." International Journal of Computer Applications,182(50), 18-22.
2. Singh, R., Yadav, S., & Gupta, P. (2018). "Identification andClassification of Medicinal Plants using Machine LearningAlgorithms." In 2018 4th InternationalConference on Computing Sciences (ICCS) (pp. 107- 111). IEEE.
3. Rathore, A., & Mishra, D. (2019). "Machine Learning Techniques for Classification of Medicinal Plants." International Journal of Computer Sciences and Engineering, 7(3), 1-5.
4. Verma, S., & Prasad, N. (2020). "Medicinal Plant Identification using Convolutional Neural Networks." In2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)(pp. 1-5). IEEE.
5. Sharma, A., Jain, N., & Sharma, S. (2017). "Automated Recognition of Medicinal Plants using Image Processing Techniques." In 2017International Conference on Computing, Communication and

Automation (ICCCA) (pp. 344-348).IEEE.

6. Patel, S., & Shah, K. (2020). "Classification of Medicinal Plantsusing Machine Learning and Image Processing Techniques: A Review." In 2020 International Conference on Advances in Computing and DataSciences (ICACDS) (pp. 1-5). IEEE.

7. Rajput, S., & Kumar, A. (2018). "Deep Learning-based Identification of Medicinal Plants using Leaf Images." In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.

8. Sharma, R., Sharma, S., & Singh, A. (2019). "Automated Classification of Medicinal Plants using Machine Learning Algorithms." In 2019International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp.1-5). IEEE.