IDENTIFICATION OF MEDICINAL PLANTS USING CONVOLUTIONAL NEURAL NETWORKS

Smt. S. Rama devi, Assistant Professor, Department Of ECE, Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh, India
A. Surekh, N. Pavani, A. Meghana, N. Jasmitha, Priyanka Sahu, StudentS, Department Of ECE, Andhra University College of Engineering for Women, Visakhapatnam, Andhra Pradesh, India.

Abstract

Accurate identification of medicinal plants is crucial for safe and effective use in Ayurvedic medicine. However, traditional identification methods relying on visual and aromatic inspection by human experts can be inconsistent. To address this challenge, a novel approach utilizing image processing and machine learning has been developed. A database of scanned images of leaves and flowers from commonly used medicinal plants in Indian Ayurveda was created. Morphological operations were then applied to process images, followed by a convolutional neural network algorithm for pixel data recognition. Testing yielded identification rates up to 93%, demonstrating high accuracy. This promising technique has the potential to significantly enhance medicinal plant identification in Ayurvedic medicine.

Keywords: Plant recognition, Image processing, Feature extraction, convolutional neural network, Machine learning

Introduction

Automated species identification was presented 15 years ago as a challenging but very promising solution for the development of new research activities in biology. Traditional methods of identifying a plant rely on human experts examining the visual features, aroma of plant species and their size, shape and texture of the species but this approach sometimes unreliable which leads to several health effects to people. Therefore, we developed an image processing-based approach using machine learning algorithm that is convolutional neural network to identify medicinal plant and for what diseases it is cured and the regional names of medicinal pant used in different states. This method uses a digital image of medicinal plant which is the leaf of particular medicinal plant as input and train those images using CNN in order to extract the features and this feature matches the leaves present in the trained dataset and displays the medicinal plant name, diseases cured and regional names. This approach is easy, fast and highly accurate in order to predict the results which requires no expensive equipment other the camera and computer.

Objectives of the study

- To train a machine learning model to accurately classify plants as medicinal or non-medicinal based on their characteristics such as morphology, chemical composition, or other relevant features.
- To develop a user-friendly interface that allows users to easily input plant data and obtain a prediction of whether the plant is medicinal or not.

Review of Literature

Machine learning is a of computers to learn and make decisions without being programmed to do. So, this is done by training the dataset. Neural networks are the model of machine learning which are used to make decisions and used to solve real life problems and it has ability to perform computations quickly. These neural networks algorithms examine the general information and learns the different highlights needed to recognize hidden patterns within it. Over the past few years, the field made significant advances in deep learning based computational power to analyze different kinds of information, especially images and ordinary language. Automatic identification is mostly

benefited from image processing and object recognition. on such as identifying plant species by simply uploading an image itself. Images are captured and given to computer aided systems to detect automatically.

There are several researches done for identification of plant species one of the methods is the image processing techniques for the identification. Color, texture, and shape are the features of identifying flowers. For the methodology they have divided their approach into four steps as image enhancement, image segmentation features extraction and classification. They have used Chan-vese image processing segmentation technique for segmenting flowers from the rest of the image in order to simplify and enhance the features extraction process. For the extraction of features the have used HSV color descriptor, Gray Level Co-occurrence Matrix (GLCM) as texture descriptor and Invariant Moments (IM) as a shape descriptor. The images of flowers are cropped to get the relevant size. They have done cropping a single flower from bunch of flowers, cropping but keeping the full bunch and keeping the same structure (No cropping). For the classification part Back Propagation ANN was used. As the overall result they have included all flower species with 20 images and acquired 80% accuracy.

Research Methodology

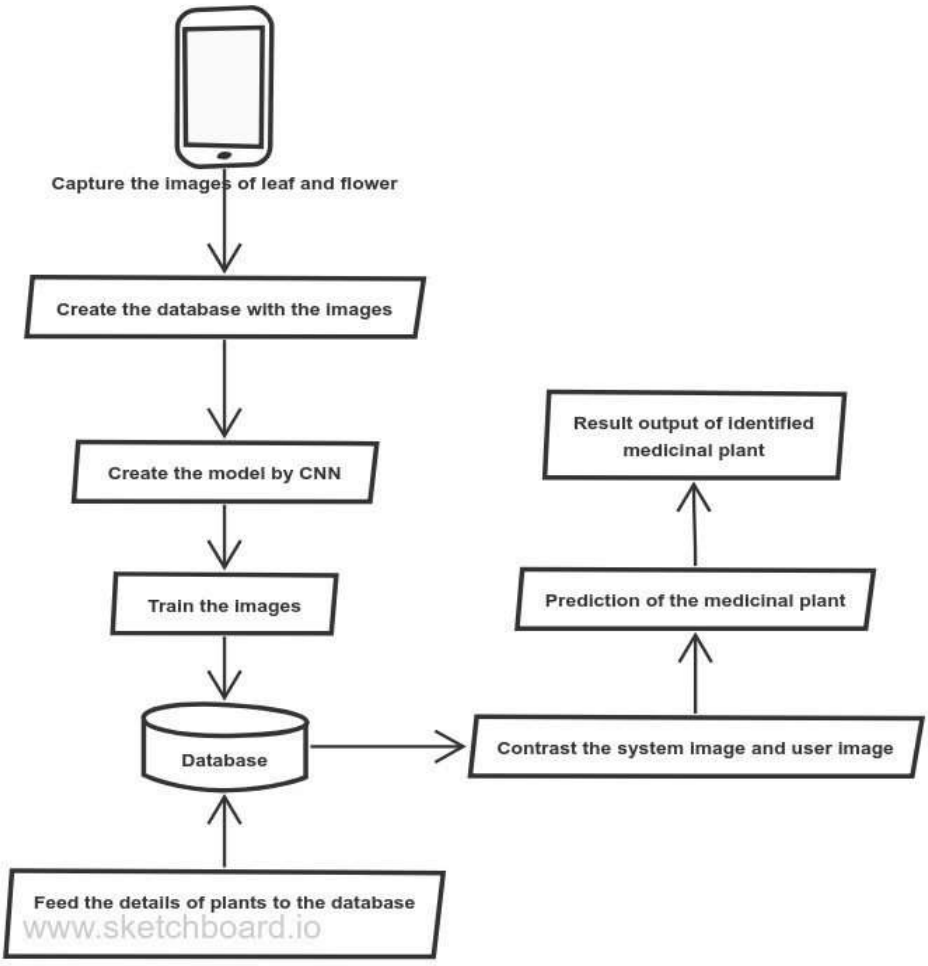

Figure 1: System model

- The objective is to create a window which is used to upload an image and start detects the medicinal plant name, diseases cured and the regional name of the medicinal plant. The first step is collecting large number of images over (5000) for different types of plant species. To train those plant species and get the features of these plants can be obtained by using the fine-tuning algorithm to produce more and more accuracy. Developing a user-friendly portal where we can upload an image and get results after processing.

- A dataset is a collection of data that is treated as a single unit by a computer . This means that a dataset contains a lot of separate pieces of data but can be used to train an algorithm with the goal of predicting the medicinal plant in the dagaset
- In this project dataset is choosen from https://www.kaggle.com/datasets/ because it is official for data related to India. Also, the data will give accurate and better result
- The complete database with leaves is extracted from kaggle dataset and the complete dataset is trained byusing convolutional neural network which is a machine learning algorithm.
- The trained model extract the features of the complete dataset and stored them in a database some details of the plants like regional names ,name of the plant and diseases cured using the predicted medicinal plant are also stored in the database
- When the input image is provided it matches the features which in included in the database and the medicinal plant is predicted along with diseases it can be cured and regional names.
-

Convolutional Neural Network (CNN)-Machine learning algorithm

CNNs are multilayered neural networks that identify visual patterns from pixel images through convolution a linear mathematical operation that creates a third function expressing changes in one functions shape caused by another. A CNN consists of convolution, pooling, and fully connected layers and uses a back propagation algorithm to learn data hierachis.

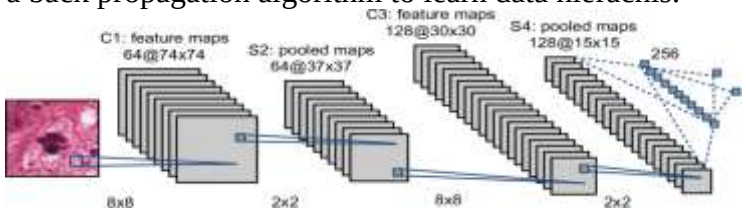

Figure 2: CNN Architecture

1. Convolutional layer:

Convolutional layer consists of three components as the input image, feature detector or mask and an output image. To get the greater accuracy mask is considered. The mask overlaps the input image and then by overlapping the input image, the maximum value of the pixel will be extracted by applying the mask continuously, then the output image is obtained. In this process some features of the input image will be neglected. But by extracting the maximum value of the pixel the identification is done with greater accuracy.

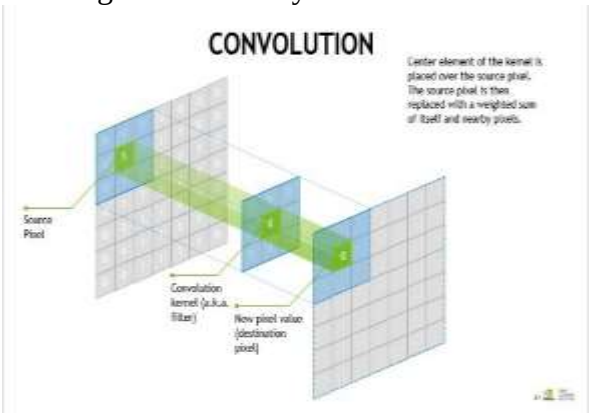

Figure 3: Convolutional layer

2. Pooling layer:

This pooling layer reduces the spatial representation which also results in loss of the information. Generally, there are two types of pooling functions which are norm rectangular neighborhood and L2 norm rectangular neighborhood where method can be chosen from the maximum output they provide.

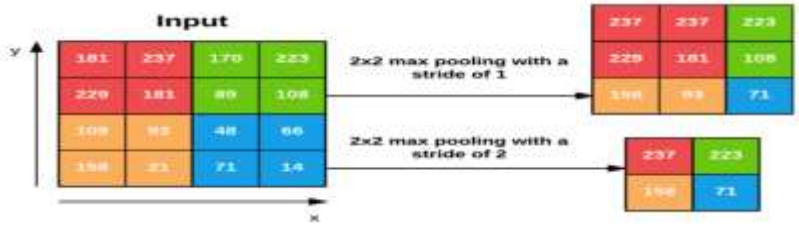

Figure 4: Pooling layer

3. Fully connected layer:

This fully connected layer connects the output of one layer as the input to the other layer which helps in increasing efficiency.

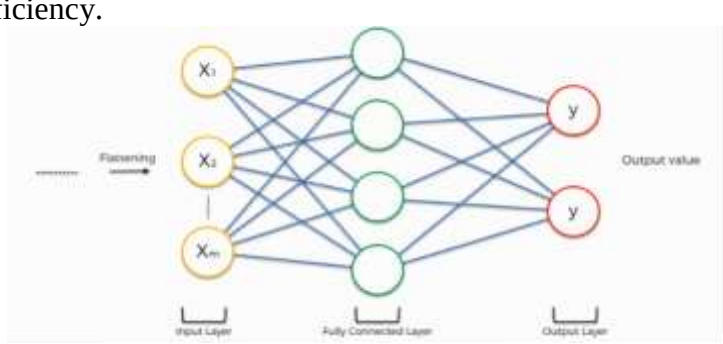

Figure 5: Fully connected layer

Results and Discussion

The following results are obtained as below:

a. Tunning parameters

Table 1. Tuning parameters

parameters	values
Activation	RELU
Optimizer	Gradient descent
Loss	Cross entropy
Epochs	10
Dropout	0.4
Batch Size	-1

b. Training Losses graph

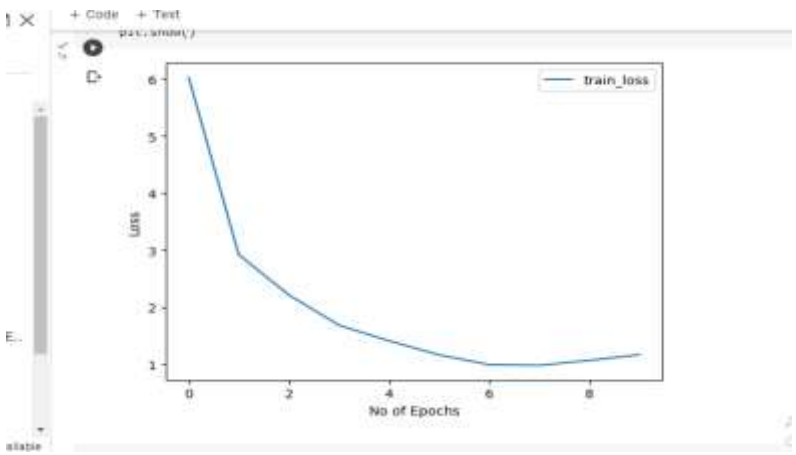

Figure 6: Training Losses graph

c. Convolutional neural network Accuracy

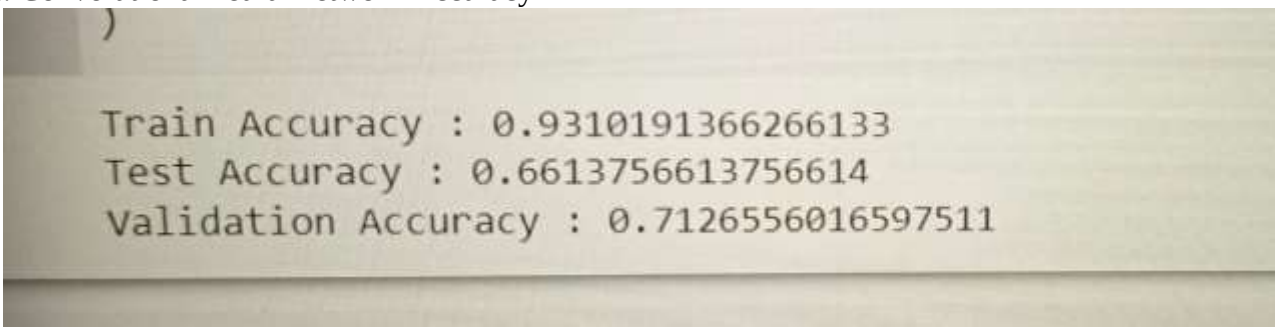

Figure 7: Accuracy

d. Result screen

Figure 8: Result screen

Conclusion

This study found the identification of medicinal plant is successfully done with the help of convolutional neural networks (CNN). Using CNN, it provided the result after training the data and also with the validation accuracy. Although this study can be done with various methodologies like visual characteristics, using CNN it provided greater accuracy than other methodologies. This approach for identification of medicinal plants provided the accuracy of 93% because of pretrained model provided to the database taken from dataset. Finally, it predicts the medicinal plant along with regional and and diseases cured for the predicted medicinal plant.The results which reduce the adverse effects of incorrect identification.

References

[1] Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H. and Tekinerdogan, B., 2019. Analysis of transfer learning for deep neural network-based plant classification models. Computers and Electronics in Agriculture, 158, pp.20-29. https://doi.org/10.1016/j.compag.2019.01.041

[2] Barbedo, J.G., 2018. Factors influencing the use of deep learning for plant disease recognition. Biosystems engineering, 172, pp.84-91. https://doi.org/10.1016/j.biosystemseng.2018.05.013

[3] Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, pp.311-318. https://doi.org/10.1016/j.compag.2018.01.009

[4] Lee, S.H., Chan, C.S., Mayo, S.J. and Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition, 71, pp.1-13. http://dx.doi.org/10.1016/j.patcog.2017.05.015

[5] Barbedo, J.G.A., 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture, 153, pp.46- 53. https://doi.org/10.1016/j.compag.2018.08.013
[6] Grinblat, G.L., Uzal, L.C., Larese, M.G. and Granitto, P.M., 2016. Deep learning for plant identification using vein morphological patterns. Computers and Electronics in Agriculture, 127, pp.418-424. http://dx.doi.org/10.1016/j.compag.2016.07.003
[7] Ghazi, M.M., Yanikoglu, B. and Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing, 235, pp.228-235. http://dx.doi.org/10.1016/j.neucom.2017.01.018
[8] Too, E.C., Yujian, L., Njuki, S. and Yingchun, L., 2018. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2018.03.032
[9] Lateef, F. and Ruichek, Y., 2019. Survey on Semantic Segmentation using Deep Learning Techniques. Neurocomputing. https://doi.org/10.1016/j.neucom.2019.02.003
[10] Zhu, X., Zhu, M. and Ren, H., 2018. Method of plant leaf recognition based on improved deep convolutional neural network. Cognitive Systems Research, 52, pp.223-233. https://doi.org/10.1016/j.cogsys.2018.06.008