Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, May : 2024

# CNN BASED MEDICINAL PLANT

**Dr.Eppakayala BalaKrishna**, Associate Professor CSE, Vaagdevi College of Engineering(Autonomous), India
**T.Kasturi**, UG Student,CSE, Vaagdevi College of Engineering (Autonomous),India
**S.Akhila**, UG Student, CSE, Vaagdevi College of Engineering(Autonomous), India **B.Nagajyothi**, UG Student, CSE, Vaagdevi College of Engineering(Autonomous), India
**D.Karthik**,UG Student,CSE,Vaagdevi College of Engineering (Autonomous), India

## ABSTRACT
Medicinal plants (herbs) are plants that are known to have certain compounds which are nutritious for health. The human body is complex and organic, while chemical medicines contain chemicals that are inorganic and pure. Therefore, chemical medicine are considered not very suitable for consumption by the human body, which if consumed continuously can even be bad for human health. However, some chemical drugs are actually symptomatic (temporary) so they must be taken for life by patients with certain diseases. Therefore ,a system is needed to be able to help the community to recognize medicinal plants better, in this case the medicinal plants are focused on the introduction of medicinal leaves. In this study identification of medicinal plant leaves was carried out using the Convolutional Neural Network method. This research will build a system of medicinal plant leaves by using Convolutional Neural Networks. Using training data that is carried out in a computer set and then implemented in mobile-based software to recognize the types and benefits of medicinal plant leaves identified.

## 1. INTRODUCTION
Medicinal plants (herbs) are plants that are known to have certain compounds that are nutritious for health. Each part of the medicinal plant is believed to have various properties to prevent, diverse or even cure a certain disease. In Indonesia there are 30,000 types of plants and 7000 of them are classified as medicinal plants (herbs). In addition to the many presence of these medicinal plants, the use of medicinal plants is also considered to be safer because it has natural ingredients compared to chemical medicine. To find out the content of medicinal plants can be done by phytochemical screening methods, so it can be known active compounds that are in certain plants that can be useful as a medicine.

The human body is complex and organic, while chemical medicines contain chemicals that are inorganic and pure. Therefore, chemical medicine is considered not suitable for consumption by the human body, which if consumed continuously can even be bad for human health. However, some chemical drugs are actually symptomatic (temporary) so they must be taken for life by patients with certain diseases. The number of medicinal plants is still not balanced with the public's knowledge about the medicinal plants themselves, so many people prefer chemical medicines because they are considered more practical and easy to obtain. Therefore,a system is needed to be able to help the community to recognize medicinal plants better, in this case the medicinal plants are focused on the introduction of medicinal leaves. Leaves can be identified from image images by colour, size, texture and shape using various methods including Neural Network.

Previous studies identified herbal medicinal plants based on leaf imagery using Artificial Neural Networks (ANN) Gray Level Co-occurrence Matrix and K- Nearest Neighbour Algorithms, Local Binary Patterns, Support Vector Machines, Multilayer Perceptron (MPL). The leaves to be identified are medicinal plant leaves which are often used for hypertension medicine, including bay leaves, avocado leaves, cat's whiskers leaves, celery leaves, soursop leaves, African leaves, starfruit leaves, grass jelly leaves, and betel leaves. These leaves were chosen because hypertension is a health problem that is considered serious and suffered by most people.

In this study identification of medicinal plant leaves was carried out using the Convolutional Neural Network method. CNN is one of the algorithms from the branch of Machine Learning that is based on Artificial Neural Networks (ANN) or its development, namely Deep Learning which is a development of the Multilayer Perceptron (MPL) to process
two-dimensional data, one of them is image. CNN is used in image data to detect and recognize objects in an image, with Backpropagation type training. The way CNN works is similar to MLP, but in CNN each neuron propagated on the network has a two-dimensional shape, so that the weight and linear operating parameters on CNN are different.

This research will build a system of identifying nine types of leaves of hypertensive medicinal plants using Convolutional Neural Networks. In order to facilitate the user in using the system, the system built will be implemented on a mobile-based software (Android) to recognize the types and benefits of medicinal plant leaves that are identified.

## 2. LITERATURE SURVEY

1. In this paper, we employ probabilistic neural network (PNN) with image and data processing techniques to implement a general purpose automated leaf recognition for plant classification. 12 leaf features are extracted and orthogonalized into 5 principal variables which consist the input vector of the PNN. The PNN is trained by 1800 leaves to classify 32 kinds of plants with an accuracy greater than 90%. Compared with other approaches, our algorithm is an accurate artificial intelligence approach which is fast in execution and easy in implementation.
2. Plant identification is an essential topic in computer vision with various applications such as agronomy, preservation, environmental impact, discovery of natural and pharmaceutical product. However, the standard and available dataset for medicinal plants have not been widely published for research community. This work contributes the first large, public and multi class dataset of medicinal plant images. Our dataset consists of total 20,000 images of 200 different labelled Vietnamese medicinal plant (VNPlant-200). We provide this dataset into two versions of size 256 × 256 and 512 × 512 pixels. The training set consists of 12,000 images and the remainder are used for testing set. We apply the Speed-Up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT) for extracting features and the Random Forest (FR) classifier is associated to recognize plant. The experimental results on the VNPlant-200 have been shown the interesting challenge task for pattern recognition.
3. In this work, we present a discriminative and effective local texture descriptor for bark image classification. The proposed descriptor is based on three factors, namely, pixel, magnitude and direction value. Unlike most other descriptors based on original local binary pattern, the proposed descriptor is conducted the changing of local texture of bark image. The performance of the proposed descriptor is evaluated on three benchmark datasets. The experimental results show that our approach is highly effective.
4. Tree identification is one of the areas that are regarded by researchers. It is done by human expert with high cost. Experts believe that tree bark has a high relation with species in comparison with other phenotype properties. Repeated textures in the bark is usually various with slight differences. So, lbp-like descriptors used in most recent works. But, most of them do not provide discriminative features. Also some texture descriptors are sensitive to noise and rotation. Local ternary pattern is one of the operators that are resistant to the noise with high discrimination. In most of descriptors,
histogram of patterns is used to extract features. But, it is rotation sensitive with high computational complexity. In this paper, the main contribution is to propose a method for bark texture classification with high accuracy based on the improved local ternary patterns (ILTP). In the proposed ILTP, the ternary patterns are coded into two binary patterns, and then each one is classified into two uniform/non- uniform groups. The extracted patterns are labeled according to the degree of uniformity.

Finally the occurrence probability of the labels is extracted as features. Also, a multilayer perceptron is designed with four theories in the number of hidden nodes. Experimental results on two benchmark datasets showed that our proposed approach provides higher classification accuracy than most well known methods. Noise-resistant and rotation invariant are other advantages of the presented method. The proposed bark texture classification, because of its high classification accuracy, can be applied in real applications and reduce the financial costs and human risks in the diagnosis of plant species.

5. Plant recognition is one of important research areas of pattern recognition. As plant leaves are extremely irregular, complex and diverse, many existing plant classification and recognition methods cannot meet the requirements of the automatic plant recognition system. A plant recognition approach is proposed by combining singular value decomposition (SVD) and sparse representation (SR) in this paper. The difference from the traditional plant classification methods is that, instead of establishing a classification model by extracting the classification features, the proposed method directly reduces the image dimensionality and recognizes the test samples based on the sparse coefficients, and uses the class-specific dictionary learning for sparse modeling to reduce the recognition time. The proposed method is verified on two plant leaf datasets and is compared with other four existing plant recognition methods The overall recognition accuracy of the proposed approach for the 6 kinds of plant leaves is over 96%, which is the best classification rate. The experimental results show the feasibility and effectiveness of the proposed method.

6. Plant species recognition using a single organ, such as flower and leaf, is not sufficiently reliable, because different species may have very similar flowers or leaves, while the same species may have rather different flowers or leaves. Combining leaves and flowers to recognize plant species can produce positive results. Based on multi- modal learning scheme, an automatic plant species recognition method is proposed by combining leaves and flowers of plant. In the method, a modified local discriminant canonical correlation analysis (MLDCCA) is designed by incorporating the idea of local discriminant embedding (LDE) into canonical correlation analysis (CCA). Firstly, two neighbour graphs are constructed based on the exploration of the manifold that the input data lie on. Then, two projection matrices for dimensionality reduction are obtained by making the within-class neighbour samples most correlated and between-class neighbour samples least correlated, and meanwhile keeping the correlation between leaves and flowers of the same species maximum. Finally, K- nearest neighbour classifier with geodesic distance is used to recognize the plant species. MLDCCA is a powerful supervised multi-modal dimensional reduction method which can extract the discriminant features from two plant organs, meanwhile preserve the discriminant information and the data structure well. Experimental results on a real leaf and flower image dataset validate the effectiveness of the proposed method.

7. The leaves of plants have rich information in recognition of plants. In general, agriculture experts accomplish information extraction from the leaves. Since the leaves contain useful features for recognising various types of plants, so these features can be extracted and applied by automatic image recognition algorithms to classify plant species. In this study, the authors investigate a novel approach for recognition of plant species using GIST texture features. Then, the principal and suitable features are selected by principal component analysis (PCA) algorithm. In the classification step, three different approaches such as Patternnet neural network, support vector machine, and K-nearest neighbour (KNN) algorithms were applied to the extracted features. For evaluation of the authors' approach, they applied their proposed algorithm on three famous datasets. In comparison to some widely used features, the results show that their approach outperforms the other methods in the case of the time and the accuracy. The best results were achieved by applying PCA algorithm to GIST feature vector and using the Cosine KNN classifier.

## 3. PROBLEM STATEMENT

Previous studies identified herbal medicinal plants based on leaf imagery using Artificial Neural Networks (ANN), Gray Level Co-occurrence Matrix and K-Nearest Neighbour Algorithms[3], Local Binary Patterns[13], Support Vector Machines, Multilayer Perceptron (MPL). The leaves to be identified are medicinal plant leaves which are often used for hypertension medicine, including bay

leaves, avocado leaves, cat's whiskers leaves, celery leaves, soursop leaves, African leaves, starfruit leaves, grass jelly leaves, and betel leaves. These leaves were chosen because hypertension is a health problem that is considered serious and suffered by most people.

### 3.1 LIMITATION OF PROBLEM STATEMENT

Using machine learning algorithm will not provide better identification of Medicinal plants. These wrong identification may becomes serious problem if they take any medical decision with the plant or leaves.

## 4. PROPOSED SYSTEM

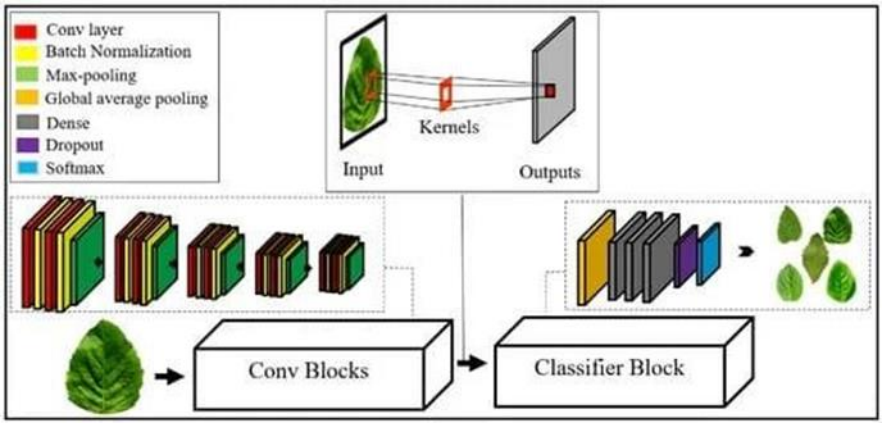
In this paper, we will identify medicinal plant/leaves was carried out using the Convolutional Neural Network method . CNN is one of the algorithms from the branch of Machine Learning that is based on Artificial Neural Networks (ANN) or its development, namely Deep Learning which is a development of the Multilayer Perceptron (MPL) to process two- dimensional data, one of them is image. CNN is used in image data to detect and recognize objects in an image, with trained models .

### 4.1 ADVANTAGES OF PROPOSED SYSTEM

We are using deep learning CNN algorithm to get the better accuracy . Using the CNN model we can accurately identify medicinal plant.

## 5. SYSTEM ARCHITECTURE

CNN is an architecture of DL algorithms containing a collection of non-linear transformation functions. These networks are an advanced version of artificial neural networks (ANNs) that can perform image processing operations such as object identification, segmentation, and classification. CNN models are made up of several layers producing an output from an input. Conventional, Pooling, and Fully Connected layers are among the most common algorithm of a CNN network implemented to process RGB images.




## 6. IMPLEMENTATION

### 6.1 UPLOAD DATASET

Using this module we can load medicinal plant dataset from the location of the project to train the CNN algorithm.

### 6.2 GENERATE TRAINING AND TESTING IMAGES

**Image Data Generator**: that rescales the image, applies shear in some range, zooms the image and does horizontal flipping with the image. This Image Data Generator includes all possible orientation of the image.

**train_datagen.flow_from_directory** is the function that is used to prepare data from the train_dataset directory Target_size specifies the target size of the image.

**test_datagen.flow_from_directory** is used to prepare test data for the model and all is similar as above.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, May : 2024

**fit_generator** is used to fit the data into the model made above, other factors used are steps_per_epochs tells us about the number of times the model will execute for the training data.
**epochs** tells us the number of times model will be trained in forward and backward pass.

### 6.3 GENERATE CNN MODEL

In this module we are generating CNN Model with train_datagen and test_datagen generated by Image Data Generator class.
Here we have training this CNN algorithm multiple time to get the better accuracy using epochs. Finally we will get the best CNN model with average accuracy above 90%

### 6.4 UPLOAD TEST IMAGE

Using this module we can upload test image AND pass the test image to the model to identify medicinal plant

### 6.5 IDENTIFY MEDICINAL PLANT

Using this model will call the CNN Model which is already generated and take the image from the $4^{th}$ step and pass to model. Then the model will identify the medicinal plant.

## 7. OUTPUT SCREENS


Fig 7.1 Home page

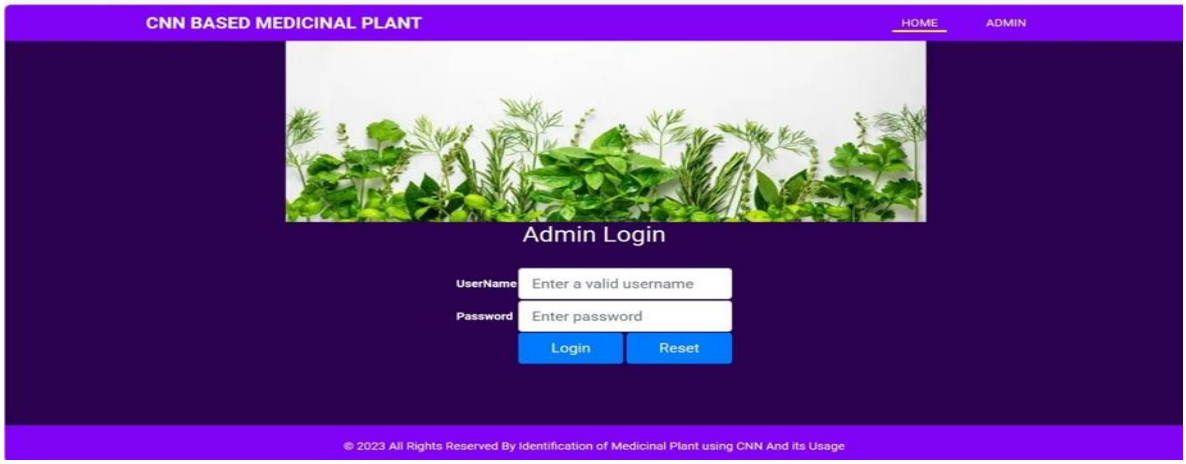



Fig 7.2 Admin login

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, May : 2024


Fig 7.3 upload dataset


Fig 7.4 Generating images as training and testing


Fig 7.5 CNN model generate successfully with training and testing image


Fig 7.6 Choose test image and upload

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, May : 2024

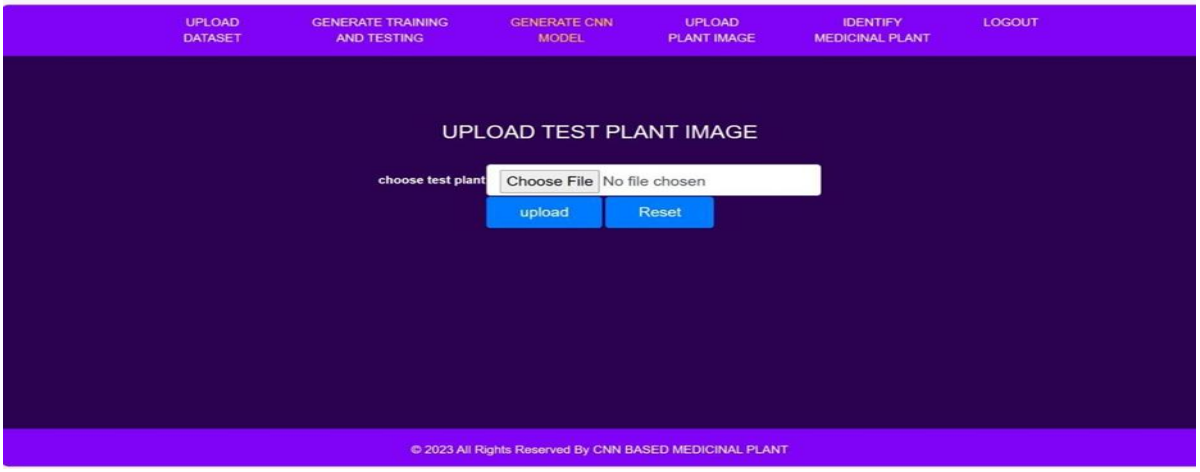



Fig 7.7 Uploading test plant

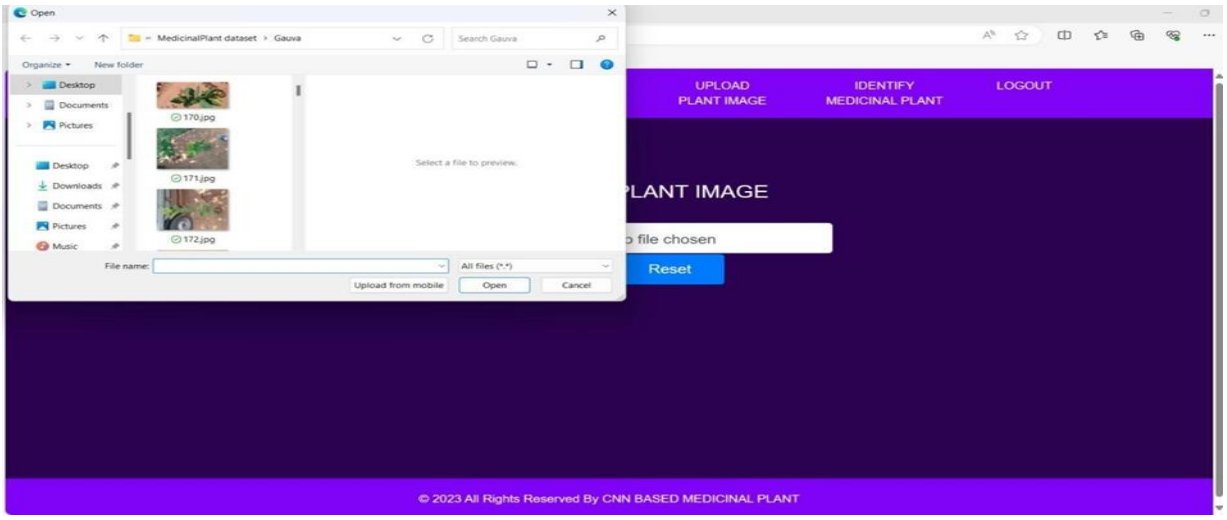



Fig 7.8 Uploaded test plant

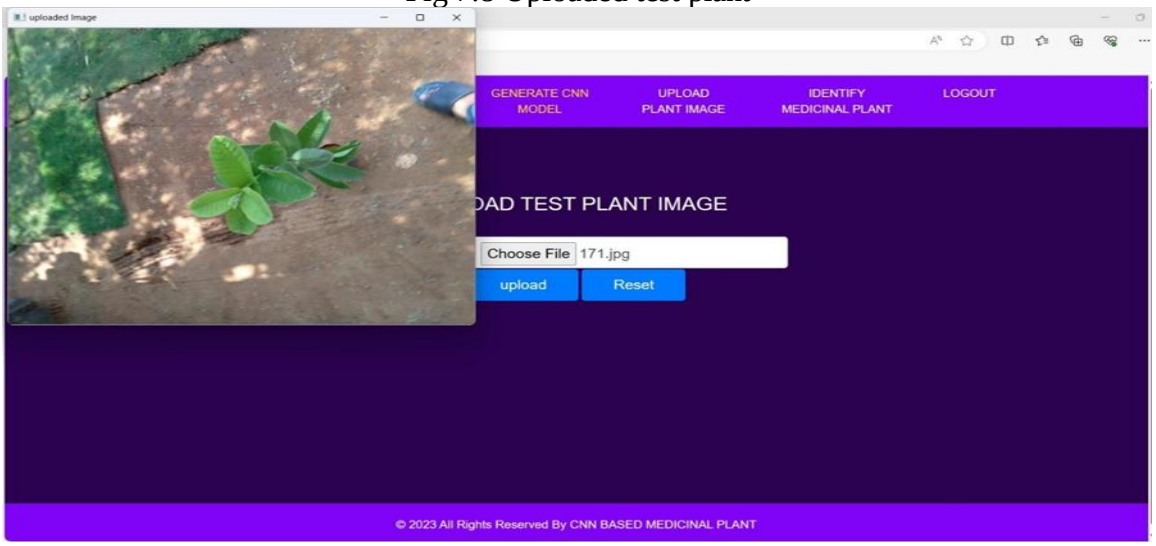



Fig 7.9 Upload status

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, May : 2024


Fig 7.10 Image identified

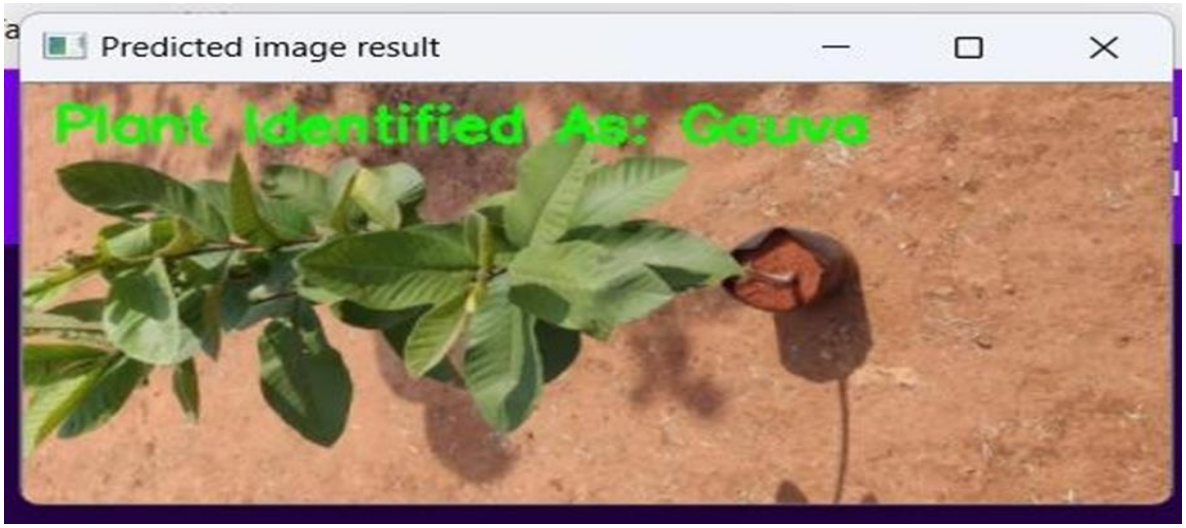



Fig 7.11 result


Fig 7.12 plant usage

## 8. CONCLUSION

This implementation of a CNN-based plant identification system within a Django framework represents a significant milestone in leveraging technology for biodiversity conservation and healthcare. By enabling users to upload images for plant identification and matching them against a trained dateset, the system provides a practical solution for identifying medicinal plants

accurately. This not only facilitates the exploration of traditional medicine but also aids in pharmaceutical research and conservation efforts. This project underscores the potential of machine learning and web development in addressing complex ecological and healthcare challenges.This research has successfully implemented the Conventional Neural Network method to extract features on medicinal plant and identify them into 40 classes of medicinal plant based on the closest value between the training data and test data.

## 9. FUTURE SCOPE

Moving forward, there are several avenues for expanding and enhancing the capabilities of the plant identification system. Firstly, continual refinement of the CNN model through additional training and the incorporation of more diverse datasets can improve accuracy and broaden the range of identifiable plant species. Secondly, the integration of user feedback mechanisms can enable iterative improvement of the system's performance over time. Additionally, exploring techniques such as transfer learning and ensemble models could further enhance the robustness of the identification process. Furthermore, extending the system to include mobile applications would increase accessibility, allowing users to identify plants on the go. Collaborations with botanical experts and institutions could enrich the dataset and provide valuable insights for refining the system's accuracy and relevance. Moreover, the inclusion of features such as geolocation integration and multi-language support would enhance the user experience and expand the system's usability on a global scale. Overall, the future scope of the plant identification system holds promise for advancing biodiversity conservation efforts and promoting sustainable healthcare practices worldwide.

## 10. BIBLIOGRAPHY

[1] Neeraj Kumar, Peter N. Belhumeur, Arijit Biswas, David W. Jacobs, W. John Kress, Ida C. Lopez, and Joao V. B. Soares. Leafsnap: A ˜ Computer Vision System for Automatic Plant Species Identification. In Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, Computer Vision – ECCV 2012, volume 7573, pages 502–516. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. Series Title: Lecture Notes in Computer Science.

[2] Alexis Joly, Herve Go ´ eau, Pierre Bonnet, Vera Baki ¨ c, Julien Barbe, ´ Souheil Selmi, Itheri Yahiaoui, Jennifer Carre, Elise Mouysset, Jean- ´ Franc¸ois Molino, Nozha Boujemaa, and Daniel Barthel´ emy. Interactive ´ plant identification based on social image data. Ecological Informatics, 23:22–34, September 2014.

[3] O. Soderkvist, Computer Vision Classification of Leaves from Swedish Trees, Dissertation, 2001.

[4] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. In 2007 IEEE International Symposium on Signal Processing and Information Technology, pages 11–16, Giza, Egypt, December 2007. IEEE.

[5] Trung Nguyen Quoc and Vinh Truong Hoang. Vnplant-200–a public and large-scale of vietnamese medicinal plant images dataset. In International Conference on Integrated Science, pages 406–411. Springer, 2020.

[6] Tuan Le-Viet and Vinh Truong Hoang. Local binary pattern based on image gradient for bark image classification. In Kezhi Mao and Xudong Jiang, editors, Tenth International Conference on Signal Processing Systems, page 39, Singapore, Singapore, April 2019. SPIE.

[7] Shervan Fekri-Ershad. Bark texture classification using improved local ternary patterns and multilayer neural network. Expert Systems with Applications, 158:113509, November 2020.

[8] Shanwen Zhang, Chuanlei Zhang, Zhen Wang, and Weiwei Kong. Combining sparse representation and singular value decomposition for plant recognition. Applied Soft Computing, 67:164–171, June 2018.

[9] Shanwen Zhang, Chuanlei Zhang, and Wenzhun Huang. Integrating leaf and flower by local

discriminant CCA for plant species recognition. Computers and Electronics in Agriculture, 155:150–156, December 2018.
[10] Fateme Mostajer Kheirkhah and Habibollah Asghari. Plant leaf classification using GIST texture features. IET Computer Vision, 13(4):369– 375, June 2019.