ISSN 2063-5346

MEDICINAL PLANT CLASSIFICATION USING A CONVOLUTIONAL NEURAL NETWORK

[1*]Nilesh Bhelkar
Research Scholar, Department of Computer Science and Engineering,
Madhyanchal Professional University,
Ratibad, Bhopal(MP),India
nileshbhelkar26@gmail.com

[2]Dr. Avinash Sharma
Professor, Department of Computer Science and Engineering,
Madhyanchal Professional University, Ratibad, Bhopal (MP),India

[3]Dr. Manoj Patil
Assistant Professor, Department of Computer Engineering,
DMCE, Mumbai,(MS), India

Article History: Received: 03.04.2023 **Revised:** 25.04.2023 **Accepted:** 15.05.2023

Abstract—
The classification of plant species used in herbal medicine is one of the most popular applications of machine learning, deep learning and image processing. There are countless species of plants, some of which have therapeutic benefits, in the world. Herbal, Ayurvedic, and Traditional Medicine all make heavy use of medicinal herbs in their production. Among the different methods for identifying plants and herbs is leaf identification, thus it is essential to accurately examine and classify plant leaf. The leaves are sorted into groups based on their shape and size. The automatic identification system for medicinal plants described in this paper will help people in the community learn more about these plants. This paper presents a CNN model to classify 45 different classes of medicinal leaf, and found accuracy of 93.53% after 50 number of epochs.

Keywords*— Classification, medicinal plant, CNN, accuracy*

I. INTRODUCTION

Medicinal plants have properties that make them beneficial for organically healing diseases. In the majority of the world, conventional medicine is still the norm. Herbal plants have leaves, stems, or roots that can be utilized to make modern or traditional medicines. The vast majority of the time, you will find these plants growing on hills or in forests. The leaves of different herbs can provide a wealth of information about the plants themselves. In the natural world, there are a great many plants that are used for medical and aromatic purposes, and these plants can be found in abundance. The growth pattern of medicinal plants is typically used as the foundation for classifying the plants. Trees, shrubs, herbs, annuals, biennials, tubers, rhizomes, and climbers are all possible forms for this plants.

We are becoming more disconnected from nature as a result of the increasing technological sophistication of our way of life. Despite the fact that we want to get away from nature, we just can't because we are a part of it. Herbal remedies have no known adverse effects, are non-toxic to humans or the environment, and are easily accessible in their original habitats because they are all-natural products. In traditional medicine, a wide variety of herbs are employed in the treatment of conditions that are directly associated with particular times of the year. In order to prevent the loss of human life, it is necessary to raise awareness about them.

For our continued existence, humans are greatly dependent on plants. A handful of them not only give us food and oxygen but also have medicinal properties that can treat numerous diseases. More than 8,000 indigenous plants to India have been shown to have promising clinical applications, including the treatment of neurological disorders, leukoderma leprosy, diarrhea, and many other conditions. Research on medicinal plants has been ongoing for a while because the medications made from them are more affordable and have no potential adverse effects. Several plants also have anti-colorectal cancer qualities that tend to be more effective than the conventional chemotherapy treatment [3]. Plant-based medications have also shown promise in treating conditions such as obesity [2] and the COVID-19 infection [4].

Several studies and works in this field have improved plant species identification. We used the CNN model, in order to achieve the best possible results in the computer vision-based processes of categorization and recognition. In total, this study employed 4795 images of leaves from 45 different categories. The primary objective of this study is to design a system for the classification of the leaves of medicinal plants utilising a CNN-based model that has been developed specifically for this purpose.

II. REVIEW OF RELEVENT WORK

This section provides a discussion of the numerous methods that can be used to identify and categorize plant species based on photographs of their leaves..

In this study [1], Author employed three separate optimizers named Adam, Nadam, and RMSProp. Each of these optimizers has a unique learning rate. When compared to Nadam and RMSProp, Xception Architecture with the Adam optimizer and a learning rate of 0.0001 achieves 99% accuracy with dataset of 11 tomato disease classes. This work [5] is able to achieve 99.66% accuracy on the test set and an average accuracy of 99.9% using pre-trained models like, ResNet50 InceptionV3, and MobileNetV2. The medicinal leaf dataset, which comprises of 30 classes.

This article [6] details a method for transfer learning. Out of six supervised learning algorithms, the average performance of DenseNet201 was noted 87%, with GoogLeNet performing at the bottom at 79%. Automatic recognition of medicinal plants is accomplished in this work [7] using machine learning and computer vision. We gathered the leaves of 24 medicinal plants. A total of 9 characteristics were retrieved from each leaf, and the best classification accuracy of 90.1% came from a random forest algorithm utilizing a 10-fold cross-validation procedure.

A classification strategy for medicinal plant leaves is proposed in this [8] study. With the help of the SVM classifier, the method classified the medicinal plant leaves based on their 10 shape and 5 texture features. We tested the classifier on 12 image sets of leaves from different medicinal plants, and it was able to correctly identify them 93.3% of the time on average.

This paper [9] examines feature vectors from a green leaf's front and back, as well as morphological traits, to discover the best combination of features for recognition. A blend of feature vectors from this work, expanded to include identification by dried leaves, has yielded identification rates exceeding 94%.

Twenty leaves from 40 different species were examined for different characteristics. A csv file for Excel is used in processing. The object's geometry, colour, and texture are examined by Weka's pre-processing and visualisation tools. Identification feature values are determined by this investigation [10]. Six colour characteristics were left out of the 12 on the front side. Leaf classification is best performed using mean-standard deviation HSV colour planes.

The proposed approach [11] classifies medicinal plants using textural properties that strongly affect leaf patterns. The suggested method involves categorization, feature extraction, and image improvement. Smartphone photos are processed to obtain leaf properties and compare them. Finally, a 60%-accurate automatic classifier was created using the K-nearest neighbor (KNN) classifier.

Researchers Yigit E and others [12] classified the leaves using a variety of artificial intelligence strategies, including artificial neural networks, KNN, SVM, and a few other artificial intelligence methods. The results of the experiment demonstrated that the SVM performed significantly better than the other classifiers that were evaluated.

Forty plant species are examined in this study [13], with Ayurvedic plant identification techniques being used; twenty image samples are collected for each species. This strategy makes use of the Support Vector Machine as well as the Multilayer Perceptron. Following the preprocessing step, the colour, morphology, and texture of the leaf become visible. When Multilayer Perceptron is utilised, the accuracy of the system's outputs is increased to 94.5 percent.

Pearline, et al [14] extracted features and trained machine learning classifiers using deep learning and conventional approaches. Folio, Swedish Leaves, and Flavia employ LBP (local binary pattern, Hu moments, Color channel statistics, Haralick texture feature linear discriminant analysis, (KNN) k-nearest neighbour classifier, bagging classifier, and Logistic Regression classifier. Bagging, Logistic Regression, Local Binary Pattern, Haralick texture feature, Hu moments Color channel statistics, and others are used. The investigation was 95% accurate.

This work makes use of the K-Nearest Neighbor classifier, as well as grey scaling, thresholding, and filtering. Munisami, et al. [15] When compared to the datasets from Folio, the system provides accuracy and testing with a precision of 83.5 percent far more quickly. Leaves are recognizable using the image processing methods [16–21]. The history that led to this assertion is detailed down below. To begin, the majority of plants used for medicinal purposes are found in dense forests and have leaves that are very similar. If someone picks the wrong plant, they run the risk of contracting a terrible sickness that could ultimately be fatal. There are a variety of techniques that can be used to identify plants. There is a high potential for making mistakes when manually identifying plants [16]. In order to circumvent this issue, researchers developed a method known as automatic system identification [17]. Various researchers are currently identifying plant illnesses, segmenting leaves, and assessing the quality of leaves. The colour and shape of the leaves of medicinal plants can, according to [18], be used to categorise the plants themselves. An accuracy of 96.66% was achieved by a Support Vector Machine classifier when applied to a dataset with optimised feature sets. A method of identifying plants that is based on their leaves was reported in [19]. Edge detection was performed on fifty medical photos sourced from Google Images. The texture patch that CNN generated was used for classification, and the results showed that it had an accuracy of 97.80 percent. The various fungal infections that might affect sugarcane leaf are categorised [20]. Using a triangle threshold, sugarcane leaves were successfully classified at a rate of 98.60% accuracy. [21] presented a CNN-based method for the classification of leaf images. 98.32% of the 12,673 soybean leaf samples can be classified using LeNet.

Using the leaf's characterised form, colour, and texture, this research proposes a classification scheme for the leaf. Cellular automata filter performs the initial processing [22]. The use of Region-of-Interest Segmentation and Histogram Equalization. This system is an integration of several different machine learning methods, including fuzzy vector machine, k-nearest neighbour (KNN and, support vector machine (SVM).

A description of an institution in Romania that recognises medicinal plants is provided in [23]. They examined colour and grayscale images (GL). LDA + PCA accuracy is 92.9%.
The authors propose a fuzzy logic-based segmentation method for plant disease control [24]. They used ML classifiers to combine colour texture and histogram data. Random forest classifier scores 98.4%. [25] suggested categorising plants by leaf surface images using Local Binary Patterns (LBP). ML removed salt and pepper noise and extracted textural information. They scored 93.5 percent.

The literature reviewed presents use of classical machine learning, ANN, CNN and other image processing based approaches to classify the leaf images and found useful for classification of leaf on the basis of various features. All the approaches provide fruitful outcomes to classify and identify plants on the basis of leaves. This paper presents a convolutional neural networks model approach to classify 45 different leafs.

III. METHODOLOGY

This work develops the CNN model and uses the Sequential () class to instantiate a sequential model. Once that's done, convolutional, pooling, and dense layers can be added with the help of the add()method and the Conv2D(), MaxPooling2D(), and Dense()classes, respectively. A method for increasing the variety of your training set by randomly applying modifications to the images, such as resizing, rescaling, rotating, flipping, and contrasting the images. These are the Keras preprocessing layers, and they are utilised for data augmentation. Create a new layer node by calling a layer on this inputs object. Create a Model by defining its inputs and outputs in the layer graph. Examine the model summary by plotting the model as a graph as shown in figure 1.

An SGD technique known as Adam optimization seeks to achieve optimal performance through the adaptive estimation of first- and second-order moments. Thus Adam optimizer is used in this model. To facilitate the training process, the data was first split into a training set and a test set. In order to train the model, we used 80% of the leaf images and 20% for testing. After training the model for a total of 50 epochs, both the accuracy and the loss of the model during training and validation were analysed and plotted.

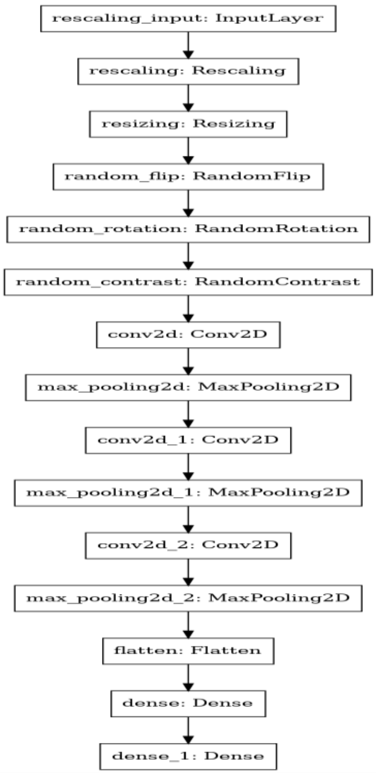

Figure 1Model Summery plotted in graphical form

Above figure represents the layered architecture of the CNN model for the classification of leaf images.

IV. Results and Discussion

This particular study only looked at the leaves of 45 different medicinal plants, however there are millions of different kinds of plants and herbs used for therapeutic purposes around the world. The dataset utilised in this work are open source platforms as Mendeley dataset [26,27] and kaggle platform. Few images were collected from internet and clubbed that images with existing dataset.

In this study, medicinal plant leaves classification model using a convolutional neural network is developed. The dataset contains 45 different types of medicinal plant leaf, the distribution of images in each class is represented as shown in figure 2.

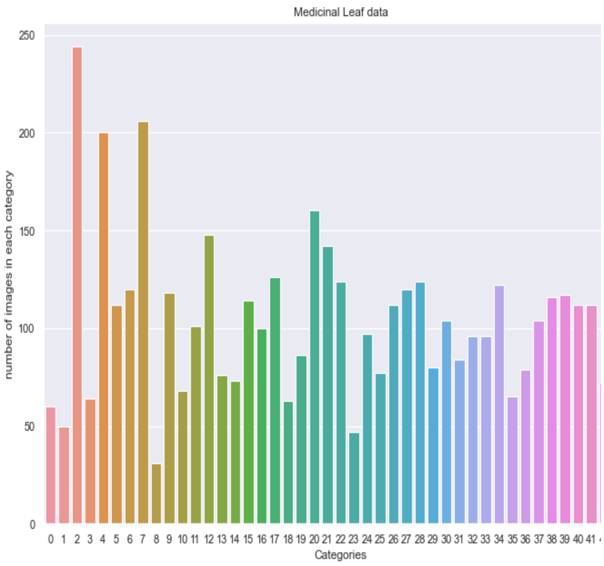

Figure 2 Image distribution plot for 45 classes of database

The plot of training and validation illustrated as shown in figure. 3, representing loss and accuracy of both training and validation after 50 epochs.

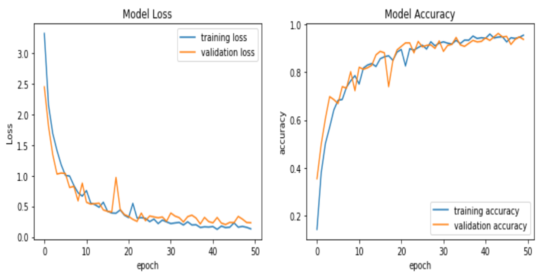

Figure 3 Validation and training Loss plot Vs epochs b) Validation and Accuracy plot

The methodology that has been proposed is significantly more efficient and effective than the one that has been detailed in the past [6,7,8,11,15,23,25]. In addition to this, it is consistent, satisfies requirements, and is an improvement over the current classification of medicinal plant leaves. Table 1 presents the results of a comparison examination of the suggested technique with the currently available works. The results obtained in this work are for 45 different classes and number of images in the dataset are also differ thus this comparison may differ when same dataset is utilized for classification purpose. Methodologies used are also differ along with the preprocessing before training, thus actual results obtained in this work are effectively sufficient for classification of the medicinal plant leaves. The training and validation loss obtained are 13.13% and 23.08% respectively. Training accuracy of 95.52% and resultant validation accuracy found to be 93.75%.

Table 1 Comparison of previous work

Ref. No	Classifier	Accuracy
[6]	DenseNet201, GoogleNet	87 % 79 %
[7]	Random forest	91.1 %
[8]	Support vector machine	93.3 %
[11]	KNN	60 %
[15]	KNN	83.5 %
[23]	LDA + PCA	92.9 %
[23]	LBP	93.5 %
This work	CNN	93.75 %

After successfully training the prediction results obtained for random image is depicted in figure 4 which represents accuracy of 99.96 % for Mexican mint leaf image.

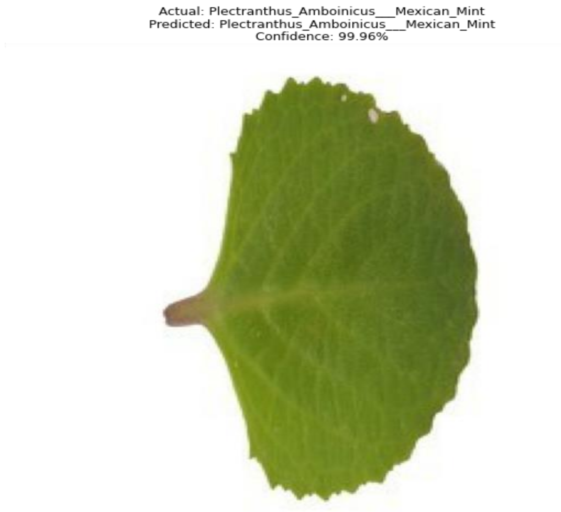

Figure 4 Random predicted Image with accuracy 99.96 %

V. CONCLUSION

Within the scope of this investigation, we build a convolutional neural network for the categorization of the leaves of medicinal plants. In this study, we use the leaves of medical plants to build a convolutional neural network for classifying them. A test is done on the size (256 x 256) of the dataset for classifying the leaves of medicinal plants. With 50 epochs, the accuracy was 93.75% and was well-organized. Automatic plant recognition is most difficult when leaf images are taken in complex backgrounds.

REFERENCES

[1] Patokar, A.M., Gohokar, V.V. (2023). Classification of Tomato Leaf Diseases: A Comparison of Different Optimizers. In: Kulkarni, A.J., Mirjalili, S., Udgata, S.K. (eds) Intelligent Systems and Applications. Lecture Notes in Electrical Engineering, vol 959. Springer, Singapore. https://doi.org/10.1007/978-981-19-6581-4_3.

[2] de Freitas Junior LM, de Almeida EB Jr.(2017) Medicinal plants for the treatment of obesity: ethnopharmacological approach and chemical and biological studies. Am J Transl Res. 2017 May 15;9(5):2050-2064. PMID: 28559960; PMCID: PMC5446492.

[3] Benarba B, Pandiella A (2018) Colorectal cancer and medicinal plants: principle findings from recent studies. Biomed Pharmacother 107:408–423. https:// doi. org/ 10. 1016/j. biopha. 2018. 08. 006

[4] Adhikari B, Marasini BP, Rayamajhee B et al (2021) Potential roles of medicinal plants for the treatment of viral diseases focusing on COVID-19: a review. Phyther Res 35:1298–1312. https:// doi. org/ 10. 1002/ ptr. 6893

[5] Sachar S, Kumar A (2022). Deep ensemble learning for automatic medicinal leaf identification. Int J Inf Technol. (2022);14(6):3089-3097. doi: 10.1007/s41870-022-01055-z. Epub 2022 Aug 12. PMID: 35975198; PMCID: PMC9373896.

[6] Oppong SO, Twum F, Hayfron-Acquah JB, Missah YM (2022). A Novel Computer Vision Model for Medicinal Plant Identification Using Log-Gabor Filters and Deep Learning Algorithms. Comput Intell Neurosci. 2022 Sep 27;2022:1189509. doi: 10.1155/2022/1189509. PMID: 36203732; PMCID: PMC9532088.

[7] Begue, Adams, et al.(2017) “Automatic Recognition of Medicinal Plants Using Machine Learning Techniques.” International Journal of Advanced Computer Science and Applications, vol. 8, no. 4, The Science and Information Organization, 2017. Crossref, https://doi.org/10.14569/ijacsa.2017.080424.

[8] Y. G. Naresh and H. S. Nagendraswamy,(2016) "Classification of medicinal plants: An approach using modified LBP with symbolic

representation", Neurocomputing, vol. 173, pp. 1789-1797.

[9] P. M. Kumar, C. M. Surya and V. P. Gopi,(2017)"Identification of ayurvedic medicinal plants by image processing of leaf samples," 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2017, pp. 231-238, doi: 10.1109/ICRCICN.2017.8234512.S

[10] Shweta Srivastava, (2014) "Weka: A Tool for Data preprocessing, Classification,Ensemble, Clustering and Association Rule Mining", International Journal of Computer Applications (0975 - 8887), vol. 88, no.10, Feb.2014.

[11] B. R. Pushpa, N. Megha and K. B. Amaljith,(2020) "Comparision and Classification of Medicinal Plant Leaf Based on Texture Feature," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154155.

[12] Yigit, E., Sabanci, K., Toktas, A., & Kayabasi, A. (2019). A study on visual features of leaves in plant identification using artificial intelligence techniques. Computers and electronics in agriculture, 156, 369-377.

[13] VijayaLakshmi, B., & Mohan, V. (2016). Kernel- based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features. Computers and Electronics in Agriculture, 125, 99-112.

[14] Pearline, S. A., Kumar, V. S., & Harini, S. (2019). A study on plant recognition using conventional image processing and deep learning approaches. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-8.

[15] Kumar, P. M., Surya, C. M., & Gopi, V. P. (2017, November). Identification of ayurvedic medicinal plants by image processing of leaf samples. In 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN) (pp. 231-238). IEEE.

[16] Zhang, F.; Zhang, X. Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation. Sensors 2011, 11, 2369–2384.

[17] Rahmatullah, M.; Azam, M.N.; Rahman, M.M.; Seraj, S.; Mahal, M.J.; Mou, S.M.; Nasrin, D.; Khatun, Z.; Islam, F.; Chowdhury, M.H. Chowdhury. A survey of medicinal plants used by Garo and non-Garo traditional medicinal practitioners in two villages of Tangail district, Bangladesh. Am. Eurasian J. Sustain. Agric. 2011, 5, 350–357.

[18] Dahigaonkar, T.D.; Kalyane, R. Identification of ayurvedic medicinal plants by image processing of leaf samples. Int. Res. J. Eng. Technol. (Irjet) 2018, 5, 351–355.

[19] Sabarinathan, C.; Hota, A.; Raj, A.; Dubey, V.K.; Ethirajulu, V. Medicinal plant leaf recognition and show medicinal uses using convolutional neural network. Int. J. Glob. Eng. 2018, 1, 120–127.

[20] Khirade, S.D.; Patil, A.B. Plant disease detection using image processing. In Proceedings of the 2015 International Conference on Computing Communication Control and Automation, Pune, India, 26–27 February 2015; pp. 768–771.

[21] Wallelign, S.; Polceanu, M.; Buche, C. Soybean plant disease identification using convolutional neural network. In Proceedings of the Thirty-First International Flairs Conference, Melbourne, FL, USA, 10 May 2018.

[22] Chaki, J., Parekh, R., Bhattacharya, S. (2015). Plant leaf recognition using texture and shape features with neural classifiers. Pattern Recognition Letters, 58, 61-68.

[23] Simion, I.M.; Casoni, D.; Sârbu, C. Classification of Romanian medicinal plant extracts according to the therapeutic effects using thin layer chromatography and robust chemometrics. J. Pharm. Biomed. Anal. 2019, 163, 137–143.

[24] Dhingra, G.; Kumar, V.; Joshi, H.D. A novel computer vision based neutrosophic approach for leaf disease identification and classification. Measurement 2019, 135, 782–794.

[25] Turkoglu, M.; Hanbay, D. Leaf-based plant species recognition based on improved local binary pattern and extreme learning machine. Phys. A Stat. Mech. Appl. 2019, 527, 121297.

[26] S, Roopashree; J, Anitha (2020), "Medicinal Leaf Dataset", Mendeley Data, V1, doi: 10.17632/nnytj2v3n5.1

[27] ikhwananda, Muhammad; Saputra, Hendra (2021), "Medicinal Leaf Dataseat", Mendeley Data, V1, doi: 10.17632/5tz79rtd39.1