How to Cite:
Bhelkar, N. S., & Sharma, A. (2022). Identification and classification of medicinal plants using leaf with deep convolutional neural networks. *International Journal of Health Sciences*, *6*(S6), 11596–11605. https://doi.org/10.53730/ijhs.v6nS6.13233

Identification and classification of medicinal plants using leaf with deep convolutional neural networks

Nilesh S. Bhelkar
Research Scholar, Madhyanchal Professional University, Ratibad, Bhopal - 462044, MP, India
*Corresponding author email: nileshbhelkar26@gmail.com

Dr. Avinash Sharma
Professor Dept of Computer Science and Engineering, Madhyanchal Professional University, Ratibad, Bhopal -.462044, MP, India
Email: avinashvtp@gmail.com

***Abstract*---**The Indian medical practise of Ayurveda has gained international renown. Herbal preparations are the basis of Ayurveda medicine. The pharmaceutical industry is beginning to pay more attention to medicinal plants because they have fewer adverse effects and reactions than modern medicine and are also less expensive. In recent years, numerous Deep learning, machine learning algorithms that are both effective and reliable have been utilised for plant classifications by using images of leaf. In this work, 45 distinct medicinal plant leaves were used, and a deep learning model was applied in order to achieve a high degree of accuracy in the classification and recognition procedures that were carried out with the help of computer vision techniques. After categorising the leaves of numerous medicinal plants, the Xception model has a 97.65% accuracy rate.

***Keywords*---**deep learning, classifier, medicinal leaf, herbal medicine, Xception.

Introduction

Ayurveda is sometimes referred to as the "mother of all medical sciences" since it is an ancient Indian system of medicine that makes use of medicinal herbs that are found naturally in the Indian subcontinent. According to historical accounts, the practice of Ayurveda dates back more than 5,000 years and was developed by wise people in ancient India. Early researchers believed herbs could cure a variety

Manuscript submitted: 9 May 2022, Manuscript revised: 18 July 2022, Accepted for publication: 27 August 2022

of illnesses. They experimented to determine the medicinal efficacy of different herbs. Thus-formulated drugs have few side effects. The manual identification of medicinal plants is a process that takes a significant amount of time and requires the assistance of trained professionals in order to be completed properly. Automatic ways of detecting and categorizing medicinal plants, which will ultimately be of greater benefit to society, are necessary to solve this problem. In the domain of image processing, one of the most active study areas at the moment is concerned with the automatic detection and classification of various medicinal/herbal plants. This required several processes, the most important of which are feature extraction and classification, both of which have an impact on the accuracy of the classification system as a whole.

The spread of medical plant agriculture can be aided by automatic identification and classification of medicinal plants, which is being done to better inform the public and provide reliable information to farmers. Suppliers, agents, pharmacy students, pharmaceutical companies, researchers, Ayurvedic medicine practitioners, botanists, and the cosmetics industry all benefit from the information and species database made available by the automatic identifying and classifying system. To distinguish between objects of similar and dissimilar types, recognition accuracy is the most crucial metric to consider. There should be a high degree of precision in the detecting procedure. Applications such as face recognition use this parameter to restrict access to authorised users, while medicinal plant recognition systems use it to locate the appropriate plant in time to save a patient's life.

In most cases, the task of gathering the plants from the forests is delegated to regular, everyday people. Due to the possibility of human error, they were not always successful in identifying the rare and significant plants. A patient's life may depend on the availability of these uncommon plant species in order to successfully treat their condition. Additionally, there is a possibility that these individuals could pick up the wrong species, which could result in a potentially harmful plant. A patient's life may depend on the availability of these uncommon plant species in order to successfully treat their condition. Additionally, there is a possibility that these individuals could pick up the wrong species, which could result in a potentially harmful plant. In these kinds of circumstances, it is essential to use a system that is capable of automatic plant recognition. Using this system, even a person with no prior knowledge of botany should be able to distinguish between the different plant species. These schemes are also helpful for mountaineers who want to collect plant species. Various studies already been conducted in this area to improve plant species identification. These methods still fall short of accurately classifying the plant species. Plants can be mostly identified by their leaves, flowers, bark, seeds, fruits, stems, roots, and other anatomical and physiological characteristics, such as their height, the region in which they grow, and the environmental factors they are exposed to. We used the Xception model, a deep convolutional neural network architecture, to achieve optimal results in the classification and recognition processes based on computer vision.

Review of previous work

This section presents review on various approaches to identify and classify plant species on their leaf images. In this work [1], the CNN architecture was employed to train the dataset that was collected and to develop a system that has a high level of accuracy. The success rate in locating the appropriate medicinal plant is 96.67% as a direct result of utilizing deep learning model. This study[2] uses Support Vector Machines, the Transfer Learning-VGG16 Model, and the You Only Look Once approach to categorise medicinal plants according to the characteristics of their leaves. In terms of accuracy, transfer learning achieved 98%, SVM achieved 97% after GridSearchCV hyper parameter tuning, and You Only Look Once achieved 84%. This paper [3] uses CNNs to distinguish Indian leaf species. Transfer Learning was used to select three pre-trained CNN architectures: InceptionV3, VGG16, and ResNet101. InceptionV3's validation accuracy and F1-score were 0.9732 and 0.9653. This study [4] combined CNN classification results dynamically using entropy impurity. VGG16, ResNet50, and Inception V3 are used. CNN's dynamic Resnet50 approach was 97.4% accurate.

This research [5] aims to develop a hybrid neural network system. "AousethNet" is a modified version of AlexNet. AousethNet's 98.61% accuracy. This paper [6] uses region-based and color-based thresholding. HOG and LBP were used to select features. Two-class and multi-class SVM give 99% accuracy. In this study [7], deep learning was utilised to classify leaf images. Five Medicinal plants. This study had 86% accuracy. This paper [8] describes how to identify medicinal plants using leaf features and preprocessing techniques. In this paper [9], authors focus on extracting image features and segmenting images of 125 different herb leaves from Malaysia, including Belalai Gajah, Rerama, Sirih, Mexican Mint, and Senduduk. For each image, a total of 14 features were extracted: 7 geometrical features and 7 textural features. The findings demonstrate that Sobel is capable of successfully segmenting the images and calculating the characteristics of the herb leaves.

In order to complete the task, ResNet50 was utilised in this work [10]. The architecture has been validated through the use of four different leaf datasets, one of which is a self-created dataset consisting of leaf images collected from the internet. The other three datasets have been taken from publicly available sources. While MK-D1 and MK-D2 have an accuracy of 99.05% and 99.89% respectively, the Flavia dataset has an impressive top-1 accuracy of 100%. This article [11] describes survey on various methods to identify medicinal plants by the shape and texture of their leaves. A computer vision method found better in plant leaf sample recognition. This study [12] classified the different types of leaves and stages of maturity using a CNN-based system. With tenfold cross-validation in a CNN-driven computer vision framework, an accuracy of 99% had been achieved in classifying leaf species and maturity stage.

In this study 13], a CNN-based system was used to identify the plant from photographs of its leaves, with an accuracy of 95.58 percent. To create this taxonomy, the author gathered pictures of leaves from six groups known by their local names: Tulshi, Sojne, Pathorkuchi Darchini, Tejpata, and Neem. The dataset used in this study [14] included data on 15 different medicinal plants, with 30 leaf samples per plant. The GLCM algorithm is utilised during the extraction process, whereas the K-NN algorithm is utilised during the classification stage. There were 15 different plant species used in this study, and the results showed that 74% accuracy. This research [15] suggests a machine learning-based system for categorising the leaves of medicinal plants. An enhanced medicinal plant leaves dataset is used to test out various machine learning models, including multi-layer perceptron, bagging, simple logistic, random forest, and logit-boost. The tulsi had a 99.10 percent accuracy rate in the multi-layer perceptron classifier, peppermint had a 99.80, bael had a 98.5 percent rate, lemon balm had a 99.5, catnip had a 98.5, and stevia had a 99.20 percent accuracy.

The purpose of this research [16] is to examine how various Laws' masks affect the labelling of images of medicinal leaves. The filter masks derived from Laws' masks of length 9 were found to have the highest classification accuracy (90.27 percent). Various techniques had been studied and suggested in this paper are those of image enhancement, feature extraction, and classification [17]. All extracted the features are compared. Finally, they found the K-nearest neighbour (KNN) classifier best to develop an automatic classifier. In this paper [18], researchers take a look back at the history of machine learning algorithms used to categorise plants based on images of their leaves and discuss which ones have proven to be the most effective and trustworthy. The techniques used in image processing to identify leaves and to extract key leaf features for use in various machine learning classifiers are discussed.

In this study [19], the pre-trained neural networks VGG16 and AlexNet were used to classify 11 distinct leaf diseases, and a comparison of the two models was also presented. The accuracy of classification provided by VGG16 was found to be superior. This work demonstrates [20] how standard classifiers, such as logistic regression[lR] and SVM may be combined with MPEG-7 colour and texture feature descriptors to produce very outstanding results across a wide variety of categories. In this research [21], a leaf classifier based on optimization approach (LCOT) is suggested for the purpose of identifying different species of the Hibiscus plant utilising a variety of leaf characteristics. The article [22] describes image processing algorithms used to determine leaf features. The performance of machine learning methods in detecting leaf photos based on typical plant features for classification.

In this study [23], they compared ML (SVM, SGD, RF) with DL (Inception-v3, VGG-19, VGG-16) for citrus plant disease identification. VGG-16 has the highest accuracy (89.5%) among DL techniques. In this article [24], the leaves are categorised according to their own distinctive combination of features. When tested across a broad spectrum of classifiers, identification rates of up to 99% have been recorded. In this research, [25] researchers take into account a Group Labelled Classification Model that analyses both the dorsal and ventral surfaces of a green leaf, as well as its morphological traits, to determine the best possible set of features for improving recognition accuracy.

Methodology

The retraining of a convolutional neural network to classify a new set of images can be accomplished with the help of transfer learning. One is able to take a network that has already been pretrained and use it as a foundation for learning a new task. It is typically much more time-consuming and challenging to train a network from scratch with randomly initialised weights than it is to fine-tune network using transfer learning, which can be done much more quickly and simply. With fewer training images needed, it is possible to quickly apply previously learned features to new tasks. The procedure for retraining a convolutional neural network to classify a new dataset of images using transfer learning is outlined in the following flowchart.

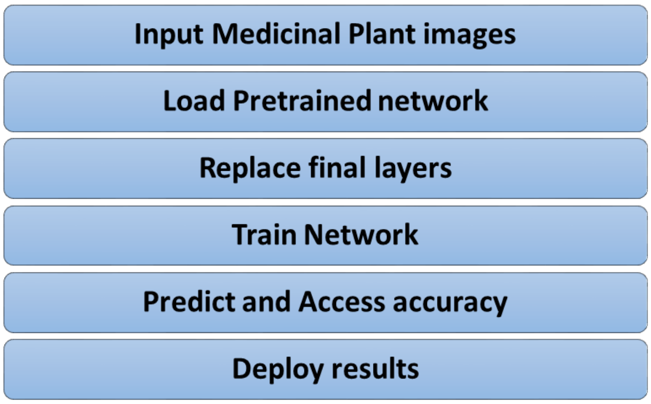

Figure 1. Flow of Implementation

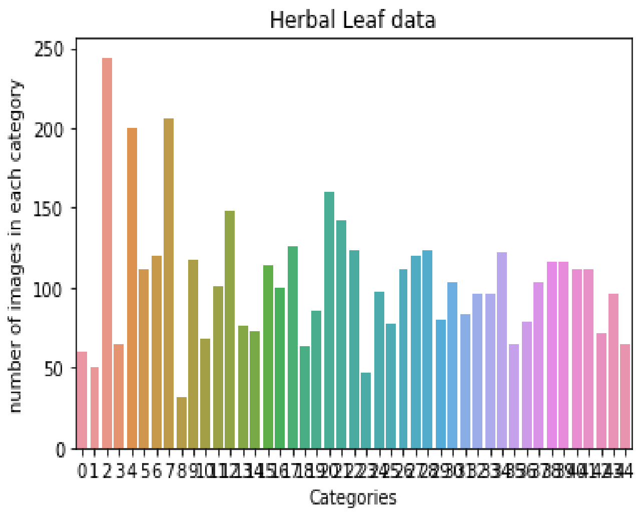

Figure 2. Count plot of the all 45 category of herbal leaf images

The process of retraining a CNN to classify on new dataset of images is illustrated in figure 1, which shows how transfer learning can be used. It is possible to quickly transfer previously learned features to a new task while using a reduced number of training images. The following is an elaborated list of the steps for using a pretrained model.

Input Medicinal Plant images

The new images are loaded into a database of images for classification. Separate the data into sets for training and validation. 70% of the images should be used for training, while 30% should be used for validation. Before proceeding with further operations, it is essential to visualise and then pre-process the images we will be working with, which are stored in a list. The figure 5 depict the visualisation of random 45 images from the dataset.

Load Pretrained network

In this stage of the process, a pre-trained version of the Xception network should be loaded. More than a million pictures drawn from the ImageNet database were used during the training process for this version. Using the Xception model to classify new images requires retraining the network.

Replace final layers

The network's convolutional layers are responsible for the extraction of image features, which are then used by the network's final learnable layer and its final classification layer to classify the input image. Most networks' last learnable weight layer is fully connected. Replace this fully connected layer with one that has as many outputs as the new data set's classes (45, in this study) Increase the layer's learning rate factors to learn faster than in transferred layers. The network's final output categories are defined in the classification layer. Change out the labelled classification layer with one that is label-free. When you train a network, the layer's output classes are determined for you automatically.

Train Network

Retraining the network with the updated data set is now possible. Weights in lower-level network layers can be "frozen" if desired. based on the specifications of the network, convert the size of the input image. Indicate any further augmentation procedures that should be applied to the training images.

Predict and Access accuracy

The training epoch length must be entered. You can determine how well the network performs at classification by applying it to the validation images Training for transfer learning does not require as many iterations. For the purposes of training, an epoch is a complete iteration over the entire dataset. Mini-batch size and validation information must be provided. Once per epoch, you should calculate the validation accuracy.

Deploy results

Show some sample validation images along with the labels that have been predicted for them and the probabilities that have been predicted for those labels.In total, 4682 images of the leaves of medicinal plants from 45 different species are collected for this study. Graphical visualization (bar graph) of the data taken for training in this study is presented as shown in the figure 2.

Results and Discussion

This section presents the outcomes obtained from retraining the Xception model for 45 classes of medicinal/herbal leaf classification. The plot of training and validation illustrated as shown in figure. 3, representing loss and accuracy of both training and validation after 50 epochs.

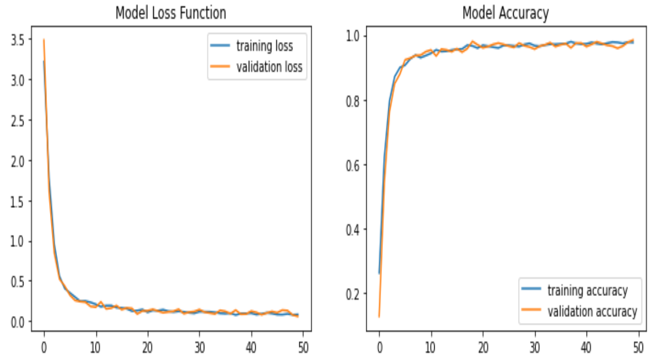

Figure 3. Training, Validation Loss and Accuracy plot

The accuracy of the training was found to be 98.40% after 50 epochs, and the accuracy of the validation was found to be 97.65%. Precision, recall, and the f1-score can all be manually calculated from the confusion matrix depicted in figure 4, which represents the classification model's performance. On newly collected test data, the trained model is evaluated for its ability to predict. The system was given a random set of images to process as an input, and it then plotted the outcomes, which are depicted in figure 6 as a prediction. The concluding portion of the prediction result presents the correct classification of the item.

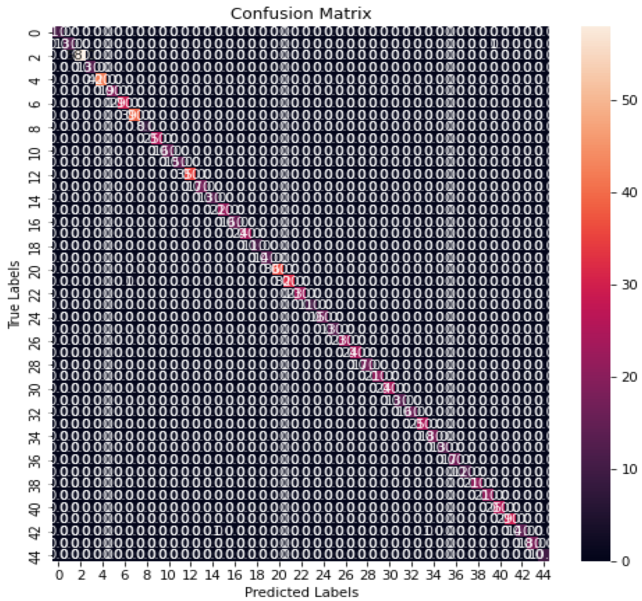

Figure 4. Confusion Matrix

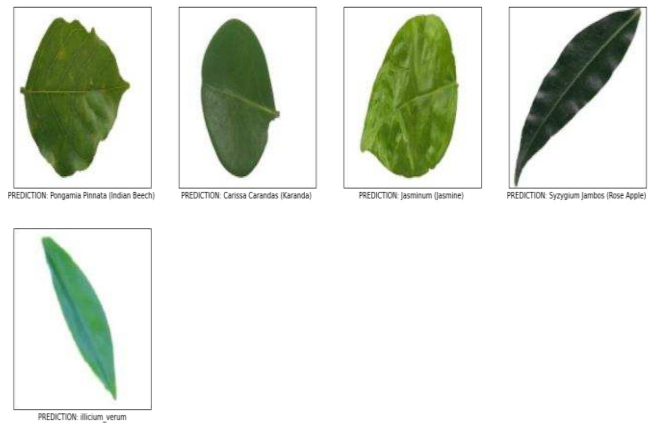

Figure 5. Prediction on Random Images

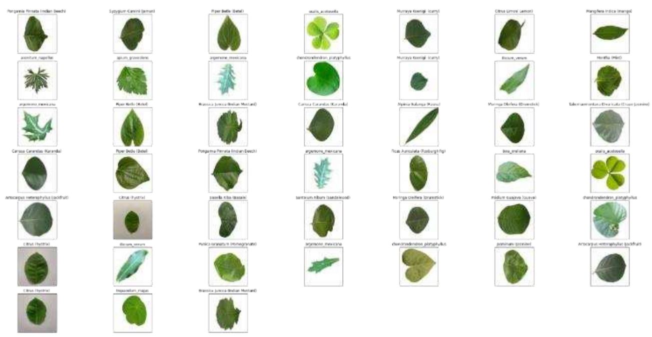

Figure 6. Random Visualization of dataset images

Conclusion

Within the scope of this study, 45 distinct plant species were analysed. To classify all 45 categories, the accuracy of the pre-trained model was determined to be satisfactory, and it achieved an accuracy of 97.65%. It is possible, as a future work, to design a customised deep learning model to implement the classification for medicinal plant recognition, and to compare the performance of various models.

References

1. A. Bahri, Y. Bourass, I. Badi, H. Zouaki, K. El moutaouakil, and K. Satori, "Dynamic CNN Combination for Morocco Aromatic and Medicinal Plant Classification," Int. J. Comput. Digit. Syst., vol. 11, no. 1, 2022, doi: 10.12785/ijcds/110120.
2. A. Gokhale, S. Babar, S. Gawade, and S. Jadhav, "Identification of medicinal plant using image processing and machine learning," in Advances in Intelligent Systems and Computing, 2020, vol. 1155, doi: 10.1007/978-981-15-4029-5_27.
3. A. Muzakir and U. Ependi, "Model for Identification and Prediction of Leaf Patterns: Preliminary Study for Improvement," Sci. J. Informatics, vol. 8, no. 2, 2021, doi: 10.15294/sji.v8i2.30024.
4. A. Rana and A. Mittal, "Herbal Plant Classification and Leaf Disease Identification Using MPEG-7 Feature Descriptor and Logistic Regression," in Advances in Intelligent Systems and Computing, 2020, vol. 1048, doi: 10.1007/978-981-15-0035-0_62.
5. B. R. Pushpa, N. Megha, and K. B. Amaljith, "Comparision and classification of medicinal plant leaf based on texture feature," in 2020 International Conference for Emerging Technology, INCET 2020, 2020, doi: 10.1109/INCET49848.2020.9154155.
6. C. P. Blesslin Elizabeth and S. Baulkani, "Novel Network for Medicinal Leaves Identification," IETE J. Res., 2022, doi: 10.1080/03772063.2021.2016504.
7. D. Joshi, V. Mishra, H. Srivastav, and D. Goel, "Progressive Transfer Learning Approach for Identifying the Leaf Type by Optimizing Network Parameters," Neural Process. Lett., vol. 53, no. 5, 2021, doi: 10.1007/s11063-021-10521-x.
8. D. Puri, A. Kumar, J. Virmani, and Kriti, "Classification of leaves of medicinal plants using laws' texture features," Int. J. Inf. Technol., vol. 14, no. 2, 2022, doi: 10.1007/s41870-019-00353-3.
9. G. Mukherjee, B. Tudu, and A. Chatterjee, "A convolutional neural network-driven computer vision system toward identification of species and maturity stage of medicinal leaves: case studies with Neem, Tulsi and Kalmegh leaves," Soft Comput., vol. 25, no. 22, 2021, doi: 10.1007/s00500-021-06139-9.
10. J. V. Anchitaalagammai, S. L. J. S. Revathy, S. Kavitha, and S. Murali, "Factors influencing the use of Deep Learning for Medicinal Plants Recognition," in Journal of Physics: Conference Series, 2021, vol. 2089, no. 1, doi: 10.1088/1742-6596/2089/1/012055.
11. K. Arunaggiri Pandian, T. S. Sai Kumar, S. Thabasum Aara, and A. Prabalakshmi, "Identification of Indian Medicinal Plants from Leaves using Transfer Learning Approach," in Proceedings of the 5th International

Conference on Trends in Electronics and Informatics, ICOEI 2021, 2021, doi: 10.1109/ICOEI51242.2021.9452917.

12. K. Pushpanathan, M. Hanafi, S. Mashohor, and W. F. Fazlil Ilahi, "Machine learning in medicinal plants recognition: a review," Artif. Intell. Rev., vol. 54, no. 1, 2021, doi: 10.1007/s10462-020-09847-0.
13. M. Diqi and S. H. Mulyani, "Implementation of CNN for Plant Leaf Classification," Int. J. Informatics Comput., vol. 2, no. 2, 2021, doi: 10.35842/ijicom.v2i2.28.
14. M. M. Singh, "A Survey on Different Methods for Medicinal Plants Identification and Classification System," Rev. Gestão Inovação e Tecnol., vol. 11, no. 4, 2021, doi: 10.47059/revistageintec.v11i4.2362.
15. M. Musa, M. S. Arman, M. E. Hossain, A. H. Thusar, N. K. Nisat, and A. Islam, "Classification of Immunity Booster Medicinal Plants Using CNN: A Deep Learning Approach," in Communications in Computer and Information Science, 2021, vol. 1440 CCIS, doi: 10.1007/978-3-030-81462-5_23.
16. P. Architha Shastry and A. P. Patil, "Classification of Medicinal Leaves Using Support Vector Machine, Convolutional Neural Network and You only Look Once," in Proceedings of B-HTC 2020 - 1st IEEE Bangalore Humanitarian Technology Conference, 2020, doi: 10.1109/B-HTC50970.2020.9297878.
17. P. Bose, S. Dutta, V. Goyal, and S. K. Bandyopadhyay, "Leaf Diseases Detection of Medicinal Plants Based on Support Vector Machine Classification Algorithm," J. Pharm. Res. Int., 2021, doi: 10.9734/jpri/2021/v33i42a32391.
18. P. K. Thella and U. Venugopal, "A group labelled classification model for accurate medical plant detection used in drug preparation," Rev. d'Intelligence Artif., vol. 35, no. 2, 2021, doi: 10.18280/ria.350208.
19. P. K. Thella and V. Ulagamuthalvi, "A Comparative Analysis on Machine Learning Models for Accurate Identification of Medical Plants," Rev. Gestão Inovação e Tecnol., vol. 11, no. 4, 2021, doi: 10.47059/revistageintec.v11i4.2309.
20. P. Kumar Thella and V. Ulagamuthalvi, "A Brief Review on Methods for Classification of Medical Plant Images," in Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021, 2021, doi: 10.1109/ICCMC51019.2021.9418380.
21. P. Shetty and R. Balasubramani, "Leaf classifier using optimization technique for hibiscus plant species identification," J. Adv. Res. Dyn. Control Syst., vol. 12, no. 3 Special Issue, 2020, doi: 10.5373/JARDCS/V12SP3/20201384.
22. Patokar Arun, and Vinaya Gohokar. "Plant Leaf Disease Identification and Classification Using Transfer Learning" Solid State Technology, 31 Mar. 2020, http://solidstatetechnology.us/index.php/JSST/article/view/2087
23. R. Sujatha, J. M. Chatterjee, N. Z. Jhanjhi, and S. N. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection," Microprocess. Microsyst., vol. 80, 2021, doi: 10.1016/j.micpro.2020.103615.
24. S. A. Halim, N. A. Hadi, and N. S. Mat Lazim, "Segmentation and features extraction of Malaysian herbs leaves," in Journal of Physics: Conference Series, 2021, vol. 1770, no. 1, doi: 10.1088/1742-6596/1770/1/012005.

S. Naeem et al., "The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach," Agronomy, vol. 11, no. 2, 2021, doi: 10.3390/agronomy11020263.