RESEARCH ARTICLE

SAMRIDDHI Volume 15, Issue 3, 2023 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Developing a Neural Network-based Model for Identifying Medicinal Plant Leaves using Image Recognition Techniques

Nidhi Tiwari*, Bineet K. Gupta

Department of Computer Science and Engineering, Shri Ramswaroop Memorial University, Lucknow, Uttar Pradesh, India

Abstract

Herbal plants play a vital role in human health and the environment, as they can provide both medical benefits and oxygen. Many herbal plants contain valuable therapeutic elements that can be passed down to future generations. Traditional methods of identifying plant species, such as manual measurement and examination of characteristics, are labor-intensive and time-consuming. There has been a push to develop more efficient methods using technology, such as digital image processing and pattern recognition techniques to address this. The proper identification of plant methods using computer vision and neural network techniques has been proposed. This approach involves neural network models such as CNN, ALexnet and ResNet for identifying the medical plants based on their respective features. Classification metrics give the 96.82 average accuracies. These results have been promising, and further research will involve using a larger dataset and going more into deep-learning neural networks to improve the accuracy of medicinal plant identification. It is hoped that a web or mobile-based system for automatic plant identification can help increase knowledge about medicinal plants, improve species identification techniques, and protect endangered species.

Keywords: Medicinal Plant, CNN, Alexnet, ResNet, ANN.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2023); DOI: 10.18090/samriddhi.v15i03.04

Introduction

There is a desire for an automated plant identification system that can assist users without specialized knowledge or training in botany and plant taxonomy to identify medicinal plants by taking pictures of the plants and inputting them into the system. Many traditional societies around the world have used the natural resources in their environment for medicinal purposes. Herbal remedies have been used since ancient times to maintain health and alleviate diseases.

From Vedic India, we can conclude that our Veda's describes a lot about medicinal plants. Rigveda describes the use of around 67 herbal plants, and Yajurveda and Atharveda mentions 81 and 289 herbal plants. The Sushruta and Charaka Samhita also holds a portion on the neuropathic use of around 500 medicinal plants. Nowadays, people from the globe are switching to traditional medicines because of their low cost and fewer side effects. Conventional medicines not necessarily based on plants but on some animal products also. In recent years, medico-ethnobotanical research has increased at the national and international levels. There is a significant lack of knowledge and scientific validation of ethnomedicine in India.

India is known for its abundance of medicinal plants, with many of them being collected for use in the production of drugs and perfumes. Roughly 8000 herbal treatments are marshelled in the Ayush system in India, i.e., Ayurveda, Unani, Siddha and folk medicines, out of which Unani and Ayurveda are the most widely practiced in India.

The WHO roughly calculates that around 80% of the global population depends on herbal medicines for their preliminary wellness requisites, under which 21000 herbal plant breeds have the capability to be used as Medicinal plants. In India, plants such as neem, tulsi, aloe,turmeric and ginger are widely used to treat numerous ailments and are frequently used as home remedies. India's warm, humid climate supports the growth of many species of flora, including 12,500 species[1] of seed plants that bear seeds, with more than 740 being indigenous. There are 2013

Corresponding Author: Nidhi Tiwari, Department of Computer Science and Engineering, Shri Ramswaroop Memorial University, Lucknow, Uttar Pradesh, India, e-mail: nidhitiw.7@gmail.com

How to cite this article: Tiwari, N., Gupta, B.K. (2023). Developing a Neural Network-based Model for Identifying Medicinal Plant Leaves using Image Recognition Techniques. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology,* 15(3), 301-310.

Source of support: Nil

Conflict of interest: None

recorded herbal plants in peninsular India, with an additional 68 species recently discovered in Gunung Ledang, Johor.[2,3] However, the study of herbal plants in India tends to focus on their utilization and treatment due to its mercantile value, is sluggish and seeks comprehensive understanding and cautious inspection of plant phenotypes. There are several methods for plant identification, including examining characteristics such as colors, flowers, leaves, and textures, but these can be complex due to a large number of plant species and similarities between plants at the brood level. It also seeks a botanist who can recognize the species that can be sluggish due to its morphological use of plant identification keys.[4, 5] An ecologist typically inspects one or more attributes of a plant such as shape of the leaf before focusing on distinct features to determine the species. This process often involves asking a series of questions before the species can be confirmed. Leaves are often used as the primary input in plant identification because every plant has them, and they can be easily identified based on distinct features such as shape, veins, and blades. The initial input in plant identification are the leaves. Nevertheless, there are examples where plants share similarities, particularly at the family level, making identification formidable even for experienced ecologists. This highlights the need for more well-organized plant identification methods, which can be used in plant identification, preservation planning and treatment of disorders.

Deep learning models, such as convolutional neural networks (CNNs),[6] can be helpful in plant identification because they extract hierarchical representations of input data and extract features such as patch size on different -different branches.[6-8] Unlike traditional methods, CNN significantly reduces the time that plant identification requires, and thus it can be enhanced to a greater extent through the use of graphical user interfaces that provide additional information about the identified plants. An attempt to model the human visual system as closely as possible is achieved by applying concepts, ideas, and techniques such as computer graphics, artificial intelligence, pattern matching, and digital image processing. Image recognition is a fundamental task in computer vision. While human vision is superior in detecting, identifying, and discriminating objects, computer vision aims to approximate this capability. When you identify a plant based on a specimen that has been previously collected, you are determining the identity of the plant. When you classify a plant, you categorize it based on characteristics that are similar to those of other plants.

Medicinal plants are necessary to balance human lives and nature as the major source of medicine. Hence, identifying medicinal plants is very important since they are a core part of life on Earth and allow human species and other organisms to survive by producing food and oxygen. A modern plant identification system can be brought to play to quickly characterize medicinal plant species without seeking specialist advice, streamlining the process. There are approximately thousands of plant species in the world, of which several plants have medicinal value, some of which are endangered, and others that are harmful to humans. Plants are essential human resources and the basis of all food chains. Hence, it is very important to classify and study plants correctly in ordrer to protect and use those plant species. Recognizing unspecified plants often depends on a specialist botanist's intrinsic knowledge and skills using traditional manual-based methods based on morphological characteristics. However, this manual recognition process can be tedious and sluggish. As a result, several researchers have conducted studies on the automatic grouping of plants based on their physical characteristics. These systems typically involve preparing the collected leaves, preprocessing to recognize certain attributes, classifying the leaves, populating a database, training for the identification and estimating the results. Since the leaves are most widely used for plant identification, other parts such as stems. Flowers, petals, seeds and even the whole plant can also be used in the automated process. Non-botanical experts can use a mechanized plant identification system to quickly identify the plant species.

Materials and Methods

When plants get infected with diseases, they commonly display visible signs such as marks or lesions on their leaves, stems, flowers, or fruits. These visible patterns are usually distinct and specific to the particular disease or pest condition, making them useful for accurate diagnosis. Extension officers are trained to diagnose plant diseases and pests based on visible symptoms. However, manual diagnosis can be subjective and time-consuming, and there may be a lack of trained personnel in certain areas. As a result, there is a need for efficient and accurate automated methods for the diagnosis of plant diseases and pests. Many researchers have attempted to address this issue using image processing, pattern recognition, and machine learning techniques. These techniques can be used to extract features from images of infected plants and classify the disease or pest based on these features.

Training extension officers can be costly and time-consuming.[9] Non-native diseases and pests may be difficult for farmers and extension officers to identify accurately.[10, 11] Distinguishing between anomalies with similar visual characteristics often requires a high level of expertise,[11-14] and even highly trained experts may make mistakes due to fatigue, poor lighting, or poor vision. In addition, individual experts often specialize in only a few types of disorders.[13, 15] Continuous monitoring is necessary for early disease detection and to prevent the spread of disease, but this can be tedious, time-consuming, costly, and inefficient to do on an ongoing basis.[4, 12-14, 16] Lab tests can be expensive due to the cost of lab equipment and are also destructive, as they involve collecting plant samples from the field and transporting them to the lab for analysis. There may also be quarantine restrictions on the transportation of samples to the lab.[16]

In a study by Begue and colleagues in 2017,[17] the characteristics of multiple leaves from a dataset of 24 plant species with 30 images each were analyzed. Various features such as the number of vertices, length, width, perimeter, area of the hull, and color were extracted. The research found that the best accuracy achieved was 90.1% by utilizing a random forest classifier. The plant species analyzed in the study were from the tropical island of Mauritius.

A study conducted by Dissanayake and Kumara[18] and their colleagues in 2021 compared the effectiveness of various machine learning algorithms in analyzing herbal, fruit, and vegetable plants based on their leaves. The research utilized 3,150 photos of leaves from 25 different plant species. Color photos were converted to grayscale and a Gaussian filter was applied to reduce image noise. The study focused on collecting 17 features that fell under three categories: shape, texture, and color. The study assessed the classification accuracy of several algorithms, including support vector machine, k-nearest neighbors, multilayer perceptron, random forest, and decision tree. The results showed accuracy rates of 85.82, 75.45, 82.88, 80.85, and 64.39%, respectively. Naeem *et al.* 2021[19] developed a machine learning-based medical plant leaf classification using multispectral and texture datasets. A total of six varieties of medicinal plant leaves were used. The study employed a chi-square feature selection strategy to select 14 features from a total of 65. Five different algorithms were tested to determine the most effective machine learning classifier for the task at hand, including multilayer perceptron, random forest, logit-boost, basic logistic, and bagging. Among these, the multilayer perceptron classifier demonstrated the highest accuracy at 99.01%.

Chung *et al.* 2021[20] created a dual-path CNN model with two independent subnetworks that receive input from either the original image or a centrally cropped image. The proposed model had an accuracy rate of 77.1% A 14-species tree dataset of Taiwan's most common trees was used to train and validate the model.

Xue and colleagues 2019[21] examined different methods for identifying 20 Chinese medicinal plant species by their leaves. Their research found that an Artificial Neural Network (ANN) model using morpho-colorimetric parameters as inputs performed better with 98.3% accuracy than a visible (VIS)/ Near Infrared (NIR) spectral analysis with an accuracy of 92.5%.

In a study by Kaur and Kaur in 2019,[22] Gaussian filtering was applied as a preprocessing technique on the Swedish Leaf dataset. The study then extracted texture and color features, which were classified using a multiclass support vector machine. The research found that this approach resulted in an accuracy of almost 93.26%.

Chouhan and colleagues in 2018 [4] proposed the use of Local Binary Patterns and Support Vector Machines (LBP-SVM) on the Swedish Leaf dataset and compared it to the K-Nearest Neighbors (K-NN) and Binarized Neural Network (BNN) classifiers. The LBP-SVM model demonstrated higher accuracy at 84%, while the existing BNN and KNN models achieved accuracies of 77 and 75%, respectively.

ANNs with backpropagation were proposed by Aakif *et al.*, 2015.[23] Their ANN classified their dataset with an accuracy of 96% based on a vector of morphological characteristics, Fourier descriptors (FD). Furthermore, after testing their efficiency, they achieved 96% accuracy for both Flavia and ICL datasets.

In Ahmed *et al.* 2016,[24] the authors developed an algorithm that extracts around 15 shape features and applies feature normalization and dimensionality reduction. Classification was performed using a Support Vector Machine (SVM), and an aggregate accuracy of 87.40% was achieved when tested on the Flavia dataset.

In 2017, Begue and colleagues[17] created a system that utilized their proprietary dataset of leaf images from 24 medicinal plant species. They obtained shape-based features from each image and applied a variety of classifiers including k-NN, naive Bayes, SVM, neural network, and random forest. The random forest classifier achieved the highest accuracy, reaching 90.1%.

Amlekar *et al.* 2018[9] developed a method that performs classification by automatically extracting shape features and using a feedforward backpropagation neural network. The method was tested on the ICL dataset and achieved accuracies of 99% for training images and 96% for testing images.

Shruthi *et al.* 2019[25] presented the stages of a general plant disease detection system and studied machine learning techniques for plant disease detection. They found that a convolutional neural network (CNN) was effective at detecting many plant diseases with high accuracy.

P. Srinivasan *et al.* 2019[26] developed software to classify and categorize groundnut leaf diseases. In their approach, groundnut crop diseases such as Early leaf spot, Late leaf spot, Rust, and Bud Necrosis were categorized into four distinct diseases using image acquisition, image preprocessing, segmentation, feature extraction, and the K-Nearest Neighbor (KNN) algorithm.

Numerous studies have employed shape feature descriptors for extracting features, as they are deemed the most distinguishing characteristic in plant identification. Prior to the ultimate classification of a pla nt image, it is vital to consider features that describe different facets of a plant leaf during the feature extraction stage in image processing. Furthermore, texture and color features can serve as more effective descriptors for a leaf image in scenarios where leaves are not fully grown or damaged. In our study, we have extracted the most optimal set of texture and color features for classification.

Dataset Collection

Object recognition research requires appropriate datasets at all stages, from training to evaluating the performance of recognition algorithms. For this research, images were collected from the Internet by searching for the names of specific medicinal plants in various languages, such as Latin, English, German, Serbian, and Hungarian. The images

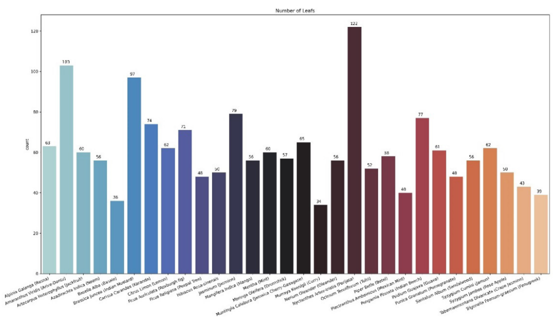


Figure 1: Plant Species for training and testing data

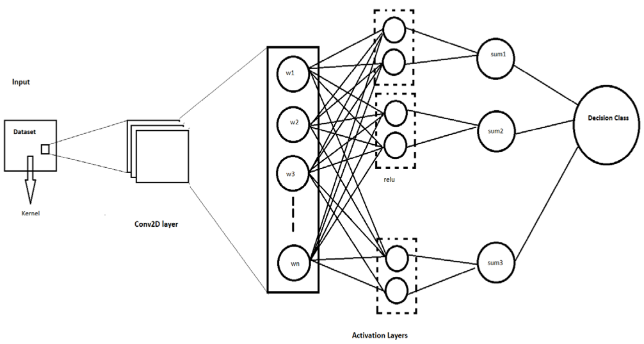


Figure 2: CNN Architecture

were grouped into 30 different classes, Many medicinal plants from various countries were included, such as Andrographispaniculata, Morindacitrophilia, Persicaria minor, Micromellumminutum, and Chromolaenaodorata, but few studies have been conducted on their automatic identification. One additional class in the dataset contained background images to improve the accuracy of classification by training the deep neural network to differentiate the leaves from the surrounding environment. The background images were taken from the Stanford background dataset.

Proposed Method

The input image is assigned to one of the labels of the classes (in this case, plant species) using a supervised technique. In particular, we have used convolutional neural networks (CNNs), which have proven effective in the field of computer vision and image analysis

Classification Methods

For classification, we have used CNN Architecture in this proposed method, following are types of Neural Network that we applied on the dataset.

- CNN 5 Layer
- CNN 7 Layer
- Alexet
- ResNet

Neural Networks

Neural Networks involves the analysis and visualization of images using computers.

Plant detection techniques can be based on two main neural network technologies: self-organizing maps and multilayer perceptrons (MLPs). A neural network is a system of hardware and/or software that is modeled after the operation of neurons in the human brain. It is also known as an artificial neural network. Self-organizing maps (SOMs) are a type of neural network that can be used for unsupervised learning, meaning that they can learn patterns in data without the need for labeled training examples. SOMs can be used to visualize high-dimensional data in a low-dimensional space, making them useful for data clustering and anomaly detection tasks. Multilayer perceptrons (MLPs) are a type of neural network that consists of multiple layers of artificial neurons, each of which receives input from the previous layer and passes it on to the next. MLPs are often used for supervised learning tasks like classification and regression. They are a widely used architecture for neural networks and have succeeded in various applications.

Deep Learning

The term "deep learning" refers to applying multiple layers of artificial neural networks (ANNs) to model high-level abstractions based on complex data structures. These layers provide different interpretations of the input data. Inspired by the way the human brain processes information, deep learning algorithms use large amounts of data to map input to specific labels. Examples of deep learning architecture include Neural networks are composed of convolutional layers, hidden belief layers, deep neural layers, and recurrent neural layers. There have been impressive results achieved using deep learning in fields such as digital audio, audio recognition, visual video recognition, computer vision, digital images, natural language processing(NLP), and automatic speech recognition, among others. In contrast to traditional machine learning algorithms, which typically involve preprocessing, feature extraction, feature selection, and classification, deep learning combines feature extraction and classification. This makes it a universal learning approach that is robust, generalized, and scalable. However, deep learning requires large amounts of data, has complex data models, and is expensive to train. It also requires a classifier to understand the learning results. Overfitting is a common problem in machine learning, particularly in deep learning, when training a model with a large number of parameters. This occurs when the trained model does not generalize well to unknown test data.Regularization prevents overfitting by allowing the model to adapt more effectively to unknown data when training on a limited training set or using imperfect optimization methods. The following strategies can be applied to regularization: batch normalization, weight decay, data augmentation, DropConnect, dropout, stochastic pooling, early stopping, as well as $\ell1$ and $\ell2$ regularization. Deep neural networks have been shown to perform particularly well in modern machine learning model training.

Convolution Neural Networks

It is a type of deep neural network that identifies and classifies specific features in images using convolutional neural networks (CNNs). They are used for recognizing images and videos, classifying them, analyzing medical images, and dealing with natural languages. As CNNs are highly accurate, they can be applied to recognize images in a wide range of fields, including medical imaging, smartphones, security systems, and recommendation systems. Convolution in CNN refers to the function of convolution, a linear operation in which two functions are multiplied to give a third function that represents how one function alters the shape of the other. A feature is extracted from an image by multiplying two images represented as matrices. There is no doubt that CNNs are capable of performing better than traditional methods when it comes to pattern classification and image processing. A CNN is a supervised feedforward artificial neural network, but there are also unsupervised and recurrent versions. The neural net is made of layers of convolutional, pooling, and fully connected (FC) layers, all of which have nonlinear activation functions, similar to the visual cortex of the eye. Many CNNs and deep learning algorithms developed as part of the ImagesNet Large Scale Visual Recognition Challenge (ILSVRC), such as DenseNet, Xception, ResNet, AlexNet and VGG have been applied to this challenge. These networks have achieved strong performance in recent years. Neural networks are composed of CNNs, which are one of their most important components. In order to detect objects and recognize faces, they use image recognition and classification. Some neurons in a CNN can be manipulated to learn their weights and biases. Neurons receive multiple inputs, calculate their weighted sum, and then activate them to produce output. Classifying images, clustering them based on similarity, and recognizing objects are some of the primary functions of CNNs. In addition to identifying people, street signs, and animals, CNNs can also be used in many other algorithms.

ResNet

The innovation of Residual Neural Network (ResNet) revolutionized the creation of Convolutional Neural Networks (CNNs) in 2015. It established a new belief that deeper layers learn new features from preceding layers by copying their connections to the input of the next layer without affecting the feature and identity extraction from the last layer. Despite having 152 layers, ResNet is less computationally complex than AlexNet and VGG, despite having 20 times more layers. The error rate of ResNet is lower than that of humans after training and implementing on the ImageNet dataset.[27]

Alexnet

AlexNet was the first deep convolutional neural network (CNN) to be widely used for image classification tasks. It was first introduced to the public at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, where it demonstrated the effectiveness of deep CNNs for image classification. AlexNet has five convolutional layers and three fully connected layers, and it introduced the Local Response Normalization (LRN) technique and the ReLU activation function. As a regularization method, it also used dropouts. AlexNet comprises five convolutional layers, one pooling layer, a ReLU activation function, and three fully connected layers, which gives it a more advanced architecture than the earlier LeNet. AlexNet requires RGB images that are 256X256 in size, so if they are not already 256X256, all training and test images must be resized.[27]

Results

This study is performed on 30 distinct species of medicinal plants, with a diverse range of leaf textures, shapes, and sizes. Performance metrics allow us to determine which neural network model is most effective on this dataset

Figure 3 displays all the leaves of the plants included in the dataset.

In medicinal plant leaf detection, a confusion matrix can be used to examine the capabilities of a classifier in identifying leaves in images. The matrix shows the number of times the classifier correctly predicted the presence of a leaf (True Positives, TP), the number of times the classifier incorrectly predicted the presence of a leaf when it was not present (False Positives, FP), the number of times the classifier failed to predict the presence of a leaf when it was present (False Negatives, FN), and the number of times the classifier correctly predicted the absence of a leaf (True Negatives, TN).

The F1 score is the performance metric that merges precision and recall to measure the classifier's performance. In other words, precision measures the percentage of positive predictions that are actually correct by dividing True Positives by False Positives. The recall measures the proportion of

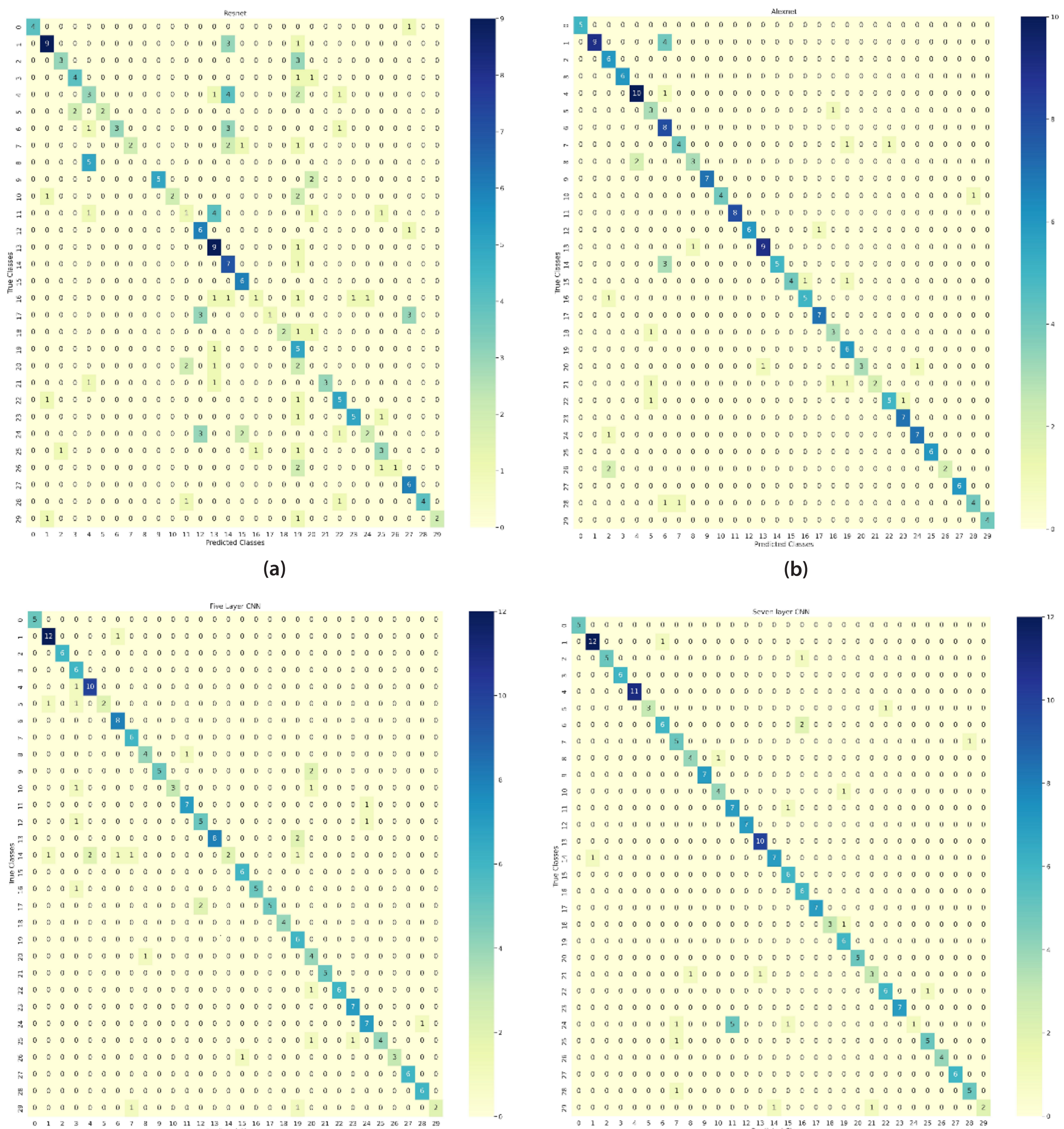
Figure 3: Same sample from the collected medicinal plant dataset.

true positives that were identified correctly by the classifier by dividing the number of True Positives by the sum of True Positives and False Negatives.

For example, if the classifier correctly identified 200 leaves out of 250 in the test dataset and also correctly identified 1000 non-leaf instances out of 1200, the confusion matrix would look like:

Actual		
Predicted	Leaf	Non-Leaf
Leaf	200	50
Non-Leaf	0	1000

The precision would be 200 / (200 + 50) = 0.8, the recall would be 200 / (200 + 0) = 1.0, and the F1 score would be 2 * (0.8 * 1.0) / (0.8 + 1.0) = 0.89. This F1 score indicates that the



(c) (d)

Figure 4: Confusion matrix for Neural Network models (a) Seven layer CNN (b) Five layer CNN (c) AlexNet (d) ResNet

classifier has good stability across precision and recall, with a high recall (100% of leaves were detected) and a relatively high precision (80% of positive predictions were correct).

In the context of leaf detection, loss function, validation loss, validation accuracy, and accuracy are important metrics used to analyze the machine learning model's performance.

A loss function is used to measure the difference between the predicted and actual outputs of the model. In the case of leaf detection, the loss function could be the mean squared error (MSE) between the predicted and actual bounding boxes for the leaves. The goal is to minimize the loss, indicating that the model is improving its performance in detecting the leaves.

Validation loss is calculated on a validation set, which is a subset of the training data. It measures the average loss over the validation set and helps to analyze if the model is overfitting or underfitting the training data.

Validation accuracy is the percentage of actual predictions made by the model on the validation set. It helps to analyze the model's accuracy on unseen data and prevent overfitting.

Accuracy is the percentage of actual predictions made by the model on a test set, which is separate from the training and validation sets. It is a commonly used metric to evaluate the model's overall performance in detecting the leaves.

By monitoring these metrics, one can understand the strengths and weaknesses of the model and make necessary improvements to achieve better performance in leaf detection.

The performance of all the models is shown in Table 1 with respect to Accuracy, Loss Function Validation Loss and Validation Accuracy.

The performance of all the models is shown in Table 2 with respect to F1 score, Precision and Recall.

In the context of medicinal plant leaf detection, F1 score with respect to macro, micro, and weighted is a generally used metric to analyze the performance of these Neural Network model.

The macro average F1 score determines the mean F1 score for each class in the leaf detection problem, treating each class equally regardless of size. This approach gives equal weight to each class in the calculation. This can be useful when the number of samples in each class is small or balanced, and you want to balance the F1 score between all classes.

The micro average F1 score aggregates the individual true positive, false positive, and false negative counts across all classes to compute the overall F1 score. This approach gives more weight to the larger classes. This can be useful when you have a class imbalance problem and you want to give more weight to the larger classes.

The weighted average F1 score is similar to the macro average but assigns different weights to each class based on its size. The weight of each class is proportional to the number of samples in the class. This can be useful when you have a class imbalance problem and you want to give more weight to the classes with more samples.

Loss function is used to calculate the variance between the predicted and actual outputs of the model. In the case of medicinal plant leaf detection, the loss function could be the mean squared error (MSE) between the predicted and actual

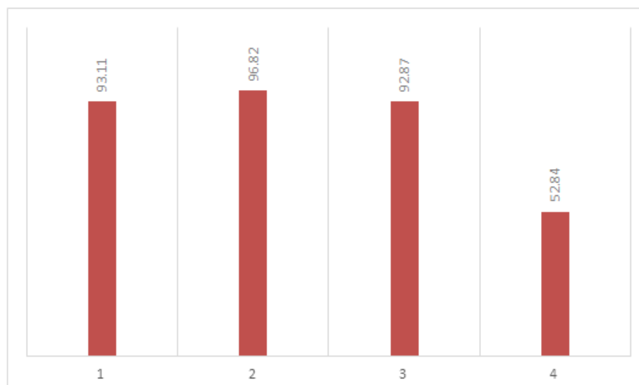


Figure 5: Performance accuracy bar chart.

Table 1: Comparison of different classification methods using CNN architecture

Method	*Accuracy*	*Loss Function*	*Val. Loss*	*Val Accuracy*	*Image Size*	*Epochos*	*Batch Size*
CNN 7 layer	93.11	0.187	1.401	77.67	256x256	100	32
CNN 5 layer	96.82	0.101	1.409	82.7	256x256	100	32
AlexNet	92.87	0.245	1.661	63.82	256x256	100	32
ResNet	52.84	1.48	1.98	49.69	256x256	100	32

Table 2: Comparison of different classification methods using CNN architecture on various performance metrics

	F1 score			*Precision*		*Recall*	
Method	*Macro*	*Micro*	*Weighted*	*Macro*	*Weighted*	*Macro*	*Weighted*
CNN 7 layer	85.93	87.24	86.21	89	89	86	87
CNN 5 layer	84.4	85.2	84.2	89	88	84	84
Alexnet	79.6	79.59	79.09	87	86	80	80
ResNet	53.5	54.08	53.33	68	66	53	54

Figure 6: Performance Loss Function bar chart.

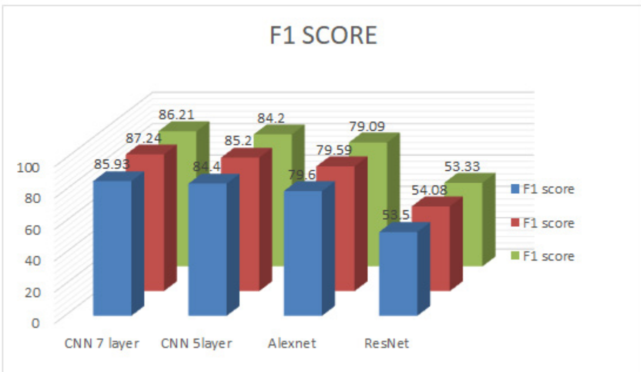


Figure 7: Performance F1 Score bar chart.

bounding boxes for the leaves. The goal is to minimize the loss, indicating that the model is improving its performance in detecting the leaves.

By monitoring these metrics, one can understand that CNN 5 layer is more reliable than others (Dias B.L., 2023).

Analysis

The obtained results of training accuracy and testing accuracy values are shown in the graphical form in figures. In the

CNN 7 layer

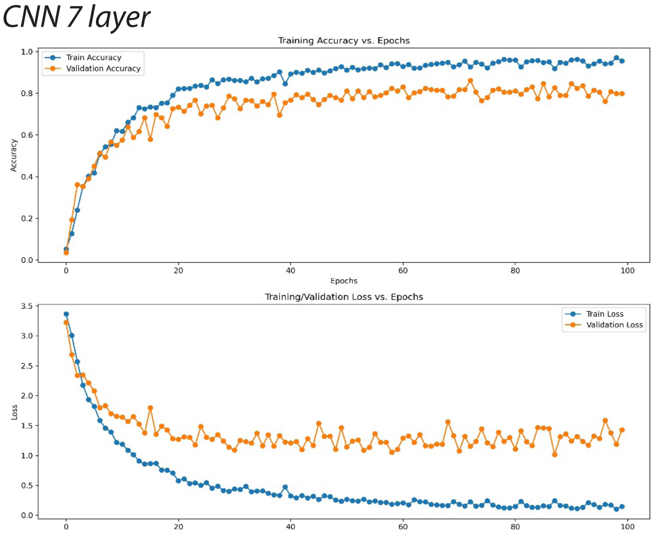


Figure 8: CNN 7-layer model training/validation accuracy vs epochs and training/validation loss vs epochs

CNN 5 layer

Figure 9: CNN 5-layer model training/validation accuracy vs epochs and training/validation loss vs epochs

Alexnet

Figure 10: Alexnet model training/validation accuracy vs epochs and training/validation loss vs epochs

ResNet

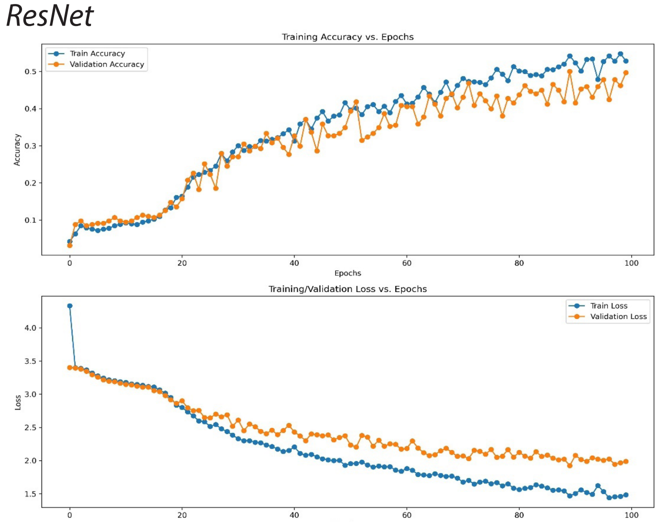


Figure 11: ResNet model training/validation accuracy vs epochs and training/validation loss vs epochs

graphs of Accuracy and Epochs we can also observe that the testing accuracy and training accuracy is increasing as the number of epochs increases. In case of more epochs, these two accuracy values are approximately comparable.

In the graphs of Loss Function and Epochs we can observe that loss function if gradually decreases as the number of epochs increases.

Discussion

An analysis of medicinal plant data was used to train and test the proposed deep CNN model. There are 30 classes of healthy medicinal plant in the dataset, divided into training and testing sets. Various models were compared with the proposed models, including CNN 5 layer, CNN 7 layer, Alexnet and ResNet. Moreover, comparisons are made based on accuracy and loss function. In this research paper each model runs for 100 epochs and batch size of 32 and a standard image size of 256x256. In conclusion, CNN 5 layer outperforms all the models.

Conclusions

This automatic classification system can help farmers and the general public to increase the production of Ayurveda provisions. It can identify medicinal plants without human assistance in various fields such as botany, taxonomy, Ayurveda manufacturing, and Ayurveda practice. In this paper, we used CNN and machine learning methods for the classification of medicinal plant species, including the ResNet, Alexnet, CNN 5 layer, and CNN 7 layer algorithms. The dataset was used for this purpose, both for training and testing with an accuracy of 96.82% was achieved. As demonstrated, the 5-layer CNN outperforms in all scenarios. The future goal is to develop a web or mobile-based automatic plant identification system, which has the potential to enhance understanding of medicinal plants, advance species identification methods, and aid in the preservation of threatened species.

References

[1] Lengguan, S. (1992). Forest resources and ecosystem conservation in Malaysia. *Recursos Geneticos Forestales (FAO)*.

[2] Milow, P., Malek, S., & Ramli, R. M. (2017). Medicinal Plants of the Indigenous Tribes in Peninsular Malaysia: Current and Future Perspectives. *Active Ingredients from Aromatic and Medicinal Plants*, 1-11.

[3] Onrizal, O., Ismail, S. N., & Mansor, M. (2019, May). Diversity of plant community at Gunung Ledang, Malaysia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 260, No. 1, p. 012068). IOP Publishing.

[4] Chouhan, S. S., Kaul, A., Singh, U. P., & Jain, S. (2018). Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. *Ieee Access*, *6*, 8852-8863.

[5] Simpson, M. G. (2019). *Plant systematics*. Academic press.

[6] Ibrahim, Z., Sabri, N., & Isa, D. (2018). Palm oil fresh fruit bunch ripeness grading recognition using convolutional neural network. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *10*(3-2), 109-113.

[7] Du, C., & Gao, S. (2017). Image segmentation-based multi-focus image fusion through multi-scale convolutional neural network. *IEEE access*, *5*, 15750-15761.

[8] Hu, J., Chen, Z., Yang, M., Zhang, R., & Cui, Y. (2018). A multiscale fusion convolutional neural network for plant leaf recognition. *IEEE Signal Processing Letters*, *25*(6), 853-857.

[9] Amlekar, M. M., & Gaikwad, A. T. (2019). Plant classification using image processing and neural network. In *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2018, Volume 2* (pp. 375-384). Springer Singapore.

[10] Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, S., Navajas, A. D., & Ortiz-Barredo, A. (2017). Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Computers and electronics in agriculture*, *138*, 200-209.

[11] Anand, R., Veni, S., & Aravinth, J. (2016, April). An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method. In *2016 international conference on recent trends in information technology (ICRTIT)* (pp. 1-6). IEEE.

[12] Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017, April). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)* (pp. 1-4). IEEE.

[13] Cruz, A. C., Luvisi, A., De Bellis, L., & Ampatzidis, Y. (2017). X-FIDO: An effective application for detecting olive quick decline syndrome with deep learning and data fusion. *Frontiers in plant science*, *8*, 1741.

[14] Padol, P. B., & Yadav, A. A. (2016, June). SVM classifier based grape leaf disease detection. In *2016 Conference on advances in signal processing (CASP)* (pp. 175-179). IEEE.

[15] Lu, J., Hu, J., Zhao, G., Mei, F., & Zhang, C. (2017). An in-field automatic wheat disease diagnosis system. *Computers and electronics in agriculture*, *142*, 369-379.

[16] Dias, B. L. (2023). Integrating Predictive Models into Public Health Policy: Forecasting Lead Exposure Risks Across the United States. International Journal of Humanities and Information Technology, 5(03), 18-38.

[17] Fujita, E., Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2016, December). Basic investigation on a robust and practical plant diagnostic system. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)* (pp. 989-992). IEEE.

[18] Begue, A., Kowlessur, V., Singh, U., Mahomoodally, F., & Pudaruth, S. (2017). Automatic recognition of medicinal plants using machine learning techniques. *International Journal of Advanced Computer Science and Applications*, *8*(4), 166-175.

[19] Dissanayake, D. M. C., & Kumara, W. G. C. W. (2021). Plant leaf identification based on machine learning algorithms. *Sri Lankan Journal of Technology*, 60-66.

[20] Naeem, S., Ali, A., Chesneau, C., Tahir, M. H., Jamal, F., Sherwani, R. A. K., & Ul Hassan, M. (2021). The classification of medicinal plant leaves based on multispectral and texture feature using machine learning approach. *Agronomy*, *11*(2), 263.

[21] Chung, Y., Chou, C. A., & Li, C. Y. (2021). Central attention and a dual path convolutional neural network in real-world tree species recognition. *International Journal of Environmental Research and Public Health*, *18*(3), 961.

[22] Xue, J., Fuentes, S., Poblete-Echeverria, C., Viejo, C. G., Tongson, E., Du, H., & Su, B. (2019). Automated Chinese medicinal plants

classification based on machine learning using leaf morpho-colorimetry, fractal dimension and visible/near infrared spectroscopy. *International Journal of Agricultural and Biological Engineering*, *12*(2), 123-131.

[23] Kaur, S., & Kaur, P. (2019). Plant species identification based on plant leaf using computer vision and machine learning techniques. *Journal of Multimedia Information System*, *6*(2), 49-60.

[24] Aakif, A., & Khan, M. F. (2015). Automatic classification of plants based on their leaves. *Biosystems Engineering*, *139*, 66-75.

[25] Ahmed, N., Khan, U. G., & Asif, S. (2016). AN AUTOMATIC LEAF BASED PLANT IDENTIFICATION SYSTEM. *Science International*, *28*(1).

[26] Shruthi, U., Nagaveni, V., & Raghavendra, B. K. (2019, March). A review on machine learning classification techniques for plant disease detection. In *2019 5th International conference on advanced computing & communication systems (ICACCS)* (pp. 281-284). IEEE.

[27] Vaishnnave, M. P., Devi, K. S., Srinivasan, P., & Jothi, G. A. P. (2019, March). Detection and classification of groundnut leaf diseases using KNN classifier. In *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)* (pp. 1-5). IEEE.

[28] Oppong, S. O., Twum, F., Hayfron-Acquah, J. B., & Missah, Y. M. (2022). A novel computer vision model for medicinal plant identification using log-gabor filters and deep learning algorithms. *Computational Intelligence and Neuroscience*, *2022*.