# Comparative Study of Convolutional Neural Networks for Leaf Classification

Sweety Kunjachan[1], Remya R[1] Santhos Kumar[1], Asish G R[2], Sheela K[2], and Kala S[1]

[1] Department of ECE, Indian Institute of Information Technology Kottayam, India,
{sweetygisat.22phd12003,rr.madathil22phd12001}@iiitkottayam.ac.in
{santhos,kala}@iiitkottayam.ac.in

[2] Vaidyaratnam Ayurveda Research Institute (VARI), Thaikkattussery, Thrissur, Kerala, India,
rd.asish@vaidyaratnam.com

**Abstract.** Ayurveda is the traditional medicine of India. Herbs are the main essence and are at core of most Ayurvedic medicines. A reduction in the number of experts in the field of identification of herbs reduces the quality of medicines. Recognition of herbs through automatic classification of leaves is the remedy for this problem. A vast number of deep learning models for classification of images are available in the present era. This paper makes a comparison between various state-of-the-art pretrained models for classification of leaves. Architectures evaluated include VGG-16, ResNet-50, DenseNet-121, MobileNetV2, EfficientB0 and EfficientV2M. This experiment was done with a data of 1835 images of leaves taken from 30 different species. ResNet-50, VGG-16, EfficientV2M achieved best results with an accuracy of 98% and EfficientV2M attained maximum accuracy with less number of epochs.

**Key words:** Deep learning, Transfer learning, Convolutional neural network, Image recognition, Artificial intelligence

## 1 Introduction

Ayurveda is the gleaming crown jewel amongst India's various contributions to the betterment of mankind [1]. Ayurveda believes that health is a mirror of harmony between nature and mankind. Ayurveda focuses mainly on root causes of diseases and tries to rectify them through natural remedies. Herbs play a vital role in the preparation of ayurvedic medicines. According to FFPRI (Forestry and Forest product Research Institute) in Japan, forest area per capita has decreased by 60% over the past 60 years [2]. As a result of decreasing forest area, the availability of medicinal plants has become scarce. General awareness and knowledge of medicinal plants is also decreasing due to multiple socio-economic reasons. A huge cache of knowledge of the medicinal value of plants found in the interiors of forest ecosystems, accumulated by the villagers and tribals by way of being passed down to them from their forefathers remains unknown and untapped to the scientists and modern community living far away in cities. Thus, the immediate concern is to preserve this knowledge in digital form through concepts of machine learning, pattern recognition and computer vision.

Identification of plants is manually done by observing the morphological characteristics such as root, stem, fruits, leaves etc. Among them, leaves are found above the ground and will be present over the life of a plant while fruits and flowers remain only for a relatively small period of time [3]. Hence researchers mainly focus on classification of leaves for identification of herbs.

Artificial intelligence has found its way to a wide range of applications [4] such as computer aided plant disease classification, herb recognition etc. Advancements in Deep learning models are very crucial for these studies. There have been many significant breakthroughs in image classification through deep convolutional neural network [5]. The pretrained models coming under this category helps developers to build AI models without being explicitly built from scratch. Transfer learning makes use of knowledge gained while solving one problem and applying it to a different but related problem. It may be used as is or further fine-tuned to fit an application's specific needs [6]. A model that has already been pre-trained on a large dataset saves training time period and reduces computational cost whilst also delivering better accuracy with a smaller number of image dataset. This study is an attempt on an empirical analysis of pretrained models commonly entrusted with the task of leaf classification. The architectures used are VGG-16, ResNet-50, DenseNet-121, MobileNetV2, EfficientB0 and EfficientV2M.

The rest of the paper is organized as follows. Section 2 looks at related work done in the field of leaf classification. In Section 3 we describe some of the existing state-of-the-art techniques required to accomplish this task. Section 4 presents the experimental setup as well as the results and Section 5 discussion and conclusion.

## 2 Recent works

Several approaches are used for the classification of leaves. The pretrained models are the saved networks that are already trained over the large set of images such as imagnet. Because of the fewer time and simple computational complexity the pre-trained models gained much popularity nowadays.

In [7] researchers classified the plant disease based on the AlexNet model. The nine different plants with two categories such as diseased and healthy leaves have been selected for study. The images were resized to the required format. Entire dataset has been divided into two: training and testing data. The pre-trained AlexNet model classified the leaves into two categories and compared performance parameters of accuracy, confusion matrix etc with the support vector machine. AlexNet gained 91.15% accuracy and SVM with linear kernel achieved 89.69% accuracy.

The authors of [8] did the quality inspection of different sugarcane varieties by using convolutional neural networks. The successful classification increases the yield as well as plant population. Two levels of transfer learning approaches have been used in this paper. In first level transfer learning, all models have been pre-trained with a large dataset of single sugarcane variety. In second level transfer learning, models are retrained with a small dataset of different varieties of sugarcane billets. Four architectures used for this classification are AlexNet, VGG-16, GoogleNet and ResNet. Last fully connected layer consists of only two neurons, as it has only two classes. The trade off between processing time and accuracy is essential for image classification. AlexNet attained high accuracy and low processing time.

In [9] researchers recognized the leaves of different medicinal plant species by using some image processing techniques and Artificial Neural Network. In this paper six different species of plants from 40 different sites are selected for the experimentation purpose. The segmented images from the background are given to the feature extractor for extracting color, texture and shape. The Artificial Neural Network classified the images and mean square error and correct classification rate are evaluated.

In [10] the authors used MobileNet architecture for the detection of apple leaf disease. Since MobileNet architecture is a lightweight deep learning method, it has become easier to deploy on mobile phones. This Low cost technique provides higher stability and precision. Training and testing dataset has been selected in the ratio of 3:1. Accuracy and average handling time are the two performance metrics used for the evaluation. MobileNet has been compared with ResNet-152 and InceptionV3. MobileNet achieved higher efficiency.

The authors of [11] uses transfer learning for the detection of leaf disease. In their work pre-trained convolutional neural networks such as ResNet 50 were developed and compared the accuracy with other pretrained models such as VGG16, VGG19, and AlexNet. The performance of ResNet 50 outperforms all other models.

In [12] researchers introduce classification techniques for the recognition of Malaysian herbs using deep learning algorithms. Self acquired images were used for this work and the coloured images were converted to gray scale image as a pre-processing step. The canny edge detector extracts the edges. Zernike and Hu are the two descriptors that collect shape features. Through GLCM techniques textural features were extracted. Classifiers such as support vector machine (SVM) with RBF kernel and deep learning neural network (DLNN) classify images and mobile app for recognition of herbs have been developed in this work. DLNN achieved higher accuracy when compared with SVM.

## 3 Materials and Methods

Convolutional neural networks (CNN) are the backbone of image classification tasks and provide dramatic improvements in performance compared to traditional image processing techniques. The advancements in CNN generate more complex and accurate computer vision models. Feature maps can be extracted automatically and by gaining knowledge the entire model can be utilized for other related tasks.

### 3.1 Dataset

This study has been done using a dataset from freely accessible Mendeley medicinal leaf dataset. A total of 1835 images, consisting of 60 to 100 segmented, colored high quality images, of 30 different species of medicinal herbs like Sandalwood, Mint, Mustard, Jamica cherry etc are available. In order to make the dataset available to deep learning techniques, data augmentation was done. The data is split into two categories: one set for training and the other for testing purposes. A splitting ratio of 80%:20% is used for training and validation. Evaluation of the model was done based on the test data.

### 3.2 Deeplearning classifiers

**VGGNet** VGGNet is a convolutional neural network proposed by A Zisserman and K Simonyan [13]. The architecture consists of convolutional layers, max pooling and fully connected layers. Convolutional layers of 3x3 kernels are followed by ReLU. ReLU stands for rectified linear unit of activation function. ReLU will produce an output when the input is positive or it results in zero. Max pooling reduces the dimensionality. Fully connected layers and softmax layers are added as final layers [14].

**ResNet** Residual Network is a deep neural network that uses the concept of skip connection [15]. In skip connection the original input is added to the output of the convolutional block and it skips certain layers to overcome vanishing gradient problem. ResNet is eight times deeper than VGG nets. It consists of convolutional layers and pooling layers. Batch normalization is introduced between convolutional layer and activation function. Two types of skip connections are used in residual networks and it depends upon the dimensions of both input and output. If the input and output activation are in the same dimension, then identity blocks are used. On the other hand, if it is not in the same dimension, convolutional blocks need to be inserted. Convolutional block uses CONV2D layer in the shortcut path.

**MobileNet** MobileNet is a lightweight deep convolutional neural network. The main aim of mobileNet architecture is to make the size of the system small and maximize speed [16]. Depthwise separable convolution used in MobileNet reduces the number of parameters. Two operations involved in this architecture are depthwise convolution followed by pointwise convolution. Depthwise convolution applies a single filter to each input channel. Pointwise convolution applies 1x1 convolution to the output of depthwise convolution. MobileNetV2 is the second version of MobileNet. Three convolution layers were used in this architecture along with residual connection. First convolution block is the expansion layer as it expands the number of channels before giving it to depthwise convolution. The second block is the same as that of MobileNet. Third block acts as a bottleneck layer that reduces data flow. Residual network removes the problem of vanishing gradient.

**DenseNet** DenseNet is a deep convolutional neural network in which all layers are interconnected [17]. Each layer receives feature maps from previous layers. Hence it reduces vanishing gradient problem. Dense blocks and Transition blocks are the main building blocks of DenseNet. Each dense layer consists of batch normalization, ReLU, 3x3 convolution and drop out [18]. Transition blocks consist of batch normalization, ReLU, 1x1 convolution, drop out and pooling layer. Size of the feature maps should remain the same when we do concatenation and so this operation is performed in the transition block. Connection between layers enhances feature propagation and concatenation encourages feature reuse.

**EfficientNet** EfficientNet is a convolutional neural network that performs scaling on depth, width and resolution. Deeper networks have the ability to capture more complex features from images but it will have the disadvantage of higher training time due to vanishing gradient. Width scaling works well only for small models. The finer features can easily be captured from high resolution images but it will result in diminishing accuracy gain. Hence balancing of all dimension scaling is very important for achieving better efficiency and accuracy. Compound scaling is an efficient remedy for this issue. Constant coefficients are determined by a small grid search technique. EfficientNet B0 is the baseline model. Mobile inverted Bottleneck convolutional layers and squeeze and excite blocks are the main building blocks of EfficientNet architecture [19]. EfficientNetV2M is the second version of EfficientNet model and provides higher training speeds and a better parameter efficiency. It introduces training aware architecture search and it chooses the best combination of MBConv, MBFusedConv, number of layers, kernel size and expansion ratio. Progressive learning with adaptive regularization is the other contribution of EfficientNetV2 model outperforming vision transformer [20].

## 4 Results

### 4.1 Experiments

Deep learning models are developing by the day and image classification tasks with smaller datasets are taking advantage of pre-trained models. Pre-trained models are the ones which were already trained on larger datasets. Knowledge gained from previous experience can be transfered to other high similarity low quantity cases thereby achieving higher efficiency and reduced training time. The block diagram of pre-trained models are shown in Fig. 1.

In Deep learning models, the first few layers are used predominantly for feature extraction and the final layers for image classification. VGG-16, ResNet-50, DenseNet-121, MobileNetV2, EfficientNetB0 and EfficientV2M are some of the pretrained models taken for this study. These models have been pretrained to the imagenet dataset which contains 1.2 million images under a 1000 different categories. The knowledge acquired from this training has been used to train new datasets. The time required for training fine-tuned models is relatively much

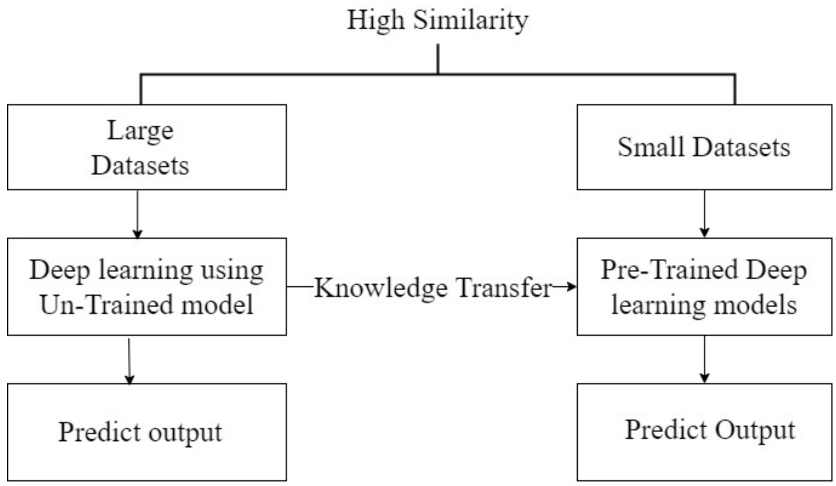


**Fig. 1.** Block diagram of pre-trained models

lesser when compared with that of models built from scratch. New datasets are taken from Mendley medicinal leaf dataset which consists of 1835 images belonging to 30 different species. 80% of images are used for training and the remaining 20% for testing adding up to 1468 and 367 images respectively. The images have been resized in order to be compatible with the deep learning models. The experiments were performed on graphics processing unit (GPU) mode. The training of images is carried out in 32 batches and uses Adam for optimization. Initial layers of all models were kept unaltered and a dense layer of 30 with softmax layer is added to final layers. Each of the experiments are carried out in 5 and 15 epochs. The performance metrics such as accuracy and loss are monitored and noted.

### 4.2 Results and Discussion

In this study, an analysis of six pre-trained deep learning models has been done for the task of leaf classification. The results of experiments are shown in Fig.2-8. Accuracy and loss per epoch of each model are plotted in graphs.

Table1 analyzed performance of all models for 5 epochs. All architectures except DenseNet attained accuracy above 80%. EfficientV2M and EfficientB0 achieved the highest results with an accuracy level of 97.5%. The validation accuracy, training accuracy, validation loss and training loss according to epochs are described in the table. It has been noted that DenseNet performs poorly with fewer iterations.

Table 2 compared different models for 15 epochs. After 15 epochs ResNet-50, VGG-16 and EfficientV2M provides maximum accuracy. Among the three

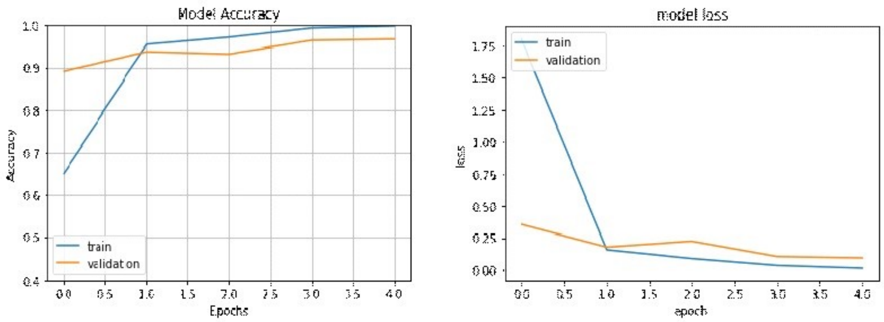


**Fig. 2.** Accuracy and loss per epoch of VGG-16

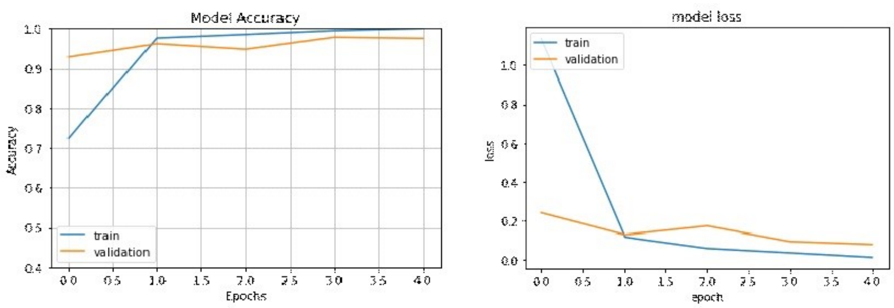


**Fig. 3.** Accuracy and loss per epoch of ResNet-50

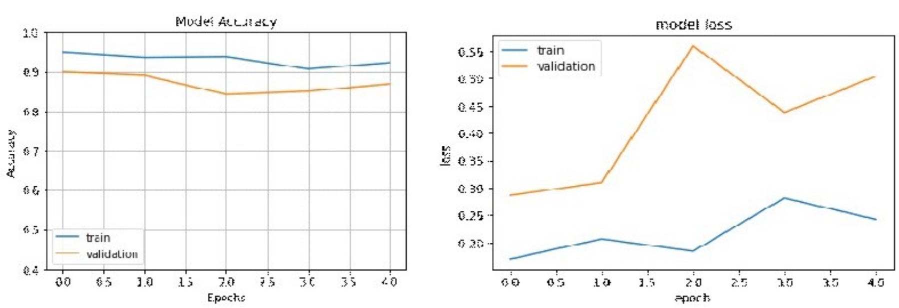


**Fig. 4.** Accuracy and loss per epoch of DenseNet-121

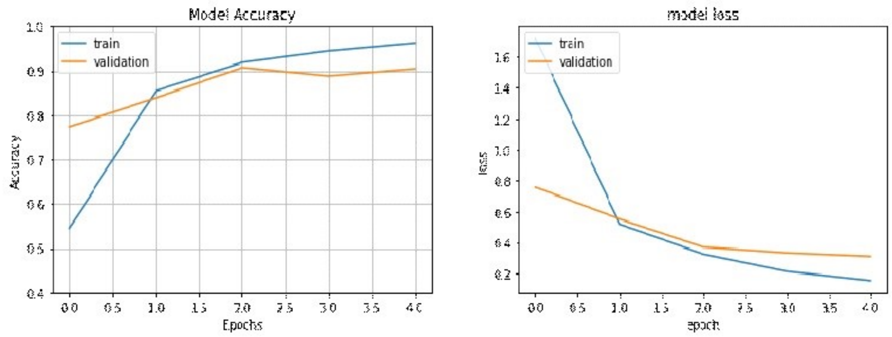


**Fig. 5.** Accuracy and loss per epoch of MobileNetV2

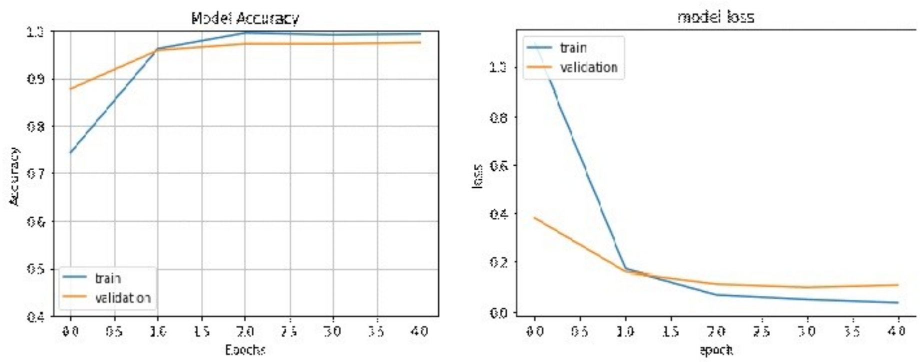


**Fig. 6.** Accuracy and loss per epoch of EfficientNetB0

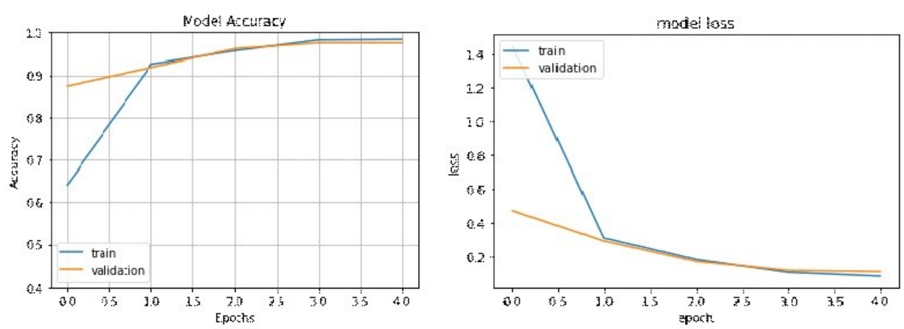


**Fig. 7.** Accuracy and loss per epoch of EfficientNetV2M

**Table 1.** Comparison of Pre-trained models for 5 epochs

| Model | Training Accuracy% | Test Accuracy% | Training loss | Test Loss |
|---|---|---|---|---|
| VGG-16 | 100 | 96.7 | 0.00597 | 0.0905 |
| ResNet-50 | 100 | 97.54 | 0.00958 | 0.09074 |
| DenseNet-121 | 91.96 | 86.92 | 0.2427 | 0.5034 |
| MobileNetV2 | 97.41 | 90.46 | 0.1282 | 0.3119 |
| EfficientNetB0 | 99.9 | 97.54 | 0.01728 | 0.1088 |
| EfficientNetV2M | 99.72 | 97.54 | 0.03641 | 0.1120 |

models EfficientV2M converges faster and hence EfficientV2M gives best results in-terms of both accuracy, time and with fewer number of epochs.

**Table 2.** Comparison of Pre-trained models for 15 epochs

| Model | Training Accuracy% | Test Accuracy% | Training loss | Test Loss |
|---|---|---|---|---|
| VGG-16 | 100 | 98.63 | 0.0011 | 0.0553 |
| ResNet-50 | 100 | 98.91 | 0014 | 0.0663 |
| DenseNet-121 | 93.32 | 86.92 | 0.2267 | 0.4204 |
| MobileNetV2 | 100 | 93.73 | 0.0139 | 0.1906 |
| EfficientNetB0 | 100 | 97.82 | 0.0027 | 0.0675 |
| EfficientNetV2M | 100 | 98.36 | 0.0044 | 0.0624 |

## 5 Conclusion

Classification of herbs into different categories is an essential field in ayurvedic research. Convolutional neural networks perform well for leaf classification. In this work, fine-tuning and evaluation of state-of-the-art deep learning models is performed. The architectures included in this study are VGG-16, ResNet-50, DenseNet-121, MobileNetV2, EfficientB0 and EfficientV2M. EfficientNetV2M achieved higher accuracy rates of 98% with fewer number of epochs. Thus deep learning networks can easily replace human interventions in the recognition of herbs. Handheld devices for recognition of herbs enhance the collection of herb from the forest area.

## References


1. Mukherjee, Pulok K., Ranjit K. Harwansh, Shiv Bahadur, Subhadip Banerjee, Amit Kar, Joydeb Chanda, Sayan Biswas, Sk Milan Ahmmed, and C. K. Katiyar. "Development of Ayurveda–tradition to trend." Journal of ethnopharmacology 197, pp: 10-24, 2017.
2. Ronald C Estoque et al, Spatiotemporal pattern of global forest change over the past 60 years and the forest transition theory, Environmental Research Letters, 2022, DOI: 10.1088/1748-9326/ac7df5
3. Roopashree S, J Anitha, T R Mahesh, V Vinoth Kumar, Wattana Viriyasitavat and Amandeep Kaur, "An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach." Measurement 200, pp.111484, 2022.
4. Ryan, Mark, "The social and ethical impacts of artificial intelligence in agriculture: mapping the agricultural AI literature." AI and SOCIETY, pp.1-13, 2022.
5. Kaya, Aydin, Ali Seydi Keceli, Cagatay Catal, Hamdi Yalin Yalic, Huseyin Temucin, and Bedir Tekinerdogan. "Analysis of transfer learning for deep neural network based plant classification models." Computers and electronics in agriculture 158, pp.20-29, 2019.
6. S. Niu, Y. Liu, J. Wang and H. Song, "A Decade Survey of Transfer Learning (2010–2020)," in IEEE Transactions on Artificial Intelligence, vol. 1, no. 2, pp. 151-166, Oct. 2020, doi: 10.1109/TAI.2021.3054609.
7. Wagle, Shivali Amit, "Comparison of Plant Leaf Classification Using Modified AlexNet and Support Vector Machine", Traitement du Signal vol. 38, no. 1, 2021
8. M. Alencastre-Miranda, R. M. Johnson and H. I. Krebs, "Convolutional Neural Networks and Transfer Learning for Quality Inspection of Different Sugarcane Varieties," in IEEE Transactions on Industrial Informatics, vol. 17, no. 2, pp. 787-794, Feb. 2021, doi: 10.1109/TII.2020.2992229.
9. Azadnia, R. and Kheiralipour, K., 2021. Recognition of leaves of different medicinal plant species using a robust image processing algorithm and artificial neural networks classifier. Journal of Applied Research on Medicinal and Aromatic Plants, 25, p.100327.
10. Bi, Chongke, Jiamin Wang, Yulin Duan, Baofeng Fu, Jia-Rong Kang, and Yun Shi, "MobileNet based apple leaf diseases identification" Mobile Networks and Applications, pp. 1-9, 2022.
11. I. Z. Mukti and D. Biswas, "Transfer Learning Based Plant Diseases Detection Using ResNet50," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, pp. 1-6, 2019, doi: 10.1109/EICT48899.2019.9068805.
12. A. Muneer and S. M. Fati, "Efficient and Automated Herbs Classification Approach Based on Shape and Texture Features using Deep Learning," in IEEE Access, vol. 8, pp. 196747-196764, 2020, doi: 10.1109/ACCESS.2020.3034033.
13. Simonyan Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition.", in ICLR, 2015.
14. S.Ghosal and K. Sarkar, "Rice Leaf Diseases Classification Using CNN With Transfer Learning," 2020 IEEE Calcutta Conference (CALCON), Kolkata, India, pp. 230-236, 2020, doi: 10.1109/CALCON49167.2020.9106423.
15. Shafiq M and Gu Z, Deep residual learning for image recognition: a survey. Applied Sciences, 12(18), p.8972, 2022

16. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C, "Mobilenetv2: Inverted residuals and linear bottlenecks". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510-4520, 2018.
17. Huang G, Liu Z, Van Der Maaten, L. and Weinberger K.Q, Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708, 2017.
18. D.O'Neill, B. Xue and M. Zhang, "Evolutionary Neural Architecture Search for High-Dimensional Skip-Connection Structures on DenseNet Style Networks," in IEEE Transactions on Evolutionary Computation, vol. 25, no. 6, pp. 1118-1132, Dec. 2021, doi: 10.1109/TEVC.2021.3083315
19. Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." In International conference on machine learning, pp. 6105-6114. PMLR, 2019.
20. Tan, Mingxing, and Quoc Le. "Efficientnetv2: Smaller models and faster training." In International conference on machine learning, pp. 10096-10106. PMLR, 2021.