Automatic Plant Identification Using Transfer Learning

S Sachar¹ and A Kumar²

¹ Research Scholar, Department of Computer Science and Applications, Panjab University, Chandigarh, India.
² Assistant Professor, Department of Computer Science and Applications, Panjab University, Chandigarh, India.

Email for correspondence: silky.sachar@gmail.com

Abstract. Plant identification is a widely researched area in the field of computer vision. Many attempts have been made to automate the process of plant identification using an image of a part of plant including flower, leaf and bark. Leaf has proven be the most reliable source of information. After exhaustive experiments, we chose to apply transfer learning to compare the feature extraction capabilities of VGG-16, Xception, MobileNetV2 Convolutional Neural Network (CNN) and DenseNet121 architectures to freely available Swedish, Flavia and MalayaKew leaf image datasets. Random Forest is used as classifier to identify the species of given leaf. The evaluations and comparisons of the specified feature extractor models are provided. DenseNet121 achieved maximum accuracy of 100%, 99% and 92.4% respectively in the three datasets.

Keywords: Feature extraction, CNN, Transfer learning, Random Forest, DenseNet

1. Introduction

Human beings rely on plants for their survival. Plants not only provide basic food and oxygen but few of them also have immense medicinal value. The main concern faced by common man is the recognition of plants. This has led many researchers to work in the field of automatic identification of plants. Since, leaf is available all the year round, it is considered to the most valuable source of information. To identify the plant from the leaf image, various features of leaf have to be taken into account. These features include shape, color, texture and vein. The computed feature vector then becomes the input of the classifier, which then provides the name of the plant as output. These classifiers are Support Vector Machines (SVM), Random Forests, K-NN etc. Another approach that has gained importance since a decade is Convolutional Neural Network (CNN) as it eliminates the need of hand-crafted feature extractors. Various layers of CNN automatically compute and learn different features of the input images and its last layer provides the output. The main requirement of CNN model is an extensive dataset to train the network as it learns better by seeing more examples. However, various CNN architectures like VGG-16 and VGG-19 [1], ResNet-50 [2] have been developed which have been pre-trained on huge datasets like ImageNet. The concept of transfer learning is, when the network has already learnt the features of the images on which it was trained and is now being trained on the required dataset. Features can be extracted from any intermediate layer of the model. The deeper layers contain more identification information. In this work, the features have been obtained from the fully connected layers of the
network. Then the feature vector has been input into the random forest classifier to obtain the target class. Section 2 shows the related work using hand-crafted feature extractors, deep learning models and transfer learning. Section 3 shows different layers used in the architecture of CNN and models used in our experiments. Section 4 shows the proposed system design, experiments and results. Section 5 concludes the paper.

2. Related Work
There has been ongoing research to identify plants from various parts of plant [3] by applying machine learning algorithms. However, the most available and reliable source of information is considered to be a leaf that is why leaf images have often been used to identify plant species. In order to support this, there are many freely available leaves datasets which help researchers to test their algorithms. These datasets include: LeafSnap [4] which consists of 23147 lab images of leaves, taken in controlled environment and 7719 images taken in field from 185 species of Northern United States, Swedish [5] which consists of 15 species/classes of plants and 75 samples in each class etc.

Main features extracted from the leaf images are shape, color, texture and venation. Various authors have used different methods of feature extraction and classification. For example, authors in[6] insisted that the combination of all these features constituted the feature vector for the Probabilistic Neural Network(PNN). It achieved an average accuracy of 93.75% on the freely available Flavia dataset [7] to identify the plant species from leaf image. Authors in [8] obtained texture features from leaf images using Grey Tone Spatial Dependency matrix and Local Binary Patterns. These features were classified using 6 classifiers where Stochastic Gradient Descent, k-NN and Decision Trees achieved 94.7% accuracy in identifying medicinal plants. Wavelet and fractal dimensions have also been proposed to extract texture features from leaf images [9]. The wavelet method tends to ignore details in the images as these are only concentrated on the low-frequency sub-bands. The fractal method tends to find patterns with different geometrical nature but equivalent fractal dimension. So, the authors propose the combination of wavelet and fractal dimension to extract the texture information. This feature vector when inputted into the Artificial Neural Network with back propagation, achieved 91.1% accuracy on Flavia dataset. To identify soybean, red bean and white bean, authors in [10] used vein pattern of the leaf images by using adaptive thresholding and Unconstrained Hit or Miss transform. These features were input into the Random forest classifier which achieved 95% accuracy. Another work which compared the performance of different classifiers was proposed by [11]. Shape and color features were used as input to Support Vector Machine(SVM), k-NN, Naïve Bayes and Random Forest to conduct experiments. Random Forest attained the highest accuracy of 96%.

Another exemplary work in the same field to identify leaves of same species is mentioned in [12]. Instead of using the whole leaf images, the authors have used only a part of the leaf to obtain the color, venation, texture and shape features. Later the features obtained from these parts were combined for a each whole image. Vein features were obtained using morphological opening operations. Texture features were obtained using Gray Level Co-occurrence Matrix (GLCM) which is based on the relationship between neighboring pixels in grey scale images. Color features were attained by using statistical measures. The shape was calculated using the Fourier descriptors based on the boundary pixels of the leaf region in the image. These features were input into the Extreme Learning Machine(ELM) classifier which is single-hidden layer feed forward network. It is considered to be fast classifier as the network parameters (weights and biases) start randomly and remain constant throughout the learning process. The Least squares method is used to calculate the output weights. This method achieved 99.1% accuracy on Flavia dataset.

Since past few years, the emphasis has shifted to Deep learning algorithm called Convolutional Neural Network (CNN). It reduces the amount of human intervention by learning the useful features from the input images and classifying them. Authors in [13] proposed a Multi-Scale Fusion CNN to identify plants from leaf images by down-sampling the input images to multiple low level images so that they be fed step-by-step to the proposed algorithm. The features obtained at different scales are then fused to obtain all the information which can help predict the species of plant. To detect diseases
that can affect plants, authors in [14] successfully developed a consistent and reliable deep CNN 9 layer model. The input data was first augmented to increase the number of samples and help CNN learn better. It was tested against various classifiers and achieved 96% accuracy. Similarly, Authors in [15] developed a 50 layer deep residual network which attained an accuracy of 98.7% with the loss of 0.0462% on the LeafSnap dataset.

Another most evolving aspect of Deep learning is transfer learning. Transfer learning is the concept of using the knowledge obtained from one task to solve other tasks. It has been explained in detail in [16] For computer vision tasks, most popular models are VGG-16, VGG-19, Inception V3 [17], Xception [18], ResNet-50. These are the state-of-art architectures which are open sourced. The parameters and weights obtained during their training, can be used by others researchers to solve the problems in their own domains. A CNN model called D-Leaf was proposed in [19] to extract features from the input images. Its performance was compared to feature extraction from Alexnet and fine-tuned Alexnet. These features were then fed into various classifiers. Artificial Neural Network (ANN) achieved maximum accuracy of 94.88%. Another application of transfer learning was observed in [20] where the authors used a pre-trained model called MobileNet to extract features of leaf images from Flavia and LeafSnap datasets and then classified them by using Logistic regression classifier. They obtained 99.6% accuracy for Flavia which has 32 classes and 90.5% for LeafSnap which has 184 classes. To compare various parameters used in Deep learning, authors in [21] evaluated the performance of GoogleNet, AlexNet and VGGNet on LifeCLEF 2015 dataset which consists of images of leaves, flowers, branch, fruit, stem from all classes. They studied the effects of network parameters like iteration size, batch size and data augmentation while training these networks from scratch and while fine-tuning them. They showed that fine-tuning GoogleNet and VGG from scratch performs better and training AlexNet from scratch provides better results due to its simple architecture than the other two. Authors in [22] used MobileNet which is comparatively lighter than other pre-trained models due to its depthwise separable convolutions which also makes the computation less expensive. They have implemented 16 different combinations based on different hyperparameters of MobileNet on their own medicinal leaf dataset. Highest accuracy of 98.7 % was reported.

Most challenging area in the field of automatic recognition of plants is when the images of leaves are obtained in complicated background. Deep learning has also been used in this field [23]. The authors proposed Inception V2 as feature extractor. Images were first divided into sub-images. Batch normalization was used as a regularization method in the proposed Fast Region Convolutional Neural Network (RCNN). It helped improve the classification accuracy. The multiscale features were then input into the region proposal network (RPN). The softmax and the bounding box regressor were used as classifier. This approach proved to be better than the conventional Faster RCNN.

3. Architecture of Convolutional Neural Networks

- Traditionally, multilayer perceptron (MLP) models were used to recognize images but due to full connectivity between the nodes the higher resolution images did not scale well. CNN overcome the issues of MLPs. CNN models assume that the input are images are converted to a single feature vector which is thus transformed through a series of hidden layers. The layers of CNN are made up of neurons. Neurons are arranged in 3 dimensions namely width, height and depth. These are arranged to a region of layer above them and not to all neurons as in fully connected fashion. All the layers in CNN transforms the 3D volume input into 3D volume output activations. Basic building blocks of CNN are described below:

  - Convolutional Layer: This is considered as the core building block of the CNN architecture. The filters are considered as learnable parameters as these are convolved across the input data. The dot product between the filter and the input produces an activation map. This process makes the network learn filters which activate on detecting a specific type of feature.

  - Pooling Layer: This layer compresses the input non linearly. Example, Max pooling divides the input into non overlapping rectangles and computes maximum for each sub region. This controls over fitting.
• ReLU Layer: It stands for rectified linear unit. The activation function used is \( f(x) = \max(0, x) \). It helps to remove negative values from activation map by setting them to zero, thereby increasing the non-linearity of decision function and of whole network. Various functions that can be applied at this layer are tanh and sigmoid.

• Fully connected Layer: Neurons in this layer are connected to activations of previous layer. These are considered to perform high-level feature recognition.

• Loss Layer: This is the final output layer in the CNN. It attains the true labels from the predicted outputs. Example, softmax is used to predict a class from K mutually exclusive classes of inputs.

Following CNN architectures are selected after an exhaustive experimentation on the selected datasets to give high accuracy. These models have stood out in ImageNet Large Scale Visual Recognition Challenges (ILSVRC) and are mentioned below:

• VGG-16: This model was proposed in [1] It won 1st and 2nd place in image detection and image classification tasks in 2014 ILSVRC. This model achieved top-1 accuracy of 0.713 and top-5 accuracy of 0.901 on ImageNet validation dataset. It consists of 138 million parameters from 16 convolutional layers with 3x3 receptive fields, 5 max pooling layers of size 2x2, 3 fully connected layers and linear layer using softmax activation function. Its architecture is simple. The convolution layers are stacked on top of each other. Max-pooling layer reduces the volume size. Then there are 2 fully connected layers followed by softmax classifier layer.

• Xception: This model is an extension of Inception V3 model [18] The Inception modules were replaced by depthwise separable convolutions. The pointwise convolutions (1x1 conv to change the dimension) is followed by depthwise convolution (channel wise nxn spatial conv). The intermediate ReLU non-linearity is also absent. It achieved Top-1 accuracy of 0.790 and top-5 accuracy of 0.945 on ImageNet validation dataset.

• MobileNetV2 [24] model is improvement of MobileNetV1 where non-linearities in narrow layers were removed. It consists of 2 types of blocks. First is residual block with stride of 1 and another for downsizing with stride of 2. Each block has 3 layers. Layer 1 is 1x1 convolution with ReLU6. Layer 2 is depthwise convolution. Layer 3 is 1x1 convolution without non-linearity. It has achieved top-1 accuracy of 0.713 and top-5 accuracy of 0.901 on ImageNet validation dataset.

• DenseNet121 [25] is shown in Figure 1. Every layer is connected directly to each other and helps to reduce vanishing gradient problem. It has about 8 million parameters. DenseNets are made up of dense blocks. There are only 12 feature maps per layer due to which only few sets of information are added to the ‘knowledge’ of the network. The dimension of feature map in the block remains constant and number of filters changes between them. Transition layers between them apply batch normalization which helps to reduce overfitting on training sets with few numbers of images, 1x1 convolution (bottleneck layer) and 2x2 pooling. This is different from ResNet as there is concatenation of feature maps which were learned by different layers. This process not only increases the variation in the input to subsequent layers but also improves the efficiency of the network. The global state is accessible from any layer which eliminates the need of replication. It has achieved top-1 accuracy of 0.750 and top-5 accuracy of 0.923 on ImageNet validation dataset.
4. Proposed System Design

The experiments are divided into three categories on three publicly available datasets namely Swedish, Flavia and MalayaKew [25]. For the purpose of experimentation, datasets Flavia and Swedish were augmented. Since MalayaKew was already balanced dataset, hence no augmentation was applied. The pre-trained models namely VGG-16, Xception, MobileNet V2 and DenseNet 121 have been trained, using the state-of-art systems, on the ImageNet dataset which consists of 1000 classes of images. Instead of using the hand-crafted feature extractors, we have made use of these pre-trained models for the same purpose. The features learnt by the fully connected layers of these models are then input into the random forest classifier as shown in Figure 2. Accuracy is used as an evaluation metric shown in Eq 1.

\[
\text{Accuracy} = \frac{\text{Correct test set predictions}}{\text{Total No of Test set predictions}}
\]

Figure 1: Architecture of DenseNet121

Figure 2: Proposed system overview
4.1. Datasets:
The datasets used in this paper are Swedish dataset consisting of 15 classes [5], Flavia consisting of 33 classes [7] and MalayaKew consisting of 44 classes [25]. Only whole images were used. Training set images from all datasets were augmented by rotating the images by 40º, translating the images horizontally and vertically by a fraction of 0.2, applying shearing transformation, zooming inside the images, horizontally flipping half of the pictures and used ‘nearest’ as the mode to fill in the newly created pixels. As it has been seen in the previous works that CNNs learn better with a greater number of examples and the number of images in some classes of leaves were fairly less. One sample image from each dataset is shown in Figure 3.

Figure 3 (A) Populus tremula from Swedish, (B) Canadian Poplar from Flavia, (C) qlobata from MalayaKew

4.2. Transfer learning for the feature extraction
Feature set obtained from leaf images can be based on shape, texture, color or venation. These features can also be used in combination. Here, we present a comparison of feature extraction capabilities of 4 pre-trained CNN models. Transfer learning is the concept of using the knowledge acquired by the network using the state-of-art process, to be applied for identification of other related images. In this work [26] it is shown that the features obtained from these pre-trained networks are also efficient in recognizing other related images. Feature extraction helps to curb over-fitting of CNN models, speed up training, improve accuracy and improve data visualization. It also reduces the features from the dataset by creating new features from the existing ones and discarding the original ones. This process creates a summary of the original features from the dataset. Features for our experiments were obtained from fully connected layer of the pre-trained models. They are summarized in Table 1 (Depth refers to the topological depth which includes activation layers, batch normalization layers etc).

Table 1: List of pre-trained models used, the input image size, depth and number of features extracted from each model.

| Model name        | Input image size | Depth | No. of features extracted |
|-------------------|------------------|-------|---------------------------|
| VGG-16            | 224x224          | 23    | 4096                      |
| Xception          | 299x299          | 126   | 2048                      |
| MobileNetV2       | 224x224          | 88    | 1280                      |
| DenseNet121       | 224x224          | 121   | 1024                      |
4.3. Classification
For the purpose of classification, Random forest classifier is used as it is robust and generalizes well. Random Forest is ensemble of decision trees. These independent decision trees are each trained on the subset of feature set from the training set to ensure that they are learning to make the predictions in different ways. Then simple voting is used to pool their outputs together. Different parameters taken into account for this classifier were: number of estimators, maximum features, maximum depth, minimum samples required to split a node, minimum samples required at each leaf node and bootstrap. All these parameters were tuned using RandomizedSearchCV for each training set. 3-fold cross validation was used which means that for every hyperparameter combination the model got trained on 2/3 of training data and 1/3 was used for validation. Thus, this eliminates the need of separate validation data.

4.4. Experiments and Results
Various experiments were conducted using different pre-trained models on the mentioned datasets. The ones which helped to achieve exemplary results are shown in this section. The images were first augmented and then resized according to requirements of the feature extractor model used. The size of feature vector for different models is shown in Table 1. The feature vector was then input into the random Forest classifier which was used to predict the species of the leaf (target class). For the purpose of evaluation, accuracy/ precision was used as a metric. The explanation of the datasets and comparative analysis is provided below.

4.4.1 Evaluation on Swedish dataset. Swedish dataset was introduced in [5] It consists of 15 classes of leaves. An image from this dataset is shown in Figure 3 (a). Images of training set are augmented so that the model could learn as many numbers of features as possible. The images are first resized and pre-processed based on the requirements of the feature selector and then classified using random forest classifier. Various hyperparameters tuned for each case and their accuracies obtained on the test set are shown in the Table 2 below:

| Feature selector | Accuracy (%) |
|------------------|-------------|
| VGG-16           | 99          |
| Xception         | 95          |
| MobileNetV2      | 99          |
| DenseNet121      | 100         |

4.4.2 Evaluation on Flavia dataset. It is one of the most widely used datasets in the field of plant recognition using leaf images. The dataset consists of 33 classes/ species of plants. The images are 1200x1600 pixels. The train dataset is augmented and images are resized and pre-processed as required by the feature extractor model. Then random forest classifier is used and hyperparameters tuned accordingly. Table 3 shows the various accuracies obtained in each case

| Feature selector | Accuracy (%) |
|------------------|-------------|
| VGG-16           | 98          |
| Xception         | 95          |
| MobileNetV2      | 98          |
| DenseNet121      | 99          |
4.4.3 Evaluation on MalayaKew dataset. This dataset consists of 44 classes of leaves images of size 256x256 pixels taken from Royal Botanic Gardens, Kew, England. The whole dataset consists of three sub datasets. In this experiment only MK-D1 was taken into account which consists of whole leaf images. Images in this dataset were not augmented. Table 4 shows the accuracies obtained after the necessary resizing and pre-processing.

Table 4: Accuracy attained from each pre-trained model on MalayaKew Dataset.

| Feature selector | Accuracy (%) |
|------------------|--------------|
| VGG-16           | 79           |
| Xception         | 75           |
| MobileNetV2      | 76           |
| DenseNet121      | 92.4         |

From the above experiments, it can be observed that pre-trained DenseNet121 architecture outperforms pre-trained VGG-16, Xception and MobileNetV2 in extracting features from leaf images. Table 5 shows the comparison of the proposed technique to the other state-of-the-art techniques in the literature. Combination of DenseNet121 and Random Forest achieved comparable/good feature extraction and classification accuracies in Flavia ,Swedish and MalayaKew datasets.

Table 5. Related state-of-the-art works and accuracies on the Flavia, Swedish and MalayaKew datasets, in percentage (RF-Random Forest, LBP-Local Binary Patterns, ELM-Extreme Learning Machines, HOG- Histogram of Oriented Gradients, SVM- Support Vector Machines,ANN- Artificial Neural Network).

| Reference | Year | Methodology | Flavia | Swedish | Malakew |
|-----------|------|-------------|--------|---------|---------|
| Our study | 2020 | DenseNet121/RF | 99     | 100     | 92.4    |
| [27]      | 2019 | LBP/ELM     | 98.9   | 99.4    | -       |
| [28]      | 2019 | VGG-16/SVM  | 96.1   | -       | -       |
| [19]      | 2018 | D-leaf/ANN  | 94.6   | 98.09   | 90.3    |
| [29]      | 2017 | HOG,LBP/SVM | 97     | -       | -       |
| [9]       | 2015 | Wavelet, Fractal/ANN | 91.1 | - | - |

5. Conclusion
The performance of pre-trained models as feature extractors was observed. DenseNet 121 as a feature extractor outperformed other CNN pre-trained models. DenseNet solves the vanishing gradient problem encountered in CNN as they grow deeper. They also have narrow layers, few parameters and allow feature re-use. In all the cases above, it is clear that the feature extracted from DenseNet when input into random forest resulted in maximum accuracy possible. It achieved 100% accuracy on Swedish dataset , 99% accuracy on Flavia 92.4% accuracy in MalayaKew dataset.

References
[1] Simonyan K and Zisserman A 2015 Very deep convolutional networks for large-scale image recognition 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–14
[2] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition vol 2016-December
[3] Nazarenko D V., Kharyuk P V., Oseledets I V., Rodin I A and Shpigun O A 2016 Machine learning for LC-MS medicinal plants identification Chemom. Intell. Lab. Syst.

[4] Kumar N, Belhumeur P N, Biswas A, Jacobs D W, Kress W J, Lopez I C and Soares J V B 2012 Leafsnap: A Computer Vision System for Automatic Plant Species Identification BT - Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 7573 LNCS 502–16

[5] Söderkvist O J O 2001 Computer Vision Classification of Leaves from Swedish Trees 74

[6] Kadir A, Nugroho L E, Susanto A and Santosa P I 2013 Leaf Classification Using Shape, Color, and Texture Features 225–30

[7] Wu S G, Bao F S, Xu E Y, Wang Y X, Chang Y F and Xiang Q L 2007 A leaf recognition algorithm for plant classification using probabilistic neural network ISSPIT 2007 - 2007 IEEE Int. Symp. Signal Process. Inf. Technol. 11–6

[8] H.Arun C, R. Sam Emmanuel W and Christopher Dunairaj D 2013 Texture Feature Extraction for Identification of Medicinal Plants and comparison of different classifiers Int. J. Comput. Appl. 62 1–9

[9] Zhang H and Tao X 2015 Leaf image recognition based on wavelet and fractal dimension J. Comput. Inf. Syst. 11 141–8

[10] Larese M G, Craviotto R M, Arango M R, Gallo C and Granitto P M 2012 Legume identification by leaf vein images classification Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) vol 7441 LNCS pp 447–54

[11] Caglayan A, Guclu O and Can A B 2013 A plant recognition approach using shape and color features in leaf images Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 8157 LNCS 161–70

[12] Turkoğlu M and Hanbay D 2019 Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts Appl. Math. Comput. 352 1–14

[13] Hu J, Chen Z, Yang M, Zhang R and Cui Y 2018 A multiscale fusion convolutional neural network for plant leaf recognition IEEE Signal Process. Lett. 25 853–7

[14] Geetharamani G and J. A P 2019 Identification of plant leaf diseases using a nine-layer deep convolutional neural network Comput. Electr. Eng. 76 323–38

[15] Bodhwani V, Acharjya D P and Bodhwani U 2019 Deep residual networks for plant identification Procedia Comput. Sci. 152 186–94

[16] Pan S J and Yang Q 2010 A survey on transfer learning IEEE Trans. Knowl. Data Eng. 22 1345–59

[17] Zeng G, He Y, Yu Z, Yang X, Yang R and Zhang L 2016 InceptionNet/GoogLeNet - Going Deeper with Convolutions Cvpr 91

[18] Chollet F 2017 Xception: Deep learning with depthwise separable convolutions Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017 2017 Janua 1800–7

[19] Tan J W, Chang S W, Abdul-Kareem S, Yap H J and Yong K T 2020 Deep Learning for Plant Species Classification Using Leaf Vein Morphometric IEEE/ACM Trans. Comput. Biol. Bioinformatics. 17 82–90

[20] Beikmohammadi A and Faez K 2018 Leaf Classification for Plant Recognition with Deep Transfer Learning Proc. - 2018 4th Iran. Conf. Signal Process. Intell. Syst. ICSPIS 2018 21–6

[21] Mehidipour Ghazi M, Yanikoglu B and Apytoula E 2017 Plant identification using deep neural networks via optimization of transfer learning parameters Neurocomputing 235 228–35

[22] Duong-Trung N, Quach L Da, Nguyen M H and Nguyen C N 2019 A combination of transfer learning and deep learning for medicinal plant classification ACM Int. Conf. Proceeding Ser. Part F1479 83–90

[23] Zhu X, Zhu M and Ren H 2018 Method of plant leaf recognition based on improved deep convolutional neural network Cogn. Syst. Res.
[24] Sandler M, Howard A, Zhu M, Zhmoginov A and Chen L C 2018 Inverted Residuals and Linear Bottlenecks: Mobile Networks for Classification, Detection and Segmentation Mark Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.

[25] Lee S H, Chan C S, Mayo S J and Remagnino P 2017 How deep learning extracts and learns leaf features for plant classification Pattern Recognit. 71 1–13

[26] Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E and Darrell T 2014 DeCAF: A deep convolutional activation feature for generic visual recognition 31st Int. Conf. Mach. Learn. ICML 2014 2 988–96

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

[28] Jasitha P, Dlleep M R and Dlvya M 2019 Venation Based Plant Leaves Classification Using GoogLeNet and VGG 2019 4th IEEE Int. Conf. Recent Trends Electron. Information, Commun. Technol. RTEICT 2019 - Proc. 715–9

[29] Ibrahim Z, Sabri N, Nabilah N and Mangshor A 2018 Leaf Recognition using Texture Features for Herbal Plant Identification Indones. J. Electr. Eng. Comput. Sci. 9 152–6