Deep Learning for Plant Species Classification Using Leaf Vein Morphometric

Jing Wei Tan, Siow-Wee Chang, Sameem Abdul-Kareem, Hwa Jen Yap, and Kien-Thai Yong

Abstract—An automated plant species identification system could help botanists and layman in identifying plant species rapidly. Deep learning is robust for feature extraction as it is superior in providing deeper information of images. In this research, a new CNN-based method named D-Leaf was proposed. The leaf images were pre-processed and the features were extracted by using three different Convolutional Neural Network (CNN) models namely pre-trained AlexNet, fine-tuned AlexNet, and D-Leaf. These features were then classified by using five machine learning techniques, namely, Support Vector Machine (SVM), Artificial Neural Network (ANN), k-Nearest-Neighbor (k-NN), Naïve-Bayes (NB), and CNN. A conventional morphometric method computed the morphological measurements based on the Sobel segmented veins was employed for benchmarking purposes. The D-Leaf model achieved a comparable testing accuracy of 94.88 percent as compared to AlexNet (93.26 percent) and fine-tuned AlexNet (95.54 percent) models. In addition, CNN models performed better than the traditional morphometric measurements (66.55 percent). The features extracted from the CNN are found to be fitted well with the ANN classifier. D-Leaf can be an effective automated system for plant species identification as shown by the experimental results.

Index Terms—Tropical tree, deep learning, convolutional network, leaf vein morphometric, feature extraction, classification, artificial neural network

1 Introduction

The number of plant species are extremely huge, with about 391,000 vascular plant species all over the world [1]. Hence, it is impossible and not practical for a botanist or an expert, to be able to identify and classify all the species. In addition, some plant species may have high similarity between each other, taking a long time to differentiate them. In addition, many plants face the problem of extinction. Endangered and non-endangered plant species need to be preserved and conserved in a proper way to reduce the risks of extinction. Hence, there is a need to develop an automated or computerized system to identify and classify plants. Leaf shape is the most commonly used feature used to develop such automated plant classification systems. Other than shape, the leaf can provide additional information such as textures, veins, and colours.

- *J. W. Tan and S.-W. Chang are with the Bioinformatics Programme, Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia. E-mail: jingwei_92@siswa.um.edu.my, siowwee@um.edu.my.*
- *S. Abdul-Kareem is with the Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia. E-mail: sameem@um.edu.my.*
- *H. J. Yap is with the Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia. E-mail: hjyap737@um.edu.my.*
- *K.-T. Yong is with the Ecology and Biodiversity Programme, Institute of Biological Sciences, Faculty of Science, University of Malaya 50603, Kuala Lumpur, Malaysia. E-mail: yongkt@um.edu.my.*

Manuscript received 13 Sept. 2017; revised 6 June 2018; accepted 9 June 2018. Date of publication 19 June 2018; date of current version 4 Feb. 2020. (Corresponding author: Siow-Wee Chang.) Digital Object Identifier no. 10.1109/TCBB.2018.2848653

With the advancement of science and technology, machine learning has been widely employed for classification and recognition tasks in many domains especially in the biological fields. Machine learning techniques, such as, the Artificial Neural Network, Support Vector Machines, k-Nearest Neighbour, and others are artificial intelligent techniques mainly employed to perform pattern recognition.

Currently, deep learning, a subfield of artificial intelligence (AI), is a popular and widely used technique, that has been applied in various domains including biology, medical, computer vision, speech recognition and others [2], [3], [4], [5]. Deep learning is a modern AI approach, which contributes a robust framework towards supervised learning [6]. It is able to map an input vector rapidly and efficiently to an output vector even in a large dataset [6]. Deep learning architecture can be further divided into Convolutional Neural Network (CNN), Deep Belief Network (DBN) and so on. Deep learning is able to extract more detailed information as compared to the conventional machine learning techniques.

In this research, CNN is applied to extract the features from leaf images of selected tree species. Three different CNN models were used, namely, the pre-trained AlexNet CNN model, fine-tuned pre-trained AlexNet CNN model and the proposed D-Leaf CNN model. The extracted features were then fed into a few classification approaches for learning and training purposes. Five classifiers were employed in this research which are CNN, Support Vector Machine (SVM), Artificial Neural Network (ANN), k-Nearest Neighbour (k-NN) and Naïve Bayes (NB). A conventional method, which segmented the leaf veins by using Sobel edge detection technique and performed vein morphological measurements, was used for benchmarking. Based on the literature review, this is one of the first few

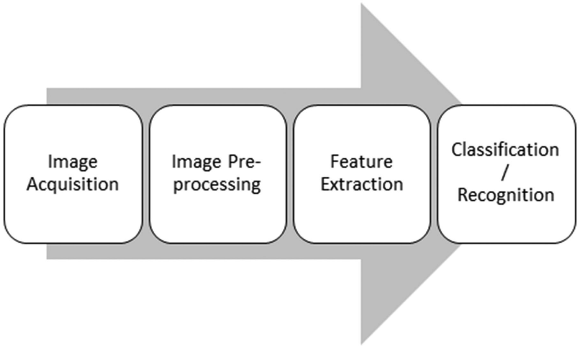

Fig. 1. General approach for automated plant classification.

studies, which have applied CNN in tropical tree species classification, by using both leaf morphometric and venation pattern approaches.

2 Background Study

Fig. 1 shows the fundamental steps of an automated plant classification system. Initially, the leaf images would be acquired using digital camera, scanner or some other equipments. The images were then pre-processed to remove noise and improve the quality. Noise occurs as pixel values which do not represent the true intensities of an image during the image acquisition. Image enhancement is a process that is used to emphasise the features of an image [7]. It is a necessary step to remove the image noises in order to highlight or enhance the important features of an image. Subsequently, the region of interest (ROI) was segmented from the images, followed by feature extraction. Finally, the extracted features are fed to classification or recognition system.

Leaves are commonly used in plant species recognition due to their availability throughout the year, especially, in the tropical areas. Many useful features can be acquired from a single leaf; such as, shape, texture, venation pattern, and colour. Each of these features could be extracted by different approaches; either through traditional morphometric measurement or machine learning methods. However, some approaches can be applied to obtain more than one feature, for example, Zernike Moment can be applied to obtain shape and texture features.

Shape is the most common feature that have been used to develop plant identification systems. The features that are commonly used to evaluate shape are, ratio between slimness, roundness, compactness, rectangularity, and other aspects. Texture is one of the important features of the plant identification system, which can be used to characterize the leaves based on the surface structure of the leaves. It is a non-consistent spatial distribution pattern of different image intensities [8], [9], which concentrates mainly on each single pixel of an image. Venation pattern is another important feature for plant identification. Leaf veins can be further categorized into three types, which are primary veins, secondary veins and tertiary veins. Identification of plant species can also be done by using the colour of the leaves since leaf colour varies from species to species. Leaves from the same species may also be of different colours. For example, the young leaves of *Cinnamomum inners* is red in color which turn to green when they are matured.

2.1 Conventional Approach

Cope et al. [10] introduced an evolved vein classifier based on genetic algorithms (GA) and Ant Colony algorithms to extract the vein structure. The evolved vein classifier is capable of extracting nearly the whole primary and secondary vein patterns with only a little noise. It was superior than the Ant Colony algorithms in extracting venation structures that are highly discontinuous. On the other hand, the Ant Colony algorithms, are more potent in continuous venation extraction, however, it may develop bigger and connected noise areas that may cause difficulty in actual vein identification. In short, the evolved classifiers outperformed the ant algorithms. Hence, a combination of both methods may be deemed to be more reliable.

A combination of colour and texture features based plant identification system had been proposed by Anami et al. [11]. The Sobel operator was used to extract the colour histogram and edge direction histogram as colour and texture features respectively from 1000 images of different types of herbs, shrubs and trees. The extracted features were then trained by using a radial basis exact fit neural network (RBENN) and a SVM.

Larese et al. [12] constructed an automated leaf identification approach for legumes based on the vein architecture only. Simple measurements were applied on the vein morphology and then identified by a Random Forests approach. 39 vein features were extracted and then classified using the Random Forests. Furthermore, the authors discovered that the performance of using 7 features subset is comparable with the performance of 39 features.

Kadir et al. [13] proposed another method on the Foliage dataset and the Flavia dataset. The shape features, represented by 3 geometric features and the Polar Fourier Transform (PFT), colour moments, texture features extracted using GLCM and vein features were used to develop several models using different combination of features.

2.2 CNN-Based Plant Identification

Lee et al. [15], [17] proposed a CNN technique to identify 44 plant species acquired from the Royal Botanic Gardens of Kew, England. These studies employed a pre-trained CNN model for feature extraction and deconvolutional network (DL) for unique features filtration and image visualisation on the extracted features. The extracted features were then classified with a Multilayer Perceptron (MLP) and a SVM. Two different datasets were used, namely, whole image (D1) and leaf patches (D2). Both datasets achieved an accuracy of more than 97 percent. Furthermore, researchers had combined both local and global features together in [17] and achieved more than 91 percent accuracy.

Furthermore, CNN was employed by Sladojevic et al. [18] for plant diseases recognition. This study was tested on 13 classes of plant by using an open source pre-trained network model known as CaffeNet model which consisted of a set of weights trained by using ImageNet. This model consists of 5 convolutional layers and 3 fully-connected layers out of 8 learning layers. The results had been improved from 95.8 percent (before fine-tuning) of accuracy to 96.3 percent of accuracy (after fine-tuning) with 100 training iterations. The best Top-1 and Top-5 success in this study was 96.3 to 99.9 percent.

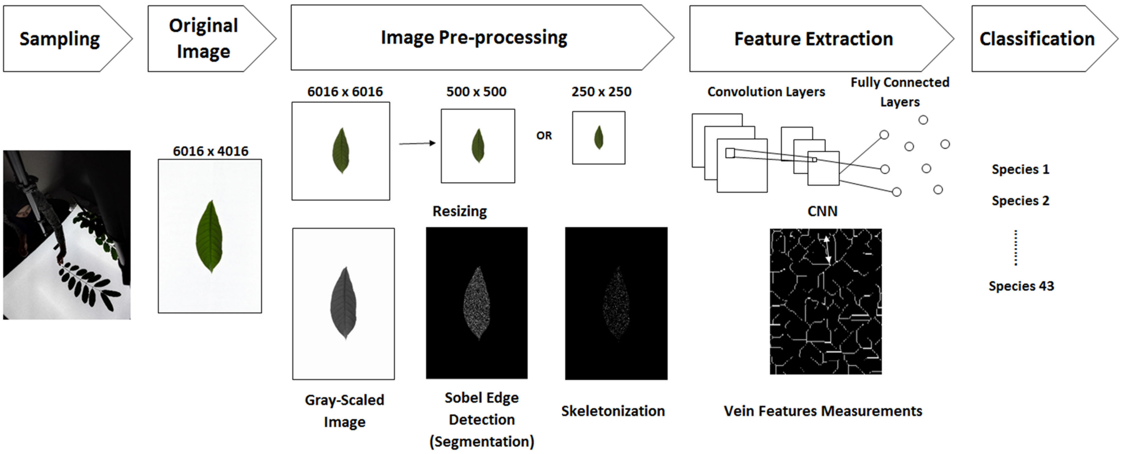

Fig. 2. Proposed architecture of the research.

Another study based on leaf vein morphological patterns and using deep learning technique for plant identification was proposed in Grinblat et al. [19]. The authors trained the CNN models with different number of layers which ranged from 2 layers (1 convolutional layer +1 Softmax layer) to 6 layers (5 convolutional layers +1 Softmax layer). The 5-layers CNN model which combined veins with three different scale factors (100, 80 and 60 percents) performed the best with an average accuracy of 96.9 percent.

3 Methodology

This research consisted of four main steps as shown in Fig. 2, which are, sampling, image pre-processing, feature extraction and classification. First, the leaf samples were collected and images were acquired. The leaf images were then pre-processed and fed into the feature extraction step to retrieve the important information from the leaves using CNN and Sobel edge detection approach. Lastly, the extracted features were trained and classified by using various machine learning methods.

3.1 Sampling

The leaf samples of this research were collected from three locations in the University of Malaya, Kuala Lumpur, Malaysia. These locations were areas around Varsity Lake (VL), main library (ML) and Dewan Tunku Canselor hall (DTC). The leaf images of this dataset were collected from the common tropical trees, which can be found easily in the University of Malaya as well as in Malaysia. The leaf is the part of the tree that is always chosen as the samples to be studied, instead of fruits, flowers or some other parts due to its availability throughout the year. 43 species of tropical trees with 30 samples per species were collected. Thus, a total of 1290 leaf images were collected for this research. The selected 43 plant species with their scientific name, common name and location are listed in Appendix I, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TCBB.2018.2848653.

The leaf samples for each species are shown as in Appendix II (a), available in the online supplemental material.

From Appendix II (a), available in the online supplemental material, we can see that some of the leaf samples in this dataset possessed a high similarity in shape and colour.

The leaf images were acquired by using a Nikon D750 model of DSLR camera. The samples were put on a box with white background and fluorescent lights were placed under the box as shown in Appendix II (b), available in the online supplemental material in order to capture good quality images with standard background. As the lighting of this setup is from the bottom, it can help to reduce glaze and shadow on the leaf.

3.2 Image Pre-Processing

Raw images are not appropriate for analysis purposes and need to be converted into the processed format, such as, jpeg, jpg and tiff for further analysis. In this research, the acquired images were stored in a format of Nikon camera, named Nikon Electronic File (NEF) with 6016 x 4016 resolution. Adobe Photoshop was used to convert these raw images into Tagged Image File Format (TIFF). In addition, the background noises of the images were reduced by using Adobe Photoshop and MATLAB R2016a was used for resizing and image conversion. Two different pre-processing methods were applied, namely, image reconstruction for CNN and vein morphometric measurements (segmented by Sobel).

3.2.1 *Image Reconstruction for CNN*

The leaf images were reconstructed into square dimension (m x m), which was required as the inputs of our selected CNN models. The original images with 6016×4016 resolution were added with 1000 padding at both top and bottom width into 6016×6016 resolution before resizing in order to maintain the ratio of the leaf shape. Then, the images were resized into 250×250 resolution to reduce the computational time. The images were then retained in the RGB format.

3.2.2 *Edge Detection Method - Sobel*

Sobel is an edge detection method, which is simple and effective in detecting the edge of an object [22], [23]. Sobel was used as a segmented method to extract the vein architectures of all the leaf samples.

First, all images were converted from RGB images (Fig. 3a) into grey-scale images (Fig. 3b). Then, Sobel was used to segment out the region of interest (ROI) from the images (Fig. 3c). After segmentation, the images were then post-processed and skeletonized to ensure a clean vein architecture could be obtained (Fig. 3d).

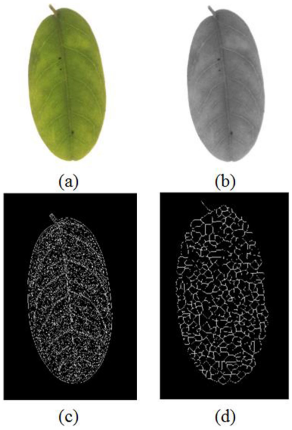

Fig. 3. (a) RGB image. (b) Gray-scaled image. (c) Segmented image. (d) Skeletonized image.

3.3 Software and Hardware

Workstation with Intel Xeon CPU E5-1603 v3 @ 2.80 GHz processor, 32 GB of RAM and Nvidia Quadro K2200 4 GB was used in the development and execution of this research. MATLAB version 2016a and 2017a were used in feature extraction and classification.

3.4 Feature Extraction

Feature extraction is the key stage in this research, which is used to extract the important features from the leaf images.

The features that are commonly used for plant identification systems are shape [24], [25], texture [9], [24], [25], [26], colour [9], [24], [25] and vein [8], [24], [25], [27]. Numerous methods can be employed for feature extraction such as Histogram of Oriented Gradient (HOG), Zernike Moments, Hu's Moment and others. A deep learning algorithm, namely, the Convolutional Neural network, was employed to extract the features in this research. A conventional approach of using morphological feature extraction to segment the leaf venation by using Sobel edge detection method was investigated and compared.

3.4.1 *Deep Learning—Convolutional Neural Networks (CNNs)*

CNN, one of the deep learning methods—is proposed for feature extraction in this study. CNN is considered as another type of multilayer perceptron since it used more than one layer of perceptron to learn the features of an image. A typical CNN layer composed of three layers [18], [19] as shown in Fig. 4. The first layer is the convolution stage which performs convolution operation on an input image by using filters and kernels to construct feature maps. Whilst, in the second layer, non-linear activation function such as Rectified Linear Unit (ReLU) is implemented into all the feature maps. This function would remain to linear closely, in the term of piecewise linear function. Thus, this will preserve many of the properties that optimize and generalize the linear models with gradient based methods easily [6]. Lastly, the third layer is the pooling stage which employs a function to simplify and summarize the information of the particular layer.

Pre-trained AlexNet Model: A pre-trained CNN AlexNet model was used in this research. AlexNet was trained with 1000 different classes of images from ImageNet [2], [26]. AlexNet is made up of nine layers which include five convolution layers, three fully connected layers, and a softmax classification layer as illustrated in Fig. 4. Softmax classification layer is referred to the softmax function which yield the predicted probability of each group and is fully-connected to the final full connected layer.

In AlexNet, the input image size was resized into 227 $\times$ 227 $\times$ 3. The parameters of each AlexNet layer are as shown in Table 1. Three fully-connected layers with 4096, 4096 and 1000 neurons respectively were included in this model. This is followed by the softmax classification layer.

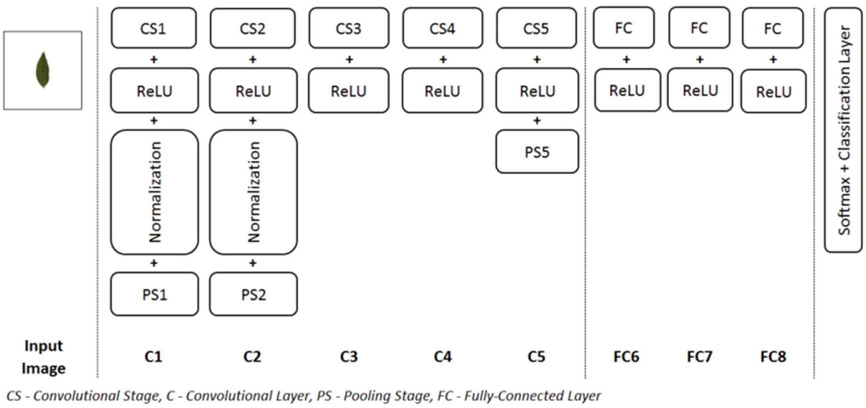

Fig. 4. AlexNet Architecture. CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, and FC - Fully-Connected Layer

TABLE 1
Parameters for AlexNet Layers

CNN Layer	CS1	PS1	CS2	PS2	CS3	CS4	CS5	PS5	FC6	FC7	FC8
Filter Size	11 x 11	3 x 3	5 x 5	3 x 3	3 x 3	3 x 3	3 x 3	3 x 3	—	—	—
No. of Kernel	96	—	256	—	384	384	256	—	4096	4096	1000
Size of Stride*	[4 4]	[2 2]	[1 1]	[2 2]	[1 1]	[1 1]	[1 1]	[2 2]	—	—	—

*— Not Applicable, * Distance between the receptive field centers of neighboring neurons [2].*

TABLE 2
Parameters for Fine-Tuned Alexnet Layers

CNN Layer	CS1	PS1	CS2	PS2	CS3	CS4	CS5	PS5	FC6	FC7	FC8
Filter Size	7 x 7	3 x 3	5 x 5	3 x 3	3 x 3	3 x 3	3 x 3	3 x 3	—	—	—
No. of Kernel	96	—	256	—	384	384	256	—	1290	1290	43
Size of Stride*	[2 2]	[2 2]	[1 1]	[2 2]	[1 1]	[1 1]	[1 1]	[2 2]	—	—	—

*— Not Applicable, * Distance between the receptive field centers of neighboring neurons [2].*

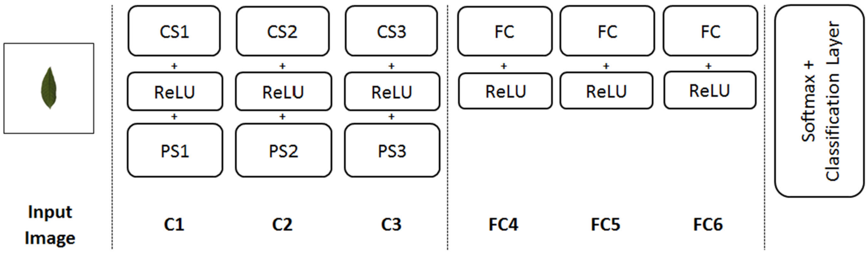

Fig. 5. D-Leaf Architecture. CS - Convolutional Stage, C - Convolutional Layer, PS - Pooling Stage, and FC - Fully-Connected Layer

TABLE 3
Parameters for D-Leaf Layers

CNN Layer	CS1	PS1	CS2	PS2	CS3	PS3	FC4	FC5	FC6
Filter Size	11 x 11	2 x 2	5 x 5	2 x 2	4 x 4	2 x 2	—	—	—
No. of Kernel	64	—	96	—	256	—	1290	1290	43
Size of Stride*	[4 4]	[2 2]	[2 2]	[2 2]	[1 1]	[2 2]	—	—	—

*— Not Applicable, * Distance between the receptive field centers of neighboring neurons [2].*

Fine-tuned AlexNet Model: The architecture of the fine-tuned AlexNet model is the same as the original AlexNet model as shown in Fig. 4. Some parameters in the AlexNet model were fine-tuned as shown in Table 2. The input size of these fine-tuned models was set in $227 \times 227 \times 3$, same as in the AlexNet models. The first convolutional stage was fine-tuned with 7 x 7 of filter size and [2 2] of stride. Besides that, the neuron number of the three fully connected layers was fine-tuned into 1290, 1290 and 43. The other parameters of this model were remained unchanged as in the AlexNet model.

Proposed Model—D-Leaf: In this research, a CNN model (Fig. 5) was developed to extract the features from the images instead of fine-tuning the AlexNet model. As show in Fig. 5, a total 6 layers of CNN model with three convolution layers ($\mathrm{Convolutionstage + ReLU + Poolingstage}$), three fully connected layers, and a softmax classification layer was developed using MATLAB.

Table 3 shows the parameters of each layer in the D-Leaf model. Three fully connected layers were employed in this model. The output of C3 was then passed into the first fully connected layer (FC4) and second fully connected layer (FC5) with 1290 neurons in each of them. The third fully connected layer (FC6) consisted of 43 neurons which denoted the number of classes of the dataset. The model ended with a softmax classification layer. The training options of the D-Leaf model were based on the Stochastic Gradient Descent with Momentum algorithm with a batch size of 100. Mini-batch is a training subset which measures the loss function gradient and update the weights [31].

3.4.2 Vein Morphometric Measurements

The vein features were extracted from the segmented images (with Sobel) by measuring the vein morphological features. 62 vein features were extracted, such as, number of branching points, number of ending points, number of branches, number of areoles, and others as listed in Appendix III, available in the online supplemental material. The leaf area was computed for the density calculation of veins, branching points, ending points and areoles.

3.5 Classifiers

Classification, as the last step for an automated plant recognition system, is an intelligent algorithm in training data to recognize the specific features of each individual plant species and categorizing a new sample as the correct species. The favoured machine learning methods for plant identification are Artificial Neural Network (ANN), Support Vector Machines (SVM) and k-Nearest Neighbour (k-NN). Five classification methods that was used in this research were SVM, ANN, k-NN, Naïve Bayes (NB) and CNN.

Support Vector Machines (SVM), a supervised machine learning approach, is conceded as one of the powerful classification methods due to its high capability in dealing with high dimensional space and data points which are not linearly separated [31]. Applying linear SVM on feature-mapped data can execute speedy with low storage and improve the classification performance [14]. Linear SVM with 'One versus all' scheme was employed in this research since it involves a multi-classes dataset.

Artificial Neural Network (ANN) is the most common and frequent algorithm that have been applied in numerous

TABLE 4
Results of Alexnet and Fined-Tuned Alexnet Models

Classifiers	Accuracy#	
	AlexNet	Fine-tuned AlexNet
SVM	79.40	87.79
ANN	**93.26**	**95.54**
k-NN	85.60	87.33
NB	83.33	87.33
CNN	—	88.30

#*Testing accuracy, — Not Applicable.*

TABLE 6
Classification Accuracy
of Morphometric Measurements

No. of Neuron	Accuracy#
20	63.81
40	64.34
60	64.22
80	**66.28**
100	66.55

#*Testing accuracy.*

researches [3], [13], [15], [17], [32]. The development of ANN was inspired by the human's brain neurons structure and behaviours. Different number of neurons and hidden layers in an ANN model may affect the classification performance. The proposed research employed a feed forward neural network with a single hidden layer which consisted of 80 neurons. The training process used the default Scaled Conjugate Gradient function and the achievement of a minimum gradient as the stopping criteria.

A k-NN is a classification approach, which classifies a sample according to the majority vote of its neighbours [33]. The number of neighbours in this research is fixed as 1 with the city block distance metric.

Bayesian classifier is a statistical classifier which attempts to make prediction of the class with an unknown sample based on probability [31]. NB assumes that all features of the samples are unrelated to each other [34]. However, the conditional independence of Bayes theorem diminishes the classification accuracy.

The architecture of CNN classifiers used for fine-tuned AlexNet model were the same as shown in Fig. 4 with parameters shown in Table 2. Meanwhile, the CNN classifier employed by the D-Leaf model was the same as shown in Fig. 5 and Table 3.

4 Results

4.1 AlexNet and Fine-Tuned AlexNet

For performance evaluation, the data was partitioned into training set and testing set with a ratio of 80:20. Each model was executed over feature extraction once and classification for 10 times to get an optimum performance. A pre-trained model—AlexNet was applied for feature extraction from FC7 layer followed by four classifiers. A fine-tuned AlexNet was implemented for feature extraction and the extracted features were trained and tested with five classifiers. The performance of each classifier is as shown in Table 4.

TABLE 5
Classification Accuracy of Different
Classifiers with D-Leaf Features

Classifiers	Accuracy#
SVM	82.75
ANN	**94.88**
k-NN	82.44
NB	81.86
CNN	79.03

#*Testing accuracy.*

As can be seen in Table 4, for AlexNet models, ANN outperformed the other models with an accuracy of 93.26 percent while all the other classifiers obtained an accuracy that less than 86 percent. However, the worst model is the SVM classifier with features extracted by AlexNet. This model obtained only 79.40 percent of accuracy.

Whereas, for fine-tuned AlexNet models, ANN classifier with the fine-tuned AlexNet features achieved the best performance of 95.54 percent for accuracy. However, the other classifiers obtained accuracy of less than 89 percent. The fine-tuned AlexNet model required additional training time to retrain the model before it can be used for feature extraction. However, the AlexNet model can be used directly for feature extraction without retrain the model.

In term of performance, fine-tuned AlexNet models performed better than AlexNet models.

4.2 D-Leaf

In the experiments related to the proposed D-Leaf model, an image input size of 250×250 was used with the CNN and the performances are compared. The extracted features from FC5 layer of D-Leaf were then classified by using five (5) classifiers, which were, SVM, ANN, k-NN, NB and CNN as shown in Table 5.

As shown in Table 5, all models obtained more than 79 percent of accuracy with the ANN model achieving the best performance of 94.88 percent of testing accuracy. On the other hand, CNN classifier had the worst performance of 79.03 percent. Note that we did not check to see if the differences in the performances of the classifiers were statistically significant.

5 Benchmarking

The performance of the D-Leaf method was then benchmarked with a conventional method and cross validation (CV) methods.

5.1 Sobel Edge Detection

A conventional method, which employed Sobel edge detection method for vein segmentation and then computed the vein morphological features was used as a benchmark. The extracted features from the Sobel method were classified using ANN and various total number of neurons in order to examine the performance of the morphometric measurements. This is shown in Table 6.

From Table 6, we can see that the morphometric measurements performed in the range of 63 to 67 percents. The best performance was achieved by the models with 80 and

TABLE 7
Validation Results with Cross-Validation

Data Partition	Accuracy#
5-CV	93.15
10-CV	93.31

#*Testing accuracy.*

100 neurons with a testing accuracy of about 66 percent. The accuracy seems to become stable after 80 neurons.

Finally, the performance of the D-Leaf method was validated by using cross-validation (CV) and benchmarked against other datasets.

5.2 Cross-Validation (CV)

First, cross-validation (CV) was conducted in order to reduce overfitting problems since the proposed dataset consisted of 30 samples per species only. 5-fold CV and 10-fold CV were investigated with the D-Leaf method. The performance of CV was show in Table 7.

From Table 7, we can see that the performance of the D-Leaf method (without CV: 94.88 percent) is relatively similar with that of both the 5-fold CV (93.15 percent) and the 10-fold CV (93.31 percent).

6 Validation

The performance of the D-Leaf method was validated on three publicly available datasets, which are, MalayaKew, Flavia and Swedish Leaf Dataset. MalayaKew dataset was collected from Royal Botanic Gardens, Kew, England with 44 classes of plant species [35]. The Flavia dataset is made up of 32 plant species from Yangtze Delta, China and 50 samples for each species [36], while the Swedish leaf dataset consists of 15 tree classes with 75 samples per species [37].

As shown in Table 8, an accuracy of 94.63 percent was achieved in the Flavia dataset while the Swedish dataset achieved the accuracy of up to 98.09 percent. This may be due to the lower number of species in the Swedish dataset. However, the MalayaKew dataset performed the least favorably at 90.38 percent accuracy. This may due to the black colour background of the leaf images provided in the MalayaKew dataset [17].

TABLE 8
Validation Results with Other Datasets

Dataset	Accuracy#
MalayaKew [35]	90.38
Flavia [36]	94.63
Swedish [37]	98.09

#*Testing accuracy.*

7 Discussion

The first part of this research was evaluated by employing a pre-trained CNN AlexNet and fine-tuned AlexNet model for feature extraction followed by classification process. Next, a new CNN model—D-Leaf was proposed. A comparison between AlexNet and D-Leaf models, showed that AlexNet is definitely more complicated than D-Leaf. A system with more convolutional layers needs a longer execution time.

In terms of performance, the testing accuracy achieved by the D-leaf (94.88 percent) is comparable with the fine-tuned AlexNet model (95.54 percent). However, AlexNet cannot be employed as a classifier since the last fully-connected layer is set with 1000 neurons as shown in Table 1 which is not compatible with the total number of the plant species—43 species. Thus, some fine-tuning processes are performed on the AlexNet model in order to use it as a classifier. D-Leaf can be an efficient automated plant species classification tool since it is able to achieve a comparable performance with the fine-tuned AlexNet model.

Generally, the ANN models achieved the best performance due to the stability of the ANN approach. In this research, we can claim that ANN is highly suitable with the CNN feature extraction model. In contrast, SVM, k-NN and NB were less suitable with the CNN. This may be due to their same basic concept – the underlying architecture of the CNN is the neural network. Additionally, three species were found to be misclassified frequently, namely, Species 19 (*Erythrina variegate*), Species 31 (*Plumeria rubra*) and Species 34 (*Saraca thaipingensis*). Yet, the misclassification of these species only occurred in the SVM, k-NN and NB. This proved that the features extracted using the CNN were more suitable with the ANN classifier. However, the performance of the CNN as a classifier was less favorable as compared to ANN as shown in Tables 4 and 5. In short, CNN would be preferable as a feature extraction approach rather than a classifier in this research.

From the results shown in Table 6, we can see that the best architecture for the ANN for leaf identification by using conventional Morphometric measurements is that with 80 neurons as there is no significant increase in performance after 80 neurons. Morphometric measurements' performance is less desirable in this research as compared to CNN models. During these conventional processes, some loss of important features may occur. Nevertheless, CNN could extract the leaf features directly from the RGB images without any segmentation or enhancement. Generally, the CNN approach would extract and accumulate the features from layer to layer and extract all the common features from a leaf such as shape, vein, and colour at one time and compact all these features together. On the other hand, a conventional feature extraction method extracts each type of feature separately and manually consuming a lot of time. For example, if the shape features are considered in this research, different sets of processes will be required for segmentation followed by shape feature extraction. Thus, deep learning approach is more practical and appropriate than conventional methods for developing an automated plant species classification system.

8 Conclusion

In this work, we can conclude that the CNN is better than conventional morphological methods for feature extraction in plant species. More pre-processing works need to be done when using the conventional methods as compared to the CNN. Besides that, CNN is found to be a favourable feature extraction method rather than a classification method in this research. The ANN classifier together with the CNN feature extractor obtained the most optimal result as compared to other classifiers. In fact, the best result in this research is, when using the D-Leaf for feature extraction and the ANN as a classifier, achieving a testing accuracy of 94.88 percent.

Additionally, the performance of the D-Leaf method is validated with CV methods as well as MalayaKew, Flavia and Swedish datasets. The validation performance (>93 percent) proved that the D-Leaf method can be used for automated plant species classification. As a future work of this research, we would like to include more tropical plant species, especially those with compound leaves, and add in more taxonomic features to further enhance the plant species identification process that can be used by botanists and the public.

ACKNOWLEDGMENTS

This research was supported by the University of Malaya Research Grant (UMRG) with the project number of RP038C-15AET. The funder had no role in study design, collection and analysis of data, publication decision, or the manuscript preparation.

REFERENCES

[1] K. J. Willis, et al., "State of the World's plants 2017," *Report*, Royal Botanic Gardens, Kew 2017.
[2] A. Krizhevsky, I. Sutskever, and G. E Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[3] I. Goodfellow, Y. Bengio, and A. Courville, "Introduction," *Deep Learning*. Cambridge, MA: MIT Press, vol. 1, pp. 1–12, 2015.
[4] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, "Deep learning for identifying metastatic breast cancer," *Int. Symp. Bio. Imaging's Grand Challenge*, arXiv:1606.05718, 2016.
[5] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82—97, Nov. 2012.
[6] I. Goodfellow, Y. Bengio, and A. Courville *Deep Learning*. Cambridge, MA, USA: MIT Press. 2016
[7] R. Malladi, and J. A Sethian, "A unified approach to noise removal, image enhancement, and shape recovery," *IEEE Trans. Image Process.*, vol. 5, no. 11, pp. 1554–1568, Nov. 1996.
[8] V. Metre and J. Ghorpade, "An overview of the research on texture based plant leaf classification," *Int. J. Comput. Sci. Net.*, vol. 2, 2013.
[9] C. Arun, W. S. Emmanuel, and D. C Durairaj, "Texture feature extraction for identification of medicinal plants and comparison of different classifiers," *Int. J. Comput. Appl.*, vol. 62, no. 12, 2013.
[10] J. S. Cope, P. Remagnino, S. Barman, and P Wilkin, "The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms," in *Int. Conf. Adv. Concepts Intell. Vis. Syst.*, Dec. 2010, pp. 135–144.
[11] B. S. Anami, S. N. Suvarna, and A Govardhan, "A combined color, texture and edge features based approach for identification and classification of indian medicinal plants," *Int. J. Comput. Appl.*, vol. 6, no. 12, pp. 45–51, 2010.
[12] M. Larese, R. Craviotto, M. Arango, C. Gallo, and P. Granitto, "Legume identification by leaf vein images classification," in *Proc. 17th Iberoamerican Congr. Progress Pattern Recognit. Image Anal. Comput. Vis. Appl.*, 2012, pp. 447–454.
[13] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Neural network application on foliage plant identification," *Int. J. Comput. Appl.*, vol. 29, no. 9, pp. 15–22, 2013.
[14] M. Sulc, and J. Matas, "Texture-based leaf identification," *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 185–200.
[15] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, "Deep-plant: Plant identification with convolutional neural networks," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 452–456.
[16] S. H. Lee, Y. L. Chang, C. S. Chan, and P. Remagnino, "Plant identification system based on a convolutional neural network for the lifeclef 2016 plant classification task," *CLEF Working Notes*, pp. 502–510, 2016.
[17] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, "How deep learning extracts and learns leaf features for plant classification," *Pattern Recognit.*, vol. 71, pp. 1–13, 2017.
[18] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Comput. Intell. Neurosci.*, vol. 2016, 2016, Art. no. 6.
[19] G. L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, "Deep learning for plant identification using vein morphological patterns," *Comput. Electron. Agriculture*, vol. 127, pp. 418–424, 2016.
[20] M. G. Larese, R. Namías, R. M. Craviotto, M. R. Arango, C. Gallo, and P. M. Granitto, "Automatic classification of legumes using leaf vein image features," *Pattern Recognit.*, vol. 47, no. 1, pp. 158–168, 2014.
[21] A. K. Reyes, J. C. Caicedo, and J. E. Camargo, "Fine-tuning deep convolutional networks for plant recognition," *CEUR Workshop Proc.*, vol. 1391, 2015.
[22] S. Gupta and S. G. Mazumdar, "Sobel edge detection algorithm," *Int. J. Comput. Sci. Manage. Res.*, vol. 2, no. 2, pp. 1578–1583, 2013.
[23] O. R. Vincent, and O. Folorunso, "A descriptive algorithm for sobel image edge detection," in *Proc. Inf. Sci. IT Edu. Conf.*, vol. 40, pp. 97–107, Jun. 2009.
[24] A. R. Yadav, M. Dewal, R. Anand, and S. Gupta, "Classification of hardwood species," in *Proc. 4th Nat. Conf. Comput. Vis. Pattern Recognit. Image Process. Graph.*, Dec. 2013, pp. 1–5.
[25] D. S. Prasvita, and Y. Herdiyeni, "MedLeaf: Mobile application for medicinal plant identification based on leaf image," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 3, no. 2, pp. 103–106, 2013.
[26] A. Ehsanirad, and Y. Sharath Kumar, "Leaf recognition for plant classification using GLCM and PCA methods," *Oriental J. Comput. Sci. Technol.*, vol. 3, no. 1, pp. 31–36, 2010.
[27] A. Aakif, and M. F. Khan, "Automatic classification of plants based on their leaves," *Biosystems Eng.*, vol. 139, pp. 66–75 (2015).
[28] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional neural networks for MATLAB," in *Proc, 23rd ACM Int. Conf. Multimedia*, 2015, pp. 689–692. http://dx.doi.org/10.1145/2733373.2807412.26
[29] DeepDreamImage, (2017). [Online]. Available: https://www.mathworks.com/help/nnet/ref/alexnet.html Retrieved, Mar. 10, 2017.
[30] Convolution2dLayer, (2017). [Online]. Available: https://www.mathworks.com/help/nnet/ref/trainingoptions.html, Retrieved, Mar. 10, 2017.
[31] A. Caglayan, O. Guclu, and A. B. Can, "A plant recognition approach using shape and color features in leaf images," in *Proc. Int. Conf. Image Anal. Process.*, Sep. 2013, pp. 161–170.
[32] S. W. Chang, S. A. Kareem, T. G. Kallarakkal, A. F. M. A. Merican, M. T. Abraham, and R. B. Zain, "Feature selection methods for optimizing clinicopathologic input variables in oral cancer prognosis," *Asian Pac. J. Cancer Prev.*, vol. 12, pp. 2659–2664, 2011.
[33] C. Pornpanomchai, S. Rimdusit, P. Tanasap, and C. Chaiyod, "Thai herb leaf image recognition system (THLIRS)," *Kasetsart J. (Nat. Sci.)*, vol. 45, pp. 551–562, 2011.
[34] A. McCallum, and K. Nigam, "A comparison of event models for naive bayes text classification," in *Proc. AAAI-98 Workshop Learn. Text Categorization*, vol. 752, pp. 41–48, Jul. 1998.
[35] C. S Chan, "MalayaKew dataset ground truth plant classification dataset," (2015). [Online]. Available: http://web.fsktm.um.edu.my/~cschan/downloads_MKLeaf_dataset.html, Retrieved Jul. 09, 2017
[36] Flavia., 2009. [Online]. Available: http://flavia.sourceforge.net/, Retrieved Apr. 27, 2017
[37] Swedish Leaf Dataset, 2016 Mar., 2008. [Online]. Available: http://www.cvl.isy.liu.se/en/research/datasets/swedish-leaf/, Retrieved Apr. 10, 2017.
[38] J. X. Du, X. F. Wang, and G. J. Zhang, "Leaf shape based plant species recognition," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 883–893, 2007.
[39] X. F. Wang, J. X. Du, and G. J. Zhang, "Recognition of leaf images based on shape features using a hypersphere classifier," in *Proc. Int. Conf. Intell. Comput.*, Aug. 2005, pp. 87–96.
[40] M. A. Hedjazi, I. Kourbane, and Y. Genc, "On identifying leaves: A comparison of CNN with classical ML methods," in *25th Signal Process. Commun. Appl. Conf.*, May 2017, pp. 1–4.

Jing Wei Tan received the BSc degree in bioinformatics form the University of Malaya, Kuala Lumpur, Malaysia, in 2016. She is currently working toward the MSc degree in bioinformatics and is a research assistant with the Institute of Biological Sciences, Faculty of Science, University of Malaya. Her research interests include machine learning and deep learning, image processing, pattern analysis, and bioinformatics.

Siow-Wee Chang received the PhD degree in computer science from the University of Malaya, Malaysia, in 2013. She is currently a senior lecturer with the Bioinformatics Programme, the Institute of Biological Sciences, Faculty of Science, University of Malaya, Malaysia. Her current research interests include machine learning, data analytics, bioinformatics, computational biology, and human computer interaction.

Sameem Abdul-Kareem received the BSc degree in mathematics (Hons) from the University of Malaya, in 1986, the MSc degree in computing from the University of Wales, Cardiff, UK, in 1992, and the PhD degree in computer science from the University of Malaya, Malaysia, in 2002. She is currently a professor of the Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya. Her research interests include biomedical informatics, machine learning, data mining and intelligent techniques, ontologies, and image processing. She has published more than 100 peer reviewed articles in journals and conference proceedings.

Hwa Jen Yap received the bachelor's degree in mechanical engineering, the master's of engineering science degree, and the PhD degree from the University of Malaya, Kuala Lumpur, Malaysia, in 2000, 2005, and 2012, respectively. Currently, he serves as an associate professor with the Department of Mechanical Engineering, Faculty of Engineering, University Malaya. He was the head for the Centre of Product Design & Manufacturing (CPDM) from 2013 to 2015. He is a member of the Institute of Engineering and Technology (IET) and a chartered engineer of the Engineering Council, UK. His area of interest in research includes virtual reality (VR), augmented reality (AR), virtual manufacturing (VM), product design, artificial intelligence, robotics, and automation.

Kien-Thai Yong received the PhD degree in botany from the University of Malaya, Malaysia, in 2010. He is currently a senior lecturer with the Ecology and Biodiversity Programme, Institute of Biological Sciences, Faculty of Science, University of Malaya, Malaysia. His main research interests include plant taxonomy and systematic, floristic, and forest ecology.