IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 4, December 2022, pp. 1344~1352
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i4.pp1344-1352

Lotus species classification using transfer learning based on VGG16, ResNet152V2, and MobileNetV2

Nachirat Rachburee, Wattana Punlumjeak
Department of Computer Engineering, Faculty of Engineering, Rajamangala University of Technology Thanyaburi, Pathum Thani, Thailand

Article Info

Article history:

Received Sep 28, 2021
Revised Jul 1, 2022
Accepted Jul 30, 2022

Keywords:

Classification
Convolution neural network
Lotus species classification
Pre-trained weight
Transfer learning

ABSTRACT

Technology has played an increasingly important role in daily life. Especially, technology in object classification that comes in to make human life more comfortable as well as to help people of all ages learning unlimited in anywhere, anytime. Lotus Museum located in Rajamangala University of Technology Thanyaburi (RMUTT) that is open to the general public to learn as well as to cultivate awareness for propagation and result in future preservation. In this paper, we proposed lotus species classification with three pre-trained weights in the ImageNet dataset: visual geometry group (VGG16), residual neural network (ResNet152V2), and MobileNetV2. Fine-turning is used in the last layer after retrained with the custom data we provided. The experimental result shows the accuracy of VGG16, ResNet152V2, and MobileNetV2 are 98.5%, 98.0%, and 99.5% respectively. Therefore, MobileNetV2 not only gives the best accuracy than others but also uses the lowest parameters which are effective in computation time and proper to mobile devices. The proposed research paper on lotus classification base on transfer learning is an effective way to encourage and support people to learn without limitations.

Corresponding Author:

Wattana Punlumjeak
Department of Computer Engineering, Faculty of Engineering, Rajamangala University of Technology Thanyaburi
Pathum Thani, Thailand
Email: wattana.p@en.rmutt.ac.th

1. INTRODUCTION

Nowadays, any learning is not limited to the classroom but can study in other learning sources such as libraries, zoos, aquariums, and museums. Thailand has many learning resources, one of them is the Lotus Museum located in Rajamangala University of Technology Thanyaburi (RMUTT). Since Thailand is in a tropical climate, there will be many flowers that bloom. The lotus is one of a flower that loves sunlight, so there are many different types of lotuses in Thailand. The Lotus Museum in RMUTT preserved more than 100 lotus species. The presence of the Lotus Museum on campus is because the university logo is a lotus flower. The presence of a lotus museum according to the university's logo will communicate to the general public about the university. Also, one of the university's commitments is to preserve the original lotus species so that the lotus can continue to exist. The Lotus Museum is also used for research studies which genetic characteristics of good traits to propagate lotus species that are beneficial to the people and the nation. Therefore, having a lotus museum will result in the general public getting to know the types of lotuses, growth, flowering, condition of the soil in which it is grown propagation, and more in order preserve the lotus species in the future. Lotus Museum has used technology to use in activities or as a channel for knowledge transfer. Being able to learn anywhere, every time is one thing that allows learning to be done endlessly. Therefore, technology is the way for educating people of all ages.

Technology in image processing, object detection and, machine learning are rapidly growing fields and popular topics in the new decade. Image processing is a way to working or perform some operations on an image, convert an image into digital form to get some useful information from it. Machine learning can be categorized into supervised, unsupervised, and reinforcement. Classification is a supervised learning concept which basically identifies new observations on the basis of training data categorizes into classes. The use of classification in object detection is to classify the objects from the image into certain classes that have been learned. Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and learning processes in the human brain called artificial neural networks that use multiple layers to progressively extract higher-level features from the raw input. Neural networks have several different forms, including artificial neural networks, feedforward neural networks, convolutional neural networks, and recurrent neural networks. In object detection and classification, lower layers of neural networks in deep learning may identify image edges, then detect objects in the image while higher layers may identify the classes belonging to the object.

In the field of object detection and classification, one of the challenging ideas of researchers is trying to come up with methods or tools that will help people as well as expanding unlimited learning that can be applied that makes more comfortable in daily life. Many researchers proposed their idea and their work which used the architecture of deep learning to solve problems that are facing. The researcher used a deep convolution neural network to predict covid-19 disease from lung x-ray and computerised tomography (CT)-scan images. They compared traditional convolution neural networks and visual geometry group (VGG16). The result showed that VGG16 gave high accuracy with two classes of classification. They proposed future work that evolved in epidemiology and many models to improves the performance of diagnostic [1], [2]. Transfer learning in deep learning with many convolution neural network architectures used to classify plant disease. Various convolutional neural network (CNN) architecture were used in transfer learning of deep learning model e.g. ResNet152, AlexNet, MobileNet, DenseNet, InceptionRestNetV2, and VGG16 as a pre-trained model in this research to classify rice disease with 7 classes. VGG16 model was the highest accuracy of disease classification [3]. Plant disease classification was the objective of the research and used plant village dataset in the experiment and used DenseNet, VGG16, VGG19, ResNet50, InceptionV3, InceptionResnetV2, MobileNet in transfer learning technique. DenseNet gave the highest accuracy in the experiment [4], [5]. This research [6], [7] compared between traditional convolution neural network and VGG16 pre-trained model with leaf disease classification. The result showed VGG16 was outperformed. Using AlexNet, GoogleNet, VGG16, Resnet50, InceptionV3, DenseNet121, Xception, and MobileNet model to identify medicinal plants [8] and to recognized the plant disease in the agricultural sector [9] which InceptionV3 was the highest performance model.

Flower classification was an objective of their research [10], [11]. They used the VGG16 pre-trained model with Linear Discriminant Loss Function and stochastic gradient descent algorithm. Oxford 102 flower dataset was used in the proposed model with 102 categories. They used stochastic gradient descent to update the weight drop-out method to an optimized model. The result showed VGG16 proposed model was better than the traditional method. [12], [13] researchers used VGG-16, VGG-19, Inception-v3, MobileNet and ResNet50 as pre-trained models to classify flowers with public datasets. They compared the performance of transfer learning models with various flower datasets such as Oxford-16, Oxford-102, and Kaggle flower datasets. They founded transfer learning outperform with traditional method and ResNet was the highest performance model. Plant leaf diseases classification, chili, grape, apple, and other agricultural plants from the public dataset were used as a dataset in plant diseases classification with pre-trained models like AlexNet, DenseNet, VGG16, VGG19, ResNet. Many papers showed performance result from ResNet model that was higher than other models [14]–[16].

The research was used ResNet as a pre-trained model deep learning with some added algorithms such as feature compensation to improved classification technique. ResNet model showed the highest accuracy and represented that ResNet was a proper model in agriculture such as counting, estimation, and prediction of production [17], [18]. ResNet152V2, MobileNet, ResNet50V2, NASNetLarge, and VGG16 were used to classify diseases from X-ray images such as Covid-19 and Pneumonia. The result showed ResNet was outperformed with other model and possible to applied model for diseases classification [19]–[21].

This research approached to compare the performance of MobileNetV1 and MobileNetV2 with several datasets from TensorFlow. The result represented that performance of MobileNetV2 was higher than MobileNetV1 [22]. MobileNetV2 model was used as a base of N-MobileNetV2 to classify various kinds of flowers. A small flower dataset of 5 classes was used in the experiment that is a faster way to train deep learning. The accuracy of MobileNetV2 Transfer learning was 71% higher than without the transfer learning model [23]. Plant identification was approached in the research. MobileNetV2, VGG16, VGG19, ResNetV2, InceptionResNetV2 were used in the experiment. The research collected a plant image dataset from a local location with 33 and 109 classes. The experiment results showed that MobileNetV2 was the highest accuracy rate of the performance model [24], [25]. The research used MobileNetV2 as a based model to improve the classification of plant diseases. They used neural network named you only look once (YOLOV3) and a modified structure model with MobileNetV2 in the experiment. The proposed model of MobileNetV2 based was a suitable model in image classification [26], [27]. Fruit classification was one of the real-world problems to help a farmer improved productivity. Fine-tuned MobileNetV2 was applied with a transfers learning technique to classify fruit images such as berry fruit, apple, kiwi, durian, and guava. MobileNetV2 gave a high performance and proper to mobile devices [28], [29]. The covid-19 epidemic era was a dramatic effect on the world. MobileNetV2, ResNet50V2, Xception, VGG16, DenseNet were used to classify a Covid-19 infect from X-ray images. MobileNetV2 was the highest accuracy model [30], [31].

The paper is organized as follows: Section I is for introducing the problem and the way of the concept that leads to our proposed model. In this section, we explained the relevant literature on the subject with other research papers. In section II, the theoretical basis is given. Section III shows the proposed methodology while section IV described the experiment, results, and discussion. In the last section, the conclusion is presented and also given some direction for future work.

2. METHODOLOGY

In the past, classification object in the image was the one of the main tasks of machine learning by creating some classifiers that could classify the object. Artificial neural networks are complex models, inspired by the biological neural networks which try to mimic the way the human brain develops classification rules. Deep learning is a class of machine learning algorithms that uses multiple layers of neural networks to extract higher-level features from the vast amount of raw input data. The CNN is a class of deep neural networks, most commonly applied to analyze visual imagery. The convolution technique is a mathematical operation that conducts matrix multiplication over a subsection of the pixels in the initial image. After that, pooling layers are used to pool feature maps and then a fully connected layer which flatted the output by changing the matrix of pixels into a vector of pixels and pass it to a regular neural network for classification. In 1998, LeCun *et al*. [32] presented the well-known neural network design known as LeNet-5, which was reputed as the founder of object recognition that combined a convolutional neural network with gradient-based learning for handwritten letter recognition. The face detector system presented by Viola-Jones in 2004 was one of the state-of-the-art for object detection [33]. Instead of utilizing a single classifier, the basic idea of this research was to develop a cascade classifier and integrate it with the AdaBoost learning algorithm. At the time, this paper was a breakthrough in computer vision [34]. AlexNet, a deep convolutional neural network architecture was present by Alex Krizhevsky and his research team in ImageNet large-scale visual recognition challenge (ILSVRC-2012) which been the famous annual computer vision competition. This model showed the best performance and achieved a winning top-5 test error rate in the competition.

At that time, deep architecture is becoming more and more popular to solve computer vision tasks such as object detection, image classification, and so on. Many structures of CNN architecture models had been adapted and improved based on the AlexNet structure. Visual geometry group (VGG) architecture presented by [35] address another important aspect of CNNs from AlexNet. An input layer, VGG uses a 224x224 pixel red green blue (RGB) image as input. In VGG, the convolutional layers use a 3x3 receptive field filter. The convolution stride is set to 1 pixel, resulting in a linear transformation of the input before a ReLU unit is used. The last layers, fully-connected layers which VGG has three fully connected layers: the first two layers have 4096 channels each for each class and the third layer has 1000 channels. The architecture of VGG depicted in Figure 1.

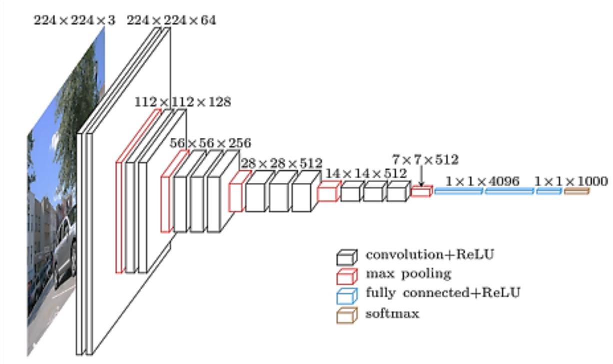

Figure 1. Architecture of VGG [35]

VGG is now still one of the most used and also outperforms baselines on image recognition architectures and datasets outside of ImageNet. In the meantime, [36] One of the popular convolutional neural networks namely ResNet proposed by a researcher group of Microsoft. ResNet architecture is the winner of ILSVRC-2015 (Image classification, localization, and detection). To solve the problem of vanishing/exploding gradients, a group of Microsoft Inventors introduces skip connection to fit the input from the previous layer to the next layer without any alteration. The Input layers of ResNet architecture are composed of multiple residual blocks regarding and their operating principle is concerned with optimizing a residual function. The following figure shows different architectures of ResNet which the number of layers varies from 18 to 152 which were used for classifying the ImageNet dataset as Figure 2.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 64\\3\times3, 64\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 64\\3\times3, 64\\1\times1, 256\end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 128\\3\times3, 128\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, 128\\3\times3, 128\\1\times1, 512\end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 256\\3\times3, 256\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1, 256\\3\times3, 256\\1\times1, 1024\end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, 512\\3\times3, 512\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, 512\\3\times3, 512\\1\times1, 2048\end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 2. The overall architecture for ResNet [36]

In 2017, [37] the researcher team from google presented a research paper that improved the weakness of the prior state-of-the-art architecture in size and complexity. The MobileNets are based on an architecture which used in Inception models that uses depthwise separable convolutions to focus on not only optimizing for latency primarily but also yield small networks. MobileNets build light weight deep neural networks and are particularly useful for mobile and embedded vision applications. The architecture of MobileNets shown in Figure 3. The improvement over MobileNetV1 was presented in 2019, namely MobileNetV2 from the google research team [38]. A MobileNetV2 architecture used inverted bottleneck blocks and residual connections to replace expensive convolutions which decrease the number of operations and memory needed while retaining the same accuracy. The convolutional blocks of MobileNetsV2 architecture are shown in Figure 4.

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5× Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
5× Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

Figure 3. MobileNets architecture [37]

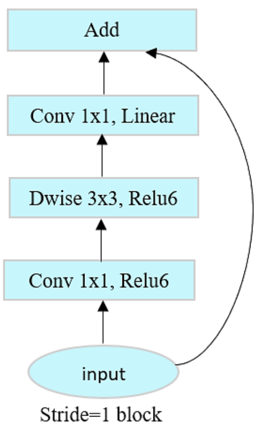

Figure 4. The convolutional blocks of MobileNetsV2 architecture [38]

Transfer learning is the idea of machine learning method where a model developed the isolated learning paradigm and utilizing knowledge acquired for a task is reused or transferred as the starting point in the process of learning for a second task or the related one. Because of the lack of sufficient training data and also the computational complexity with required more computational resources, transfer learning needs not start from scratch and is now more popular in deep learning. An annual computer vision contest, the ImageNet large scale visual recognition challenge (ILSVRC) challenge, ImageNet dataset were used to find the winning of the year in image classification, single-object localization, and object detection. ImageNet, consisting of more than 1 million labeled high-resolution images belonging to 1000 classes. The pre-trained network model is a saved network model that was already trained on a huge benchmark dataset. All of the states of the art of CNN architectures like VGG, ResNet, MobileNet, Inception, and DenseNet, are already have pre-trained networks or deep transfer learning available which leverage knowledge and use as a starting point for a new one. As a matter of knowledge, transfer learning means the process of incorporating pre-trained models into a new model. The most popular strategies for deep transfer learning are pre-trained models as feature extractors and fine tuning pre-trained models.

3. PROPOSED METHODOLOGY

Lotus species classification using transfer learning based on VGG16, ResNet152V2, and MobileNetV2 has been proposed and experimented with vital data. The aim of the proposed model is to apply image detection and classification by transfer deep learning technology to increase opportunities for learning about lotus species and their information in order to be preserved and be sustainable in nature. Therefore, our proposed methodology had 3 tasks: the first task was pre-processing which we worked with our related lotus image by taking photos of the lotus as much as possible. Then, resize or rescale to adjust the size of the image to be suitable as input in each model. Moreover, image augmentation is used to altering the existing data through different ways of processing such as random rotation, shifts, shear, and flips, to create some more data for the training process. After that, we split our dataset into 70% of data for training and 30% for testing. Secondly, we applied the pre-trained of VGG16, ResNet152V2, and MobileNetV2 to transfer learning with our custom lotus dataset. And finally, a fine-tuning task, the pre-trained weights were frozen while re-training the image and added top layers which some parameters to the rest to optimize in order to achieve better performance. The proposed ideas of our work are depicted in the Figure 5.

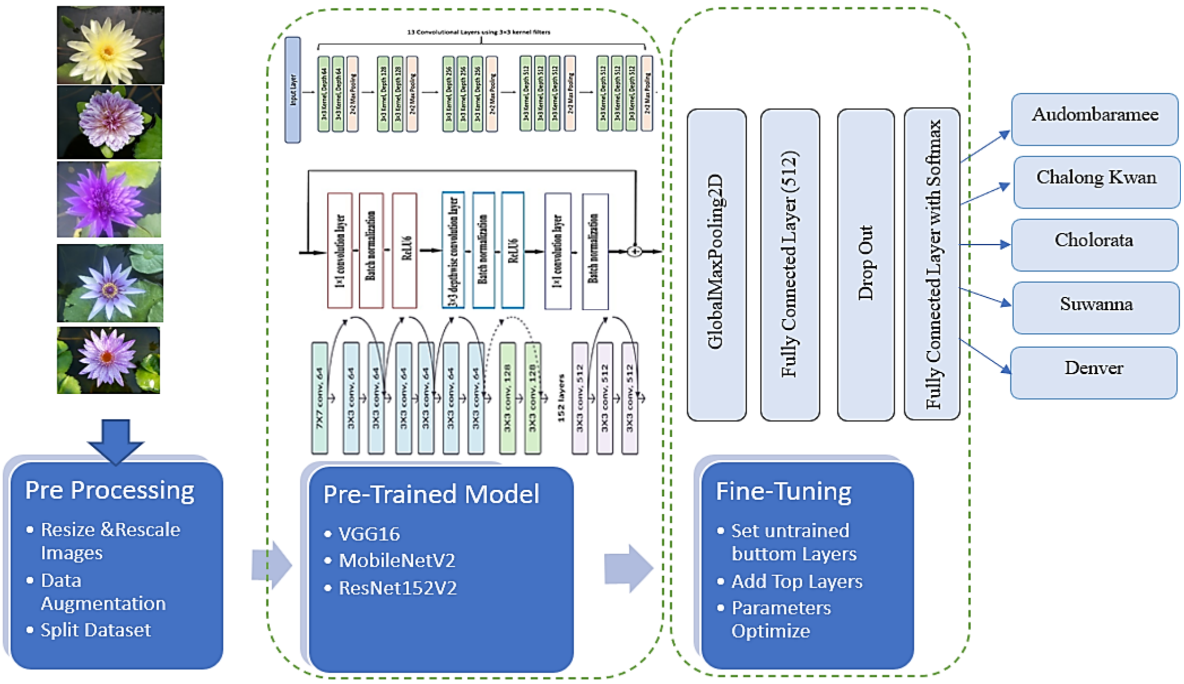

Figure 5. The proposed methodology

4. RESULTS AND DISCUSSION

The proposed methodology experimented on Google Colaboratory: Colab 12GB-RAM graphical processing units (GPU) with the following steps. Firstly, we prepared the data used by taking photos of five types of lotuses. The name of lotus species that have been selected are, (i) Audombaramee, the scientific name is Nymphaea Audombaramee; (ii) King of Siam also known as ChalongKhwan, the scientific name is Nymphaea 'King of Siam'; (iii) Cholorata, the scientific name is Nymphaea Cholorata; (iv) Suwanna, the scientific name is Nymphaea 'Suwanna'; (v) Denver, the scientific name is Nymphaea 'Denver HB'.

The following pictures are the example of the each lotus flower as shown in Figure 6. We collected a picture of the lotus flowers by taking the photo in the Lotus Museum which is located in the Rajamangala University of Technology and other lotus planting areas. After collecting the lotus images, the total number of pictures of all five lotus species is 720. The next steps of preprocessing e.g. resize or rescale rotation, shifts, shear, and flips, are used to create some more data. We split the data after pre-processing step into 70% for training and 30% for testing.

Figure 6. The example of each type of lotus flower

Secondly, we used pre-trained networks to transfer the knowledge in our custom dataset. The pre-trained architecture networks we used in our experiment are VGG16, ResNet152V2, and MobileNetV2 with the dataset we prepared from the process. In addition, the pre-trained weight of all architecture which used in this experiment is ImageNet dataset. We froze or fixed weights in extracted layers while retraining and fine-tune the rest of them to suit our needs. A fine-tuning task is involved in this step e.g. dropout some network layer to reduced overfitting, change an optimizer to Adam, try to change the variables (learning rate = 0.0001, epochs = 100), used softmax activation function as classifier function in a fully-connected layer to get the best accuracy and performance. The result of pre-trained architecture: VGG16 in term of accuracy, loss, and confusion matrix are shown as Figure 7-8, pre-trained architecture: ResNet152V2 are shown as Figure 9-10, and pre-trained architecture: MobileNetV2 are shown as Figure 11-12.

Furthermore, all parameters used in each architecture are presented in Table 1. The overall accuracy and loss are presented in Table 2 respectively. After the experiments, it appeared that the experimental results on MobileNetV2 consistently outperform very specialized task-focused not only achieved the best overall accuracy (99.5%) but also used the number of lowest parameters which the best performance.

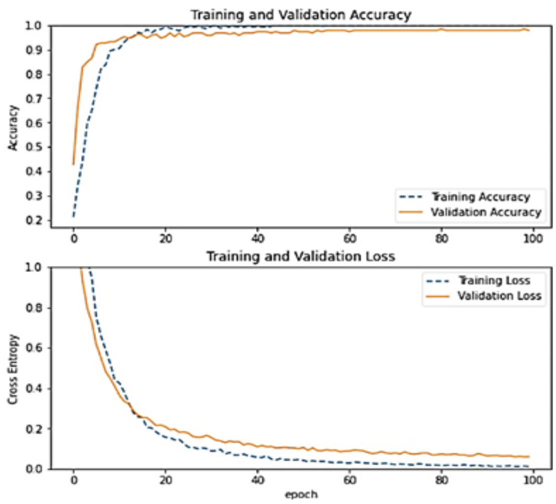

Figure 7. An accuracy and loss in VGG16

```
['audombaramee', 'chalongkwan', 'colorata', 'denver', 'suwanna']
[[38  0  0  0  0]
 [ 0 49  0  0  0]
 [ 0  1 52  0  0]
 [ 0  2  1 37  0]
 [ 0  0  0  0 30]]
              precision    recall  f1-score   support

audombaramee       1.00      1.00      1.00        38
 chalongkwan       0.94      1.00      0.97        49
    colorata       0.98      0.98      0.98        53
      denver       1.00      0.93      0.96        40
     suwanna       1.00      1.00      1.00        30

    accuracy                           0.98       210
   macro avg       0.98      0.98      0.98       210
weighted avg       0.98      0.98      0.98       210
```

Figure 8. Confusion matrix in VGG16

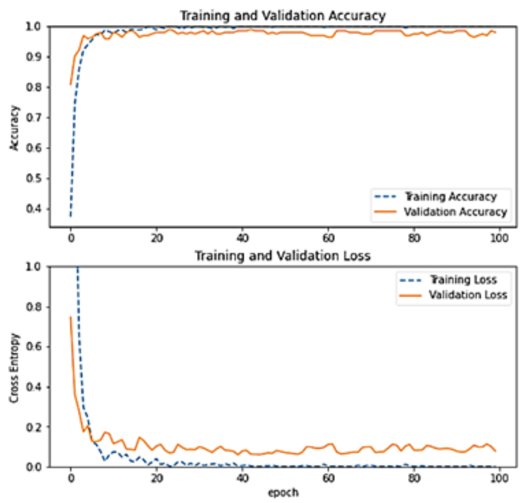

Figure 9. An accuracy and loss in ResNet152V2

```
['audombaramee', 'chalongkwan', 'colorata', 'denver', 'suwanna']
[[38  0  0  0  0]
 [ 0 48  1  0  0]
 [ 0  1 51  1  0]
 [ 0  0  0 40  0]
 [ 0  0  0  2 28]]
              precision    recall  f1-score   support

audombaramee       1.00      1.00      1.00        38
 chalongkwan       0.98      0.98      0.98        49
    colorata       0.98      0.96      0.97        53
      denver       0.93      1.00      0.96        40
     suwanna       1.00      0.93      0.97        30

    accuracy                           0.98       210
   macro avg       0.98      0.98      0.98       210
weighted avg       0.98      0.98      0.98       210
```

Figure 10. Confusion matrix in ResNet152V2

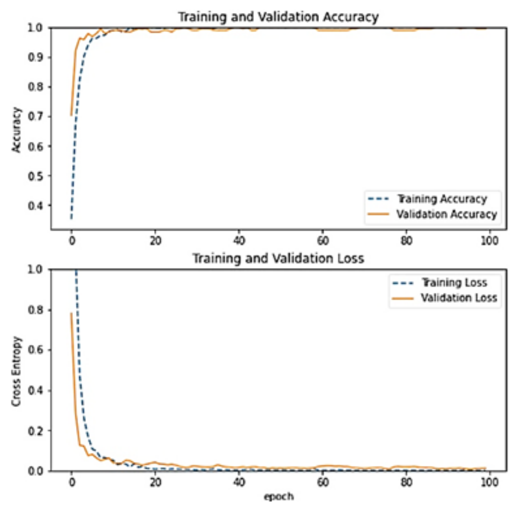

Figure 11. An accuracy and loss in MobileNetV2

```
['audombaramee', 'chalongkwan', 'colorata', 'denver', 'suwanna']
[[38  0  0  0  0]
 [ 0 49  0  0  0]
 [ 0  1 52  0  0]
 [ 0  0  0 40  0]
 [ 0  0  0  0 30]]
              precision    recall  f1-score   support

audombaramee       1.00      1.00      1.00        38
 chalongkwan       0.98      1.00      0.99        49
    colorata       1.00      0.98      0.99        53
      denver       1.00      1.00      1.00        40
     suwanna       1.00      1.00      1.00        30

    accuracy                           1.00       210
   macro avg       1.00      1.00      1.00       210
weighted avg       1.00      1.00      1.00       210
```

Figure 12. Confusion matrix in MobileNetV2

Table 1. Parameters used in each architecture

Pre-trained weight	Parameters		
	Trainable parameters	*Non-trainable parameters*	*Total parameters*
VGG16	265,221	14,714,688	14,979,909
ResNet152V2	1,051,653	58,331,648	59,383,301
MobileNetV2	658,437	2,257,984	2,916,421

Table 2. The overall accuracy and loss in each architecture

Pre-trained weight	*Accuracy*	*Loss*
VGG16	0.981	0.055
ResNet152V2	0.976	0.081
MobileNetV2	0.995	0.011

5. CONCLUSION

In this paper we have proposed a novel framework of transfer learning in pre-trained deep learning architectures based on VGG16, ResNet152V2, and MobileNetV2 to classify the lotus species. We collected a real-world dataset by taking the photo of the lotus flowers in the Lotus Museum in RMUTT and other lotus planting areas. Our proposed methodology used fine-tuning pre-trained models that weighted in ImageNet dataset. The experiment results prove the effectiveness of pre-trained MobileNetV2 which is not only able to classify the name in the class of the lotus species in the highest accuracy (99.5%) and lowest of loss (1.4%) but also time and speed in terms of parameter usage. The results of the experiment can be useful in applying the model obtained in mobile phones. People can take benefit from learning easily without limitations. An important point is that the people know more about the types of lotuses and the knowledge of planting and propagate lotus affecting the preservation of lotus species. In future work, we will study and experiment with the other CNN architectures and another famous one-stage detector like a YOLO or Siamese neural network to find a suitable one for classifying the lotus species.

REFERENCES

[1] N. Alrefai and O. Ibrahim, "Deep learning for COVID-19 diagnosis based on chest X-ray images," IJECE, vol. 11, no. 5, p. 4531, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4531-4541.

[2] Hanafi, A. Pranolo, and Y. Mao, "Cae-covidx: Automatic covid-19 disease detection based on x-ray images using enhanced deep convolutional and autoencoder," *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 1, pp. 49–62, 2021, doi: 10.26555/ijain.v7i1.577.

[3] V. K. Shrivastava, M. K. Pradhan, and M. P. Thakur, "Application of Pre-Trained Deep Convolutional Neural Networks for Rice Plant Disease Classification," *Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, pp. 1023–1030, 2021, doi: 10.1109/ICAIS50930.2021.9395813.

[4] M. A. Ihsan Aquil and W. H. Wan Ishak, "Evaluation of scratch and pre-trained convolutional neural networks for the classification of Tomato plant diseases," IJ-AI, vol. 10, no. 2, p. 467, Jun. 2021, doi: 10.11591/ijai.v10.i2.pp467-475.

[5] U. B. Patayon and R. V. Crisostomo, "Peanut leaf spot disease identification using pre-trained deep convolutional neural network," IJECE, vol. 12, no. 3, p. 3005, Jun. 2022, doi: 10.11591/ijece.v12i3.pp3005-3012.

[6] J. Gu, P. Yu, X. Lu, and W. Ding, "Leaf species recognition based on VGG16 networks and transfer learning," *IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 2189–2193, 2021, doi: 10.1109/IAEAC50856.2021.9390789.

[7] S. B. Jadhav, "Convolutional Neural Networks for Leaf Image-Based Plant Disease Classification," IJ-AI, vol. 8, no. 4, p. 328, Dec. 2019, doi: 10.11591/ijai.v8.i4.pp328-341.

[8] T. Nguyen Quoc and V. Truong Hoang, "Medicinal Plant identification in the wild by using CNN," *International Conference on ICT Convergence*, vol. 2020-October, pp. 25–29, 2020, doi: 10.1109/ICTC49870.2020.9289480.

[9] A. Julianto and A. Sunyoto, "A performance evaluation of convolutional neural network architecture for classification of rice leaf disease," IJ-AI, vol. 10, no. 4, p. 1069, Dec. 2021, doi: 10.11591/ijai.v10.i4.pp1069-1078.

[10] M. Qin, Y. Xi, and F. Jiang, "A New Improved Convolutional Neural Network Flower Image Recognition Model," *2019 IEEE Symposium Series on Computational Intelligence, SSCI 2019*, pp. 3110–3117, 2019, doi: 10.1109/SSCI44817.2019.9003016.

[11] R. Lv, Z. Li, J. Zuo, and J. Liu, "Flower Classification and Recognition Based on Significance Test and Transfer Learning," *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, pp. 649–652, 2021, doi: 10.1109/ICCECE51280.2021.9342468.

[12] Y. Wu, X. Qin, Y. Pan, and C. Yuan, "Convolution neural network based transfer learning for classification of flowers," *2018 IEEE 3rd International Conference on Signal and Image Processing, ICSIP 2018*, pp. 562–566, 2019, doi: 10.1109/SIPROCESS.2018.8600536.

[13] C. Narvekar and M. Rao, "Flower classification using CNN and transfer learning in CNN-Agriculture Perspective," *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems, ICISS 2020*, pp. 660–664, 2020, doi: 10.1109/ICISS49785.2020.9316030.

[14] K. P. Akshai and J. Anitha, "Plant disease classification using deep learning," *2021 3rd International Conference on Signal Processing and Communication, ICPSC 2021*, pp. 407–411, 2021, doi: 10.1109/ICSPC51351.2021.9451696.

[15] A. Alsayed, A. Alsabei, and M. Arif, "Classification of Apple Tree Leaves Diseases using Deep Learning Methods," IJCSNS International Journal of Computer Science and Network Security, vol. 21, no.7, pp324-330, July 2021.

[16] A. D. A. Aldabbagh, C. Hairu, and M. Hanafi, "Classification of chili plant growth using deep learning," *2020 IEEE 10th International Conference on System Engineering and Technology, ICSET 2020 - Proceedings*, pp. 213–217, 2020, doi: 10.1109/ICSET51301.2020.9265351.

[17] W. J. Hu, J. Fan, Y. X. Du, B. S. Li, N. Xiong, and E. Bekkering, "MDFC-ResNet: An Agricultural IoT System to Accurately Recognize Crop Diseases," *IEEE Access*, vol. 8, pp. 115287–115298, 2020, doi: 10.1109/ACCESS.2020.3001237.

[18] V. Kumar, H. Arora, Harsh, and J. Sisodia, "ResNet-based approach for Detection and Classification of Plant Leaf Diseases," *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020*, pp. 495–502, 2020, doi: 10.1109/ICESC48915.2020.9155585.

[19] P. T. Selvy, S. K. L. Sai Kumar, R. Thirumalraj, and G. Vimal, "Retraction: Deep learning methods to analyse and detect the presence of COVID-19 using X-rays," *Journal of Physics: Conference Series*, vol. 1916, no. 1, 2021, doi: 10.1088/1742-6596/1916/1/012170.

[20] N. M. Elshennawy and D. M. Ibrahim, "Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images," *Diagnostics*, vol. 10, no. 9, 2020, doi: 10.3390/diagnostics10090649.

[21] A. Musha, A. Al Mamun, A. Tahabilder, Md. J. Hossen, B. Hossen, and S. Ranjbari, "A deep learning approach for COVID-19 and pneumonia detection from chest X-ray images," IJECE, vol. 12, no. 4, p. 3655, Aug. 2022, doi: 10.11591/ijece.v12i4.pp3655-3664.

[22] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "MobileNetV2 Model for Image Classification," *Proceedings - 2020 2nd International Conference on Information Technology and Computer Application, ITCA 2020*, pp. 476–480, 2020, doi: 10.1109/ITCA52113.2020.00106.

[23] W. Dai, Y. Dai, K. Hirota, and Z. Jia, "A Flower Classification Approach with MobileNetV2 and Transfer Learning," *The 9th International Symposium on Computational Intelligence and Industrial Applications*, pp. 1–5, 2020, [Online]. Available: https://isciia2020.bit.edu.cn/docs/20201114083020836285.pdf.

[24] N. Van Hieu and N. L. H. Hien, "Automatic plant image identification of Vietnamese species using deep learning models," *International Journal of Engineering Trends and Technology*, vol. 68, no. 4, pp. 25–31, 2020, doi: 10.14445/22315381/IJETT-V68I4P205S.

[25] T. Akiyama, Y. Kobayashi, Y. Sasaki, K. Sasaki, T. Kawaguchi, and J. Kishigami, "Mobile leaf identification system using CNN applied to plants in Hokkaido," *2019 IEEE 8th Global Conference on Consumer Electronics, GCCE 2019*, pp. 324–325, 2019, doi: 10.1109/GCCE46687.2019.9015298.

[26] S. Z. M. Zaki, M. Asyraf Zulkifley, M. Mohd Stofa, N. A. M. Kamari, and N. Ayuni Mohamed, "Classification of tomato leaf diseases using MobileNet v2," IJ-AI, vol. 9, no. 2, p. 290, Jun. 2020, doi: 10.11591/ijai.v9.i2.pp290-296.

[27] J. Liu and X. Wang, "Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model," *Plant Methods*, vol. 16, no. 1, 2020, doi: 10.1186/s13007-020-00624-2.

[28] Venkatesh, Nagaraju, S. U. Hegde, and Stalin, "Fine-tuned MobileNet Classifier for Classification of Strawberry and Cherry Fruit Types," *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*, 2021, doi: 10.1109/ICCCI50826.2021.9402444.

[29] A. L. Sabarre, A. S. Navidad, D. S. Torbela, and J. J. Adtoon, "Development of durian leaf disease detection on Android device," IJECE, vol. 11, no. 6, p. 4962, Dec. 2021, doi: 10.11591/ijece.v11i6.pp4962-4971.

[30] M.-L. Huang and Y.-C. Liao, "A lightweight CNN-based network on COVID-19 detection using X-ray and CT images," Computers in Biology and Medicine, vol. 146, p. 105604, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105604

[31] H. Mary Shyni and E. Chitra, "A comparative study of X-ray and CT Image in COVID-19 detection using image processing and deep learning techniques," Computer Methods and Programs in Biomedicine Update, vol. 2, p. 100054, 2022, doi: 10.1016/j.cmpbup.2022.100054..

[32] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

[33] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004, doi: 10.1023/B:VISI.0000013087.49260.fb.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.

[35] K. S. and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *International Conference on Learning Representations*, 2015, doi: 10.48550/arXiv.1409.1556.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[37] A. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *Computer Vision and Pattern Recognition*, 2017, doi: 10.48550/arXiv.1704.04861.

[38] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

BIOGRAPHIES OF AUTHORS

Nachirat Rachburee is a lecturer at Department of Computer Engineering, Faculty of Engineering, Rajamagala University of Technology, Pathum Thani, Thailand. His research interests include Data Mining, Big data analytics, Deep Learning, Neural Networks, and Predictive analytics. He can be contacted at email: nachirat.r@en.rmutt.ac.th.

Wattana Punlumjeak is a lecturer at Department of Computer Engineering, Faculty of Engineering, Rajamagala University of Technology, Pathum Thani, Thailand. Her research interests include Data Mining, Big data analytics, Deep Learning, Neural Networks, and Predictive analytics. She can be contacted at email: wattana.p@en.rmutt.ac.th.