Comparison of Transfer Learning Approaches for Image Classification

Ei Phyu Myint
University of Computer Studies, Yangon
eiphyumyint@ucsy.edu.mm

Thin Lai Lai Thein
University of Computer Studies, Yangon
tllthein@ucsy.edu.mm

***Abstract*—Deep learning technologies have been successful in many fields in recent years. The deep learning model owns a powerful learning ability by integrating the feature extraction and classification process into a whole in order to complete the image classification test. In this way, it can effectively improve the image classification accuracy. From a deep learning perspective, the image classification problem can be solved through transfer learning. This research shows how to implement a transfer learning solution for image classification problems by using Pytorch framework and python programming language. This research also shows the comparison of two pre-trained deep learning models such as ResNet-50, and VGG-16 by using transfer learning based on fine-tuning, co-tuning and behavioral regularization.**

Keywords—Image classification, Deep learning, Transfer learning

I. Introduction

One of the fields of computer science is computer vision that enables computers to focus on identifying and processing objects in images and videos as the same way as humans do. Computer vision techniques can make a computer to extract, analyse and understand useful information from a single or sequence of images. Image classification, image classification with localization, object segmentation, object detection and so on are some of these techniques.

Image classification technique is mainly focused on this paper. Image classification is a process that can predict a specific class or label for something that is defined by a set of data points. The classification process of a specific image may include distinguishing an image among potentially countess categories of images. In addition, there are hundreds of thousands of images for respective categories. A manual classification includes comparing images and grouping similar ones according to their characteristics respectively. It is a very tedious and time-consuming process. So, in order to classify images quickly and efficiently, an automatic system is needed.

Deep learning has had a tremendous impact on various field in science over the last years. For everything related to computer vision, this technology is also highly relevant obviously. The objects in digital images from cameras and videos can be accurately identified and classified by using deep learning models. However, deep learning systems require a large amount of data to get an effective model. In addition, it may take days or even weeks to train on very large datasets.

Transfer learning is the idea of how to utilize knowledge acquired for one task to solve related ones. In computer vision domain, there are various deep learning networks with state-of-the-art performance that have been developed and tested. The deep transfer learning technique uses these pre-trained networks or models as the basis of transfer learning in the context of deep learning. In transfer learning for image classification, a model will effectively serve as a generic model of the visual world if it is trained on a large and general enough dataset. This pre-trained model can be used to customize for a given task by applying transfer learning without having to start from scratch by training a large model on a large dataset. A well-known transfer learning approach is fine-tuning.

The studying of how to take advantage of knowledge from other fields by deep neural networks is called deep transfer learning. A significant amount of deep transfer learning methods have been proposed since deep neural networks have become popular in various fields. Deep transfer learning can be classified into four categories such as network-based deep transfer learning, instances-based deep transfer learning, adversarial-based deep transfer learning, and mapping-based deep transfer learning based on the techniques used in deep transfer learning [1]. Network-based deep transfer learning means that it refers to the reuse of the partial network that pre-trained in the source domain, consisting of its network structure and connection parameters, and then transfer it to be a part of deep neural network which is used in target domain. This paper compares three deep transfer learning methods based on network-based deep transfer learning technique. These three deep transfer learning methods are fine-tuning, co-tuning and deep learning transfer using feature map with attention for convolutional neural network.

This paper is organized as follows: Section II presents similar works. In section III, the deep transfer learning strategy is presented and three transfer learning approaches are also described. Experimental results are analyzed in section IV. Finally, conclusion is presented in section V.

II. Related Works

In computer vision, object classification is an important task. Emine Cengıl and Ahmet Çinar [2] presented flowers classification system based on transfer learning by using the Alexnet, Googlenet, VGG16, DenseNet and ResNet pre-trained models. A forecast model for classifying fast food images in Thailand is presented in N. Hnoohom et al. [3]. Natural images (GoogLeNet dataset) was used to train the model and the predictable Thailand fast-food model is created by using a fine tuned deep learning process. Treesukon Treebupachatsakul and Suvit Poomrittigul [4] utilized LeNet CNN method to classify the bacteria images in standard resolution by using deep learning process. [5] presented rocket classification model by using deep learning process based on the residual network ResNet18. In [6] Magnetic resonance (MR) images were classified by utilizing deep transfer learning process based on ResNet34 model.

III. METHODOLOGY

A. *Transfer Learning for Deep Learning*

Deep learning models can be built only if there is a million of data. But a customized model can be built without having large data set by using transfer learning.

A network trained on a different domain for a different source task is taken and a customized network for target domain and target task can be adapted instead of training a deep network from scratch for target task.

B. *Transfer Learning based on fine-tuning*

It is a common practice to take advantage of features learned by a model trained on a larger dataset in the same domain when working with a small dataset. In fine-tuning (partially transfer) approach, pre-trained models are only partially transferred by discarding task-specific layers and fine-tuning bottom layers as a feature extractor. Removing task-specific top layers can cause the loss of 20% of the total parameters. To achieve full transfer of DNNs, a co-tuning framework is also used in this paper. In fine-tuning (freezing some layers) approach, customized models are built by fixing weights in lower convolution layers and re-training weights in upper layers using data from the target domain. This can cause over-fitting because parameters of the target model may be driven far away from initial values. In order to solve this issue, deep learning transfer using feature map with attention for convolutional neural network framework is also used based on behavioral regularization.

C. *Transfer Learning by using co-tuning framework*

The most important part of co-tuning framework is learning category relationship. So, as shown in figure 3, it learns the relationship between categories and target categories firstly, then one-hot target labels are translated into probabilistic source labels which collaboratively supervise the fine-tuning process.

Suppose that f_0 is a probabilistic model, y_s is a source category and y_t is a target label to learn category relationship. As shown in figure 3, categories in the pre-trained dataset are diverse enough to serve as basic categories to compose target category. So, the mapping y_s->y_t is firstly learned from $(f_0(x_t),y_t)$ pairs, where $f_0(x) \approx p(y_s|x)$ is a probability distribution over source categories y_s. In this way, $P(y_s| \ y_t)$ can be computed from $p(y_t\,|y_s)$ by Bayes's rule.

D. *Transfer Learning by using feature map with attention for convolutional neural network framework*

The "Behavior", the outer layer (convolutional layer) outputs (e.g. the feature maps) produced by each layer should be regularized. Through aligning the behaviors of the convolutional layers of the target network to the source one, which has been pre-trained using an extremely large dataset by using constrained feature maps, the generalization capacity could be improved.

Through re-weighting the feature maps with a novel supervised attention mechanism, deep learning transfer using feature map with attention for convolutional neural network framework selects the discriminative features from outer layer outputs. The distance between source/target networks using the outer layer outputs is characterized and such distance is incorporated as the regularization term of the loss function. The optimization for weights of deep neural network is finally affected by such regularization and with the back-propagation, the target network generalization capacity inherited from the source network is also awarded. Specifically, using an attention mechanism with feature map regularization makes the approach to identify the transferable channels and preserve such filters through regularization and identify the untransferable channels and reuse them.

IV. EXPERIMENTAL RESULTS

A. *Data Set*

Three thousand images are used for dataset. The data set is split as training, validation and testing datasets as follows:

Dataset	Number of images
Training	1920
Validation	600
Testing	480

There are six classes to classify such as bicycle, bus, car, motorcycle, person and truck.

B. *Using Fine-tuning for transfer learning*

For fine-tuning, stochastic gradient descent optimizer of pytorch is used with learning rate (0.001) and momentum (0.9).

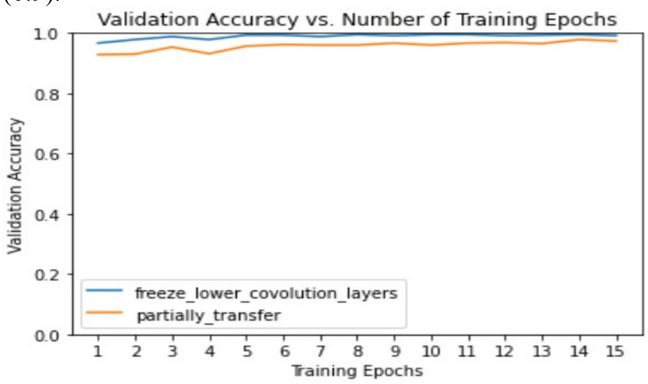

Figure 1. Accuracy of finetuning ResNet-50 model

In figure 1, the learning curve shows that fine-tuning (partially transfer)'s accuracy higher than fine-tuning (freezing some layers) by using pre-trained ResNet-50 model. Figure 2 shows that the accuracy curve for fine-tuning of VGG-16 pre-trained model with fine-tuning (partially transfer) and fine-tuning (freezing some layers). In transfer learning by using VGG-16 pre-trained model, the accuracy of fine-tuning (freezing some layers) can classify images more accurately than fine-tuning (partially transfer). So, "partially transfer method" is more suitable than "freezing some layers method" if fine-tuning approach is used for transfer learning.

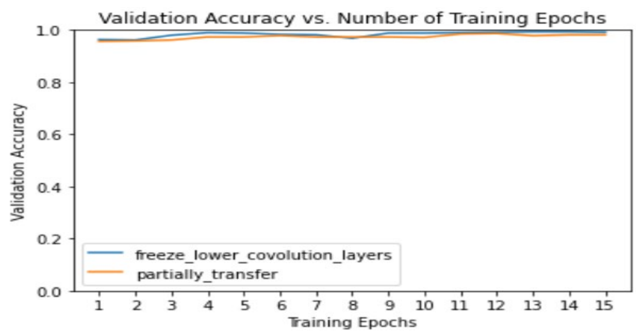

Figure 2. Accuracy of finetuning VGG-16 model

C. *Using Co-tuning for transfer learning*

In figure 3, the category relation between ImageNet dataset classes and Proposed dataset classes. As shown in figure, "bus" class in Proposed dataset has higher category relationship with "minibus" class in ImageNet dataset.

Proposed Dataset clas	Top 3 ImageNet Dataset class		
bicycle	jinrikisha, ricksha, rickshaw	milk can	bicycle-built-for-two, tandem bicycle, tandem
bus	minibus	recreational vehicle, RV, R.V.	police van, police wagon, paddy wagon, patrol wagon, wagon, black Maria
car	minivan	sports car, sport car	knee pad
motorcycle	half track	moped	black grouse
person	red-breasted merganser, Mergus serrator	abaya	bearskin, busby, shako
truck	moving van	convertible	bulletproof vest

Figure 3. Category Relationship between Proposed Dataset and ImageNet Dataset Classes

D. *Using Behavioral regularization for transfer learning*

The network behaviors are intended to regulate and some layers of the target network are forced to behave as similar as to the source networks. The distance between the outer layer outputs of the two networks is considered to regularize the behavior of the networks. So, feature maps are needed to extract from outer layer outputs (convolutional layers). For weighting feature maps with supervised attention models, a supervised attention method is proposed, in which the potential performance loss characterizes the weights of the features when removing these features from the network.

Figure 4 shows that the comparisons of accuracy for transfer learning methods such as fine-tuning (partially transfer), fine-tuning (freezing some layers), co-tuning and behavioral regularization. Behavioral regularization approach has the highest classification accuracy.

Pre-trained model	Fine-tuning(partially transfer)	Co-Tuning	Fine-tuning(fixing weights)	Behavior-based Regularization
ResNet50	0.82	0.97	0.70	0.98
VGG16	0.15	0.46	0.46	0.53

Figure 4. Classification Accuracy of fine-tuning (partially transfer), fine-tuning (freezing some layers), co-tuning and behavioral regularization

Figure 5 also shows the classification results of some of the images in the proposed dataset.

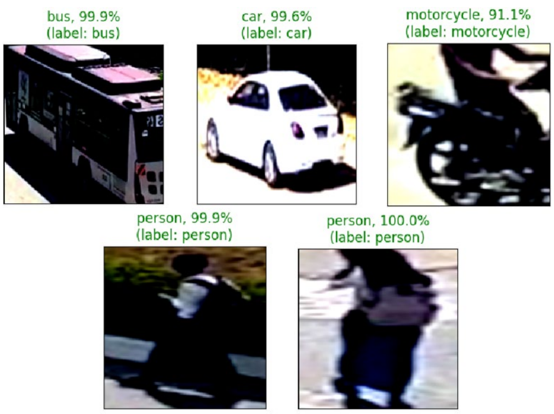

Figure 5. Classification Result of some images in the proposed dataset

V. Conclusion

This research presents the comparison of pre-trained models by customizing based on transfer learning with fine-tuning, co-tuning and behavioral regularization. The issues of fine-tuning approach can be solved by using co-tuning and behavioral regularization. Moreover, a customized model can be built by using pre-trained models based on transfer learning even if there is a small data set. It also makes us to save time for training of data set.

References

[1]Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C. (2018). "A Survey on Deep Transfer Learning". Artificial Neural Networks and Machine Learning – ICANN 2018. Vol. 11141, pp 270-279.

[2]Cengıl, Emine and Çinar, Ahmet, "Multiple Classification of Flower Images Using Transfer Learning," 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), IEEE, 2019.

[3]Hnoohom, Narit, and Sumeth Yuenyong, "Thai fast food image classification using deep learning", 2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON), IEEE, 2018.

[4]Treesukon Treebupachatsakul, Suvit Poomrittigul, "Bacteria Classification using Image Processing and Deep learning", 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), IEEE, 2019.

[5]Liang Zhang, Zhenhua Chen, Jian Wang and Zhaodun Huang, "Rocket Image Classification Based on Deep Convolutional Neural Network", 2018 10th International Conference on Communications, Circuits and Systems (ICCCAS), IEEE, 2018.

[6]Talo, M., Baran Baloglu, U., Yıldırım, O., Rajendra Acharya, U., Application Of Deep Transfer Learning For Automated Brain Abnormality Classification Using Mr Images, Cognitive Systems Research (2018).

Acknowledgment

I would like to thank my teachers, parents, colleagues and friends.