PROCEEDINGS OF
INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGIES
https://proceedings.icatsconf.org/
11th International Conference on Advanced Technologies (ICAT'23), Istanbul-Turkiye, August 17-19, 2023.

Utilizing Transfer Learning on Landscape Image Classification Using the VGG16 Model

Abubakar MAYANJA [1], İlker Ali ÖZKAN [2], Şakir TAŞDEMİR [3]

[1] *Institute of science, Selcuk University, Konya, Türkiye*
abubakar.mayanja@karatay.edu.tr, 0000-0003-4576-5771

[2] *Department of Computer Engineering, Selcuk University, Konya, Türkiye*
ilkerozkan@selcuk.edu.tr, 0000-0002-5715-1040

[3] *Department of Computer Engineering, Selcuk University, Konya, Türkiye*
stasdemir@selcuk.edu.tr, 0000-0002-2433-246X

***Abstract*— In recent times, the need for the use of image classification techniques of machine learning to solve worldly problems in various areas such as agriculture, the health sector, and tourism is rocketing up day by day. Traditionally, one of the most used techniques in image classification is the use of deep neural networks called convolution neural networks (CNN). To come up with a good network model, one needs to have an enormous quantity of data in the form of images and design a network model from scratch in a trial-and-error way. This not only takes a lot of time but also requires very powerful computation equipment such as graphical processing units (GPU). To overcome such barriers, a machine learning technique called transfer learning enables the use of already trained network models in the form of fine-tuning them to solve related issues. In this work, the 2014 ImageNet winner model called Vgg16 was adopted to classify landscape images in the Intel dataset. The dataset contains 5 categories of images namely buildings, forest, glacier, mountain, sea, and street. The performance of Vgg16 was compared to that of a 7-layer ordinary convolution neural network and the results showed that transfer learning with Vgg16 outperformed the ordinary network by 90.1% for Vgg16 compared to 62.5% for the ordinary convolutional neural network model.**

***Keywords*— classification, deep neural networks, landscape image, transfer learning, VGG16 model**

I. INTRODUCTION

As humans, we possess the ability to transfer our mastery skills from one activity to another with ease and in minimal time. This means that a learner's knowledge in a single environment or background enables them to apply the same knowledge in various contexts [1][2]. For instance, a mechanic who specializes in a particular car brand can easily utilize their knowledge to fix faults in almost every other car brand, despite differences in appearance and manufacturers.

In contrast, machine and deep learning algorithms are designed differently. Conventional operation on a particular feature-space distribution is a burdensome task, and building a model from scratch requires a large amount of data and time to train the network. As a result, it is almost impossible to train a network that performs well with only a few classified data for supervised learning [3]. To address the challenge of inadequate training data, time, and computational resources such as GPUs, a technique called Transfer Learning [4], which involves the re-application of pre-trained model knowledge for another application, is used.

The concept of transfer learning in machine learning neural networks was first introduced by Boznovsik and Fulgosi in 1976 [5]. The idea of transfer learning is better explained in Figures 1(a) and (b).

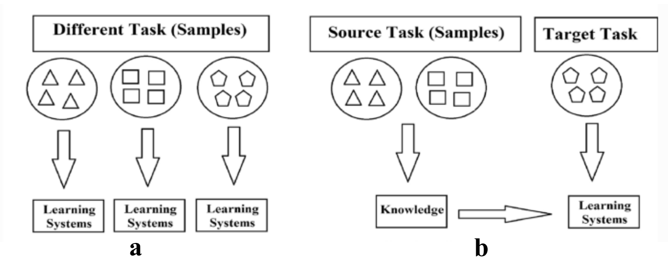

Figure 1. (a) Classical Machine Learning (b) Transfer Learning

Transfer learning can be categorized into various techniques depending on "How What and When" to use it.

The “When” is about when it is suitable to use transfer learning since some sources and the targeted domain may differ which makes its applicability worthless.

The "What" comes into play when it is relevant to know the type of information that will be exchanged between domains. Finally, the "How " is about the procedures to be followed and the algorithms with which the information will be transmitted.

There are many universally known deep learning trained models for image classification, such as Xception [7], ResNet101V2[8], DenseNet, etc. In this study, we decided to use the knowledge transferred from a Keras pre-trained model called VGG16 to classify a dataset of five categorized landscape images called the Intel dataset.

In the next part of this paper, a review of the related work will be discussed in chapter two, the anatomy of the VGG16 model that will be used and the dataset to be applied will be discussed in section 3, the algorithm implementation in section and finally the results and discussion in section 5.

II. Review of Related Works

A lot of widely recognized work has been done on transfer learning in the deep learning section of artificial intelligence. Not only was it done using the Vgg16 model, but also others in the same discipline.

Srikhant Tamina [9] used the Vgg16 model in transfer learning to classify the images of dogs and cats. She compared the performance of the traditional convolution network to that of a fine-tuned Vgg16. The former gave a 72.4% validation accuracy while the latter scored 95.4%, which greatly showed the benefit of using transfer learning.

Similarly, Pardede et al. [10] fine-tuned the Vgg16 model to classify ripened fruits. They adopted the model by removing the last layer and adding a multilayer perceptron that included the drop-out technique. Doing so improved the results by 18.42% compared to the ordinary model. In addition, Jiayi Li et al. [11] used Vgg16 to analyze images of hurricane-damaged buildings in the year 2020, combining satellite images and convolution neural network transfer learning.

In 2021, Ming-Ai Li et al. [12], in the field of brain computers (BCT) that helps the disabled to recover their neural function, showed a deficit in the acquisition of adequate images, which resulted in unsatisfactory results. They applied transfer learning with Vgg16 to overcome this problem and succeeded with 96.59% accuracy. In the same year, Sakr et al. [13] adopted three image classification models, Vgg19, vgg16, and InceptionV3, to compete in the detection of cancer-infected and uninfected breasts in women. The three models were subjected to a monographic image analysis society dataset and the experiments showed outstanding results of 98%, 96.8%, and 96.099% for the Vgg19, Vgg16, and InceptionV3 models, respectively.

III. Intel Image Dataset

In this work, transfer learning with Vgg16 is applied to the Keras Imagenet dataset called the Intel Image dataset, which contains six types of natural and artificial scenery. These include buildings, mountains, glaciers, mountains, the sea, and streets. The dataset has an approximation of 25000 images of 150 × 150 pixels and is indexed as tabulated below.

TABLE I
Image composition of Intel dataset

Image category	Index
Buildings	0
Forest	1
Glacier	2
Mountain	3
Sea	4
Street	5

The training, test, and prediction sets contain 14000, 3000, and 7000 images, respectively. Samples of the dataset images are shown in Figure 2.

Figure 2. Intel Image Dataset image samples

A. The anatomy of the Vgg16 model

The visual graphic group commonly known as Vgg16 [14] is one of the greatest milestones in the history of image vision. This model was proposed by Simanyon and Zisserman at the University of Oxford. The model became famous after winning the 2014 ImageNet Challenges. Figure 3 shows the general structure of the Vgg16 model.

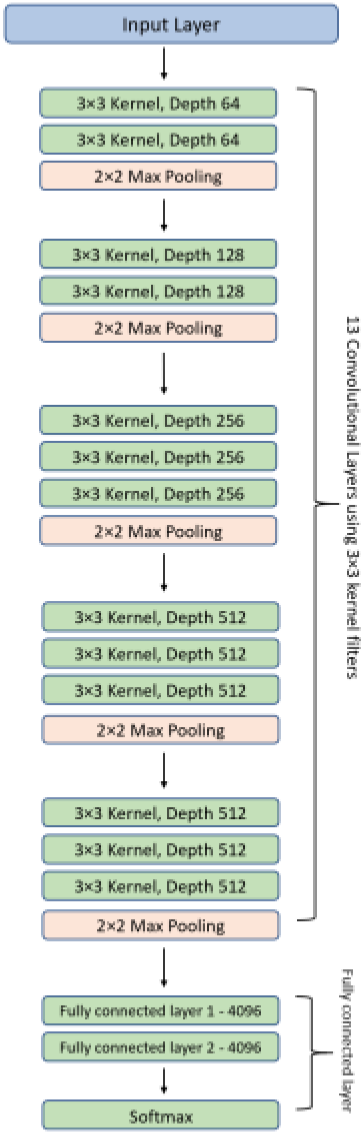

Figure 3. Vgg16 structure

This is a widely used architecture used for ImageNet. It is designed to take an input image of 224 x224 and finally outputs an array of size 1 × 1000. The model consists of 16 layers, 13 of which are convolutions, three are fully connected, and five are pooling layers. Furthermore, the convolution layers have a 3 × 3 filter with a 1 px stride. Additionally, the model has a total of 134,285,126 parameters, 24,582 of which are nontrainable.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 block1_conv1 (Conv2D)       (None, 224, 224, 64)      1792

 block1_conv2 (Conv2D)       (None, 224, 224, 64)      36928

 block1_pool (MaxPooling2D)  (None, 112, 112, 64)      0

 block2_conv1 (Conv2D)       (None, 112, 112, 128)     73856

 block2_conv2 (Conv2D)       (None, 112, 112, 128)     147584

 block2_pool (MaxPooling2D)  (None, 56, 56, 128)       0

 block3_conv1 (Conv2D)       (None, 56, 56, 256)       295168

 block3_conv2 (Conv2D)       (None, 56, 56, 256)       590080

 block3_conv3 (Conv2D)       (None, 56, 56, 256)       590080

 block3_pool (MaxPooling2D)  (None, 28, 28, 256)       0

 block4_conv1 (Conv2D)       (None, 28, 28, 512)       1180160
...
Total params: 134,285,126
Trainable params: 24,582
Non-trainable params: 134,260,544
```

Figure 4. Different blocks of Vgg16 with parameters.

As seen from Figure 3 and 4, Vgg16 uses a 3 × 3 kernel size at every layer except for the input, output, and max pooling layers. The depth of the convolution layers was double incremented from 64 to 512. At least after every two convolution layers, a 2 × 2 Max Pooling layer is applied. The last two layers are fully connected dense layers with softmax as the activation function.

B. Data Preparation and Model Fine Tuning.

As mentioned above, the dataset contains close to 25000 images of RGB nature. Because this work aims to transfer the knowledge of the Vgg16 model to be applied to the classification of different sets of images that it has never been applied to before, only 3900 images were randomly selected equally from all categories. That is, 500 × 6, which makes a total of 3000 images for training, 600 images with 100 in each category for validation, and 300 images with 50 from each category for testing. This block of images is subjected to a transitional process from 150 × 150 to 224 × 224, for which the Vgg16 model was designed. This was performed using the Keras imageDataGeneration model.

The second step was to fine-tune and adopt the model. This was done by freezing some of the model layers and then changing the last multi-perceptron layer to form 1000 original outputs to suit the desired six categories of the dataset, as shown in Figure 5.

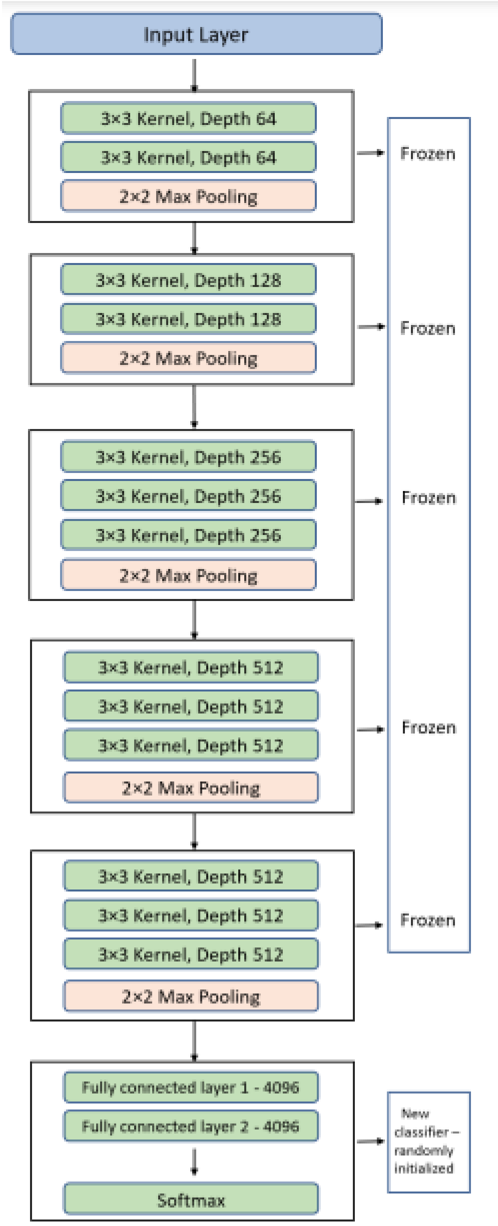

Figure 5. The structure of Vgg16 model

The model is turned into a sequential model, except for the last layer. To avoid changes in the weights of the already trained Vgg16 model, all layers were frozen by having their trainability feature false. Finally, a softmax activation function with six output nodes was used in the last layer.

Vgg16 is a large model in size and complexity; to run very well, it requires very good and powerful computational power, such as the use of GPUs. However, our experiments were performed on a core i3 processor with 4 gigabytes of random-access memory, which is why it was only run for 10 epochs.

Under the same conditions, the adopted dataset was used on a simple CNN with an input layer, two 2DConv layers accompanied by a pooling layer, and two fully connected layers with softmax as the activation function, as shown in Figure 6.

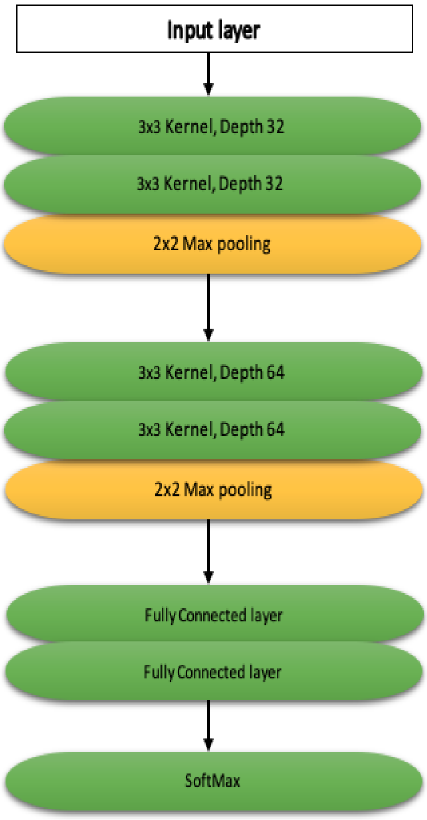

Figure 6. The structure of the ordinary CNN

IV. Experimental Results.

After experimenting with our sample dataset on both the Vgg16 transfer-learned model and the ordinary 7-layer convolutional neural network, the following graphical results were obtained after 10 epochs.

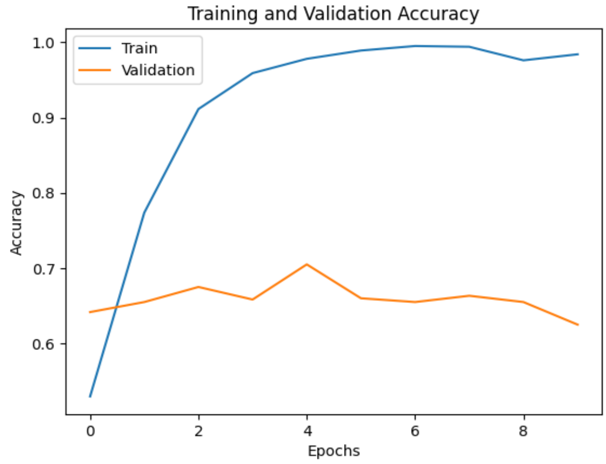

Figure 7. Accuracy of the ordinary CNN

Figure 8. Loss function results of the ordinary CNN

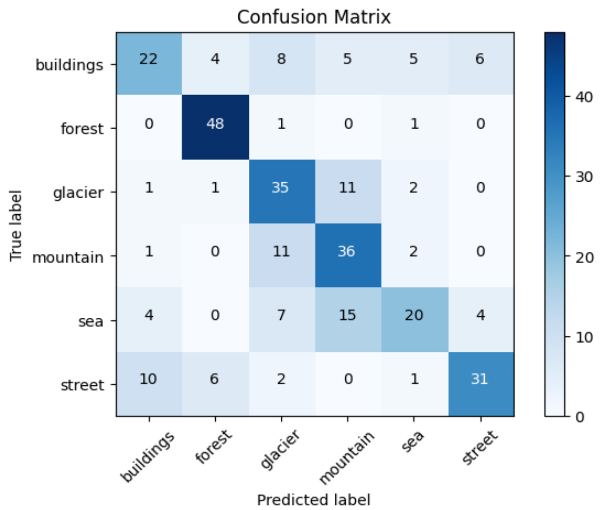

Figure 9. Prediction results of the CNN

From Figures, 6,7, and 8, the validation accuracy of the ordinary CNN training on a small quantity of data was barely beyond 65%, while the training accuracy almost reached 99%, which is a sign of overfitting. Furthermore, the loss was extremely large and continued to rocket as the jumper of the epochs accumulated. Finally, of the 50 true images in each category, the network only predicted 22 buildings, 48 forests, 35 glaciers, 36 mountains, 20 sea, and 31 street images according to the confusion matrix image in Fig 9.

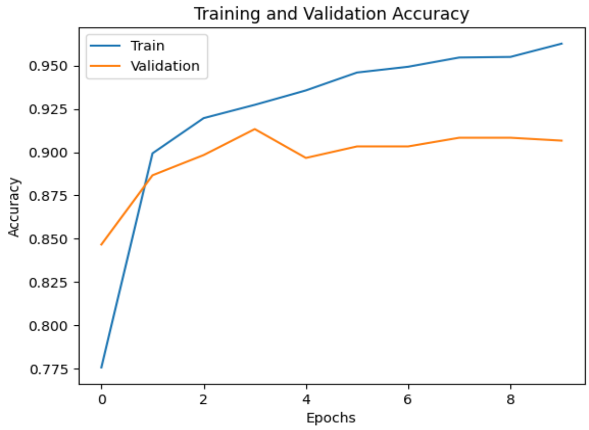

Figure10. Accuracy of the Vgg16 Transfer Learning

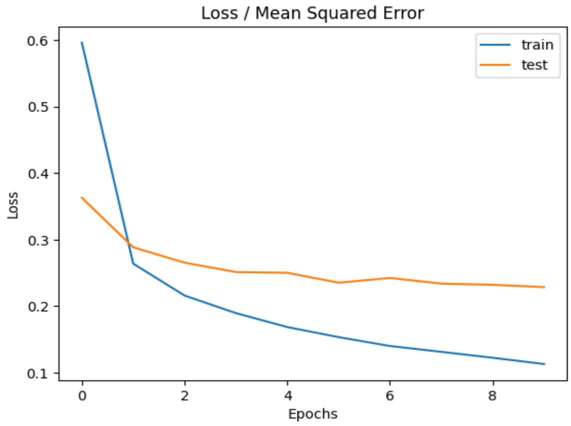

Figure 11. Loss function results for Vgg16 model

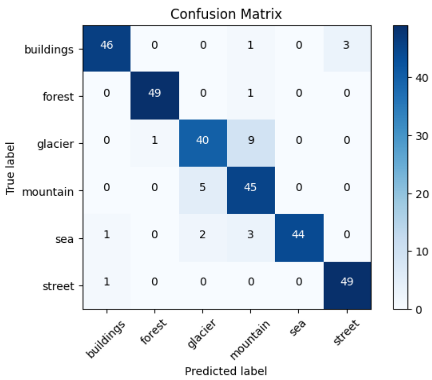

Figure 12. Confusion matrix for Vgg16 model

With a close look at the graphical results from Figures 10, 11, and 12, the validation accuracy exceeded 90.1% in just 10 epochs, while the loss gradually descended below 0.27. At the testing stage, the model correctly predicted 45 buildings, 49 forests, 40 glaciers, 45 mountains, 44 seas, and 49 streets out of the 50 images in each category.

V. Discussion and Conclusion

From the results above, it can be concluded that the network where transfer learning with Vgg16 was applied performed better than that of the ordinary CNN by almost a 30% difference under similar conditions, as shown in Table 2 below.

TABLE 2.

EXPERIMENT RESULT COMPARISON

Model	**Average Time Epochs**	**Accuracy**	**MSE**
Ordinary CNN	750 seconds	62.50%	1.9444
Vgg16 with TL	3100 seconds	90.10%	0.2723

The only advantage that small ordinary CNN models have over large ones such as Vgg16 is their requirement of a small time to execute the task, as seen in the table where Vgg16 takes almost five times the time required by an ordinary CNN to run a simple epoch.

In conclusion, in situations where the availability of training data is limited, transfer learning emerges as a valuable tool. This is due to the fact that pre-trained models like VGG16 possess pre-existing knowledge of specific object features and finely tuned parameters, resulting in highly favourable outcomes.

REFERENCES

[1] Jackson, Denise & Fleming, Jenny & Rowe, Anna. (2019). Enabling the Transfer of Skills and Knowledge across Classroom and Work Contexts. Vocations and Learning. 12. 1-20. 10.1007/s12186-019-09224-1.

[2] De Corte, Erik. (2003). Transfer learning is the Productive Use of Acquired Knowledge, Skills, and Motivations. Current Directions in Psychological Science. 12. 142-146. 10.1111/1467-8721.01250.

[3] Khamis, Azme. (2001). The Effects of Outliers Data on Neural Network Performance. Journal of Applied Sciences. 1394-1398. 10.3923/jas.2005.1394.1398.

[4] Sharma, Chandeep. (2022). Transfer Learning and its Application in Computer Vision: A Review.

[5] Bozinovski, Stevo. (2020). Reminder of the First Paper on Transfer Learning in Neural Networks, 1976. Informatica. 44. 10.31449/inf.v44i3.2828.

[6]

[7] Chollet, Francois. (2017). Xception: Deep Learning with Depth Wise Separable Convolutions. 1800-1807. 10.1109/CVPR.2017.195.

[8] Bilal, Muhammad & Maqsood, Muazzam & Yasmin, Sadaf & Ul Hasan, Najam & Rho, Seungmin. (2022). A transfer learning-based efficient spatiotemporal human action recognition framework for long and overlapping action classes. The Journal of Supercomputing. 78. 10.1007/s11227-021-03957-4.

[9] Tammina, Srikanth. "Transfer Learning Using VGG-16 with Deep Convolutional Neural Network for Classifying Images." International Journal of Scientific and Research Publications (IJSRP) 9.10 (2019): p9420. Web.

[10] Pardede, Jasman & Sitohang, Benhard & Akbar, Saiful & Khodra, Masayu Leylia. (2021). Implementation of Transfer Learning Using VGG16 on Fruit Ripeness Detection. International Journal of Intelligent Systems and Applications. 13. 52-61. 10.5815/ijisa.2021.02.04.

[11] J. Li, Z. Liang, and C. Xiao, "Transfer Learning Performance Analysis for VGG16 in Hurricane Damage Building Classification," 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Zhuhai, China, 2021, pp. 177-184, doi: 10.1109/ICBASE53849.2021.00041.

[12] M. -A. Li and D. -Q. Xu, "A Transfer Learning Method based on VGG-16 Convolutional Neural Network for MI Classification," 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 2021, pp. 5430-5435, doi: 10.1109/CCDC52312.2021.9602818.

[13] Saber, Abeer & Sakr, Mohamed & Abou-Seida, Osama & Keshk, Arabi. (2021). A Novel Transfer-Learning Model for Automatic Detection and Classification of Breast Cancer Based Deep CNN. Kafrelsheikh Journal of Information Sciences. 2. 1-9. 10.21608/kjis.2021.192207.

[14] Sugata, T & Yang, C. (2017). Leaf App: Leaf recognition with deep convolutional neural networks. IOP Conference Series: Materials Science and Engineering. 273. 012004. 10.1088/1757-899X/273/1/012004.