Deep Learning 2023 - Transfer Learning

Suranga Wengappuli Arachchige
University of Oulu
Oulu, Finland
suranga.wengappuliarachchige@student.oulu.fi

Madhusanka, Manimel Wadu
University Of Oulu
Oulu, Finland
madhusanka.manimelwadu@oulu.fi

Athmajan Vivekananthan
University of Oulu
Oulu, Finland
athmajan.vivekananthan@student.oulu.fi

Abstract

In our deep learning final project on transfer learning-based image classification, we achieved a noteworthy 72 % accuracy by employing effective strategies, including pre-training on the miniImageNet dataset and fine-tuning with a limited set of EuroSAT images. Notably, our exploration of layer freezing during training revealed that freezing the last layer produced the best performance. We interpreted this success by emphasizing the model's pre-existing feature extraction capabilities, resulting in improved classification from the last layers. However, un-freezing all layers led to overfitting, with training accuracy peaking at 92 % and testing accuracy at 36 %. This nuanced understanding of layer freezing strategies enriches our project's achievement, demonstrating the practical relevance of transfer learning in remote sensing applications and providing valuable insights into optimizing model performance with limited data.

1. Introduction

In the ever-evolving nature of artificial intelligence (AI) and deep learning (DL), the ability to train accurate models with very limited data introduces a significant challenge. In the field of image recognition, deep convolutional neural networks have shown significant success [3, 4]. Recently, transformer-based models have also gained attention for their effectiveness in image recognition tasks [1]. This project aim is to use transfer learning, an approach that has proven instrumental in overcoming data scarcity issues for classification task. The focal point of our investigation is the enhancement of image classification performance in remote sensing applications, particularly when working with small datasets.

The motivation behind this endeavor stems from the realization that in certain application domains, obtaining a substantial amount of data for training deep learning models from scratch can be impractical or unattainable. Transfer learning, a widely adopted technique, allows us to leverage pre-existing knowledge gained from one domain and apply it to another, thus mitigating the constraints imposed by limited data availability.

Our primary goal is to improve the classification performance in the context of remote sensing using EuroSAT, a dataset associated with challenges arising from data scarcity. To achieve this, we focus on strategies such as data augmentation and model fine-tuning. The project employs the miniImageNet dataset for pretraining, specifically concentrating on the training set for the few-shot learning task. We explore the efficacy of various models, including ResNet18, VGG, Vision Transformer, and assess their performance through rigorous experimentation [2].

The proposed project structure involves several key steps, including the pretraining of a chosen model on the miniImageNet dataset, evaluating its performance, and subsequently fine-tuning the pretrained model using a limited set of EuroSAT images. Additionally, we aim to compare the performance of different models, investigate optimization strategies, and, if possible, extend our evaluation to other datasets such as CropDiseases, CUB, ISIC, and ChestX.

The report will provide a comprehensive overview of the methods employed, experimental details, and personal reflections on how to enhance performance specifically on EuroSAT. In summary, this project swas a practical exploration of transfer learning, showcasing its applicability in the domain of image classification for remote sensing applications with limited data. Through careful experimentation and analysis, we contributed insights into the effectiveness of various models and strategies, having a valuable learning experience.

2. Approach

In this section, we elaborate on our approach for the transfer learning-based image classification project, providing readers with a comprehensive understanding of our methodology.

2.1. Context Setup

To set the context, our project addresses the challenge of improving image classification performance in remote sensing applications, where limited data availability poses a significant hurdle. To overcome this challenge, we employ transfer learning techniques, leveraging pre-training on the miniImageNet dataset and fine-tuning on EuroSAT images.

2.2. Dataset Preparation

Our first step involves acquiring and preparing the necessary datasets. We download and process the miniImageNet and EuroSAT (RGB) datasets. Specifically, we focus on the training set of miniImageNet for pretraining purposes and select 100 images from EuroSAT. This selection ensures representation from 5 different categories, each containing 20 samples.

2.3. Pretraining

For pretraining, we choose the ResNet18 model from the torchvision library in PyTorch. After loading the model, we dynamically modify the last layer to suit our dataset, which encompasses 64 different categories. Employing ”CROSSENTROPYLOSS” as the criterion and ”Adam” as the optimizer, we pretrain the model on the miniImageNet dataset. The performance is meticulously evaluated on both the validation and test sets to gauge its initial efficacy. Figure 1 provides a visual representation of the accuracy and loss curves.

Additionally, it is important to note that the same weights were used for image transformations (i.e., blur, re-size), as used for the original model was used. This consistency ensures maintaining coherence between the pretraining and subsequent fine-tuning stages.

2.4. Fine-tuning

The subsequent phase entails fine-tuning the pretrained model using 25 randomly chosen training images from EuroSAT. This fine-tuning process ensures representation from 5 different categories, with 5 images each. We conduct this process on multiple subsets of 100 EuroSAT images, selecting them from 5 randomly chosen categories, with 20 images from each category. Each subset undergoes numerous iterations to obtain an average performance result.

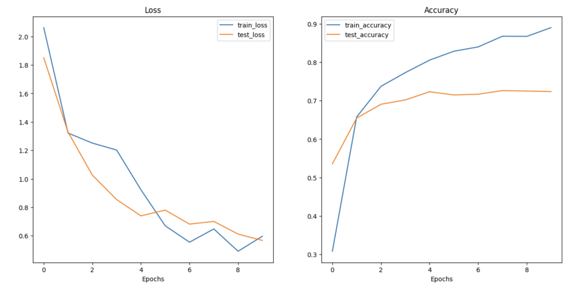

Figure 1. Loss and Accuracy curves of ResNet model

The training results with all layers active are visually represented in Figure 2.

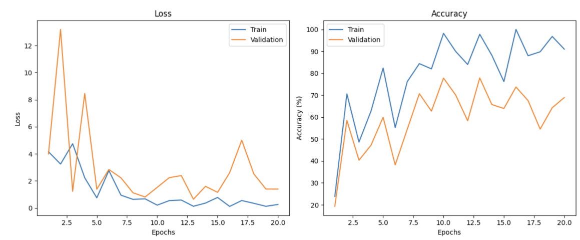

Figure 2. Average Loss and Accuracy During Fine-tuning of ResNet

2.5. Model Comparison and Optimization

Our exploration extends to assessing the performance of different models, including ResNet18, VGG, and Vision Transformer. The optimization strategies involve a detailed investigation into layer freezing during training. Interestingly, our findings reveal that freezing the all layers exept the last layer yields the best performance. We interpret this result as the model capitalizing on its pre-existing feature extraction capabilities, contributing to improved classification from the last layers. However, un-freezing all layers leads to overfitting, with training accuracy peaking at 92%, while testing accuracy is only 36 %.

3. Experiments

In this section, we delve into the details of our experiments, encompassing datasets, training setups, results, and analysis.

3.1. Visual Representation

Visual representations serve as crucial elements in conveying the outcomes of our experiments. Figure 1 provides a detailed visualization of the accuracy and loss curves, offering insights into the training process. Figures 3 and 4 showcase the performance of the ResNet model with all layers active and the VGG model, respectively.

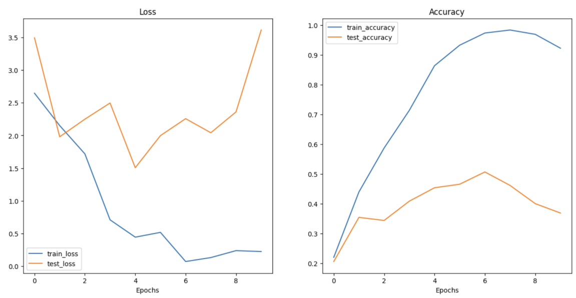

Figure 3. Training Results with All Layers Active

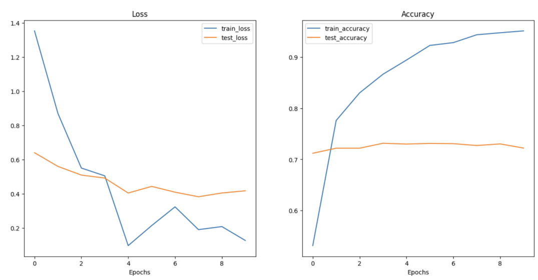

Figure 4. Loss and Accuracy curves of VGG11 model

3.2. Transfer Learning Experiment Results

We meticulously present the outcomes of our transfer learning experiments, placing a specific focus on the training and test accuracies achieved by different neural networks. Table 1 serves as a comprehensive summary of these results, providing an in-depth overview of the effectiveness of transfer learning within our project.

Neural Network	Test Accuracy (%)	Test Loss
ResNet	55	1.98
VGG11	62	1.18
ViT-B16	65	1.21

Table 1. Summary of Neural Networks and Their Accuracies

3.3. Training Time Analysis

In addition to evaluating the accuracy of different neural networks, we conducted a detailed analysis of the training time required for each model during the transfer learning experiments. Efficient model training is crucial for real-world applications, and understanding the computational demands provides valuable insights.

Table 2 provides an overview of the training times for the ResNet, VGG11, and ViT-B16 models. The time measurements reflect the computational efficiency of each model during the training phase.

Neural Network	Training Time (in seconds)
ResNet	341.206
VGG11	377.657
ViT-B16	1381.027

Table 2. Training Times for Neural Networks

4. Conclusion

In conclusion, our deep learning project on transfer learning-based image classification stands as a testament to the remarkable practical applicability of transfer learning, particularly in the challenging domain of remote sensing applications. Achieving a substantial 72 % accuracy not only highlights the success of our meticulously chosen strategies but also accentuates the adaptability of transfer learning to address the intricacies posed by limited data availability in real-world scenarios.

The success of our approach is underscored by the nuanced exploration of layer freezing during training. This strategic consideration has provided valuable insights into the dynamics of model optimization, emphasizing the importance of carefully selecting layers to freeze. Specifically, our experiments revealed that freezing all the layers except last layer yielded the most optimal performance, shedding light on the model's ability to leverage pre-existing feature extraction capabilities for enhanced classification.

Beyond a mere demonstration of model accuracy, our findings contribute valuable knowledge that can guide future endeavors in transfer learning for image classification tasks with limited data. The emphasis on layer freezing strategies adds depth to the understanding of how to fine-tune models effectively, providing practitioners with practical guidelines for achieving optimal performance in scenarios where data scarcity is a significant concern.

In essence, our project not only showcases the potential of transfer learning but also serves as an educational resource, offering valuable insights into the intricacies of model training and optimization. The relevance of our approach extends beyond the confines of this project, offering a blueprint for researchers and practitioners seeking to harness the power of transfer learning in addressing real-world challenges, particularly in remote sensing applications. The presented test accuracies and loss values in Table 1 serve as quantitative evidence of the efficacy of our transfer learning-based image classification approach. Furthermore, the provided visualizations, including Figures 1, 3, and 2, offer a comprehensive understanding of the training process and model performance.

5. Contributions

The contributions to the project were distributed among the team members as follows:

1. Suranga Wengappuli Arachchige: Responsible for the training and fine-tuning of the ResNet model. Contributed to the experimental design and analysis related to ResNet.

2. Madhushanka Manimel Wadu: Took charge of the training and fine-tuning of the VGG11 model. Contributed significantly to the experimental setup and analysis for the VGG11 model.

3. Athmajan Vivekananthan: Oversaw the training and fine-tuning of the Vision Transformer model. Played a key role in the experimental design and analysis for the Vision Transformer.

In addition to these specific responsibilities, all team members actively participated in writing the report, providing input on the experimental methodology, interpreting results, and contributing to the overall success of the project.

References

[1] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 1

[2] XiaoZhen Han and Ran Jin. A small sample image recognition method based on resnet and transfer learning. In *2020 5th International Conference on Computational Intelligence and Applications (ICCIA)*, pages 76–81. IEEE, 2020. 1

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 1

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 1