Self-Supervised And Transfer learning On CIFAR10

M114030007 財管碩二 張梓揚

1 Introduction

This report investigates the efficacy of utilizing self-supervised learning as a preliminary training phase for the ResNet18 model, emphasizing the rotation task. We examine the impact of predicting image rotations in an unsupervised context on enhancing feature learning, thereby benefiting subsequent supervised tasks

2 Data preparation

1. Making labels for rotation task

The first thing to do is implement a dataset class to load rotated CIFAR10 images with matching labels. Each rotation label should be an integer in the set {0, 1, 2, 3} which correspond to rotations of 0, 90, 180, or 270 degrees respectively.

2. Image preprocessing

For train and test image, the preprocessing rules as illustrated below:

```
transform_train = transforms.Compose([
        transforms.RandomCrop(32, padding=4),
        transforms.RandomHorizontalFlip(),
        transforms.ToTensor(),
        transforms.Normalize((0.4914, 0.4822, 0.4465), (0.2023, 0.1994, 0.2010)),
])

transform_test = transforms.Compose([
        transforms.ToTensor(),
        transforms.Normalize((0.4914, 0.4822, 0.4465), (0.2023, 0.1994, 0.2010)),
])
```

3 Models

1. **Self-Supervised rotation task model**:
   The input is a rotated image and the model predicts the rotation label, enabling a model to learn meaningful representations or features from unlabeled data, it’s call self-supervised learning.

2. **Fine-tuning on the pre-trained rotation task model for image classification:**
   Load the ResNet18 model pre-trained on the rotation task and fine-tune on the classification task, and freeze all previous layers except for the 'layer4' block and 'fc' layer.

3. **Fine-tuning on the randomly initialized rotation task model for image classification:**
   Randomly initializing a ResNet18 model and fine-tune on the classification task, freezing all previous layers except for the 'layer4' block and 'fc' layer

4. **Supervised training on the pre-trained rotation task model for image classification:**
   Load the ResNet18 model pre-trained on the rotation task and re-train the whole model on the classification task.

5. **Supervised training on the randomly initialized rotation task model for image classification:**
   Randomly initialize a ResNet18 model and re-train the whole model on the classification task.

4 Training

I chose CrossEntropyLoss as the criterion, and for the optimizer, I opted for Adam with a learning rate of 0.001. It's important to note that the learning rate should not be too high, such as 0.005,especially when dealing with deeper models like RESNET34 and RESNET50. This precaution ensures effective model training. Futhermore, we use a learning rate scheduling strategy using a decay factor of 0.1 for every 10 epochs so as to improve convergence and model performance.

5 Results

1. **Self-Supervised rotation task model:**
   The model achieved 78.3% accuracy on the test images after training for 45 epochs. It laid a strong foundation for feature extraction that was instrumental in downstream tasks.

2. **Fine-tuning on Pre-trained vs. Fine-tuning on Randomly Initialized Models:**
   After training for 20 epochs, the pre-trained model achieved 61.88% accuracy, while the randomly initialized model achieved 44.75% accuracy on the image test set.

3. **Full Supervision on Pre-trained vs. Fine-tuning on Randomly Initialized Models:**
   In the fully supervised scenario, the pretrained model achieved an accuracy of 84.03%, while the randomly initialized model exhibited an accuracy of 82.79%.

6 Conclusion

Our experiments delved into different model training methods, particularly focusing on applying self-supervised learning to ResNet18. Here's a concise summary of the findings:

The self-supervised rotation task model achieved a solid 78.3% accuracy on test images, laying a crucial foundation for downstream tasks. This underscores the effectiveness of self-supervised learning in extracting meaningful features from unlabeled data.

Fine-tuning the pre-trained rotation task model for image classification surpassed random initialization, with 61.88% and 44.75% accuracy, respectively, after 20 epochs. This supports the advantage of pre-trained models in enhancing generalization.

In full supervision, the pre-trained model slightly outperformed the randomly initialized model, achieving 84.03% and 82.79% accuracy on the CIFAR10 classification task, respectively.

In conclusion, our experiments highlight the significance of self-supervised learning, particularly through the rotation task, in improving model performance and generalization. Pre-trained models consistently outshine randomly initialized models, emphasizing the benefits of leveraging learned representations for enhanced supervised tasks.

7 Extra: Does Deeper models have better performance ?

1. **Self-Supervised rotation task model:**
   The RESNET18 model achieved 78.3% accuracy on the test images. However, after training for 70 epochs, the performance on RESNET34 was 79.70 %, and 75.3% for RESNET50 , both worse than RESNET18.

2. **Supervised training on the pre-trained rotation task model for image classification:**
   The RESNET18 supervised pretrained model achieved 84.03% accuracy on the test images. However, after training for 70 epochs, the performance on RESNET34 was 80.82 %, and 83.01% for RESNET50 , both worse than RESNET18.