《A4 報告》

學號:B094020007 姓名:陳文薇

1. (13 points) Train a ResNet18 on the Rotation task and report the test performance. Discuss why such a task helps in learning features that are generalizable to other visual tasks.

test performance大約落在78.5%左右。

透過隨機旋轉可以增加資料的多樣性(data augmentation)，比較能防止overfitting的情形，藉此提高generalizable to other visual tasks的能力。

旋轉不同角度有助於generalizable to other visual tasks還有一個原因是，可抓取不同旋轉角度的features，因為其他圖片可能會旋轉不同角度。總而言之，rotation task能幫助model學習到更多features

2. (12 points) Initializing from the Rotation model or from random weights, fine-tune only the weights of the final block of convolutional layers and linear layer on the supervised CIFAR10 classification task. Report the test results and compare the performance of these two models. Provide your observations and insights. You can also discuss how the performance of pre-trained models affects downstream tasks, the performance of fine-tuning different numbers of layers, and so on.

Performance of Fine-tuning on the pre-trained model : test acc around 61%
Performance ofFine-tuning on the randomly initialized model:test acc around 43.6%

Based on the test results, it's obvious that fine-tune on the pretrained model has a better performance. Because pretrained model has learned visual features from dataset, thus if fine tune and use the model to do other related task, it can performs better.(because it has already learned features.)Pretrained model is beneficial to transfer learning(some other similar tasks).
Fine tune on different layers has different effect. Fine tune on lower layer maintain some common features. While fine tune on higher layers can be easier to know the detail of certain task.
If fine tune overly, it might cause overfitting. And because fine tune more layers need more computation resources. thus we should tradeoff and consider which layers to be fine tuned.

3. (12 points) Initializing from the Rotation model or from random weights, train the full network on the supervised CIFAR10 classification task. Report the test results and compare the performance of these two models. Provide your observations and insights.

performance of Supervised training on the pre-trained model:test acc around 84%
performance of Supervised training on the randomly initialized model:test acc around 75%
My insights for compare these two models are similar to the previous question. But it's obvious that the performances are much better than previous two models. if downstream task is not that similar to tasks while pretraining, then the fine tune might not fitting well to the new task, thus the performance is lower than supervised training.In summary, supervised training might fitting new tasks better than fine tune training. supervised training can also learn or tag more features.