Transfer Learning

Basic Image Classification

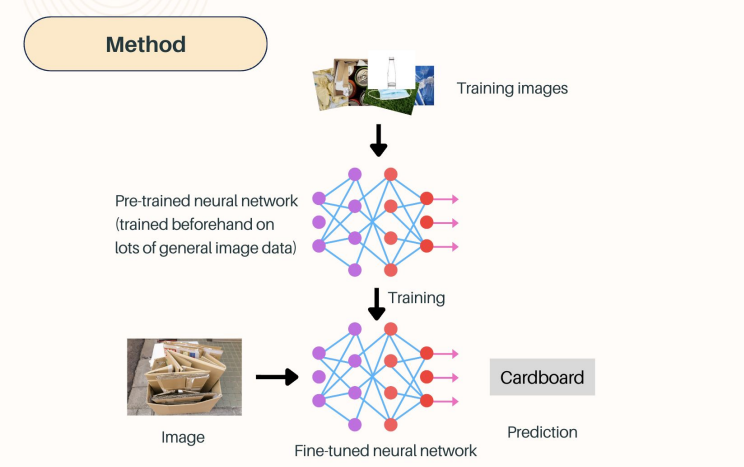
- (continuation of the CIFAR-10 image task)
- It’s not easy to get the accuracy over 85%!
- But when we see the images it’s usually quite obvious to us…
- Why?

Basic Image Classification

- The network has not been trained with enough data!
- Neural networks are data-hungry, and 50000 images (even with augmentation) from a constrained dataset is not enough to produce a quality image classifier
- We need more data?

Transfer Learning

- Solution: “pretrain” the model with a large collection of images (e.g., from ImageNet)



Transfer Learning

- To equip the model parameters with “generic knowledge” about how to extract good features from data
- And then “fine-tune” it on downstream tasks
- To adjust the model parameters from a generic-knowledge state to suit the exact dataset (the downstream task)
- This is transfer learning
- These terms are going to be very important!

Transfer Learning

- ImageNet has over 1 million images and 1000 classes, training on it is going to be very time-consuming!
- Fortunately, there are many “pretrained models” released out there
- e.g., torchvision.models.resnet50(weights=’DEFAULT’) # weights=’DEFAULT’ loads the weights pretrained using ImageNet

Transfer Learning

- Advantage: leverages general knowledge of pretrained models
- Disadvantage: might not matter much for classical machine learning (e.g., tabular data)
- the improvements might be less than expected if the data or task is too specific

Transfer Learning

- Experience using pretrained models for image classification here:
- https://colab.research.google.com/drive/11PQ-g1Ge5058biSuFjB7fj9AijQwOYAT
- https://www.kaggle.com/code/carsoncheng/transfer-learning

Transfer Learning

- For most AI applications (dealing with images, text, audio, etc.), we don't train models from scratch
- Instead, we fine-tune existing models (often foundation models with general knowledge) to fit our specific use cases

Transfer Learning

- Pretrained models may be pretrained on image data (such as ImageNet) and text data (such as the text from the internet), but less likely on tabular data