ML2021Spring HW3 Report

Department of Physics 林展慶 B05202044

Public Score	Private Score
0.84229	0.84219

The methods I used to pass the strong baselines include:

1. VGG16
2. pseudo labeling
3. weighted loss
4. augmentation (random padding + random rotation + color jitter + normalization)
5. transfer learning
6. batchnorm in convolutional layers
7. batchnorm + dropout in dense layers
8. ensemble

My model architecture is almost the same as VGG16. The only difference is that I use a smaller dense layer since the original architecture is too big in a Colab environment.

There are two parts of my workflow. First, I train a model until validation accuracy > 0.79. Secondly, I use the model as the pretrained model to train 5 more models for ensemble. To be more specific, I only use the convolutional layer of the pretrained model to initialize those 5 models, and I also modify the dropout probability from 0.5 to 0.8.