

Homework 3 Report - DeepQ

B05901022 電機三 許睿洋

1. 請分項說明你的 `train.py` 中所有的設定，包含

(a)做了那些比較的實驗

(b)最後選擇此值 / 此 `scheduling` 的原因

(1) `Pretrained=False/True`

一旦使用了 `Pretrained`，便可以直接過 `simple baseline`。

(2) `Optimizer=SGD(lr=0.1/0.01) / Adamax(lr=0.002/lr=0.001)`

經過測試，感覺 `SGD` 的表現還是較為穩定，而且一開始就可以衝到很高的數值。在 `SGD` 的 `learning rate` 中，感覺還是原本設定的 0.1 表現更為出色。

(3) `Batch size=32/16`

將 `Batch size` 從 32 下降到 16 在實驗中可以將表現變得更好。

(4) `Data Augmentation:`

`transforms.RandomAffine`

`transforms.RandomHorizontalFlip`

`transforms.RandomCrop`

在這三項 `Data Augmentation` 的方法中，`RandomAffine(degree=5, translate=(0.02,0), scale=(0.8,1.2), shear=5)`跟 `RandomHorizontalFlip(0.5)`的效果都能有很好的結果。因為原本的圖形是環境圖，在 `RandomAffine` 控制的範圍中，這樣的變化就能有不錯的表現(試過將範圍變大，但其結果較為不好，可能是因為 `Data set` 的圖形沒有嚴重扭曲的環境圖)；

`RandomHorizontalFlip` 也是較為合理的圖形，故而擇之。`RandomCrop` 的部分可能是沒有試出一個效果成功的參數，因此最後放棄使用。

(5) `Batch norm=None/0.1`

將 `Batch norm` 的值變成 0.1 後，同時配合 `Pretrained` 與 `Batch size=16` 可以有機率性的推過 `Strong Baseline(90.8% acc)`，因此選擇此項改變。

(6) `Dropout(mobilenetv2)=None/0.1/0.2/0.3/0.7`

`Dropout` 如若設成 `None`，則會 `Default` 成 0.5 的數值。0.1/0.2/0.3 的效果都會惡化結果，而 0.7 可以稍微提升準確率，故而擇之。

(7) `Epoch=(50 for resnet50)/(40 for mobilenetv2)`

經過觀察三次 `training` 的 `val acc` 數值，發現 `resnet50` 最佳的結果很有可能落在最末次結果，而 `mobilenetv2` 則會在 `epoch=40` 附近來到最佳解，並在 `epoch=50` 處嚴重惡化，故而擇之。