

HW3 Intro - Hyperparameter Tuning Report

B05902074 資工三 魏佑珊

a. 實驗過程：

一開始先把 `pretrained = True`，`weight_decay = 0.0001`，效果不彰。

便決定修一下 `batch_norm`，我試過 `batch_norm = 0.91`，`0.93`，`0.97` 左右，覺得 `0.93` 最適合，此時 `resnet_50` 的 `validation` 正確率大概可以到 `0.89` 左右。後來決定再調一下 `batch_size`，我試過將 `batch_size` 往上調，用 `64`，可是正確率反而下降，因此試著把 `batch_size` 往下調，用到 `20` 時，`resnet_50` 的 `validation` 正確率大約可以到 `0.90` 左右。

但此時 `mobilenetv2` 的正確率仍大約在 `0.87` 左右，於是開始著手調 `mobilenetv2` 的正確率。我將 `batch size` 往下調到 `12`，此時 `validation` 正確率最高可到 `0.88` 附近。希望再將正確率提升一些，因此在 `transform` 的地方傳入 `4*244*244` 的圖片(原本沒有裁切圖片，直接傳入 `244*244`)，但沒有顯著提升。正確率還是在 `0.88`。

b. 最後選擇此值 / 此 scheduling 的原因

我最後選擇的參數值是

Resnet50:

`random seed 7122` (沒有特別調，以前在其他程式用過就沿用了)

`pretrained: True, weight_decay: 0.0001`

`optimizer = GradientDescentOptimizer`(有試過想用 `ada` 等等其他的 `optimizer`，但似乎用了就會出現執行錯誤，就不用了)

`batchsize=20, transform 不變`

`lr: 0.01, batch_norm: 0.93, drop_out: None` (不修改 `drop_out`，因為我覺得雖然 `validation accuracy` 還是偏低，但 `drop out 0.5` 以經算蠻多了，就不再加了)

Mobilenetv2: 同上，`batchsize` 改成 `12`

Ps. 因為 `resnet50_best` 中的 `variables` 中的 `variables.data-00000-of-00001` 較大，我用 `split` 切分它了，需用 `cat variables.data-00000-of-00001.* > variables.data-00000-of-00001` 來接回