

# Machine Learning HW8 Report

學號：R07943107 系級：電子所碩一 姓名：徐晨皓

## 1. 請比較你本次作業的架構，參數量、結果和原 HW3 作業架構、參數量、結果做比較。(1%)

### ● 本次作業(HW8)

#### (1) 模型架構

本次作業使用 PyTorch 實作。Model 分為兩部分，CNN 與 DNN。在 CNN 部分，4 層 pointwise convolution layers，6 層 depthwise convolution layers，以及 3 層 max pooling layers 交錯組成。DNN 部分使用兩層 linear layers。其中 activation function 使用 relu，層與層之間也有使用 batch normalization。此模型有使用 dropout 來避免 overfitting。

#### (2) 參數量：102,700

#### (3) 結果

Private score: 0.62357

Public score: 0.62412

### ● 原 HW3

#### (1) 模型架構

本次作業使用 keras 實作。Model 分為兩部分，CNN 與 DNN。首先，由五層 convolution layers 與三層 max pooling layers 交錯組成 CNN 的部分，每層 convolution layer 都有使用 zero padding 的技術，讓圖片不會損失原有大小。接著再以三層 fully connected layers 組成 DNN 部分，最後再接上 output layer。除了 output layer 的 activation function 為 softmax，convolution layer 與 dense layer 都使用 relu 作為 activation function。此模型有使用 dropout 來避免 overfitting。

#### (1) 參數量：7,255,815

#### (2) 結果

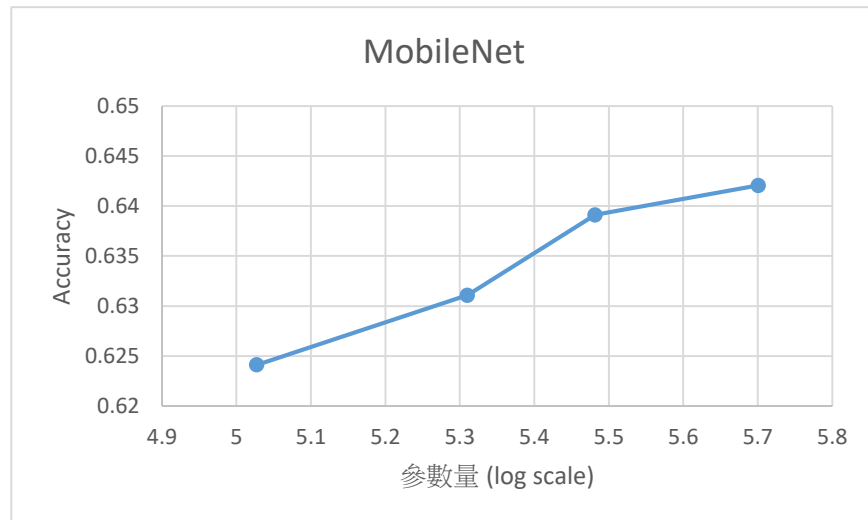
Private score: 0.70103

Public score: 0.71217

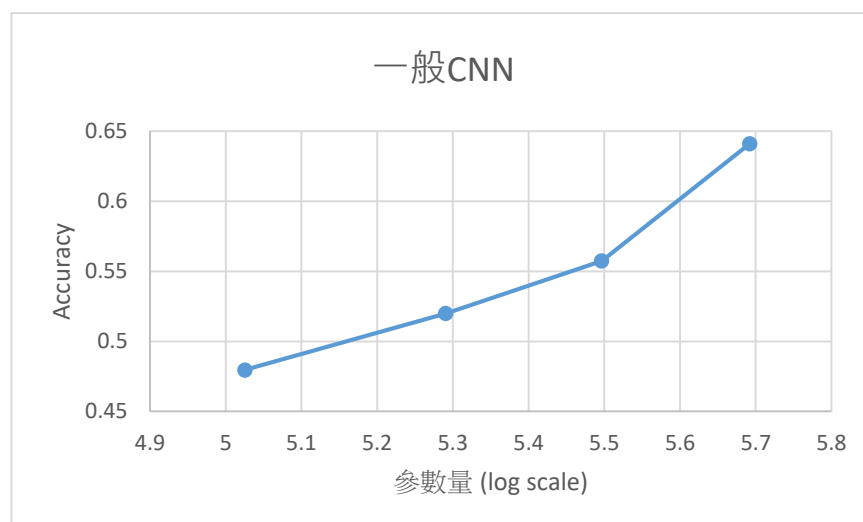
### ● 兩者比較

在模型參數量方面，HW3 的約有七百萬，然而 MobileNet 只有約十萬，相差約 70 倍。在預測準確率方面，HW3 有約 0.70 的準確率，而 MobileNet 有約 0.62 的準確率，相差約 8 的百分點。由此可見，儘管 MobileNet 只有非常少量的參數，但準確率沒有大幅下降。

2. 請使用 MobileNet 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。）（1%）



3. 請使用一般 CNN 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。）（1%）



4. 請你比較題 2 和題 3 的結果，並請針對當參數量相當少的時候，如果兩者參數量相當，兩者的差異，以及你認為為什麼會造成這個原因。(2%)

● 實驗設定

- (1) 使用的參數量約為 10 萬、20 萬、30 萬、50 萬。
- (2) 兩模型都只訓練約 30 至 40 個 epochs。
- (3) 兩模型訓練皆有使用 data augmentation。

● 兩者差異與原因探討

從下圖可以看出隨著參數的增加，一般 CNN 的準確率有大幅的提升，然而 MobileNet 的準確率卻只有緩慢的上升。我認為這是因為當一般 CNN 有較多的參數時，其 convolution 的範圍或方式更多元，因此能更偵測到更多的圖片特徵。反觀 MobileNet，其使用 depthwise convolution 因此我認為這樣的 filter 多樣性不及原來的多樣性，導致 MobileNet 隨著參數增加，準確率卻沒有大幅增加。

另外，我們可以看到，當參數量約為 10 萬時，一般 CNN 準確率只有約 0.47，然而 MobileNet 確有 0.62，這表示 MobileNet 這樣的結構在參數量少時，也能有不錯的準確率，不像一般 CNN 參數量減少時，準確率會大受影響。

