

學號：B05902128 系級：資工三 姓名：鄭百凱

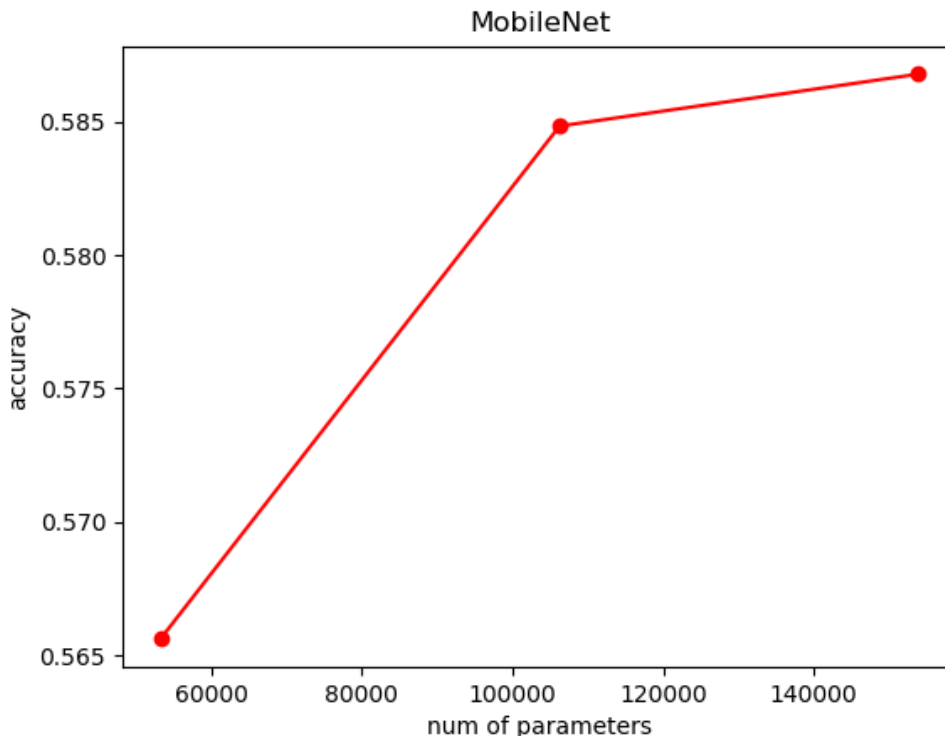
1. 請比較你本次作業的架構，參數量、結果和原 HW3 作業架構、參數量、結果做比較。(1%)

我這次作業疊了 4 層 CNN，channel 數量是 $1 \rightarrow 24 \rightarrow 32 \rightarrow 48 \rightarrow 64$ ，然後接一層 DNN 到 7 維的 output。每一層 CNN 的 kernel size 都是 3×3 ，且之後都有接一個 2×2 的 pooling，也有做 batchnormalize。Data 的部分則是有做 augmentation，其中包含隨機左右翻轉、旋轉正負 30 度、放大 0.8~1.2 倍，batchsize 是 32 (batchsize 小一點對於 public score 有顯著的進步，但會稍微降低 private score)。Train 1000 個 epoch 之後，用 model.half() 將 model 的參數從 float32 變成 float16，所以 model 的大小從原本的 386KB 變成 196KB。而在 kaggle 上的 Public score 是 0.64001。

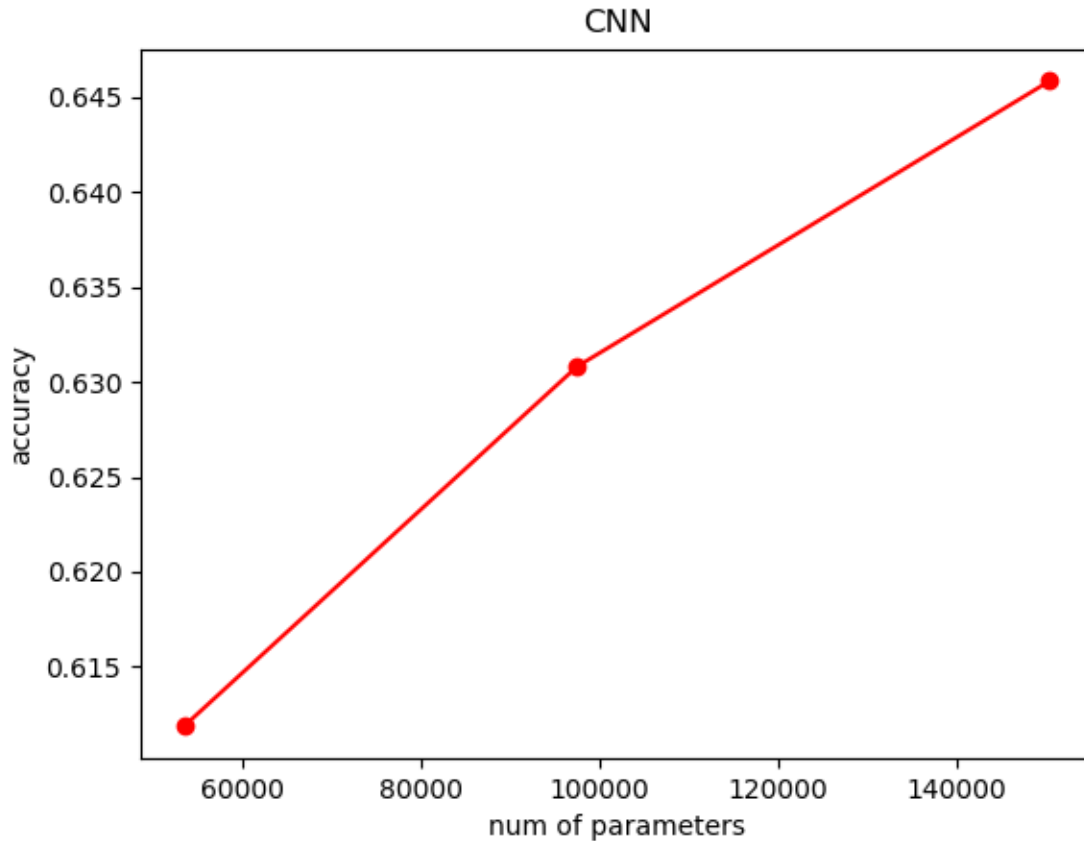
HW3 時，我則是疊了 6 層 CNN，channel 最多到 512 個，還有 3 層 DNN。Kernel size 跟 augmentation 的部分都跟本次作業相同。Epoch 次數跟 batchsize 分別是 600 和 100。最後 Model 的大小是 44MB，在 kaggle 上的 Public score 是 0.66870。

兩者相比，本次作業 model 的大小差了超過 100 倍，準確率才差 2.8% 左右。

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%)

當參數量相當少且參數量相當時，mobilenet 皆輸給 CNN 5 個百分點左右。我認為是因為我們這次作業的圖片很小，所以 CNN model 的參數可以不用太多，就有不錯的效果。根據 <https://github.com/marvis/pytorch-mobilenet/blob/master/main.py>，可以看到這篇的 mobilenet 的原始圖片大小是 224*224，所以或許當原始圖片的大小夠大時，在參數量相當少且參數量相當的條件下，mobilenet 的表現會比 CNN 好。