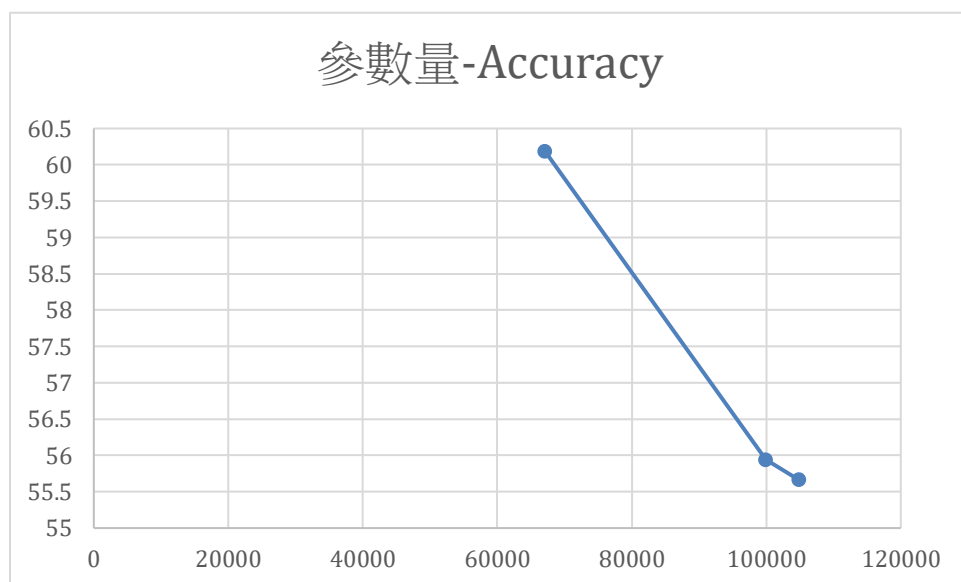


學號：B04501073 系級：土木四 姓名：李利元

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%)

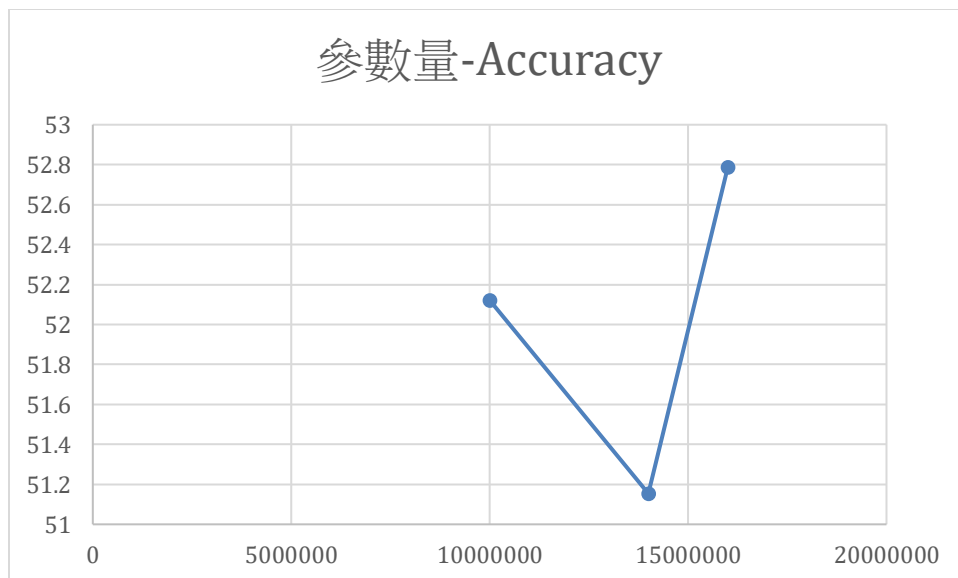
這次作業參考 sample code 給出的連結來建出我的 model，在 cnn 架構上少了 relu、dropout 和 batchnormalization，而最後 dnn 的部分，作業 3 我總共有 5 層 dnn，相較在這次作業的 model 上我只有使用一層，參數量部分：HW3 有 14,000,000 左右，HW8 有 99,911，準確率：HW3 大約 0.69，HW8 約 0.63。

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%)



以上 model 都接近 train 到 30 個 epoch，所以可以明顯看出參數量越少的模型反而準確率越高，因為他能更快的趨近收斂，但若真正 train 到每個模型接收斂的話，我相信結果應該完全相反。

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%)



由於時間關係，CNN 大概只 train 了 20 個 epoch，基本上也還看不出 model 之間的準確率哪個比較好，因為這三個都還沒到收斂的狀態，準確率都很差

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%)

題 2 和題 3 由於沒有 train 到完全收斂而且參數量並未控制在同一範圍所以不能直接比較，但在參數量相當且非常少的時候，我認為兩者應該是沒有差別的，因為其實 MobileNet 其實就是一般 CNN 架構的最精簡狀態，所以我認為如果一般 CNN 架構要做到縮減參數至一定的量，他的架構會變得跟 MobileNet 非常相似，因此在表現上應該也不會有太大的差別。