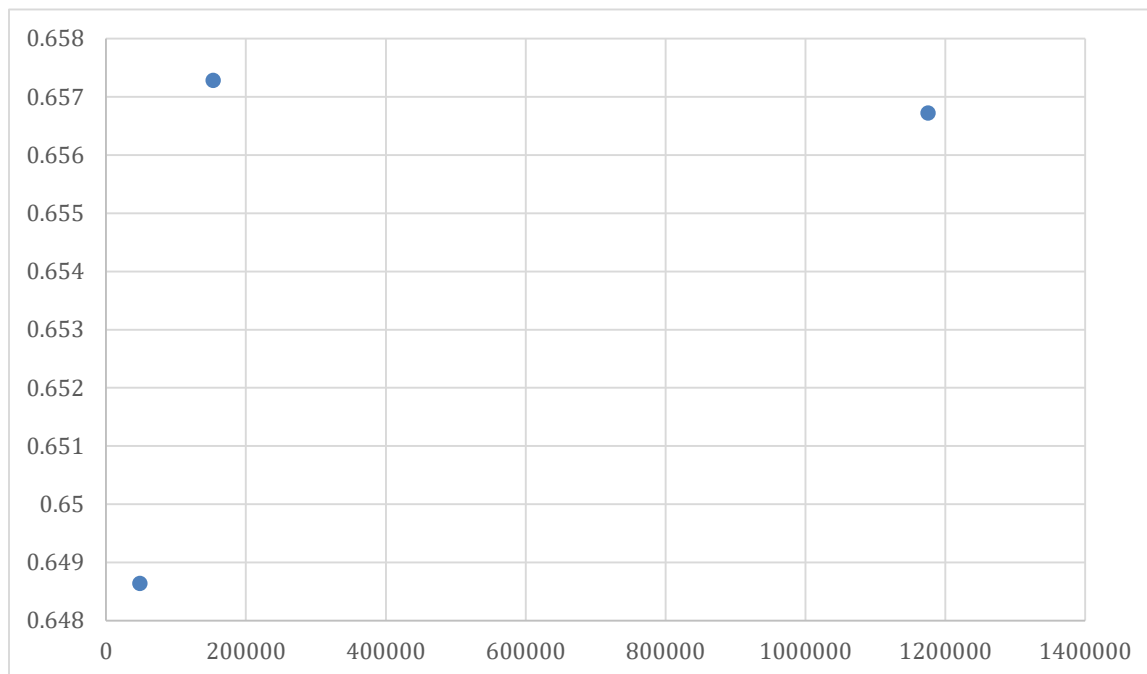


學號：R07922104 系級：資工碩一 姓名：林傳祐

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%)

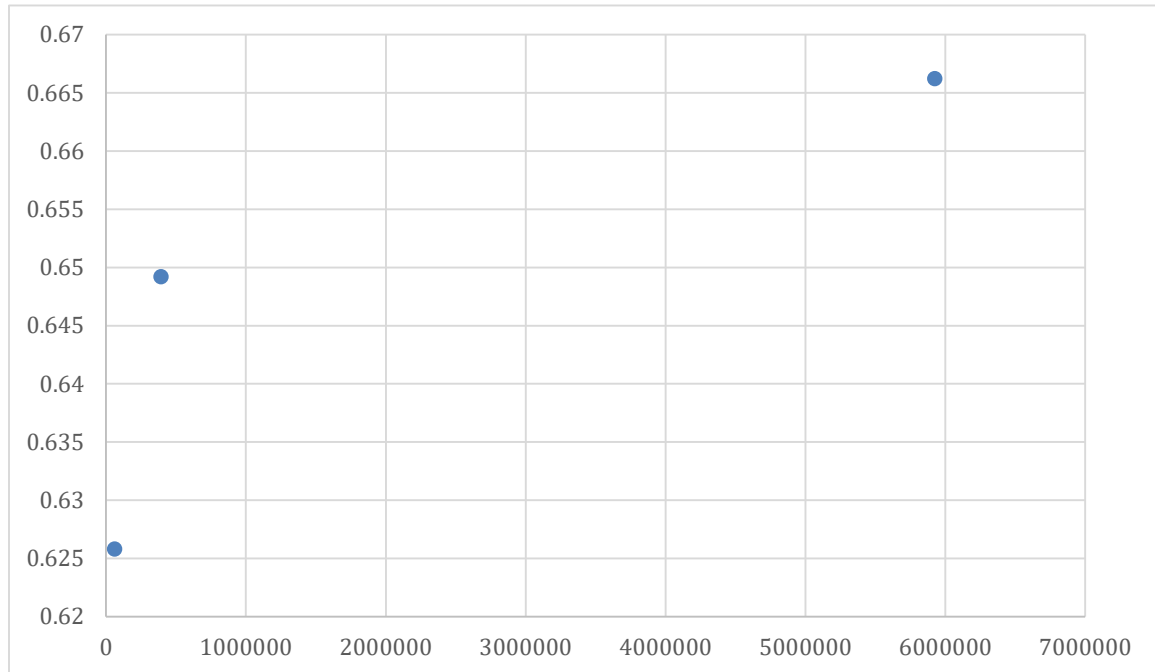
	MobileNet	CNN
架構	Conv2d(1,16), kernel_size = 3*3 (Depthwise)Conv2d(16,16), kernel_size = 3*3, stride = 2 Conv2d(16,32), kernel_size = 1*1 (Depthwise)Conv2d(32,32), kernel_size = 3*3, Conv2d(32,32), kernel_size = 1*1 (Depthwise)Conv2d(32,32), kernel_size = 3*3 Conv2d(32,32), kernel_size = 1*1 (Depthwise)Conv2d(32,32), kernel_size = 3*3, stride = 2 Conv2d(32,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3 Conv2d(64,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3 Conv2d(64,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3, stride = 2 Conv2d(64,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3 Conv2d(64,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3 Conv2d(64,64), kernel_size = 1*1 (Depthwise)Conv2d(64,64), kernel_size = 3*3, stride = 2 Conv2d(64,128), kernel_size = 1*1 (以上皆含 relu,batchnorm) Avgpooling(2) Linear(128,7) Softmax()	Conv2d(1,64), kernel_size = 5*5 Conv2d(64,128), kernel_size = 3*3 Conv2d(128,512), kernel_size = 3*3 Conv2d(512,512), kernel_size = 3*3, padding = 1 (此層無 maxpooling) Conv2d(512,512), kernel_size = 3*3 (以上皆含 relu,batchnorm,maxpooling) Linear(512,512) Linear(512,512) Linear(512,7) Softmax()
參數量	48478	5925390
結果	0.64864	0.66620

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%)



48478	153116	1175324
0.64864	0.65728	0.65672

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%)



61643	392075	5925390
0.6258	0.6492	0.6662

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%)

在參數量只有 50000 左右的時候，MobileNet 有 0.64864 的 accuracy，但一般 CNN 只有 0.6258

MobileNet 將原本的 convolution 拆成兩部分，depthwise 與 pointwise convolution

Depthwise 不像一般 CNN 是跨通道的計算 convolution，透過將每個通道獨立，參數量會是一般 CNN 的  $\frac{1}{N}$  ( $N$ =輸出維度)

Pointwise 部分，則是將 Filter\_size 改為  $1 \times 1$ ，一般 CNN 可能是  $3 \times 3$ ， $5 \times 5$  以上，參數量也比一般 CNN 更少

綜合以上兩點，在差不多參數量下，MobileNet 可以比一般 CNN 疊的更深，在同樣從  $48 \times 48 \times 1 \rightarrow 1 \times 1 \times 64$  的一連串 Convolution 中，我的 MobileNet 疊了 12 層，而我的一般 CNN 只能疊 4 層，從而影響了兩個 Model 對於預測的表現差異