

學號：R07943095 系級：電子碩一 姓名：劉世棠

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%)

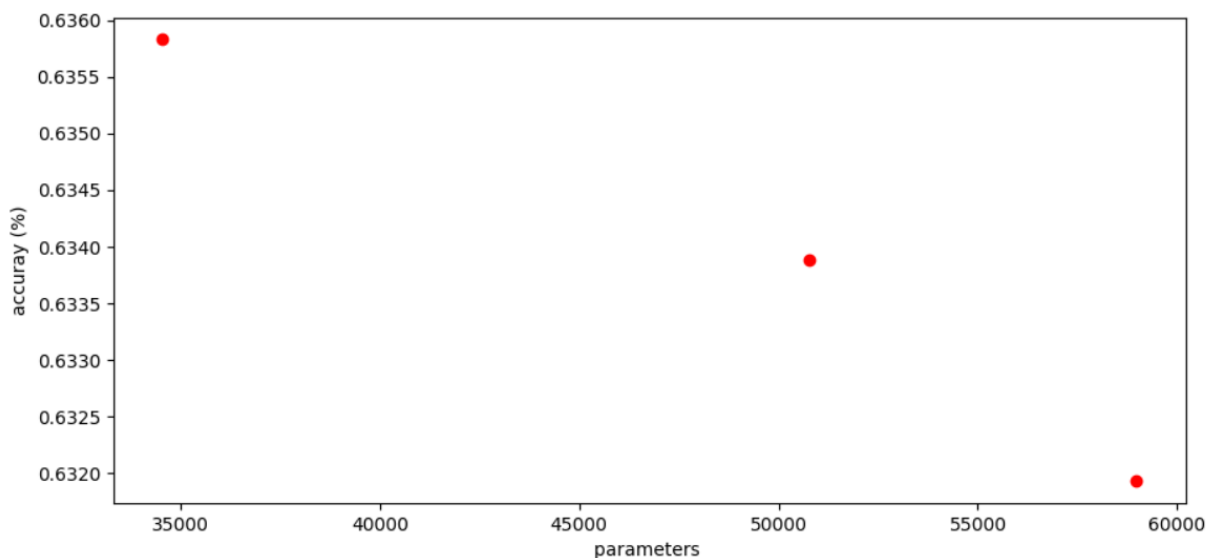
(1)架構：

在作業三因為對參數量沒有要求，卷積疊了四層並且有兩層全連接層，最後使用 ensemble 的技術，但是在作業八因為有了參數數量的要求故只使用一個 model，卷積層除了第一層之外皆改為 mobile net 的形式，全連接層也從原本的 512x512x7 變成 32x7。

(2)參數量與結果：

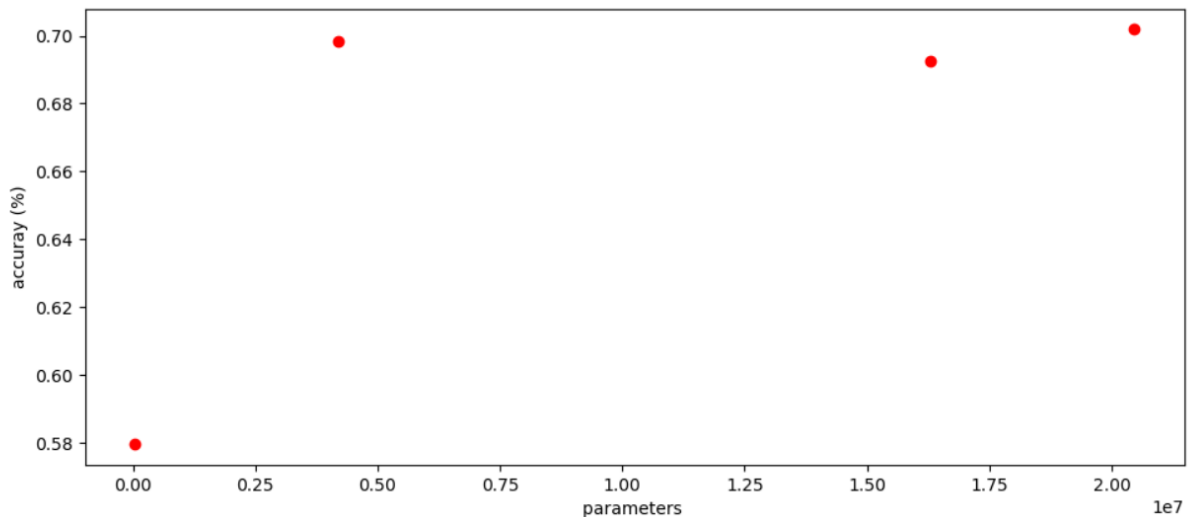
	參數量	Accuracy(public)	Accuracy(private)
HW3	20,470,798	71.329%	70.186%
HW8	34,535	64.168%	63.583%

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%)



⇒ 由於我的 model 一直在改進，所以看起來好像參數越少就越準確，但其實是因為我捨棄掉的部分可能是比較多餘的，故模型沒有多餘的力氣去學雜訊，另外種可能是原本的 model 太爛，所以才會造成訓練不起來的情形。

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%)



⇒ 以上分別為 3 個 model 與 ensemble(準確率最高)的結果。

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%)

	Parameters	Kaggle accuracy (private)
CNN	38,699	57.954%
Mobile Net	34,535	63.583%

⇒ 我將原本的 CNN 參數量減少至與 mobile net 相當，發現 CNN 的準確率下降很多，雖然兩者都有 convolution layer，但是他們的結構是不一樣的，這代表 mobile net 這個結構的假設(刪除掉某些 solution space)對準確率不會影響太多，而傳統的 CNN 因為多了很多不必要的結構(數量少時很難描述一個模型)，所以造成他參數利用率下降，進而傳統的 CNN 在與 mobile net 有差不多數量級時會有較差的結果。