

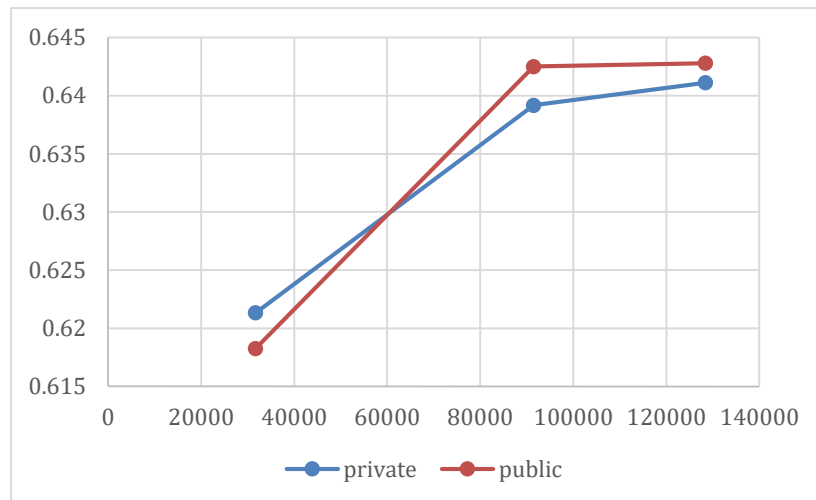
學號：B06901063 系級：電機二 姓名：黃士豪

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%)

	HW3	HW8
架構	Conv2D(64) Conv2D(64) MaxPooling2D() Conv2D(128) Conv2D(128) MaxPooling2D() Conv2D(128) Conv2D(128) MaxPooling2D() Flatten() Dense(512) Dense(512) Dense(7)	Conv2D(32) SeperableConv2D(64) SeperableConv2D(64) SeperableConv2D(128) SeperableConv2D(128) AveragePooling2D() Dense(128) Dense(128) Flatten() Dense(7)
參數量	2532551	91495
結果(public/private)	0.69155/0.67623	0.64502/0.64753

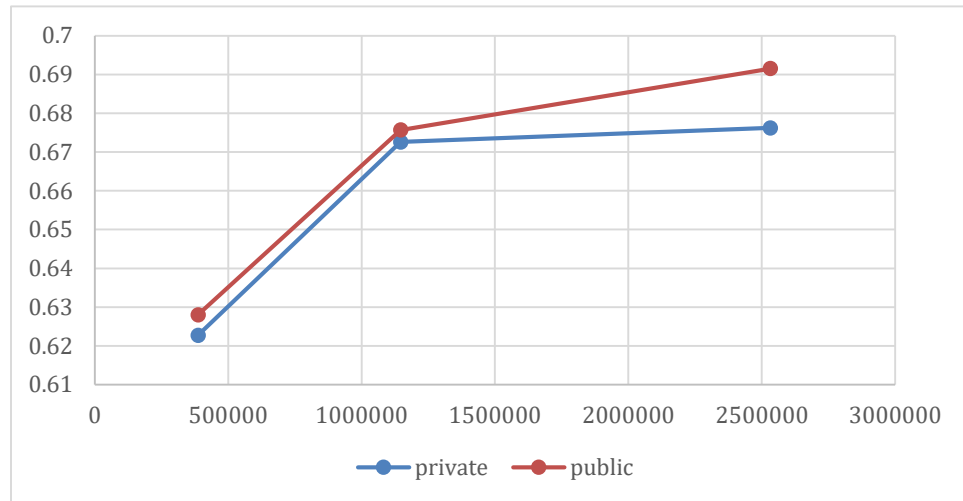
可發現 HW3 的 acc 只多了 4-5%，參數量超過了 25 倍以上，是一個 trade off。

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%)



可以發現參數量越多基本上表現越好。

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%)



可以發現參數量越多基本上表現越好。

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%)

	CNN	Mobile Net
參數量	90311	91495
結果(public/private)	0.61939/0.61855	0.64502/0.64753

可以發現參數量極少的情況下，Mobile Net 的準確率明顯比 CNN 好上一截，推測是因為 CNN 跟 Mobile Net 相比同 filter 數的情況下所需的參數量多很多，導致固定參數量會讓 CNN 能夠實作的 filter 數變成很少，所存的資訊不足以讓 model 辨識出分類的結果，因此得到較不好的結果。