

學號：B06705057 系級：資工二 姓名：黃資翔

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%)

	Mobile Net	CNN
參數量	35,327	4,040,263
Public score	0.61939	0.69462
Private Score	0.60991	0.68904

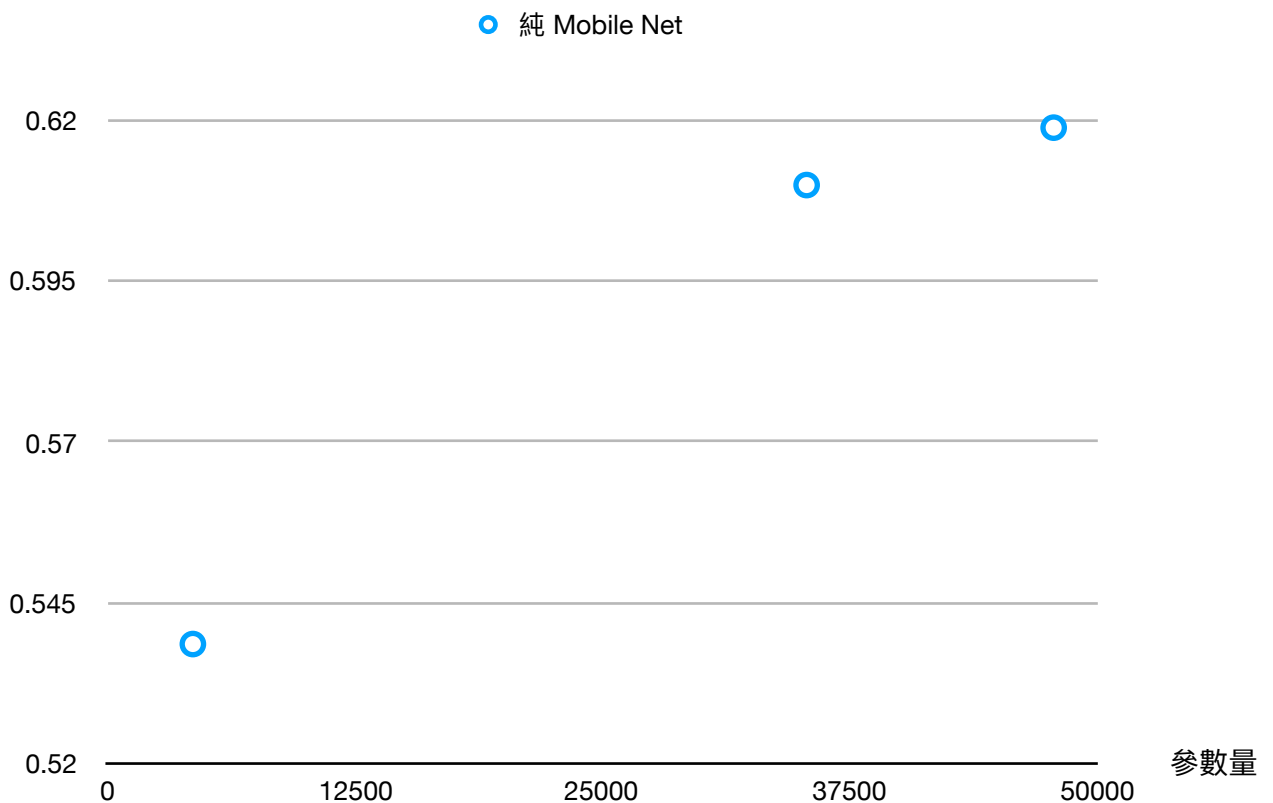
Mobile Net : Conv2D(16) -> SeparableConv2D(16) -> Conv2D(128) -> SeparableConv2D(128) -> Conv2D(64) -> SeparableConv2D(64) -> Conv2D(7) -> flatten -> Dense(7) 。

除了第一個 Conv2D 外，其餘的 Conv2D 中的 kernel 都是 (1, 1)

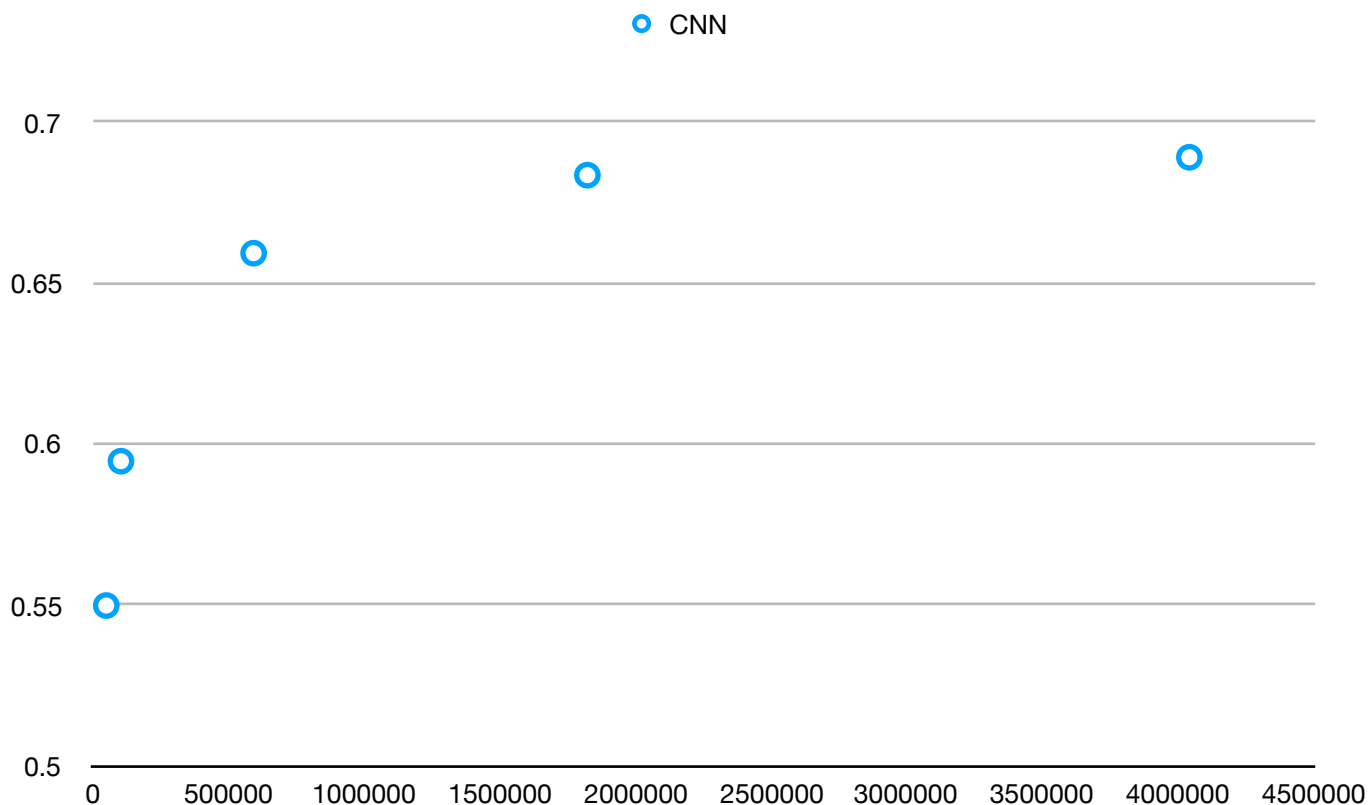
CNN : Conv2D(64) -> Conv2D(128) -> Conv2D(256) -> Conv2D(512) -> flatten -> Dense(256) -> Dense(512) -> Dense(7) 。

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%)

Accuracy



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%)



合作夥伴：B06705058 劉品桷 (參數量：1,823,239，593,799，105,191)

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%)

題二與題三的結果類似，皆是參數量越多，準確率上升幅度會越小，唯有的差別是，在相同的準確率下，題二所需的參數較少。當參數量約為五萬時，Mobile Net 的準確率可以達62%，CNN 卻只有 55%。可以想像當我們利用 Mobile Net 來模擬 CNN 時，即當每一層的 filters 數目相同，他們的表現差不多，然而 Mobile Net 卻可以有更低的參數量。雖然 Mobile Net 是用 depthwise 的 Convolution Network，但其隨後有 $1 \times 1 \times \text{filters}$ 的架構，讓它也可以考慮 filters 間的關係。也就是在相同參數量下，Mobile Net 比起 CNN 可以有更多的 filters 數目，準確率自然較高。搭配 Mobile Net，我們亦可以把權重較小的神經元刪除，因為該神經元沒那麼重要，對準確率影響甚小，卻能進一步大幅減少參數量。