

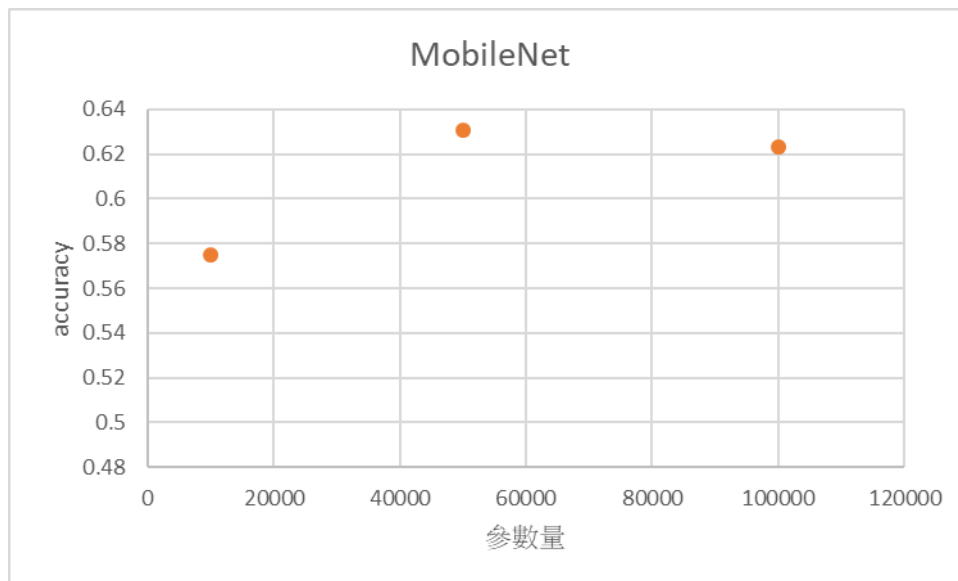
學號：B06902028 系級：資工二 姓名：林柏劭

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%)

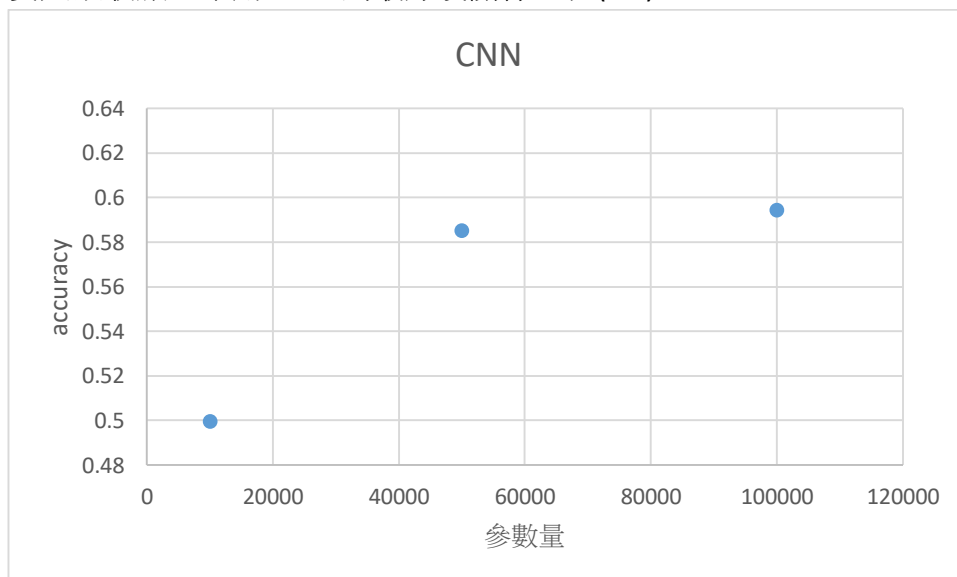
Hw8-MobileNet	Hw3-CNN
SeparableConv(128,3,3)+BN+relu	Conv(64,3,3)+BN+relu
SeparableConv(128,3,3)+BN+relu	Conv(64,3,3)+BN+relu
MaxPooling(2, 2)	MaxPooling(2, 2)
SeparableConv(64,3,3)+BN+relu	Conv(128,3,3)+BN+relu
SeparableConv(64,3,3)+BN+relu	Conv(128,3,3)+BN+relu
MaxPooling(2, 2)	MaxPooling(2, 2)
SeparableConv(64,3,3)+BN+relu	Conv(256,3,3)+BN+relu
SeparableConv(64,3,3)+BN+relu	Conv(256,3,3)+BN+relu
MaxPooling(2, 2)	MaxPooling(2, 2)
SeparableConv(32,3,3)+BN+relu	Conv(512,3,3)+BN+relu
SeparableConv(32,3,3)+BN+relu	Conv(512,3,3)+BN+relu
MaxPooling(2, 2)	MaxPooling(2, 2)
SeparableConv(32,3,3)+BN+relu	Flatten()
SeparableConv(32,3,3)+BN+relu	Dense(512)+BN+relu+dropout(0.5)
MaxPooling(2, 2)	Dense(512)+BN+relu+dropout(0.5)
Flatten()	Dense(512)+BN+relu+dropout(0.5)
Dense(7)+softmax	Dense(7)+softmax
about 50000 params	about 7600000 params
0.63806/0.63081	0.70799/0.69824

(public score/private score)

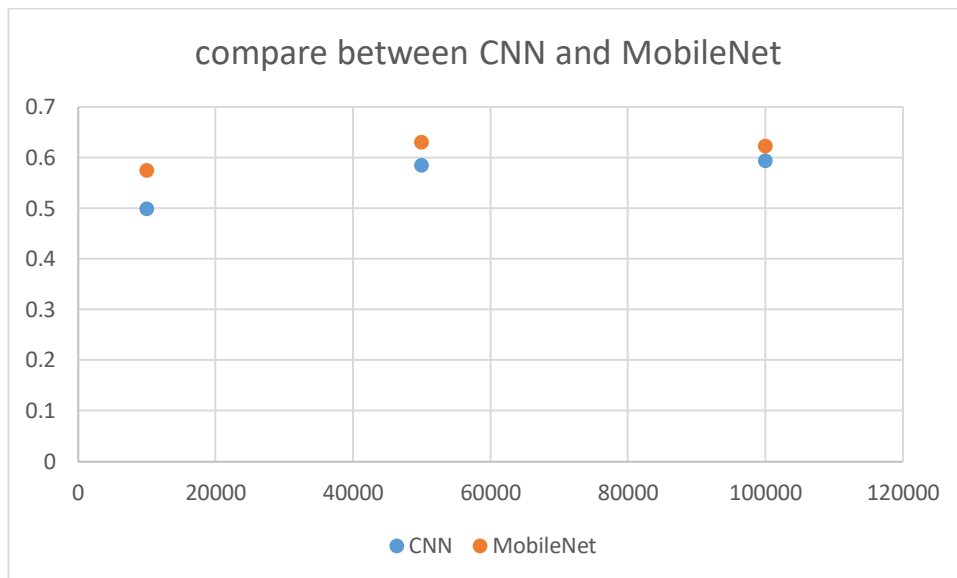
2. 請使用 MobileNet 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。） (1%)



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%)



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%)



將題 2 和題 3 結果組合即可得到上圖。可看出在在我所選擇的參數量的情況下(10000, 50000, 100000，MobileNet 都有較好的結果。會造成這樣的原因是有幾項：首先是因為就算是有 100000 個參數，對於一個一般的 CNN model 來說依舊是偏少，因此他們的準確率都不高(皆尚未超過 0.6)。但由於 MobileNet 是使用了 Depthwise CNN 的架構，便可使用較少的參數量去達到較好的結果，在參數量相當少(10000)時尤其明顯。

我也另外試驗若將我 MobileNet 的結構改回一般的 CNN，則參數量會從 10000 提升至 40000，也就是說他僅用了 10000 的參數便達到了原先要使用 40000 參數的成果，若從圖上來看更可以證實這項推論(MobileNet\_1 與 CNN\_2 結果相近)