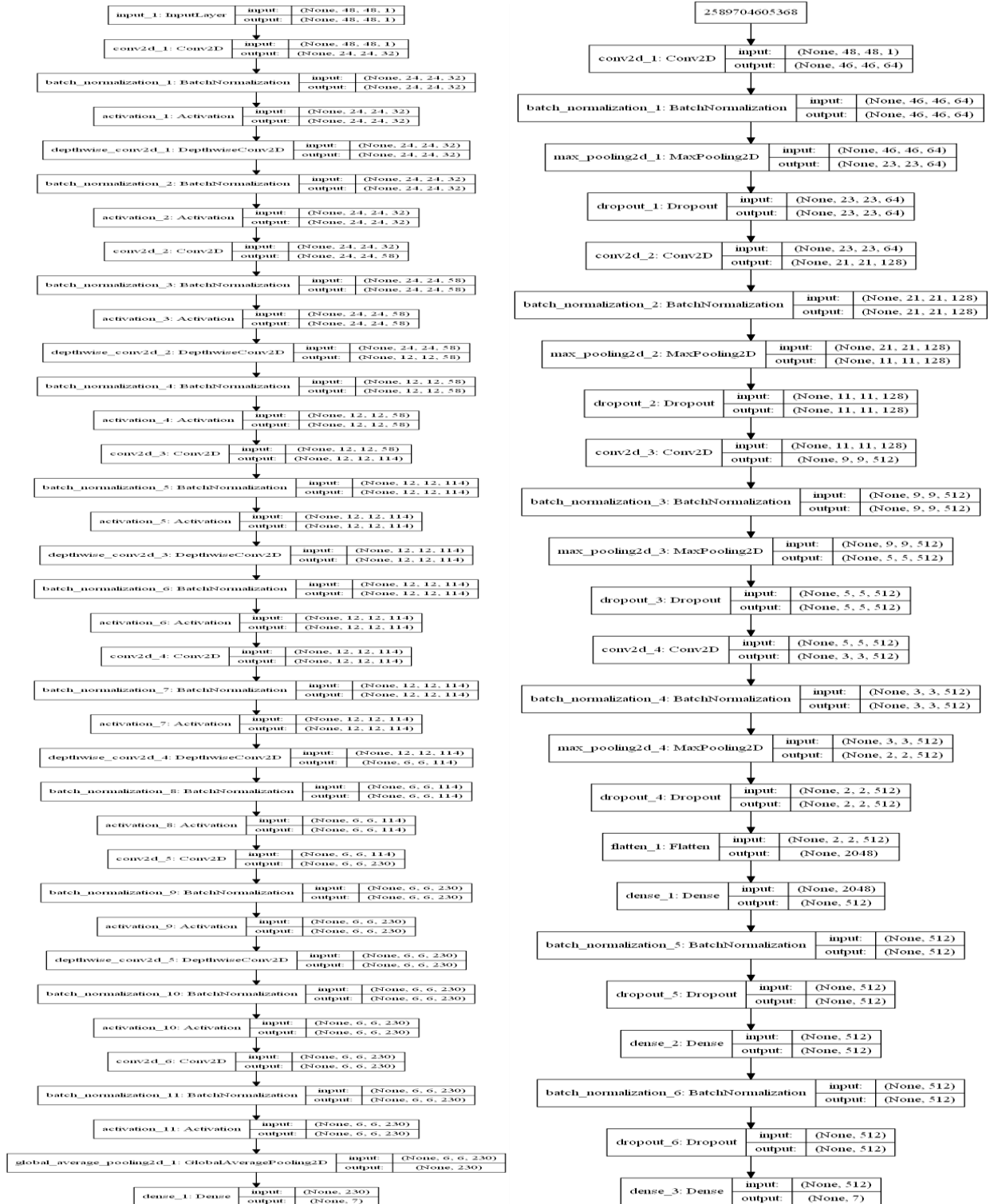


學號：B05502145 系級：電機三 姓名：林禹丞

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%)

本次作業(左)與 HW3(右)作業的架構：



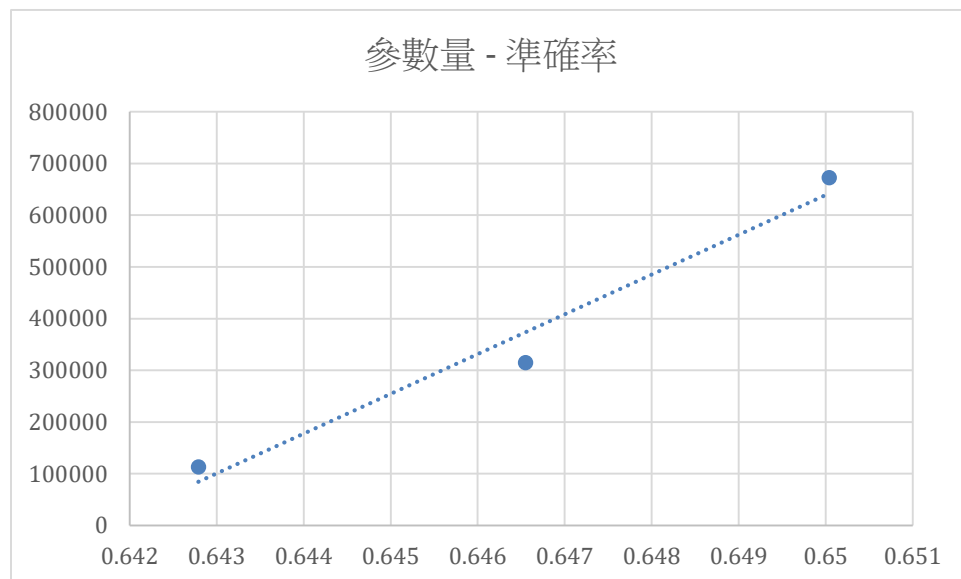
本次作業與 HW3 的比較：

	本次作業	HW3
參數量	11, 2725	434, 8935
正確率(public/private/mean)	0.64474/0.64084/0.64279	0.67177/0.66815/0.66996

- 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%)

三個 model 的架構類似，只有第四層和第五層 neuron 的各數變少而已。都 train 400 個 epoch 並取 validate 最高者，都用 batch size 128 與 data augmentation。

Model name	Total parameters	Trainable Parameter	Non-trainable parameters	Public accuracy	Private accuracy	Mean accuracy
Model3	11, 2725	11, 0073	2652	0.64474	0.64084	0.64279
Model2	31, 4775	31, 0391	4384	0.65617	0.63694	0.64655
Model4	67, 2750	66, 6594	6156	0.65004	0.65004	0.65004

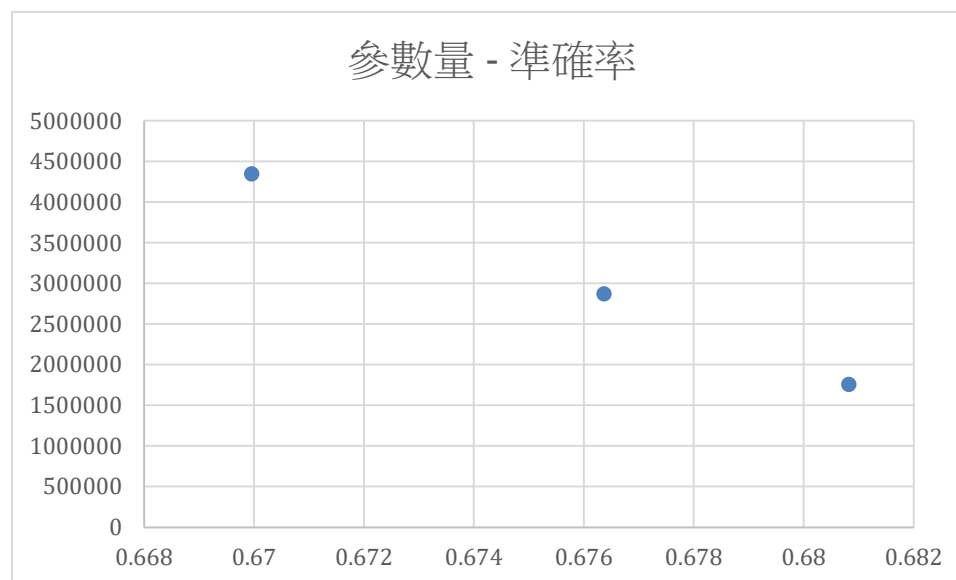


- 請使用一般 CNN 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。）(1%)

前三個 model 的架構類似，只有第三層和第四層 neuron 的各數變少而已，最後一個 model 把每層的 neuron 都調到非常少。都 train 400 個 epoch 並取 validate 最高者，都用 batch size 128 與 data augmentation。

Model name	Total parameters	Trainable Parameter	Non-trainable parameters	Public accuracy	Private accuracy	Mean accuracy
Model2	434, 8935	434, 4455	4480	0.67177	0.66815	0.66996
Model4	287, 3095	286, 9127	3968	0.68208	0.67066	0.67637
Model3	175, 7703	175, 4247	3456	0.68514	0.67651	0.68082

可以看到參數較少時反而正確率上升了，這是因為當參數量多的時候很容易造成 overfitting 的現象，這時候如果我們提升 dropout，那麼參數量比較多的 model 應該可以 train 的比較好。



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%)

Model	參數量	準確率
MobileNet	11, 2725	0.64279
CNN	11, 4601	0.60183

可以觀察到在參數量差不多的時候，CNN 的準確率比 MobileNet 的方式低了很多。這是因為 MobileNet 使用的 Depthwise covolution 加上 Pointwise convolution 可以用很少的參數量達到和 CNN 差不多的效果。因此如果 CNN 的參數量減少到和 MobileNet 差不多，效果會大幅的下降。