

學號：B05902018 系級：資工三 姓名：張凱程

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 model : Conv2D(filter=16) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=24) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=32) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=48) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=84) → Dense(7,softmax)

參數量：68,503

正確率：0.64224 + 0.62970 (public + private)

(每層後面均有加 BatchNormalization 跟 LeakyRelu，且除了第一層 Conv2D 之外，其他 Conv2D 後面均加 Maxpooling 跟 Dropout)

HW3 model: Conv2D(filter=32) → Conv2D(filter=64) → Conv2D(filter=128) → Conv2D(filter=128) → Conv2D(filter=256) → Conv2D(filter=512) → Conv2D(filter=512) → Dense(1024,relu) → Dense(512,relu) → Dense(7,softmax)

參數量：5,139,719

正確率：0.69935 + 0.69963 (public + private)

(每層後面均有加 BatchNormalization、LeakyRelu 和 Dropout，且除了前兩層 Conv2D 之外，其他 Conv2D 後面均加 Maxpooling)

比較：前者的參數較少，在結果上很難比 HW3 的大架構準確，但是依舊可以過 60% 的準確率，而下一題會發現 DepthwiseConv2D 可以幫助模型在相同參數量下提高正確率。

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%)

Ans:

model 1: Conv2D(filter=16) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=32) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=48) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=64) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=128) → Dense(7,softmax)

參數量：131,543

正確率：0.65032 + 0.64251 (public + private)

model 2: Conv2D(filter=16) → DepthwiseConv2D(depth\_multiplier=1) →  
Conv2D(filter=24) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=32) →  
DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=48) →  
DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=84) → Dense(7,softmax)

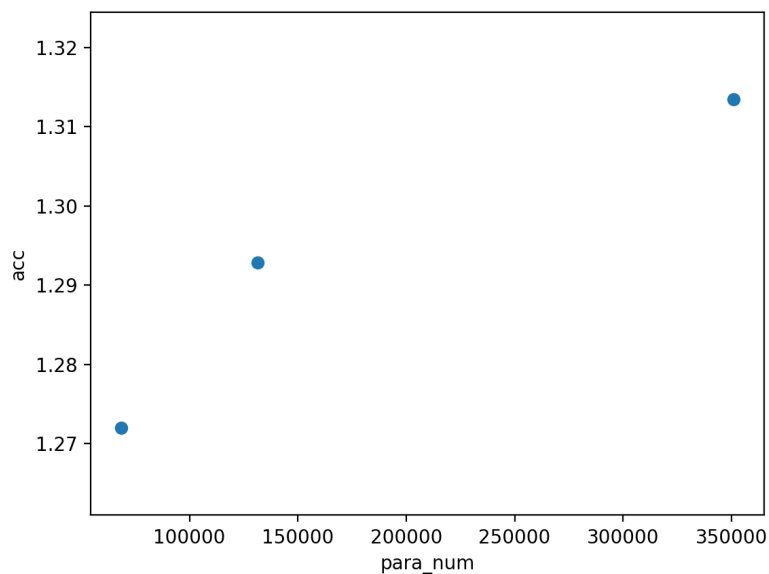
參數量：68,503

正確率：0.64224 + 0.62970 (public + private)

model 3: Conv2D(filter=24) → DepthwiseConv2D(depth\_multiplier=1) →  
Conv2D(filter=48) → DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=64) →  
DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=128) →  
DepthwiseConv2D(depth\_multiplier=1) → Conv2D(filter=192) → Dense(7,softmax)

參數量：350,959

正確率：0.65672 + 0.65672 (public + private)



可以看出參數量增加，正確率也會增加，但增加的幅度隨參數量增大而減少。

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%)

Ans:

model 1: Conv2D(filter=16) → Conv2D(filter=24) → Conv2D(filter=32) →  
Conv2D(filter=48) → Conv2D(filter=84) → Dense(7,softmax)  
參數量：66,943

正確率：0.62803 + 0.62329 (public + private)

model 2: Conv2D(filter=16) → Conv2D(filter=32) → Conv2D(filter=64) →  
Conv2D(filter=64) → Conv2D(filter=128) → Dense(7,softmax)

參數量：143,367

正確率：0.63833 + 0.63945 (public + private)

model 3: Conv2D(filter=24) → Conv2D(filter=48) → Conv2D(filter=64) →  
Conv2D(filter=128) → Conv2D(filter=192) → Dense(7,softmax)

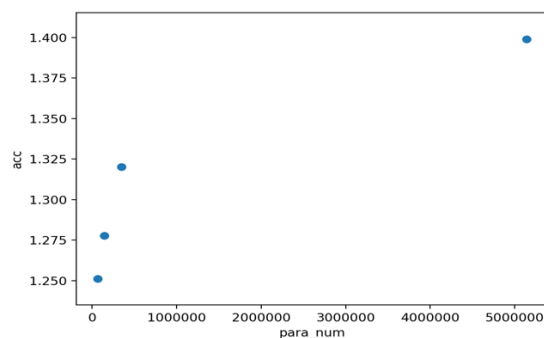
參數量：347,527 (mobile 8)

正確率：0.66090 + 0.65923 (public + private)

model 4: Conv2D(filter=32) → Conv2D(filter=64) → Conv2D(filter=128) →  
Conv2D(filter=128) → Conv2D(filter=256) → Conv2D(filter=512) →  
Conv2D(filter=512) → Dense(1024,relu) → Dense(512,relu) → Dense(7,softmax)

參數量：5,139,719

正確率：0.69935 + 0.69963 (public + private)



結果大致同第 2.題

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%)

Ans: 當參數量很小的時候，有 Depthwise Convolution 的 model 準確率會比純 CNN 高個 1% 的準確率。很可能是因為 Depthwise Convolution 對輸入的每個 channel 有多做一次 Convolution，且參數的個數只有  $(N\_channel * (3*3))$ ，對比 convolution layer 參數個數有約  $(N\_in * N\_out * (3,3))$  個，相對省空間且有 convolution 的效果。