

學號：B06902052 系級：資工二 姓名：張集貴

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 架構：

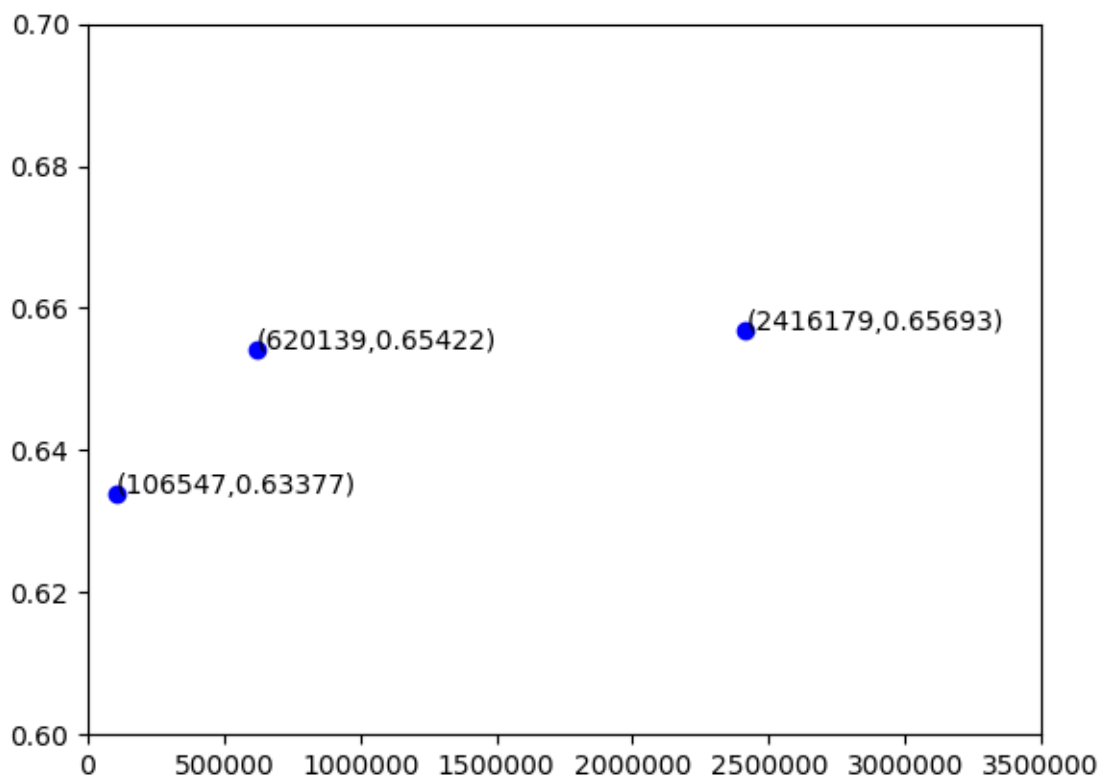
```
1 model = Sequential()
2
3 model.add(Conv2D(15, (3,3), padding='same', input_shape=(48,48,1)))
4 model.add(BatchNormalization())
5 model.add(ELU())
6 model.add(SeparableConv2D(15, (3,3), padding='same'))
7 model.add(ELU())
8 model.add(MaxPooling2D((2,2)))
9
10 for f in [50,100]:
11     model.add(Conv2D(f, (3,3), padding='same'))
12     model.add(BatchNormalization())
13     model.add(ELU())
14     model.add(SeparableConv2D(f+20, (3,3), padding='same'))
15     model.add(ELU())
16     model.add(MaxPooling2D((2,2)))
17
18 model.add(SeparableConv2D(120, (3,3), padding='same'))
19 model.add(GlobalAvgPool2D())
20 model.add(Dense(units=7, activation='softmax'))
```

HW3 架構：

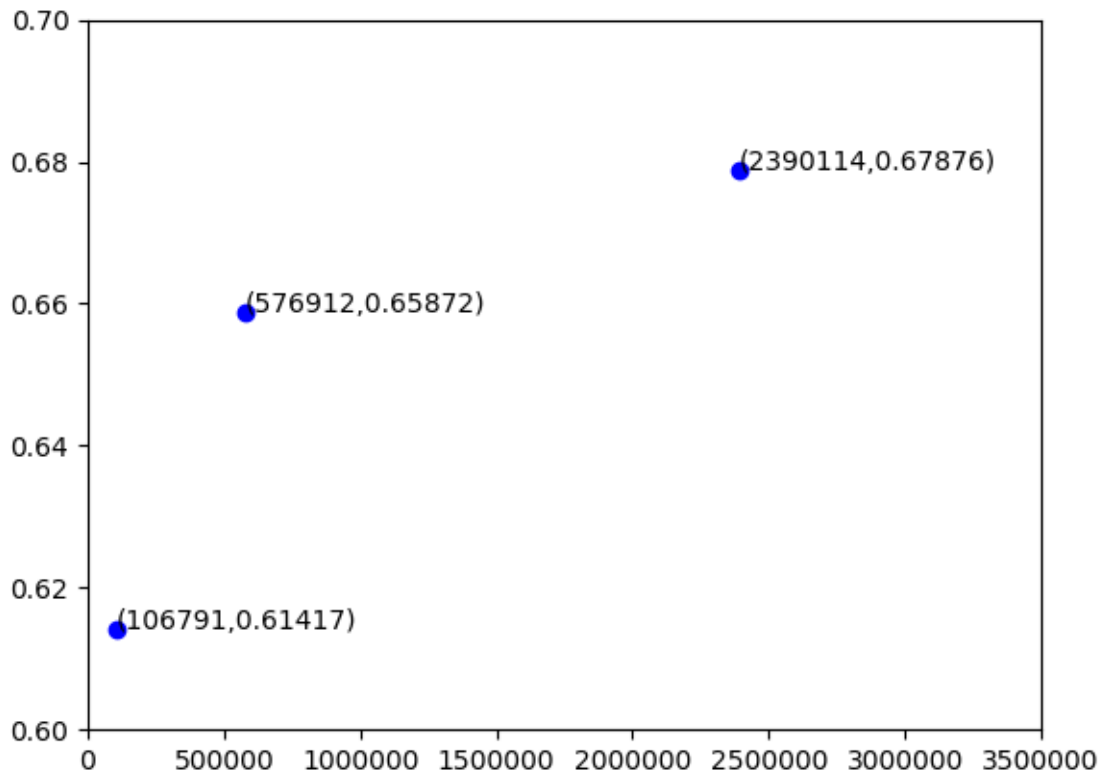
```
1 model = Sequential()
2 model.add(Conv2D(128, (3,3), input_shape = (48,48,1), activation = 'elu'))
3 model.add(BatchNormalization())
4 model.add(Dropout(0.2))
5 model.add(Conv2D(128, (3,3), activation = 'elu'))
6 model.add(BatchNormalization())
7 model.add(MaxPooling2D(pool_size = (2, 2)))
8 model.add(Dropout(0.2))
9
10 model.add(Conv2D(256, (3,3), activation = 'elu'))
11 model.add(BatchNormalization())
12 model.add(Dropout(0.3))
13 model.add(Conv2D(256, (3,3), activation = 'elu'))
14 model.add(BatchNormalization())
15 model.add(MaxPooling2D(pool_size = (2, 2)))
16 model.add(Dropout(0.3))
17
18 model.add(Conv2D(512, (3,3), activation = 'elu'))
19 model.add(BatchNormalization())
20 model.add(Dropout(0.4))
21 model.add(Conv2D(512, (3,3), activation = 'elu'))
22 model.add(BatchNormalization())
23 model.add(MaxPooling2D(pool_size = (2, 2)))
24 model.add(Dropout(0.4))
25
26 model.add(Flatten())
27 model.add(Dense(units = 512, activation = 'elu'))
28 model.add(BatchNormalization())
29 model.add(Dense(units = 512, activation = 'elu'))
30 model.add(BatchNormalization())
31 model.add(Dense(units = 7, activation = 'softmax'))
```

	參數量	Kaggle Public	Kaggle Private
HW8	104572	0.61549	0.60741
HW3	5900679	0.68514	0.68765

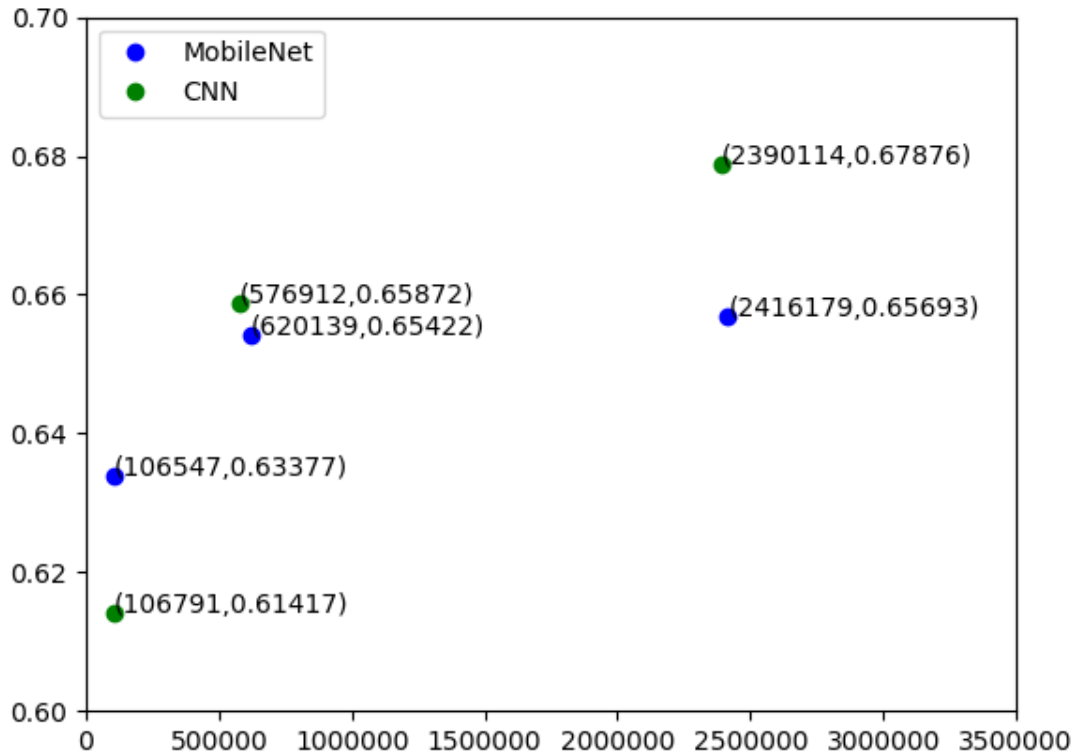
2. 請使用 MobileNet 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。） (1%)



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%)



當參數量相當少且兩者參數量相當的時候，MobileNet 的表現會比 CNN 好。

我認為在參數量相當時，Depthwise Separable Convolution 輸出的 features 數會多於一般的 CNN，因此有較好的表現。