

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

採用 mobilenetV2，該方式相較 mobilenetV1 多了 inverted residual。該層特點在於先利用 1x1 conv2D 升維，再用 Depthwise 做 deep conv，最後採用 linear bottleneck。依據論文所說[1]，由於前後 1x1 conv2d，即使中間層的計算量大，仍可有效降低參數。此外最後一層利用線性變換省去 RELU，可以防止非線性破壞過多的訊息。此外一開始先經過一層 conv 2d kernel size=3，先取得較好 feature，經過 inverted residual 後再一層 1x1 conv 2d 和 AvgPool2d。最後輸出 fully connected layer。model 採用 RELU6，減少精度換取較小的 model。但在 inverted residual 部分測試後仍採用 RELU6，發現效果較好。

```
self.conv = nn.Sequential(
    nn.Conv2d(inp, inp * expand, 1, 1, 0, bias=False),
    nn.BatchNorm2d(inp * expand),
    nn.ReLU(inplace=True),
    nn.Conv2d(inp * expand, inp * expand, 3, stride, 1, groups=inp * expand, bias=False),
    nn.BatchNorm2d(inp * expand),
    nn.ReLU(inplace=True),
    nn.Conv2d(inp * expand, oup, 1, 1, 0, bias=False),
    nn.BatchNorm2d(oup),
)
```

計算參數:

sum(s.numel() for s in
mobilenet.parameters() if s.requires_grad)

總訓練參數:365272

準確率:

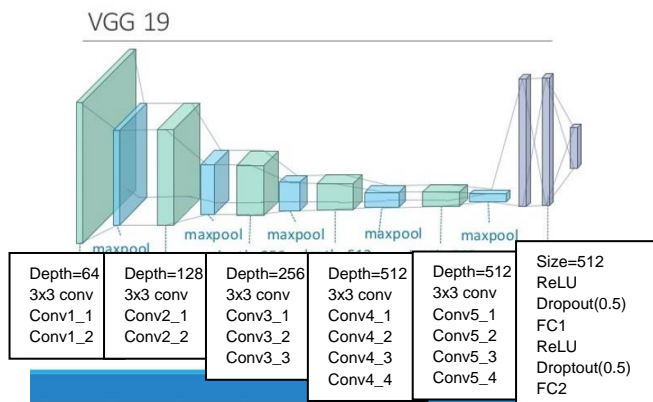
Train:0.71

Kaggle Public:0.64446

HW3

CNN model: 參考 VGG-19 的架構，並進行了調整:

MaxPool(2,2)、EPOCH=300、BATCH_SIZE= 256、LR = 0.001、activation=ReLU



計算參數:

sum(s.numel() for s in cnn.parameters() if
s.requires_grad)

總訓練參數:19526727

準確率:

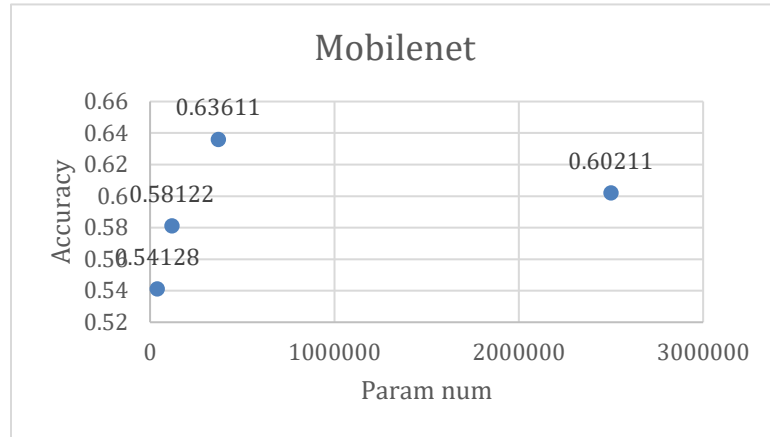
Train:0.80

Kaggle Public:0.69183

由結果可以看到，mobilenetV2 確實大大減少了參數量，原因是在架構上，mobilenet 分成 depthwise convolution 和 pointwise convolution，將 model 切分較小塊去訓練並利用 conv2d 1x1 有效降低參數，且準確率並不會下降太多，這樣比傳統的 CNN model 減少不少參數。

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%）

EPOCH=30 的結果，accuracy 為 kaggle public。主要調整中間 con2d 的層數，得到結果如下圖：

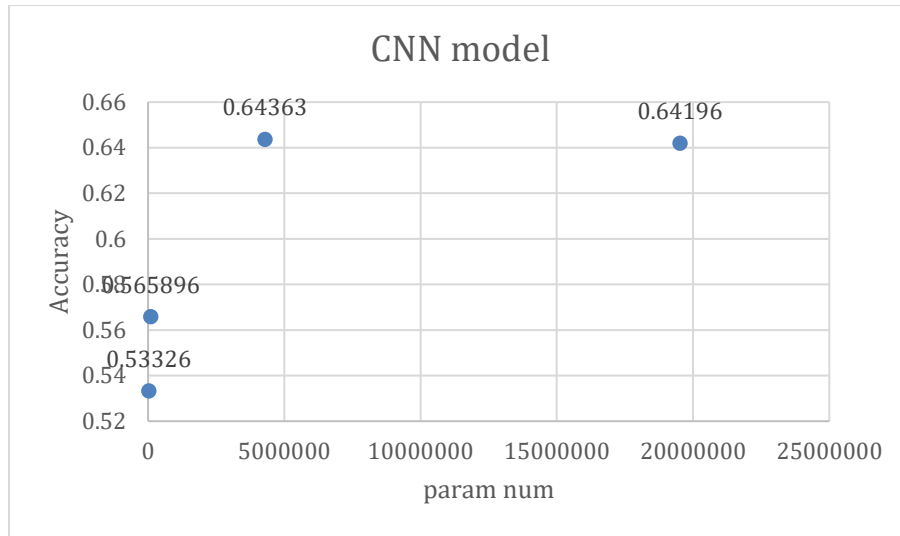


參數量	準確率
39506	0.54128
119164	0.58122
370360	0.63611
2500192	0.60211

由結果可以知道，低參數情況下，層數越多，結果有較好，但若層數太多，並不一定有更高的準確率。

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%）

EPOCH=30 的結果，accuracy 為 kaggle public。



參數量	準確率
33111	0.53326
110279	0.56896
430023	0.64363
19526727	0.64196

CNN 結果其實超乎我的預期，寫第三次作業當時想法是越多層越強大，疊的跟 VGG 一樣高，準確率就會飛高高。但實際似乎層數降低並未對結果有太大影響。

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时候，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%)

由前兩題結果可以看到，當參數量較少的时候，mobilenet 有稍微好一點，但不排除因參數量有稍微高一點才有較高準確率。而兩個結果相近的原因，我認為和 task 的 data 有關。我們 image 大小不大，所擁有的資訊有限，即使做 data preprocess 和 data augmentation，準確率也大約落在 65%~75% 的區間。因此，mobilenet 雖然有效降低參數量，但如果將 cnn model 降低層數，減少 fully connected layer，也能學到差不多的 feature，導致準確率也差不多。

Reference:

[1] MobileNetV2: Inverted Residuals and Linear Bottlenecks

<https://arxiv.org/pdf/1801.04381.pdf>