

1. (1%) 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？

答：

本次 CNNmodel 我主要採用 VGG 架構，利用增加積卷層的深度來達到提高效能的目的，但是由於積卷層越深，model 就越不容易收斂，因此在本次作業中不能使用 pretrain 參數的情況下只作了 VGG-8，activate function 都是使用 reLu，如下表 (Table 1) 為本次模型的架構。總參數為 895221 個。訓練過程的準確度如 Figure 1。

Table 1 CNN model

VGG-8
Conv3-32 Conv3-32
MaxPooling2
Conv3-64 Conv3-64
MaxPooling2
Conv3-128 Conv3-128
Flatten
Fully Conect128
Fully Conect128
Softmax



Figure 1 CNN Training Process

訓練過程中利用 mini-batch gradient 做 100 個 epoch 的訓練，但是這有一個問題是後面幾個 layer 容易會有梯度發散，因此訓練模型中有加入 Batch Normalization 來防止這個問題，以加速收斂。除此之外也加入 Dropout 防止訓練 Overfitting，最後得到結果如下：

Table 2 CNN accuracy

Training accuracy	Validation accuracy	Testing accuracy
87.4%	64.57%	65.65%

2. (1%) 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model。其模型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

答：

DNN 的部分我只利用越接近 input 越少 Neuron 來實做，如下表(Table 3)為 DNN 模型，總參數量為 879733。訓練過程的準確度如**錯誤！找不到參照來源。**。

Table 3 DNN Model

DNN
Fully Conect32
Fully Conect64
Fully Conect128
Fully Conect128
Fully Conect128
Fully Conect256
Fully Conect256
Fully Conect512
Fully Conect512
Fully Conect512
Softmax

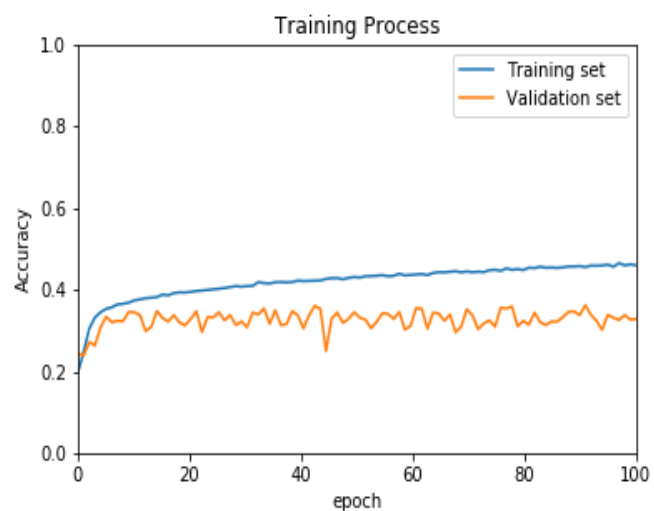


Figure 2 DNN Training Process

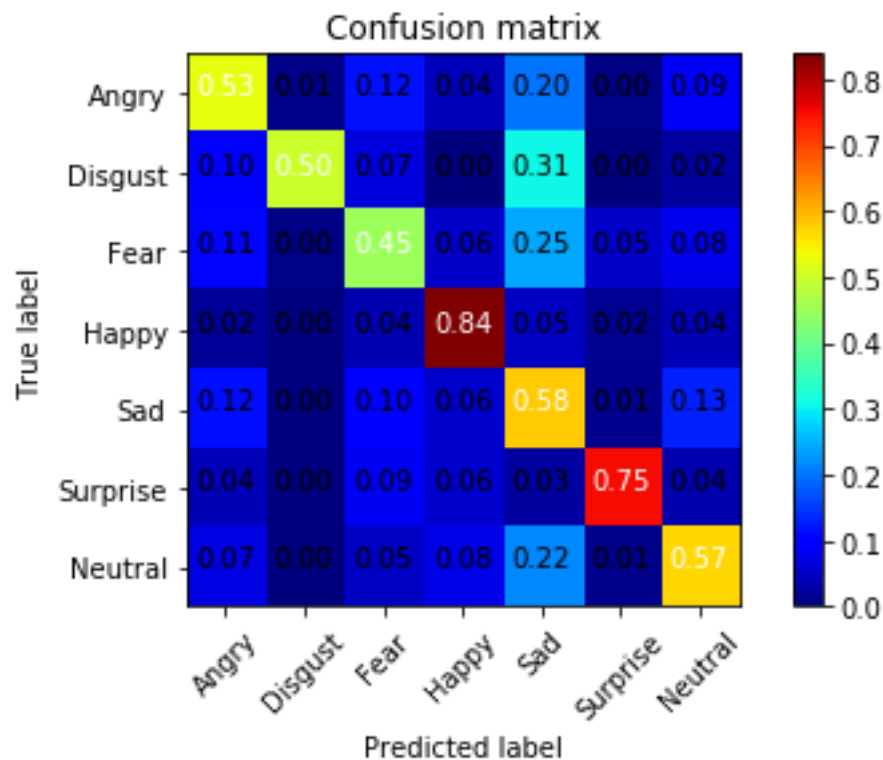
其他的部分皆跟 CNN 一樣，最後得到準確度如下：

Table 4 DNN accuracy

Training accuracy	Validation accuracy	Testing accuracy
45.95%	32.81%	28.81%

從訓練過程以及準確度來看，可以看出 DNN 比起 CNN 更不容易收斂，而且即便收斂了預測準確度也相當的不穩定。

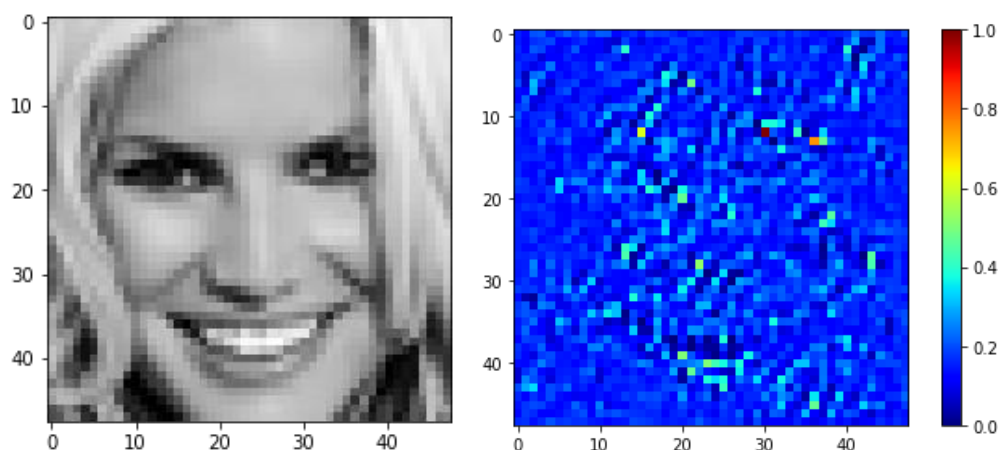
3. (1%) 觀察答錯的圖片中，哪些 class 彼此間容易用混？[繪出分析]
答：

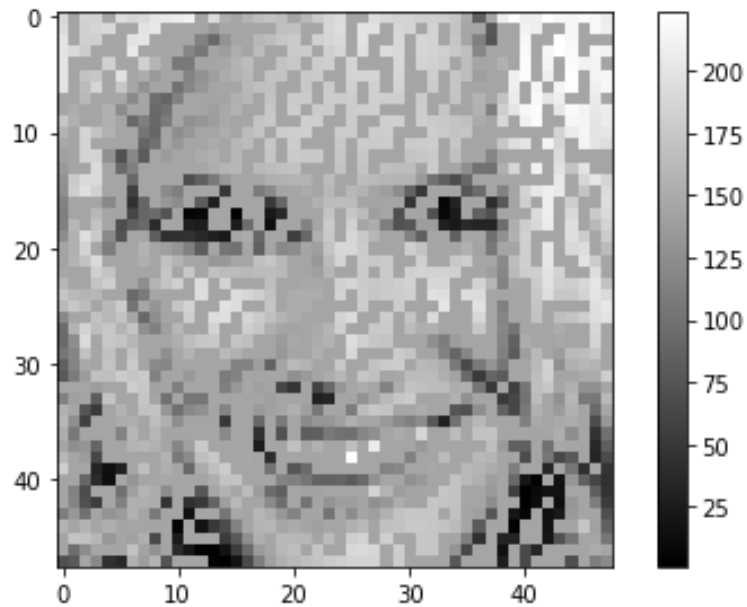


從圖中可以看出 Happy 以及 Surprise 比起其他臉部表情比較不容易被辨認成其他表情，而 Angry、Disgust、Fear 這些負面臉部特徵都很容易被辨認成 Sad。Sad 本身也容易被辨認成 Angry、Fear 等其他負面的表情。比較有趣的是 Neutral 這個無表情狀態跟 Sad 有很大的機率被混淆。

4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？

答：





從上面三個個圖看起來大概是嘴巴，臉頰以及眼睛比較容易被觀察出來

5. (1%) 承(1)(2)，利用上課所提到的 gradient ascent 方法，觀察特定層的 filter 最容易被哪種圖片 activate。

答：

[Bonus] (1%) 從 training data 中移除部份 label，實做 semi-supervised learning

[Bonus] (1%) 在 Problem 5 中，提供了 3 個 hint，可以嘗試實作及觀察（但也可以不限於 hint 所提到的方向，也可以自己去研究更多關於 CNN 細節的資料），並說明你做了些什麼？ [完成 1 個：+0.4%，完成 2 個：+0.7%，完成 3 個：+1%]