

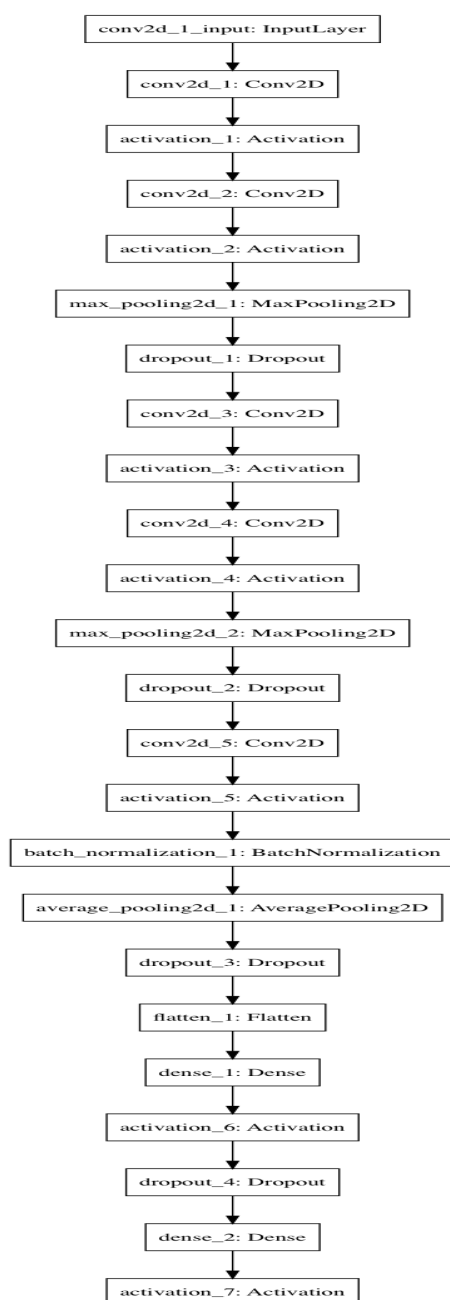
學號：B04902051 系級：資工二 姓名：林承豫

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？

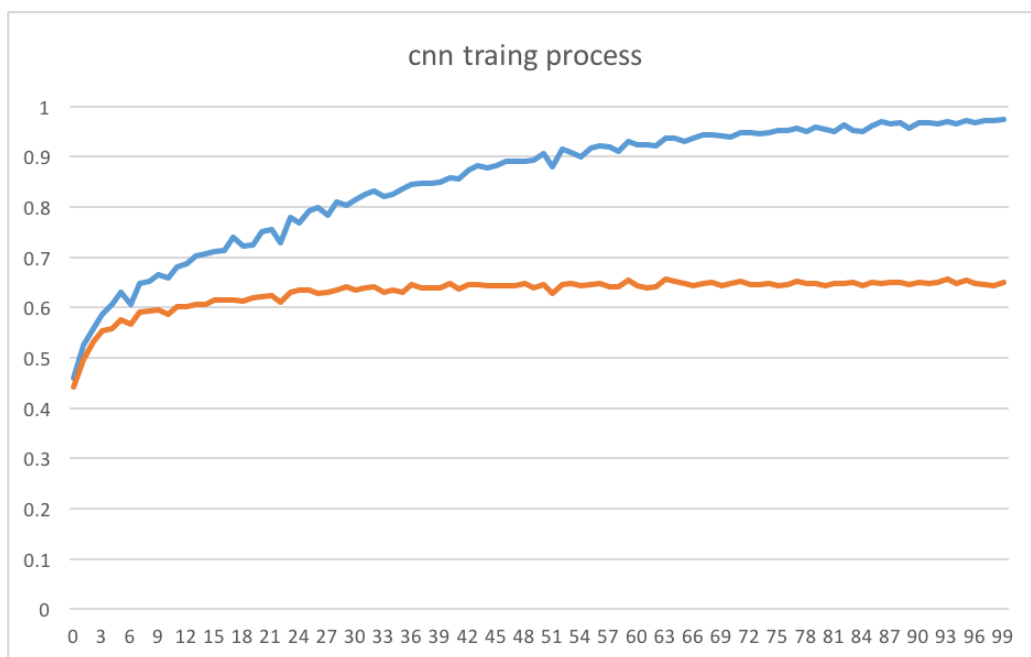
答：

先把資料切出 5500 筆當作 validation data，剩下的資料再透過把圖片上下反轉，以及透過 datagen 隨機把圖片翻轉 10 度，透過如此方法可以把資料量增為六萬多筆，之後再進行訓練。經過測試增加了 datagen 之後正確率大約多了 2%

模型架構為兩層 32 x3x3 -> maxpooling 2x2 -> 兩個 128 x5x5 -> maxpooling 2x2 -> 256 x3x3 -> batchnormalization-> averagepooling -> flatten -> dense(512)-> dense(7) -> softmax



訓練過程如下（途中藍線為 training data 橘線為 validation data）



最後結果為訓練 100epoch 取 valid 結果最好的 model，在 kaggle 上準確率大約到 65-66%之間

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？

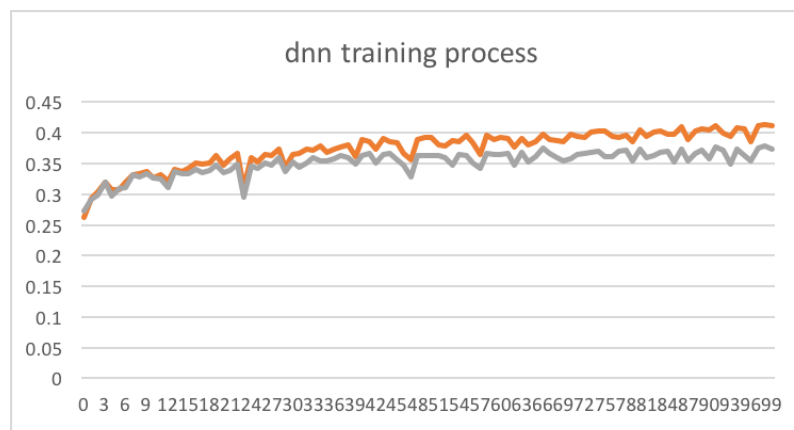
答：下圖為 cnn model 參數

```
Total params: 2,919,271.0  
Trainable params: 2,918,759.0  
Non-trainable params: 512.0
```

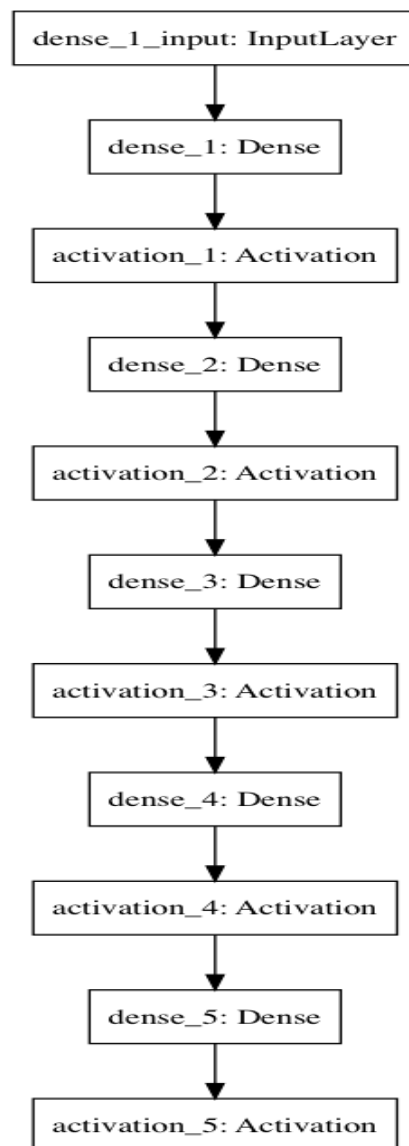
下圖為 dnn model 參數

```
Total params: 3,050,247.0  
Trainable params: 3,050,247.0  
Non-trainable params: 0.0
```

訓練過程



dnn model 架構

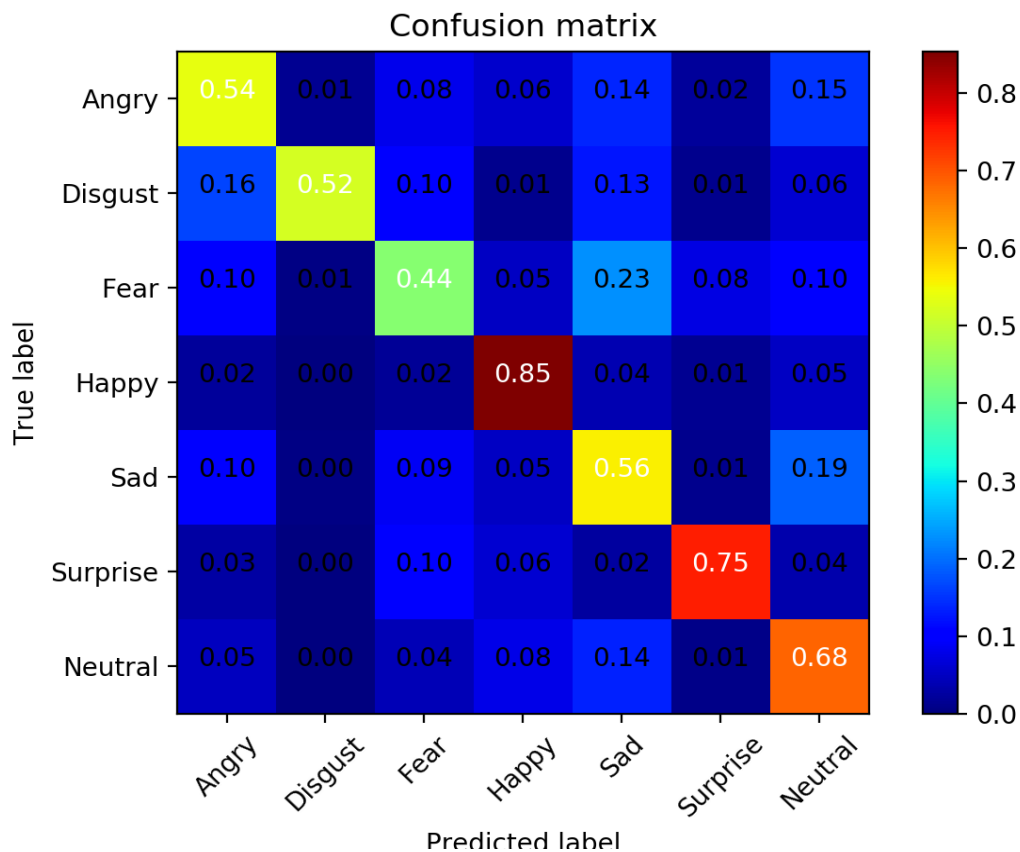


可發現同樣參數量，在同樣 epoch 數下（100） dnn model 比 cnn model 層數少很多，正確率也只有在 36%附近，比 cnn 少了 30%，我認為透過 cnn 比較容易取得圖片中某個區域的特徵，因此正確率較高。

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

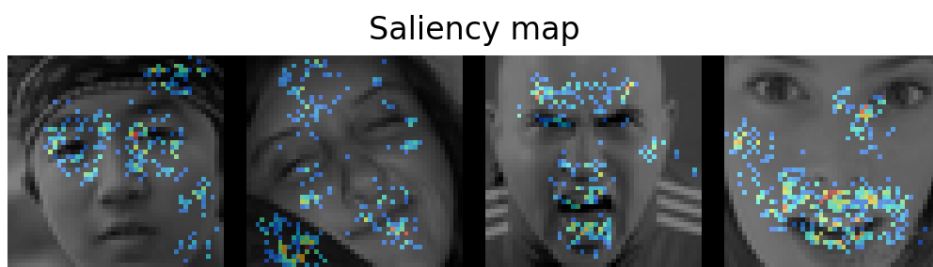
答：

下圖是用 cnn model 做出來的 confusion matrix



可以發現 angry 容易被錯認為 sad 和 neutral 而 disgust 容易被錯認為 angry 和 fear，fear 易被錯認為 sad 和 neutral，happy 最不容易被錯認，而 sad 極容易被誤認為 neutral，surprise 易被錯認為 fear，neutral 易被錯認為 sad，可發現 sad 和 neutral 與 fear 之間的界線似乎不太明顯，三者之間容易互相混淆。

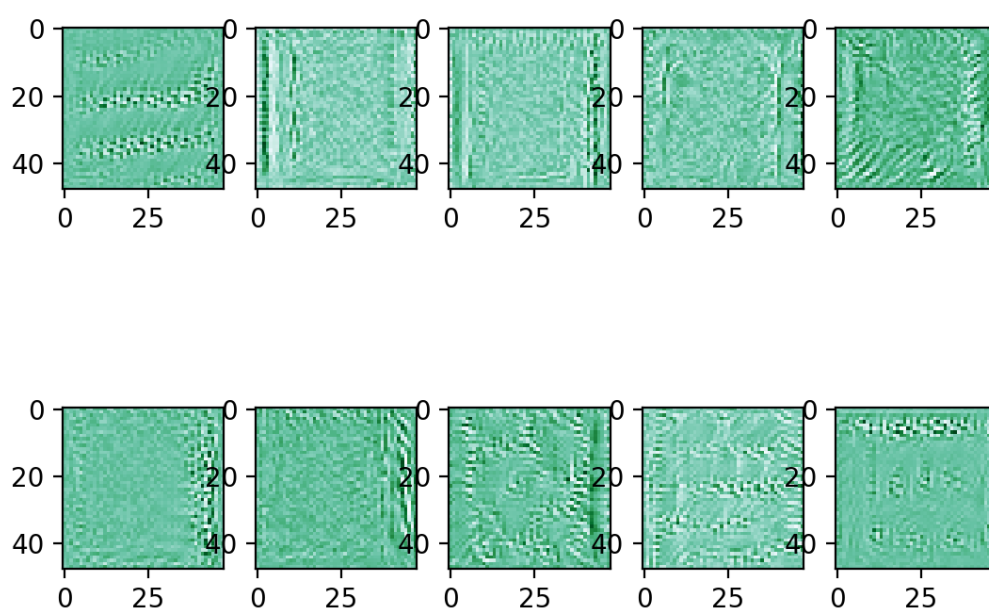
4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 **saliency maps**，觀察模型在做 **classification** 時，是 **focus** 在圖片的哪些部份？
答：



由這四張圖可以發現 model 在做 **classification** 時，是 **focus** 在眼睛跟嘴巴的部分，向右邊數來第二張圖可以看出在生氣時嘴巴張大是一個特徵，而右邊數來第三張圖的眼睛閉起來，故 filter 沒辦法找到眼睛的部分，只能辨認嘴巴的部分。

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**。

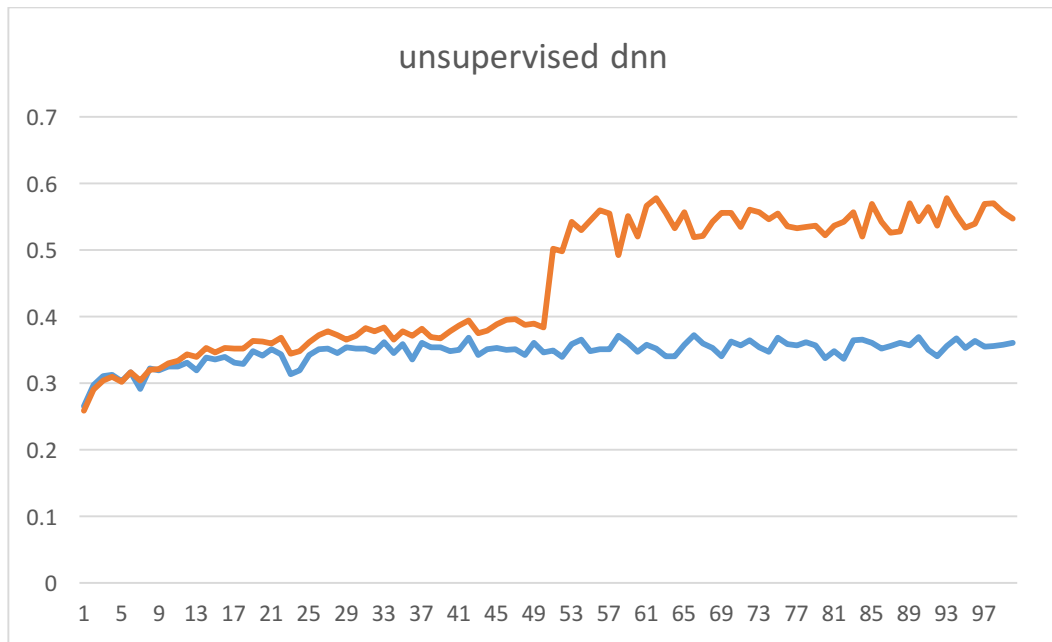
答：



model 選取的是 cnn model，而這題取的是第四層的 filter，取出十張 filter，可發現有幾張圖是辨認斜線的 filter，有些是橫線，有些是把邊框去掉的 filter，每張 filter 都有不同的功能。

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

這題因用 cnn model training 時間較長（沒有 gpu），故採用第二題的 dnn model，把一半 data label 拔除，先用一半的 data train 50epoch 之後，把沒有 label data predict 出 label，再一起 train 50 個 epoch，因為 50 epoch 後加入 unlabeled data 故 training data 正確率上升，但 validation data 並沒有上升，但有發現正確率大概跟第二題的正確率差不多，代表我們可以只用一半的 data 便可達到同樣的正確率。



[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%】