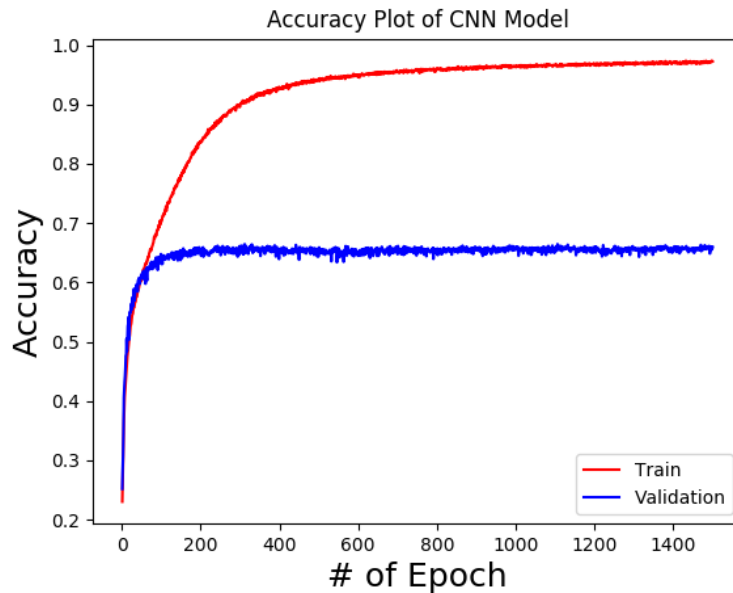


1. 請說明你實作的 CNN model,其模型架構、訓練過程和準確率為何？

答：所使用的模型主要架構為 Sequential，一共用了4層 Conv2D，4層 Dense。

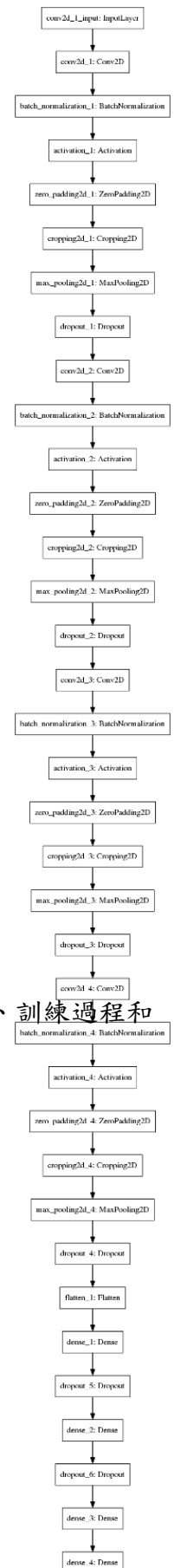
batch size:64 epoch:1800 在 kaggle 上 public set 的準確率為0.64642。



```
=====  
Total params: 1,896,757  
Trainable params: 1,894,873  
Non-trainable params: 1,884  
=====
```

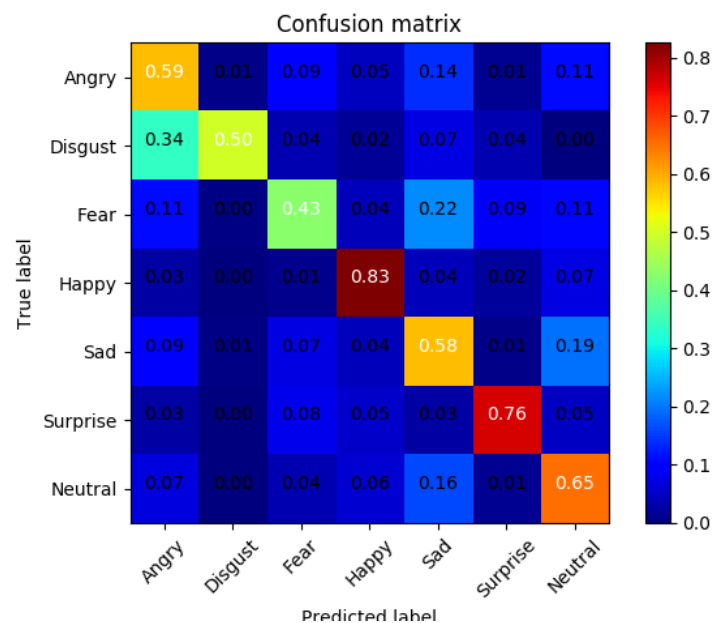
2. 承上題, 請用與上述 CNN 接近的參數量, 實作簡單的 DNN model,其模型架構、訓練過程和準確率為何？試與上題結果做比較,並說明你觀察到了什麼？

答：



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

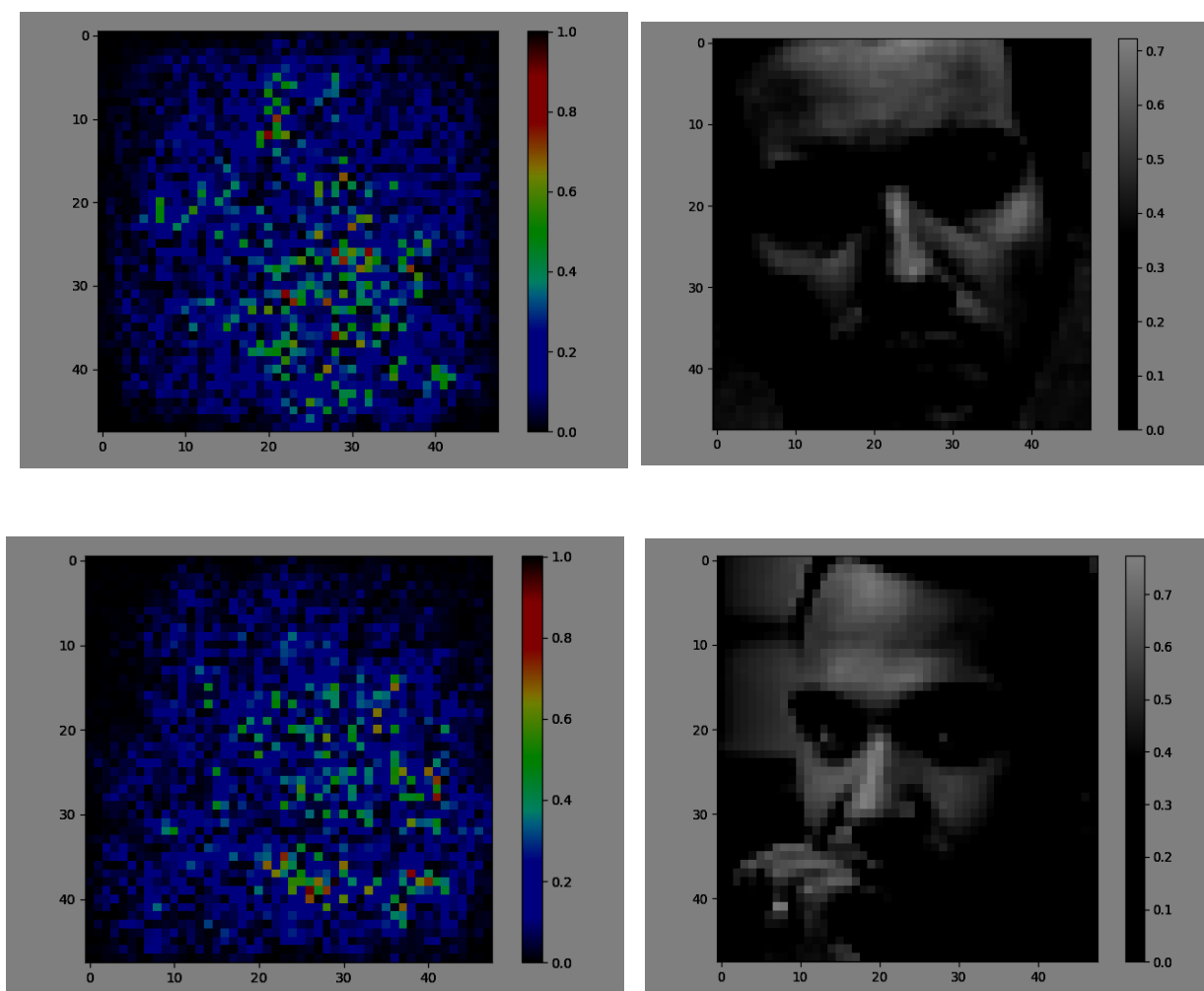
答: 從 train.csv 選取後半部15%的圖片(大概4306張)來做 confusion matrix 分析, 輸出結果如下:



可看到對角線上顏色較為鮮艷, 每種 class 被歸類到正確所屬類別可能性還算高。{Angry: 0.59; Disgust: 0.50; Fear: 0.43; Happy: 0.83; Sad: 0.58; Surprise: 0.76; Neutral: 0.65}predicted model 辨識錯誤的機率大多落在10%以下, 比較會辨識失敗的有將 Disgust 分類成 Angry、將 Fear 看作是 Sad。

4. 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：

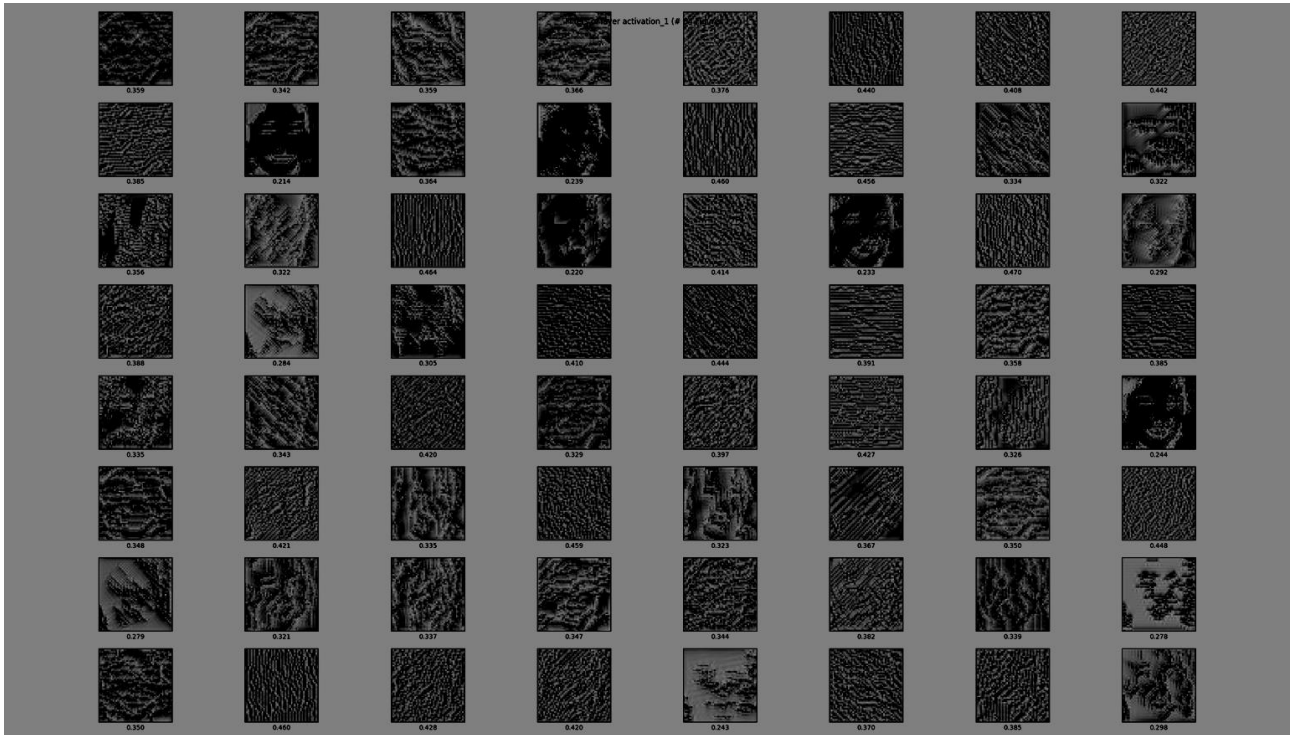


我的觀察是訓練出來的模型主要會專注在人臉的五官，眼睛、鼻子、嘴巴周圍的訊號通常較強烈。

5. 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：

(image 1) label->Happy



(image 10) label->Happy

