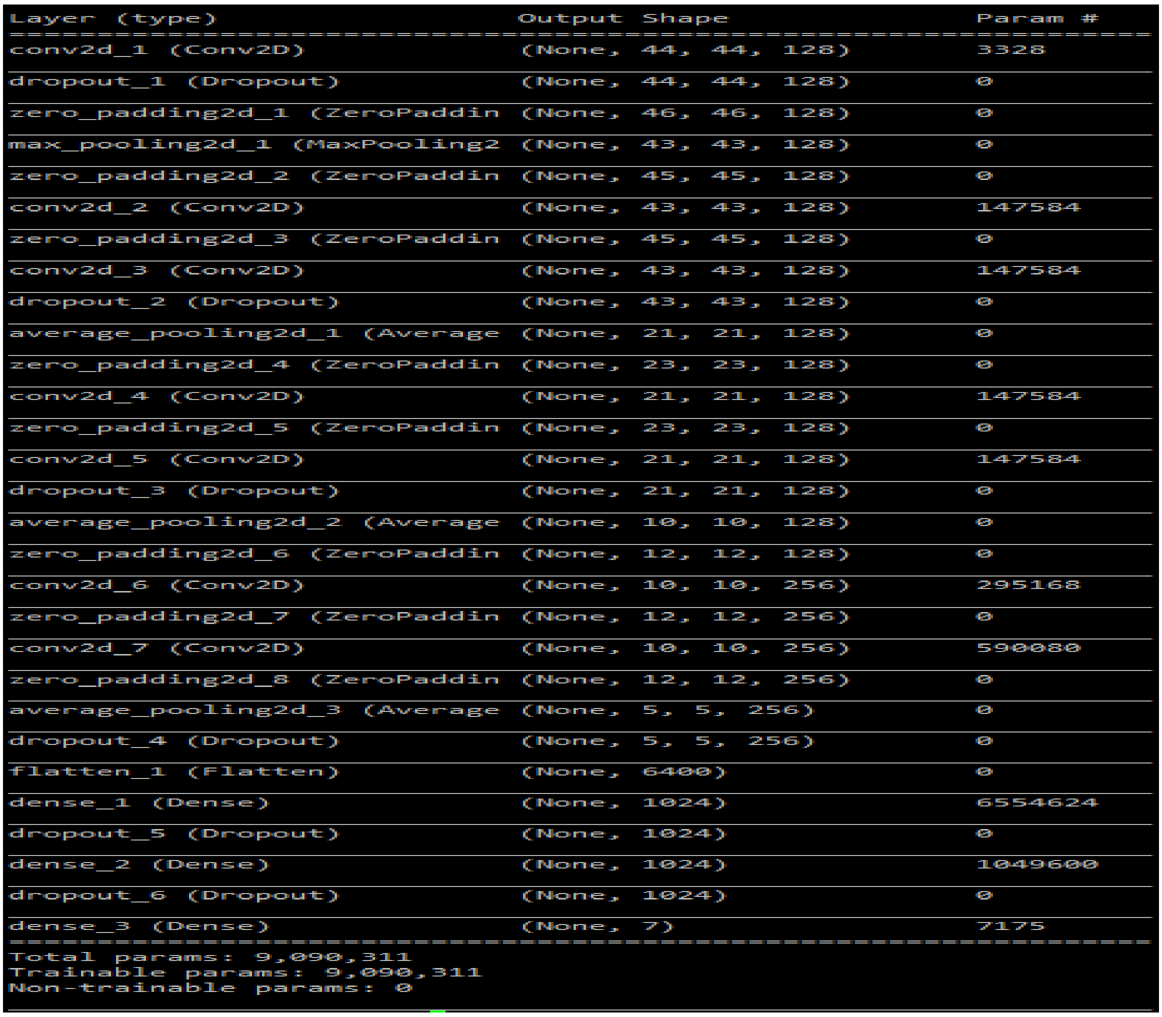
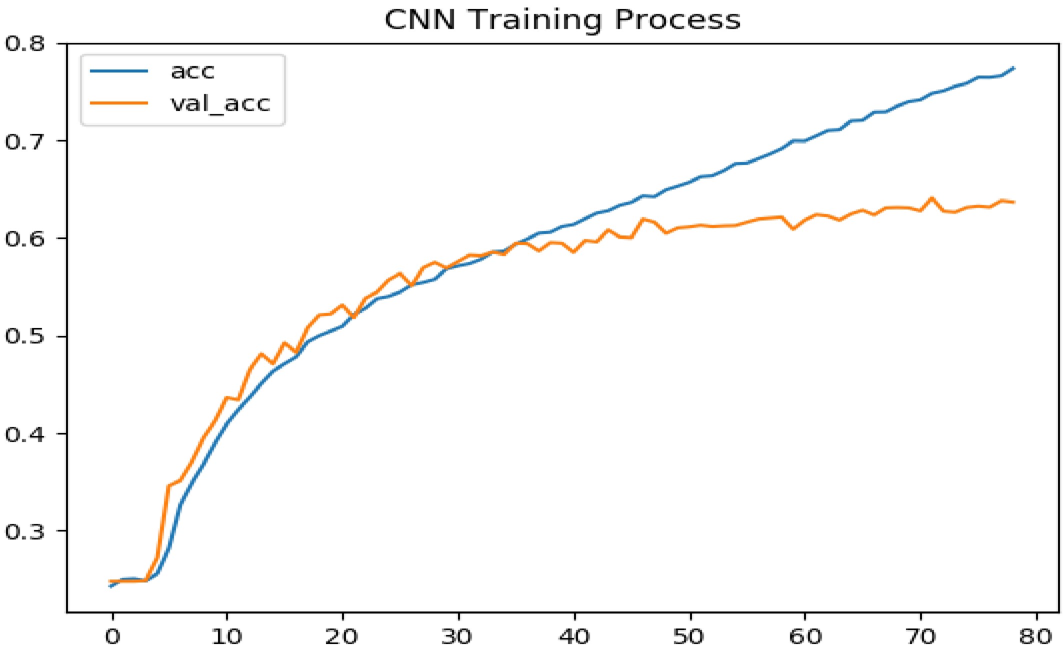
學號：B04901019 系級： 電機三 姓名：梁書哲

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？(Collaborators: 無)

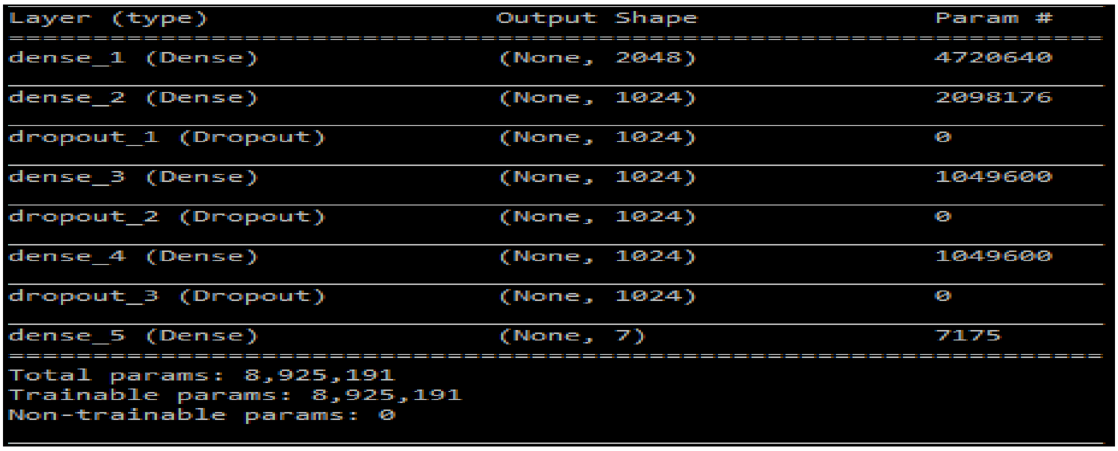
   答：Public Score:0.62022 / Private Score:0.63443



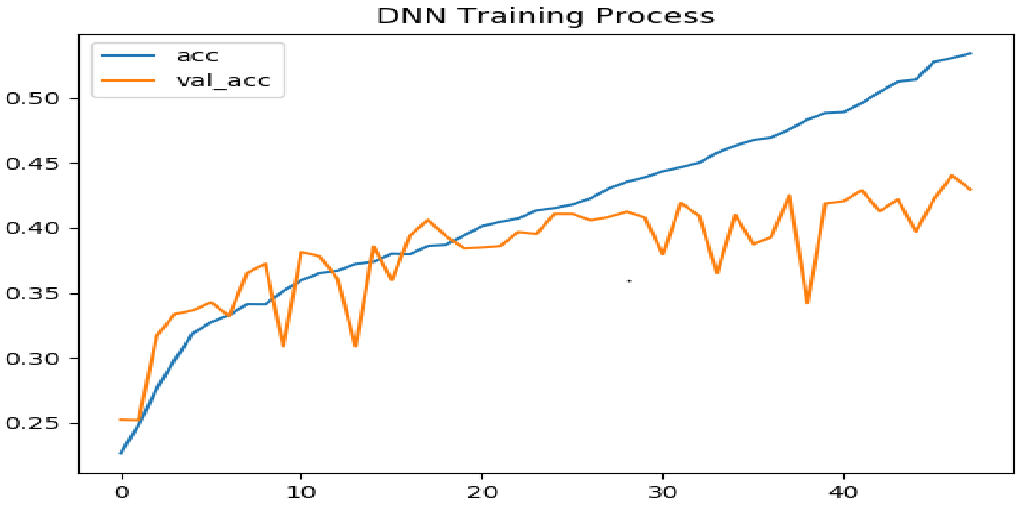
```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 44, 44, 128)       3328
_________________________________________________________________
dropout_1 (Dropout)          (None, 44, 44, 128)       0
_________________________________________________________________
zero_padding2d_1 (ZeroPaddin (None, 46, 46, 128)       0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 43, 43, 128)       0
_________________________________________________________________
zero_padding2d_2 (ZeroPaddin (None, 45, 45, 128)       0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 43, 43, 128)       147584
_________________________________________________________________
zero_padding2d_3 (ZeroPaddin (None, 45, 45, 128)       0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 43, 43, 128)       147584
_________________________________________________________________
dropout_2 (Dropout)          (None, 43, 43, 128)       0
_________________________________________________________________
average_pooling2d_1 (Average (None, 21, 21, 128)       0
_________________________________________________________________
zero_padding2d_4 (ZeroPaddin (None, 23, 23, 128)       0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 21, 21, 128)       147584
_________________________________________________________________
zero_padding2d_5 (ZeroPaddin (None, 23, 23, 128)       0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 21, 21, 128)       147584
_________________________________________________________________
dropout_3 (Dropout)          (None, 21, 21, 128)       0
_________________________________________________________________
average_pooling2d_2 (Average (None, 10, 10, 128)       0
_________________________________________________________________
zero_padding2d_6 (ZeroPaddin (None, 12, 12, 128)       0
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 10, 10, 256)       295168
_________________________________________________________________
zero_padding2d_7 (ZeroPaddin (None, 12, 12, 256)       0
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 10, 10, 256)       590080
_________________________________________________________________
zero_padding2d_8 (ZeroPaddin (None, 12, 12, 256)       0
_________________________________________________________________
average_pooling2d_3 (Average (None, 5, 5, 256)         0
_________________________________________________________________
dropout_4 (Dropout)          (None, 5, 5, 256)         0
_________________________________________________________________
flatten_1 (Flatten)          (None, 6400)              0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              6554624
_________________________________________________________________
dropout_5 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_2 (Dense)              (None, 1024)              1049600
_________________________________________________________________
dropout_6 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_3 (Dense)              (None, 7)                 7175
=================================================================
Total params: 9,090,311
Trainable params: 9,090,311
Non-trainable params: 0
```



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？
(Collaborators: 無)
答：Public Socre: 0.41877 / Private Score: 0.41766

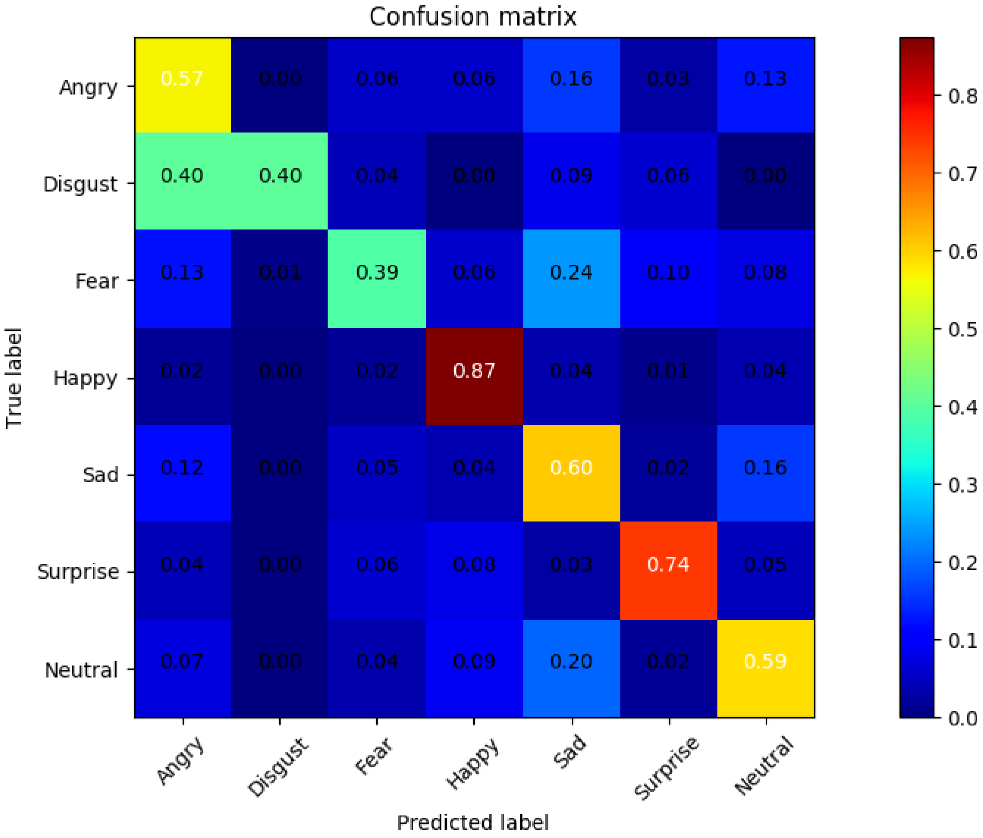


```
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 2048)              4720640
_________________________________________________________________
dense_2 (Dense)              (None, 1024)              2098176
_________________________________________________________________
dropout_1 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_3 (Dense)              (None, 1024)              1049600
_________________________________________________________________
dropout_2 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_4 (Dense)              (None, 1024)              1049600
_________________________________________________________________
dropout_3 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_5 (Dense)              (None, 7)                 7175
=================================================================
Total params: 8,925,191
Trainable params: 8,925,191
Non-trainable params: 0
_________________________________________________________________
```



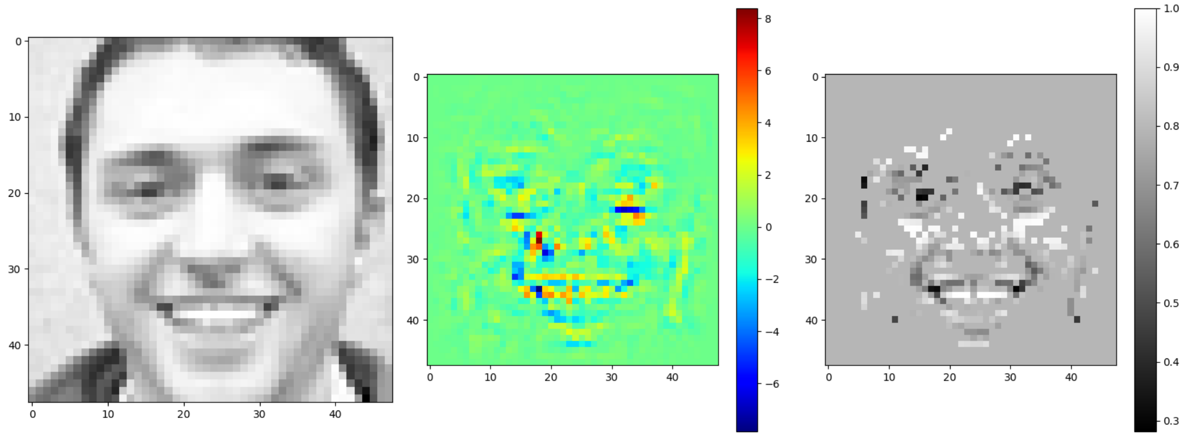
分數與 CNN 相比有顯著的下降，在訓練過程中，雖然 Training Set 的 Accuracy 持續上升，但 Validation Set 的 Accuracy 呈劇烈震盪，大約在 8 個 epoch 後就沒有穩定上升趨勢，在 50 個 epoch 左右因為 Early Stopping Callback 停止訓練。

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]
(Collaborators：無)
答：

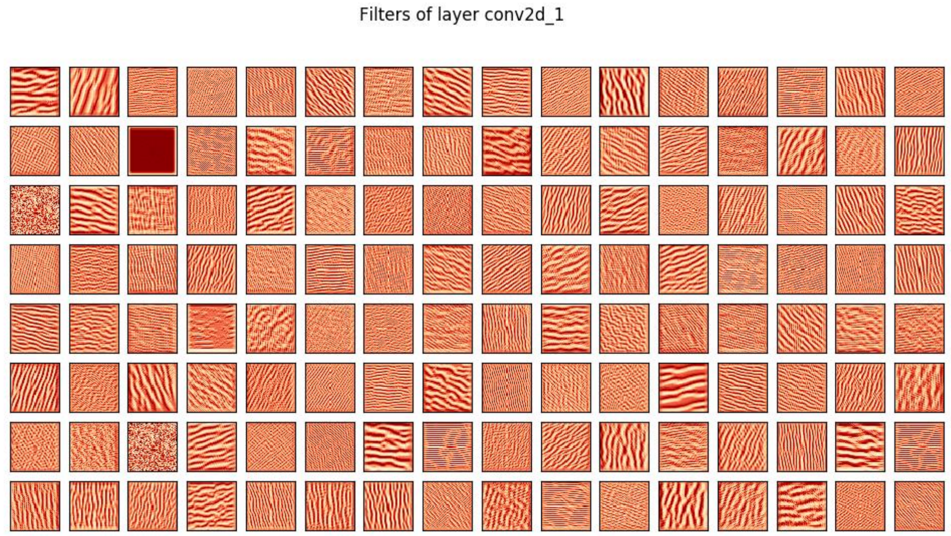


['angry','Disgust']最為明顯，大約 40%的'Disgust'Data 被 Model 辨認為'Angry'，而['Fear','Sad']也有 24%的'Fear'Data 被 Model 辨認為'Sad'，其他部分回答正確的比率皆明顯高於其他錯誤的選項，其中'Happy'的準確率達到 87%為所有 Label 之中最高。

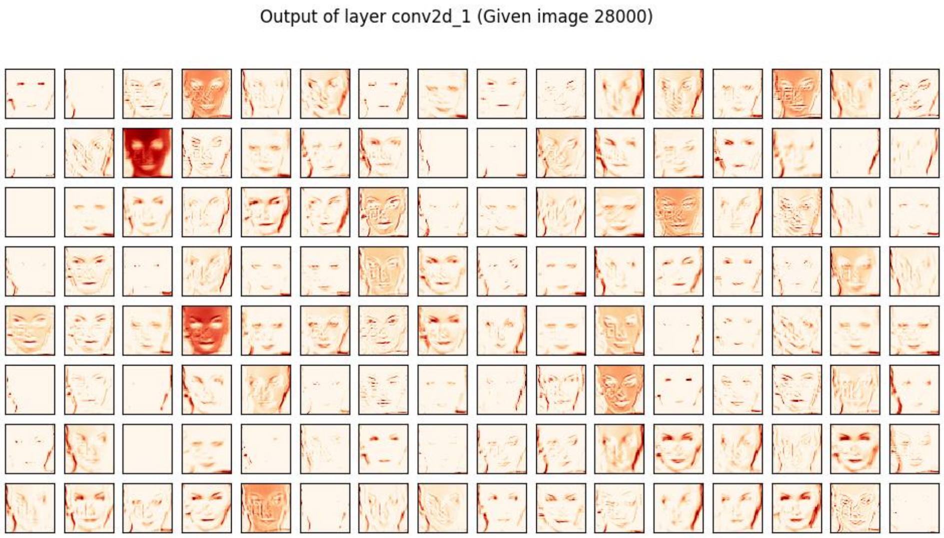
4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？
(Collaborators：無)
答：從 saliency Map 可以看出 Model 主要注意在眼睛及嘴巴的部分。



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。
(Collaborators：無)
答：



第一層的 Filter 從結果看主要是條紋的紋理能夠 activate，而各 Filter 對應到各種粗細、角度的紋理。



實際將 Validation Set 的照片餵入 Filter 後發現，幾乎每個 Output 眼睛的部分都有被 activate，而嘴巴及臉頰輪廓也有在大部分的 Output 中被 activate。