

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: 蔡孟庭、葉韋辰、鄭克宣、丁縉楷)

答：

共有五層Conv2D (每層後面都接BatchNormalization、ZeroPadding2D)

另Flatten後有兩層Dense(1024)(且有做Dropout)

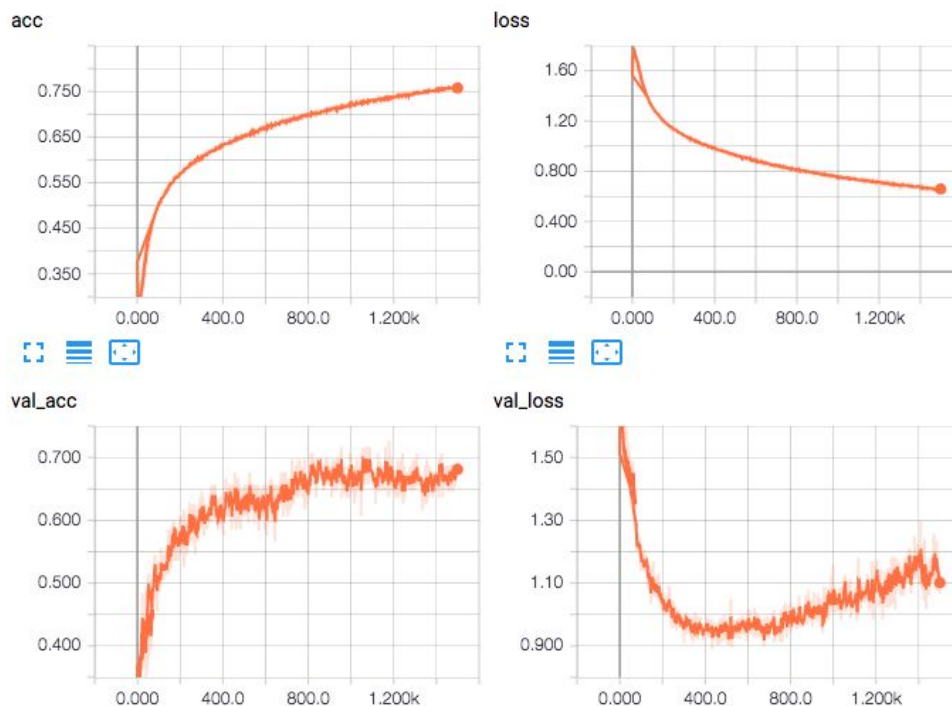
最後輸出七種結果(Dense(7))

optimizer使用adam

loss function使用categorical_crossentropy

訓練過程的部分，因為使用adam 所以val_loss動盪較大 (跟ada比起來)

比較有趣的是 val_loss到epoch 600開始 就往上飆升了，但是其實預測結果，還是epoch最多次的預測出來越好，覺得應該是因為validation set我只有切100個 太少的關係



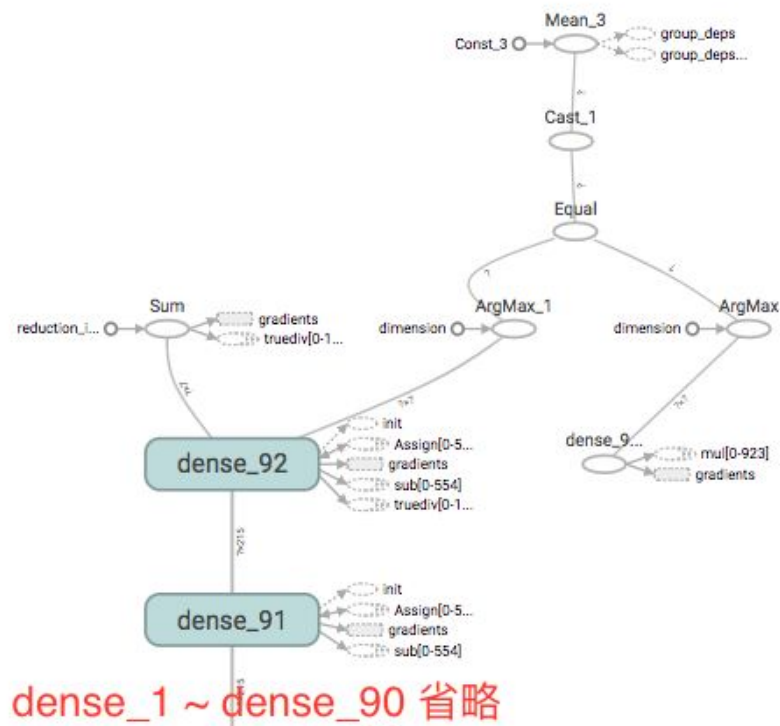
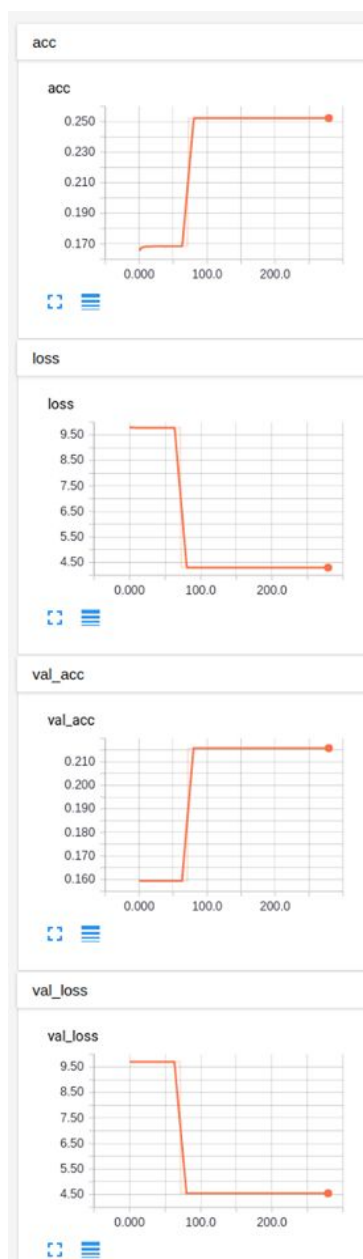


2. (1%) 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model。其模型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

(Collaborators: 蔡孟庭、葉韋辰、鄭克宣、丁縉楷)

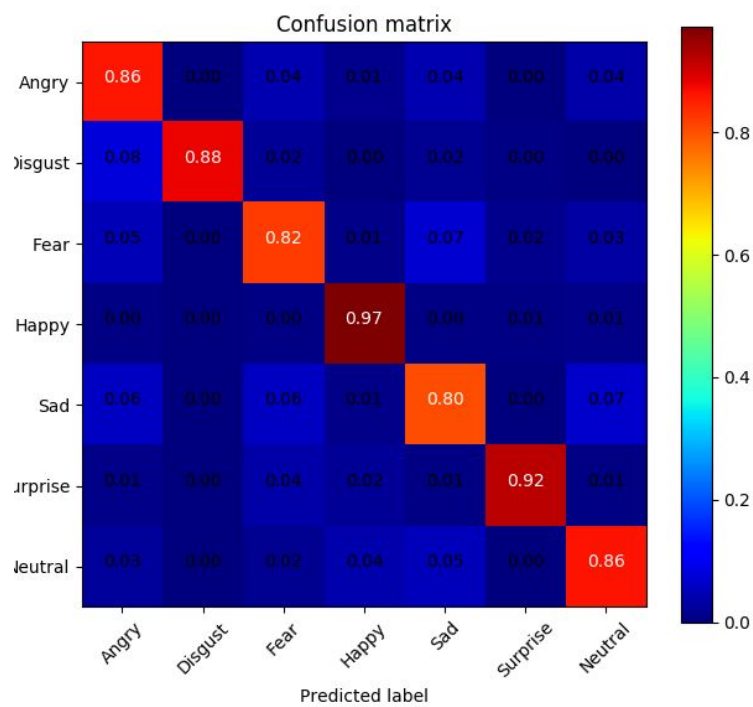
答：架了90+2層 (input、output)、每一層215個neuro、共四百六十萬 parameters

很明顯可以看出來無論是train data還是test data，準確度皆沒有比CNN來得好



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: 蔡孟庭、葉韋辰、鄭克宣、丁晉楷)

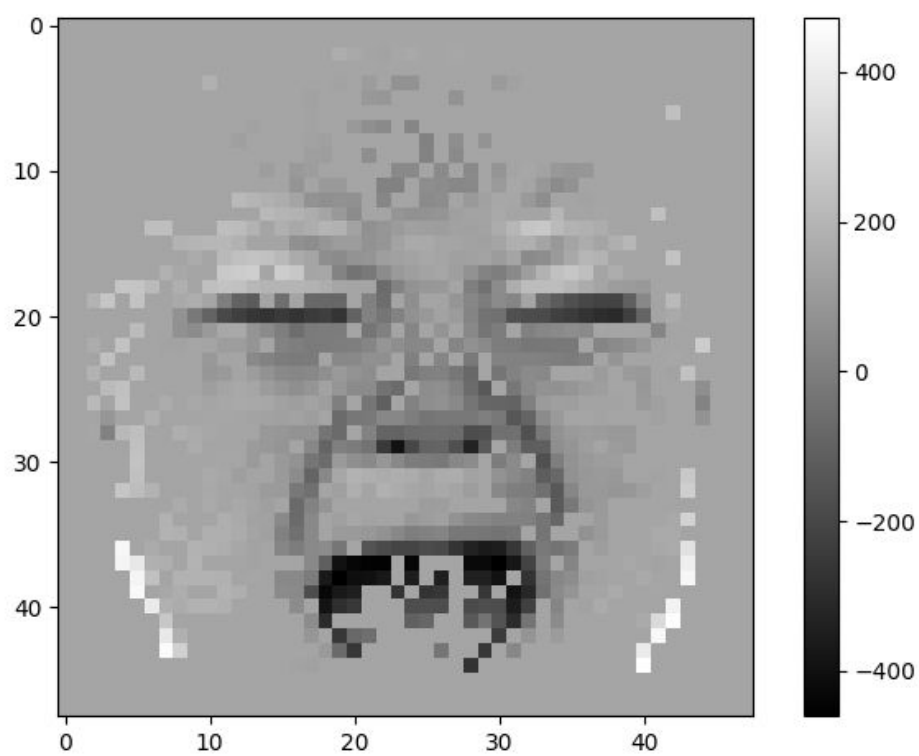
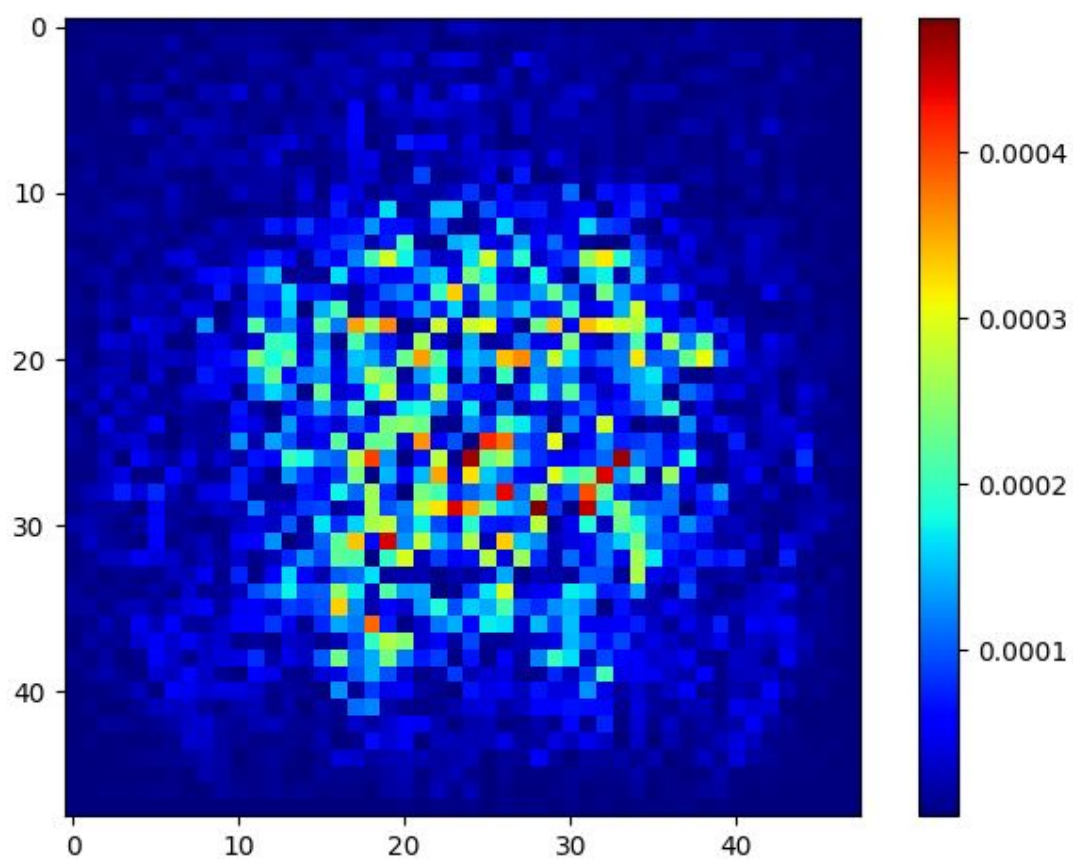
答：angry 被判成disgust、sad被判成fear、neutral被判成sad的機會高



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: 蔡孟庭、葉韋辰、鄭克宣、丁晉楷)

答：比較focus五官的部分

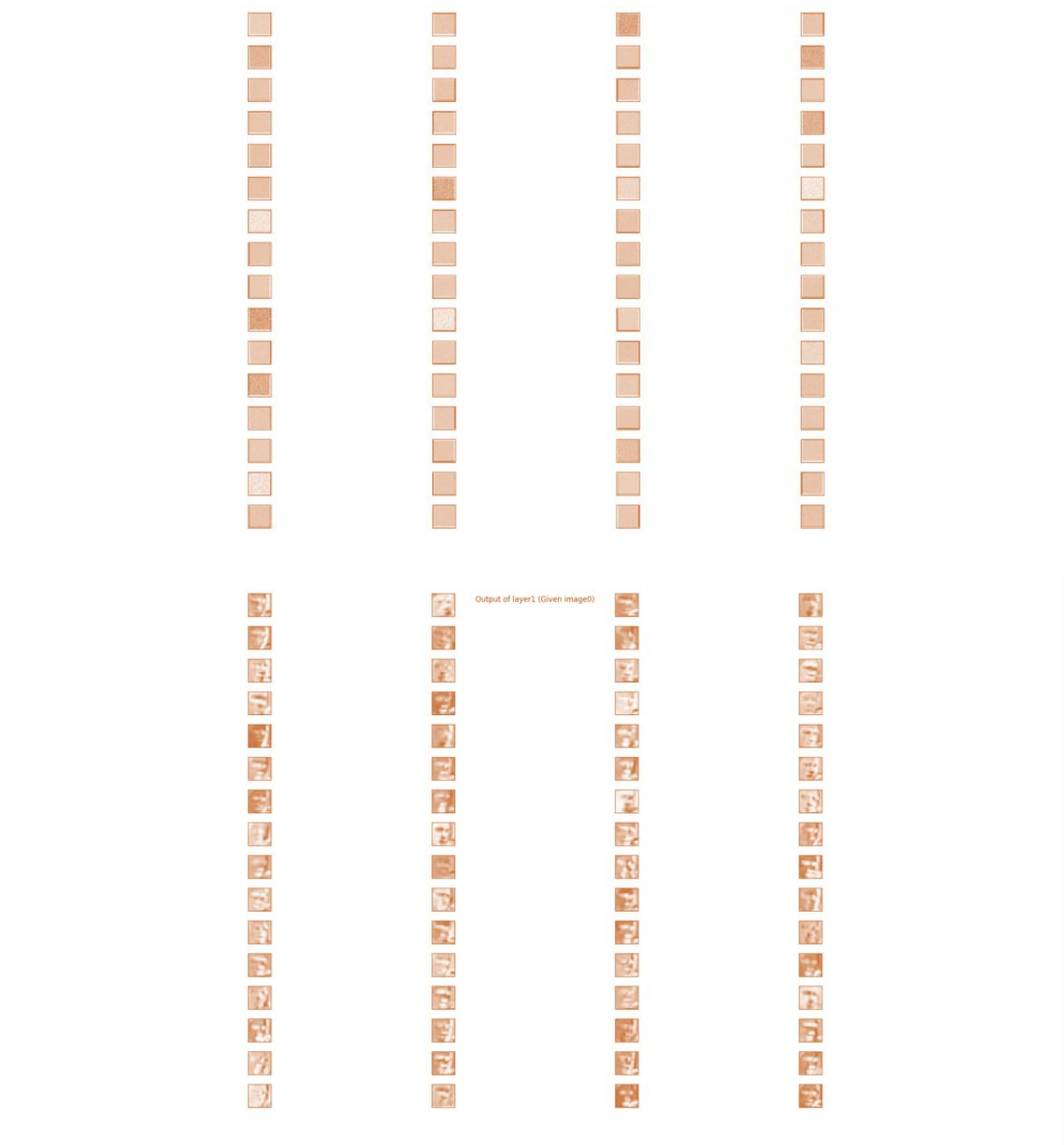




5. (1%) 承(1)(2)，利用上課所提到的 gradient ascent 方法，觀察特定層的filter最容易被哪種圖片 activate。

(Collaborators: 蔡孟庭、葉韋辰、鄭克宣、丁晉楷)

答：





Output of layer2 (Given image0)



Output of layer0 (Given image0)

