

學號：R05944019 系級：網媒一 姓名：張嘉豪

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：

這次的圖片我有先做預處理，用 ImageGenerator 將圖片做左右平移和鏡像，有試過旋轉 1 度，但是增加旋轉之後的效果沒有比較好，因此就將旋轉的部分拿掉。我覺得是因為有些影像經過旋轉之後，反而會判斷不出原本屬於哪個 class 而直接影響到 train 的結果。

	旋轉 + 平移 + 鏡像	平移 + 鏡像
Validation Set 準確率	58.3%	63.8%

表 1: 對圖片做不同預處理的效果比較

這次實做的 CNN model 我實作了兩種不同的 model，一個 neural network 比較深但是每一層的 neuron 比較少 (CNN1)，第二種是每一層的 neuron 比較多，但是深度比較淺 (CNN2)，結果可以看出來深度較深的效果會比較好。

	CNN1	CNN2
Validation Set 準確率	65.4%	63.8%

表 2: 兩種 CNN 效果比較

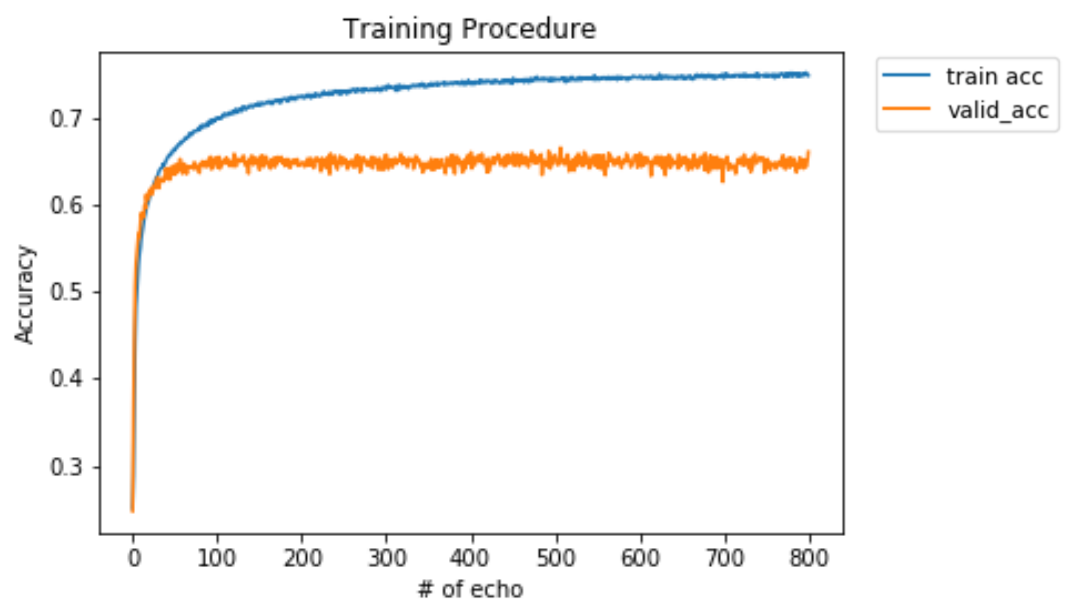


圖 1: CNN Training Procedure

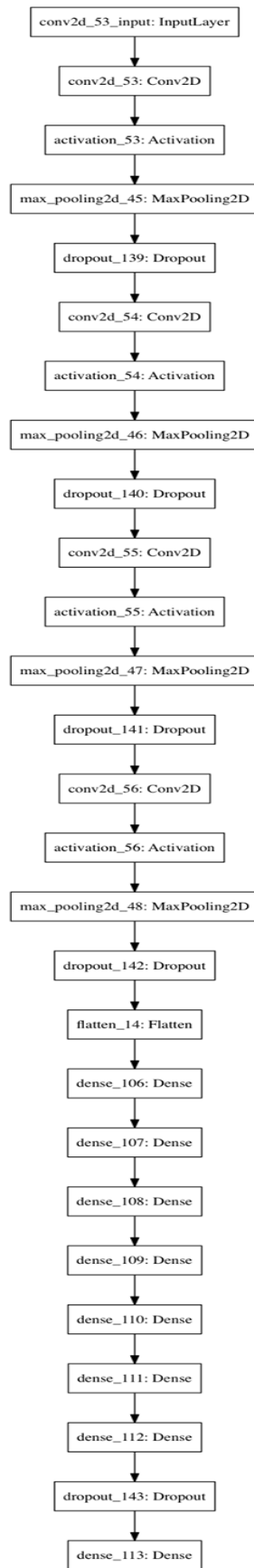


圖 2: CNN 架構

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：

第一題的 CNN 模型參數大概在 300 萬左右，這一題的 DNN 模型使用了大約 290 萬個參數，可以發現 DNN 跑出來的結果 training data 的準確度會上升，但是 validation data 的準確度一直都上不去。

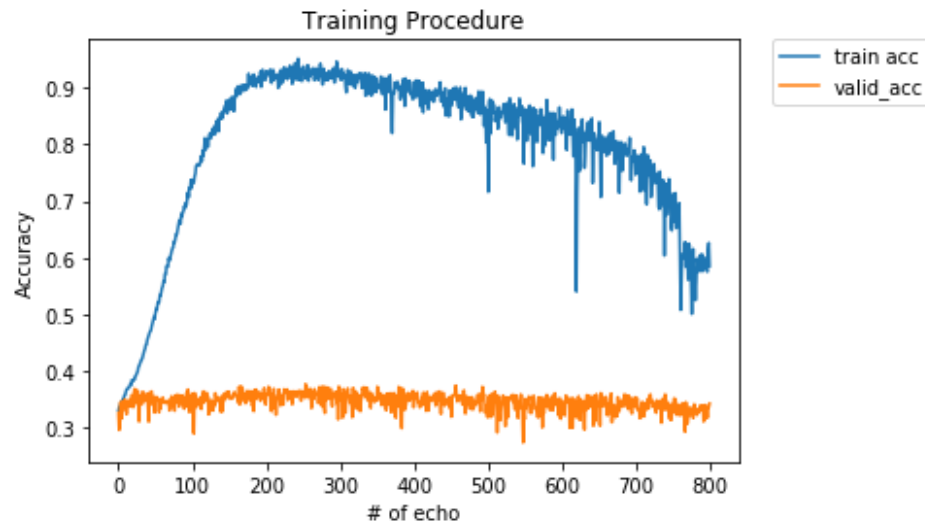


圖 3: DNN Training Procedure

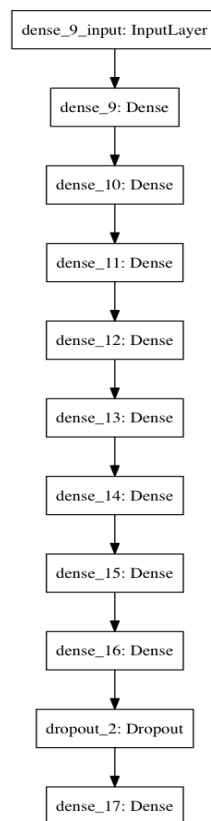


圖 4: DNN 架構

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：

我將 training data 的 15%當作是 validation set，然後將 validation set 套在跑好的 model 後，畫出一個 confusion matrix，由這張圖可以發現這個 model 會將 Disgust 誤判為 Angry 的機率最高，其次是把 Fear 誤判成 Sad，我猜想是因為這兩兩之間的表情和情緒大致上都蠻接近的，所以才會有這樣的結果。

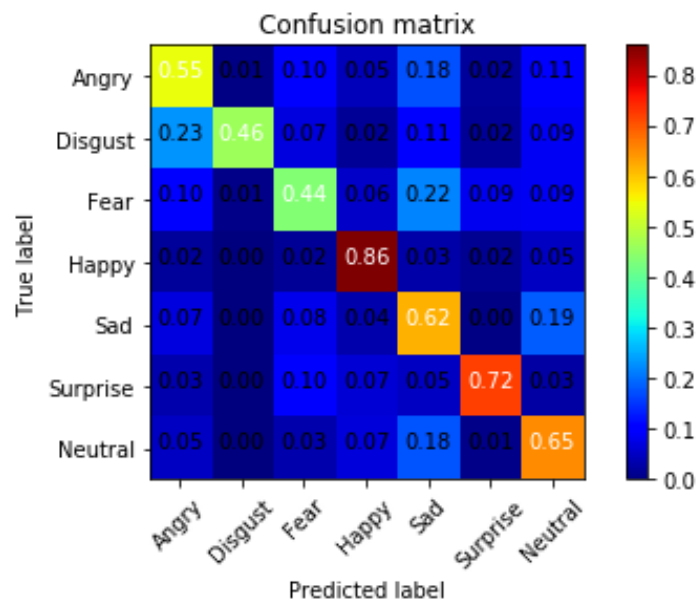


圖 5: Confusion matrix

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：

我是用第二層 CNN 輸出的結果來分析，我發現在做分類的時候，會 focus 在臉頰的部分，可見得臉頰的部分是判斷一個人表情蠻重要的位置。

Saliency map

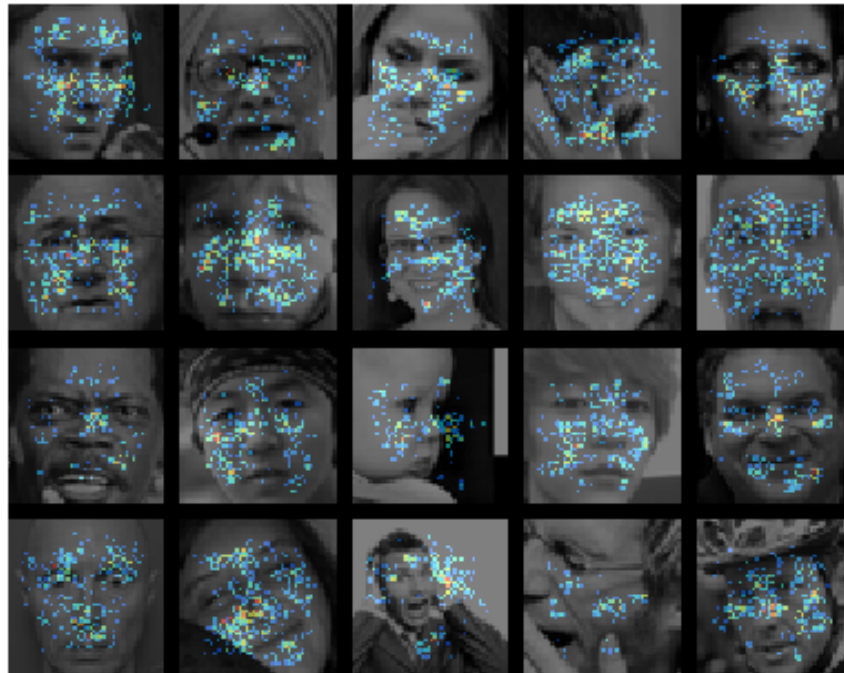


圖 6: Saliency map

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：

一開始我用第一層 activate filter，但是我發現第一層跑出來的 filter 感覺都不太明顯，後來就改用第二層的 filter，可以發現到 filter 變得比較明顯也比較有變化，直條紋跟橫條紋的 filter 變得多了很多。

Filters of layer conv2d_54

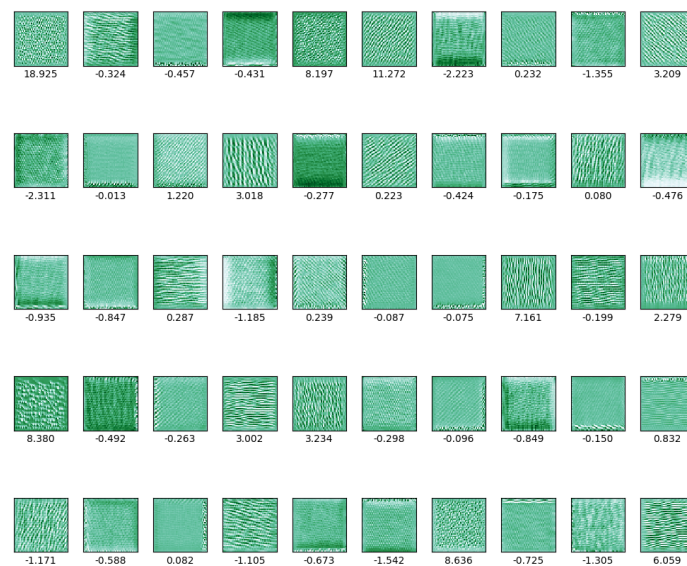


圖 7: Filter

Output of layer0 (Given image200)

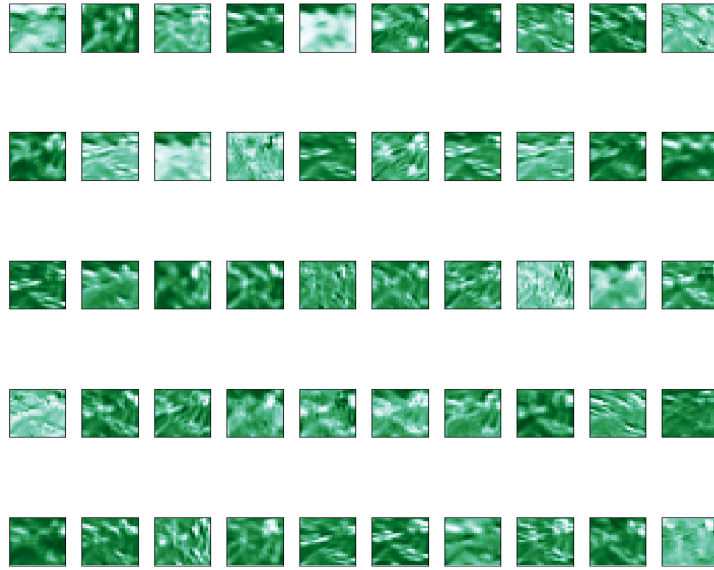


圖 8: Image after filtering

[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%]