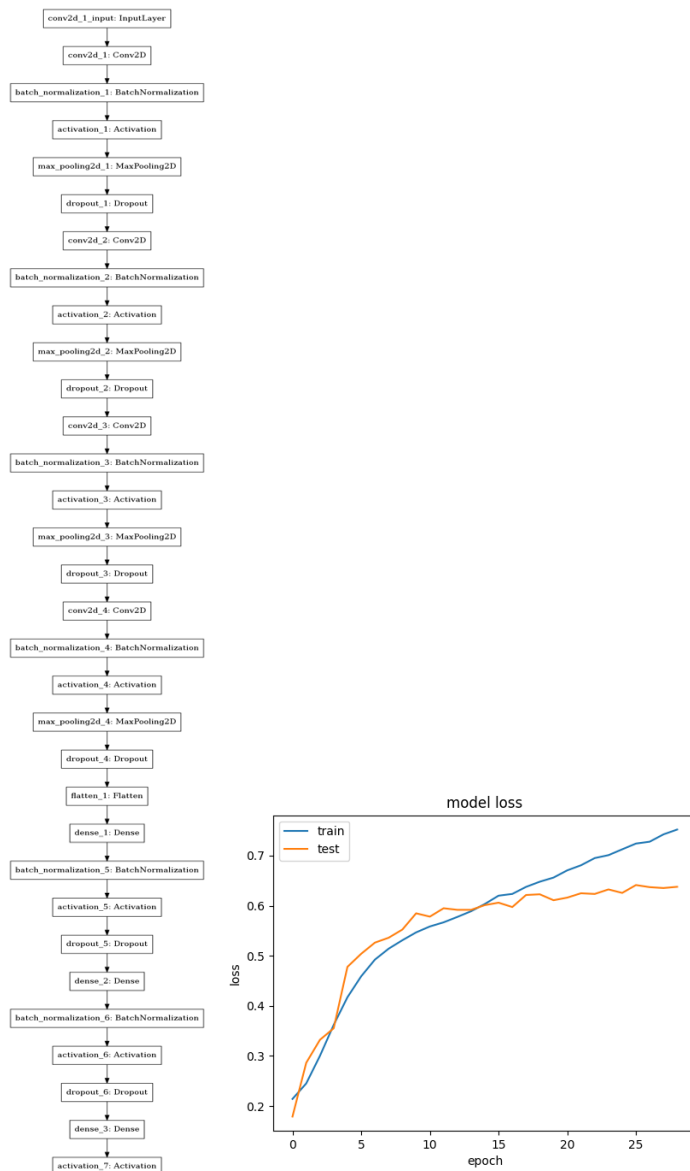


學號：R06922116 系級：資工碩一 姓名：賴柏恩

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:NONE)

答：



此 model 在 kaggle 上分數大約為 61%的準確率

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:NONE)

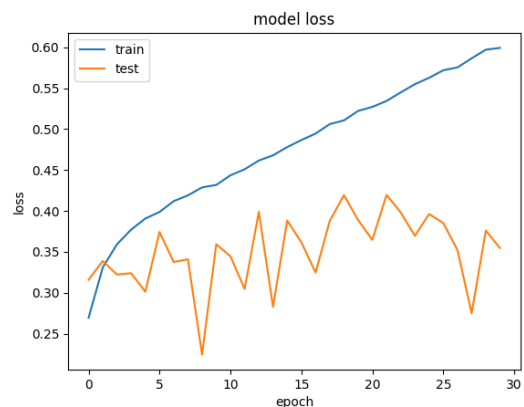
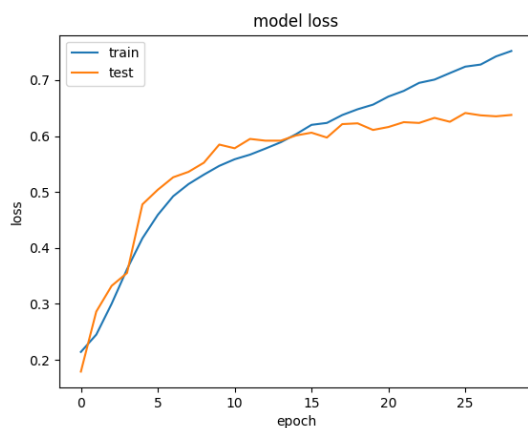
答：

CNN: 4448623 parameters, DNN: 4522727 parameters

CNN

DNN

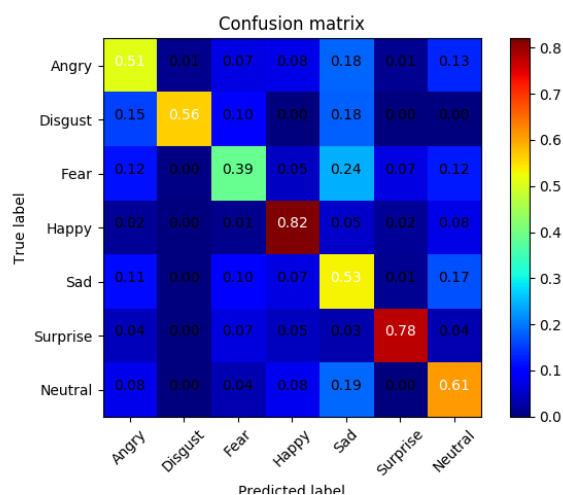
Total params: 4,448,623 Trainable params: 4,445,371 Non-trainable params: 3,252	Total params: 4,522,727 Trainable params: 4,517,279 Non-trainable params: 5,448
---	---



D N N 的 model 為三層 fully connected 的 N N 疊起來，kernel 數分別為 1200, 1024, 500, 而光從 val 的表現就可看出 D N N 的表現較差，所以我並沒有丟到 kaggle 上去看，但應該表現並不會很好，但可以從這裡看出，在辨識影像上，將周圍資訊一起去考慮是會增加其表現的。

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:NONE)

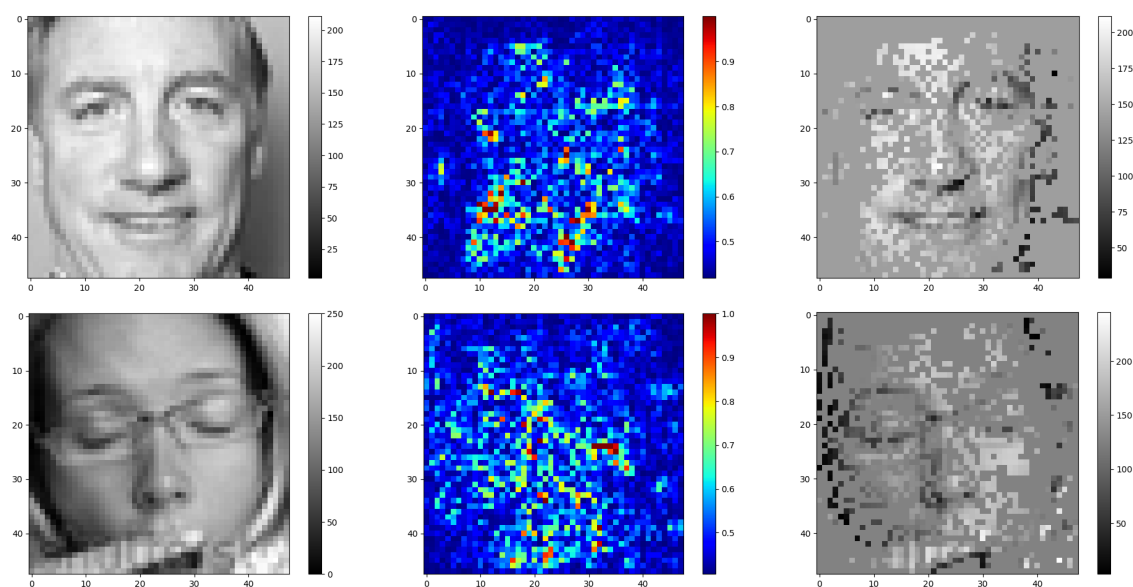


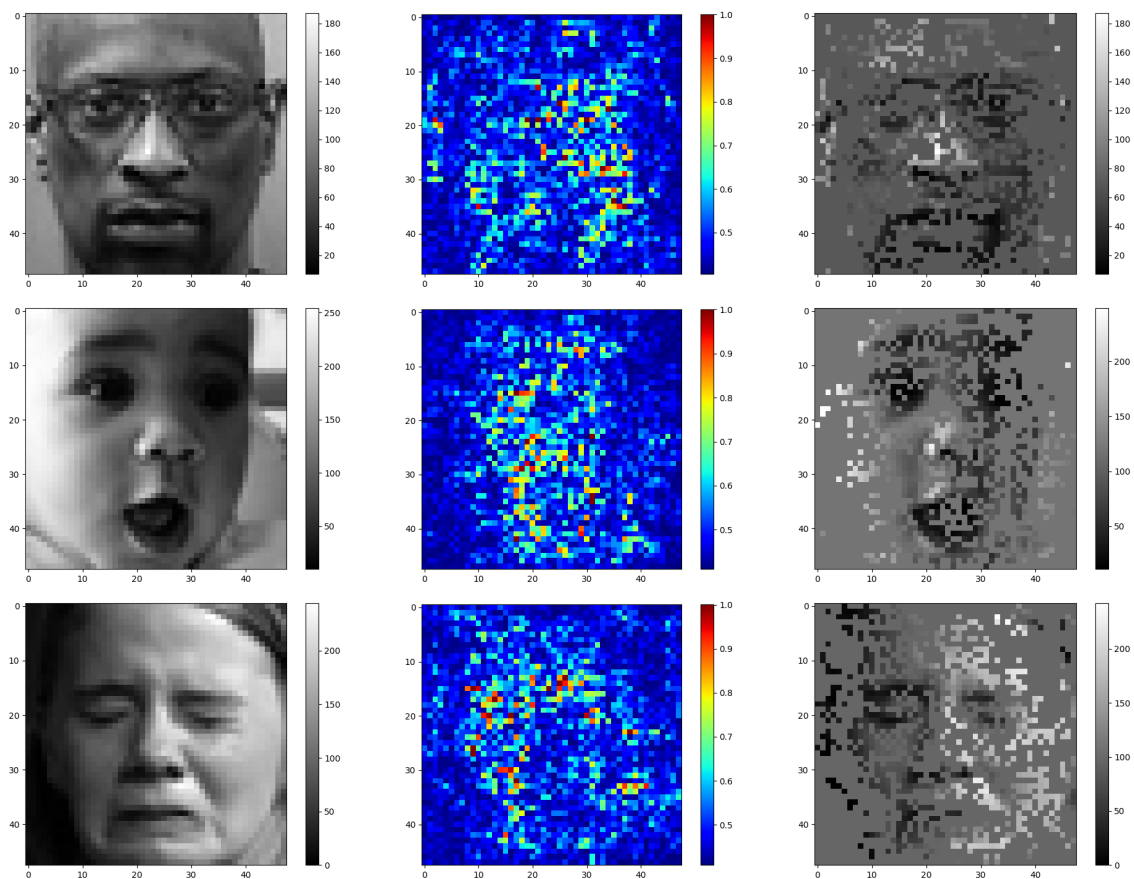
答：

從 confusion matrix 可以看出，model 在辨識 happy 與 surprise 時準確率較佳，而在辨識 fear 時經常將他與 sad 搞混，可能是因為這兩個表情的特徵相較之下較不明顯，導致 model 較無法正確的 predict 出。

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:NONE)

答：



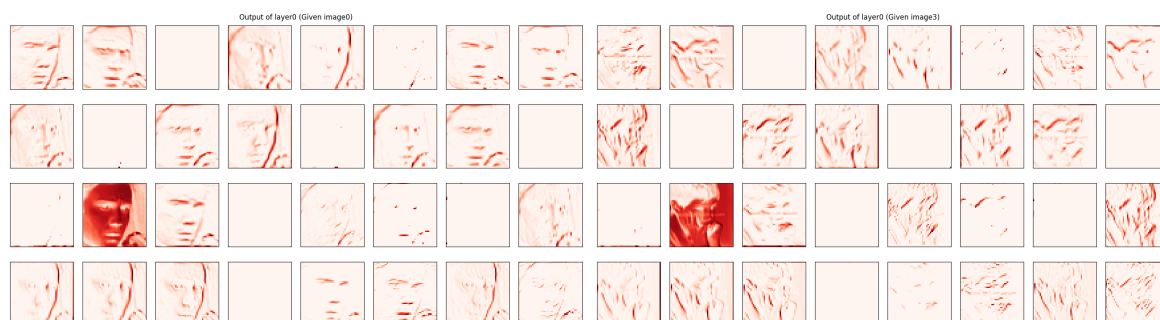


從 saliency map 中，我們可以看出 model 大部分都會保留了嘴巴及眼睛的部分，因此 model 進行 classification 時，應該是 focus 在嘴巴及眼睛的部分去進行分類。

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:NONE)

答：





圖中皆為圖片被第一層 cnn 內 filter activate 得結果，我們可以看出在第一層中，每一個圖片被 activate 的其實都差不多。

Note11/12: 在撰寫完這份報告後幾天，因為我一直都有在改善 model，我接下來有使用了 data generator 去改善 model 準確率，然後不停改變 model 但始終沒有突破 strong baseline。

Note11/14: 因為一直無法突破 strong baseline，後來使用了 ensemble 的方法，將過去 train 的 model 組合再一起，終於突破了 strong baseline

Note 11/15: 又重 train 了一次 model，改變了一下 data generator 的參數，這次終於突破 strong baseline