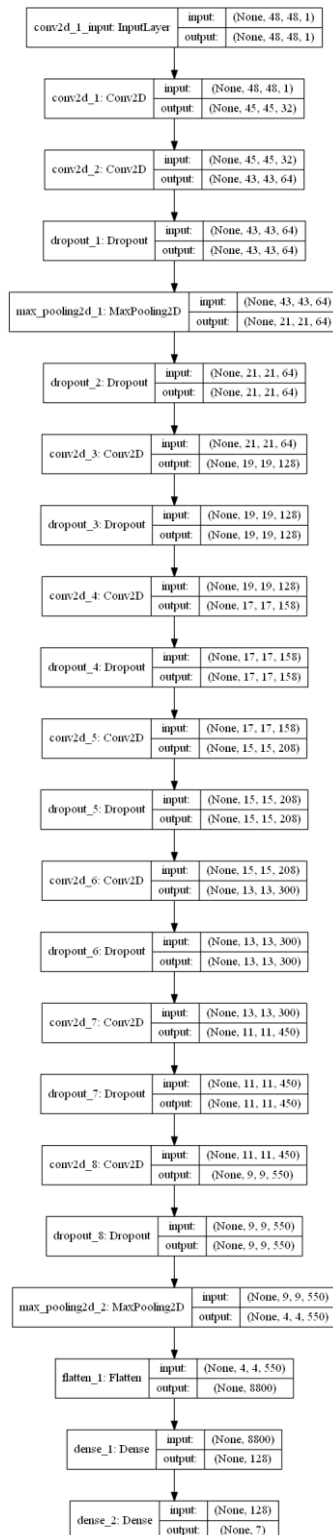
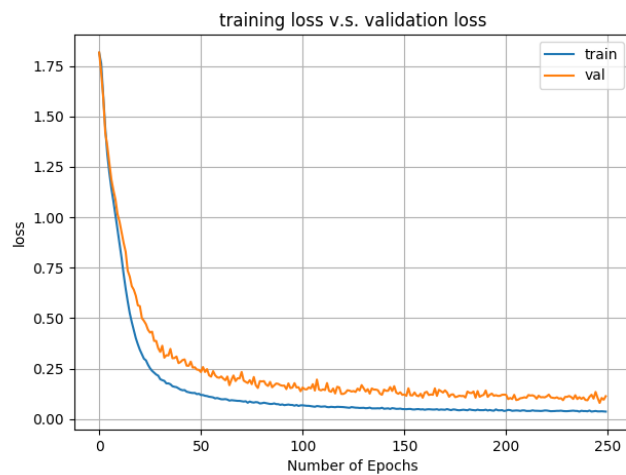


1. (1%) 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？

答：

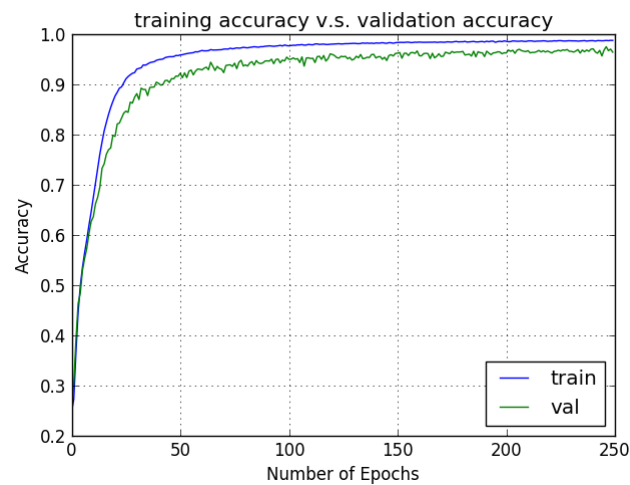


CNN 所使用的模型如左圖所示：使用了七層 conv2D、2 層的 maxpooling、2 層的 dense、中間穿插 dropout



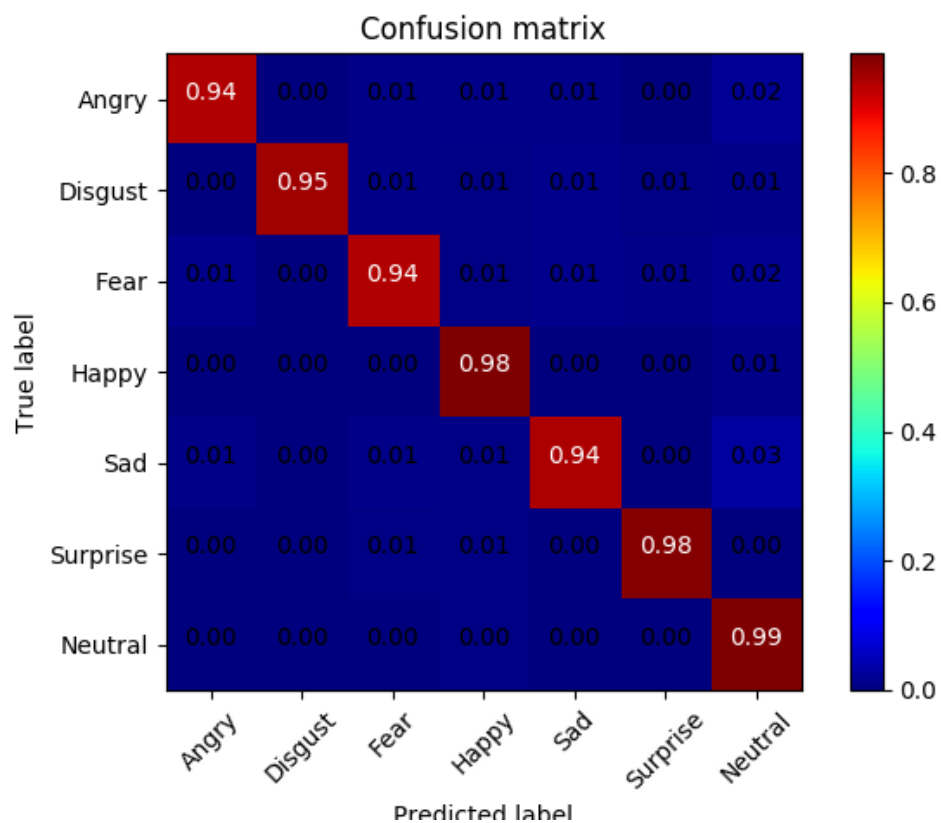
上圖為使用 cross entropy 的 training/validation loss

(validation set 為總資料的 20%)



上圖為 training/validation 的 accuracy

在 training 及 validation set 中，均有非常高的準確率

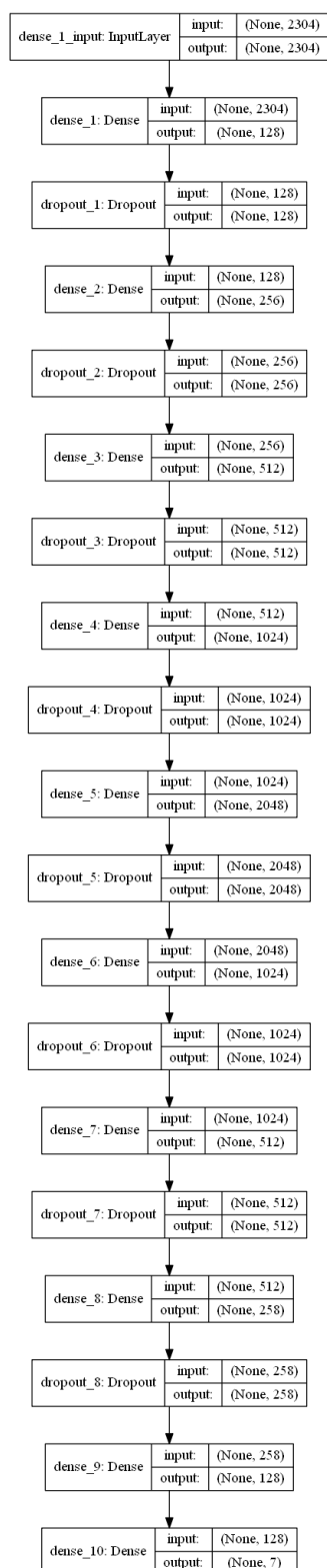


Confusion matrix 也印證這點。但是在 kaggle 上卻只有 0.6 的預測準確率。

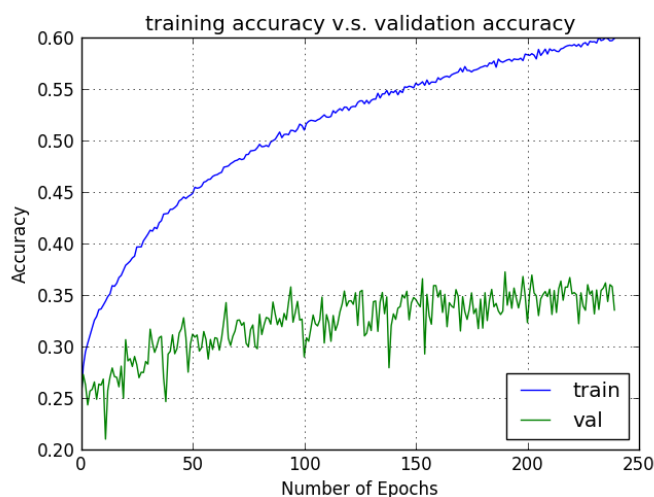
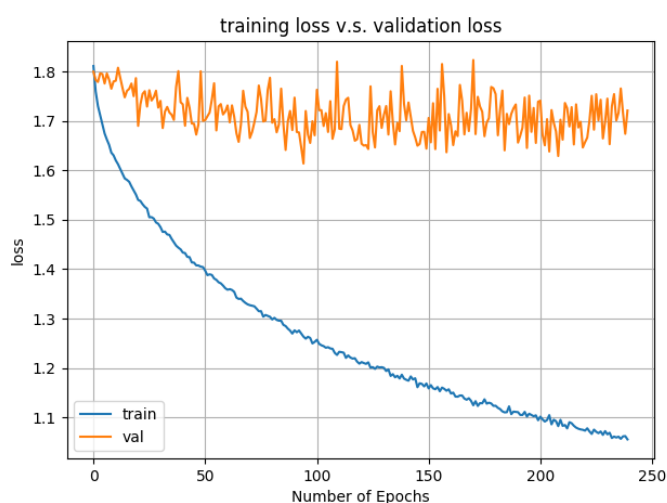
因為我有對圖片做 `zca whiten` 的處理，一開始以為是在 testing set 上的處理有問題才造成這麼大的落差，但多次驗證後排除此種可能性。後來想有可能是 keras 對於 `validation_split` 的處理跟文件上的說明有誤，才造成 validation set 被拿去當成 training 的資料，但是在 DNN 的 training 中卻又否定了這論點。因此目前還不知道主要造成此現象的原因。

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：



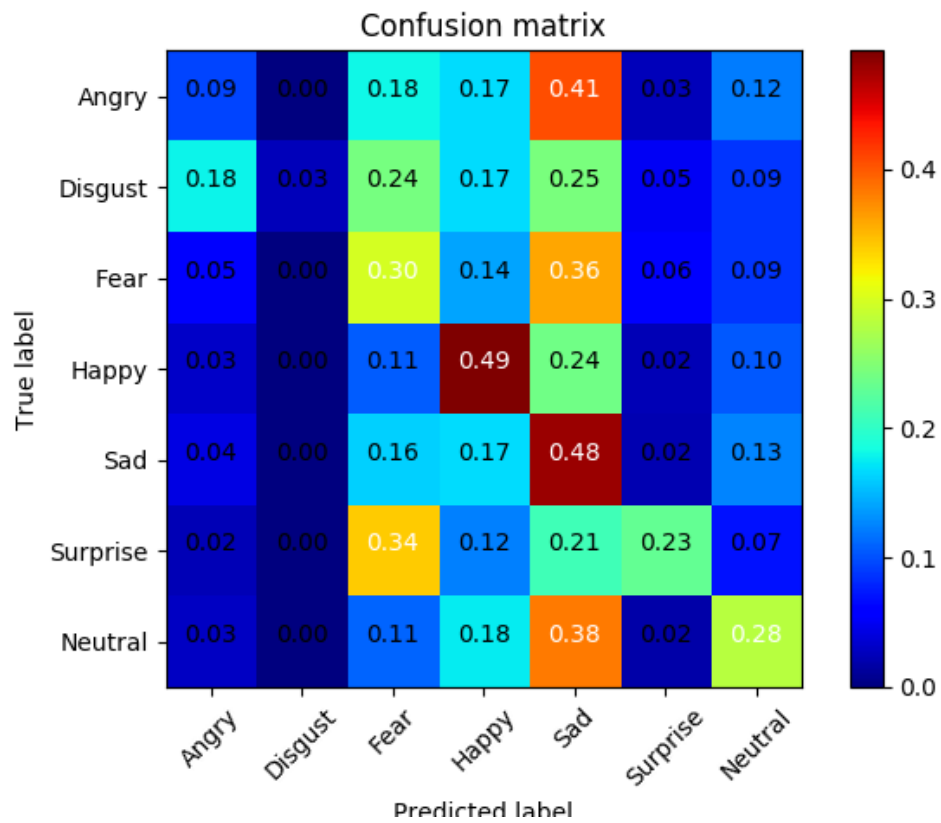
在參數量與 CNN 差不多的情況下，嘗試許多不同的組合後，預測準確率大致上在 0.3~0.4 之間，找不到一個有效率提升準確率的方式。



由上面 2 張圖可以看出，在 validation 中的落差起伏相當大，而且上升的幅度緩慢。相較於 CNN，更難找出一個適合用在影像辨識的 model。

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]

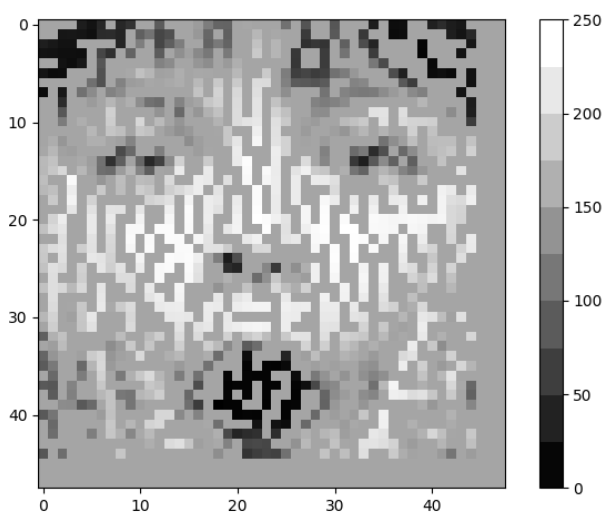
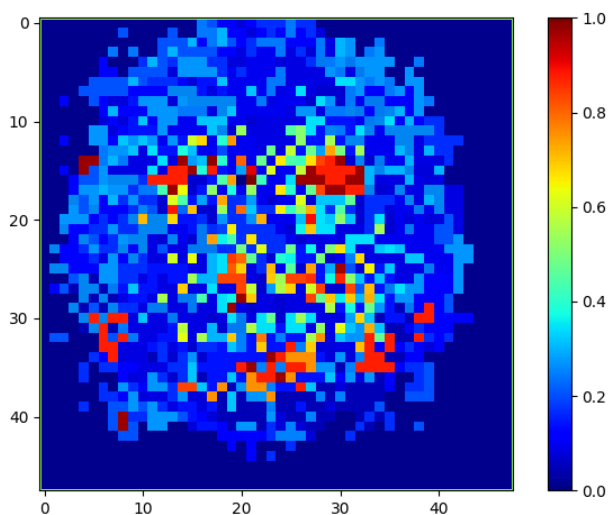
答：



Disgust 在整體的資料量中算相對少，所以無法有效分辨其類別。
而 Fear 容易與 disgust 和 surprise 混淆。Sad 也容易和 neutral、fear 甚至 angry 混淆。

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：



大致上可以看出熱點集中在臉上，又以眼睛、嘴巴感覺比較明顯。

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：