

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

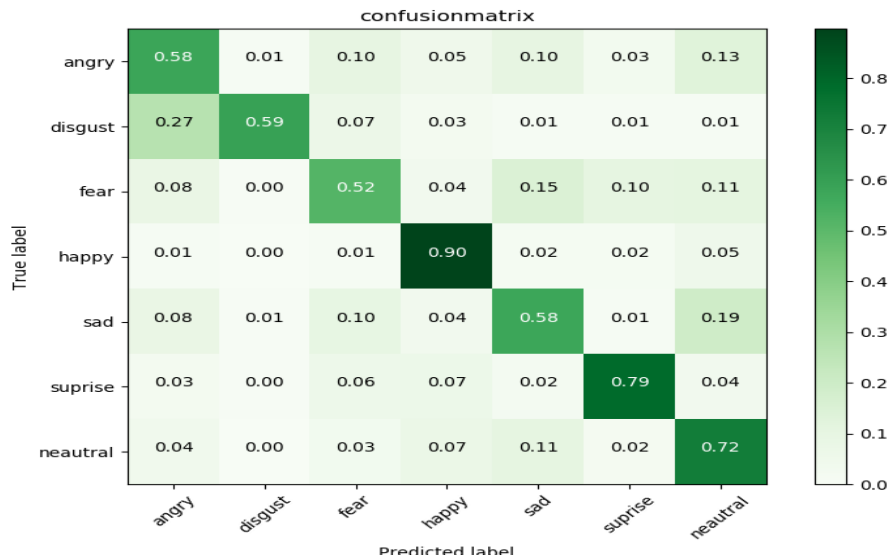
這次的模型我使用了四層的 convolution 加上四層的 Fully connected 模型，前四層分別有 50，200，400，400 個 filter，在 Fully connected 分別有 500，300，200 的隱藏層，在每一層的連接我都有做 normalization 以及 dropout，沒有做 dropout 的話在樣本內的 accuracy 可以來到 0.97，但最後發現會有 overfitting 的問題，因此我在每一層都 dropout 了 30 percent，outsample 準確度提高了 2 percent 左右，另外我的 batch size 一次抓 128 筆資料，每一個 epoch 中更新八百多次，我已 callback 存下 outsample 表現最好的模型，最後存下為第 117 的 epoch，上傳 kaggle public score 為 0.686 跟我切最後五千筆來做 validation 的準確度差不多，模型參數為 4,272,107 個。

2. (1%) 請嘗試 data normalization, data augmentation, 說明實行方法並且說明對準確率有什麼樣的影響？

data normalization 這邊使用 20 個 epoch 作為比較，如果未做 normalization 跑完 20 個 epoch，樣本內預測準確度級樣本外準確度為 0.39 及 0.42，做完 normalization 樣本內預測準確度級樣本外準確度為 0.54 及 0.58，我們可以看出做 normalization 除了準確度可以提升為，收斂數目也會較快。

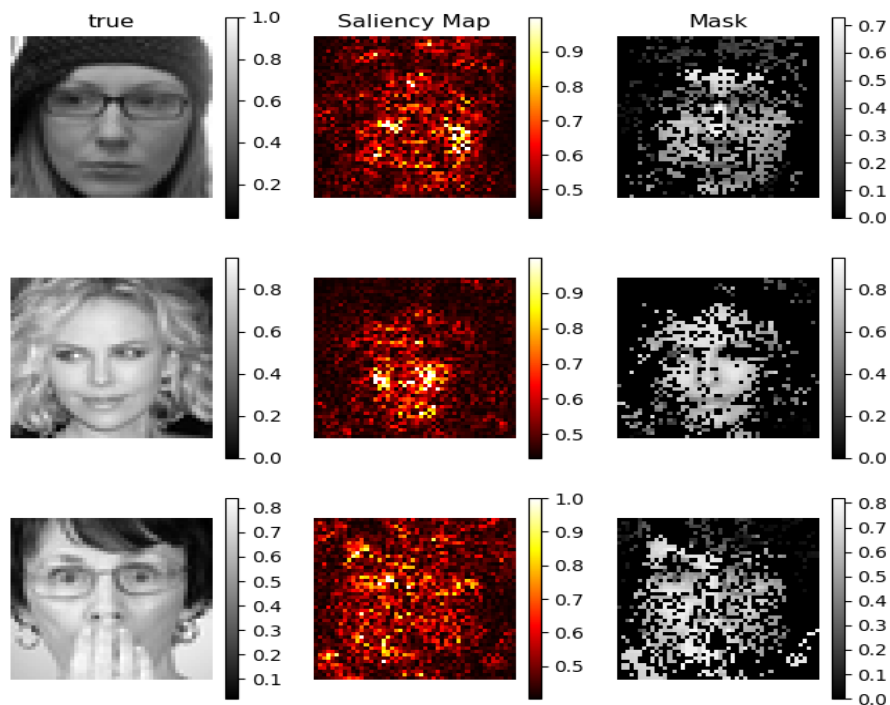
在一開始未做 data augmentation 前，train 模型的速度很快，但是上傳 kaggle 上的分數在 0.55 左右徘徊，後來對圖片做平移、翻轉、選轉、縮放、推移之後，上傳 public score 的成績就到 0.66，準確度提升了很多，然而 train 模型的時間相對較久。

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]



可以看到開心的準確度為最高，噁心最容易誤判成生氣，機率約為 0.27；第二高為傷心易誤判為中立，機率約為 0.19；第三名為恐懼誤判傷心，機率為 0.15，第四第五高為生氣易誤判為中壢以及恐懼，機率分別約為 0.13 及 0.1。

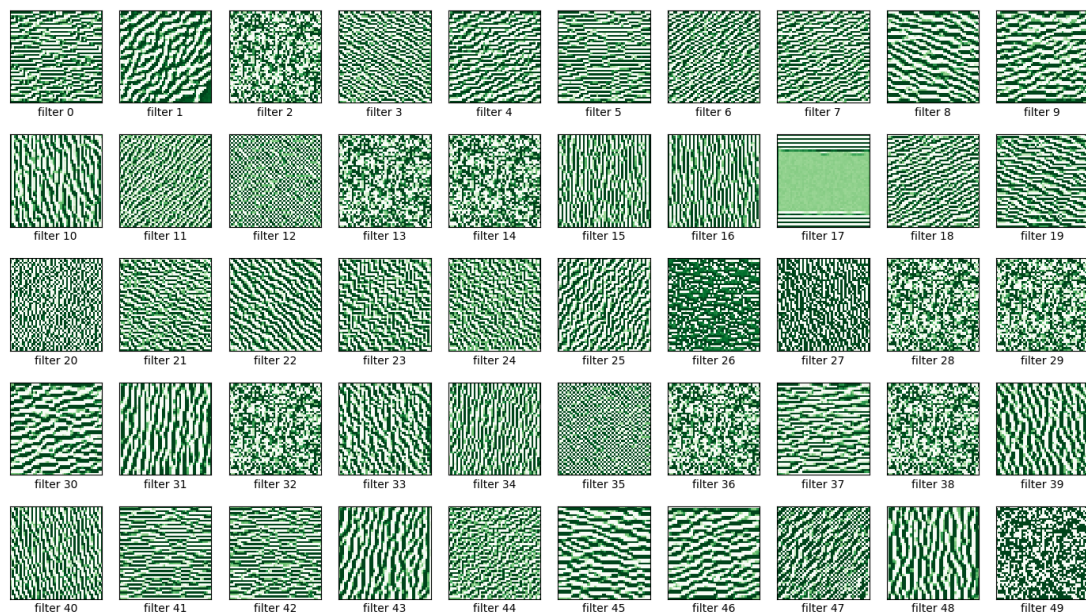
4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？

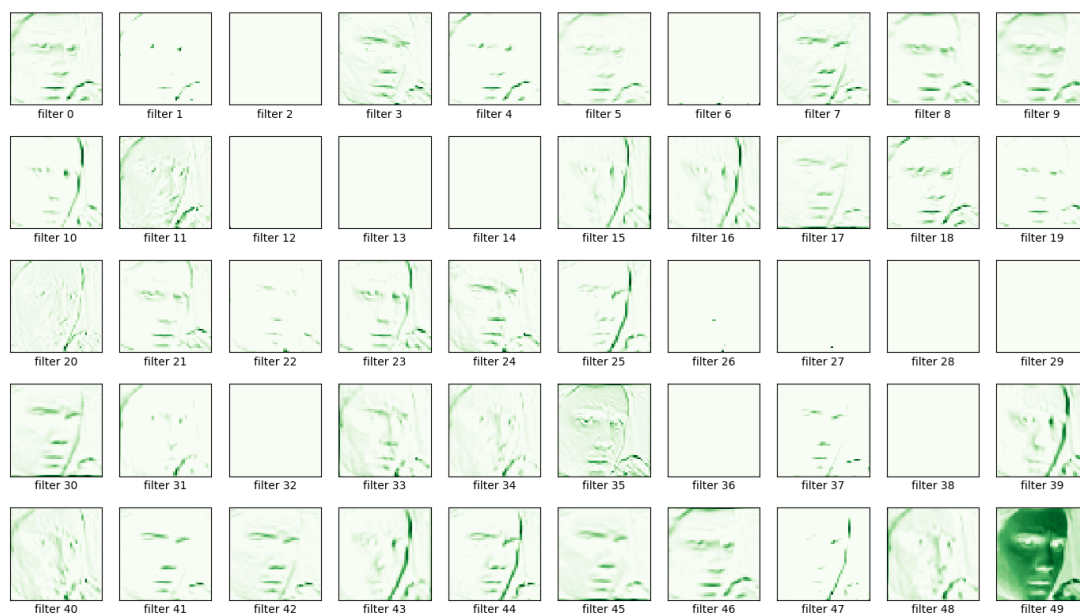


我們可以看到在 saliency maps 中越敏感的地方約在臉頰靠近鼻子兩端，而在外

為通常較敏感，在這裡我把 saliency maps 低於 0.5 的直接帶 0 mask 掉，可以看到第二章圖嘴型以及鼻子旁兩側臉頰很重要，其餘兩張圖也是臉頰表情的變動很敏感。

5. (1%) 承(4) 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 與觀察 filter 的 output。





此題我去拆解 best model convolution 的第一層，第一層看起來主要是在抓取大範圍的紋路以及線條，**filter** 中的線條似乎傾斜角度也有不同，可能是因為照片中為側看，我們可以想成第一層的 **layer** 被基本的條紋給 **activate**，下圖則為被相對應 **filter** 所產生在這一層的 **output**。