

學號：R06723025 系級：財金所碩一 姓名：林耘寬

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: 翁瑋)

整個模型概略為：

{ Zeropadding2D(1,1)

Conv2d : Filters 依序為 32~256，kernel size = (3,3)

...repeating

Batchnormalization

Max pooling: (2,2), strides = (2,2)

Dropout：從 0.25~0.5 不等 }

...repeating

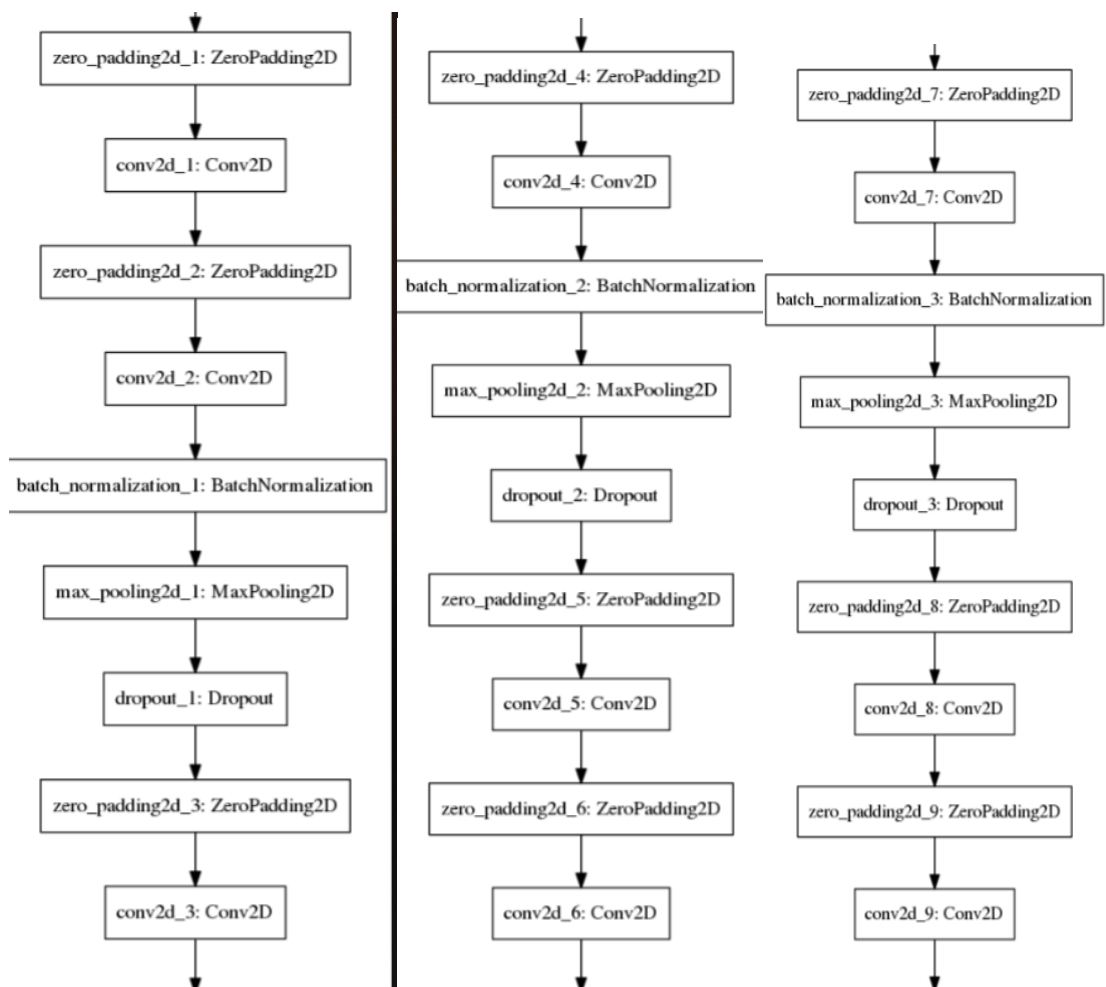
Dense: 2048, activation function 皆為 relu。

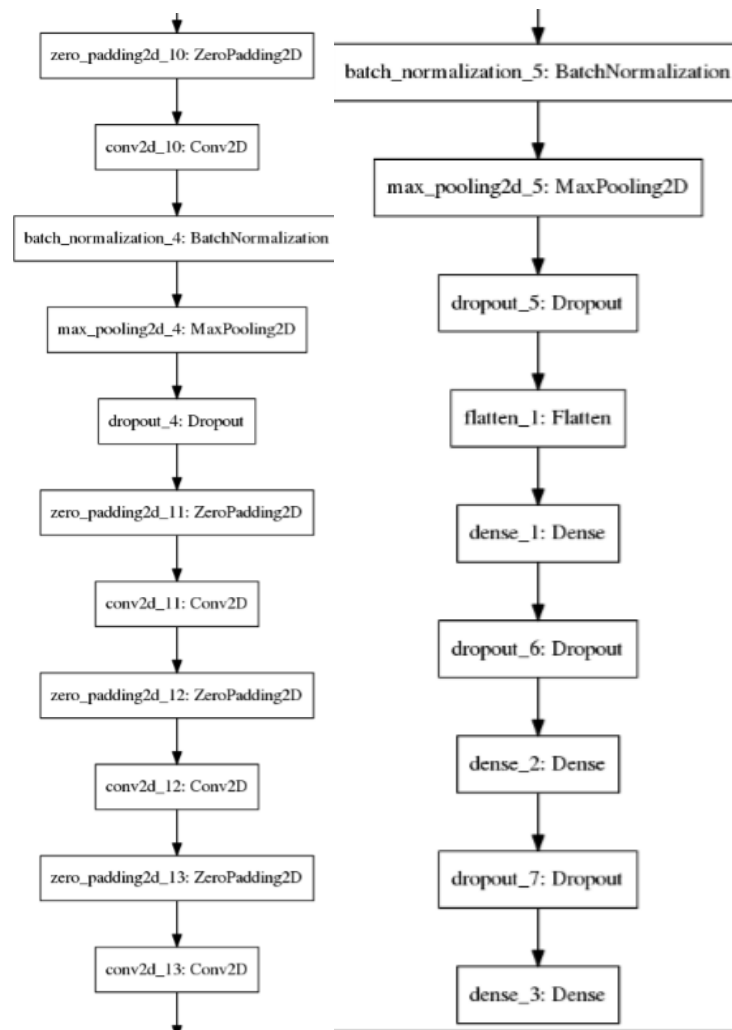
Dropout:0.5

...repeating

Loss function: Cross entropy, Optimizer: Adam

使用 Imagegenerator 來增加訓練資料，整個模型總共有 8,419,559 個參數





此模型的準確率最高 public and private score 皆超過 0.68，但我不幸遺失這個模型，目前使用的是 private score: 0.67 的模型，模型架構相同，僅是參數值以及 epoch 次數的差異。我認為該模型最後的準確率還能再提高，因為我把一部分 data 設成 validation set，根據經驗應該準確率可以再增加 0.1~0.2。

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？

DNN 模型的架構如下：

Dense: 2048, Activation function: ReLU

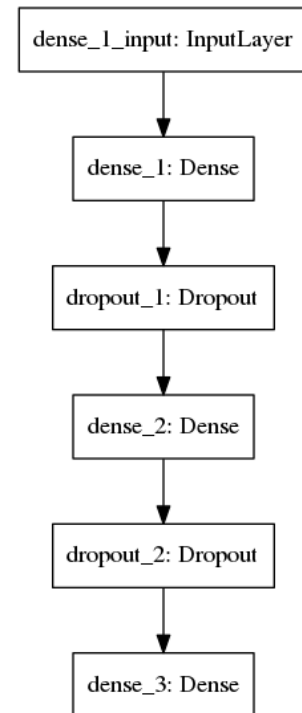
Dropout: 0.5

Dense: 2048, Activation function: ReLU

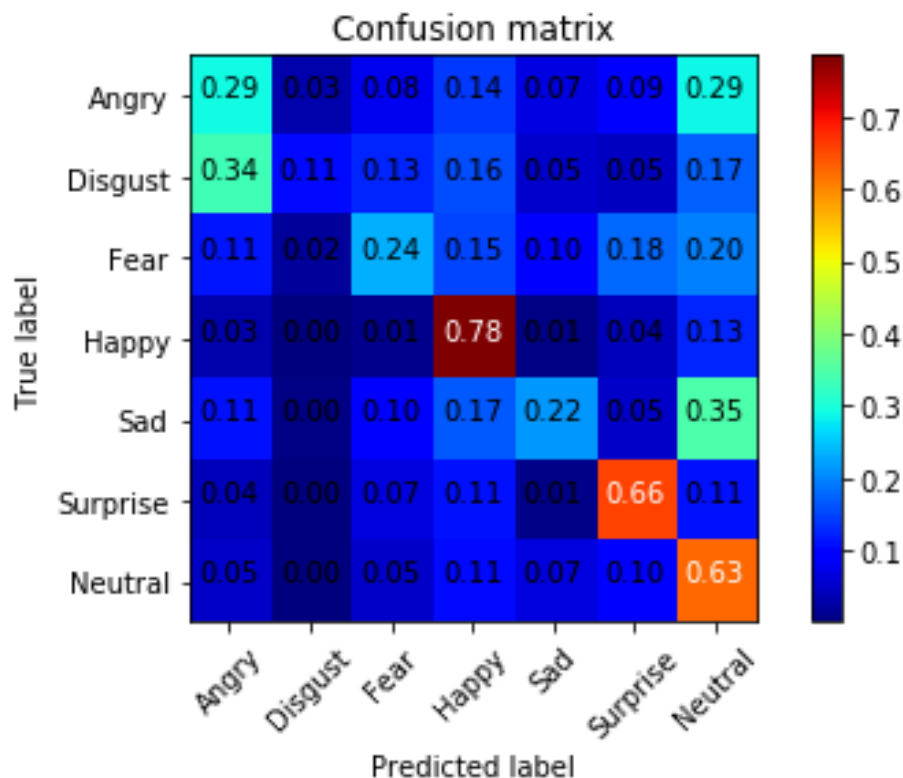
Dropout: 0.5

總共有 8,931,335 個參數，與上述 CNN 模型相去不遠。

但此模型在 kaggle private score 只有 0.40，而相較於 CNN，DNN 的訓練速度也快很多，即使在兩者參數差不多的情況下。



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]

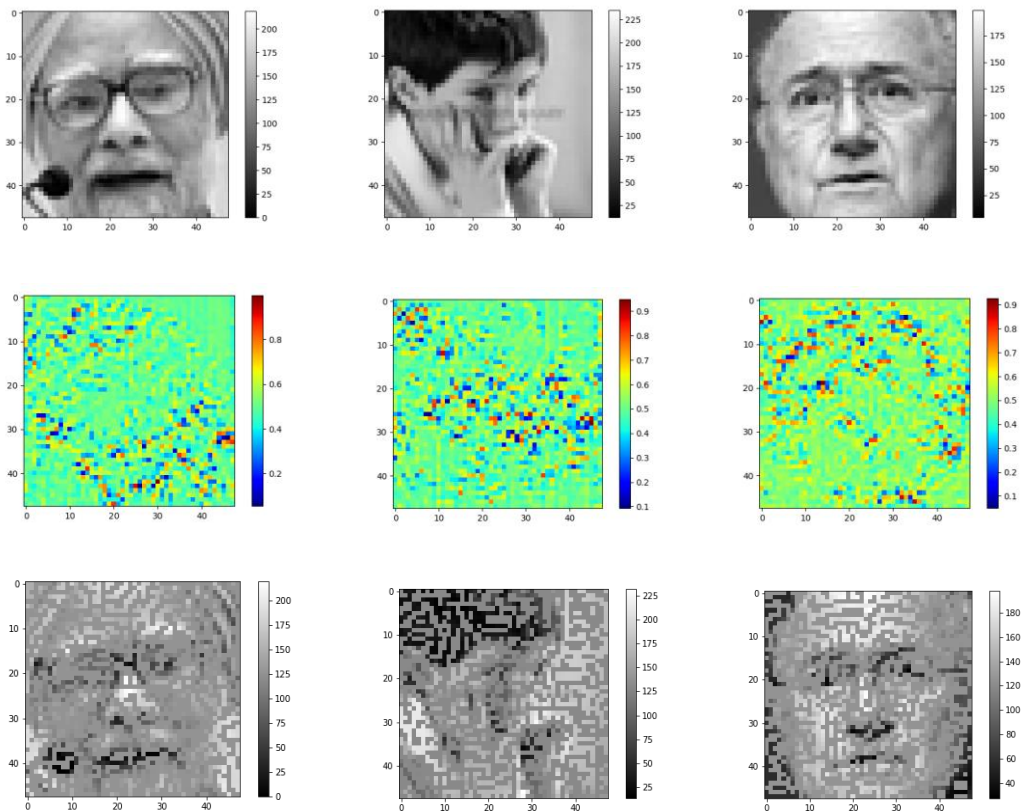


由上圖可以得知，Happy, Surprise, Neutral 的圖示較容易答對的。

容易混淆的有(True/Predicted)：Sad/Neutral, Disgust/Angry, Angry/Neutral, Fear/Neutral.

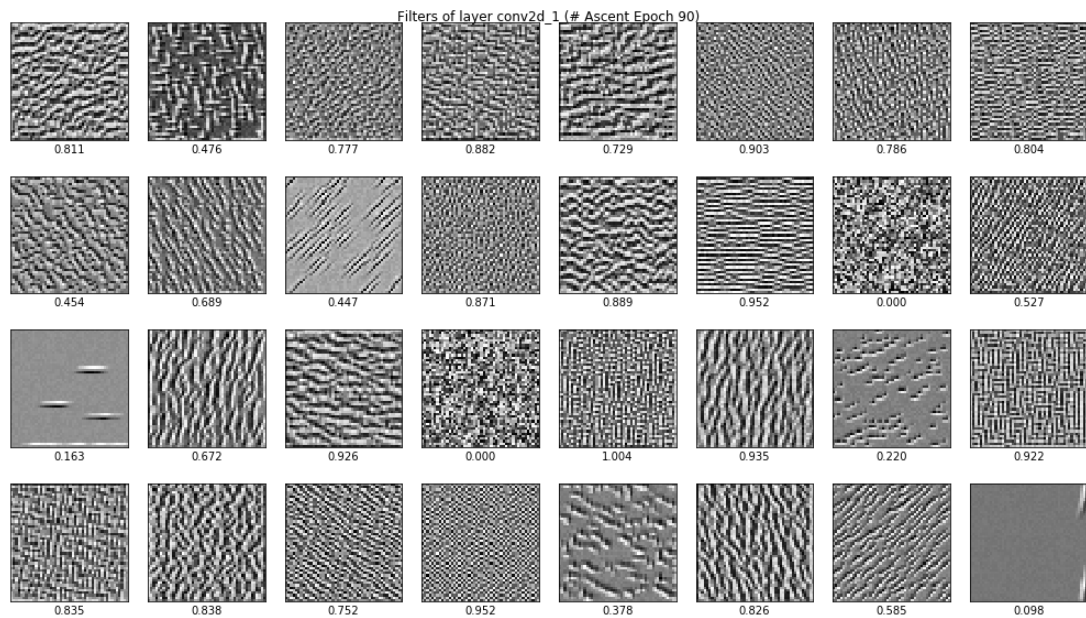
由上述推測，可能在表情較不明顯的情況下，機器容易把圖片歸在 Neutral 的類別上。

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？

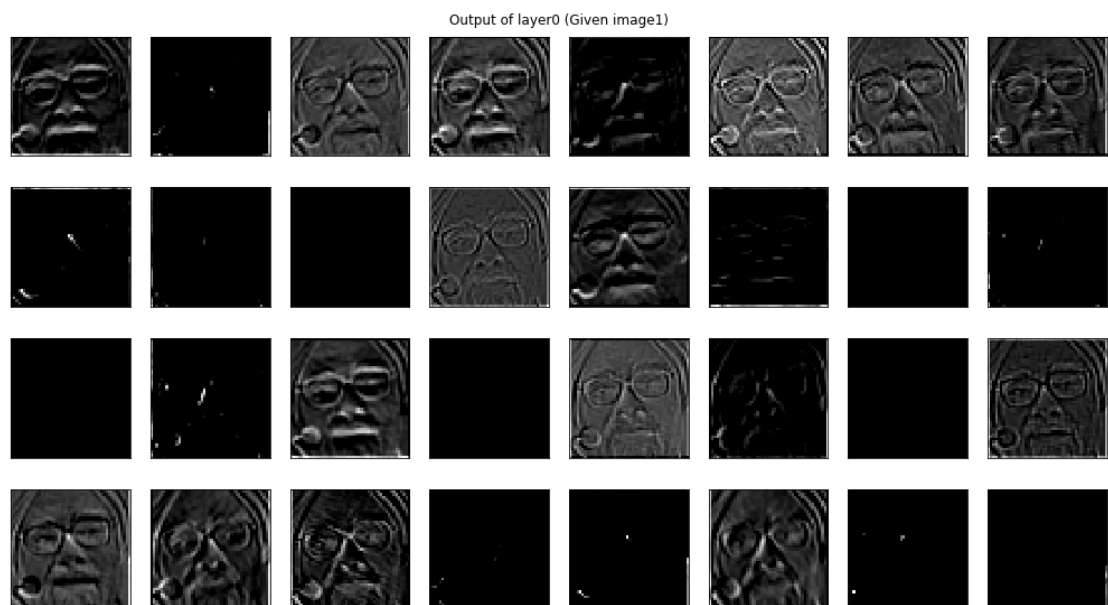


上圖為我隨意選出來的三張圖片，可以發現臉的輪廓在圖內有較高的值，而從 mask 掉 heat 小的部分的圖來看，嘴型可能也是評斷的一個標準。

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**。



我觀察的特定層為 **conv2d_1**，可以發現圖片中分辨的 **filter** 主要是分成直條紋以及橫條紋，其中條紋也有粗細疏密之分，因為是第一層 **conv2d**，故推測可能是分析大方向為主。



我使用的圖片是第四題最左邊的那張圖片，可以發現主要分析的應該是人的輪廓，故上述關於此層的推測可能是正確的。