

學號：R06921005 系級：電機碩一 姓名：陳昱文

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:)

答：

Accuracy	Public	Private	Total
CNN	0.68096	0.67456	0.67776

架構：Conv2D(32)→Conv2D(32)→MaxPooling2D((2,2))→ Conv2D(64)→Conv2D(64)→MaxPooling2D((2,2))  
→ Conv2D(128)→Conv2D(128)→MaxPooling2D((2,2))→Flatten()→Dense(256)→Dense(64)→Dense(7)

訓練過程：使用 Keras 內建 datagenerator 產生 augmented data 以幫助 model 學習，亦加入 Dropout(), BatchNormalization(), Kernel\_initializer 以增進效能，最後亦手動對 data 進行對比度放大調整以期望 model 能夠更容易 focus 在重要資訊上

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:)

答：

Accuracy	Public	Private	Total
CNN	0.68096	0.67456	0.67776
DNN	0.37837	0.37280	0.37559

此 DNN 模型採用的參數量與 CNN 接近，但由上列圖表可知，純 DNN 的效能相較於 CNN 遜色許多，可見 CNN 在影像處理佔有極重要的地位，因為他保有一些與影像相關的特色，比如說偵測特定 pattern 等等。而在此次實驗中(1,2 小題)，層數的差異或許也會是影響 performance 的因子之一。

架構：Dense(input\_dim=48\*48,128)→Dense(384) →Dense(384)→Dense(384)→Dense(384)→Dense(7)

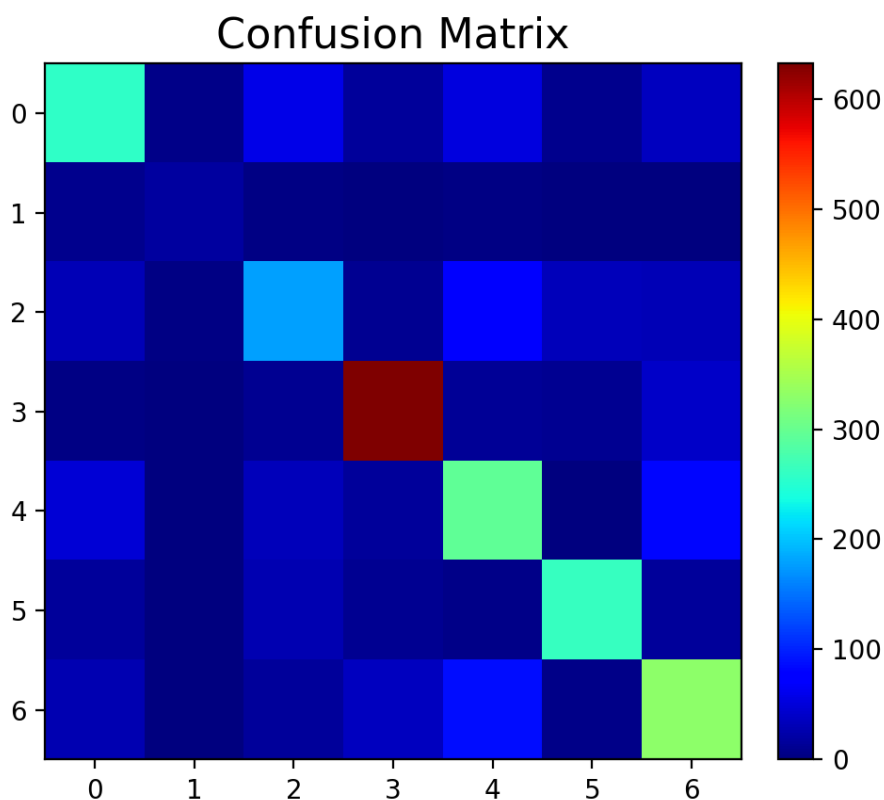
訓練過程：類似 CNN 訓練步驟

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: )

答：

如下圖可明顯看出在 Class 1(厭惡的表情) 其分類最差，推測其沒有明顯判讀的標準

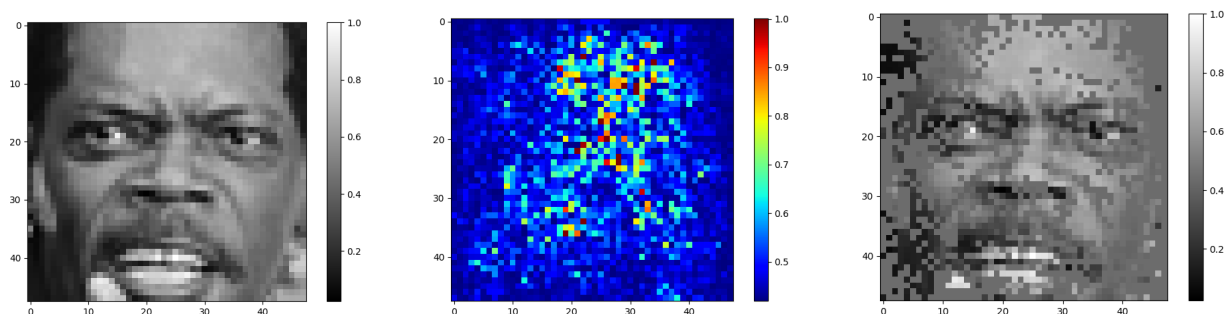


4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？

(Collaborators: )

答：

如下三圖所示，第二張為 heatmap 而第三張則將 heat 小的部分給 mask 掉，可以觀察出其有部分能力能夠 focus 在五官或者是高低起伏(對比度高)等地方



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: )

答：

如下圖所示，此為第二層 Convolution Layer filters 利用 gradient ascent 的方法實作，可以觀察出他好像在辨別各種紋路，這也是 convolution 之所以有好效能的原因之一，而在下圖中可見部分圖形類似接近空白，猜想也許是資料量太少等等原因，可能不需要用到那麼多 filter 數

