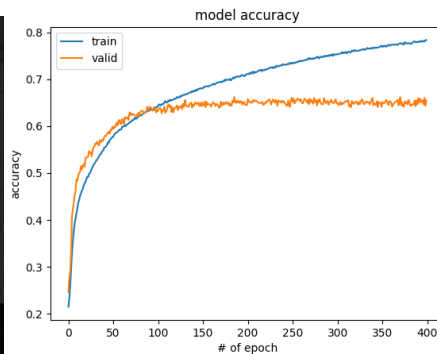


學號：R06944031 系級：網媒碩一 姓名：林蔚廷

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？

Model summary

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 48, 48, 32)	320
conv2d_2 (Conv2D)	(None, 48, 48, 32)	9248
conv2d_3 (Conv2D)	(None, 48, 48, 48)	13872
max_pooling2d_1 (MaxPooling2D)	(None, 16, 16, 48)	0
conv2d_4 (Conv2D)	(None, 16, 16, 48)	20784
conv2d_5 (Conv2D)	(None, 16, 16, 48)	20784
conv2d_6 (Conv2D)	(None, 16, 16, 64)	27712
max_pooling2d_2 (MaxPooling2D)	(None, 8, 8, 64)	0
conv2d_7 (Conv2D)	(None, 8, 8, 64)	36928
conv2d_8 (Conv2D)	(None, 8, 8, 64)	36928
conv2d_9 (Conv2D)	(None, 8, 8, 96)	55392
max_pooling2d_3 (MaxPooling2D)	(None, 4, 4, 96)	0
flatten_1 (Flatten)	(None, 1536)	0
dense_1 (Dense)	(None, 64)	98368
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 48)	3120
dropout_2 (Dropout)	(None, 48)	0
dense_3 (Dense)	(None, 7)	343
Total params: 323,799		
Trainable params: 323,799		
Non-trainable params: 0		



準確度: (Kaggle)

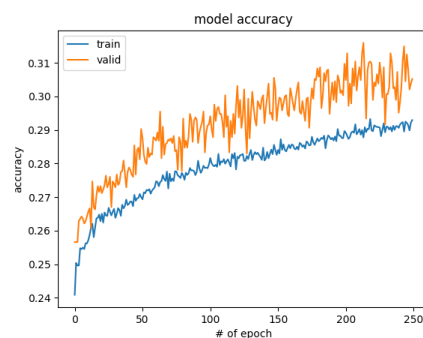
Private Score: 0.67985

Public Score: 0.66425

答: 九層的 CNN，每隔三層作一次 max pooling，再接上三層 fully-connected layer，中間 dropout 兩次，第一次丟 0.3 第二次丟 0.2。我的 model 大概到 100 個 epoch 時候 validation 就不太增加了，大約停在 64%~65% 的準確度，而在 Kaggle 上得到比較高的結果(66%、67%)

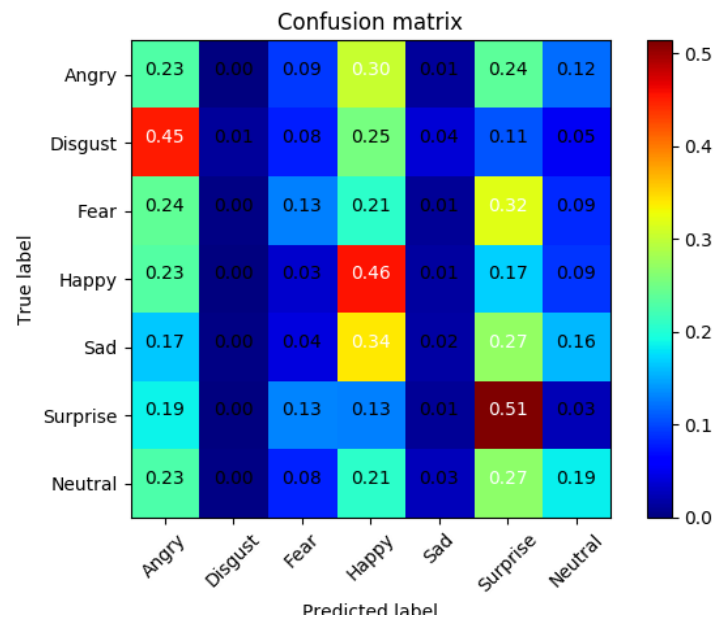
2. (1%) 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model。其模型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

flatten_1 (Flatten)	(None, 2304)	0
dense_1 (Dense)	(None, 48)	110640
dense_2 (Dense)	(None, 48)	2352
dense_3 (Dense)	(None, 48)	2352
dense_4 (Dense)	(None, 96)	4704
dense_5 (Dense)	(None, 128)	12416
dense_6 (Dense)	(None, 128)	16512
dense_7 (Dense)	(None, 128)	16512
dense_8 (Dense)	(None, 128)	16512
dense_9 (Dense)	(None, 192)	24768
dense_10 (Dense)	(None, 256)	49408
dense_11 (Dense)	(None, 256)	65792
dense_12 (Dense)	(None, 7)	1799
Total params: 323,767		
Trainable params: 323,767		
Non-trainable params: 0		



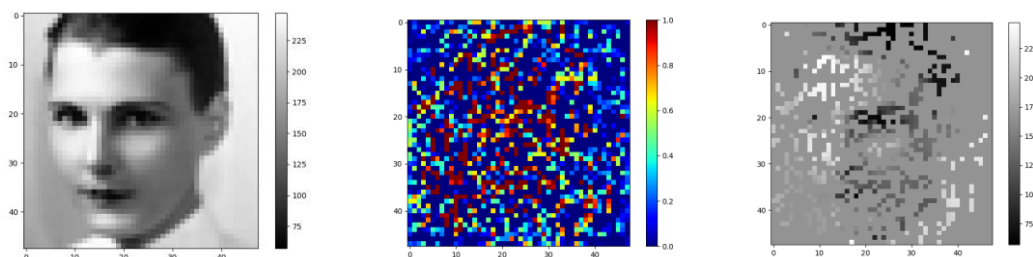
答: 我用了 12 個 Layer，323,767 個參數，與我的 CNN model 接近，可以看到 train 了 400epoch 花了一個半小時，仍然沒有提升準確度，在 29%~31% 之間震盪。結論: DNN 模型來預測，效果比起 CNN 差很多。

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: 我發現在我的 model 中，Disgust 容易被判斷成 Angry，還有預測 Happy 實際卻是 Angry 或 Sad 的機率蠻高的，然後我的 model 幾乎不猜 Disgust 和 Sad

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: 有抓到頭部的部分，看起來 focus 在眼睛和頭髮的部分

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：這題我做不出來 抱歉了