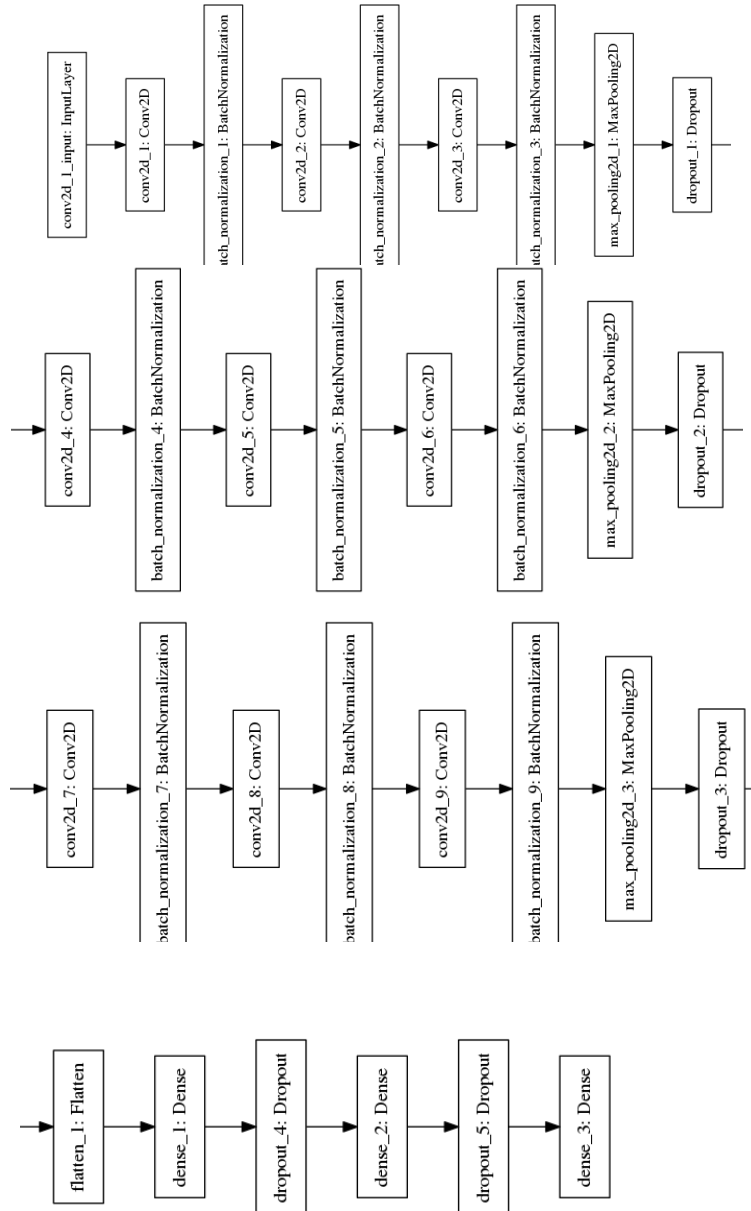


1. (1%) 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？  
(Collaborators: 柯志霖)

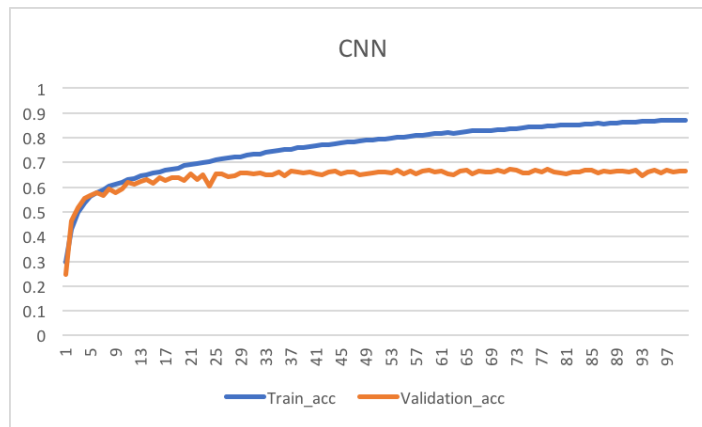
答： 模型架構：



model 實做 9 convolution layers 而且每個 convolution layer 會加上 batch normalization  
並且每三層 convolution layers 做一次 maxpooling 以及 dropout  
經過九個 convolution layers 之後 後面接兩層 dense layers 以及 dropout

利用 DataGenerator 調整 training data 和 validation data 提昇準確度。

訓練過程：



最後 Epoch = 100: loss: 0.3531 - acc: 0.8724 - val\_loss: 1.2416 - val\_acc: **0.6634**

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: )

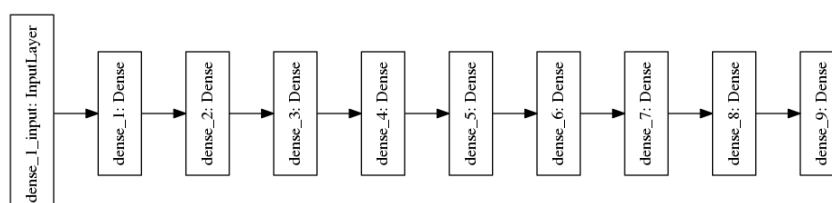
答： CNN model 參數量：

```
Total params: 1,696,583
Trainable params: 1,695,239
Non-trainable params: 1,344
```

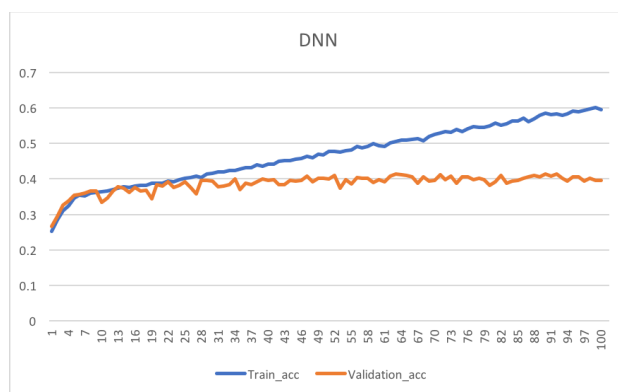
DNN model 參數量：

```
Total params: 1,698,183
Trainable params: 1,698,183
Non-trainable params: 0
```

模型架構：



訓練過程：

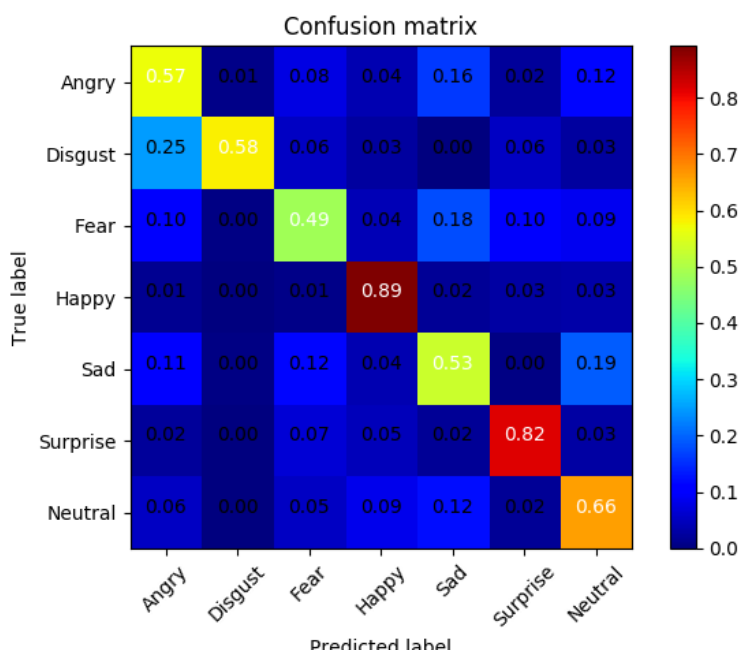


最後 Epoch = 100: loss: 1.0565 - acc: 0.5956 - val\_loss: 1.8777 - val\_acc: 0.3967

在相同的參數量下，DNN 相較 CNN，DNN 在做表情辨識上的效果較差。

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：由 confusion matrix 可以看出，將 Disgust 辨識成 Angry 或是 Fear 辨識成 Sad 以及 Sad 辨識成 Neutral 的混淆機率較高

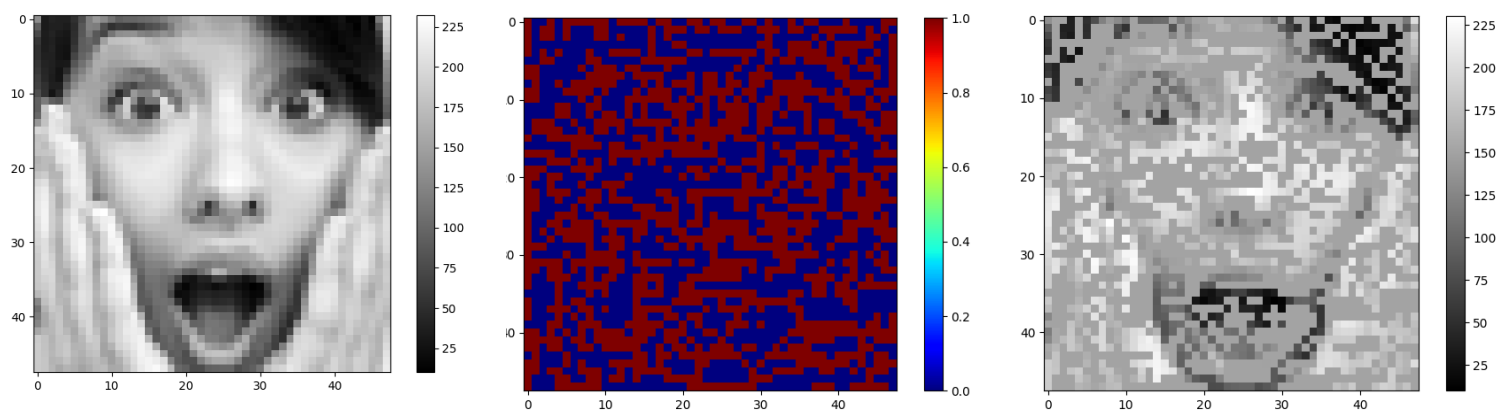


4. (1%) 從  
(1)(2)可  
以發現，

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：我認為它會 focus 在輪廓比較深的部份。例如五官或是陰影處



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:

Filters of layer conv2d\_6 ( # Ascent Epoch 190 )

