

Homework 3 Report - Image Sentiment Classification

學號：r06921081 系級：電機碩一 姓名：張邵瑀

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？

| vgg-16改 | vgg-16 mini | 改進小老師模型 |
|---|---|---|
| Conv2D 64x(3x3), Acti: relu, Drop:0.10 Conv2D 64x(3x3), Acti: relu, Drop:0.10 MaxPooling2D 2x2, Drop:0.10 Conv2D 128x(3x3), BatchNormalization(), Acti: relu, Drop:0.25 Conv2D 128x(3x3), Acti: relu, Drop:0.25 MaxPooling2D 2x2, Drop:0.25 Conv2D 256x(3x3), Acti: relu, Drop:0.35 Conv2D 256x(3x3), Acti: relu, Drop:0.35 Conv2D 256x(3x3), BatchNormalization(), Acti: relu, Drop:0.35 MaxPooling2D 2x2, Drop:0.35 Conv2D 512x(3x3), Acti: relu, Drop:0.4 Conv2D 512x(3x3), Acti: relu, Drop:0.4 Conv2D 512x(3x3), Acti: relu, Drop:0.4 MaxPooling2D 2x2, Drop:0.4 Conv2D 512 x(3x3), Acti: relu, Drop:0.45 Conv2D 512 x(3x3), Acti: relu, Drop:0.45 Conv2D 512 x(3x3), Acti: softplus, Drop:0.45 MaxPooling2D 2x2, Drop:0.45 Flatten() Dense 4096, Acti: relu, Drop:0.5 Dense 4096, Acti: softplus, Drop:0.5 | Conv2D 64x(3x3), Acti: relu, Conv2D 64x(3x3), Acti: relu, MaxPooling2D 2x2, Drop:0.10 BatchNormalization() Conv2D 128x(3x3), Acti: relu, MaxPooling2D 2x2, Drop:0.25 Conv2D 256x(3x3), BatchNormalization(), Acti: relu, MaxPooling2D 2x2, Drop:0.35 Conv2D 512x(3x3), Acti: relu, MaxPooling2D 2x2, Drop:0.4 Conv2D 512 x(3x3), Acti: relu, Conv2D 512 x(3x3), Acti: softplus, MaxPooling2D 2x2, Drop:0.45 Flatten() Dense 4096, Acti: relu, Dense 4096, Acti: softplus, Drop:0.5 Dense 7, Acti: softmax | Conv2D 64x(5x5) LeakyReLU:0.05 BatchNormalization MaxPooling2D (2x2) Dropout:0.10 Conv2D 128x(3, 3) LeakyReLU:0.05 BatchNormalization MaxPooling2D (2x2) Dropout:0.25 Conv2D 512x(3x3) LeakyReLU:0.05 BatchNormalization MaxPooling2D (2x2) Dropout:0.30 Conv2D 512x(3x3) LeakyReLU:0.05 BatchNormalization MaxPooling2D (2x2) Dropout:0.35 Flatten() Dense 512, Acti:relu, BatchNormalization Dropout(0.5)) Dense 512, Acti:softplus, BatchNormalization() Dropout(0.5)) Dense 7, Acti:softmax, |

| | | |
|------------------------------|------------------------------|------------------------------|
| Dense 7, Acti: softmax | | |
| optimizer: adam lr = 1e-4 | optimizer: adam lr = 1e-4 | optimizer: adam lr = 1e-4 |
| acc:0.68904 | acc:0.68292 | acc:0.71022 |

結論：越深不代表精準度會飆升

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:

先減掉mean(data)，然後做 normalization，有兩種方式：

- 1.一種直接在讀完檔之後將該變數直接除以255
- 2.用keras ImageDataGenerator 中設定rescale=1./255 (我採用這種)

| | 實作前 | 實作後 |
|-----|---------|---------|
| acc | 0.65645 | 0.68904 |

data augmentation:

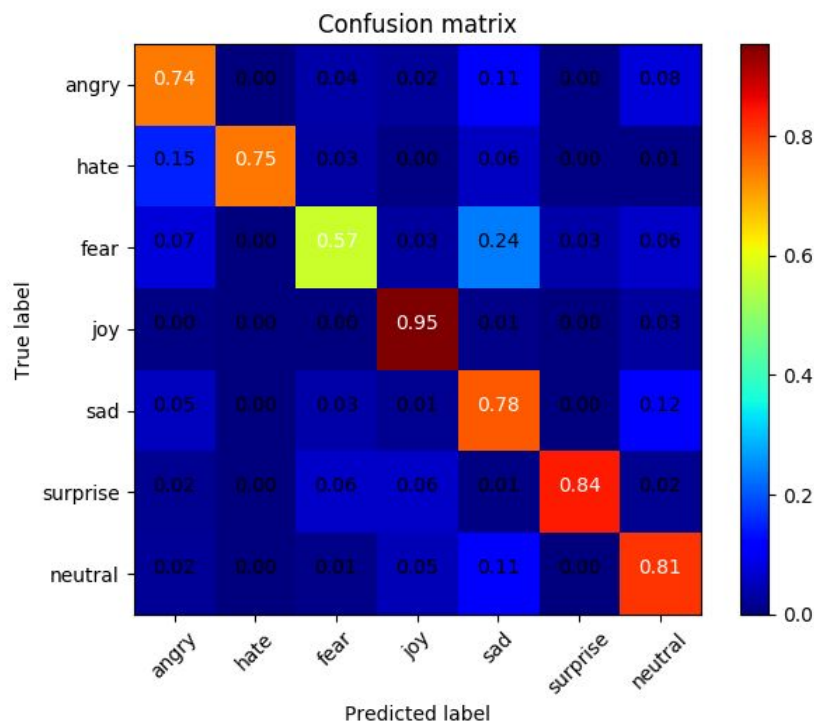
有試著手寫產生DATA的code但發現不會比keras的好，所以果斷放棄

ImageDataGenerator(rotation_range=20, width_shift_range=0.2,
height_shift_range=0.2, rescale=1./255, shear_range=0.2,
zoom_range=0.2, horizontal_flip=True, fill_mode='nearest')

| | 實作前 | 實作後 |
|-----|---------|---------|
| acc | 0.64864 | 0.68904 |

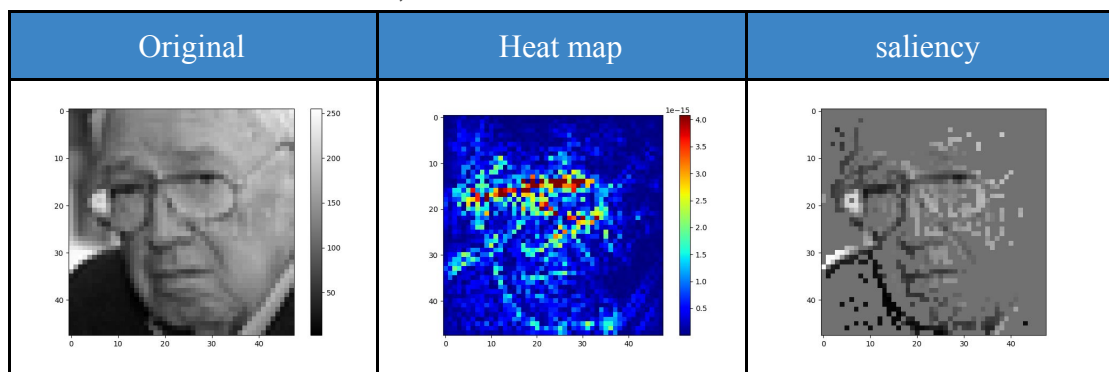
這兩種作法對準確率都有顯著的幫助，在訓練時若是沒有使用data augmentation則會在training正確率拿到8,90%的正確率validation正確率卻很低6X%的情況，normalization則可以有效的提高正確率，我猜是因為relu的對於有正規化的資料有比較好的資訊保留。

3. (1%) 觀察答錯的圖片中，哪些 class 彼此間容易用混？



可以看到在我的model裡fear會搞混成sad的機率最大hate到angry次之，但其實都不大

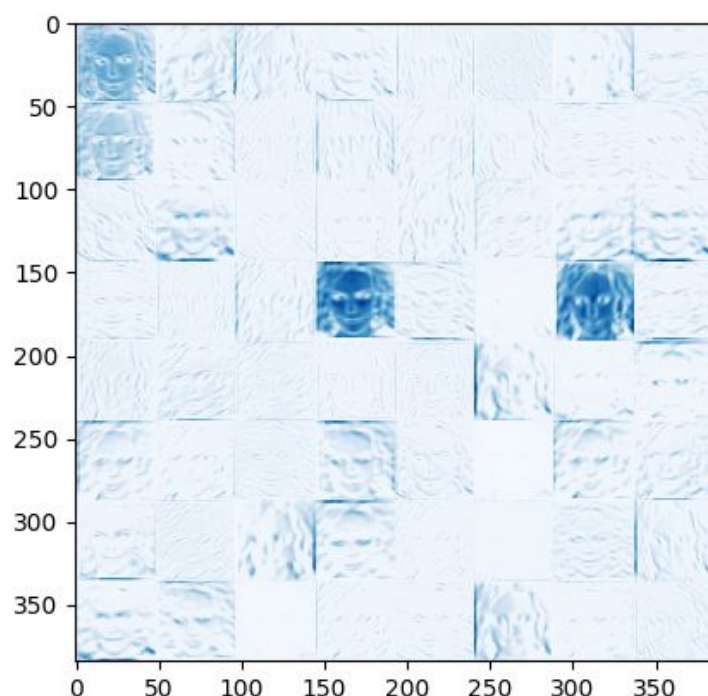
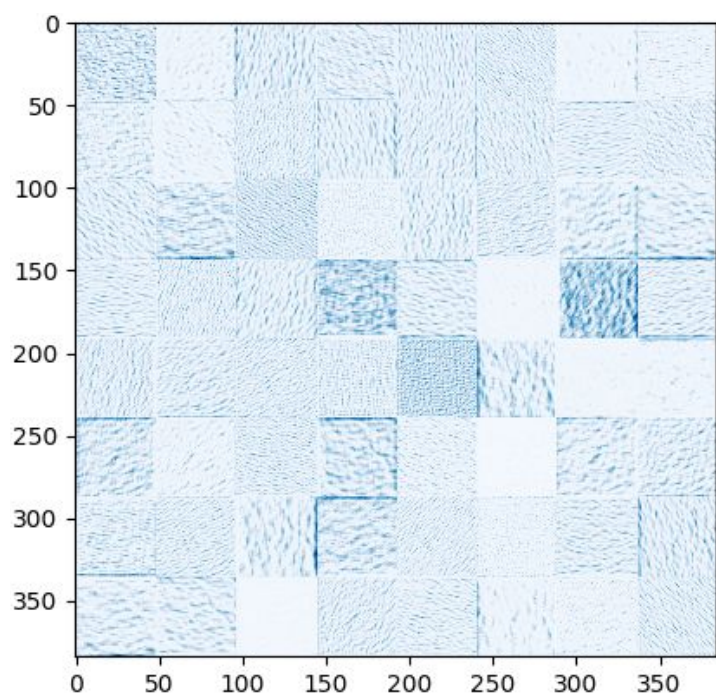
4. (1%) 從(1)(2)可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？



很明顯得出的結果還是著重在辨識眼睛與嘴巴還有些許眉毛的部位，還有一些臉頰上部，還有因為是sad所以法令紋也是個很重要的feature.

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。

MaxPooling2D



這是觀察model的第1層對 test 10 作圖可發現對眉毛眼睛和嘴型特別的清楚