## 機器學習作業四 report

學號:B05901040 系級:電機三 姓名:蔡松達

1. (2%)從作業三可以發現,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: Ref: https://reurl.cc/0Kvx9)

答:圖片如下表所示,基本上在進行分類時,生氣(0)聚集在額頭,可能是生氣時額頭皺褶特別深;厭惡(1)集中於臉頰輪廓的皺褶,表達很討厭某事物的感覺;恐懼(2)主要集中在眉梢之間,眉頭緊繃表達出緊張害怕的感受;高興(3)與前三者的區域明顯不同,集中於鼻子至嘴巴處的微笑;難過(4)涵蓋眼睛、鼻子、嘴巴,較不容易表達出來;驚訝(5)也較不顯著,只能大概推斷臉頰兩側及眼睛下方可能是在驚嚇時臉部肌肉收縮;中立(6)明顯集中在額頭處,說明在沒有表情的狀況下額頭是平整而沒有任何皺褶的。由此可見大多數的特徵集中在額頭、眼睛、臉頰、嘴巴之間,分類的方式以明顯的臉部皺褶處居多。

| 圖片編號     | fig1_0.jpg  | fig1_1.jpg  | fig1_2.jpg  | fig1_3.jpg  |
|----------|---|---|---|---|
| training | 0   | 299   | 5   | 7   |
| data     |   |   |   |   |
| 圖片       | Saliency map for feature 0, picture number 0  - 0.003 - 0.002 - 0.001 - 0.000 - 0.001 - 0.000 - 0.001 - 0.002 | Saliency map for feature 1, picture number 299  -0.00100 -0.00075 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 -0.000050 | Saliency map for feature 2, picture number 5 -0.002 -0.001 -0.001 -0.001 -0.002 -0.001 -0.002 -0.004                                    | Saliency map for feature 3, picture number 7 -0.00006 -0.00004 -0.00002 -0.000002 -0.000004 -0.000004 -0.000004 -0.000006 -0.000006 |
| 圖片編號     | fig1_4.jpg  | fig1_5.jpg  | fig1_6.jpg  |   |
| training | 6   | 15  | 4   |   |
| data     |   |   |   |   |
| 圖片       | Saliency map for feature 4, picture number 6 0 -0002 -0001 -0001 -0001 -0002 -0002 -0002 -0002 -0003          | Saliency map for feature 5, picture number 15  10 - 0002  - 0001  - 0001  - 0 001  - 0 002  - 0 002  - 0 002  - 0 003   | Saliency map for feature 6, picture number 4  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008  0 008 |   |

2. (3%)承(1),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 與觀察 filter 的 output。(Collaborators: Ref: https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html)

答:此處觀察  $conv2d_7$  這層,含有 128 個 filter,此處選擇 16 個 filter,分別是編號 4, 6, 18, 20, 21, 22, 24, 25, 41, 42, 43, 44, 47, 53, 57, 59,  $fig2_1$  (下圖)呈現出這些 filters,其實這些 filter 呈現出的輪廓雖然各有不同,但大多數的圖片是具有規則性的,而且其中有幾張圖感覺只是翻轉過後的結果而已,代表這些 filter

在進行辨識的時候喜歡去找紋路較深,特徵較為顯著的地方;此外我也嘗試過對 於較為前面的 conv 層進行 gradient ascent,得到的結果其實較不明顯,紋路比較 淡,表示前面層比較沒辦法描繪出比較深刻的紋路,越深層紋路也越深。

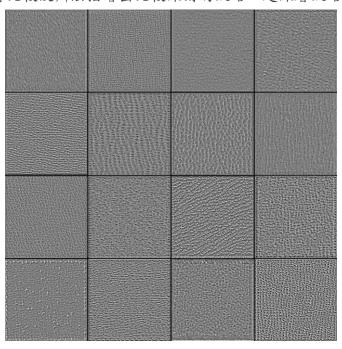


fig2\_1.jpg

而 fig2\_2 (下圖)呈現出 conv2d\_7 這層的 output (取前 64 個 filter),由此層可見圖片大多模糊,卻各自強調不同的重點,有些強調眼睛,有些強調嘴巴,也有些強調背景,而較前面 layer 的 output 則仍可較為清晰的看見人臉的樣子,兩相比較下越深層的 output 較能夠 focus 在圖片中的某些地方。

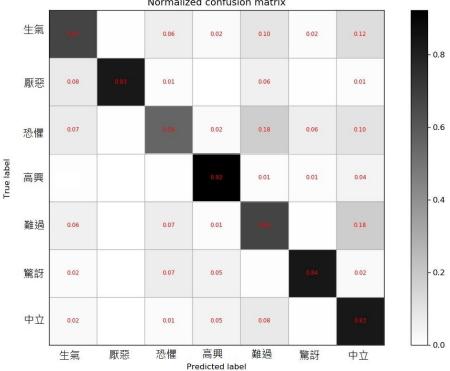


fig2\_2.jpg

3. (3%)請使用 Lime 套件分析你的模型對於各種表情的判斷方式,並解釋為何你的模型在某些 label 表現得特別好(可以搭配作業三的 Confusion Matrix)。 答:從 confusion matrix 可以發現最佳的 label 是高興,從圖片中可以發現重點多在圖片下半部分,也就是下巴處可能微笑的特徵較明確,是其他表情甚少會有的;而其次較好的是中立,全都 focus 在圖片上緣,代表中立的圖片表情通常不

明顯,所以可能將重點擺在不是臉部的部分,而其餘的 attribute 重點較為分散,不過在眼睛、嘴巴的部分都有重點的存在,代表可能從眼睛的表達上來判定成功的機會較大;但較為分散也說明可能重點並不明確,比較不容易分辨正確。

| 17 M                        |   |  |  |   |  |  |
|-----------------------------|---|--|--|---|--|--|
| 圖片編號                        | fig3_0.jpg  | fig3_1.jpg   | fig3_2.jpg   | fig3_3.jpg  |  |  |
| training                    | 0   | 473  | 5  | 7   |  |  |
| data                        |   |  |  |   |  |  |
| 圖片                          | 200<br>213<br>225<br>207<br>207<br>207<br>207<br>207<br>207<br>207<br>207<br>207<br>207 | 20 100 100 100 100 100 100 100 100 100 1   | 10 10 10 10 10 10 10 10 10 10 10 10 10 1                           | 200<br>23<br>20<br>30<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40 |  |  |
| 圖片編號                        | fig3_4.jpg  | fig3_5.jpg   | fig3_6.jpg   |   |  |  |
| training                    | 6   | 15   | 11   |   |  |  |
| data                        |   |  |  |   |  |  |
| 圖片                          |   | 200<br>215<br>20<br>30<br>30<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40<br>40 | 200<br>200<br>200<br>200<br>200<br>200<br>200<br>200<br>200<br>200 |   |  |  |
| Normalized confusion matrix |   |  |  |   |  |  |



4. (2%) [自由發揮] 請同學自行搜尋或參考上課曾提及的內容,實作任一種方式來觀察 CNN 模型的訓練,並說明你的實作方法及呈現 visualization 的結果。答:這裡使用與第二題類似的觀念進行觀察,得到下列兩張不同的 output,第一張圖 (fig4\_1.jpg) 是使用訓練好的 model;第二張圖 (fig4\_2.jpg) 是使用只 train一個 epoch 的爛 model,而這裡是取出最後 model 中最後一次進行 Maxpooling前的 Leaky\_relu 層的 output,128 個 output 中選出前 64 個的 output,可以發現第一張圖的 output 幾乎都還能看出臉型,而且圖中黑色都只有出現在部分區域,代表每個 filter 都有各自不同的作用;然而可以從第二張圖發覺黑的部分就佔了一大半,幾乎沒辦法看出每個 filter 各自重點在哪,因此由較為深層(例如此層)的 output 來判斷分類的結果好壞是觀察 CNN 模型訓練的好方法。



fig4\_2.jpg