

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: 無)

答：

fig1_0.jpg (Angry)

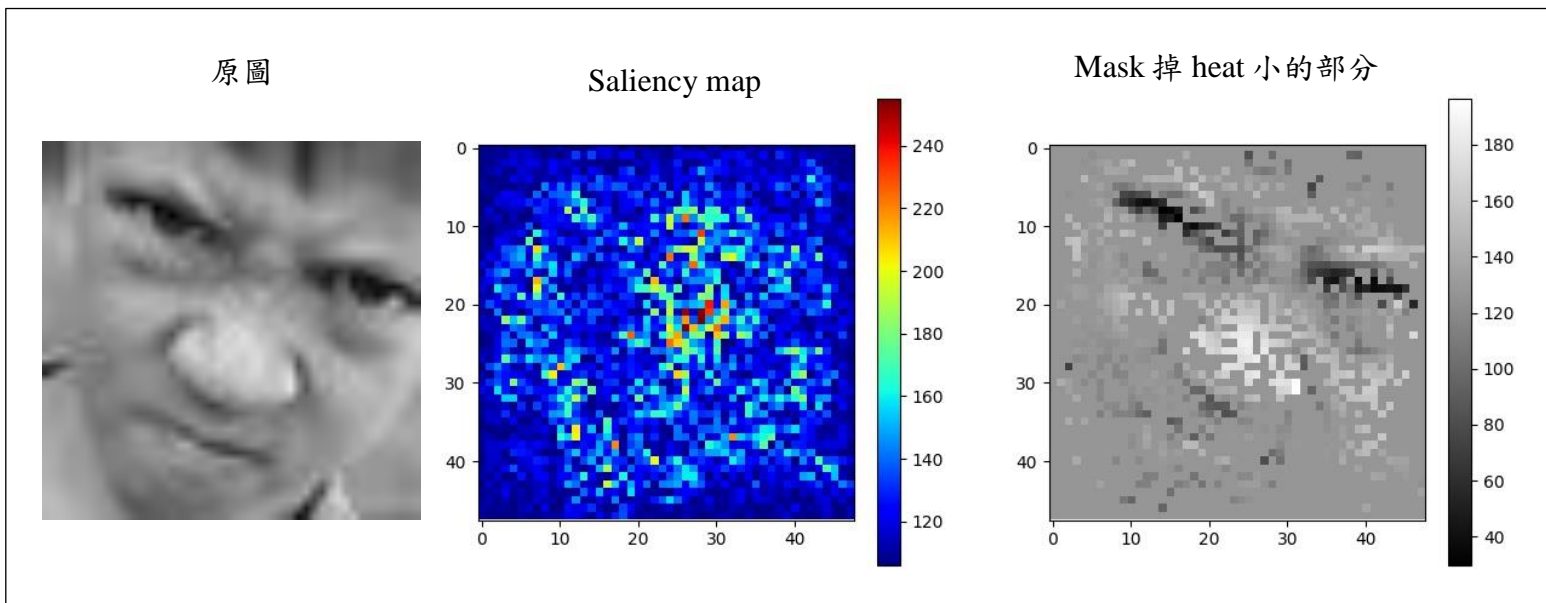


fig1_1.jpg (Disgust)

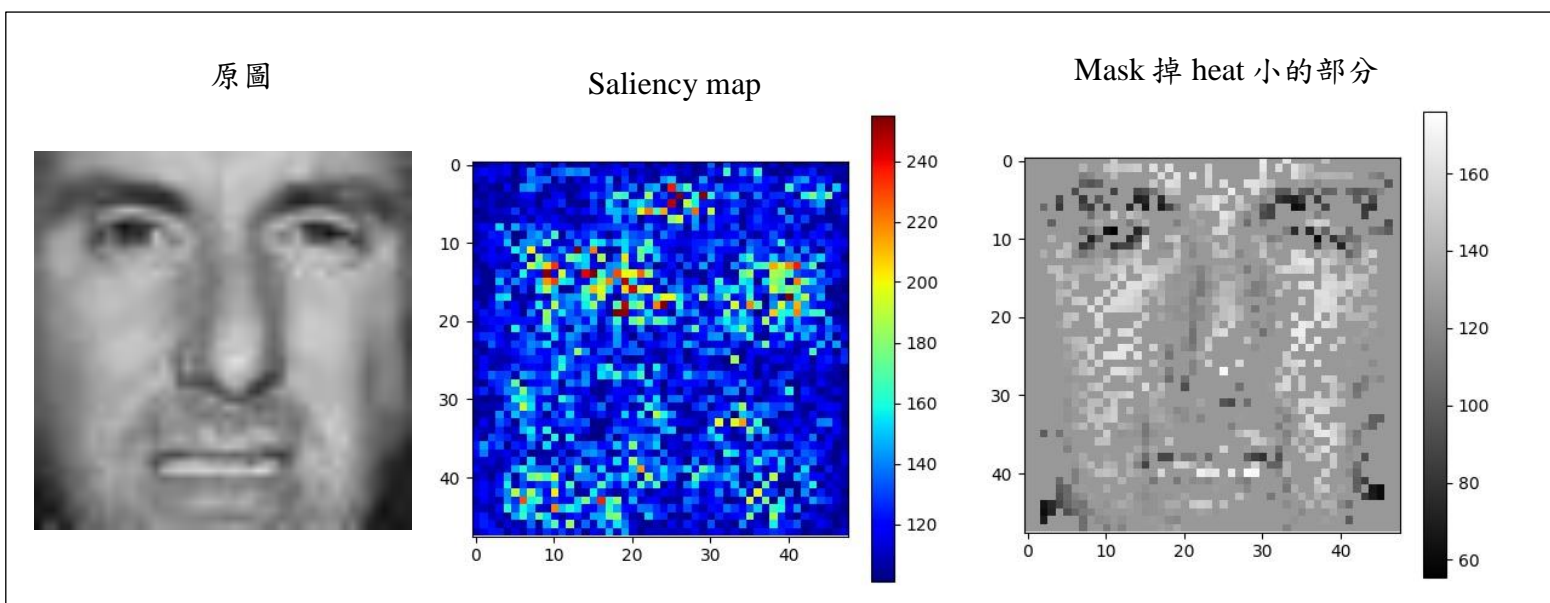


fig1_2.jpg (Fear)

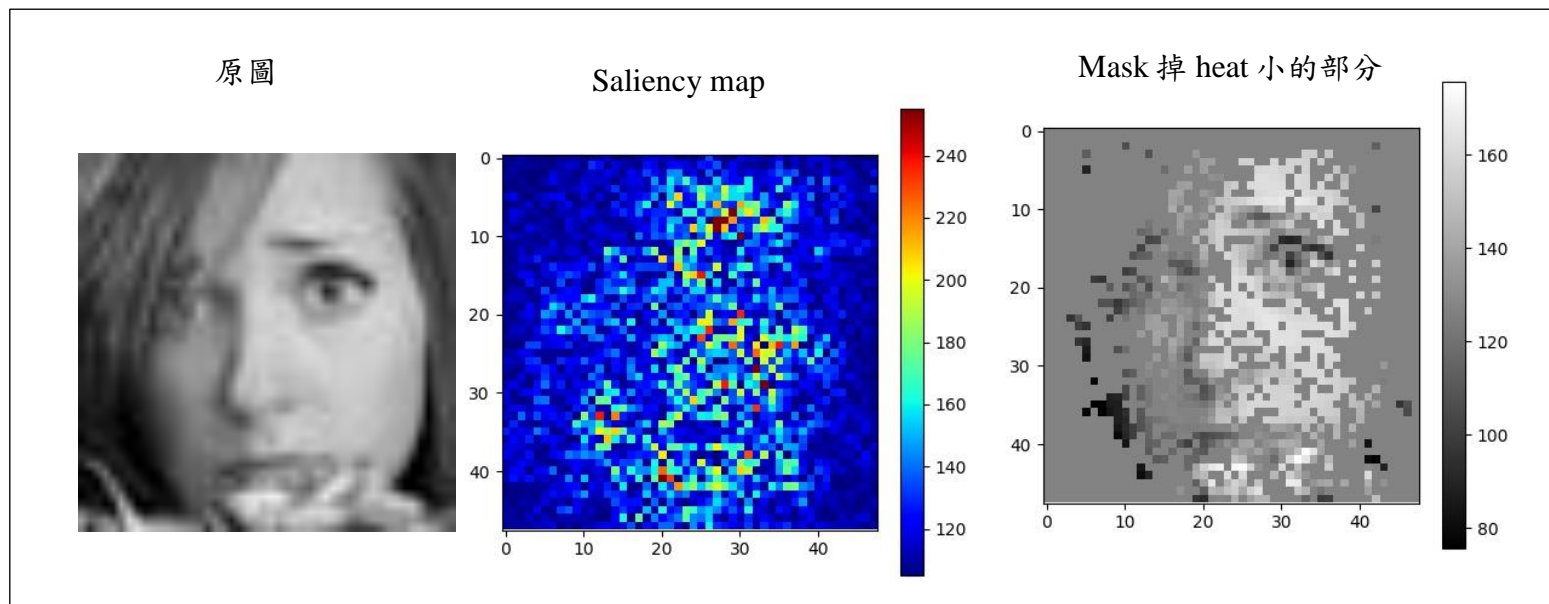


fig1_3.jpg (Happy)

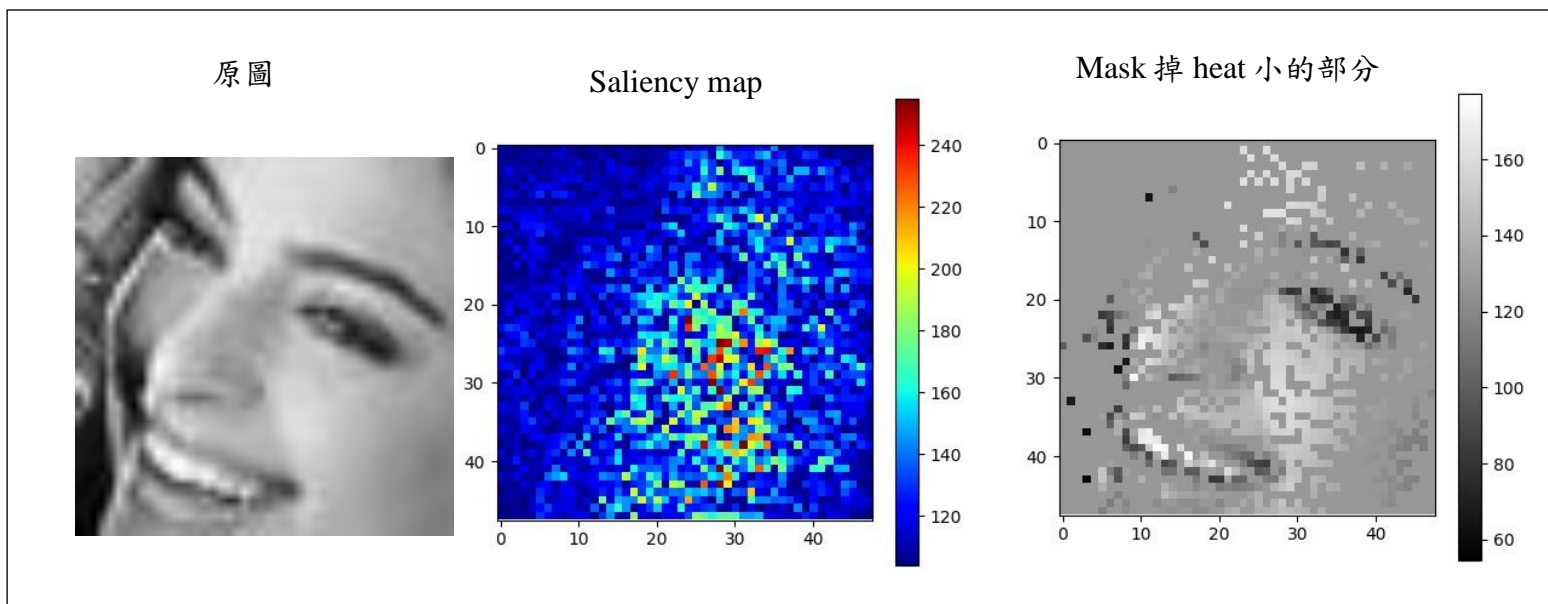


fig1_4.jpg (Sad)

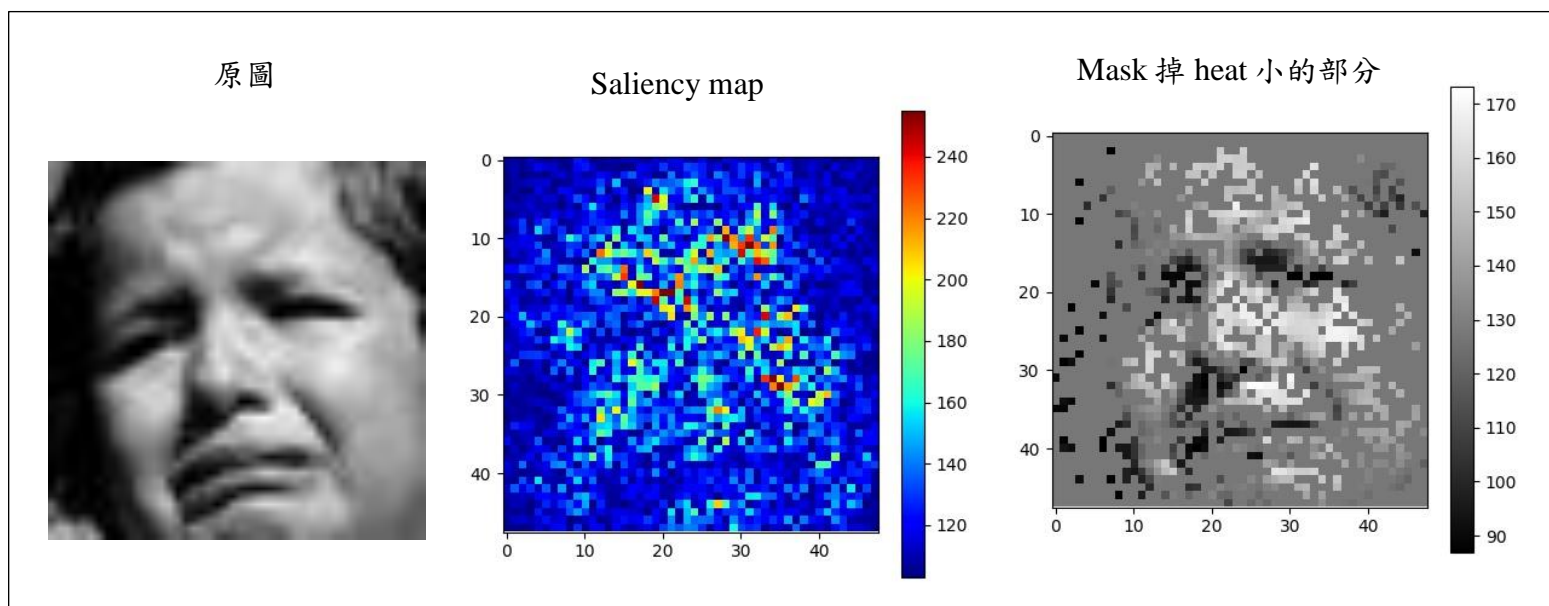


fig1_5.jpg (Surprise)

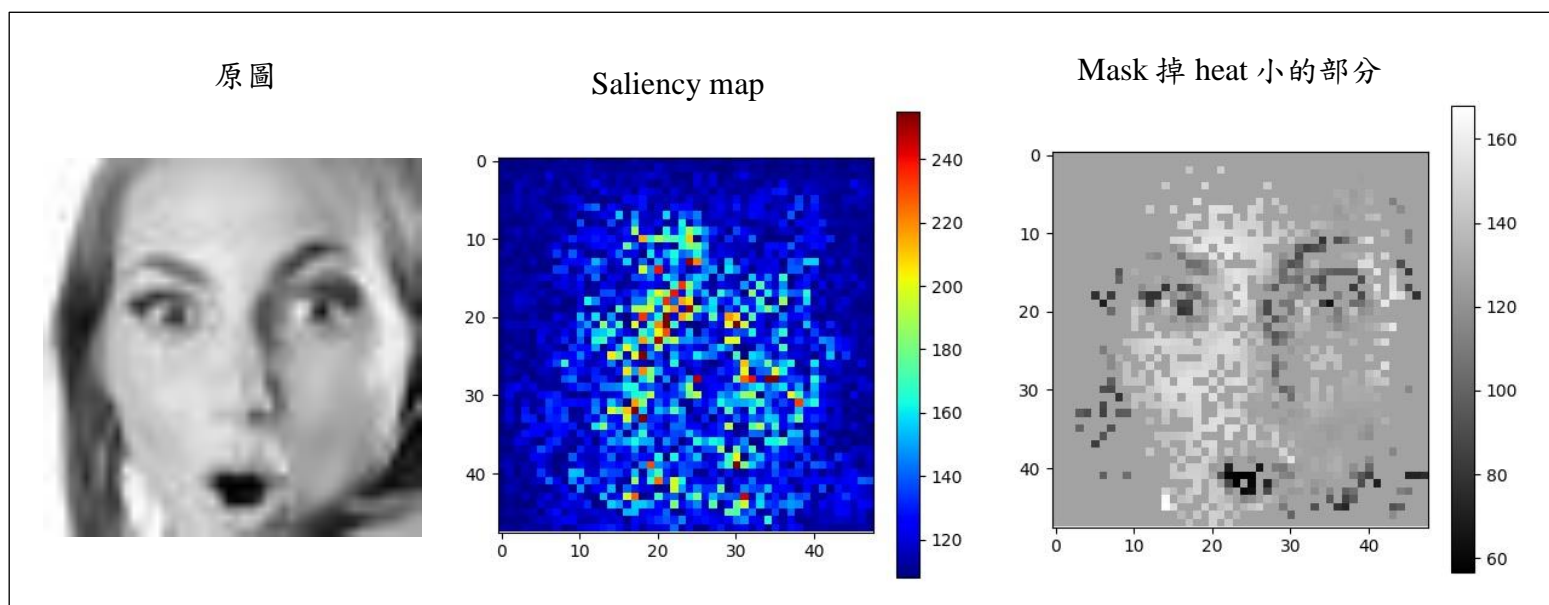
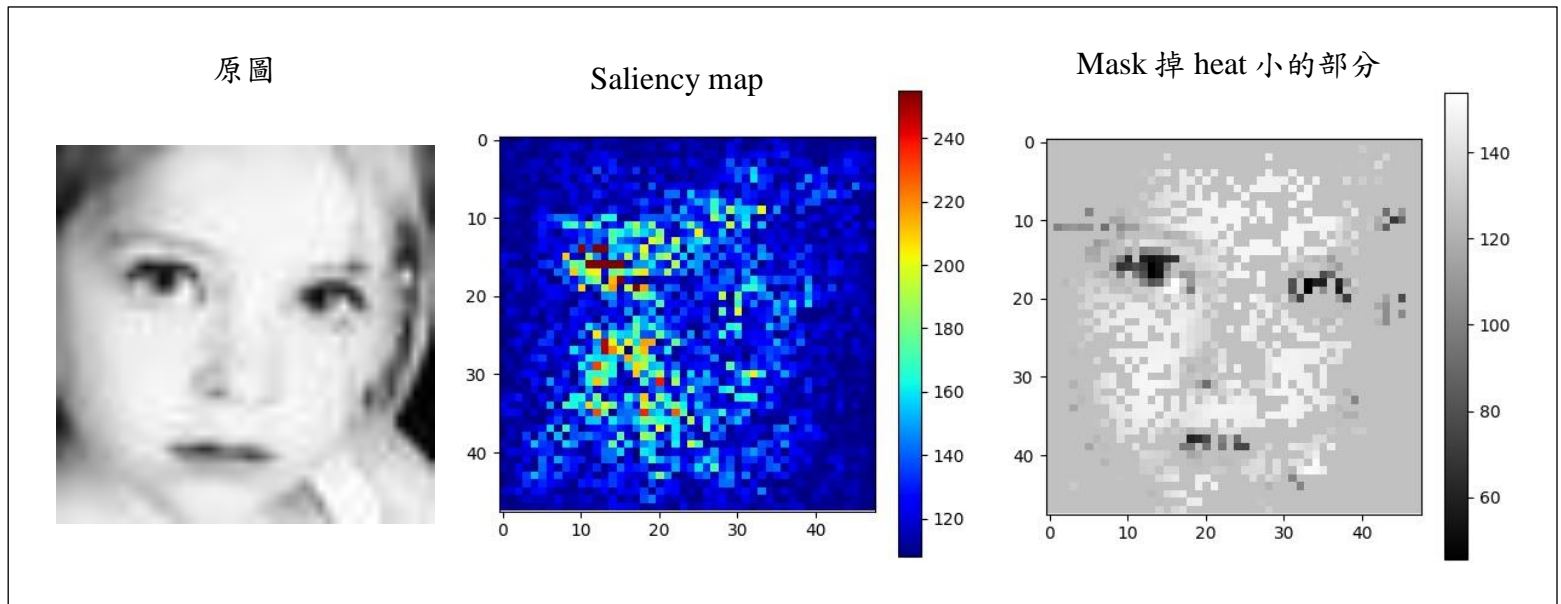


fig1_6.jpg (Neutral)



由上面的 saliency map 可看出模型在做分類時主要 focus 在五官，眼睛和嘴巴尤其明顯，而這也符合我們的常識，眼睛和嘴巴本來就是平常我們判斷一個人的情緒的最主要指標。

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: 無)

答：

fig2_1.jpg

The first 32 filters of conv2d_2

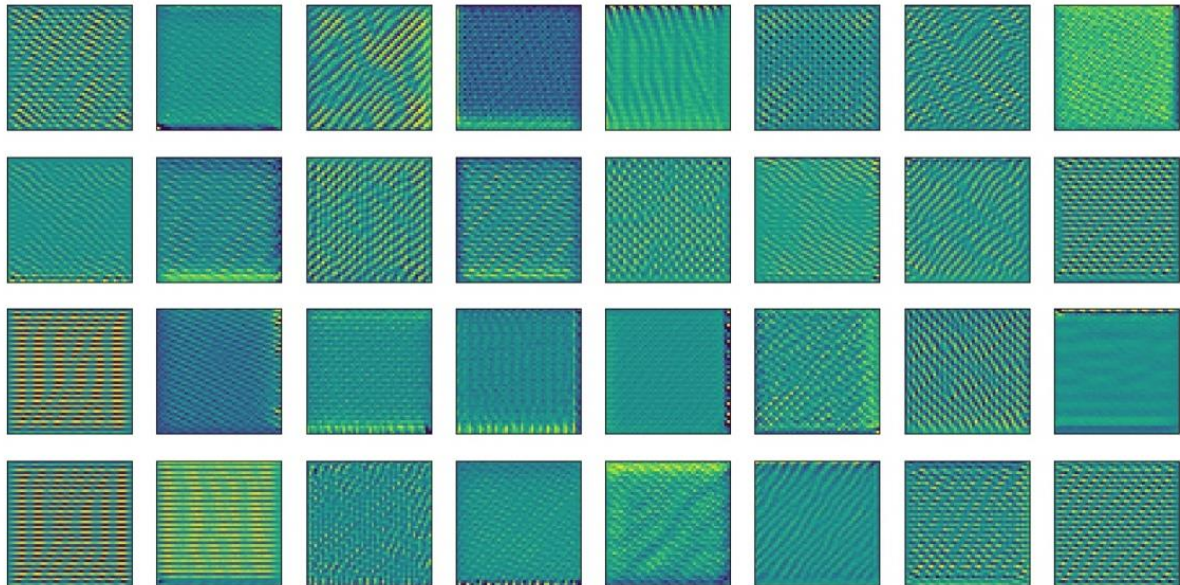
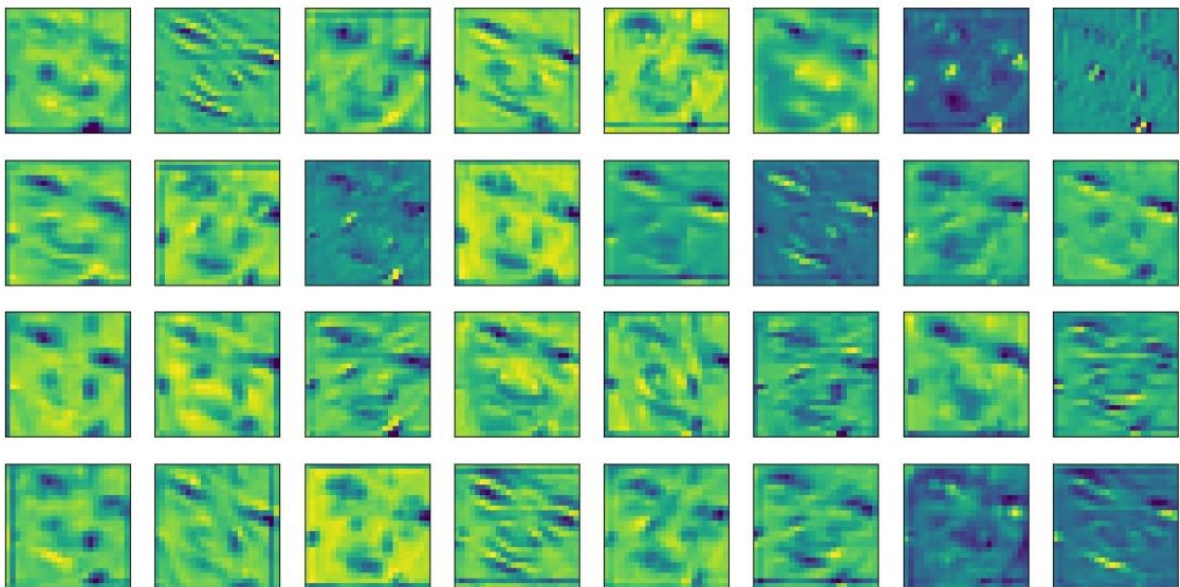


fig2_2.jpg

The output of the first 32 filters of conv2d_2



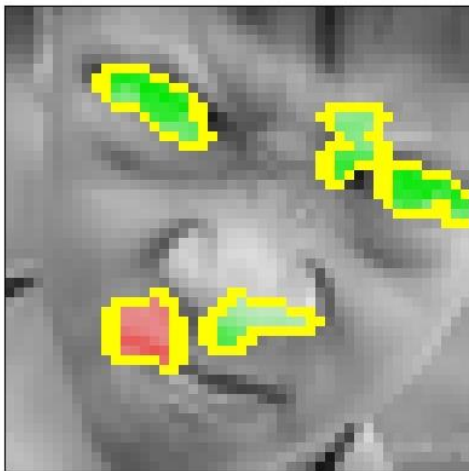
Conv2d_2 是我模型中的第二層 convolution layer，在此之前已通過一層 maxpooling layer，故其輸出的圖片大小為 24x24。觀察 fig2_1.jpg 可看出這一層主要是在偵測一些顆粒與紋理，包含了斜條紋、橫條紋、直條紋等。而我在 fig2_2.jpg 丟進去的圖片為我在 fig1_0.jpg 中使用的那張圖片，如下圖，觀察 output 可發現這一層的 filter 的輸出主要對眼睛、鼻子、嘴巴等十分敏感，這些器官在 fig2_2.jpg 中十分清晰。



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：

fig3_0.jpg (Angry)



我的模型覺得這張圖的眼睛、鼻子部分讓它覺得這張圖是生氣，嘴角則是讓它覺得這張圖不是生氣，我推測可能是生氣時眼睛、鼻子會有皺紋。嘴角則是有可能誤判成中立或其他表情。

fig3_1.jpg (Disgust)



我的模型覺得這張圖的眼睛、鼻子部分讓它覺得這張圖是厭惡，我推測可能是厭惡主要是從眼神看出。

fig3_2.jpg (Fear)



我的模型覺得這張圖的眼睛部分讓它覺得這張圖是恐懼，我推測可能是恐懼主要是從眼神看出。搭配 **confusion matrix**，我的模型對恐懼的辨識成功率偏低，推測是恐懼的其他部分很像中立、難過等。

fig3_3.jpg (Happy)



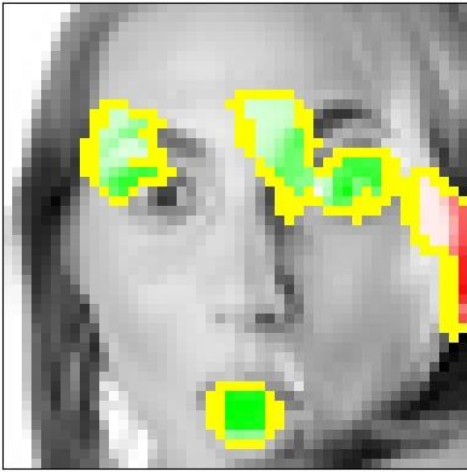
我的模型覺得這張圖的嘴巴部分讓它覺得這張圖是開心，我推測可能是開心時很常露出牙齒。搭配 **confusion matrix**，我的模型對開心的辨識成功率偏高，推測應該是其他的表情比較不會露出牙齒。

fig3_4.jpg(Sad)



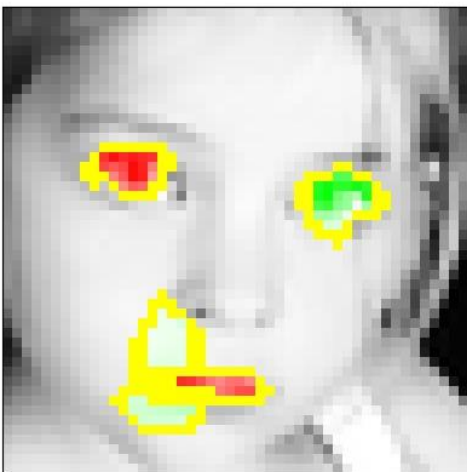
我的模型覺得這張圖的法令紋、嘴角、左眼部分讓它覺得這張圖是難過，我推測可能是難過時會有嘴角下垂、法令紋出現等特徵。另外右眼對我的模型來說是讓他覺得這張圖不是難過，推測是難過時較常眯眼而這張圖的右眼張得比較開。

fig3_5.jpg (Surprise)



我的模型覺得這張圖的眼睛、嘴巴部分讓它覺得這張圖是驚訝，我推測可能是驚訝時常瞪大眼睛、張大嘴巴。搭配 **confusion matrix**，我的模型對驚訝的辨識成功率偏高，推測應該是其他表情比較不會瞪大眼睛和張大嘴巴。

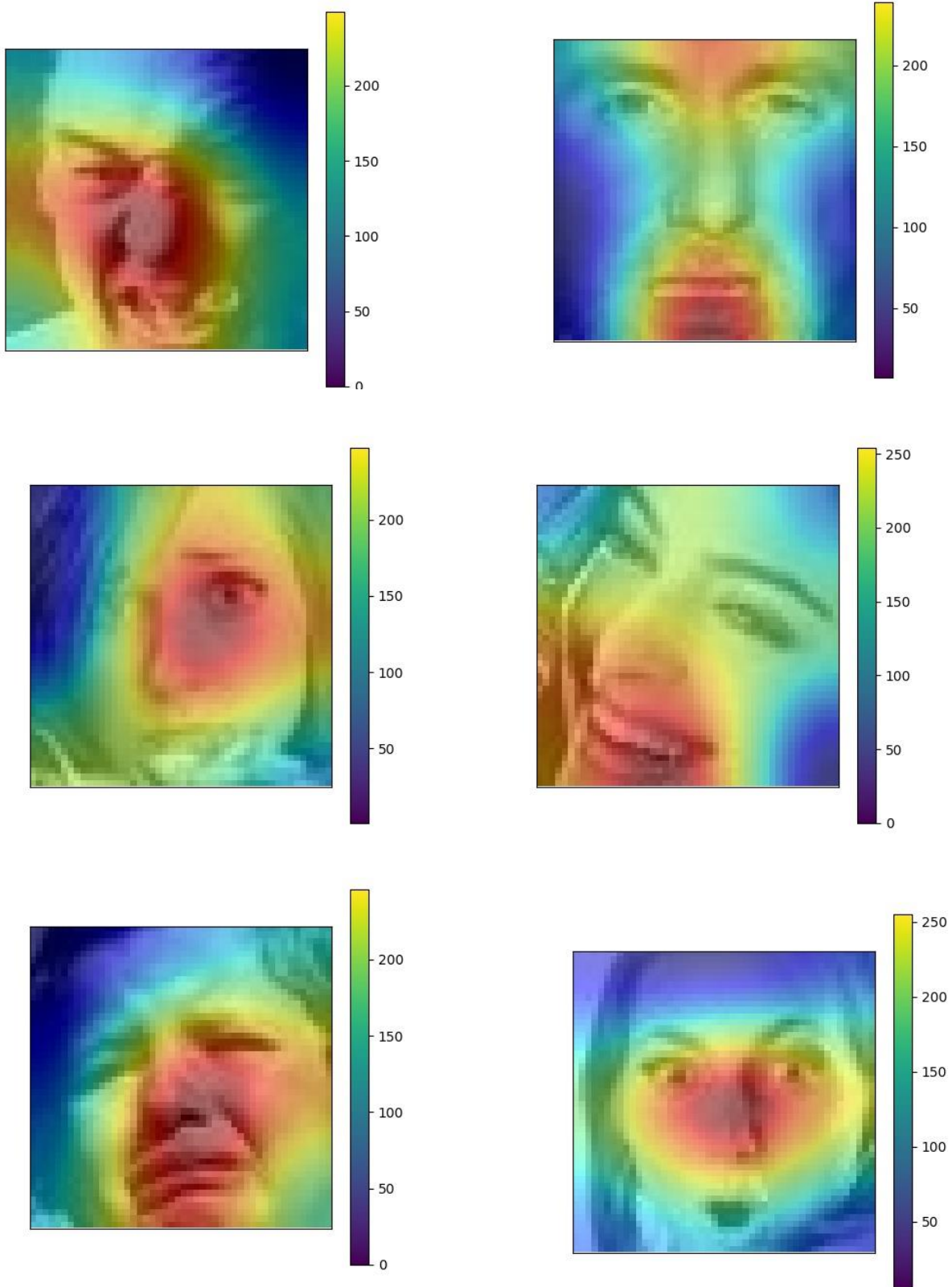
fig3_6.jpg (Neutral)

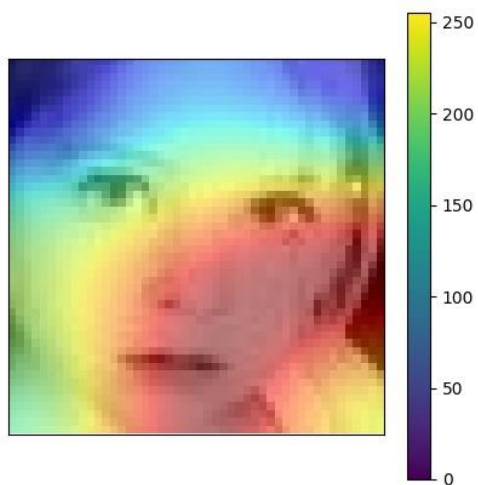


我的模型覺得這張圖的眼睛、嘴巴部分讓它覺得這張圖是不是中立，我推測可能是很多表情的眼睛或嘴巴和中立很像的關係所導致。

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：





以上是我使用 keras-vis 套件的 `guided visualize_cam` 對上面所選的照片畫出來的結果，再次驗證了我的模型的確對五官較為敏感，尤其是眼睛、鼻子和嘴巴。