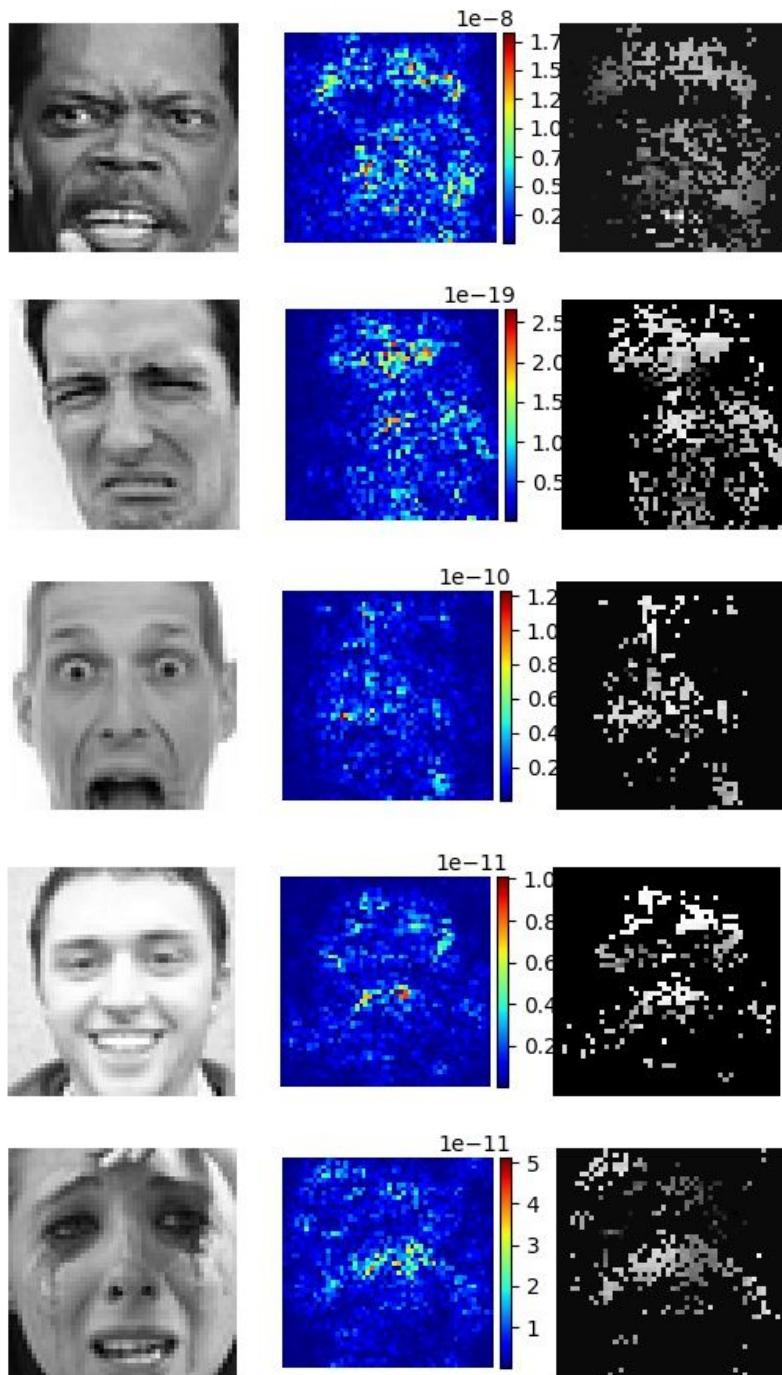
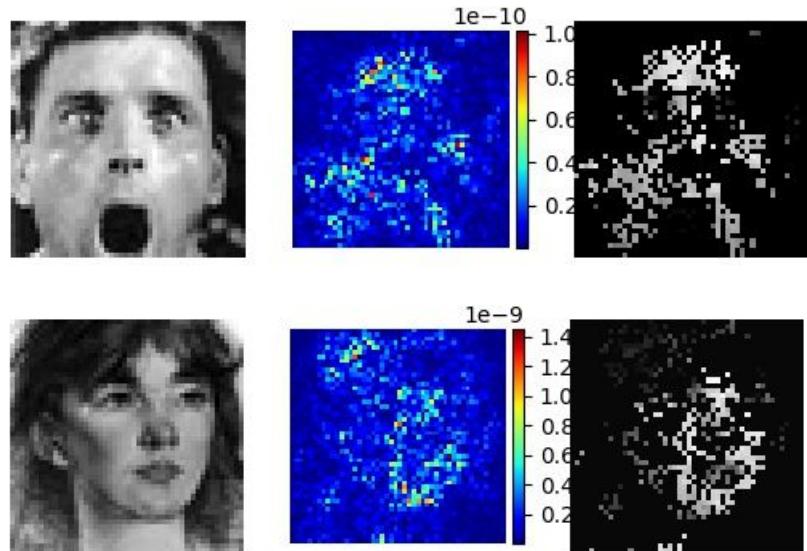


1. (2%) 從作業三可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？(Collaborators:)

答：

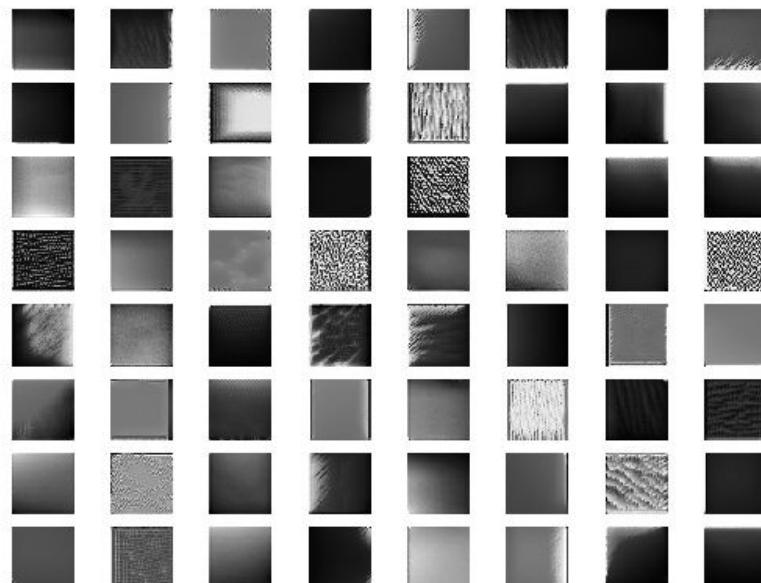




由圖可以發現，模型主要都focus在眉毛與額頭、鼻子與鼻子下方、嘴巴上緣等部分。由每張原圖可以發現這些圖的眉毛與額頭等focus到的部位都因為表情而有明顯的變形，模型可能就是依據這個特性來判斷表情。

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:)

Filters of layer conv2d_2



答：

我以第二個conv2d來做。可以看到有些圖片沒有特別的紋路，而比較像不同角度亮暗分界，猜測這可能是代表這些filter在判斷照片的光源位置，因此這些filter比較容易被對應方向的光源activate。另外有些層有不同角度的紋路，這可能是作比較基本的圖形分割，因此比較容易被filter比較容易被對應角度的紋路activate；。下圖是一張經過filter後的output及其原圖。

Filters of layer conv2d_2



由上圖可以發現，一些在activate時幾乎全黑的圖與有漸層在照片通過後陰影消失(如(第1列, 第3行)的filter, (第3列, 第3行)的filter)，可能是在擷取臉部表情；有些則是臉部表情模糊化，而陰影被突顯(如(第4列, 第5行)的filter)，可能是在擷取臉部面塊。有些則是把表情做少許角度扭曲(如(第8列, 第5, 6行)的filter)，可能是為了讓模型有更多判斷表情的資訊。

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)。



上圖分別為Confusion Matrix跟代表生氣、厭惡、恐懼、高興、難過、驚訝、中立的lime分析。我們可以由Confusion Matrix得知模型容易把生氣、恐懼、難過混淆。其中生氣的dataset中，大部分的臉部動作為張大嘴巴或緊閉嘴巴，這兩個動作分別跟恐懼或難過的臉部動作相似，因此推論可能混淆。由lime的生氣分析可以發現，決定生氣表情的因素集中在眼睛邊緣、眉毛、鼻子的張大皺摺與嘴巴，而恐懼與難過的決定因素也類似，因此可以推論這些臉的這些區塊是決定他是不是這三種表情的主要原因。厭惡較難分析高準確度的原因，他跟生氣有著類似的辨識依據，但可能因為細微表情變化剛好能被模型所辨識，才能得到較高的準確度，另外，也可能與資料過少有關造成計算出的準確度不準有關。lime分析高興的決定因素在於嘴巴與臉頰，這或許是因為笑的時候臉頰會有被抬起來的變化，由於高興時臉部表情會明顯與其他表情不同，可能因此較容易分辨。驚訝集中較多在嘴巴與眼睛，可能因為驚訝時的眼睛睜大與其他表情不同，因此較容易分辨。lime分析模型在中立的時候判斷主要集中在額頭跟鼻樑、嘴角，可能因為沒有表情時這些地方比較不會有擠壓變形有關。

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：

我用教授上課所提到的方法：用一個block box去遮住圖的某些部位來看他的預測結

果與分數，進而判斷模型是以哪邊為依據判斷的。最下面的圖為結果：

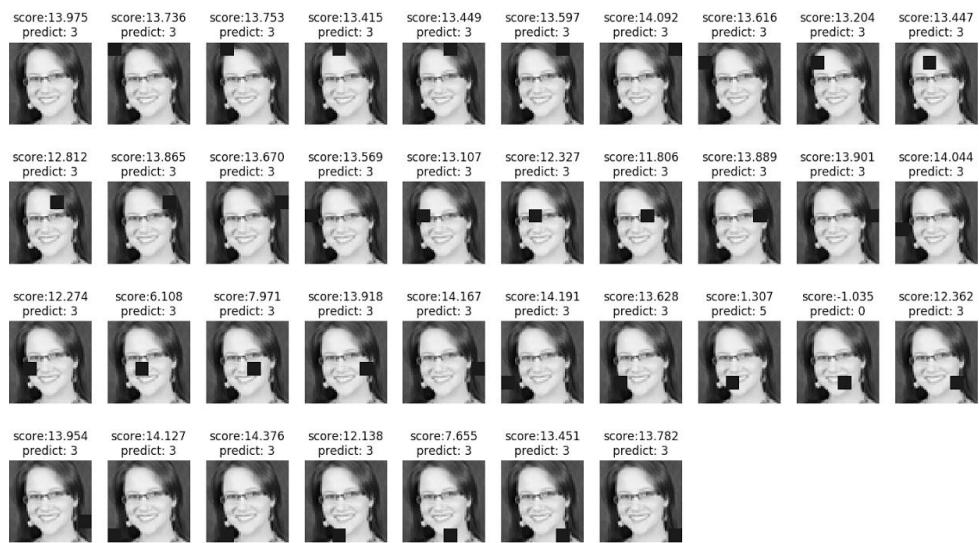
生氣的圖中，可以發現遮住嘴巴附近所得到的分數會最低，有時甚至會造成判斷錯誤；厭惡的圖中，遮住鼻子、臉頰與嘴巴上緣時，有時後會使模型判斷成生氣；恐懼、高興、難過則分別是遮住額頭、嘴巴跟鼻子造成分數降低最明顯；驚訝、中立則是在遮住鼻子與人中附近造成分數下降最明顯。結果與saliency maps、lime圖大致符合。



ground truth: 2



ground truth: 3



ground truth: 4



ground truth: 5



ground truth: 6

