
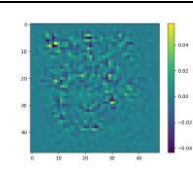
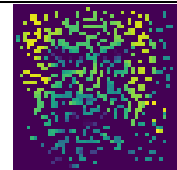

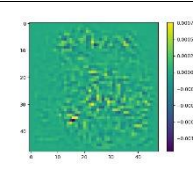
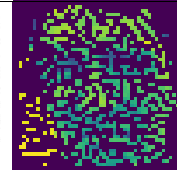

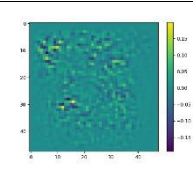
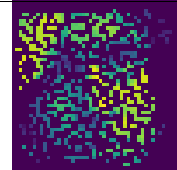

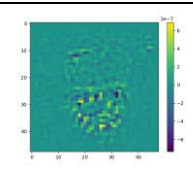
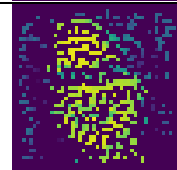

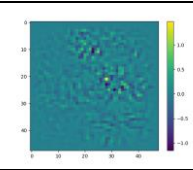
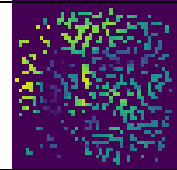

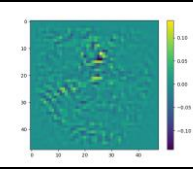
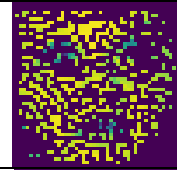

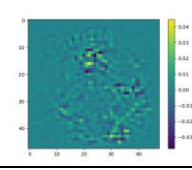
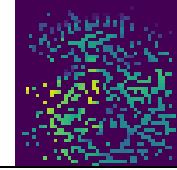


1. (2%) 從作業三可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？

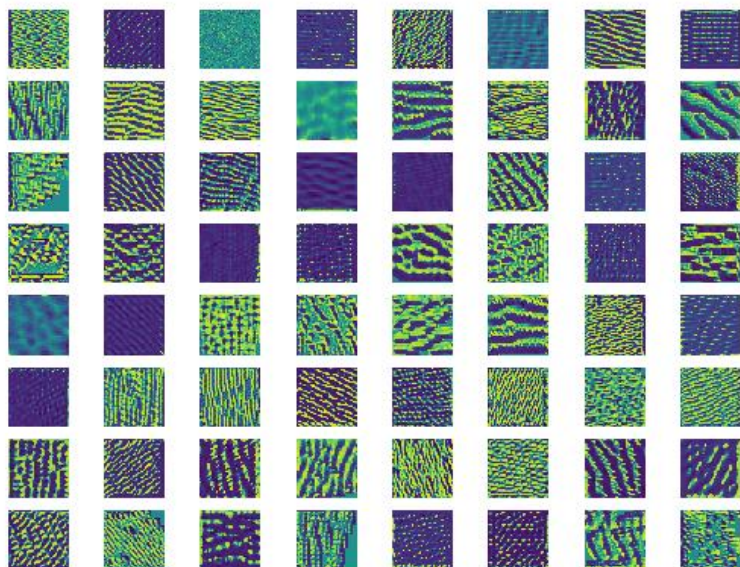
(Collaborators: 何俊緯)

答：由 saliency maps 中可發現，大部分亮點皆集中在人臉上，且在五官及臉的輪廓上特別明顯，由此可知此 model 的確有於正確的位置進行人情緒辨識。

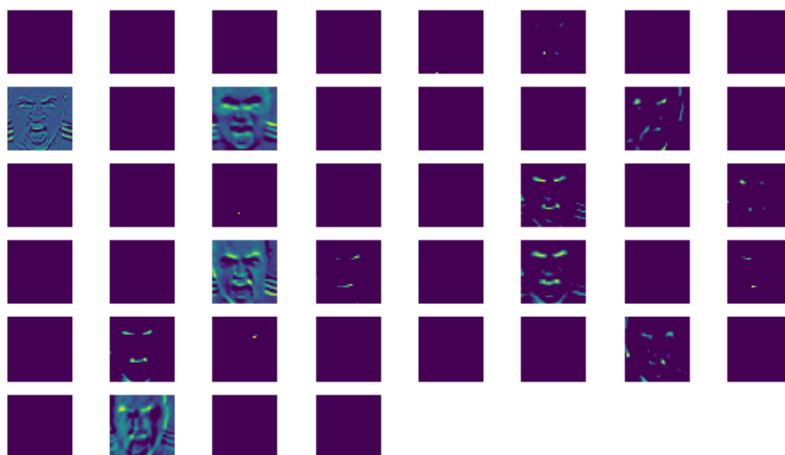
種類	origin img / saliency map / mask			
0				
1				
2				
3				
4				
5				
6				

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: 何俊緯)

答：下圖為第一次 maxpooling 層的 filter 對應圖片種類。由圖片以及輸出時計算的 loss 可以知道，其中有些 filter 對於 input 的改變是幾乎沒有反應的，如第一排左邊數來第三個 filter 的 loss 就幾乎沒有受到影響，得出的結果也是近乎無紋理。

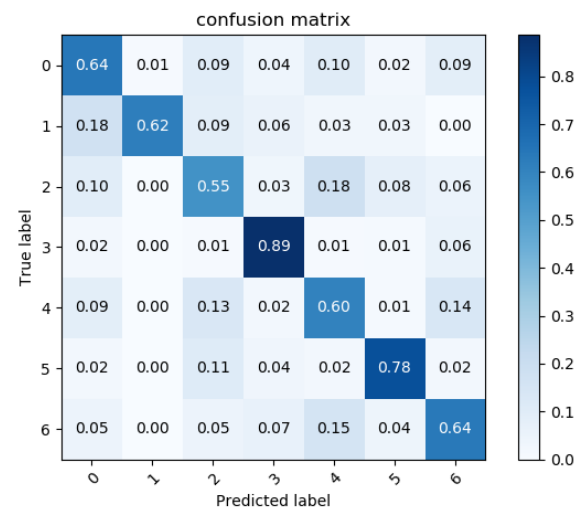
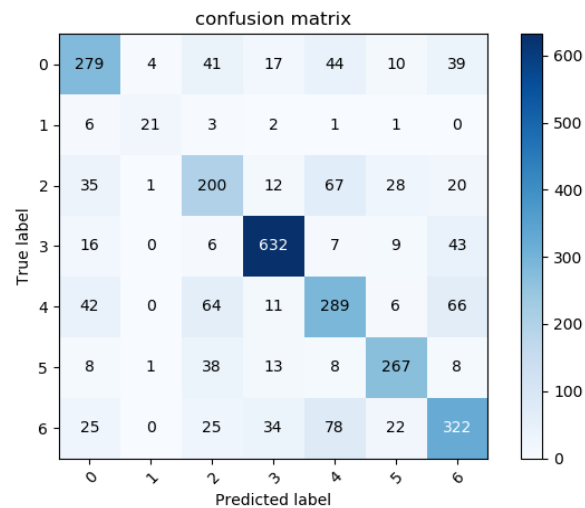


另外，我取了第一層 conv2D 層的 output 進行分析，可以發現這 44 個輸出(因為 padding 的關係減少了幾個)都幾乎有臉的輪廓，只有一些近乎全黑，也就是對於輸入的圖像毫無反應。



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：下圖由左而右為 0-6 種不同情緒分析得出的結果，其中黃色圈起來的部分為對於推測結果較為有幫助的部分，可以發現幾乎都聚集在五官處，只有 3 (高興) 是特別 focus 在嘴部的表情，極為明顯，也有頗高的準確率。其中 1 (厭惡) 的 data 比起其他組資料相對少很多，因此造成 focus 的位置沒有確切的在五官上。2 (恐懼) 4 (難過) 兩者各 focus 的點雖大略在五官上，但仍有偏差，造成 predict 結果並非非常高。



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：因為在進行 training 的時候有發現一些並非臉譜的 training data，因此我想藉由 saliency map 來觀察究竟我的 model 在碰到爛掉的 training data 時表現的狀況。下面兩張圖，左圖是將所有 data(包括非臉譜)進行 training 得到的 model，右圖則是刪去非臉譜 data 的 model，雖然不明顯，但是藉由第一張和第四張圖發現左 saliency map 圖有略為勾勒出虛假的臉譜表情，由此推測若 testing data 中有爛掉的 data，仍要在 training data 中加入同樣壞掉的 data，才能更準確的預測爛掉的 data 結果。

