

學號：B05203050 系級：化學三 姓名：陳品翰

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：

原圖（左至右為 class 0 ~ 6，皆判斷正確且 probability > 0.9）：



saliency map (mask 掉 > 215 者)：



可見機器似乎較不注重眼睛、嘴巴本身，而是額頭、臉部紋路，與下巴等周圍部位。

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:)

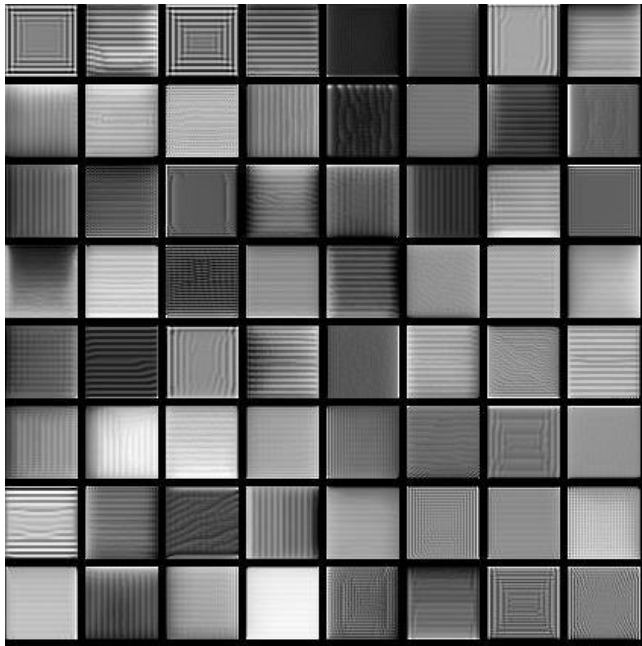
答：

可見 filter 容易被有特定直、斜線等條紋，或長方形等形狀 activate。

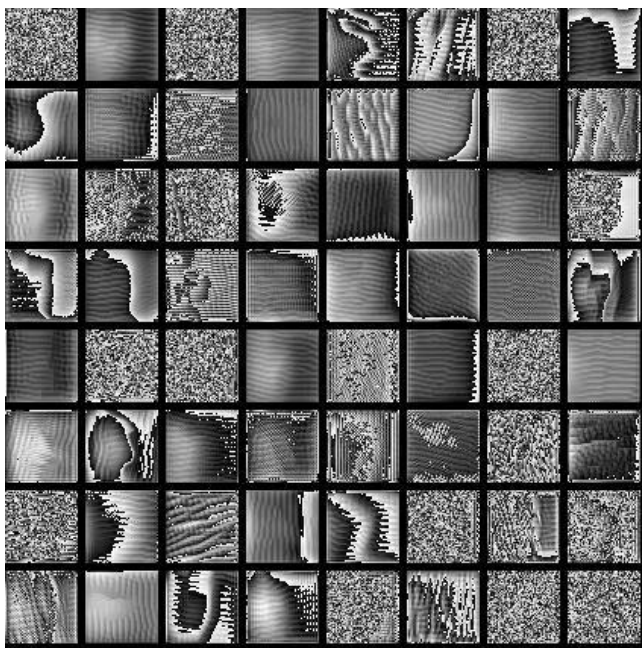
從 output 則可發現有的 filter 呈現與原圖類似之臉型輪廓，應是易被此種圖片 activate 者。有的 filter 則呈現雜訊或單一顏色之條紋，應是不易被 activate 者。



第五層其中 64 個 filter



以上 filter 對原圖之 output



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：

(左至右為 class 0 ~ 6，皆判斷正確且 probability > 0.9)：



分別列出對不同特徵著重判斷的 class：

眼睛	眼睛周圍	鼻子	鼻孔	嘴巴	嘴巴周圍	臉頰	下巴
0, 2	1, 4, 6	3, 6	0, 1, 4, 5	2, 5	2, 4, 6	1	0

可見偶數 class (生氣、恐懼、難過、中立) 皆著重判斷眼睛、眼睛周圍、嘴巴周圍等範圍較小，且差距可能較不明顯之部位，故彼此間較難以辨別，容易混淆。而奇數 class (厭惡、高興、驚訝) 則可判斷臉頰、嘴巴、鼻子等範圍較大，且差距較明顯之部位，故表現較佳。

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：

使用課堂上提過之 sensitivity to occlusion，對原圖的不同區塊進行遮蔽，觀察得到 probability 之改變，若遮蔽後正確之 class probability 下降，代表該區塊是機器識別圖片的重點，以此繪出 saliency map。

使用的 mask 為 11*11 之黑色正方形，stride = 1，對第一題的同樣六張圖進行分析，並比較兩種方法：

原圖 (左至右為 class 0 ~ 6，皆判斷正確且 probability > 0.9)：



第一題之 saliency map：



本題之 saliency map：



仍可見與第一題之方法相似的輪廓，但效果較差，應是因為圖片太細，mask 容易一次遮到複數個特徵。可見此方法較適合用在大範圍的圖片內，機器著重在哪个部分。

（註：本題之 code 在 others 資料夾之 hw4_4.py，由於怕時間超過並未於 hw4.sh 執行。）