

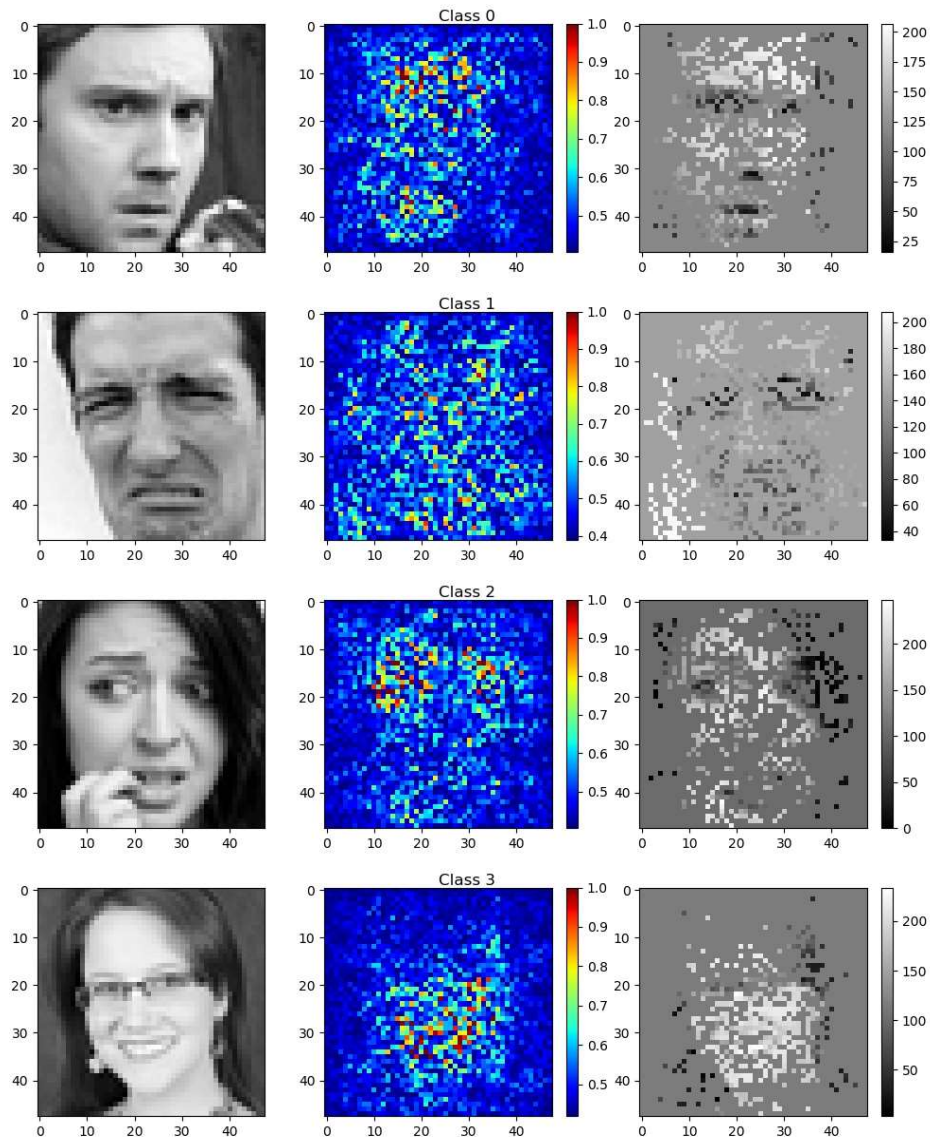
學號：B06902066 系級：資工二 姓名：蔡秉辰

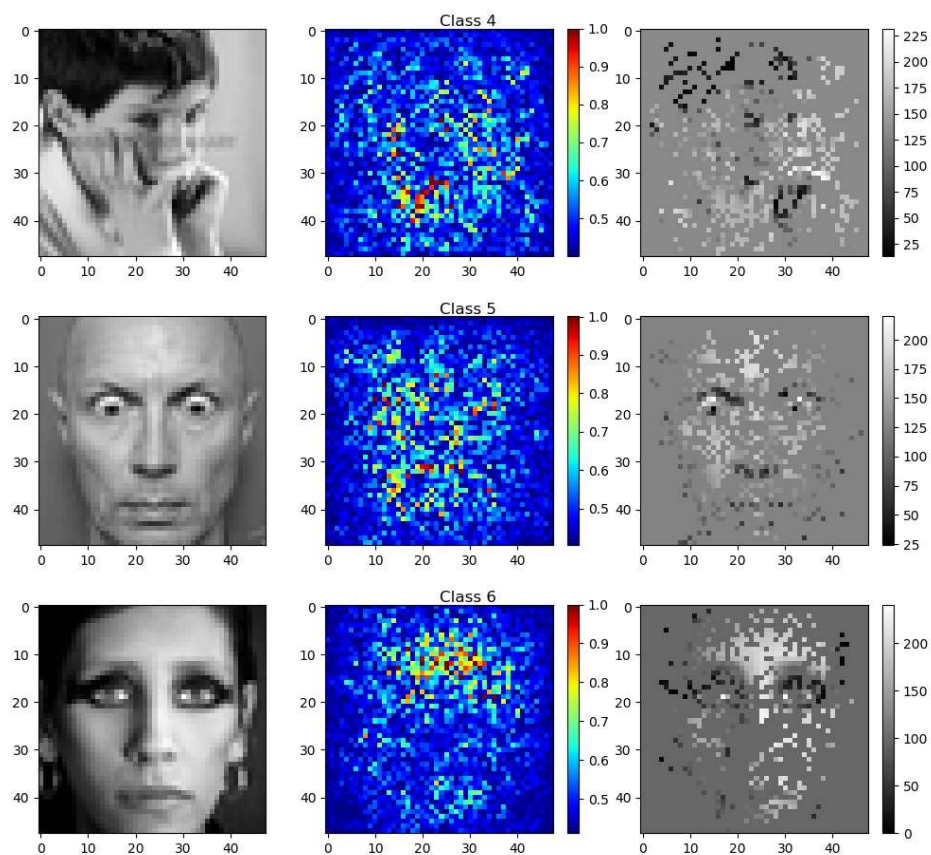
註：本份作業使用之 model 和 hw3 之 model 並不相同，model 構成如下：

```
Conv2D( 32, (3, 3) ) + Conv2D( 32, (3, 3) ) + MaxPooling2D( 2, 2 ) + BatchNormalization( ) +  
Conv2D( 64, (3, 3) ) + Conv2D( 64, (3, 3) ) + MaxPooling2D( 2, 2 ) + BatchNormalization( ) +  
Conv2D( 128, (3, 3) ) + Conv2D( 128, (3, 3) ) + MaxPooling2D( 2, 2 ) + BatchNormalization( )  
Dropout(0.2) + Flatten( )  
Dense( 256, 'relu' ) + BatchNormalization( ) + Dropout( 0.5 )  
Dense( 128, 'relu' ) + BatchNormalization( ) + Dropout( 0.5 )  
Dense( 64, 'relu' ) + BatchNormalization( ) + Dropout( 0.5 )
```

且有使用 normalization 及 augmentation。

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？

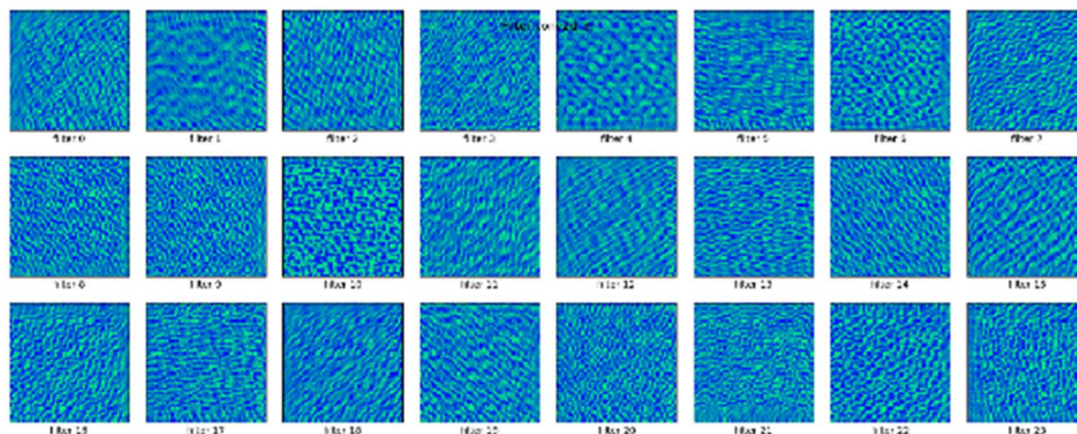


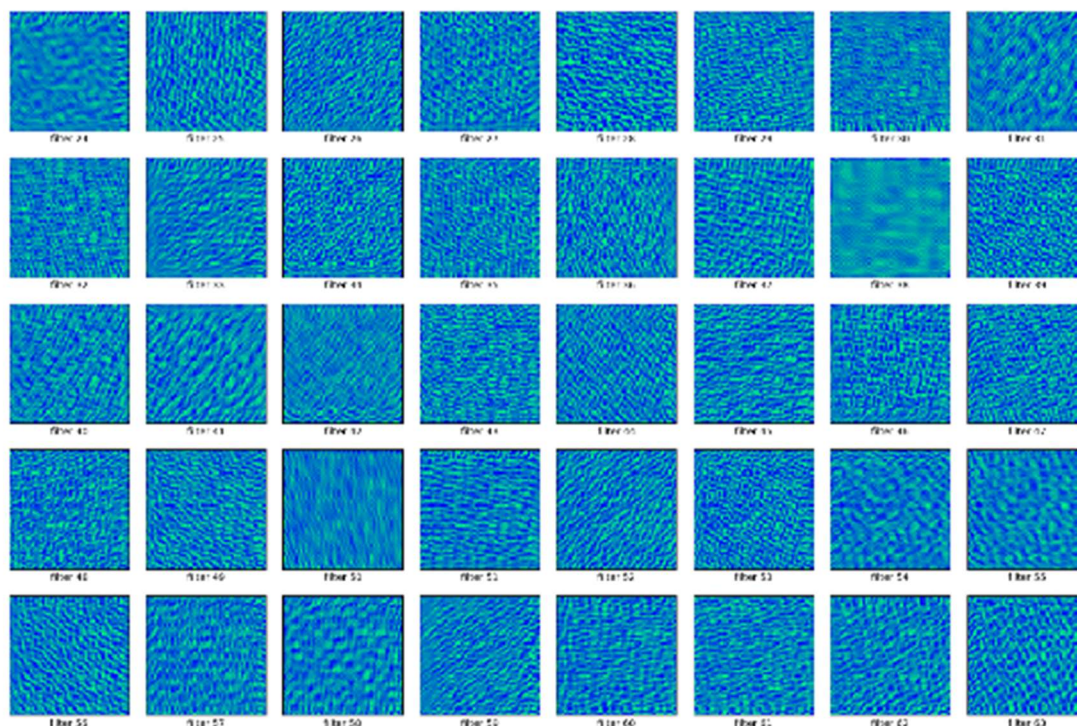


從以上結果可看出，其主要是注重在五官的部分(如眼睛、嘴巴、眉毛)；而像是皮膚、頭髮的部分就較不注重。這和我們平常辨識表情的情況相似。

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: B06902029, B06902049)

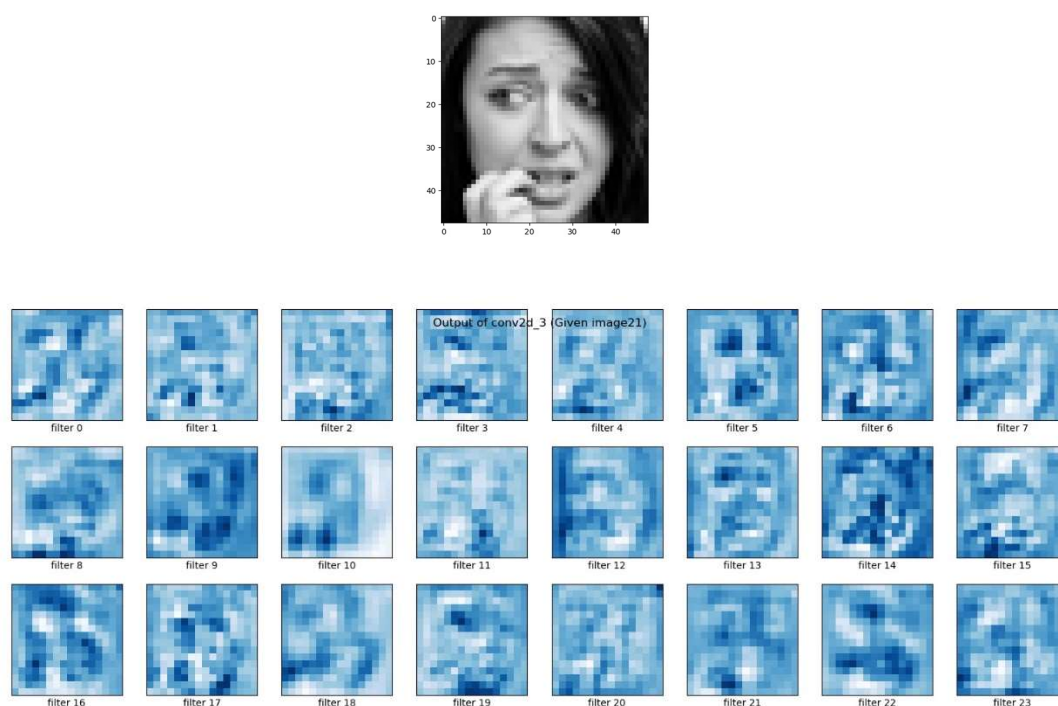
以第三層的 Convolution Layer 來做分析。
首先先觀察最容易被哪種圖片 `activate`：

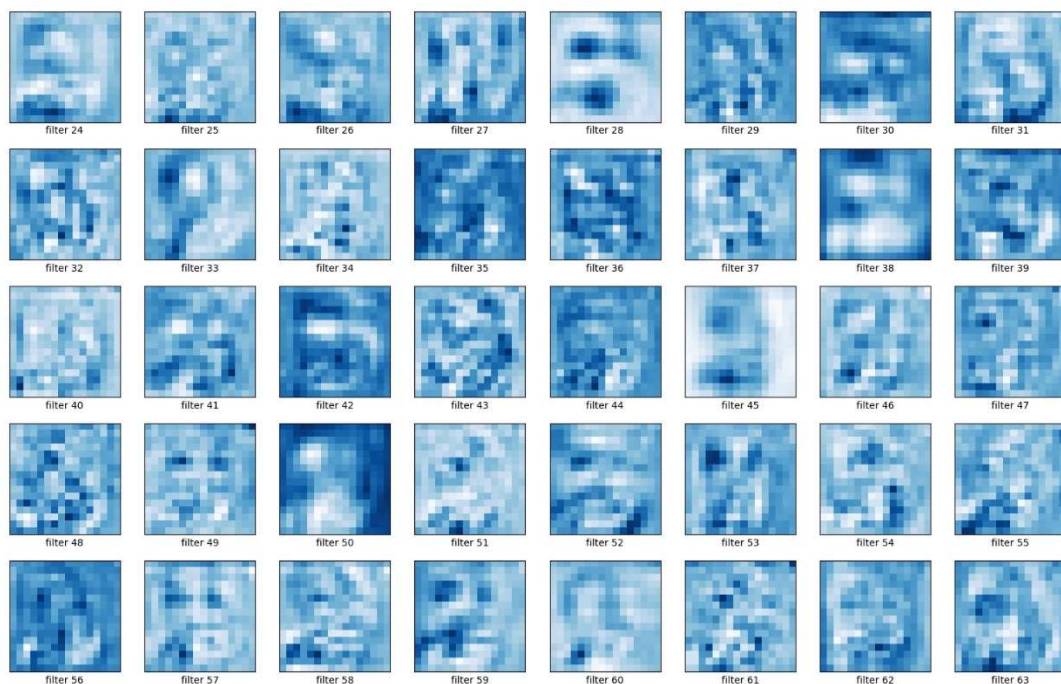




從以上圖片可以看出，在第三層的所有 filter 最容易被 activate 的圖。其中，顏色越區藍色表越深，越趨綠色表越淺。可以看出每種 filter 都有其對應偵測不同的紋路，不過偵測的紋路都不算太複雜，多數趨勢為直/橫/斜紋。畢竟這才第三層，沒有過於複雜的紋路(e.g. 橢圓、多次曲線)也屬正常。

接著觀察 filter 的 output，使用的 21 張圖片。對應之影像及結果如下：





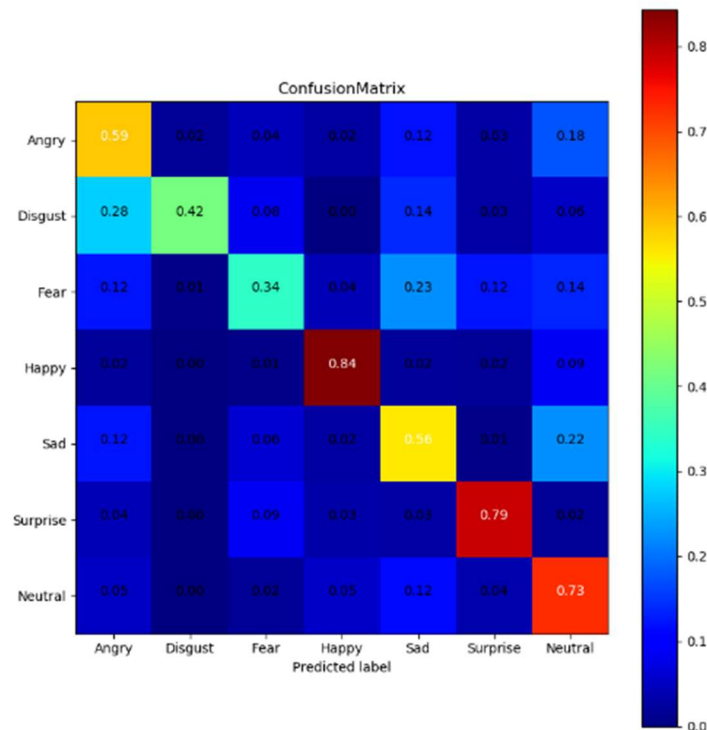
第三層的 Conv2D 是已經經過亦曾 Maxpooling 後的結果。可以看出人臉基本上已經有被簡化的樣子，但仍然能夠辨識出眼睛、嘴巴的樣子，這也符合前面做辨識時，model 較注重之地方。

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)。



可以看到基本上較多判定點(無論正/負相關均是如此)均在雙眼、嘴巴的地方。而從此 model 的 Confusion Matrix(在下頁表示)中可看出，在 Happy、Surprise、Neutral 的判定能力較好。對應到 lime 的分析，可看到 Happy 的圖片的確在嘴巴微笑的地方有正相關之判定；Surprise 在眼睛瞪大的地方亦有正相關之判定，和我們一般人認定之”驚訝時眼睛會睜大”相似；最後是 Neutral 的部分，除了眼

睛為正相關外，在鼻子、眉心附近亦出現正相關，推測是因為如 Angry、Disgust 這類的表情，鼻子周圍會有皺紋。對此，Neutral 在那附近是平滑的，固有正相關。而嘴巴呈負相關，推測是因其有略略上揚所導致。



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。(Collaborator: B06902049、B06902028)



此為使用 Grad-CAM 做出之結果與表情疊圖之後產生之圖片。實作方式為呼叫 keras-vis 內的套件 visualize_cam。

從以上結果可以看出 model 將每張圖分類為對應之表情的依據，多數也都是注重在五官的部分，和實際上我們辨識表情的方式相近。