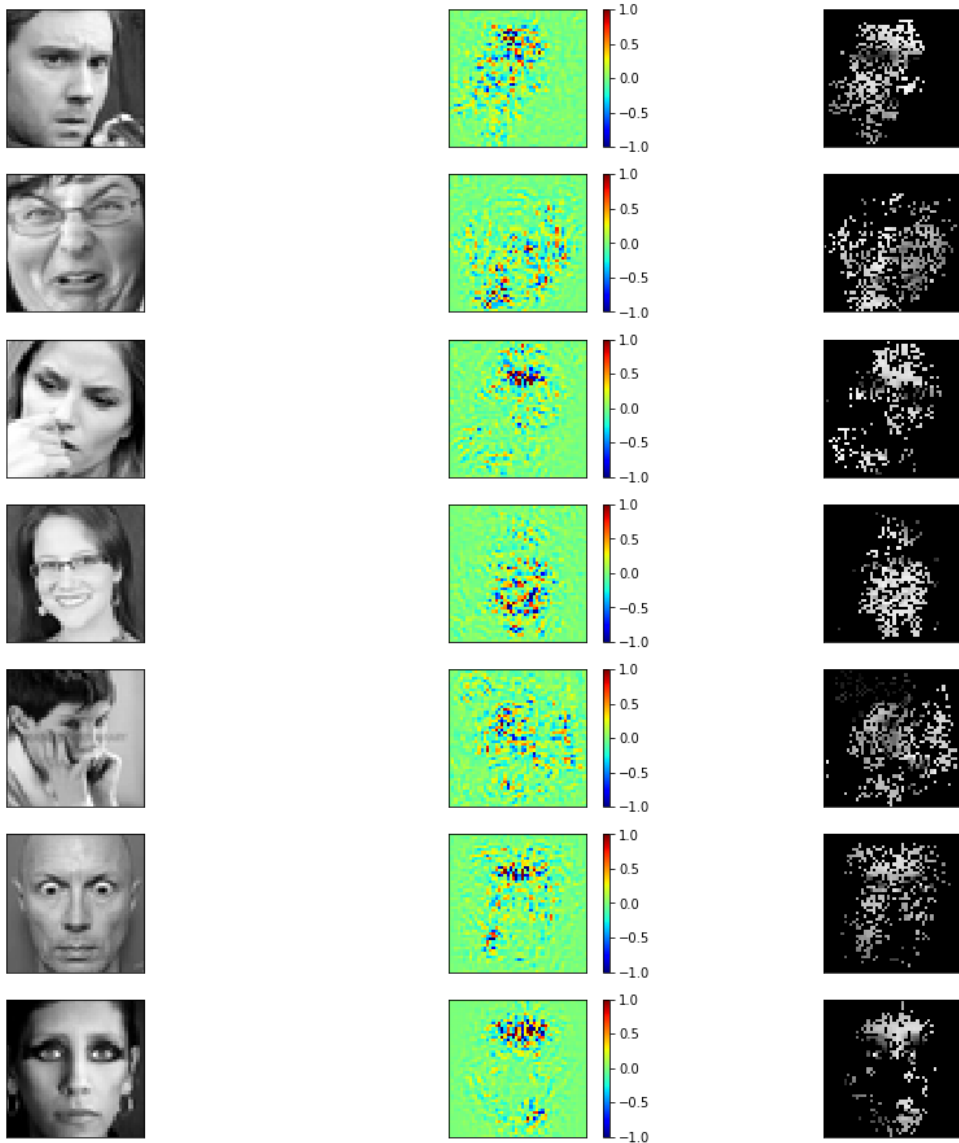
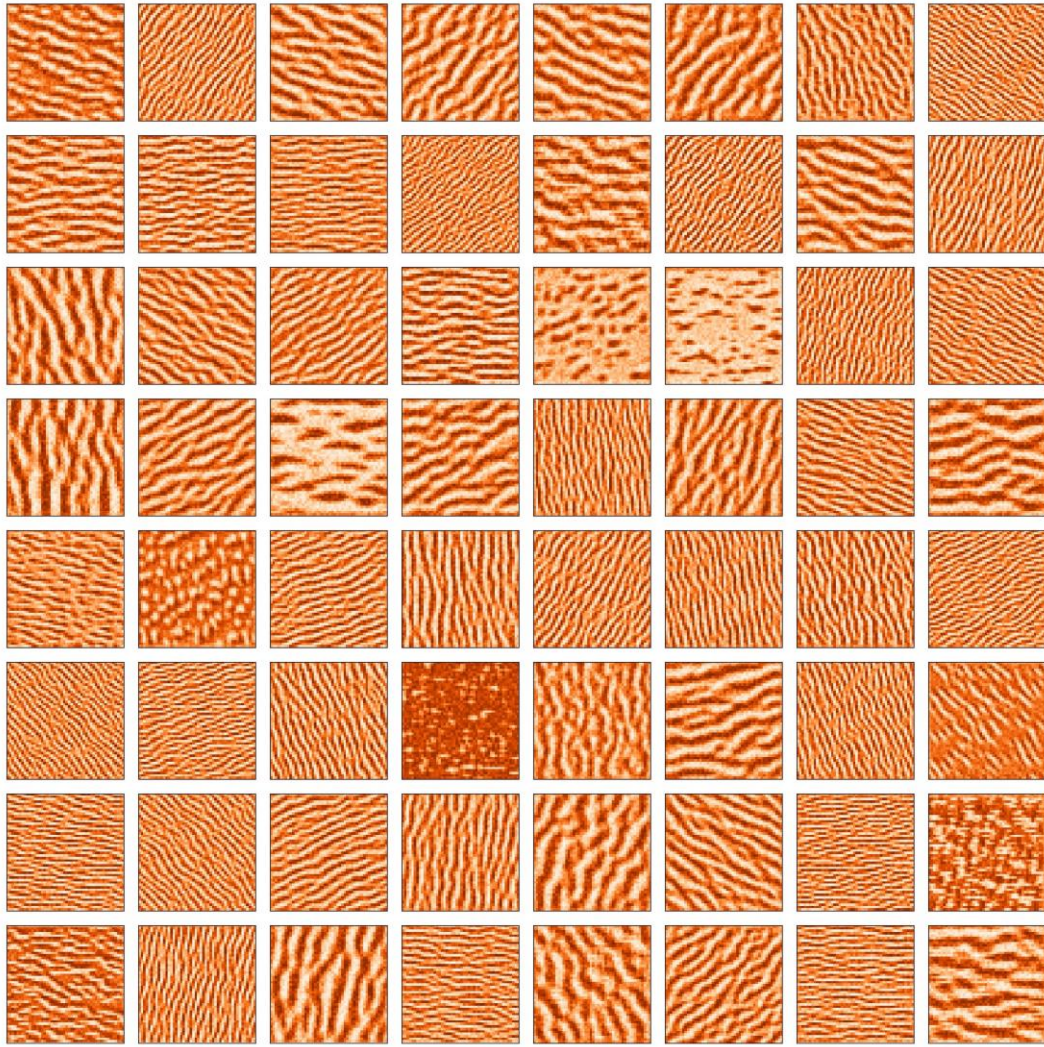


1. (2%) 從作業三可以發現，使用 CNN 的確有些好處，試繪出其 saliency maps，觀察模型在做 classification 時，是 focus 在圖片的哪些部份？

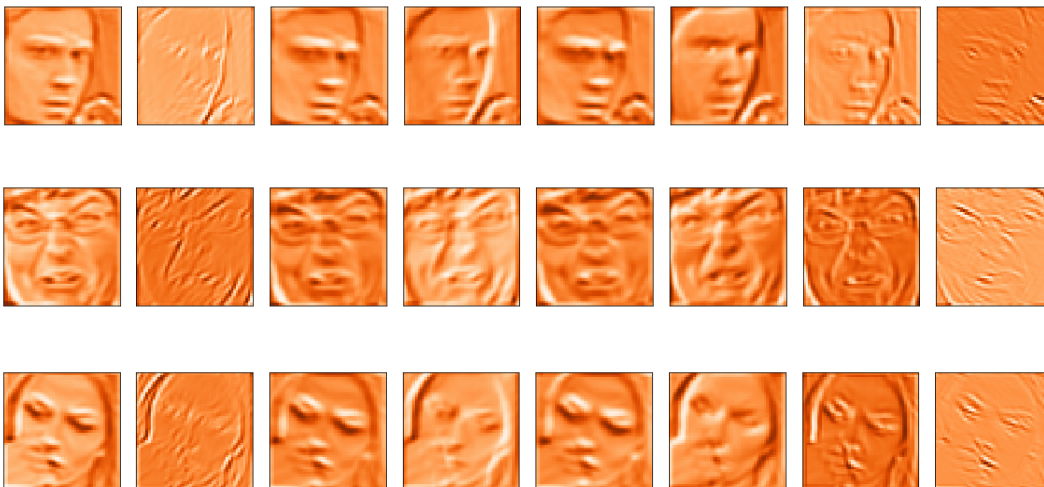


幾乎所有類別都有涵蓋到大部分的眉毛、眼睛那一部分，而有些類別還有涵蓋到些許嘴巴。

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。



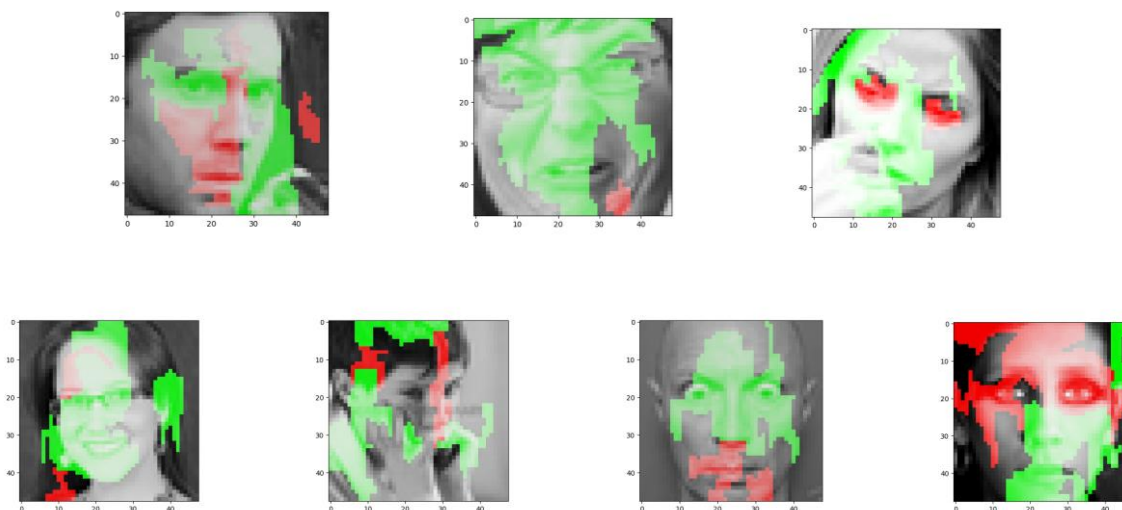
上圖為 conv2d\_1 的 64 個 filter，都具有特定的紋理。





從上到下依序是生氣、厭惡、恐懼、高興、難過、驚訝、中立七種表情在 conv2d\_1 經過前八個 filter 所產生的 output，可以看到對於不同的表情有些 filter 會產生出值較大(偏白)的 output 出來，因此便可以用這些差異做表情的辨別。

(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)。



從左到右、從上到下，分別為生氣、厭惡、恐懼、高興、難過、驚訝、中立七種表情

主要辨別的區域，第三次作業中，難過、中立和驚訝這三個表情很容易搞混，我想就是因為他們主要辨別的區域有很大部分沒有重複，或是 **gradient** 方向完全相反，這樣一張圖就有可能被辨別出三種情況。

3. (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：