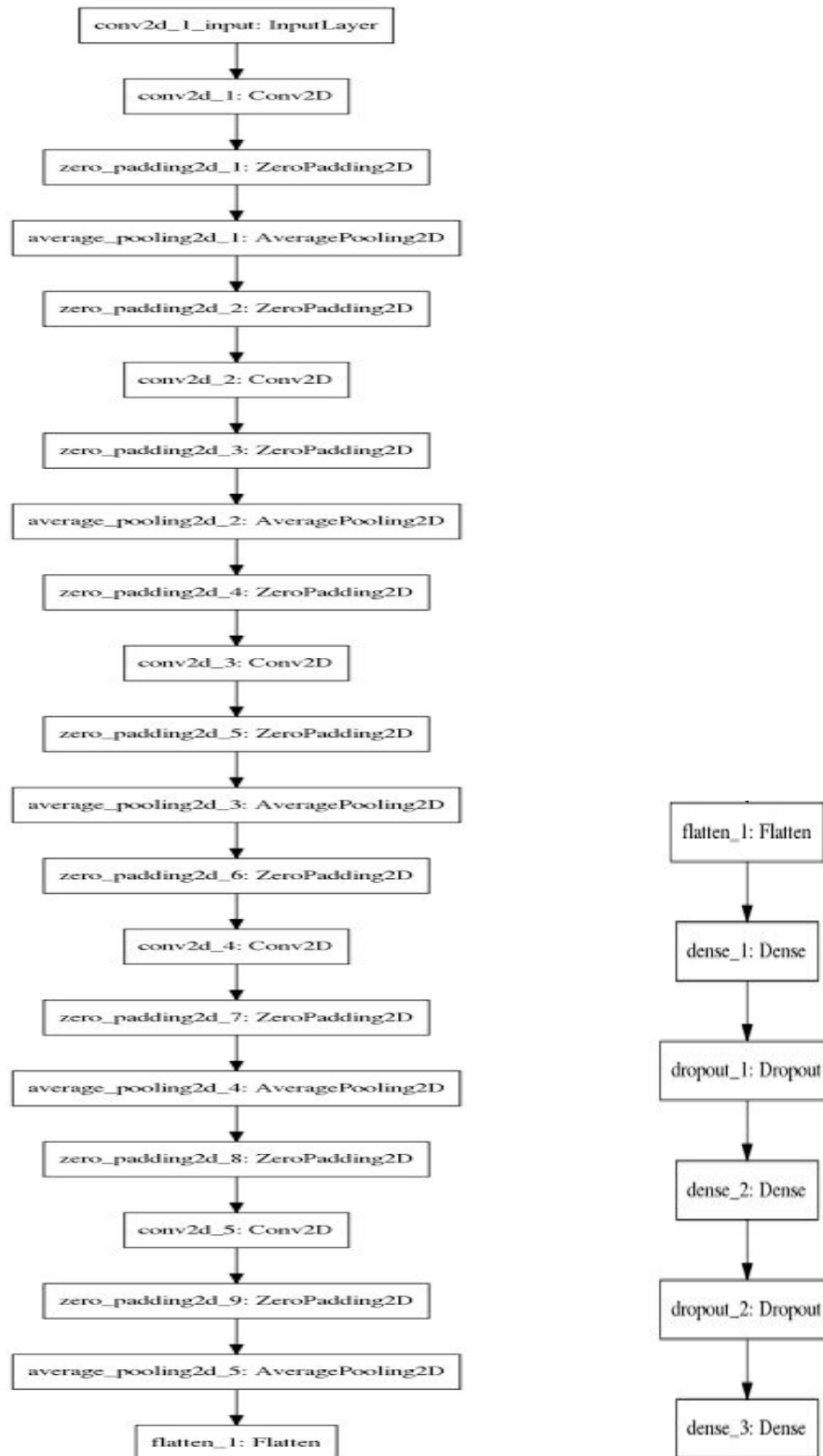
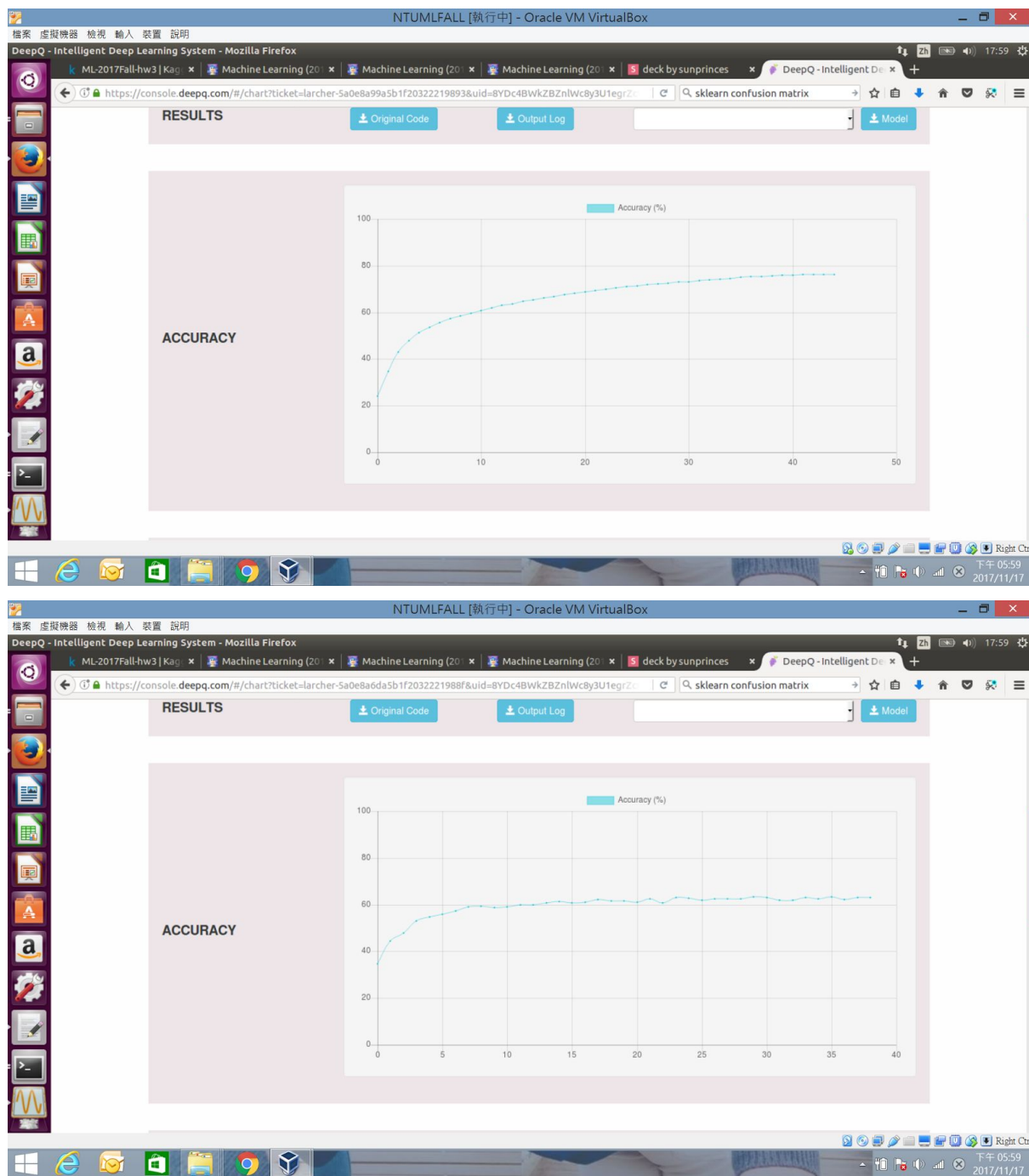


1. (1%) 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？  
(Collaborators: 無)



答：



分別為training及validation accuracy

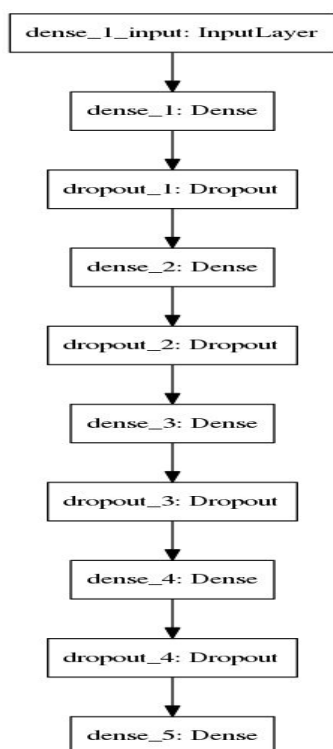
private=0.65589, public=0.65282

參數量約五百萬

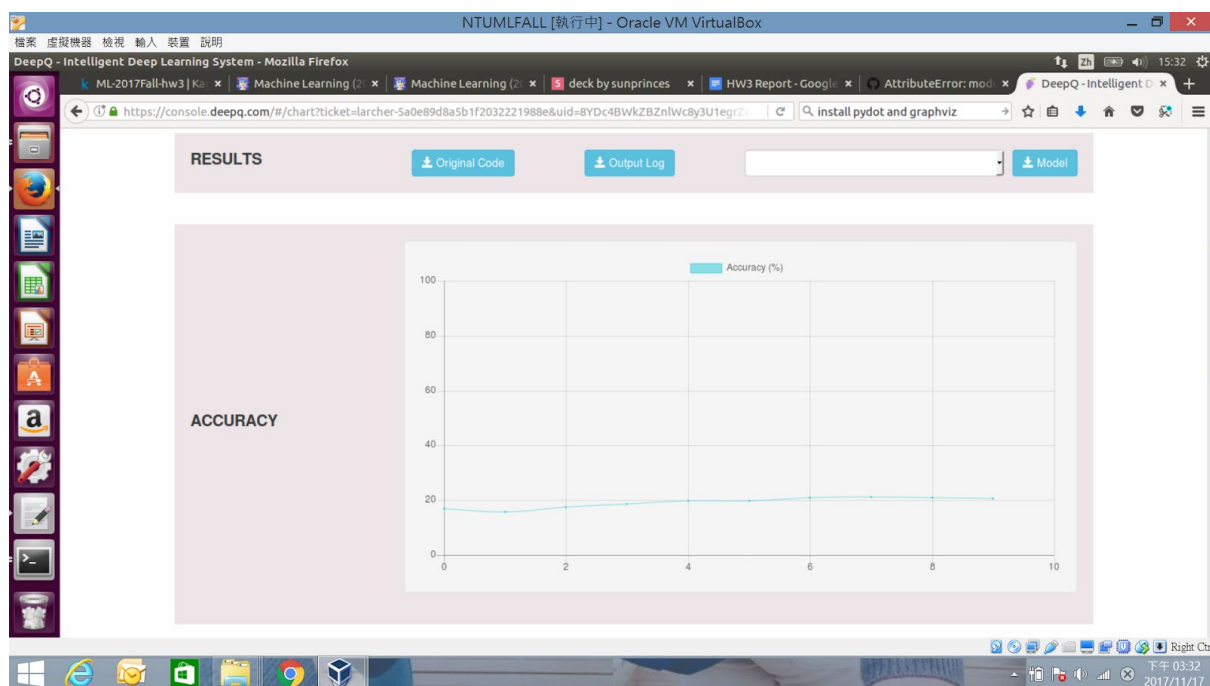
2. (1%) 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model。其模

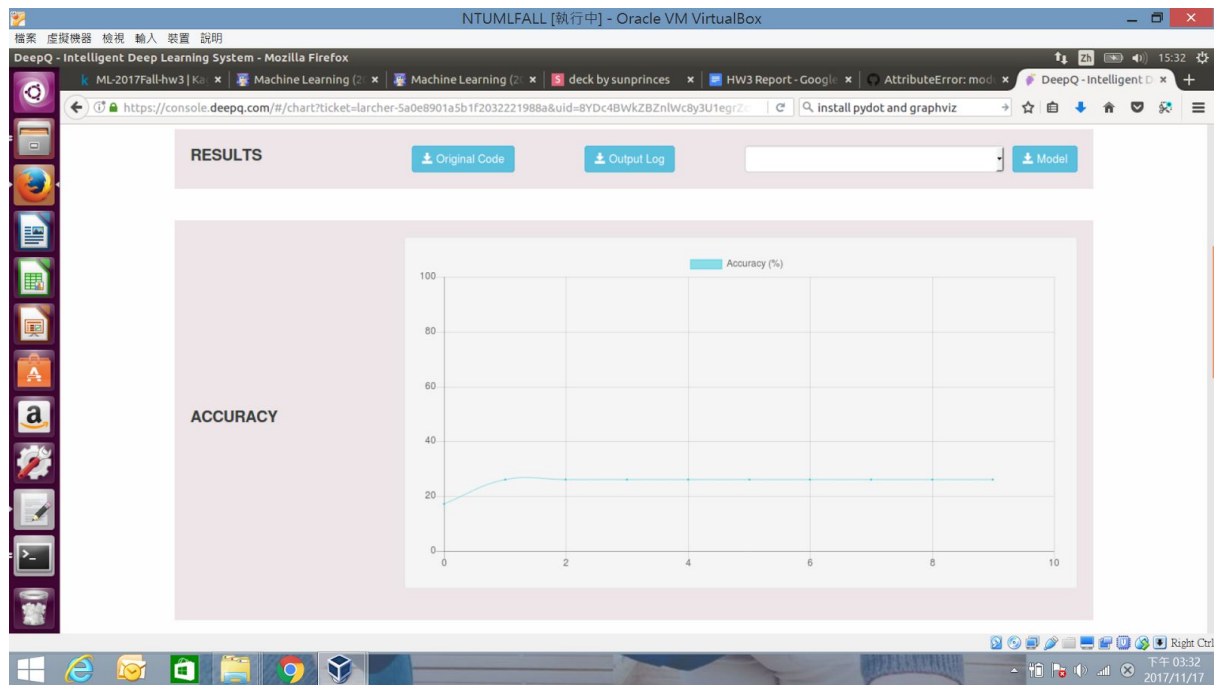
型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

(Collaborators: 無)



答：





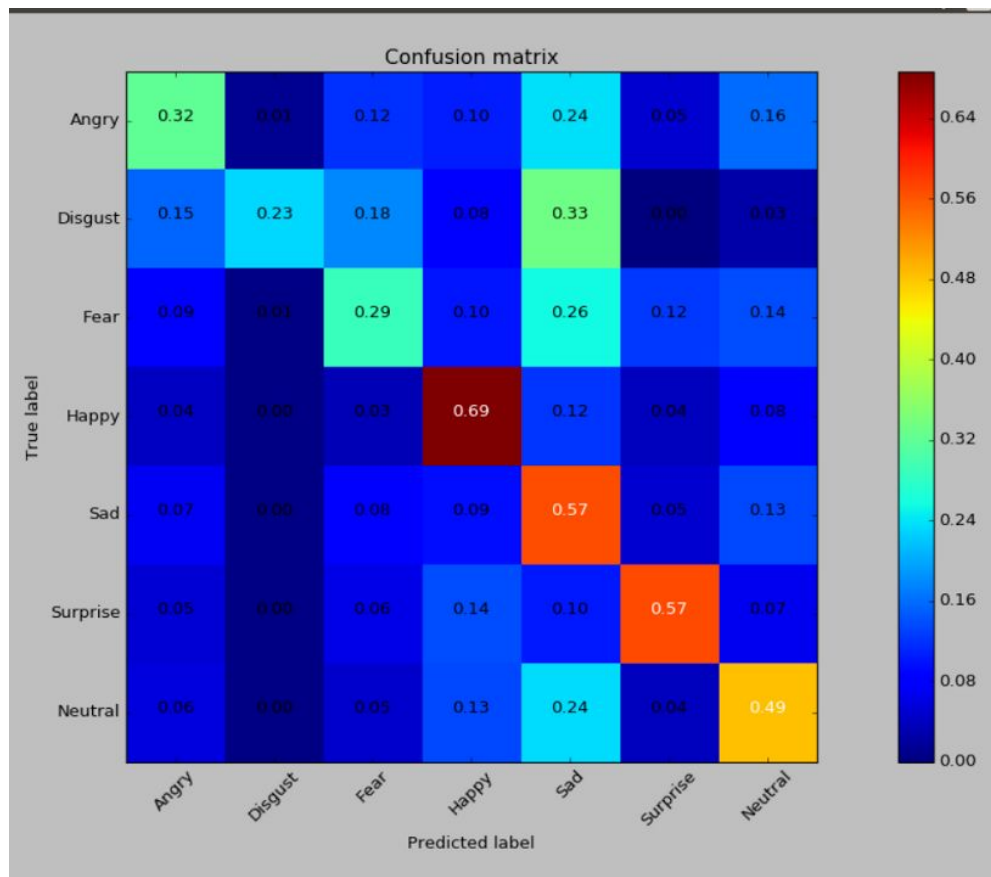
分別為training及validation accuracy

private=0.24881, public=0.24547

參數量約五百萬

比較下可以發現DNN模型的performance明顯較差，但有較快的收斂速度(但是是收斂值十分不理想)。

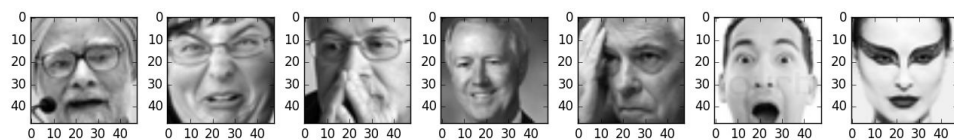
3. (1%) 觀察答錯的圖片中，哪些 class 彼此間容易用混？[繪出 confusion matrix 分析]  
(Collaborators: 無)

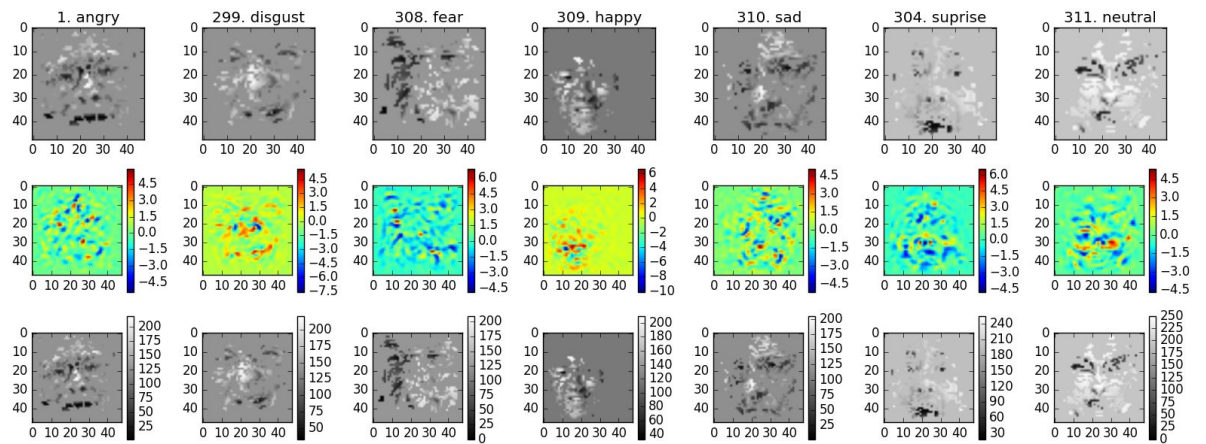


答：

最容易混淆的是所有跟sad.fear.angry以及disgust相關的部份，但前三者的原因應該是因為負面表情的區別不明顯，最後者則是因為資料量過少(在20000多筆資料中僅有1000筆左右)。此外happy的準確率相當高，應該是因為表情較明確的關係。

4. (1%) 從(1)(2)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: B04901136 張家銘)

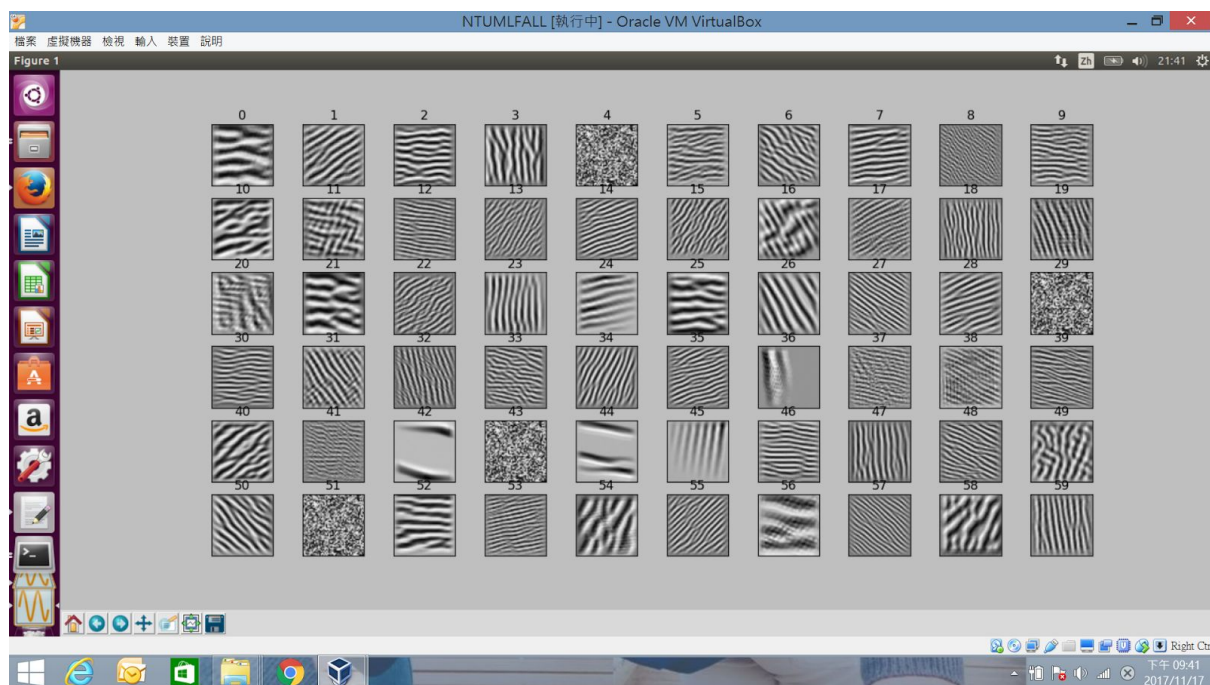




答：明顯focus在臉部相關的輪廓，特別是五官的分界線.形狀以及陰影，髮型部分則幾乎完全被忽略。但注意第一張圖中的眼鏡也被focus，應該是被誤認為五官的一部分

5. (1%) 承(1)(2)，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。  
(Collaborators: B04901136 張家銘, B04901056 張承洋)

答：



第一張圖為conv1\_2d的最敏感圖形，可看出主要感應各種角度以及粗細的線條分布。  
第二張是zeropadding2d\_2的输出，可看出經過兩層convolve之後，臉部的表情已經被許多filter所挑出，可以看出主要五官清楚的輪廓。