

學號：R06921077 系級：電機碩一 姓名：黃詩凱

1. (1%) 請說明你實作的 CNN model，其模型架構、訓練參數和準確率為何？

答：資料有先Normalization((x-x.mean)/x.std)

疊五層Convolution，每一層的filter依序為：64個(4,4)、128個(4,4)、256個(3,3)、512個(3,3)、512個(3,3)，其中Activation都是用ReLU，然後都有BatchNormalization，Dropout都是0.5

後面Flatten之後就接Fully Connected兩層，Neural數都是512個，Dropout都是0.5
optimizer用adam，validation set切0.1，batch_size設192，epoch設1000

最後準確率：training accuracy:0.9546, validation accuracy: 0.6956, public score:0.68041

2. (1%) 請嘗試 data normalization, data augmentation,說明實行方法並且說明對準確率有什麼樣的影響？

答：data normalization: 採用(x - x.mean()) / x.std() 和 BatchNormalization

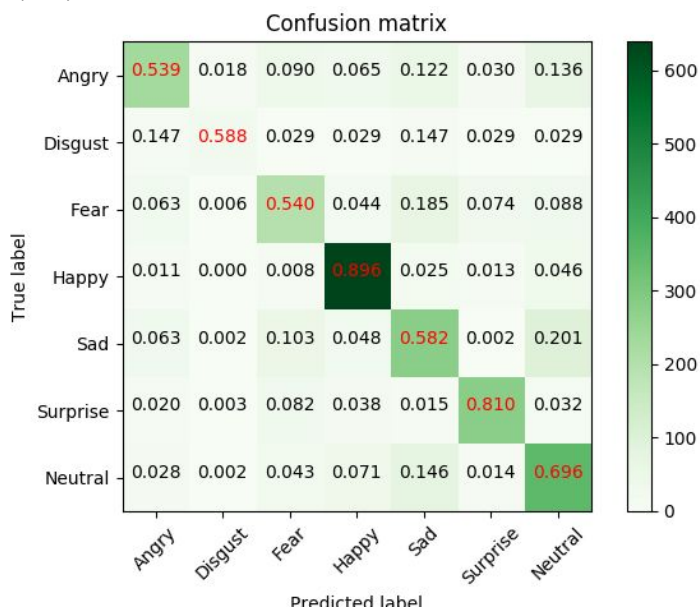
實作前：0.65617；實作後：0.67344 (public score) (其他比較的參數不變)

data augmentation: 只採用水平翻轉 (使用ImageDataGenerator)

實作前：0.63611；實作後：0.64028

我認為影響最大的是Normalization，至於Augmentation常常容易莫名的愈Train愈差，也有可能是參數調得不好，花的時間也比較久(要處理圖片)

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]

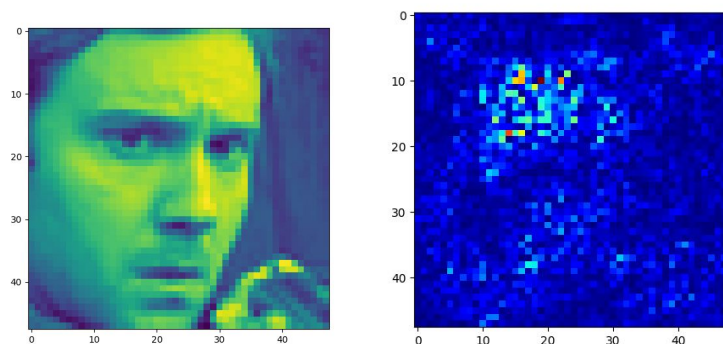


答：

容易混用的有：Sad/Neutral, Fear/Sad等

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？

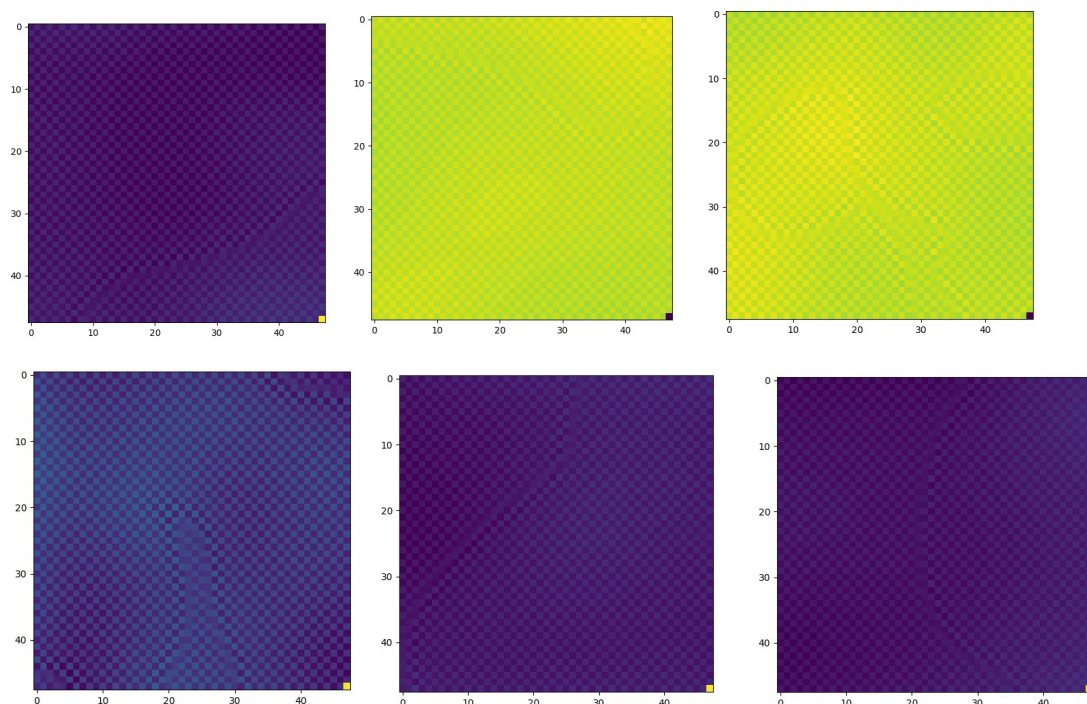


答：

看起來他focus的是他額頭的部分，有可能是眉頭深鎖能夠反映他目前的心情

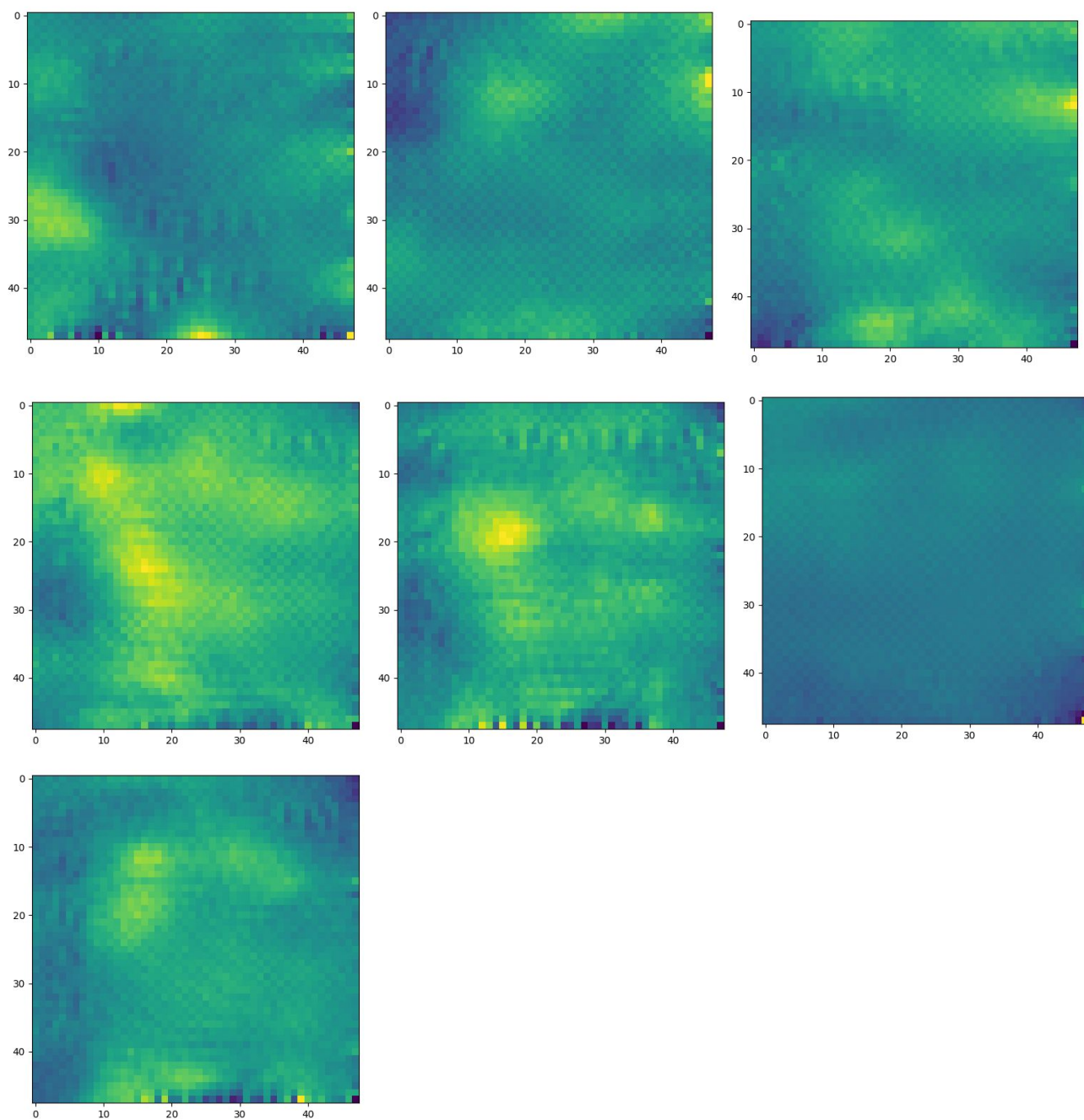
5. (1%) 承(4) 利用上課所提到的 gradient ascent 方法，觀察特定層的filter最容易被哪種圖片 activate與觀察filter的output。(Collaborators:)

答：Filter of Layer 1:



雖然不太明顯，不過圖中可以看到一些模糊的彎曲線條，應該是一些人臉的微小特徵

Filter of Last Layer:



最後一層，相較於第一層，很明顯就能看出每個心情對應到不同臉部範圍，雖然光看這幾張圖片很難直接分辨是哪種心情就是了