## Machine Learning HW3

學號:B03901109 系級: 電機四 姓名:陳緯哲

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:

在本次作業中,一共使用了 4 層 CNN 與 3 層 DNN,並在訓練一開始進行 data augmentation,以下分別說明各層參數:

- 1. 第一層 CNN:80 個 neuron,filter 的大小為 5\*5,再經由 LeakyReLU 進行 Activate,接著進行 Batch 的 normalization,經由 2\*2 大小的 maxpooling,並使用 Dropout 的方法。
- 2. 第二層 CNN:160 個 neuron,filter 的大小為 3\*3,再經由 LeakyReLU 進行 Activate,接著進行 Batch 的 normalization,經由 2\*2 大小的 maxpooling,並使用 Dropout 的方法。
- 3. 第三層 CNN: 500 個 neuron,filter 的大小為 5\*5,再經由 LeakyReLU 進行 Activate,接著進行 Batch 的 normalization,經由 2\*2 大小的 maxpooling,並使用 Dropout 的方法。
- 4. 第四層 CNN: 500 個 neuron,filter 的大小為 3\*3,再經由 LeakyReLU 進行 Activate,接著進行 Batch 的 normalization,經由 2\*2 大小的 maxpooling,並使用 Dropout 的方法,最後進行 Flatten。
- 5. 第五層 DNN: 500 個 neuron,經由 LeakyReLU 進行 Activate,接著進行 Batch 的 normalization,並使用 Dropout 的方法。
- 6. 第六層 DNN:500 個 neuron,經由 ReLU 進行 Activate,接著進行 Batch 的 normalization,並 使用 Dropout 的方法。
- 7. 第七層 DNN: 進行 softmax。

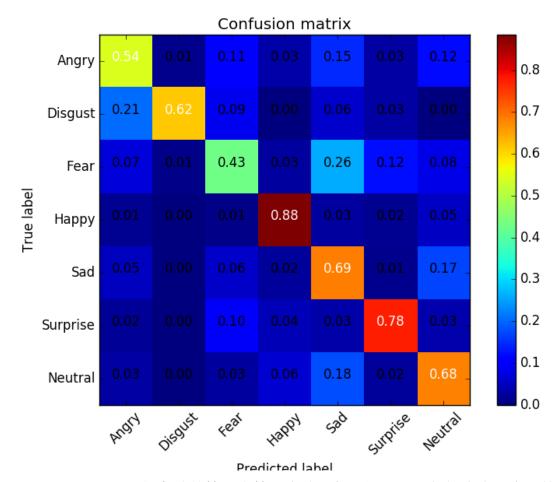
在訓練過程中也使用了 Learning rate reduction、Early stop、adam 等方式以提高準確率,最後的準確率為 0. 69601(public)、0.69713(private)。

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?(Collaborators: 王建翔)

答:

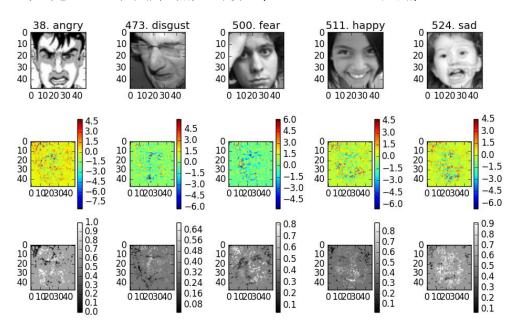
- a. Data normalization:將各張圖片中,每個位置相同位置的畫素進行 normalize  $(\frac{x-min}{max-min})$ ,有進行 normalization 的 Model valid data 的準確率為 0.69279,沒有進行 normalization 的 model 準確率為 0.69558,考量到 CNN 的隨機起始 weight 會對最後的準確率產生影響,我認為是 否進行 normalization 對於 model 的影響並不會太大。
- b. Data augmentation:利用 keras 內建的函式 ImageDataGenerator,將圖片進行以下操作:左右旋轉 20 度、水平位移、垂直位移、放大、水平翻轉、横移等,最終準確率可達 0.69558,若沒有進行 data augmentation,則準確率只有 0.65169,可以看出利用 data augmentation 可以有效避免 overfitting 的問題,改善準確率。

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: 王建翔)



(Fear, Sad)、(Disgust, Angry)是對角線外第一與第二高的比例,說明此兩個組合十分容易搞混,推測是因為這些表情有一定的相似度,並且在五官上的表現也十分相像,才會讓神經網路搞混。

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: 王建翔)

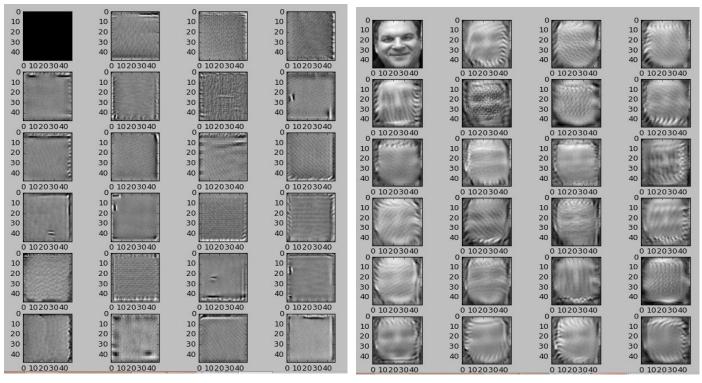


在進行分類時,神經網路會 focus 在人臉的五官與輪廓上,儘管有些人的五官較不清晰(如第二張圖),但仔細看還是能發現五官的位置。

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: 王建翔)

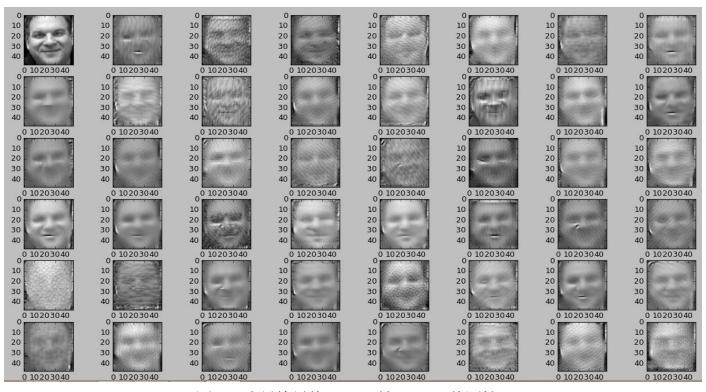
答

以下為第二層 CNN 的 filter 所看到的畫面,圖三中,可以看到人臉的部分,嘴巴微笑的部分被好幾個 filter 所 activate,可以對應到 confusion matrix 的部分,開心的正確率最高,而輸入各種 class 圖片進入 model 可以發現開心的 class 依然可以保留較多的圖片的特徵,因此開心的 class 最容易被 filter 所 activate;而圖二為較差的 model 所 activate 的圖片,臉部表情十分模糊,不如圖三清晰且輪廓明顯,因此可以看出圖三是有較正確的對圖片進行 activate。



圖一: filter 形狀

圖二:較差 Model 所 activate 的圖片



圖三:由最終最佳 Model 所 activate 的圖片