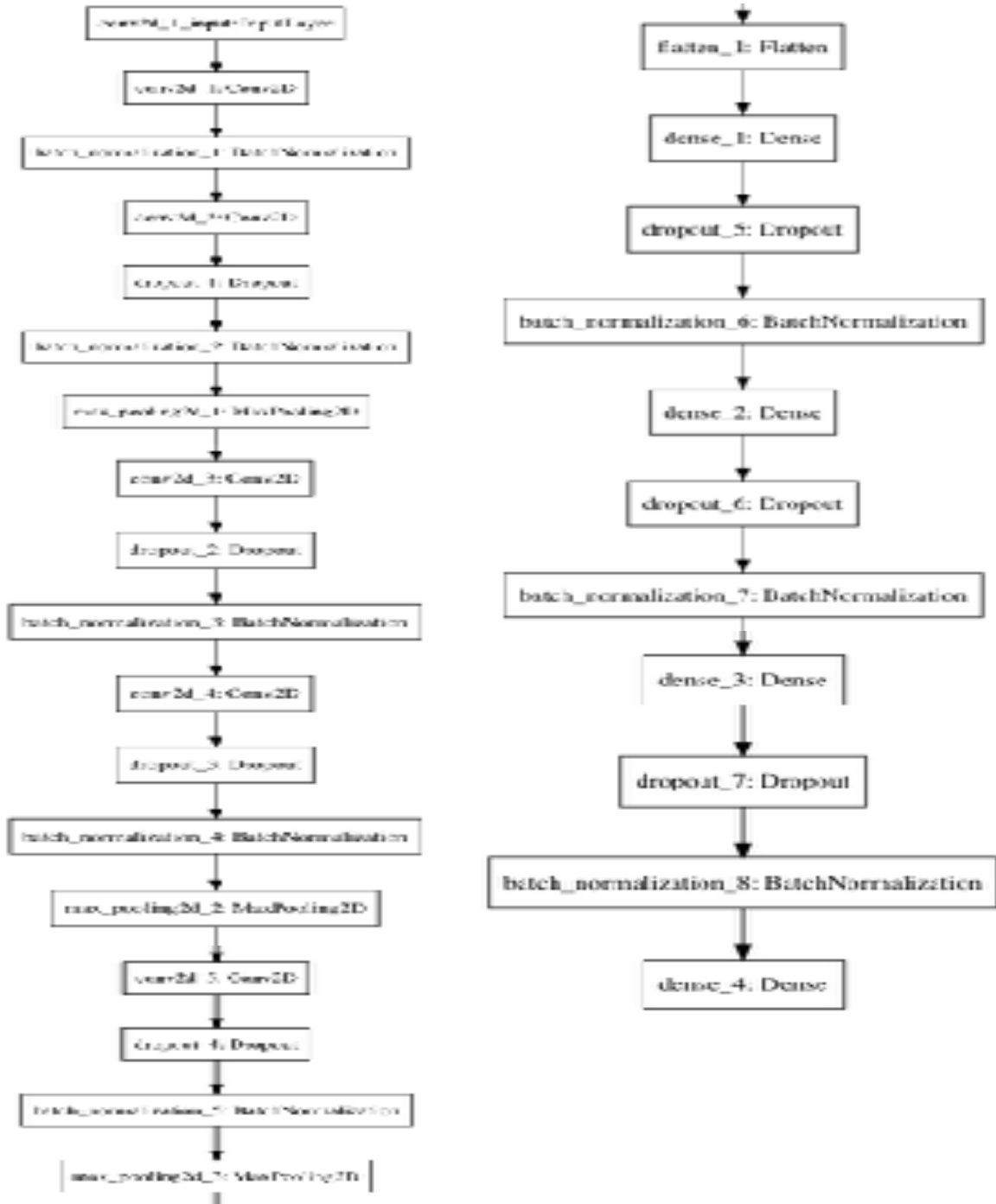


1. (1%) 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？

(Collaborators: No)

答：

(a)模型架構(塞不下圖因此切成兩半)：



(b)訓練過程：

loss function：categorical_crossentropy

optimizer：adadelta

batch size：100

epoch：80

因為臉孔圖形左右翻轉不影響意義，因此有把training data做翻轉處理，已增加兩倍的data量，並取validation set accuracy最高的model當輸出模型。

得到最高valid acc=0.68 kaggle分數 public : 0.64725 private : 0.65533

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: No)

答：

CNN使用參數量為2138823。DNN使用參數量為2337383。

(a)模型架構：



(b)訓練過程：

loss function : categorical_crossentropy

optimizer : adam

batch size : 100

epoch : 80

也有將data翻轉處理，並取validation set accuracy最高的model作為輸出模型。

最高valid acc=0.37 kaggle分數 public : 0.35859 private : 0.35831

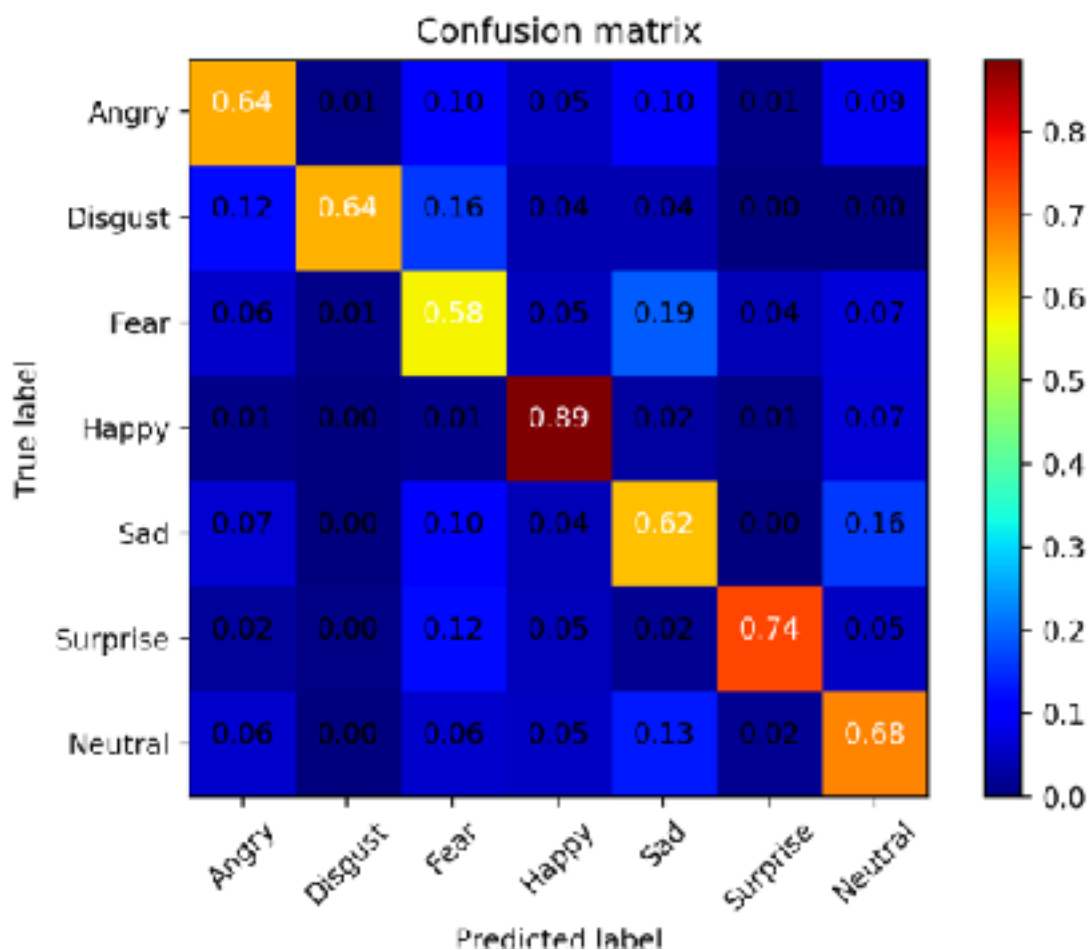
(c)觀察與比較

發現在沒有做convolution的DNN比起做CNN的結果差得很多，因為圖片中相鄰或附近兩個像素之間通常是很有關係的，不論是上下層或是左右的像素會有關聯。所以在結構上CNN比DNN多了convolution與pooling，也就是一次把附近的像素都丟進NN之中，不僅保留了圖片像素相鄰的關係特徵，還可以壓平前降低維度；而DNN則是把整張圖所有像素壓平一起訓練，因此抓不到圖片的特徵且需要大量參數。

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: No)

答：

(a)confusion matrix



(b)觀察

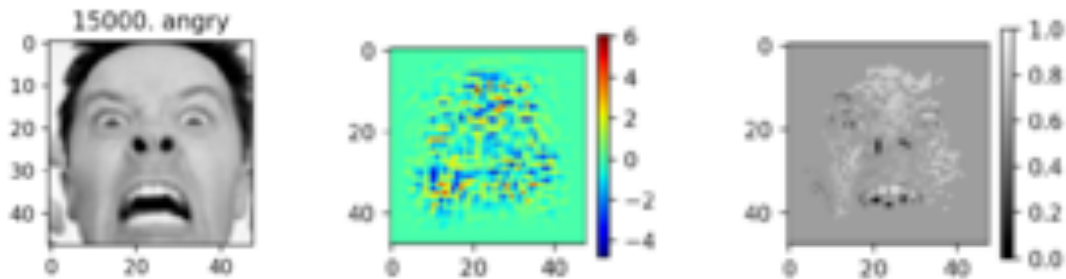
happy類別的正確率特高，可能是特徵與其他圖片較不相同且佔了data中較大的比例；容易把Fear誤認為其他的類別(Fear只有0.58正確率)，尤其是把Fear誤認成sad(0.19)，還有把Disgust誤認成Fear(0.16)，把Sad誤認成Neutral(0.16)而Neutral也常被誤認為Sad(0.13)，此兩個類別可能比較容易混淆。

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: b03902093張庭維)

答：

(a)saliency maps



(b)觀察

CNN會注重於圖片上的輪廓。

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: b03902093張庭維)

答：

觀察第一層5x5，64個filter。

