

Homework 3 Report - Image Sentiment Classification

學號：r06922086 系級：資工所 姓名：林凡煒

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？

(Collaborators：黃詩凱、葉日豐、陳代穎、林宇軒)

(1) 模型架構：參考 VGG16

註 1：BN 為 BatchNormalization

註 2：所有的 Conv2D Layers 與 Dense Layers 的 kernel_initializer 皆為 “glorot_normal”

註 3：loss function: cross entropy, optimizer: adam

CNN Structure:

Layer 1:

Padding(1)Conv2D(64, 5)BN()PReLU()

Padding(1)Conv2D(64, 3)BN()PReLU()

MaxPooling2D(2, strides=(2, 2))

Layer 2:

Padding(1)Conv2D(128, 3)BN()PReLU()

Padding(1)Conv2D(128, 3)BN()PReLU()

MaxPooling2D(2, strides=(2, 2))

Layer 3:

Padding(1)Conv2D(256, 3)BN()PReLU()

Padding(1)Conv2D(256, 3)BN()PReLU()

Padding(1)Conv2D(256, 3)BN()PReLU()

MaxPooling2D(2, strides=(2, 2))

Layer 4:

Padding(1)Conv2D(512, 3)BN()PReLU()

Padding(1)Conv2D(512, 3)BN()PReLU()

Padding(1)Conv2D(512, 3)BN()PReLU()

MaxPooling2D(2, strides=(2, 2))

Layer 5:

Padding(1)Conv2D(512, 3)BN()PReLU()

Padding(1)Conv2D(512, 3)BN()PReLU()

Padding(1)Conv2D(512, 3)BN()PReLU()

MaxPooling2D(2, strides=(2, 2))

DNN Structure:

Dense(4096)PReLU()Dropout(0.5)

Dense(4096)PReLU()Dropout(0.5)

Dense(7)SoftMax()

(2) 訓練參數：

Batch Size: 16

EPOCH: 500

Validation Data: 10 % Training Data

Model Check Point: monitor="val_acc", save_best_model

Early Stopping: monitor='loss', patience=30

(3) 準確率：

Validation data accuracy: 0.667

Training data accuracy: 0.8901

Kaggle's private/public score: 0.67679 / 0.68152

2. 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：黃詩凱、陳代穎)

(1) Data Normalization:

作法： $(\text{train_x} - \text{mean}(\text{train_x} + \text{test_x})) / \text{std}(\text{train_x} + \text{test_x})$ ，此外除了 normalization，還有作 / 255.0 的前處理。

實作前：after 40 epochs, the Kaggle's private/public score: 0.54750/0.54778

實作後：after 40 epochs, the Kaggle's private/public score: 0.62496/0.62273

明顯作了 data normalization 之後，不論是在 private or public data 表現上，都是比較好的。

(2) Data Augmentation:

作法：

1. 利用原圖產生左右翻轉後的圖片。
2. 利用原圖產生上下左右各位移 0.1 倍的圖片。
3. 利用原圖產生旋轉 30 度的圖片。
4. 利用原圖產生放大 0.2 倍的圖片。

實作前：after 40 epochs, the Kaggle's private/public score: 0.54165/0.53970

實作後：after 40 epochs, the Kaggle's private/public score: 0.62496/0.62273

一樣是作了 data augmentation 後的結果比較好。

3. 觀察答錯的圖片中，哪些 class 彼此間容易用混？

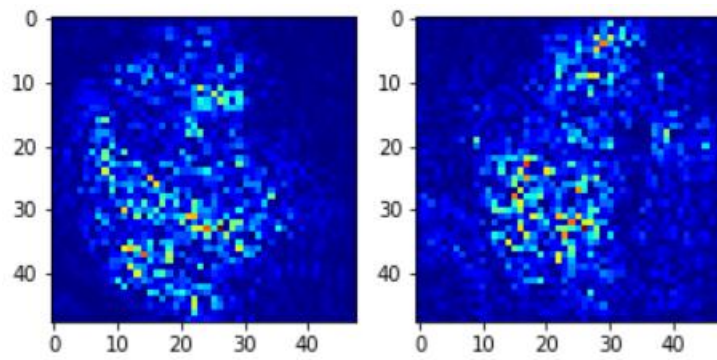
(Collaborators：陳代穎)

predict	生氣	厭惡	恐懼	高興	難過	驚訝	中立
correct							
生氣	187	3	35	23	110	10	55
厭惡	7	23	2	2	5	0	0
恐懼	32	0	214	16	98	21	50
高興	7	0	15	633	16	8	32
難過	15	2	58	16	298	1	76
驚訝	11	0	30	17	4	227	18
中立	8	1	16	51	82	1	334

由上表可以看出，生氣的情緒很容易被誤判成難過。

4. 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: 黃詩凱)



可以由上圖對照組發現，saliency maps 主要是觀察臉的下半部分：諸如鼻子與嘴巴的地方。

5. 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: 黃詩凱)

Prediction Filter Input:



Prediction Filter Output: `[array([[0., 0., 0., 0., 0., 1., 0.]])]`

結果告訴我們，驚訝這個情緒是最容易被 activate 的。