# Homework3 Report

Professor Pei-Yuan Wu
EE5184 - Machine Learning

姓名： 陳俊翰
學號：R07522814

Note:1~3 題建議不要超過三頁

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？

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv2d_1 (Conv2D) | (None, 48, 48, 64) | 640 |
| conv2d_2 (Conv2D) | (None, 48, 48, 64) | 36928 |
| max_pooling2d_1 (MaxPooling2 | (None, 24, 24, 64) | 0 |
| dropout_1 (Dropout) | (None, 24, 24, 64) | 0 |
| conv2d_3 (Conv2D) | (None, 24, 24, 128) | 73856 |
| conv2d_4 (Conv2D) | (None, 24, 24, 128) | 147584 |
| max_pooling2d_2 (MaxPooling2 | (None, 12, 12, 128) | 0 |
| dropout_2 (Dropout) | (None, 12, 12, 128) | 0 |
| conv2d_5 (Conv2D) | (None, 12, 12, 256) | 295168 |
| conv2d_6 (Conv2D) | (None, 12, 12, 256) | 590080 |
| conv2d_7 (Conv2D) | (None, 12, 12, 256) | 590080 |
| max_pooling2d_3 (MaxPooling2 | (None, 6, 6, 256) | 0 |
| dropout_3 (Dropout) | (None, 6, 6, 256) | 0 |
| flatten_1 (Flatten) | (None, 9216) | 0 |
| dense_1 (Dense) | (None, 512) | 4719104 |
| dropout_4 (Dropout) | (None, 512) | 0 |
| dense_2 (Dense) | (None, 512) | 262656 |
| dropout_5 (Dropout) | (None, 512) | 0 |
| dense_3 (Dense) | (None, 7) | 3591 |
| activation_1 (Activation) | (None, 7) | 0 |

Total params: 6,719,687
Trainable params: 6,719,687

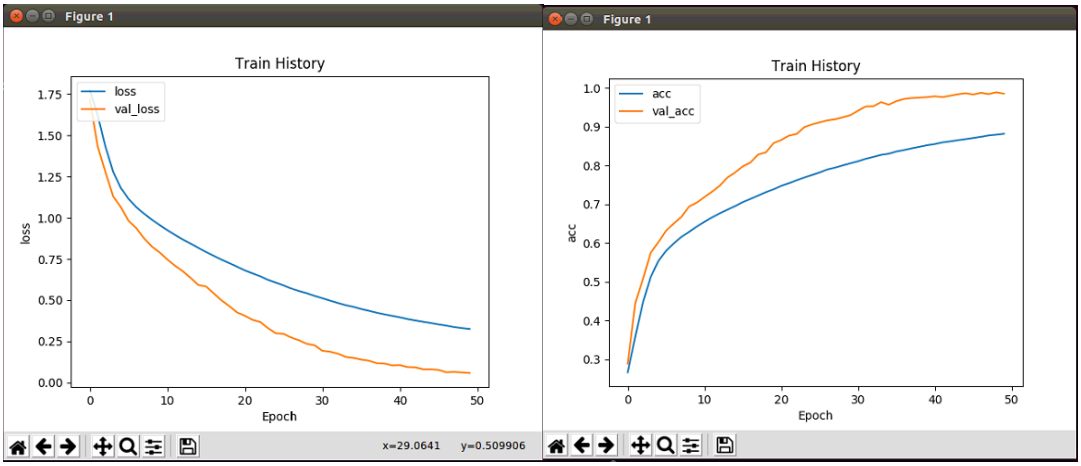
Figure 1 CNN model

模型架構 :

參考 VGG-16 架構，先兩層深度 64 的 convolution 層(使用 relu 激活)搭配 1 層 max pooling 層與 dropout 層避免 overfitting，接下來用類似的架構，架了兩層深度 128 的 convolution 層(使用 relu 激活)搭配 1 層 max pooling 層與 dropout 層，與層深度 128 的 convolution 層(使用 relu 激活)搭配 1 層 max pooling 層與 dropout 層，再來架 flatten 層讓資料扁平化，進入 Neural Network，架了 2 層 512 的全連接層與 dropout 層，最後 output layer 輸出 7 個分類，總參數:6719687 個。

訓練過程 :

首先切 training data 前 4000 筆資料為 validation data，訓練過程採用 stochastic gradient descent ( learning rate=0.005, decay=0.00001, momentum=0.9 )，loss 採用 cross entropy，衡量標準為 accuracy，並加上 callback，隨時儲存訓練過程中表現好的 model，下圖為訓練過程的 loss 與 accuracy 變化。




準確率 : 0.68347 (public leader board)

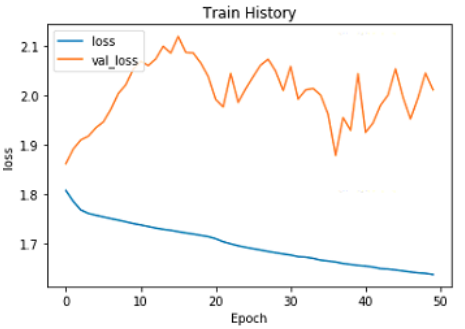
2. (1%) 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model，其模型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

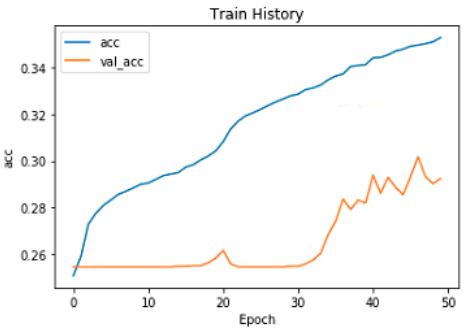
模型架構 :

直接經過 Flatten 層，再加 4 個深度 512 的層，3 個深度 1024 的層，4 個深度 2048 的層，最後 output layer 輸出 7 個分類，總參數:6736910 個。

訓練過程 :

同上題




Validation loss 很不穩定會震盪。Accuracy 只有稍微提升一點，效果很差

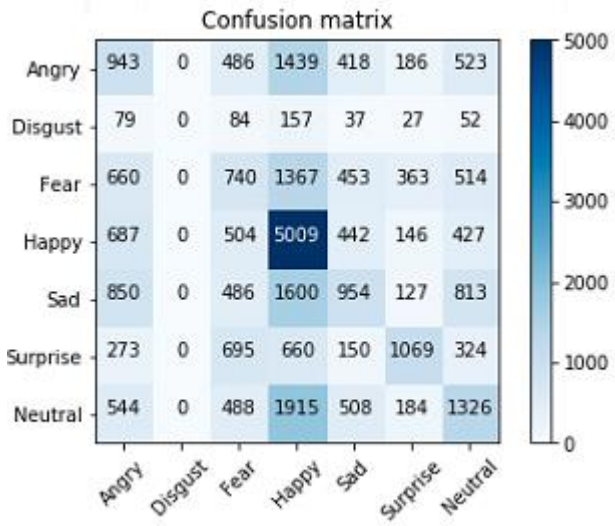
準確率 :

0.27807

比較 :

DNN 只有 loss 再下降對 accuracy 提升很有限，performance 非常不好。

3. (1%) 觀察答錯的圖片中，哪些 class 彼此間容易用混？ 並說明你觀察到了什麼？ [繪出 confusion matrix 分析]




Disgust 資料少因此無法正確識別，Happy 數量最多且和 Angry, Fear, Sad, Neutral 常常混淆，每個種類的資料數量差很多，需要特別處理。

**4.**

**a. how many parameters are there in each layer?**

$$\#\ of\ parameters\ =\ filters\ *\ (kernelsize * channels + 1)$$

**Layer A:**
**filters = 6**
**kernel size = (2,2)**
**channels = 5, because input shape = (8,8,5)**
**parameters = 6*(2*2*5+1) = 126**

**Layer B:**
**filters = 4**
**kernel size = (2,2)**
**channels = 6 because output shape of layer A = (3,3,6)**
**parameters = 4*(2*2*6+1) = 100**

**b. How many multiplications/additions are needed for a forward pass?**

**number of multiplications = 111 filters * steps * (kernel size*channels)**
**number of additions = filters * steps * (kernel size*channels-1)**

**Layer A:**
**filters = 6**
**steps = (3,3) because ((input size-kernel size)/strides)+1=3.0**
**kernelsize = (2,2)**
**channels = 5**
**multiplications = 6*(3*3)*(2*2*5) = 1080**
**additions = 6*(3*3)*(2*2*5-1) = 1026**

**Layer B:**
**filters = 4**
**steps = (1,1) because ((input size-kernel size)/strides)+1=1.5,**

**by padding=valid, we drop incomplete stride and get steps=1**
**kernelsize = (2,2)**
**channels = 5**
**multiplications = 4*(1*1)*(2*2*6) = 96**
**additions = 4*(1*1)*(2*2*6-1) = 92**

**c. What is the time complexity of convolutional neural networks?**

**time complexity = number of multiplications + number of additions**

**Layer i:**
**given**
**kernel size =** $(k_i, k_i)$
**channelsize =** $c_i$
**input shape =** $(n_i, n_i) \ = \ (n, n)$
**Then**
**steps = (n,n) since steps of layer i = input shape of layer i+1**
**filters =** $c_{i+1}$ **since filters of layer i = channelsize of layer i+1**
**Then**
**multiplications =** $c_{i+1} * (n * n) * (k_i * k_i * c_i)$
**additions =** $c_{i+1} * (n * n) * (k_i * k_i * c_i - 1)$
**time complexity =** $c_{i+1} * (n * n) * (2 * k_i * k_i * c_i - 1) \ = \ O(c_i c_{i+1} k_i^2 n^2)$
**Finally,**
**CNN:**
$number\ of\ layers \ = \ l$
**time complexity =** $\sum_{i=1}^{l} O(c_i c_{i+1} k_i^2 n^2)$ **, where** $c_{l+1}$ **is the desired output channelsize.**

## 5.

PCA practice:Problem statement : Given 10 samples in 3D space.(1,2,3),(4,8,5),(3,12,9),(1,8,5),(5,14
,(3,8,1),(11,5,6),(10,11,7)
(1) What are the principal axes?
(2) Compute the principal components for each sample.
(3) Reconstruction error if reduced to 2D.
(Calculate the L2-norm)

(1)

$$\begin{bmatrix} & & \\ u1 & u2 & u3 \\ & & \end{bmatrix} = \begin{bmatrix} 0.5035 & 0.7791 & 0.3735 \\ 0.7370 & -0.6129 & 0.2849 \\ 0.4509 & 0.1318 & -0.8828 \end{bmatrix}$$

(2)

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 131.32 & 0 & 0 \\ 0 & 11.76 & 0 \\ 0 & 0 & 5.52 \end{bmatrix}$$

(3)

$$\text{投影誤差(project error)}=\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}-x_{approx}^{(i)}\|^2 = 5.5173$$

$$\text{總變差(total variation)} =\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}\|^2 = 148.6000$$

$$\text{誤差率(error ratio)} =\frac{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}-x_{approx}^{(i)}\|^2}{\frac{1}{m}\sum_{i=1}^{m}\|x^{(i)}\|^2} = 0.0371$$

Reference : https://zhuanlan.zhihu.com/p/32953274?fbclid=IwAR0R86kHjM66-UaUWLvarBUAuS3MXDLqnHtS4OoK-BatRB-FWLYEdvaQS6Q