

# Homework3 Report

Professor Pei-Yuan Wu  
EE5184 - Machine Learning

姓名：楊宗賢  
學號：B06901031

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？

我訓練了三個 CNN model，命名為 model\_3、model\_5、model\_6，細節如下：

model_3	model_5	model_6
Conv, 3x3, 120	Conv, 3x3, 160	Conv, 3x3, 180
Conv, 3x3, 120	Conv, 3x3, 160	Conv, 3x3, 180
Maxpooling, 2	Maxpooling, 2	Maxpooling, 2
Dropout: 0.1	Dropout: 0.2	Dropout: 0.2
Conv, 2x2, 240	Conv, 2x2, 320	Conv, 2x2, 360
Conv, 2x2, 240	Conv, 2x2, 320	Conv, 2x2, 360
Maxpooling, 2	Maxpooling, 2	Maxpooling, 2
Dropout: 0.1	Dropout: 0.2	Dropout: 0.2
Conv, 3x3, 480	Conv, 2x2, 480	Conv, 2x2, 540
Conv, 3x3, 480	Conv, 2x2, 480	Conv, 2x2, 540
Maxpooling, 2	Conv, 2x2, 480	Conv, 2x2, 540
Dropout: 0.1	Maxpooling, 2	Maxpooling, 2
Dense: 720	Dropout: 0.3	Dropout: 0.2
Dropout: 0.3	Conv, 2x2, 640	Conv, 2x2, 720
Dense: 720	Conv, 2x2, 640	Conv, 2x2, 720
Dropout: 0.3	Conv, 2x2, 640	Conv, 2x2, 720
Dense: 7	Maxpooling, 2	Maxpooling, 2
	Dropout: 0.3	Dropout: 0.3
	Dense: 1280	Dense: 2560
	Dropout: 0.3	Dropout: 0.3
	Dense: 960	Dense: 960
	Dropout: 0.3	Dropout: 0.3
	Dense: 7	Dense: 7
(Conv padding:none)	(Conv padding:same)	(Conv padding:same)
訓練方式		
80 epochs	65 epochs	100 epochs
batch size: 100	batch size: 150	batch size: 144
準確率		
training: 0.9607	training: 0.9386	training: 0.9639
validation: 0.5800	validation: 0.6169	validation: 0.6263
Kaggle private: 0.60490	Kaggle private: 0.61493	Kaggle private: 0.62412

三個 model 以相同的 weight 合取 ensemble 後，  
validation accuracy=0.7950, Kaggle private score=0.67038。

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？

我修改了 model\_6，作成了約有 2900 萬個參數的 DNN model：

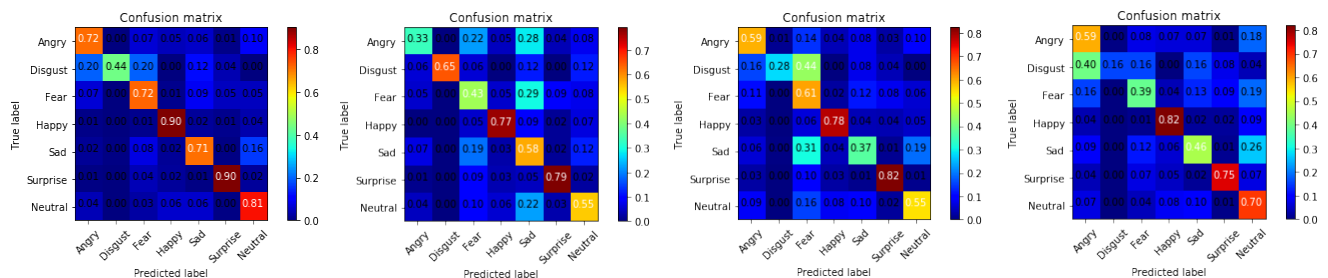
模型架構：[input=2304]→880→1760→2640→3520→2560→960→7

訓練過程：同 model\_6，100 epochs，batch size=144

準確率：training=0.3959 validation=0.3404 Kaggle private=0.32738

此 DNN 的 Confusion Matrix 顯示，Surprise 認出 0.76，Happy 認出 0.53，其餘的 class 都認不出來，連 training data 都無法 fitting。上題中，training data 至少都有 0.9 以上，顯見卷積層的有無對圖像辨識的成敗有多重要。

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？



最左為 ensembled model，其餘依序為三個 model 各自的 Confusion Matrix。

Disgust 的 class 最容易與其他 class 混淆，各有 0.2 的 Disgust 被分成了 Angry 和 Fear。另外 Fear 被認成 Sad、Sad 或 Angry 被認成 Neutral 的情形也很顯著，反倒是 Happy 和 Surprise 不太會被認錯。

按照一般社會常識來推論，正面情緒的表情會有很顯著的特徵，而負面的情緒(如 Angry、Fear、Sad 等)可能都同樣表現出一個「激動」的表情，自然會互相混淆，只能到 0.7 的 Accuracy。

training set 中的 Disgust class 佔比實在太小了，層數高的 model\_5 與 model\_6 都幾乎完全認不出來，因此我將層數雖少卻能將 Disgust 認到 0.6 的 model\_3 加入 ensemble，平衡了另兩個 model 在 Disgust class 上的重大缺點。

4. (1.5%, each 0.5%) CNN time/space complexity:

5. (1.5%, each 0.5%) PCA practice: