學號：B05901025 系級： 電機三 姓名：王鈺能

1. (2%) 請說明你實作的 CNN model，其模型架構、訓練參數和準確率為何？並請用與上述 CNN 接近的參數量，實做簡單的 DNN model，同時也說明其模型架構、訓練參數和準確率為何？並說明你觀察到了什麼？
(Collaborators:)

答：

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 46, 46, 128)	1280
batch_normalization_7 (Batch	(None, 46, 46, 128)	512
conv2d_3 (Conv2D)	(None, 44, 44, 128)	147584
batch_normalization_8 (Batch	(None, 44, 44, 128)	512
max_pooling2d_1 (MaxPooling2	(None, 22, 22, 128)	0
conv2d_4 (Conv2D)	(None, 20, 20, 256)	295168
batch_normalization_9 (Batch	(None, 20, 20, 256)	1024
conv2d_5 (Conv2D)	(None, 18, 18, 256)	590080
batch_normalization_10 (Batc	(None, 18, 18, 256)	1024
conv2d_6 (Conv2D)	(None, 16, 16, 256)	590080
batch_normalization_11 (Batc	(None, 16, 16, 256)	1024
max_pooling2d_2 (MaxPooling2	(None, 8, 8, 256)	0
dropout_6 (Dropout)	(None, 8, 8, 256)	0
conv2d_7 (Conv2D)	(None, 6, 6, 512)	1180160
batch_normalization_12 (Batc	(None, 6, 6, 512)	2048
conv2d_8 (Conv2D)	(None, 4, 4, 512)	2359808
batch_normalization_13 (Batc	(None, 4, 4, 512)	2048
max_pooling2d_3 (MaxPooling2	(None, 2, 2, 512)	0
dropout_7 (Dropout)	(None, 2, 2, 512)	0
flatten_2 (Flatten)	(None, 2048)	0
dense_7 (Dense)	(None, 1024)	2098176
batch_normalization_14 (Batc	(None, 1024)	4096
dropout_8 (Dropout)	(None, 1024)	0
dense_8 (Dense)	(None, 512)	524800
batch_normalization_15 (Batc	(None, 512)	2048
dropout_9 (Dropout)	(None, 512)	0
dense_9 (Dense)	(None, 7)	3591

Total params: 7,805,063

Layer (type)	Output Shape	Param #
dense_131 (Dense)	(None, 1024)	2360320
batch_normalization_137 (Bat	(None, 1024)	4096
dense_132 (Dense)	(None, 1024)	1049600
batch_normalization_138 (Bat	(None, 1024)	4096
dropout_102 (Dropout)	(None, 1024)	0
dense_133 (Dense)	(None, 1024)	1049600
batch_normalization_139 (Bat	(None, 1024)	4096
dense_134 (Dense)	(None, 1024)	1049600
batch_normalization_140 (Bat	(None, 1024)	4096
dropout_103 (Dropout)	(None, 1024)	0
dense_135 (Dense)	(None, 512)	524800
batch_normalization_141 (Bat	(None, 512)	2048
dense_136 (Dense)	(None, 512)	262656
batch_normalization_142 (Bat	(None, 512)	2048
dense_137 (Dense)	(None, 512)	262656
batch_normalization_143 (Bat	(None, 512)	2048
dropout_104 (Dropout)	(None, 512)	0
dense_138 (Dense)	(None, 256)	131328
batch_normalization_144 (Bat	(None, 256)	1024
dropout_105 (Dropout)	(None, 256)	0
dense_139 (Dense)	(None, 256)	65792
batch_normalization_145 (Bat	(None, 256)	1024
dropout_106 (Dropout)	(None, 256)	0
dense_140 (Dense)	(None, 1024)	263168
batch_normalization_146 (Bat	(None, 1024)	4096
dropout_107 (Dropout)	(None, 1024)	0
dense_141 (Dense)	(None, 7)	7175

Total params: 7,055,367

CNN 架構如左圖所示，主要為 8 層 CNN 和 3 層 FC，每層後都有 batch normalize，並且有 3 次 max pooling 和一些 dropout。我使用 adam、learning rate=0.0005，batch size = 32 訓練 200 個 epoch。單一 model train_acc = 0.93、val_acc = 0.68，用 8 個 models 做 bagging 後 val_acc = 0.74，最後 test_acc = 0.718。

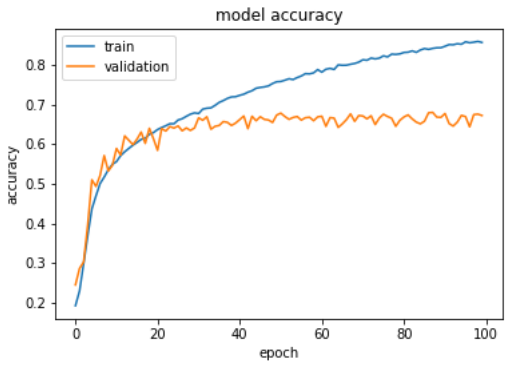
DNN 架構如右圖所示，由 11 層 256~1024 個 neuron 的 FC 構成，每層後都有 batch normalize 和一些 dropout。我使用 adam、learning rate=0.0005，batch size = 32 訓練 100

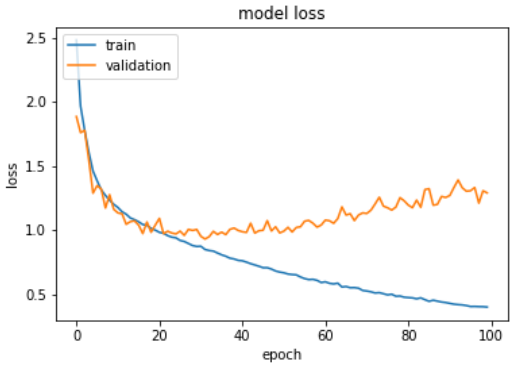
個 epoch。單一 model train_acc = 0.70、val_acc = 0.42。

在相近的參數量下，CNN 表現的明顯比單純 DNN 好，如下題圖示，DNN 不僅 training set 表現較差，testing set 表現也很快就卡住了，相較之下 CNN 兩者都比 DNN 好的多。

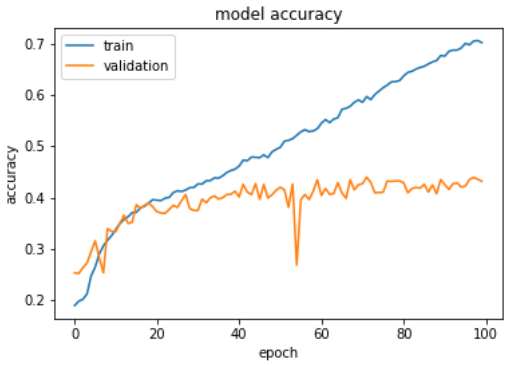
2. (1%) 承上題，請分別畫出這兩個 model 的訓練過程 (i.e., loss/accuracy v.s. epoch) (Collaborators:)

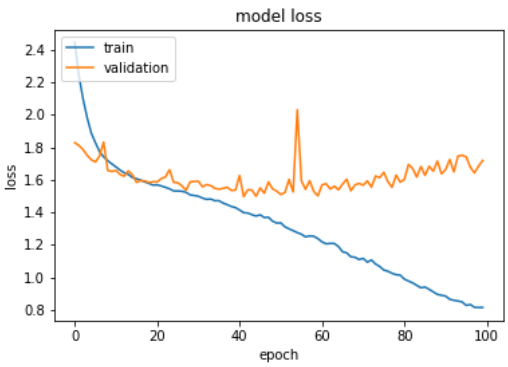
答：





(CNN)

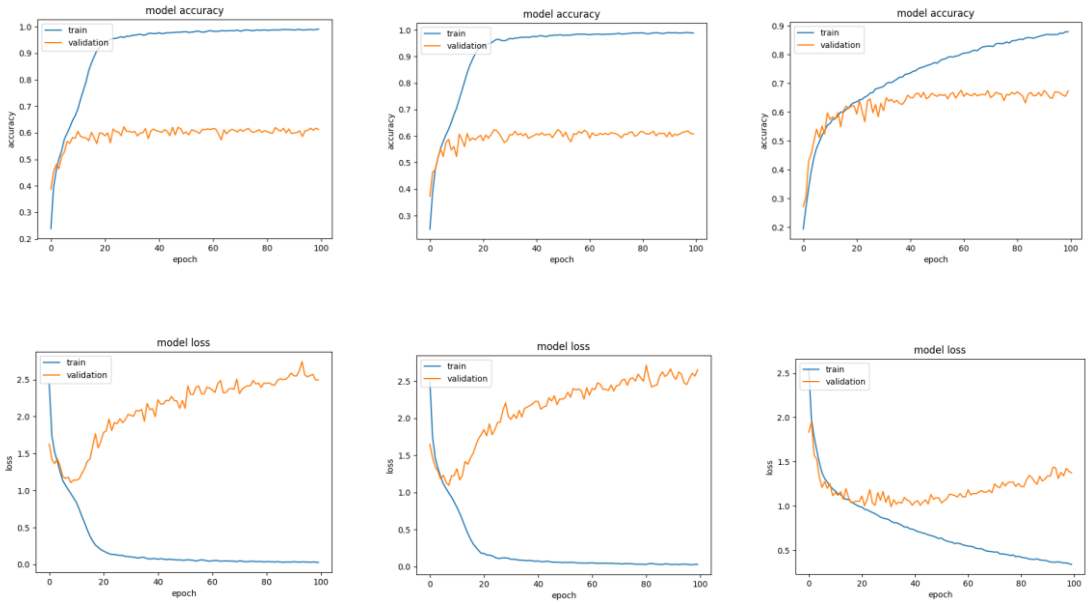




(DNN)

3. (1%) 請嘗試 data normalization, data augmentation,說明實作方法並且說明實行前後對準確率有什麼樣的影響？
(Collaborators:)

答：

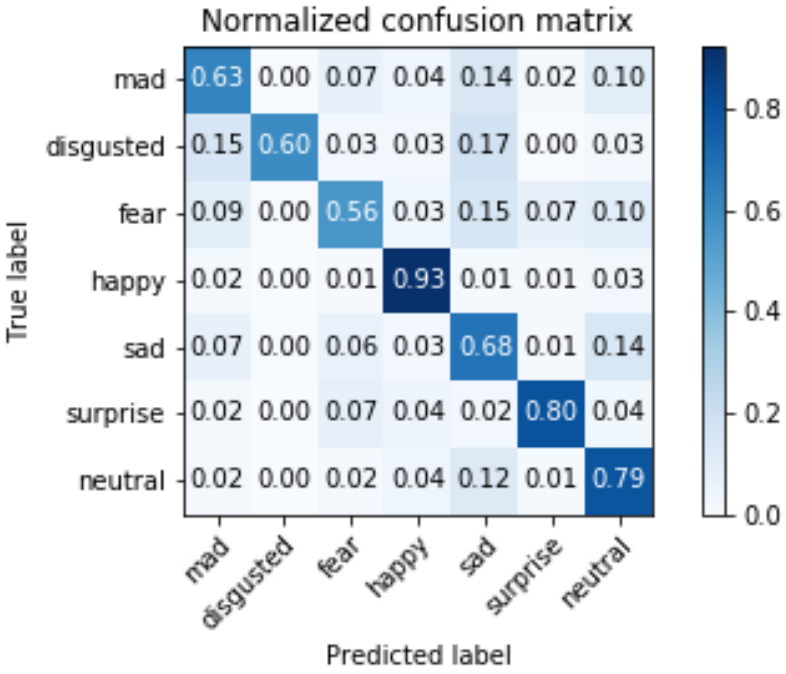


(original data) (normalize only) (augmentation only)

我 normalize 方法為將所有 pixel 除以 256，augmentation 為使用 Keras 套件進行 rotate、left/right shift、horizontal flip 等等，從上圖看起來的結論是：normalize 對 training 影響不大，而 augmentation 會讓 training 成長變慢(data 變多比較難 fit 的合理作用)，而 val_acc 有明顯的上升。

4. (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:)

答：



mad、disgusted、fear 是辨識率最差的三個，並且最常錯誤辨識成 sad，這是合理的結果，我認為現實中相較於前三者常常會有誇張明顯的臉部特徵變化，sad 算是臉部表現比較不明顯的負面情緒，所以抓 sad 的部分可能就比較敏感，任何臉部肌肉的牽動都會刺激到它，自然就容易在其他負面表情的時後有反應導致錯誤辨識。