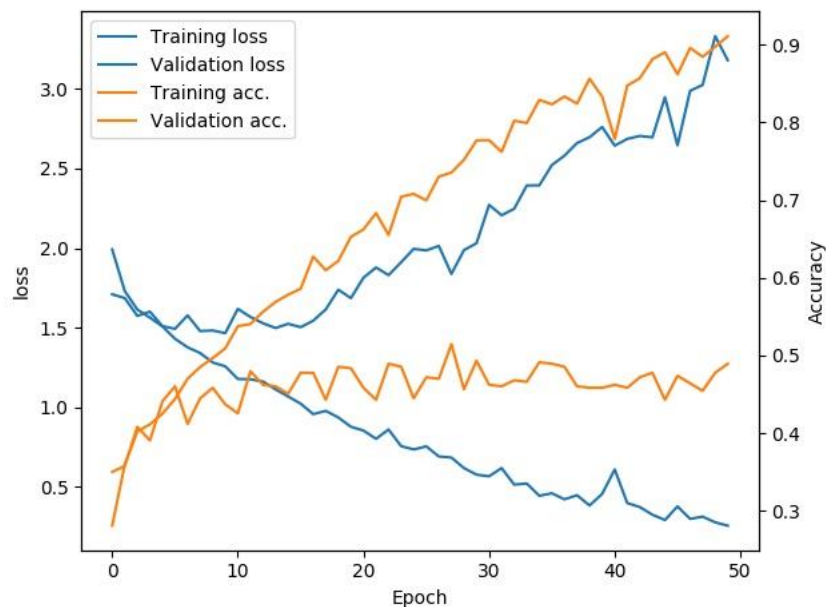


1.

- a. 我直接使用 Keras.applications 所提供的 pre-trained resnet50, 並且將 fully connected layer 拿掉(include\_top=False), 再將 Resnet 過完 average pooling 的 2048 維 feature 過兩層 fully connected, 維度分別為 512 及 128, 最後 output 出過完 softmax 的 11 維 label, 並沒有另外 fine-tune 前面的 conv.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 2048)	0
dense_1 (Dense)	(None, 512)	1049088
dense_2 (Dense)	(None, 128)	65664
dropout_1 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 11)	1419
Total params: 1,132,683		
Trainable params: 1,132,683		
Non-trainable params: 0		

- b. Accuracy 為 0.5048, 下圖為 training 過程中的 training 及 validation accuracy 及 loss.

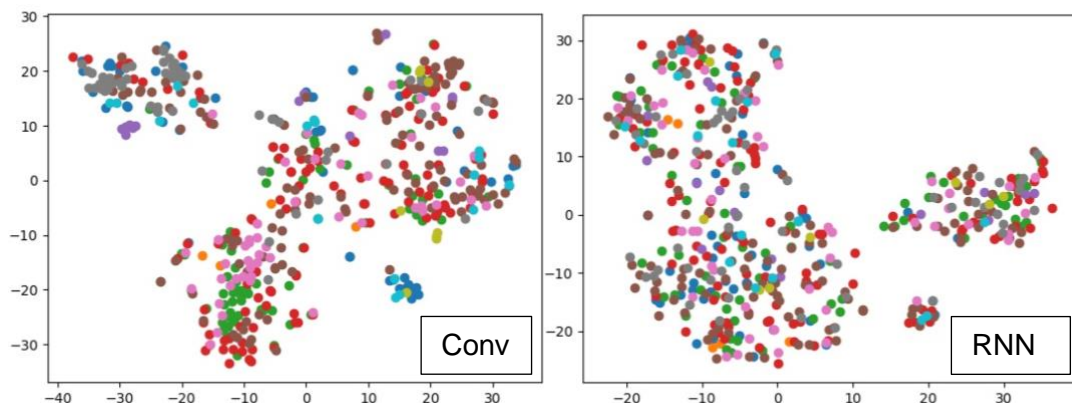


2.

- a. Model 如下, 而對於 video 長度(或說 #frames)不一, 直接使用 Keras 所提供的 pad\_sequences, 將一個 list 不同長度的 numpy array padd 成一個  $N * \text{max\_seq\_len} * 2048$  的 array, 就可以丟進 LSTM 進行訓練.

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	(None, 267, 2048)	0
masking_2 (Masking)	(None, 267, 2048)	0
lstm_2 (LSTM)	(None, 256)	2360320
dense_5 (Dense)	(None, 11)	2827
Total params: 2,363,147		
Trainable params: 2,363,147		
Non-trainable params: 0		

- b. 左圖為 validation data 的 convolution feature (pre-trained ResNet) 經過 TSNE 所畫出來的, 而右圖則為 RNN feature 所跑出來的結果, 可以看到雖然說 RNN 的 accuracy 為 0.52 左右, 稍微比 Conv 的結果好一些些, 但不同 label 之間依然沒有辦法區分的開來, 可能肇因於 RNN 對於 data 的 representativeness 並沒有太顯著的進步, 這部分也可以從 accuracy 僅進步 2% 來佐證.



3.

- a. 由於每個 full video 的 length 都太長了, 若像 part 2 一樣整個 video 下去 train, 不僅 model 較難 capture 住這麼長的資訊而且 data 只會有 5 筆, 因此我將每個 video 從開頭每 20 個 frame 切成一個 sequence, 並每次位移 10 個 frame (兩個 sequences 會有 10 個 frame 重複), 如此得到的 sequences 再下去 train.

這裡的 RNN Model 跟 part 2 一樣, 差別僅在於這邊必須每個 timestamp 都要 output label 以及算 error, 這直接透過 Keras LSTM 的 return\_sequences=True 即可以完成.

- b. 如此所得到的 accuracy 會是 0.59 (計算方式為所有 video 的 frame 算 accuracy), learning curve 如下, 比較特別的是雖然 validation loss 隨著 epoch 會往上竄, 但 accuracy 卻可以維持在 0.6 左右.

當然這邊還有許多問題沒有處理, 包括說切 sequence 的長度以及是不是要 random 切, 還有很明顯的 overfitting 的問題.

