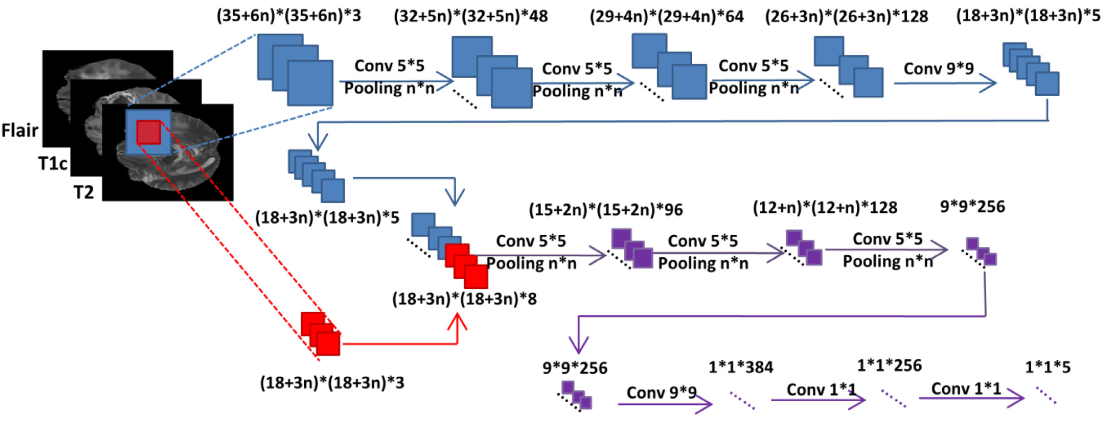
**Final Project－****Human Protein Atlas Image Classification**

隊名：NTU\_r06521504\_隊名我想想  
隊員：陳譽仁 土木所交通組 R06521504  
 趙浩雅 土木所交通組 R06521511

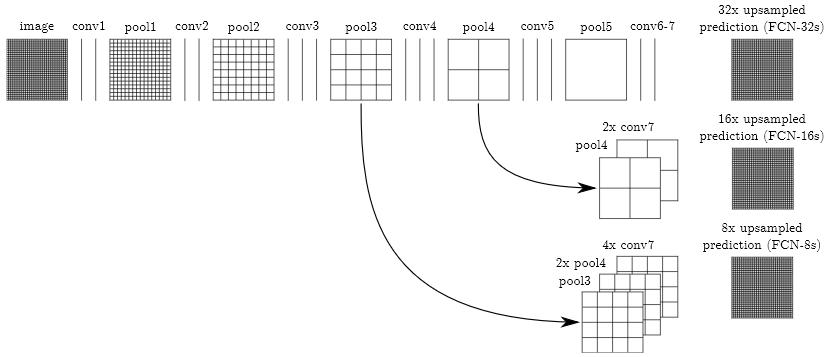
Introduction & Motivation

近幾年，深度學習成為機器學習中最受重視的一區，而藉由此技術完成的影像辨識系統被廣泛的運用在各個領域；其中，藉由影像辨識的系統，可輔助醫療判斷細胞狀況，藉以進行醫療診斷之輔助。考量到台灣未來AI的潛在姓，本次報告我們便選擇相關題目，利用CNN架構去進行圖像辨識，並用ensemble的架構去加強模型。在建立模型部分，本次報告參考兩個方面，分別為paper與競賽上參賽者分享的Kernel。

* 1. Paper方面，本報告參考了Zhao et al., 2018的研究[1]，本篇論文是對大腦腫瘤進行辨識與圖像的分區(segmentation)，在模式比較底層的部分引用了Fully Convolutional Neural Network所衍伸的架構，在模式中的一些部份疊上前幾層的結果，如下圖。



雖然分區不是本報告所選主題要做的事情，但是因為該主題是要辨識散布在圖片上的一些特徵，其中的架構或許可以拿來參考。因此，參考了FCN的文章[2]，這個架構將模式較後面的部分upsampling，再與前面maxpooling前的卷積層結果相加以保留一些特徵在圖片上的位置資訊。



* 1. 參考的Kernel則是來自Kaggle上的Kernel[3]，其中包含完整的程式碼，其模式為數個卷積層的疊加，其中一個卷積層有四個平行的卷積層，各自使用不同的kernel大小，應該是為了方便取出不同大小的特徵並疊加在一起。

Data Preprocessing \ Feature Engineering

目前讀取資料的方法、f1 score的計算、Data generator參考本競賽其他參加者在Kernel分享的程式碼[3]，對圖片資料先用np.stack的方式將同一細胞之紅、綠、藍、黃圖連接，以便在之後training過程讀取特徵值。為了增強數據，本報告用imgaug的套件將圖片進行平移、縮放、錯切等動作並加入少量的噪音，由於擔心其數據會造成過大偏差，未使用像素平移。

考慮到有部分種類數量過少，會在訓練過程中被忽略或著是被切除，在最開始切資料的時候我們另外進行篩選，避免出現未訓練該項目的狀況。

Model Description

本報告於開始先建立一個CNN+DNN的基礎模型，其架構如下圖所示：

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type)Output Shape Param # Connected to

=====================================================

input\_1 (InputLayer) (None, 256, 256, 4) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (BatchNor (None, 256, 256, 4) 16input\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_1 (Conv2D) (None, 254, 254, 32) 1184 batch\_normalization\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (BatchNor (None, 254, 254, 32) 128 conv2d\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2D) (None, 127, 127, 32) 0 batch\_normalization\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 127, 127, 32) 0 max\_pooling2d\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (BatchNor (None, 127, 127, 32) 128 dropout\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 63, 63, 64) 18496 batch\_normalization\_3[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_4 (BatchNor (None, 63, 63, 64) 256 conv2d\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 61, 61, 64) 36928 batch\_normalization\_4[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_5 (BatchNor (None, 61, 61, 64) 256 conv2d\_3[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 59, 59, 64) 36928 batch\_normalization\_5[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_6 (BatchNor (None, 59, 59, 64) 256 conv2d\_4[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2D) (None, 29, 29, 64) 0 batch\_normalization\_6[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 29, 29, 64) 0 max\_pooling2d\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_7 (BatchNor (None, 29, 29, 64) 256 dropout\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_5 (Conv2D) (None, 27, 27, 128) 73856 batch\_normalization\_7[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_8 (BatchNor (None, 27, 27, 128) 512 conv2d\_5[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_6 (Conv2D) (None, 25, 25, 128) 147584 batch\_normalization\_8[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_9 (BatchNor (None, 25, 25, 128) 512 conv2d\_6[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_7 (Conv2D) (None, 23, 23, 128) 147584 batch\_normalization\_9[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 23, 23, 128) 0 conv2d\_7[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

global\_average\_pooling2d\_1 (Glo (None, 32) 0 dropout\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

global\_average\_pooling2d\_2 (Glo (None, 64) 0 dropout\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

global\_average\_pooling2d\_3 (Glo (None, 128)0 dropout\_3[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

concatenate\_1 (Concatenate) (None, 224)0 global\_average\_pooling2d\_1[0][0]

global\_average\_pooling2d\_2[0][0]

global\_average\_pooling2d\_3[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_10 (BatchNo (None, 224)896 concatenate\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256)57600 batch\_normalization\_10[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_4 (Dropout) (None, 256)0 dense\_1[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_11 (BatchNo (None, 256)1024 dropout\_4[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 256)65792 batch\_normalization\_11[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_5 (Dropout) (None, 256)0 dense\_2[0][0]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

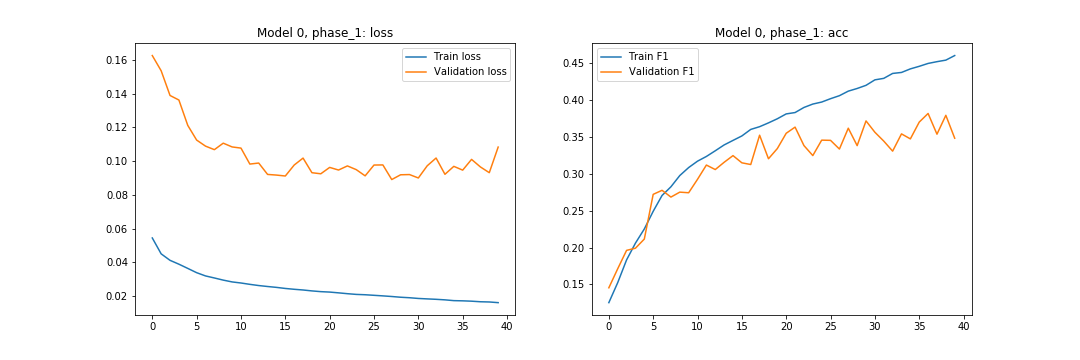
dense\_3 (Dense) (None, 28) 7196 dropout\_5[0][0]

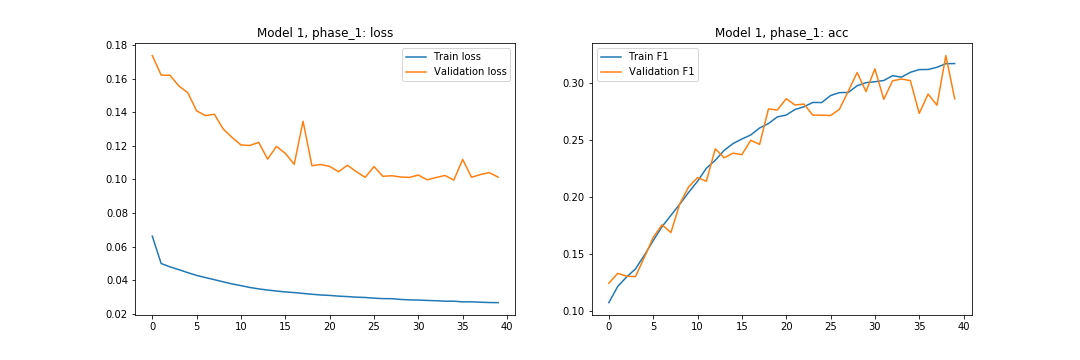
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

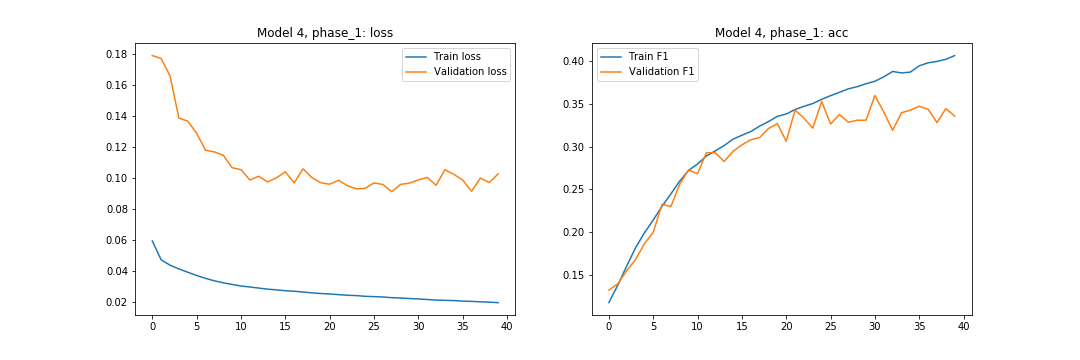
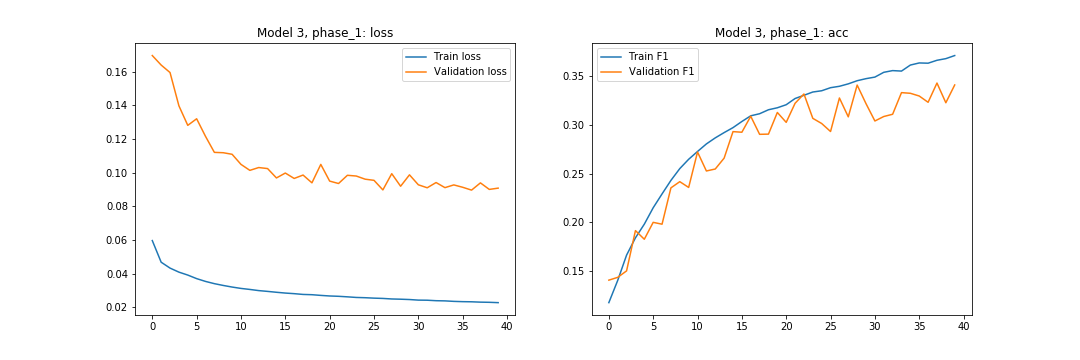
activation\_1 (Activation) (None, 28) 0 dense\_3[0][0]

Total params: 597,388 Trainable params: 595,268 Non-trainable params: 2,120

並以此模型架構做參數調整、包含將kernelsize設定與其他dropout的數值調整，另外建置出八個基礎模型，進行40次epoch訓練。下圖為模型在訓練過程中lose與F1數值變化：



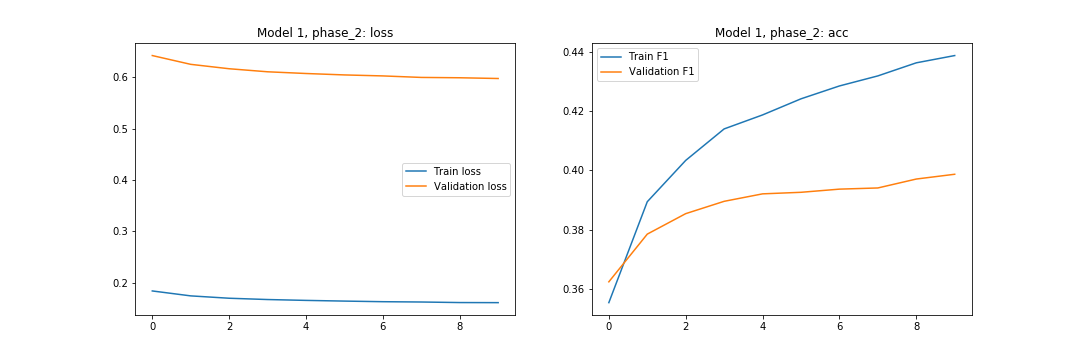




圖一、Base model 0-4

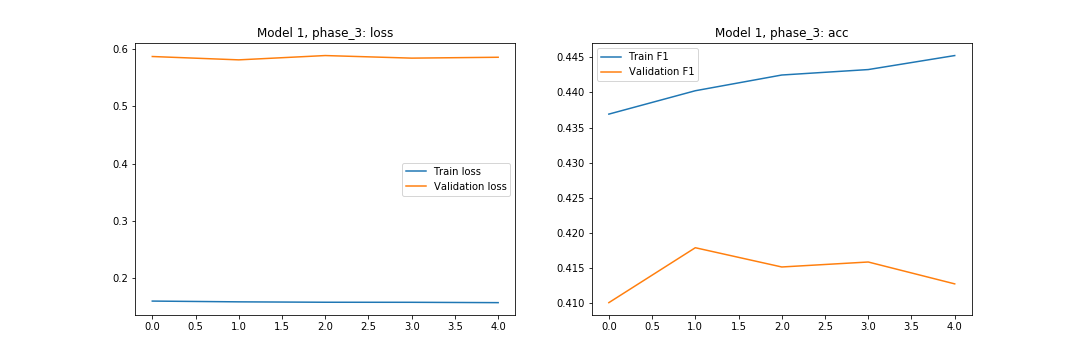
Train lose、Validation lose、Train F1，Validation F1變化過程

從圖片中可發現，原本的模型在經過四十次epochs後Validation F1大約落在0.3-0.35之間，而且有些還有上升趨勢。所以本報告後續用pretrain model用微調(Fine-tune)的方式，對CNN與DNN部分進行重新訓練，另外產生24個model，兩者過程分別如下：



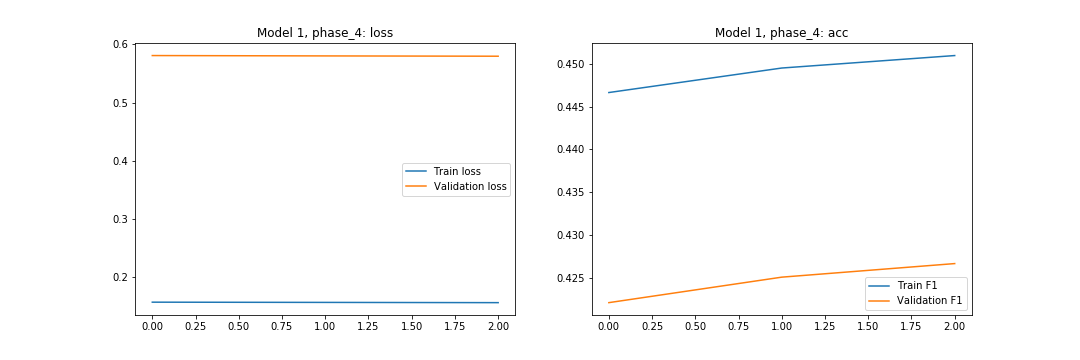
圖二 Base model 1 pretrain後用CNN繼續訓練10個epochs後

Train lose、Validation lose、Train F1，Validation F1變化過程



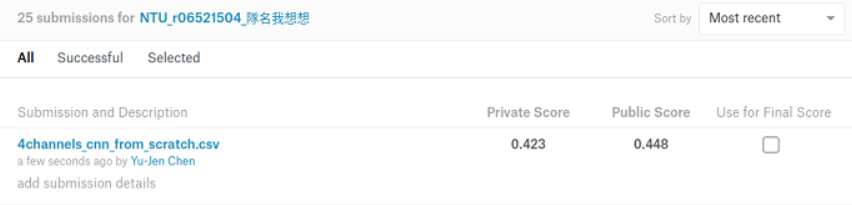
圖三 Base model 1 pretrain+CNN後用DNN繼續訓練4個epochs後

Train lose、Validation lose、Train F1，Validation F1變化過程



圖四 Base model 1 pretrain+CNN+DNN後用DNN繼續訓練2個epochs後Train lose、Validation lose、Train F1，Validation F1變化過程

從圖片趨勢中可以發現，後續訓練已經有一些overfitting的現象，不過其validation F1依舊有些許上升。最後本報告依照試驗後採用model 0, model 1, model 2,model 3,model 4與model 6，權重部分先是參考Validation F1的數值（約在0.42-0.47之間）後並進行平移(-0.34)，最後依照得到的權重比用ensemble的方式進行預測。在kaggle上得到public score 0.448與private score 0.423的成績。



Experiment and Discussion

在最後ensemble的部分，從之前測試中可發現平移權重大小其實對最後數據相差極高。以0.34與0.36做平移值，其相差的public score可以差到0.02。

另外，我們也曾經將32個模型中前期未經過CNN與DNN的模型一併放入ensemble，在權重都為1的狀況下，雖然public score並沒有比較高(0.428)，但是後來private score出來後可發現其數值較其他模型的private score要高(0.428)，推測此部分是因為權重未調整到最佳數據，如果進行調整，或許可以得到更高的分數。

Conclusion

本次報告我們先用CNN與DNN的模型進行訓練，並使用ensemble的方式加強模型強度，最終得到不錯的成果。

**Reference**

1. X. Zhao, Wu, Y., Song, G., Li, Z., Zhang, Y., and Fan, Y., “A deep learning model integrating FCNNs and CRFs for brain tumor segmentation,” *Medical image analysis*, 43, 98-111, 2018
2. Jonathan Long, Evan Shelhamer, Trevor Darrell; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440
3. Michal Haltuf, “CNN 128x128x4, Keras from scratch [LB 0.328],” *kaggle.com*, Oct. 30, 2018. [Online]. Available: <https://www.kaggle.com/rejpalcz/cnn-128x128x4-keras-from-scratch-lb-0-328?fbclid=IwAR2SbpFcM0WhZfs7I9xeIukyE6e9bMoUKX_kjxo3fUhaFeS2UUM6QQ2LpT8>. [Accessed Dec. 14, 2018].