LAB3 Diabetic Retinopathy Detection

陽明大學不分系二年級張凱博

1. Introduction

這次lab的內容是要利用ResNet (一種Convolutional Neuron Network)辨別視網膜病變的圖片，最後要將每一張圖片分類到相對應的種類，類型共有5種，這次會先利用28,099圖片做training，再利用7025張照片做testing，所以這是屬於supervise learning。

ResNet在2015年得到多次圖像辨識比賽的冠軍並刷下多項紀錄，並在2016年使其作者得到CVPR2016最佳論文獎。ResNe利用residual learning解決傳統CNN發生的gradient vanishing和exploding問題，所謂的reisdual learning就是在layer最後的output 加上最一開始的input經過convolution過後的的x，這樣還可以保有一開始input的特性。

1. Experiment setups
2. The details of your model (ResNet)
3. ResNet18

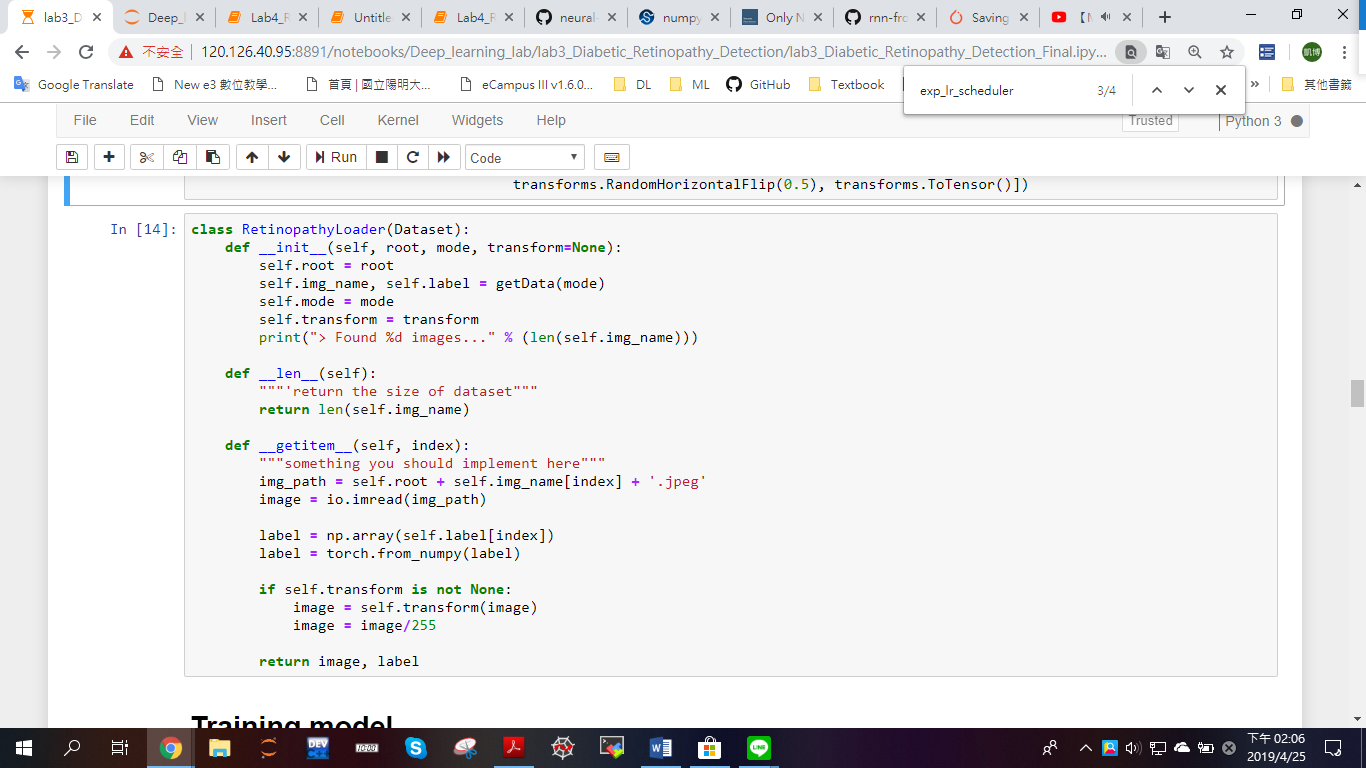
ResNet18最重要的特性是他那四層layer主要的架構為BasicBlock，BasicBlock就是很直接的將layer與layer之間的深度單純乘以2的倍數，kernel size等於3x3進行convolution，，沒有進行任何升維或降維的步驟，直接用[2, 2, 2, 2] 輸出，共有(2+2+2+2)\*2共16層convolutional layer。

1. ResNet50

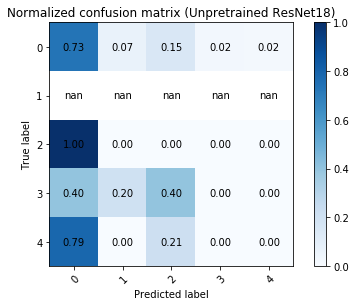
相對的ResNet50最主要的內容是使用了BottleNeck，BottleNeck比BasicBlock多了中間一層，它先用1x1的convolutional layer先降維進行1x1再進行主要3x3的convolution，這樣可以使得feature map數量變少，增加運算效率，最後一層layer會再升維，但在ResNet50的layer和layer間的深度是以4的倍數相乘。

1. The details of your Dataloader

我的dataloader依照助教給的提示，共有3個method，分別是\_\_init\_\_、\_\_len\_\_、\_\_getitem\_\_，\_\_init\_\_就是在一開始的時候就會自動設好我們裡面寫進的變數，\_\_len\_\_會return dataset的size，\_\_getitem\_\_會將每張圖的像素值讀出並經過transformation在除上225 (Normalization)，最後返回影像的向素質陣列以及的它的ground truth label，我的dataloader設計如下:

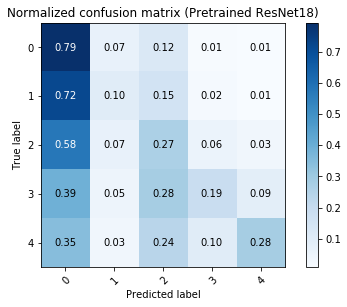


1. Describing your evaluation through the confusion matrix
2. Resnet18



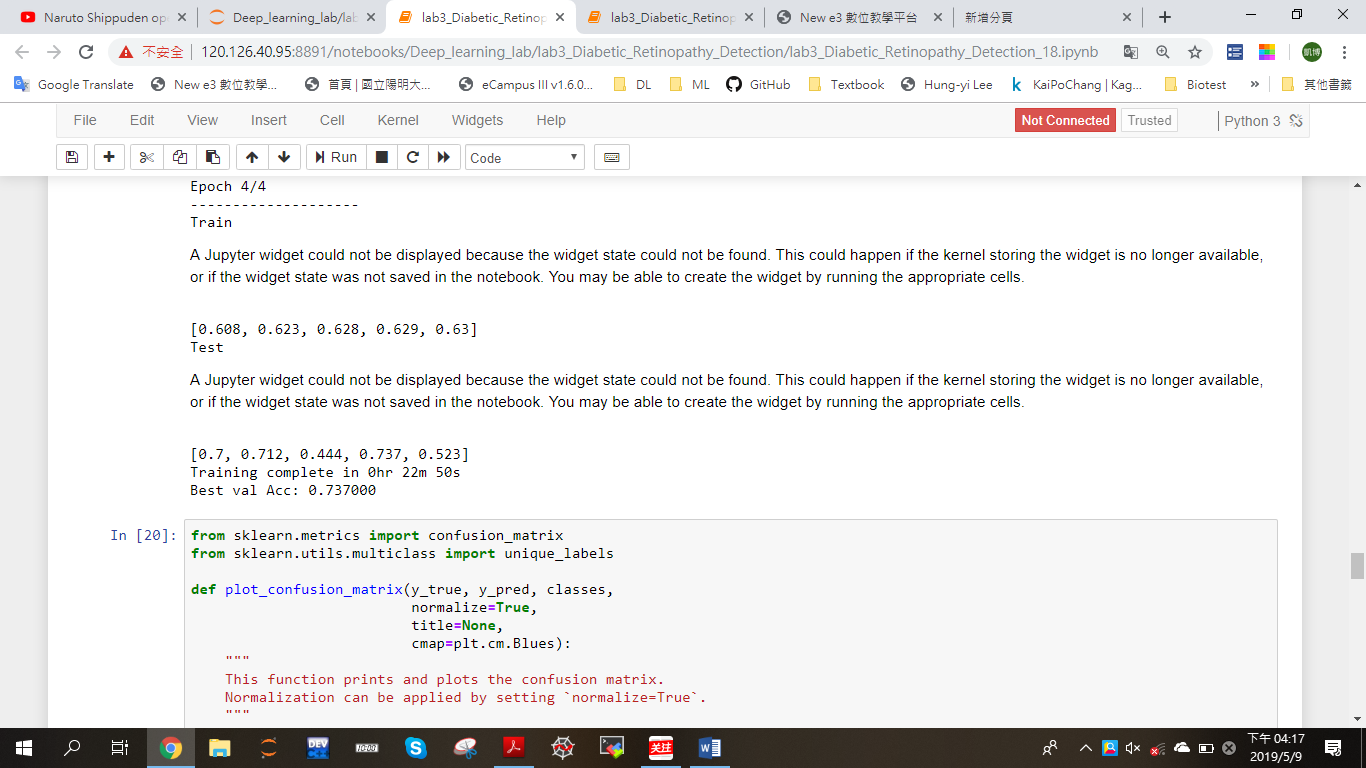
比較兩張有pretrain和沒有pretrain的confusion matrix圖可以發現，有pretrain過後的ResNet18會比unpretrain的ResNet18還要可以精準分類，在對角線上的數值 (正確配對的機率) pretrained ResNet18都比unpretrained ResNet18還要大的多，在unpretrained ResNet18上甚至沒有預測到1的這個label，所以整體來說符合預期，有pretrain過後的model預測結果比unpretrained還要好。

1. Resnet50

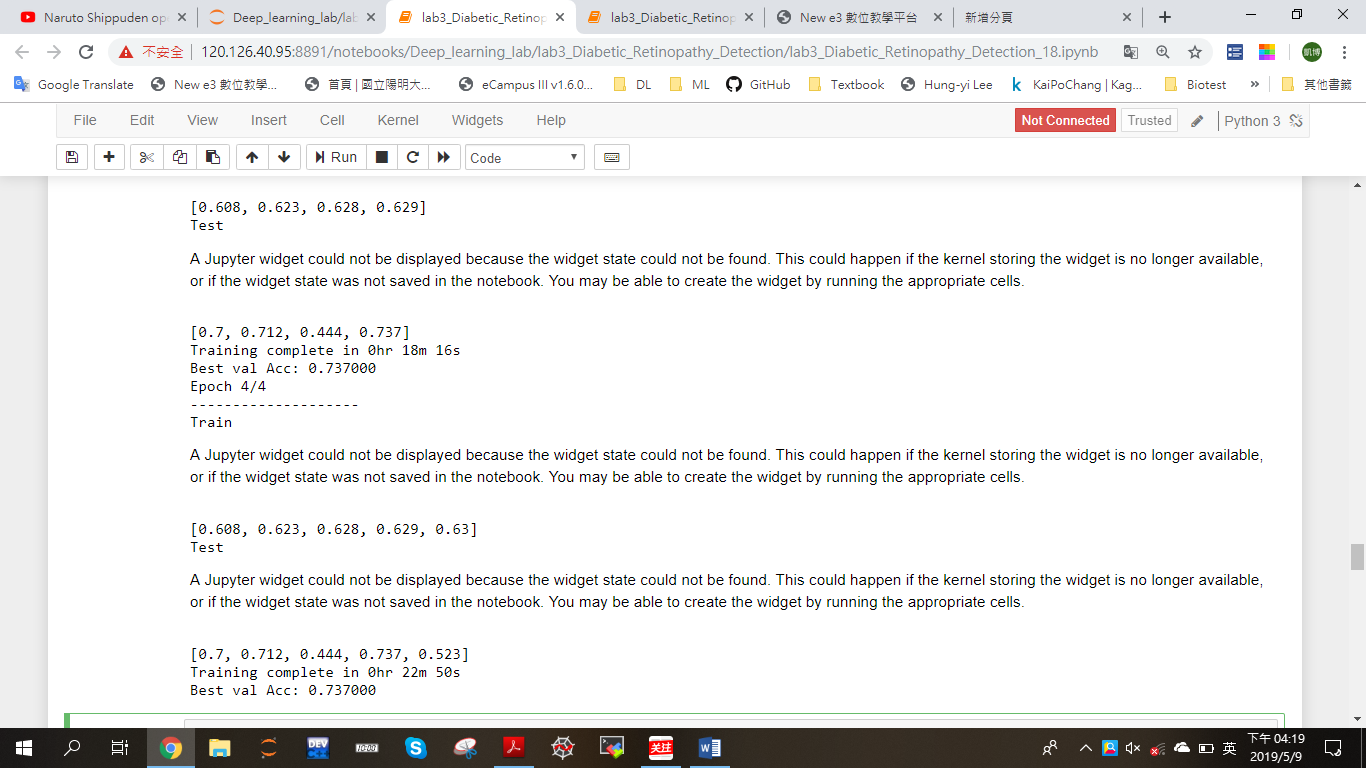


比較兩張有pretrain和沒有pretrain的confusion matrix圖可以發現，有pretrain過後的ResNet50會比unpretrain的ResNet50還要可以精準分類，在對角線上的數值 (正確配對的機率) pretrained ResNet50都比unpretrained ResNet50還要大的多，所以整體來說也是符合預期，有pretrain過後的model預測結果比unpretrained還要好。

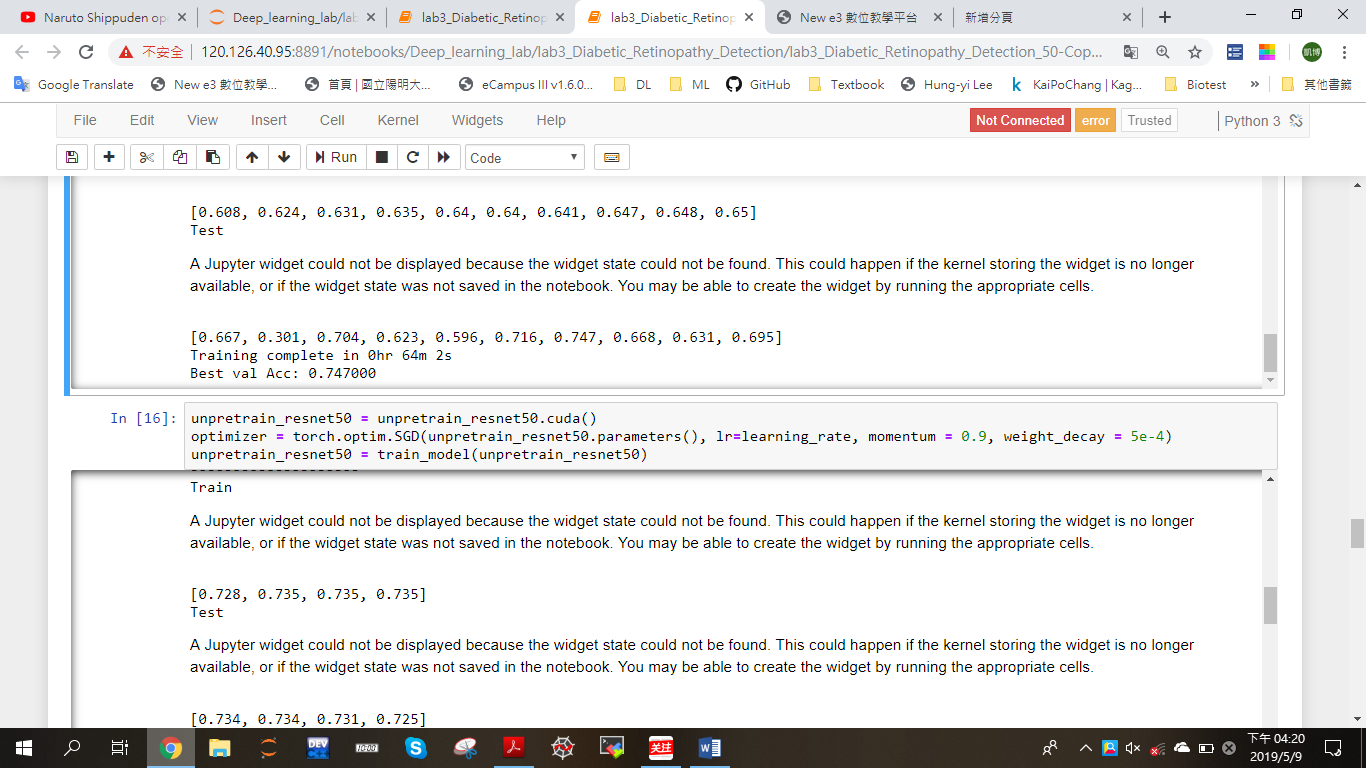
1. Experimental results
2. The highest testing accuracy
3. Screenshot
4. Pretrained ResNet18



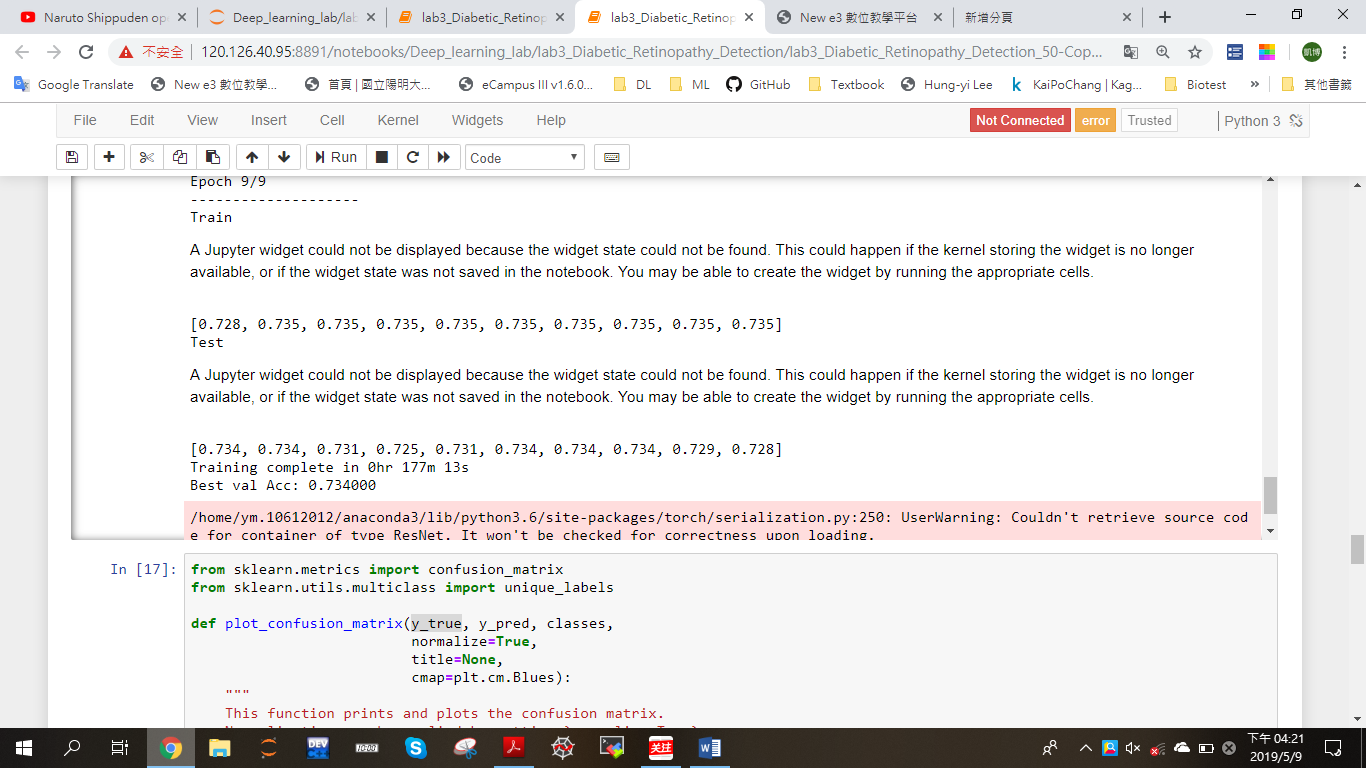
1. Unpretrained ResNet18



1. Pretrained ResNet50



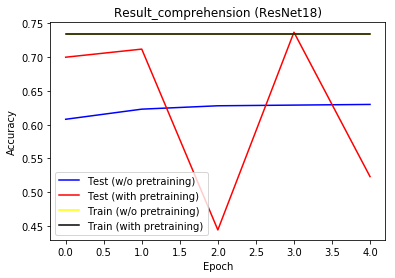
1. Unpretrained ResNet50

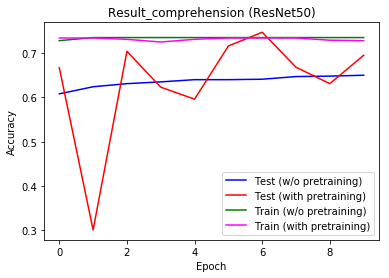


1. Anything want to present

我在ResNet18總共預測了5個epoch，在ResNet50共預測了

10個epoch。在pretrain的部分，無論是ResNet18或ResNet50，雖然model在training的預測能力有在上升，但是在test的結果卻是忽高忽低，我估測是training的epoch太少，所以沒有辦法使得model分類的testing能力穩定提升。相對的，在unpretrain的部分，也一樣無論是ResNet18或ResNet50，符合預期在只有5輪或10輪的epoch，都沒有辦法train起來。

1. Comparison figures



1. Discussion

我一開始並沒有使用GPU而是直接使用CPU，所以在training和testing ResNet18的時候，各一輪就要花上一個小時半，若是在ResNet50的時候，各一輪甚至要花上近乎23小時的時間，所以一開始真的是覺得跑超級慢，但是當使用GPU後，速度就節省了將近20倍的時間。

此外，若仔細觀察pretrained ResNet18的Confusion matrix，會發現在最後分類結果，model常常將影像分類成第零類；pretrained ResNet50也有一樣的情況，但是情況相較pretrained ResNet18有改善很多，仔細分析，原因有可能是training 的epoch太少，觀察unpretrained ResNet50之後，和pretrained ResNet50做比較，unpretrained model在training的次數不夠的時候，會將大多數的predicted model歸類為零，因此，若是training的次數到達百次之後，相信準確率必能提升至8成以上。