1. Introduction

A picture containing diagram

Description automatically generated

這次的實驗室利用深度學習的模型 (ResNet) 去辨識圖片中的眼球是否有視網膜病變，最後的分類結果為0~4等級分別從輕症到重症。從這次的實驗我們可以學習到以下的目標：

1. 實作Dataloader的細節，如何將資料集當中的圖片轉成Tensor的形式，並且透過不同的Data Augmentation 的方法，在現有的資料集中增加翻轉過後圖片。

2. 比起從頭訓練一個新的模型，也可以學習如何將一個已經訓練得不錯的模型參數載入進來繼續訓練，並且比較有經過Pretrained以及沒有經過Pretrained的模型差別。

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

補充：Bottleneck block V.S Basic block兩個的差異及好處。

Residual為何可以解決Gradient vanished的問題，使用後Gradient會如何改變。（至少還會有1）

Hyperparameters的設定都是default，除了有嘗試把batch size加大成8以外都沒有改變。

Text

Description automatically generated

ResNet18當中第一層是由左圖構成，剩下的二到五層是由右邊的基本區塊重複堆疊而成。

Text

Description automatically generated Text

Description automatically generated

ResNet50當中第一層是由左圖構成，剩下的二到五層是由右邊的基本區塊重複堆疊而成。

Text

Description automatically generatedText

Description automatically generated

ResNet最重要的一個架構就是利用Residual的方式來改善Gradient vanished 的問題，透過將前面層數的參數直接拿到後面層數進行相加，這樣就不會因為神經網路太深的關係導致Gradient到最後太小而沒辨法更新參數的問題。

Diagram

Description automatically generated

但是觀察上圖可以發現一個問題我們需要去處理，那就是Residual在將比較前面層數的參數傳到後面層數進行相加的時候，可能會因為我們神經網路的input cannel以及output cannel不相符的關係導致沒辦法直接進行相加，這時候就需要多一層網路來幫助我們將input cannel以及output cannel調整成一致以利相加。

Text

Description automatically generated

* 1. The details of your Dataloader

補充：要做Normalization才會讓mean可以移到0.5。

且Testing時通常不會做Augmentation，所以要用mode來區分。

讀進來的圖片檔可以透過Transforms() 的函式進行調整，首先是經過垂直以及水平的翻轉，再來透過ToTensor() 可以將原本 [H, W, C] 的圖片轉為

[C, H, W]，並且將pixel的值域範圍scaling到 [0.0, 1.0] 的區間裡。

Text

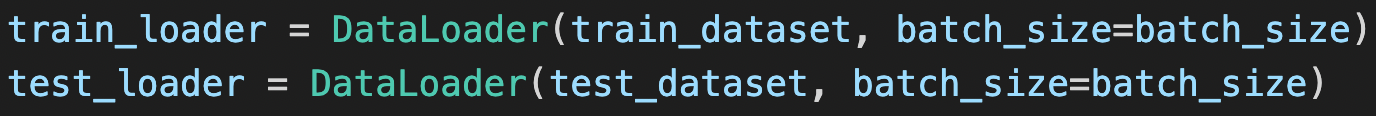
Description automatically generated

再來只需要使用\_\_getitem\_\_就可以根據index匯入進需要的tensor。

Text

Description automatically generated

在主程式當中就可以透過DataLoader套件根據batch\_size將資料讀進來。



接下來就可以從train\_loader中根據index一筆一筆的把訓練資料讀進來。

.long().view(-1) 是在tensor當中要把向量攤開的一個函式，然後就是將我們的輸入資料用forward的方式丟進去網路建出一個Graphic model，並且根據輸出跟正確答案計算Loss function，再來就是進行Backpropagation去跟新網路參數。

Text

Description automatically generated

* 1. Describing your evaluation through the confusion matrix

下方分別是ResNet18以及ResNet50 的confusion matrix，都有做過Row normalize所以每一row加總起來的機率會是100%，觀察後可以發現他們兩個都有一個共同點就是很容易會將Class1的圖片辨識成其他類別，我想可能是因為Class1的照片本身就是輕症，很難從中找到特徵。

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated

1. Experiment results
   1. The highest testing accuracy
      1. Screenshot

Text

Description automatically generated

Text

Description automatically generated

* + 1. Anything you want to present

下圖是有經過pretrained過後的ResNet50訓練的結果，可以看到在testing時中間有非常劇烈的上下起伏，也許是在第4個epoch時出現了overfitting的現象。

Chart, line chart

Description automatically generated

* 1. Comparison figures
     1. Plotting the comparison figures

(ResNet18/50, with/without pretrained)

Chart, line chart

Description automatically generated

1. Discussion

這次的ResNet真的不是普通難訓練，可能因為層數很深，所以光訓練時間就是一個很大的問題，沒辦法去嘗試很多不同的超參數，有試著使用不同的optimizer (e.g, SGD, Adam, AdamW)，但似乎對結果的幫助都沒有特別大，也有試著把batch size加大成default的兩倍，但似乎也不起什麼作用，因為資料本身就已經是一個高維度的向量，如果batch size太大整體的維度會太大。

另外其實從結果可以看到，沒有經過pretrained的模型基本上有點難訓練起來，其實也蠻合理的，畢竟有經過pretrained的模型參數並不是隨機設定，而是從一個比較好的起點開始，在起跑點就已經領先沒有pretrained的模型許多了。但如果單看下圖未經pretrained的模型其實正確率也是有在上升，只是可能epoch以及資料都不夠多，沒辦法到最後使模型參數收斂。

Word

Description automatically generated with medium confidence