LAB3 Diabetic Retinopathy Detection

陽明大學不分系二年級張凱博

(一) 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。

(二) Experiment setups

A. The details of your model (ResNet)

1. 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。

2. 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。

B. The details of your Dataloader

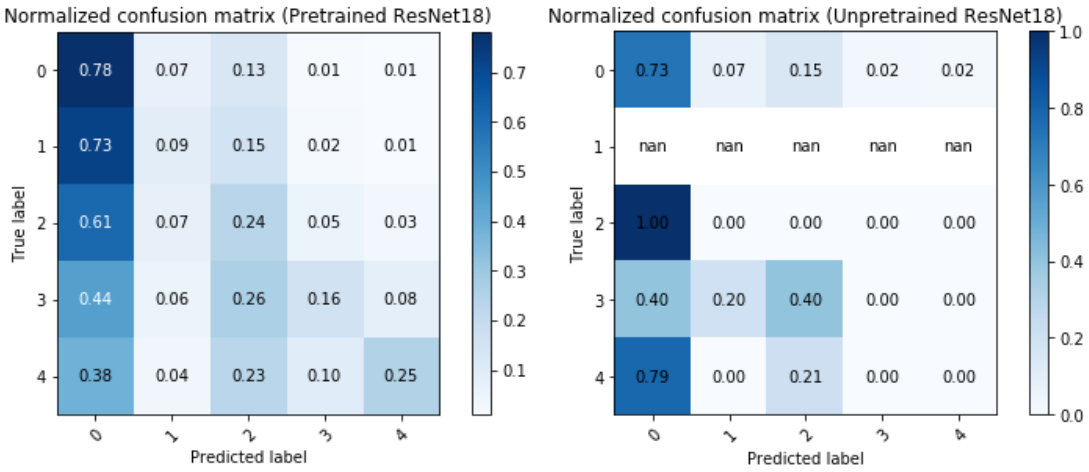
我的 dataloader 依照助教給的提示，共有 3 個 method，分別是 __init__、__len__、__getitem__，__init__就是在一開始的時候就會自動設好我們裡面寫進的變數，__len__會 return dataset 的 size，

__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:

```python
class RetinopathyLoader(Dataset):
    def __init__(self, root, mode, transform=None):
        self.root = root
        self.img_name, self.label = getData(mode)
        self.mode = mode
        self.transform = transform
        print("> Found %d images..." % (len(self.img_name)))

    def __len__(self):
        """'return the size of dataset"""
        return len(self.img_name)

    def __getitem__(self, index):
        """something you should implement here"""
        img_path = self.root + self.img_name[index] + '.jpeg'
        image = io.imread(img_path)

        label = np.array(self.label[index])
        label = torch.from_numpy(label)

        if self.transform is not None:
            image = self.transform(image)
            image = image/255

        return image, label
```

C. Describing your evaluation through the confusion matrix

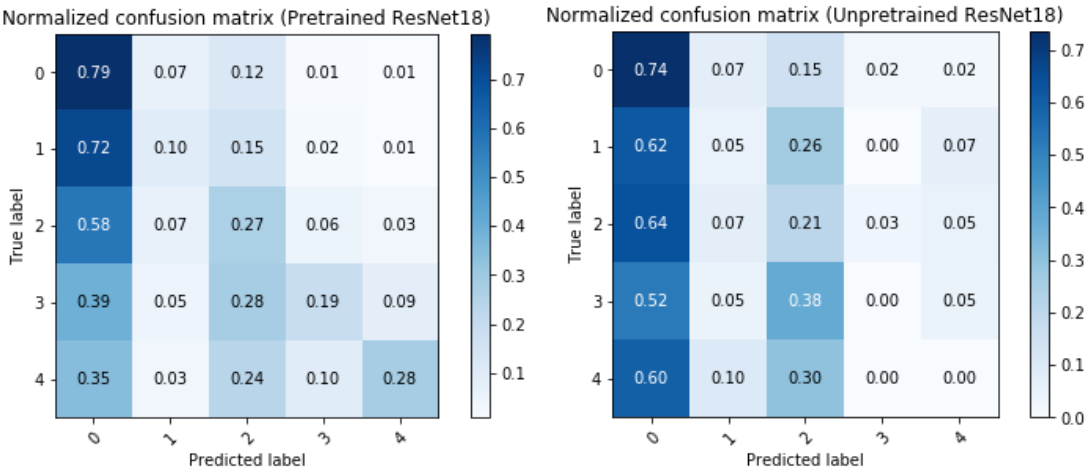
1. Resnet18



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。

2. 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。

(三) Experimental results

A. The highest testing accuracy

1. Screenshot

(1) Pretrained ResNet18

```
[0.7, 0.712, 0.444, 0.737, 0.523]
Training complete in 0hr 22m 50s
Best val Acc: 0.737000
```

(2) Unpretrained ResNet18

```
[0.7, 0.712, 0.444, 0.737, 0.523]
Training complete in 0hr 22m 50s
Best val Acc: 0.737000
```

(3) Pretrained ResNet50

```
[0.667, 0.301, 0.704, 0.623, 0.596, 0.716, 0.747, 0.668, 0.631, 0.695]
Training complete in 0hr 64m 2s
Best val Acc: 0.747000
```

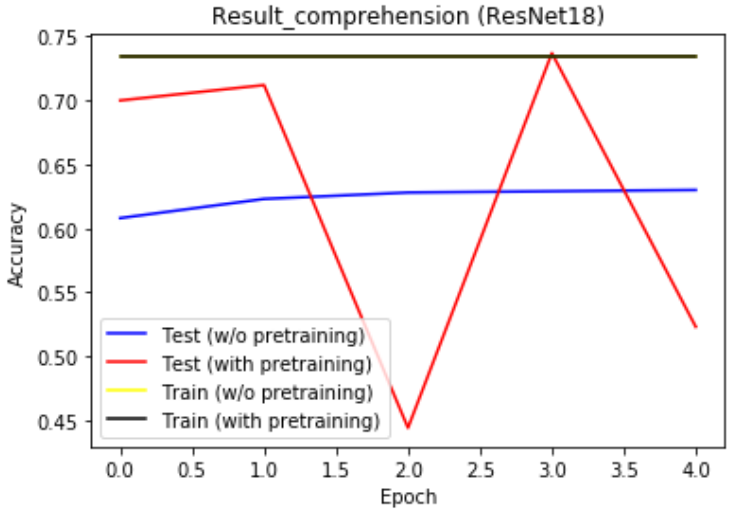
(4) Unpretrained ResNet50

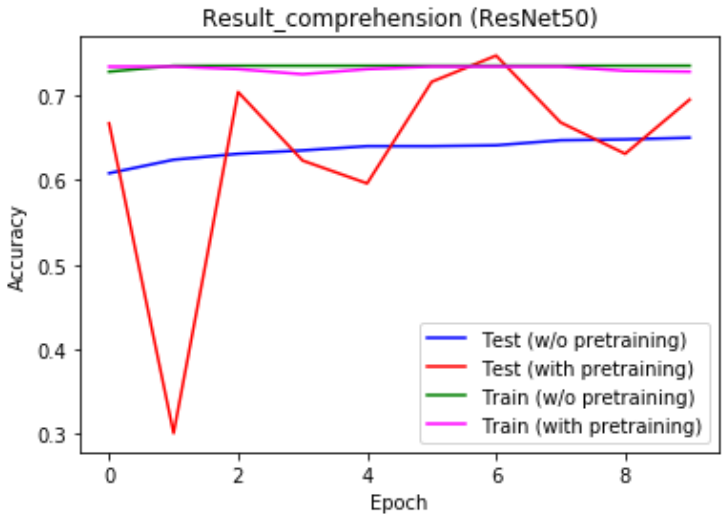
```
[0.734, 0.734, 0.731, 0.725, 0.731, 0.734, 0.734, 0.734, 0.729, 0.728]
Training complete in 0hr 177m 13s
Best val Acc: 0.734000
```

2. 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。

B. Comparison figures





(四) 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。