**Computer Vision HW2 Report**

Student ID: R11522810

Name: 曾建堯

**Part 1. (10%)**

**• Plot confusion matrix of two settings. (i.e. Bag of sift and tiny image) (5%)**

**Ans:**

|  |  |
| --- | --- |
|  |  |
| Tiny image | Bag of SIFT |

**• Compare the results/accuracy of both settings and explain the result. (5%)Ans:**

|  |  |  |
| --- | --- | --- |
|  | Tiny image | Bag of SIFT |
| accuracy | 0.222666 | 0.604666 |

由結果可以發現，相較於利用SIFT為特徵點，直接利用整張相片的Tiny image作為特徵較不精確。因為利用整張照片作為特徵其中也會包含背景等其他部分的干擾，造成辨識錯誤。SIFT特徵點則更有可能可以代表所要辨識的物件，提升辨識成功率。

**Part 2. (25%)**

**• Report accuracy of both models on the validation set. (2%)**

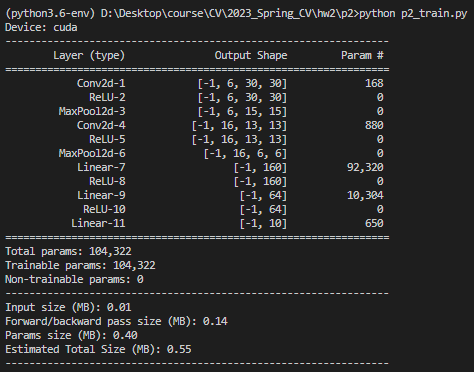
**Ans:**

|  |  |
| --- | --- |
| **mynet** | **Resnet18** |
| **0.6576** | **0.9144** |

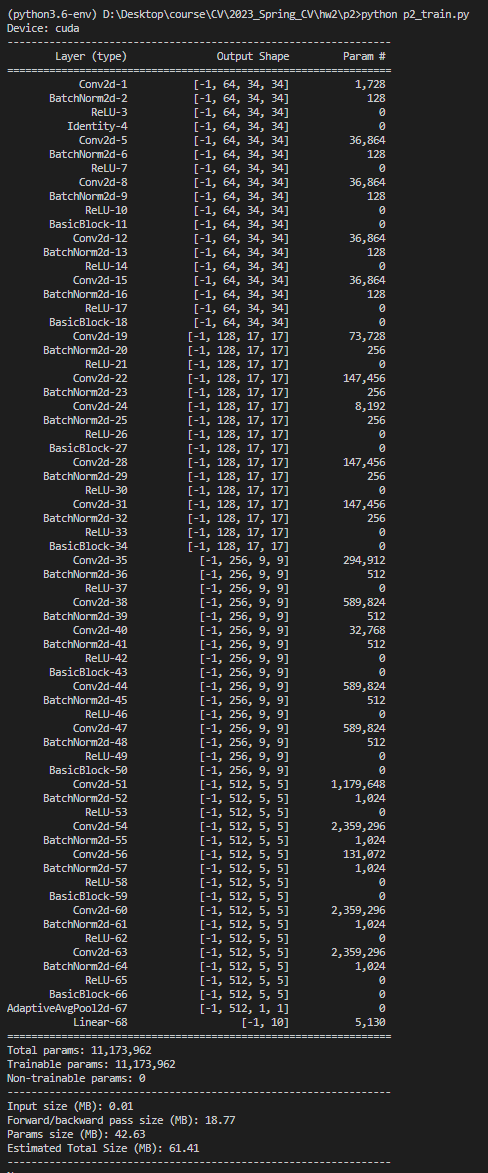
**• Print the network architecture & number of parameters of both models. What is the main difference between ResNet and other CNN architectures? (5%)**

**Ans:**

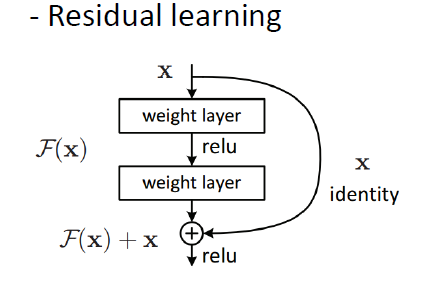
Mynet: numbers of parameters: 104,322

****

Resnet18: numbers of parameter: 11, 173,962

****

Resnet和其他CNN模型最大的差異是在Resnet利用Residual learning利用額外的edge將前面層數的值串到後面的層數，在更新參數時較能傳遞回上面的層數，使較深的模型也能訓練起來。



**• Plot four learning curves (loss & accuracy) of the training process (train/validation) for both models. Total 8 plots. (8%)**

**Ans:**

|  |  |
| --- | --- |
| **mynet** | **Resnet18** |
|  |  |
|  |  |
|  |  |
|  |  |

**• Briefly describe what method do you apply on your best model? (e.g. data augmentation, model architecture, loss function, etc) (10%)**

**Ans:**

我最好的model是利用resnet18做修改的，在模型的部分，按照助教的提示，將第一層的conv2中的kernel改為(3,3)且stride改為(1,1)，並將maxpooling改為identify。訓練的部分，有增加data augmentation, 使用包含 RandomCrop, RandomHorizontalFilp.並且在training中發存在明顯的overfit, 將batch size改為64, optimizer採用sgd、weight\_decay改為1e-4，相較於其他training，overfit情況有所減緩。