**Computer Vision HW2 Report**

Student ID: R12921059

Name: 鄧雅文

**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 |
|  |  |

Bag of sift因為是提取圖片的特徵(SIFT)再丟到KNN，所以accuracy會高很多，相較之下tiny image是直接把整張圖縮小丟到KNN，所以accuracy不會太高。

**Part 2. (25%)**

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

**Ans:**

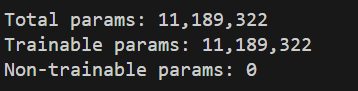
Resnet18: Accuracy = 0.9174

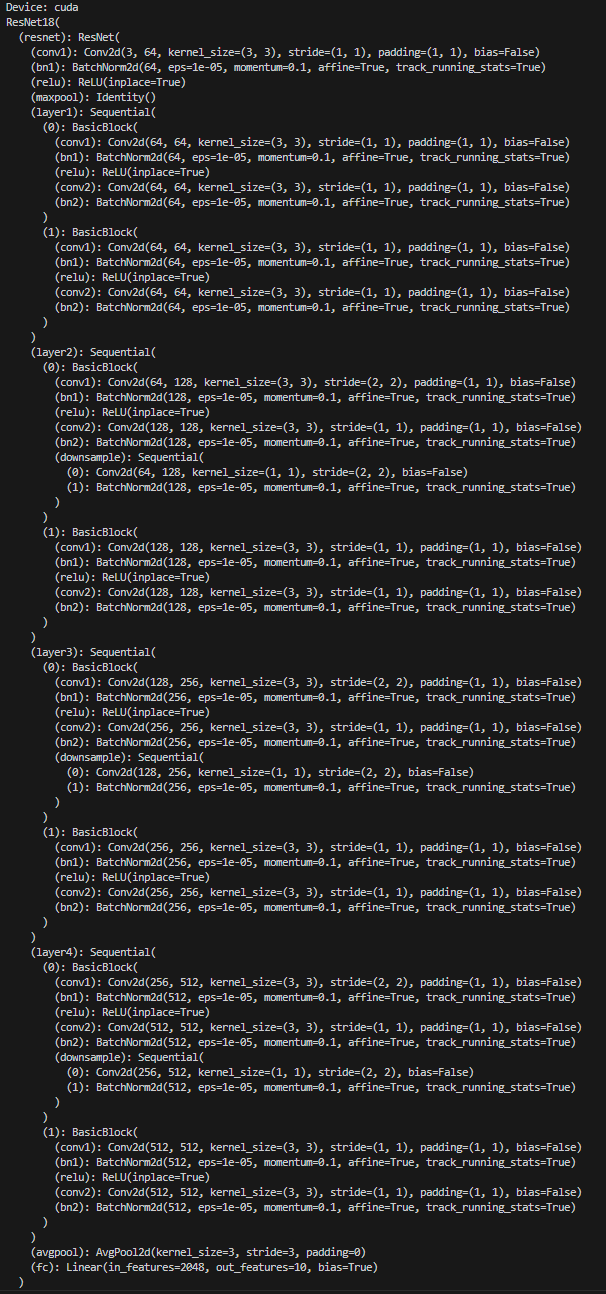
Mynet: Accuracy = 0.865

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

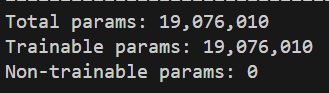
**Ans:**

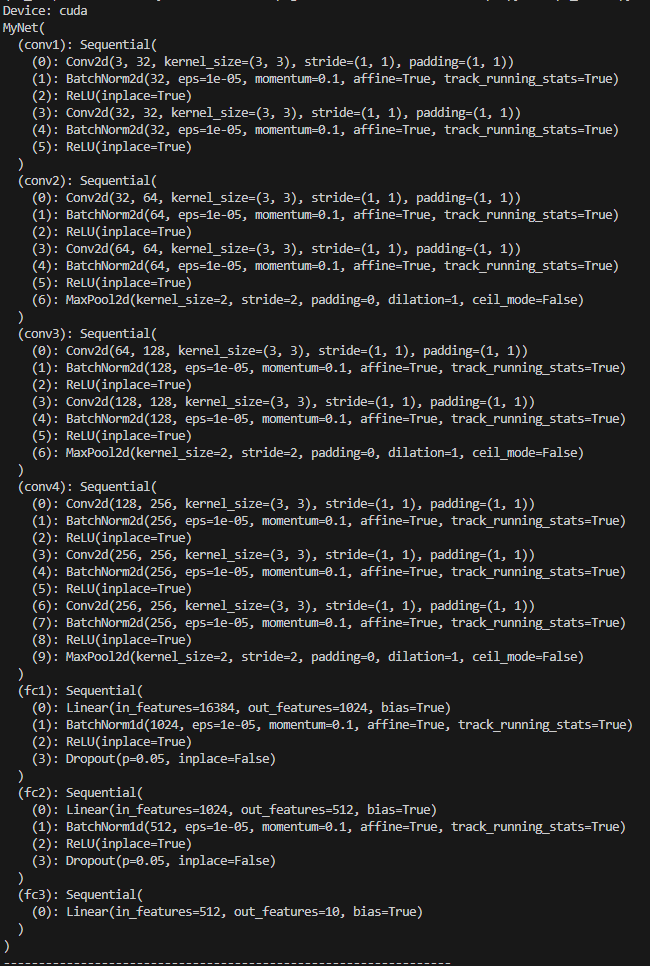
**Resnet18:**

1. Number of parameters:
2. Network architecture:



**Mynet:**

1. Number of parameters:
2. Network architecture:



ResNet最主要的差別是：他是用「殘差映射」(Residual Mapping)訓練深層神經網路，所以它要學習的目標映射函數是F(x)+x，i.e. 是學習輸入和輸出之間的殘差映射，而不是直接學習輸出。

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

**Ans:**

**Resnet18:**

|  |  |
| --- | --- |
| Train accuracy | Train loss |
|  |  |

|  |  |
| --- | --- |
| Validation accuracy | Validation loss |
|  |  |

**Mynet:**

|  |  |
| --- | --- |
| Train accuracy | Train loss |
|  |  |

|  |  |
| --- | --- |
| Validation accuracy | Validation loss |
|  |  |

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

**Ans:**

**Resnet18:**

**目前它的總體表現最好**，用val data測出來的Accuracy = 0.9174。使用的方法：

1. Data augmentation:
   1. Resize到64\*64
   2. Horizontal flip
   3. RandomCrop
2. Model architecture，對ResNet18做以下更改:
   1. 將第一層convolution的kernel size調到3\*3，stride=1
   2. 將第一個maxpool拿掉
   3. 將最後一個avgpool的kernel size調到3\*3
3. Optimizer用SGD

**Mynet:**

用val data測出來的Accuracy = 0.865。使用的方法：

1. Data augmentation 同上
2. 建9層的convolution，每層都加上Batch Normalization，然後中間穿插kernel size=2的max pooling。最後的三個fully connect layers中間加上dropout，避免overfitting。
3. Optimizer用Adam
4. (原本有嘗試用semi-supervised，但結果沒有比較好QQ)

**• Reference**

[1] <https://medium.com/@rossleecooloh/%E7%9B%B4%E8%A7%80%E7%90%86%E8%A7%A3resnet-%E7%B0%A1%E4%BB%8B-%E8%A7%80%E5%BF%B5%E5%8F%8A%E5%AF%A6%E4%BD%9C-python-keras-8d1e2e057de2>

[2] <https://hackmd.io/@Johnsonnnn/ry2lp8JRq>

[3] <https://tigercosmos.xyz/post/2020/06/cv/bag-of-visual-words/>

[4] <https://github.com/Offliners/NTUEE-CV-2022Spring>

[5] <https://github.com/Louislar/NTU_CV_HW>

[6] <https://ithelp.ithome.com.tw/articles/10321837?sc=rss.iron>