# Computer Vision HW2 Report

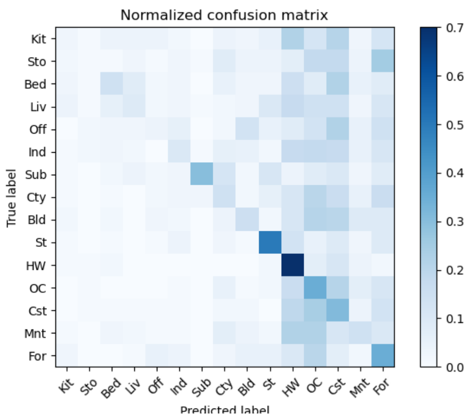
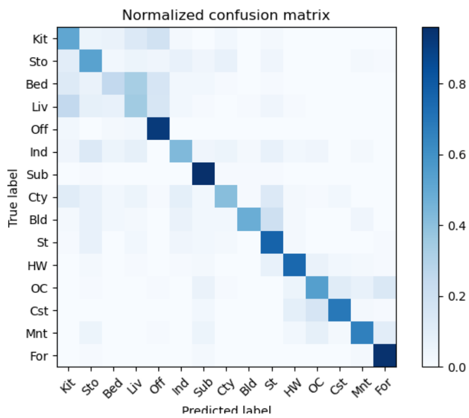

Student ID: R10921061

Name: 袁肇謙


## Part 1. (10%)

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

**Ans:**

<table>
<tr><td></td><td>Tiny image</td><td>Bag of sift</td></tr>
<tr><td>confusion matrix</td><td>
</td><td>
</td></tr>
<tr><td>k</td><td>1</td><td>1</td></tr>
<tr><td>accuracy</td><td>0.22333333333333333</td><td>0.6133333333333333</td></tr>
</table>

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

**Ans:**

As the result shown above, the performance of bag of sift is much better than tiny image. Tiny image is not a particularly good representation, because it discards all of the high frequency image content and is not especially invariant to spatial or brightness shifts.

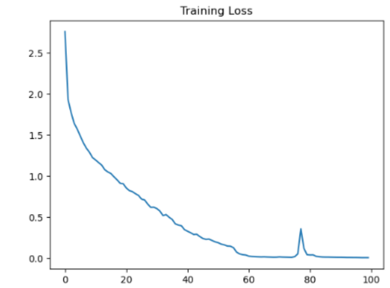
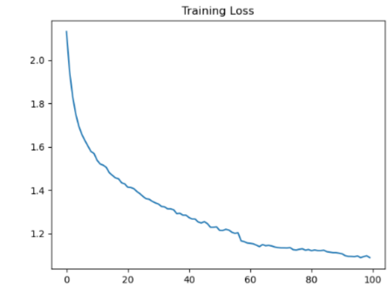
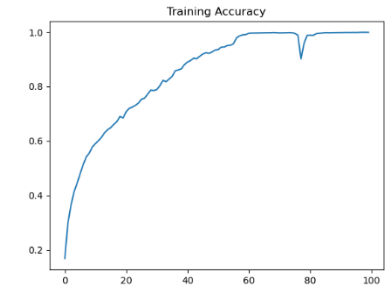
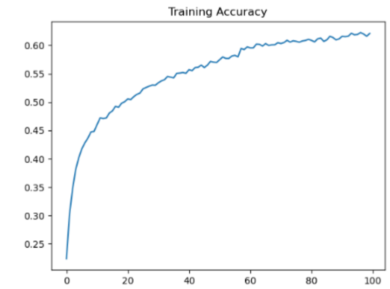
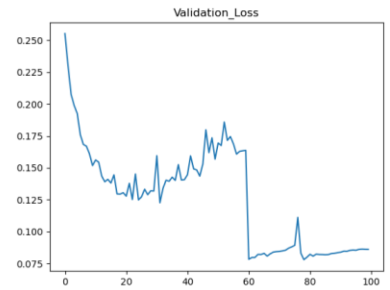
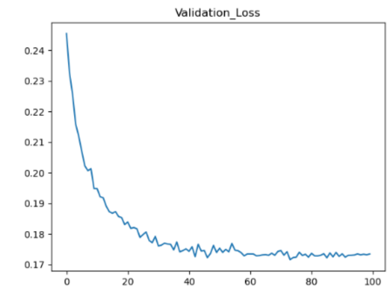
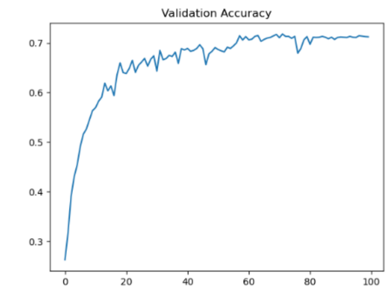
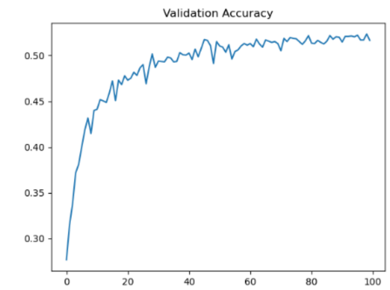
## Part 2. (35%)

**• Compare the performance on residual networks and LeNet. Plot the learning curve (loss and accuracy) on both training and validation sets for both 2 schemes. 8 plots in total. (20%)**

**Ans:**

From the following plots, we can observed that the residual network converged a bit earlier than LeNet.

| | Residual networks | LeNet |
|---|---|---|
| Accuracy (public test) | 0.814 | 0.5976 |

| | | | |
|---|---|---|---|
| Training | Loss |   |   |
| | Accuracy |   |   |
| Validation | Loss |   |   |
| | Accuracy |   |   |

**• Attach basic information of the model you use including model architecture and the number of the parameters. Besides, report the accuracy you performed on the public test set. (5%)**

**Ans:**

- accuracy on the public test set: 0.823
- model architecture and the number of the parameters: 24,562,250

```
==========================================================================================
Layer (type:depth-idx)                                       Param #
==========================================================================================
├─Conv2d: 1-1                                                9,408
├─BatchNorm2d: 1-2                                           128
├─ReLU: 1-3                                                  --
├─MaxPool2d: 1-4                                             --
├─Sequential: 1-5                                            --
│    └─Bottleneck: 2-1                                       --
│    │    └─Conv2d: 3-1                                      4,096
│    │    └─BatchNorm2d: 3-2                                 128
│    │    └─Conv2d: 3-3                                      36,864
│    │    └─BatchNorm2d: 3-4                                 128
│    │    └─Conv2d: 3-5                                      16,384
│    │    └─BatchNorm2d: 3-6                                 512
│    │    └─ReLU: 3-7                                        --
│    │    └─Sequential: 3-8                                  16,896
│    └─Bottleneck: 2-2                                       --
│    │    └─Conv2d: 3-9                                      16,384
│    │    └─BatchNorm2d: 3-10                                128
│    │    └─Conv2d: 3-11                                     36,864
│    │    └─BatchNorm2d: 3-12                                128
│    │    └─Conv2d: 3-13                                     16,384
│    │    └─BatchNorm2d: 3-14                                512
│    │    └─ReLU: 3-15                                       --
│    └─Bottleneck: 2-3                                       --
│    │    └─Conv2d: 3-16                                     16,384
│    │    └─BatchNorm2d: 3-17                                128
│    │    └─Conv2d: 3-18                                     36,864
│    │    └─BatchNorm2d: 3-19                                128
│    │    └─Conv2d: 3-20                                     16,384
│    │    └─BatchNorm2d: 3-21                                512
│    │    └─ReLU: 3-22                                       --
├─Sequential: 1-6                                            --
│    └─Bottleneck: 2-4                                       --
│    │    └─Conv2d: 3-23                                     32,768
│    │    └─BatchNorm2d: 3-24                                256
│    │    └─Conv2d: 3-25                                     147,456
│    │    └─BatchNorm2d: 3-26                                256
│    │    └─Conv2d: 3-27                                     65,536
│    │    └─BatchNorm2d: 3-28                                1,024
│    │    └─ReLU: 3-29                                       --
│    │    └─Sequential: 3-30                                 132,096
│    └─Bottleneck: 2-5                                       --
│    │    └─Conv2d: 3-31                                     65,536
│    │    └─BatchNorm2d: 3-32                                256
│    │    └─Conv2d: 3-33                                     147,456
│    │    └─BatchNorm2d: 3-34                                256
│    │    └─Conv2d: 3-35                                     65,536
│    │    └─BatchNorm2d: 3-36                                1,024
│    │    └─ReLU: 3-37                                       --
│    └─Bottleneck: 2-6                                       --
│    │    └─Conv2d: 3-38                                     65,536
│    │    └─BatchNorm2d: 3-39                                256
│    │    └─Conv2d: 3-40                                     147,456
│    │    └─BatchNorm2d: 3-41                                256
│    │    └─Conv2d: 3-42                                     65,536
│    │    └─BatchNorm2d: 3-43                                1,024
│    │    └─ReLU: 3-44                                       --
│    └─Bottleneck: 2-7                                       --
│    │    └─Conv2d: 3-45                                     65,536
│    │    └─BatchNorm2d: 3-46                                256
│    │    └─Conv2d: 3-47                                     147,456
│    │    └─BatchNorm2d: 3-48                                256
│    │    └─Conv2d: 3-49                                     65,536
│    │    └─BatchNorm2d: 3-50                                1,024
│    │    └─ReLU: 3-51                                       --
├─Sequential: 1-7                                            --
│    └─Bottleneck: 2-8                                       --
│    │    └─Conv2d: 3-52                                     131,072
│    │    └─BatchNorm2d: 3-53                                512
│    │    └─Conv2d: 3-54                                     589,824
│    │    └─BatchNorm2d: 3-55                                512
│    │    └─Conv2d: 3-56                                     262,144
│    │    └─BatchNorm2d: 3-57                                2,048
│    │    └─ReLU: 3-58                                       --
│    │    └─Sequential: 3-59                                 526,336
│    └─Bottleneck: 2-9                                       --
│    │    └─Conv2d: 3-60                                     262,144
│    │    └─BatchNorm2d: 3-61                                512
│    │    └─Conv2d: 3-62                                     589,824
│    │    └─BatchNorm2d: 3-63                                512
│    │    └─Conv2d: 3-64                                     262,144
│    │    └─BatchNorm2d: 3-65                                2,048
│    │    └─ReLU: 3-66                                       --
│    └─Bottleneck: 2-10                                      --
│    │    └─Conv2d: 3-67                                     262,144
│    │    └─BatchNorm2d: 3-68                                512
│    │    └─Conv2d: 3-69                                     589,824
│    │    └─BatchNorm2d: 3-70                                512
│    │    └─Conv2d: 3-71                                     262,144
│    │    └─BatchNorm2d: 3-72                                2,048
│    │    └─ReLU: 3-73                                       --
│    └─Bottleneck: 2-11                                      --
│    │    └─Conv2d: 3-74                                     262,144
│    │    └─BatchNorm2d: 3-75                                512
│    │    └─Conv2d: 3-76                                     589,824
│    │    └─BatchNorm2d: 3-77                                512
│    │    └─Conv2d: 3-78                                     262,144
│    │    └─BatchNorm2d: 3-79                                2,048
│    │    └─ReLU: 3-80                                       --
│    └─Bottleneck: 2-12                                      --
│    │    └─Conv2d: 3-81                                     262,144
│    │    └─BatchNorm2d: 3-82                                512
│    │    └─Conv2d: 3-83                                     589,824
│    │    └─BatchNorm2d: 3-84                                512
│    │    └─Conv2d: 3-85                                     262,144
│    │    └─BatchNorm2d: 3-86                                2,048
│    │    └─ReLU: 3-87                                       --
│    └─Bottleneck: 2-13                                      --
│    │    └─Conv2d: 3-88                                     262,144
│    │    └─BatchNorm2d: 3-89                                512
│    │    └─Conv2d: 3-90                                     589,824
│    │    └─BatchNorm2d: 3-91                                512
│    │    └─Conv2d: 3-92                                     262,144
│    │    └─BatchNorm2d: 3-93                                2,048
│    │    └─ReLU: 3-94                                       --
├─Sequential: 1-8                                            --
│    └─Bottleneck: 2-14                                      --
│    │    └─Conv2d: 3-95                                     524,288
│    │    └─BatchNorm2d: 3-96                                1,024
│    │    └─Conv2d: 3-97                                     2,359,296
│    │    └─BatchNorm2d: 3-98                                1,024
│    │    └─Conv2d: 3-99                                     1,048,576
│    │    └─BatchNorm2d: 3-100                               4,096
│    │    └─ReLU: 3-101                                      --
│    │    └─Sequential: 3-102                                2,101,248
│    └─Bottleneck: 2-15                                      --
│    │    └─Conv2d: 3-103                                    1,048,576
│    │    └─BatchNorm2d: 3-104                               1,024
│    │    └─Conv2d: 3-105                                    2,359,296
│    │    └─BatchNorm2d: 3-106                               1,024
│    │    └─Conv2d: 3-107                                    1,048,576
│    │    └─BatchNorm2d: 3-108                               4,096
│    │    └─ReLU: 3-109                                      --
│    └─Bottleneck: 2-16                                      --
│    │    └─Conv2d: 3-110                                    1,048,576
│    │    └─BatchNorm2d: 3-111                               1,024
│    │    └─Conv2d: 3-112                                    2,359,296
│    │    └─BatchNorm2d: 3-113                               1,024
│    │    └─Conv2d: 3-114                                    1,048,576
│    │    └─BatchNorm2d: 3-115                               4,096
│    │    └─ReLU: 3-116                                      --
├─AdaptiveAvgPool2d: 1-9                                     --
├─Sequential: 1-10                                           --
│    └─Linear: 2-17                                          1,049,088
│    └─ReLU: 2-18                                            --
│    └─Dropout: 2-19                                         --
│    └─Linear: 2-20                                          5,130
==========================================================================================
Total params: 24,562,250
Trainable params: 24,562,250
Non-trainable params: 0
==========================================================================================
```

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

**Ans:**

- data augmentation: RandomHorizontalFlip(p=0.5), RandomRotation(degrees=20),
- model architecture: add a dropout layer in fully-connected layers to prevent model from overfitting
- loss function: CrossEntropyLoss()
- semi-supervised: After the accuracy on validation set reached 0.7, I used the model at that point to generate pseudo labels. The images and the corresponding pseudo labels whose confidence was greater than 0.9 were added to the training set. This method gave me approximately 0.012 improvement on accuracy.