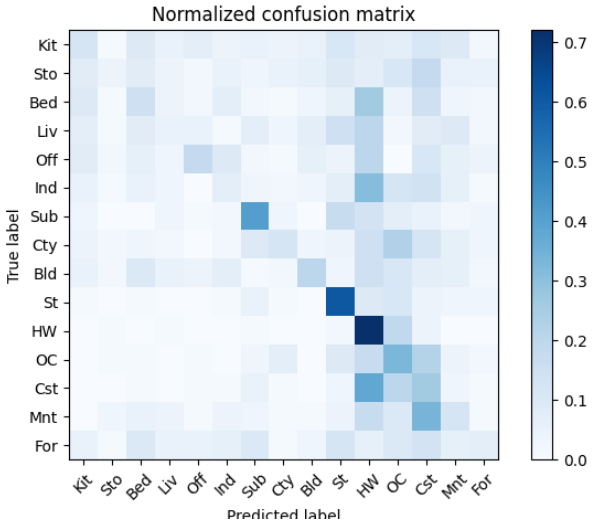
Computer Vision HW2 Report

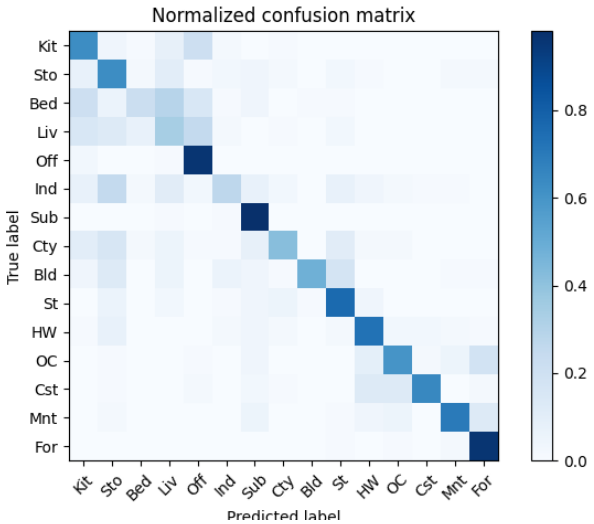
Student ID: R11921041

Name: 蔣沅均

Part 1. (10%)

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





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

tiny image 的 accuracy 是 0.243，而 bag of sift 的是 0.622。兩個方法使用的 knn 分類器設定都一樣，因此可以比較出 bag of sift 相較於 tiny image 更能有效的抽取出利於場景分類的圖像特徵。

Part 2. (25%)

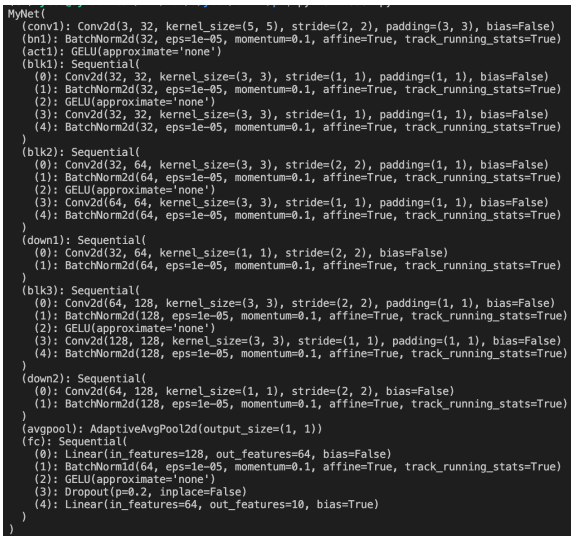
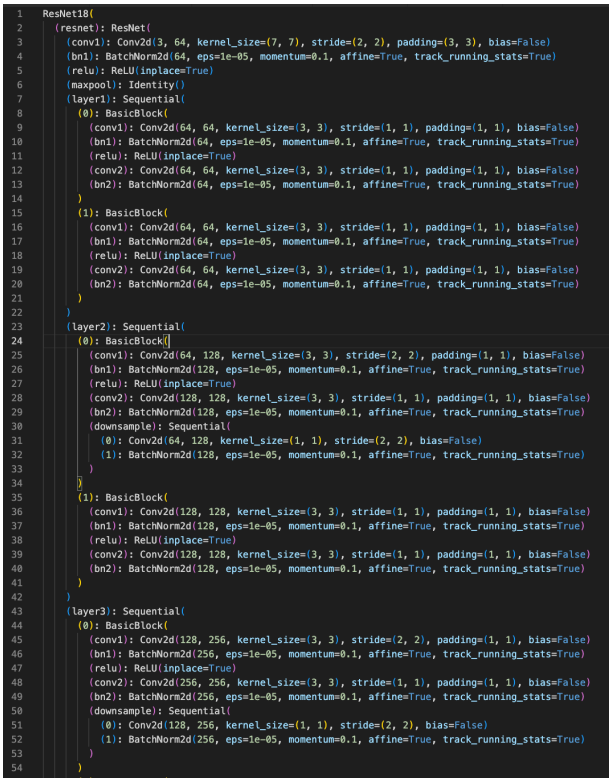
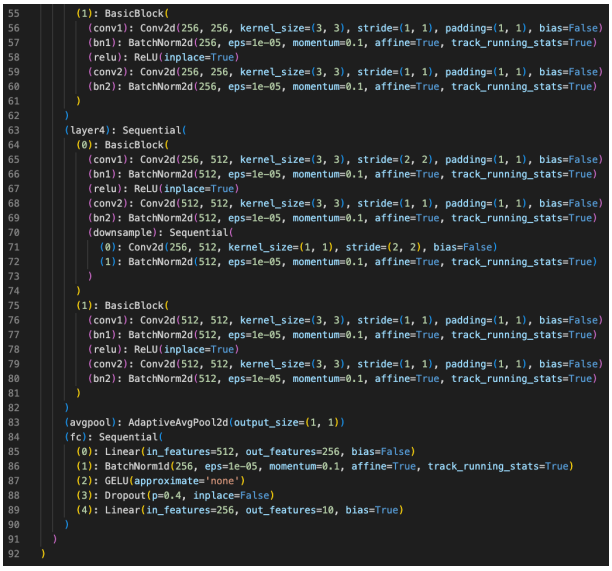
• Report accuracy of both models on the validation set. (2%)

Ans:

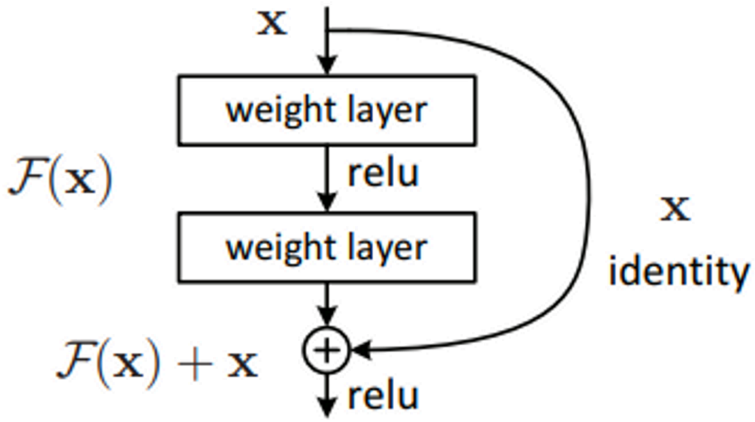
	A - mynet	B - resnet18
accuracy	0.8182	0.8824

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

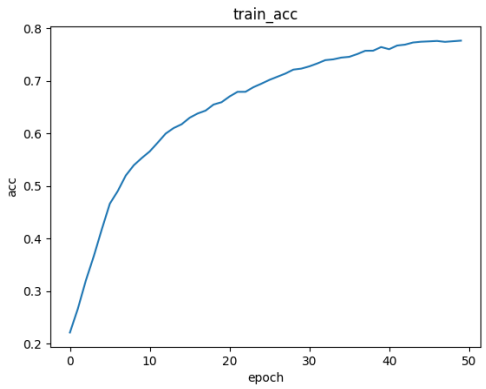
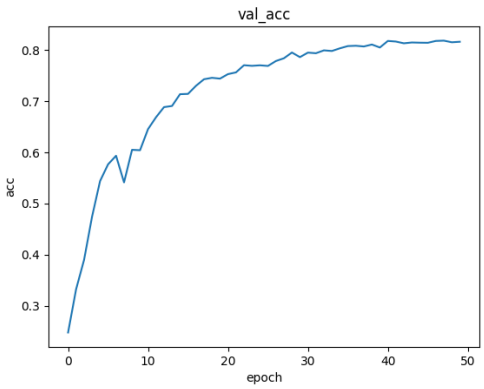
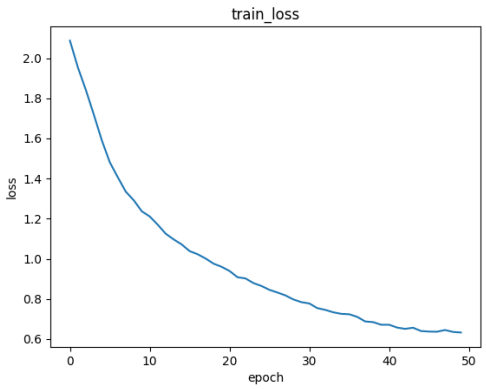
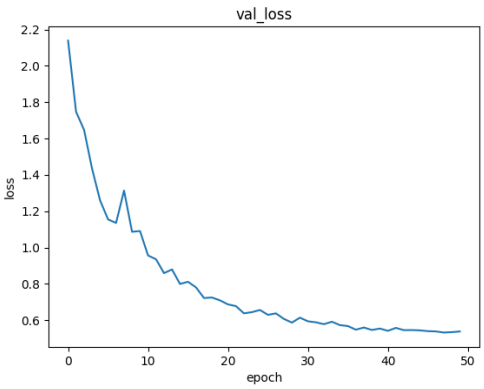
Ans:

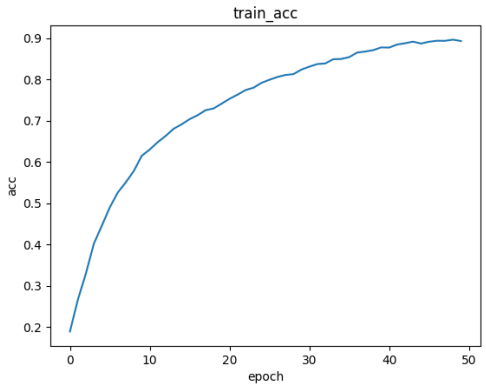
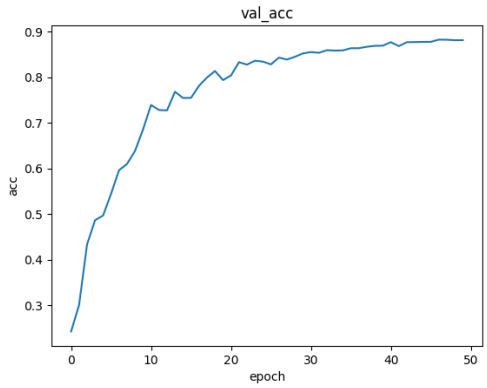
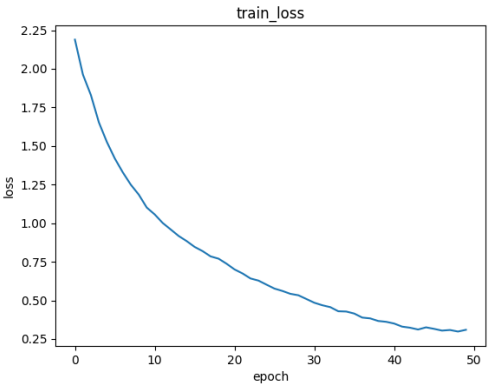
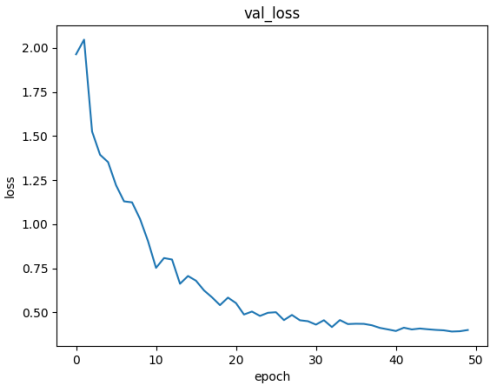
	A - mynet	B - resnet18
# of params	317,866	11,310,666
arch		 

ResNet use residual connection.
It should be easier for the solver to find the perturbations(F(x)) with reference to an identity mapping(x) than to learn the function as a new one.



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

	train	validation
A - mynet accuracy		
A - mynet loss		

B – resnet18 accuracy		
B – resnet18 loss		

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

Ans:

- **data augmentation**
  - **Random Crop**
  - **TrivialAugment (https://arxiv.org/abs/2103.10158)**
  - **RandomErasing(https://arxiv.org/abs/1708.04896)**
- **model architecture**
  - **Based on Resnet18**
  - **Remove the first maxpool layer**
  - **Modify the fc module**
- **loss function**
  - **cross entropy**