Assignment I: Computer Vision

Sun Yibo 3036047195

**Hyperparameters setting**

The following hyperparameters were changed:

* Image size – 640
* YOLO grid num – 20
* Number of epochs – 20
* Batch size – 6
* Learning rate – 4e-5

The improvement of image size from 448 to 640 is increasing the amount of information in the data. And the number of YOLO grids follows the image size, so it should be 20 instead of 14.

The loss on validation data almost stop descending after 15 epochs. (See following results)

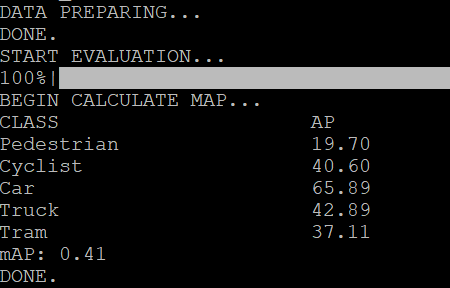
Batch size greater or equal to 8 may lead to GPU out of memory. (Because the image size became bigger)

The reason for increasing the learning rate is to rapidly reduce the loss at the initial stage of training. The step of gradient descent would be adjusted by Adam later.

**Results of experiment**

Under the above settings, the results after training are as follows.

The mAP on validation dataset is 0.41. (The loss data lost)



And the prediction results on some test data are as follows.





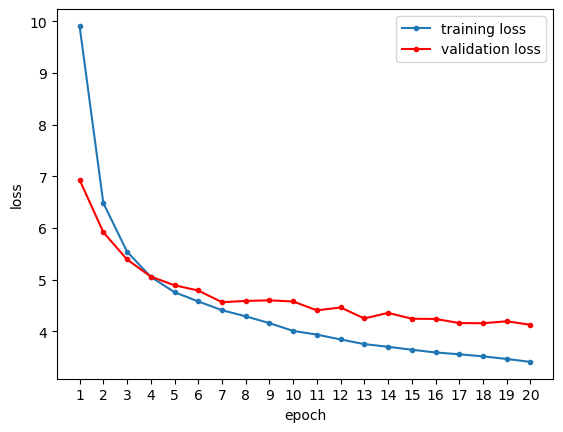
**Improvement**

During the above training, the rate of decline of loss became very slow after 5 epochs. There has even been a rebound. This may have something to do with the learning rate. The parameters may swing around the local minimum.

Therefore, I used learning rate decay to optimize the training process. Here I used Step Learning Rate scheduler.

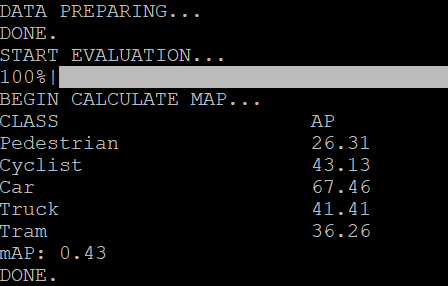
* Gamma – 0.9
* Step size - 2

Under the above adjustments, the results after training are as follows.



Though the rate of decline in training has also slowed, the loss continued to fall. And the loss on validation data became convergent after 15-17 epochs.

This time, the mAP became 0.43, which is higher than the former one.



There are also some different in prediction.





In general, the scores of each box are a little bit higher. And the size of boxes are a little bit smaller. There's less space in the boxed that doesn't contain the target.