VRDL – HW3 Report

Model link:
https://drive.google.com/drive/folders/1MnJsJZSrwpDuAWgiOJjYzKrX8v2Qgkof?usp=sharing

1 GitHub link: https://github.com/andychan8877/2021_VRDL_HW3.git

2 Reference:
- 2.1 https://github.com/cocodataset/cocoapi/blob/master/PythonAPI/pycocotools/mask.py
- 2.2 https://github.com/matterport/Mask_RCNN
- 2.3 https://github.com/matterport/Mask_RCNN/tree/master/samples/nucleus
- 2.4 https://blog.csdn.net/weixin_30718391/article/details/100001331?spm=1001.2101.3001.6661.1&utm_medium=distribute.pc_relevant_t0.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-1.no_search_link&depth_1-utm_source=distribute.pc_relevant_t0.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-1.no_search_link
- 2.5 https://blog.csdn.net/weixin_38345523/article/details/106786626
- 2.6 https://www.twblogs.net/a/5db2c49cbd9eee310d9ff8da

3 Introduction:
- 3.1 Package: I use Keras-MaskRCNN to do this homework.
- 3.2 Model: I use MaskRCNN, and backbone model pretrained Resnet-101 of Pytorch.

4 Methodology:
- 4.1 Pre-process: I adjust some parameters in the config based on nucleus sample, including setting size of image resizing to 256, RPN_ANCHOR_SCALES = (16, 32, 64, 128, 256), RPN_NMS_THRESHOLD = 0.5
- 4.2 Checkpoints: I use imagenet pre-trained model of resnet101.

5 Summary:

Since there is an example for this assignment, I mainly spent time getting more familiar with how to apply the customized data set to the package. In the process of training the model, although the loss has a gradual decline, the

performance in inference is not very good. It may be that the setting of config is still not correct, or the number of trainings is not enough. And I observed the result of inference and found that the numbers of bbox are all integers, which seems unreasonable. I'm not sure if it is a problem of resizing or training times.

In spite of the above problems, I tried to visualize the recognition results during this operation. I can see that there are results. Although the content of image recognition is more difficult, it makes me feel that I know what I am doing when I do the task of image recognition.