**《计算机视觉》线上实验/课程大作业 报告**

**(2021-2022 学年第 2 学期)**

**第三单元 图像分割**

**学生姓名：**

**提交日期：2022 年 7 月 10 日 学生签名：陈卓文**

|  |  |  |  |
| --- | --- | --- | --- |
| **学 号** | **201936380215** | **任课教师** | **彭绍武** |
| **学 院** | **软件学院** | **专业班级** | **软件工程2班** |
| **课程名称** | **《计算机视觉》** |  |  |
| **教师评语：** | | | |
| **本论文成绩评定：** **分** | | | |

**说 明**

1、本次大作业报告需要配套完成线上实验的ipynb文件一同提交

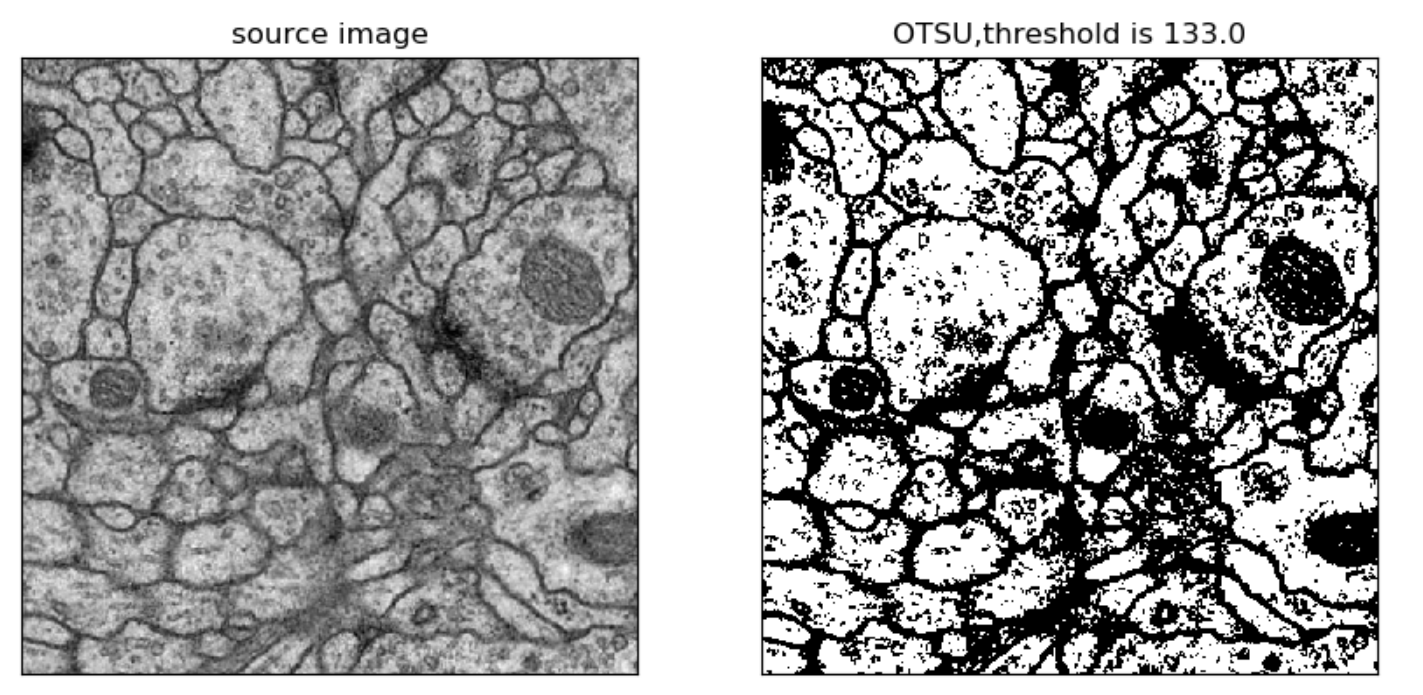
2、ipynb需要包括实验手册1~3节所对应的3个语义分割实验和4、5节对应的实例分割实验，包括各实验的中间结果，最后结果，以及开放题的实现与结果。

3、本课程大作业报告，重点摘录ipynb笔记中3个语义分割实验和一个实例分割实验的运行情况，结果分析，以及各自开放性问题。

4、针对开放性问题，大津阈值法和UNet部分需要按小题做对应的实现与分析说明，Mask-RCNN需要完整阐述所用模型或网络方法的修改方案，具体的参数设置，训练及测试的配置情况，数据使用情况，实验结果与分析评价。

6、请记录实验完成的费用情况。如果费用不足，请参照第二、第三单元的实验手册，部署本地mindspore运行。

1. **实验一 复现大津阈值法的语义分割**

****

1. **大津阈值法思考题**

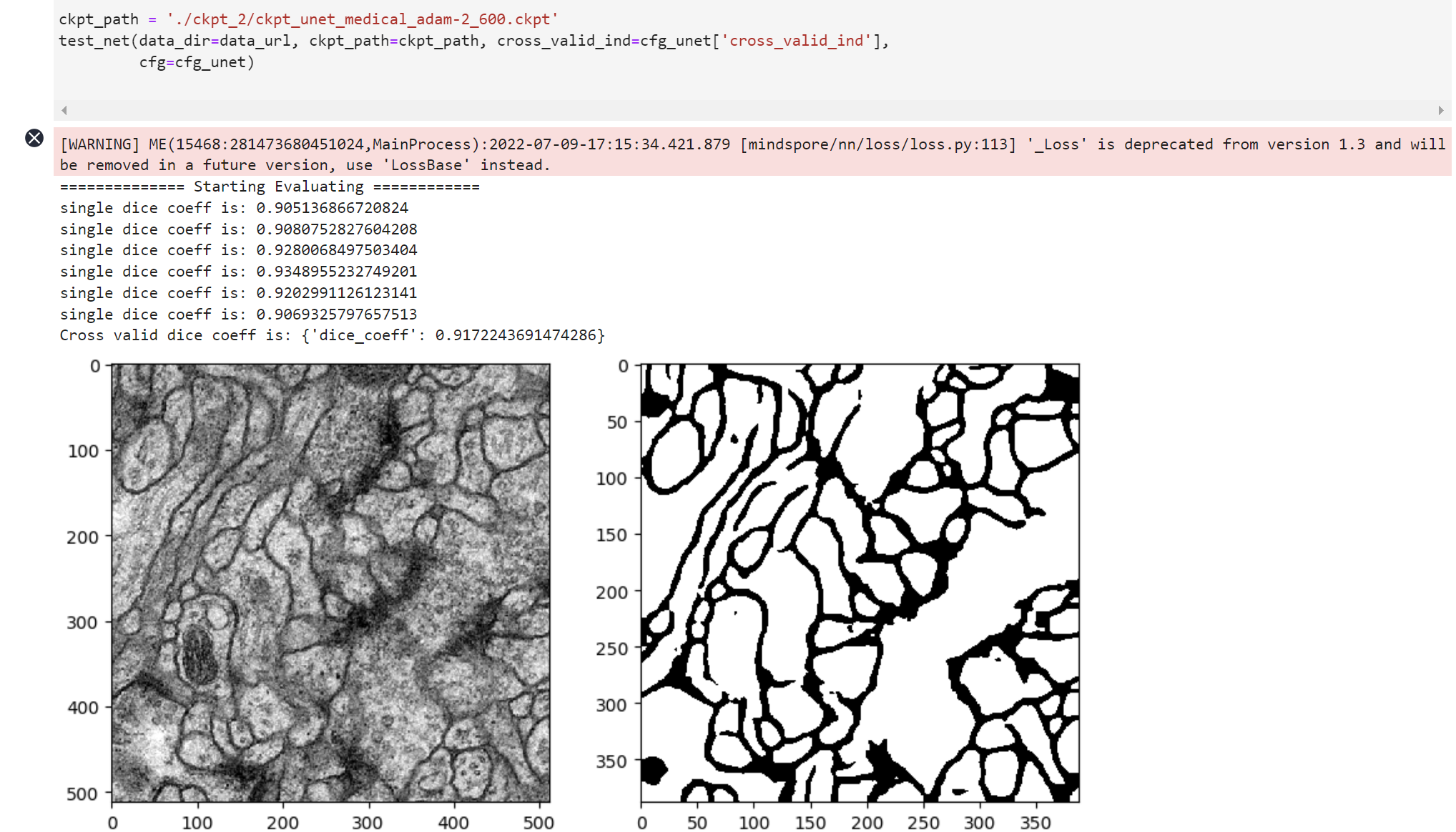
**1：请用OpenCV和matplotlib取出一张训练集中的图像，并绘制图像直方图。**

****

**2：请用OpenCV中的python接口实现基于大津阈值法的图像二值化分割。**

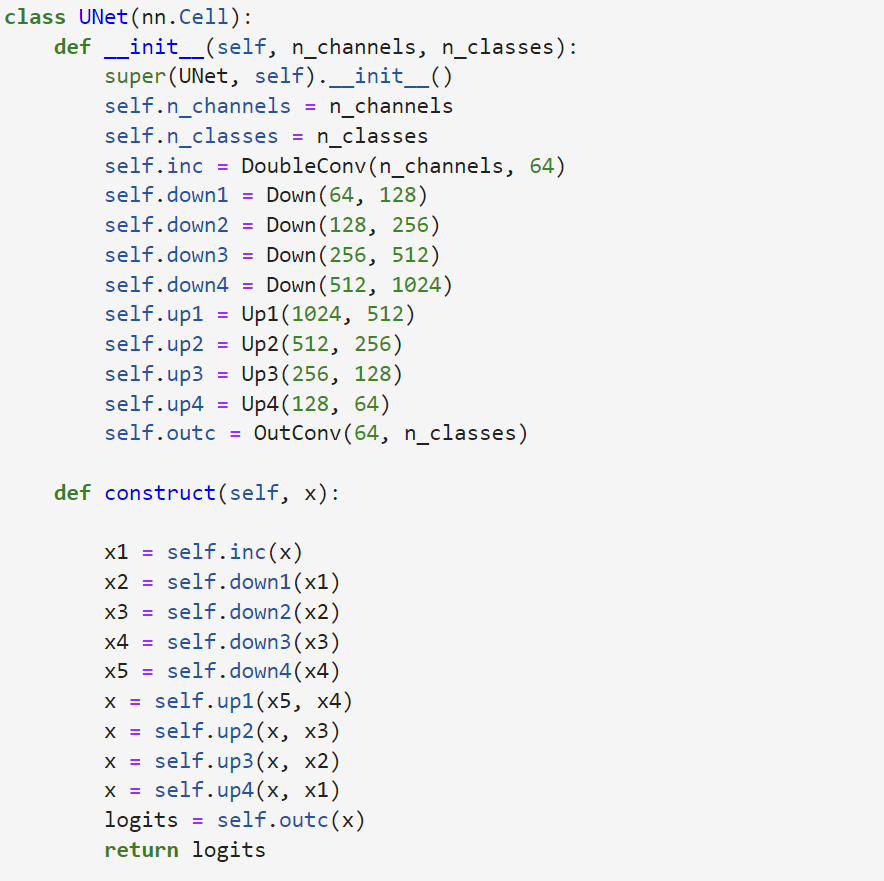
****

1. **实验二 复现UNet的语义分割**

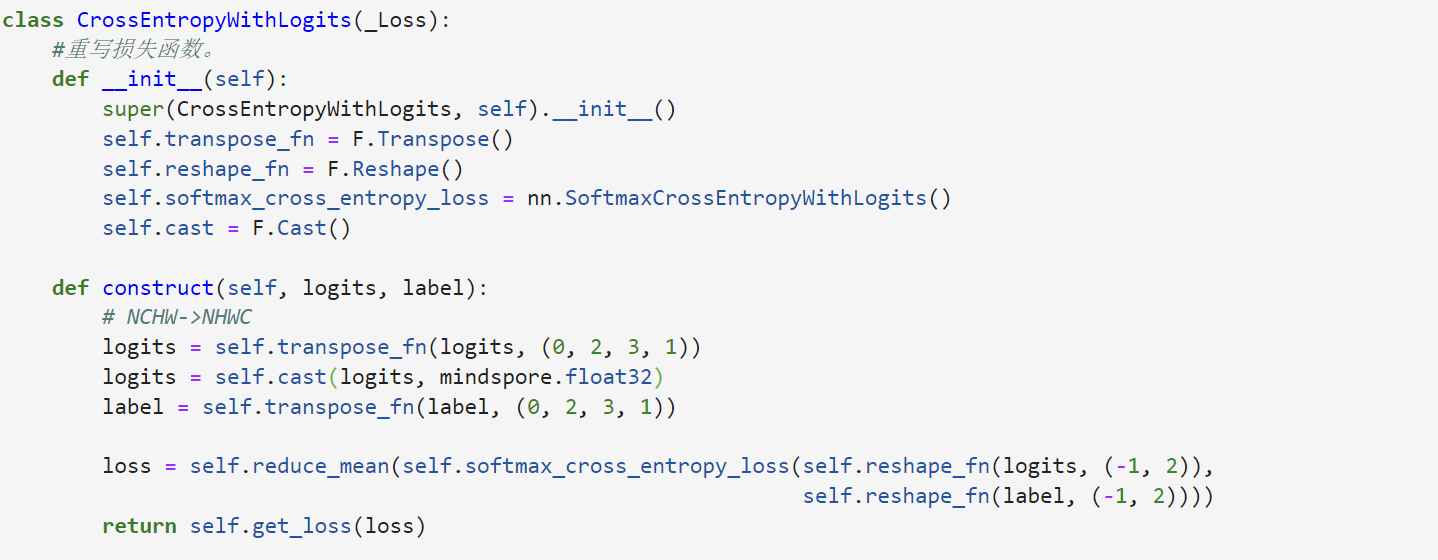
****

1. **UNet方法思考题**

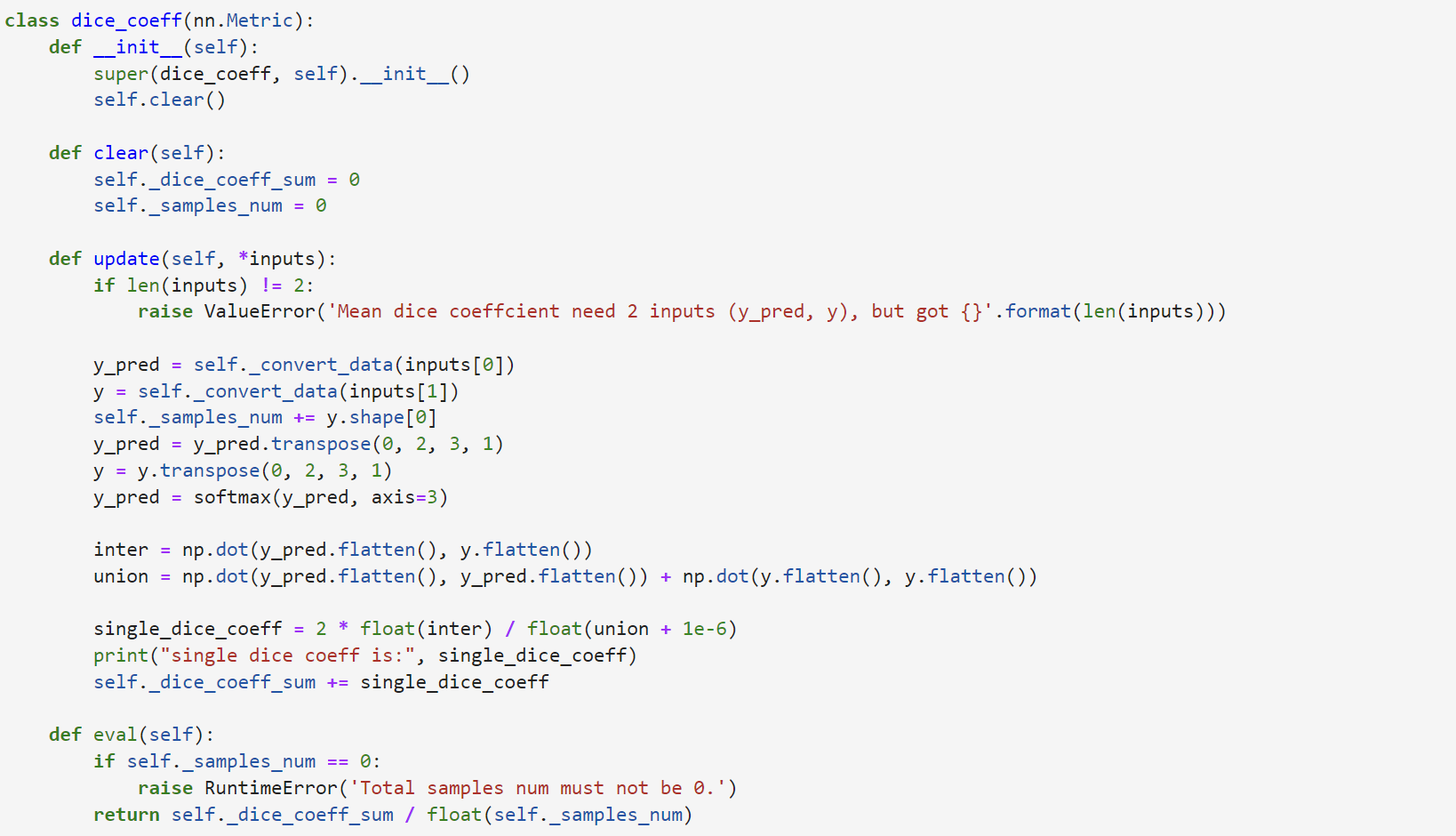
**3：使用MindSpore搭建Unet类，用于构建网络。要求：网络结构和原论文保持一致。**

****

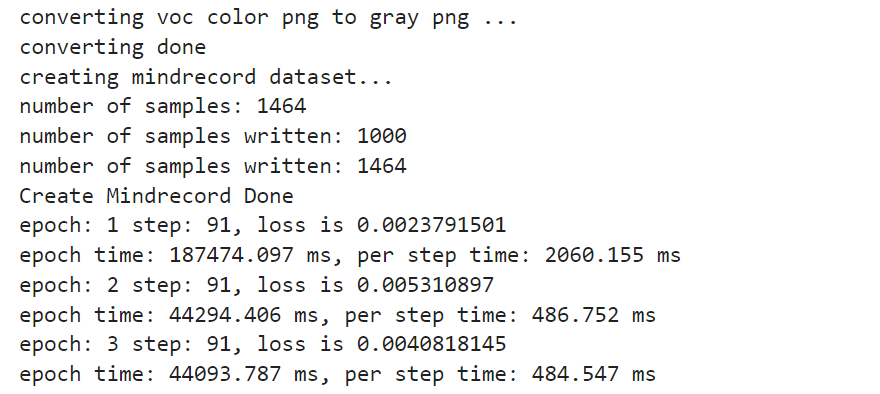
**4：请重新构建一个类，将MindSpore的nn.SoftmaxCrossEntropyWithLogits损失函数用于Unet，计算输出特征图各个位置平均的损失值。**

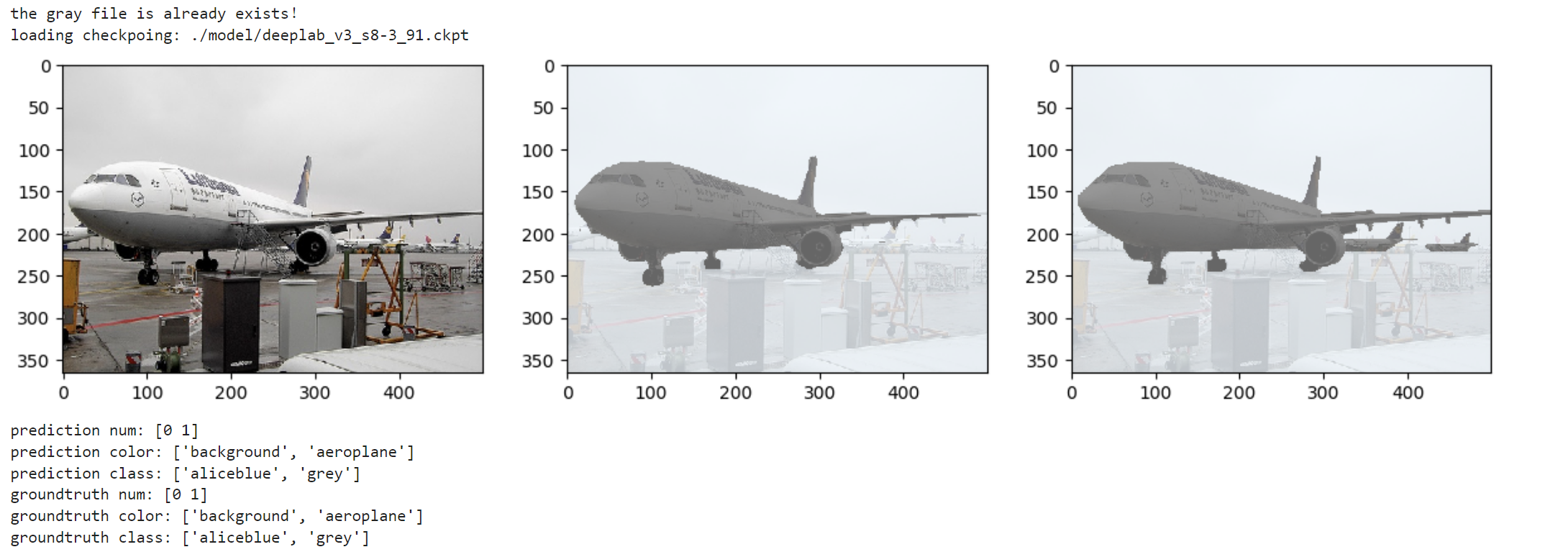
****

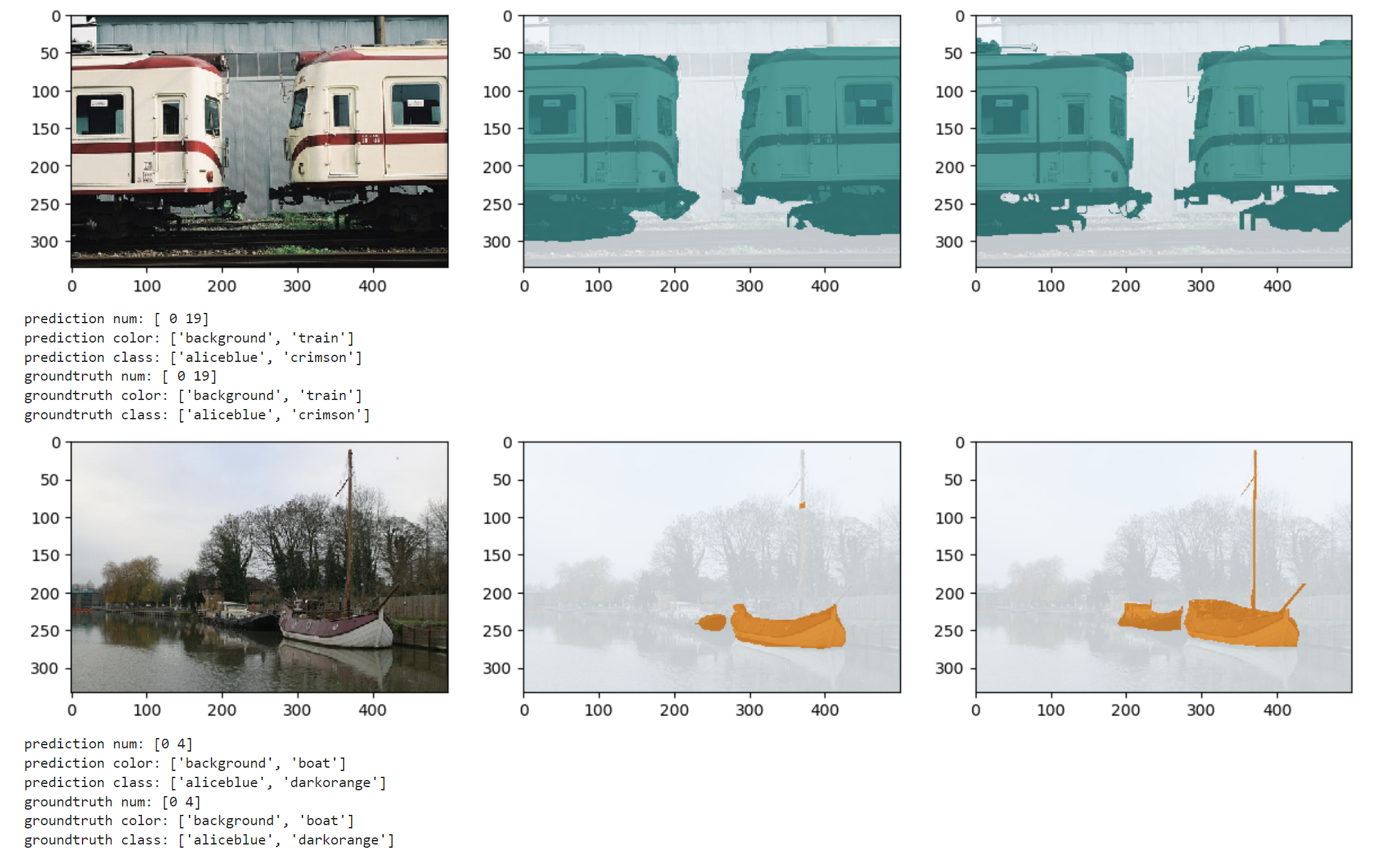
**5: 定义一个名为dice\_coeff的类，用于计算每张验证集图像的Dice以及返回验证集中Dice的均值。**

****

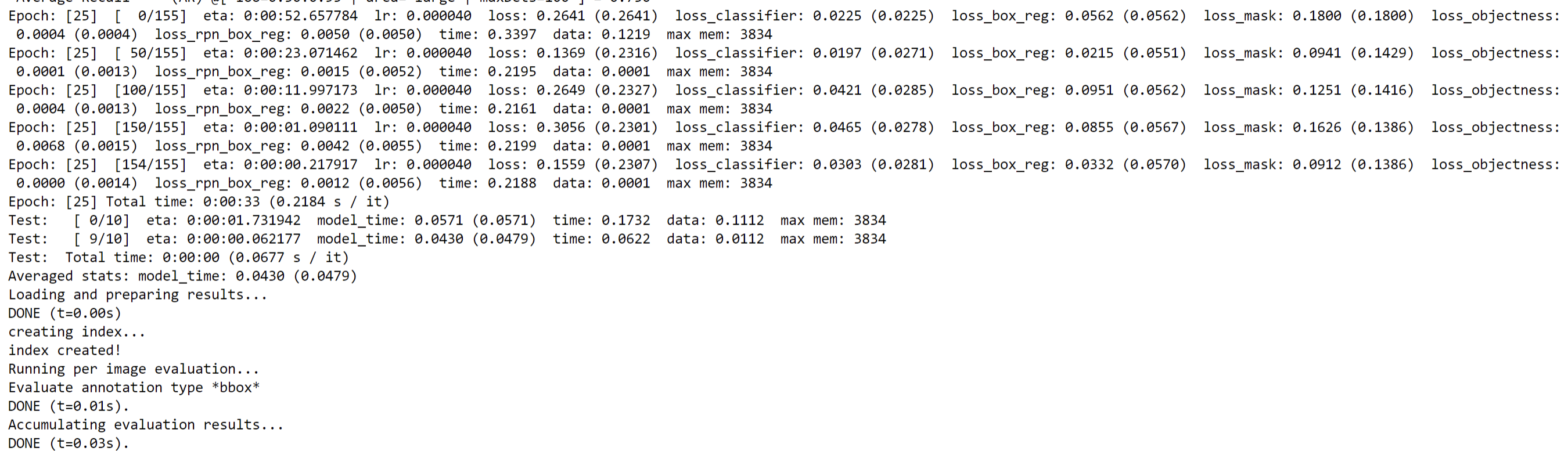
1. **实验三 复现DeepLabv3的语义分割**

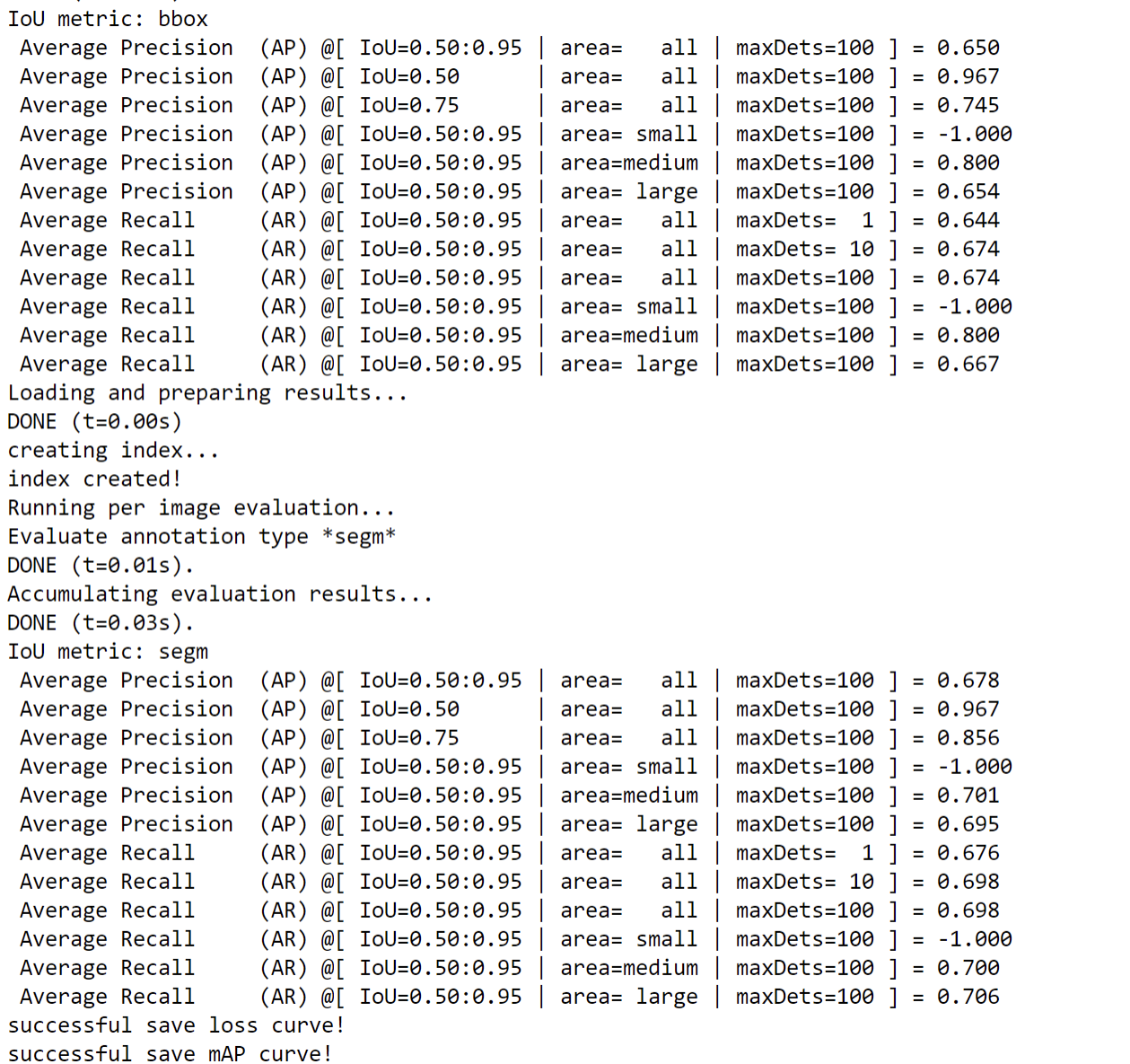
****

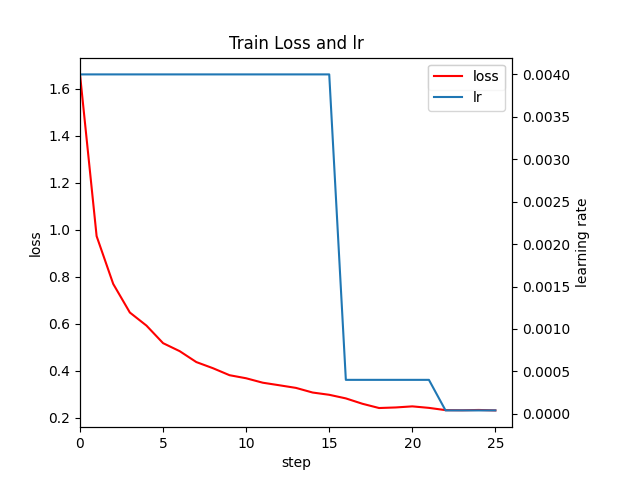
****

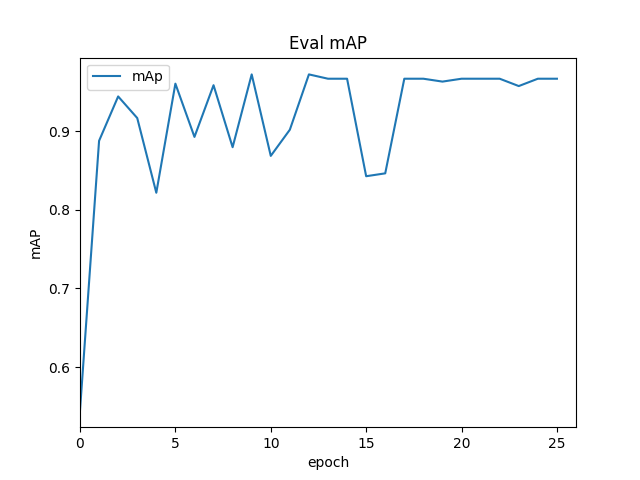
****

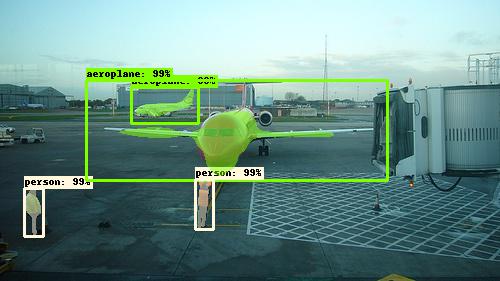
1. **实验四 复现Mask-RCNN的实例分割**

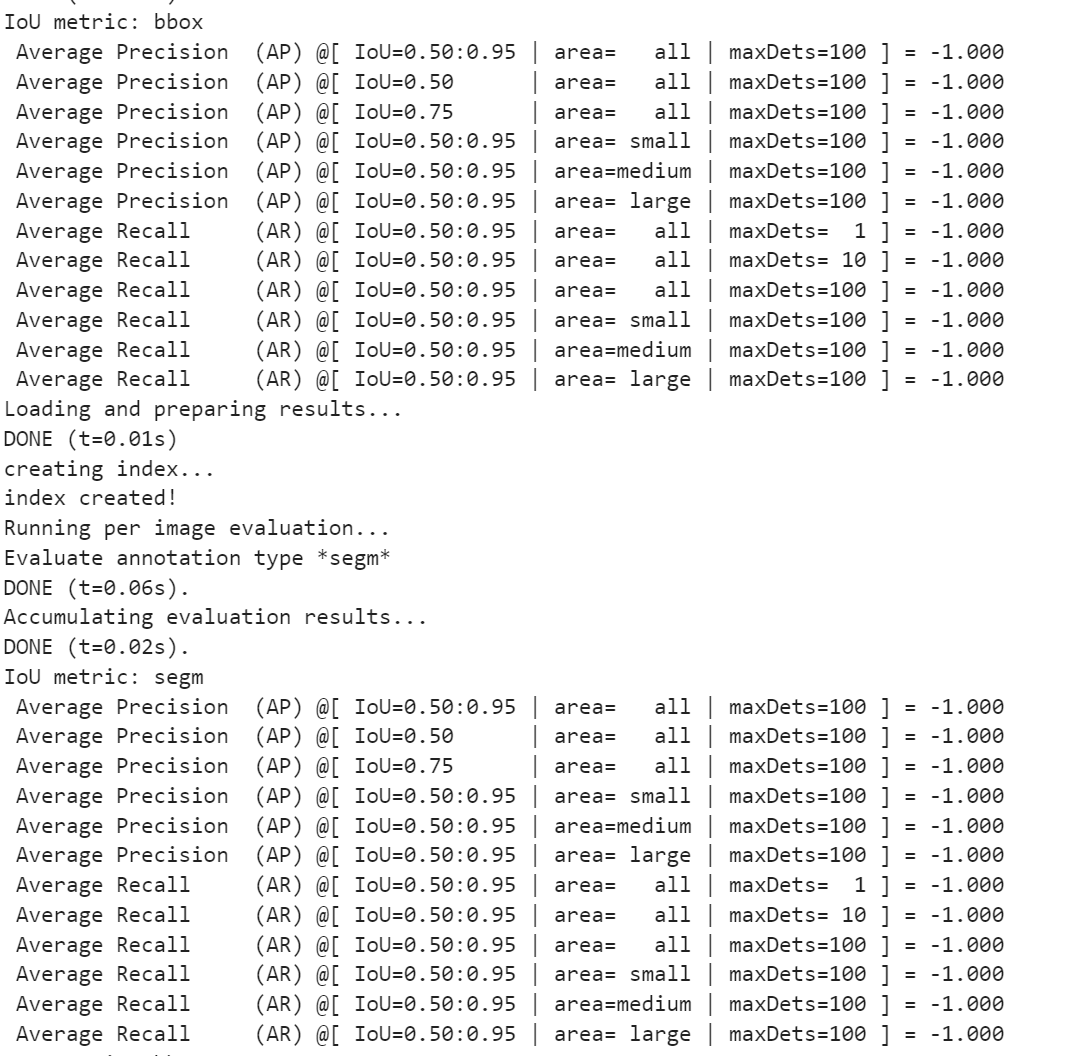
****

****

****

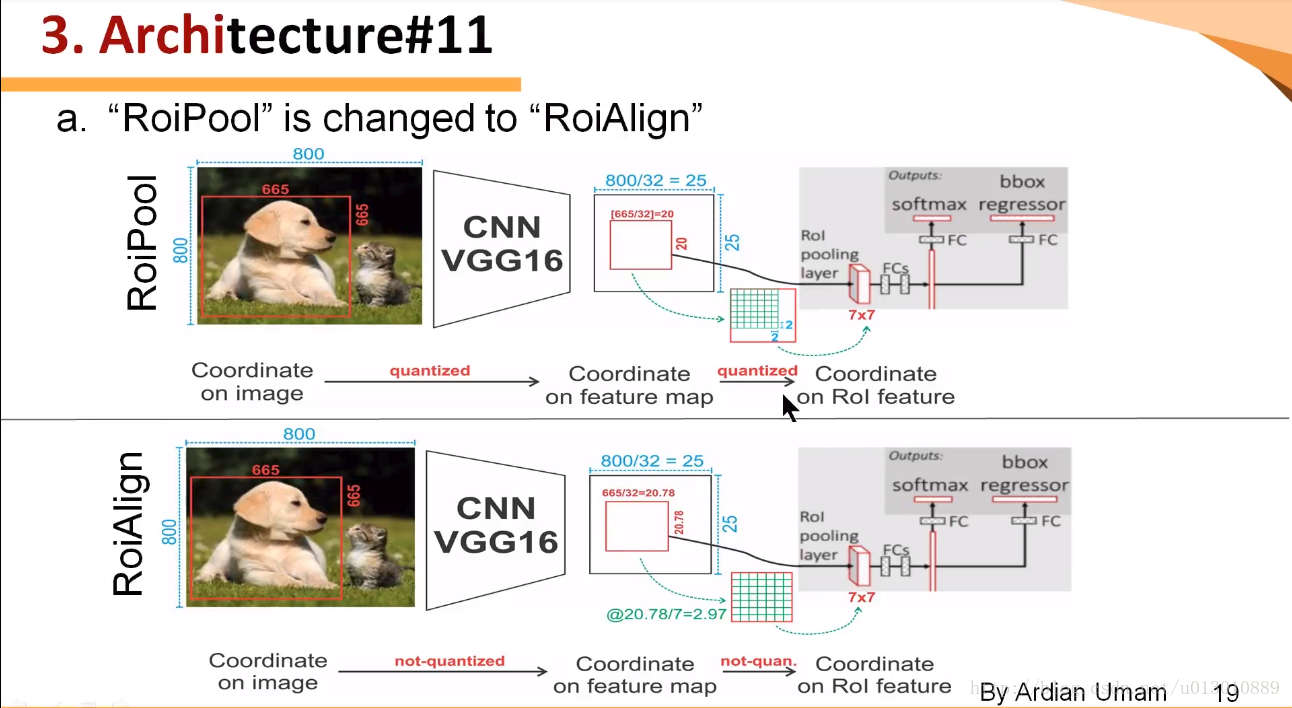
****

** **

****

1. **Mask-RCNN开放题**

**6：Mask-rcnn是在目标检测算法faster-rcnn的基础上进行改进，其中改进的一点为ROIpooling层改为了ROIAlign，请在Mask-rcnn代码上分别使用ROIAlign和ROIpooling，比较两者的效果，并解释出现这种状况的原因。**



RoiPool：

1.通过VGG16网络，对原图尺寸为800×800进行卷积池化的操作，最终生成25×25；

2.其中有一张665×665的图片进行同样的卷积池化操作，理论上产生20.78×20.78的特征图，但计算机计算的时候，自动取整，生成了20×20的操作（第一次量化）；

3.通过ROI pooling层，需要将所有的图片卷积成7×7的图片时，理论上要生成7×7的图片，20/7=2.86，以2.86×2.86形成49个格子，然后每个格子里取最大值（以最大池化为例子），但2.86无法实现，进行了 第二次量化 ，以2×2为一个小格子；

最终得到7×7的特征图，但经过了两次池化，对于分类效果来说，并不会影响，但是对于要在原图中定位目标的情况下，会产生较大的误差。

RoiAlign：

1.同样VGG16，有图片800卷积池化成25 ×25的特征图；

2.有一张665×665的图片，其中有ROI经过卷积池化后，变为20.78×20.78，不取整数；

3.最终生成的是7×7的特征图，以20.78/7=2.97 2.97×2.97为一个小格子，生成49个同等大小的小格子；

4.每个小格子的四个角如果没落到真实点上，就对最近的真实点采用双线性插值法，取虚拟点位；

5.然后在每个小格子以2为采样样本，再分成四份，取每一小份的中心点，求得该点的值，最后每四份取最大值。这样对于ROI pooling来说，去掉了两次量化的误差，使得边框精度得到极大的提升。

所以RoiAlign能够提供更高的精度，可以极大的减弱偏移现象。

**7：思考什么方法或技巧可以提升网络的分割效果，选取其中一种或两种进行验证并出示验证结果，对结果的提升原因作出解释。**

* 1. **方法介绍**

增加网络的训练轮数，使得模型得到更好的训练。

* 1. **实现**

使用以下命令进行训练。

python3.7 train.py --resume save\_weights/model\_25.pth --start\_epoch 25 --epochs 61

* 1. **数据、参数与设置说明**

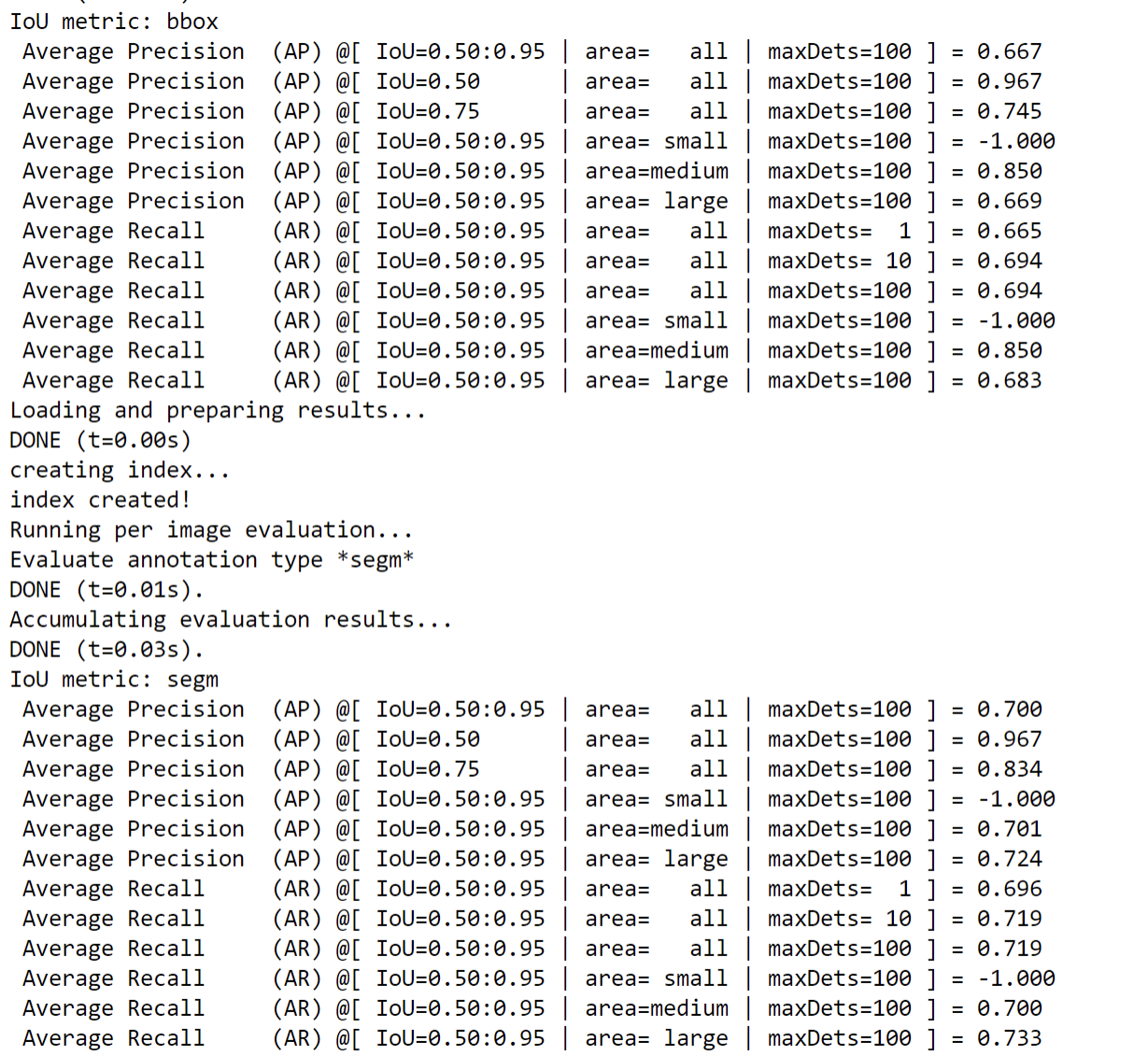
同前述复现实验相同。

* 1. **实验结果及分析**

使用以下命令进行评估：

python3.7 validation.py --weights-path ./save\_weights/model\_60.pth

得到评估结果：



可以看出，模型分割效果得到了一定的提升，但提升较为有限。

**补充：请在此记录本次大作业完成时的机时券剩余金额：￥ 679.34 元**