

物数融合智能算法挑战赛

赛题六

自动驾驶场景的2D语义分割

2024/08/08-2024/08/09



智能汽车安全技术
全国重点实验室

State Key Laboratory of
Intelligent Vehicle Safety Technology



重庆中科汽车软件创新中心

AUTOMOTIVE SOFTWARE INNOVATION CENTER CHONGQING



中国科学院软件研究所

Institute of Software Chinese Academy of Sciences



广州市智能软件产业研究院

Institute of Intelligent Software, Guangzhou



深信科创

SYNKROTRON



西南大学

计算机与信息科学学院 软件学院

COLLEGE OF COMPUTER AND INFORMATION SCIENCE, SOUTHWEST UNIVERSITY
SCHOOL OF SOFTWARE, SOUTHWEST UNIVERSITY

目录 CONTENTS

01. 赛题介绍

02. 团队介绍

03. 2D语义分割

04. 3D目标检测

05. 总结与展望

1. 赛题介绍

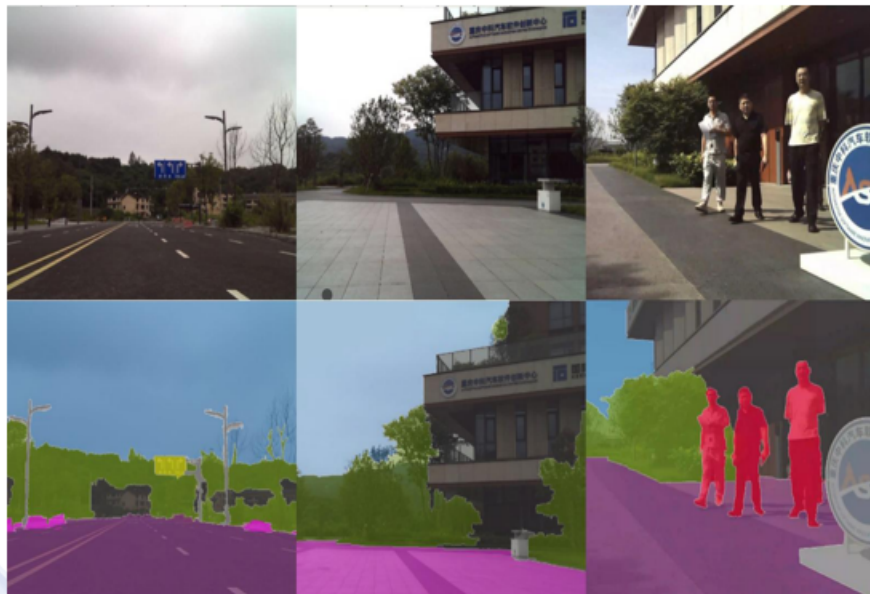
1. 赛题介绍

自动驾驶场景 2D 语义分割

任务描述：自动驾驶系统在进行环境感知过程中需要检测和识别出不同目标的类别和位置。在此任务中，需对给定输入图像，预测每个像素的类别。

评分指标： Average IoU ($\text{IoU} = \text{TP}/(\text{TP}+\text{FP}+\text{FN})$)

IoU (Intersection over Union) TP (True Positive) FP (False Positive) FN (False Negative)



2. 团队介绍

2. 团队介绍

怀揣自驾梦的超龄小孩队



杨弘毅

学校: 西南大学

专业: 计算机科学与技术

负责任务: 模型调研, 模型训练, IsaacSim调研及部署, PPT制作与演讲



罗兰茜

学校: 西南大学

专业: 智能科学与技术

负责任务: 数据采集, 数据打标, 实车调试, 模型调研, 模型部署



孙喆焱

学校: 西南大学

专业: 网络工程

负责任务: 数据打标, 模型调研, 模型训练, PPT制作

3.2D 语义分割

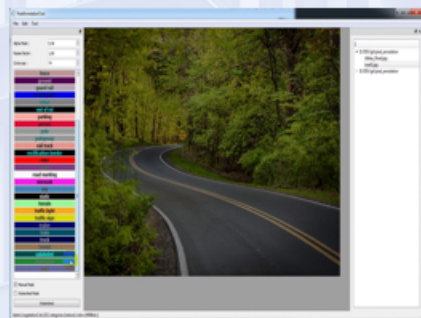
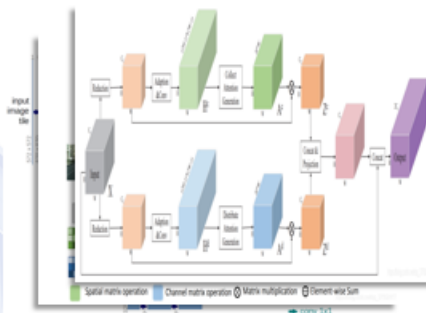
3. 2D语义分割

3.1 流程图



HONOR 90 Pro

Zhen 91.4 17166 8036



3. 2D语义分割

3.2 图片采集

团队将数据集按照

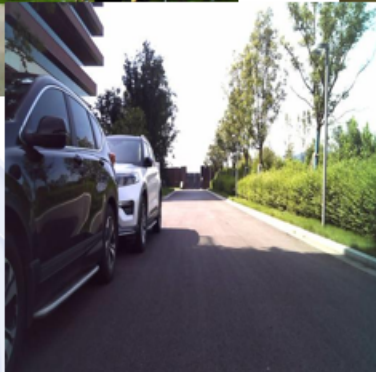
- 天气
- 行人
- 路况: 园区 公路 农村小道等外部环境分为不同难度, 采集大量样本后筛选出相对优质图片组成数据集最终版本。



天气



行人



路况

3. 2D语义分割

3.3 图片标记

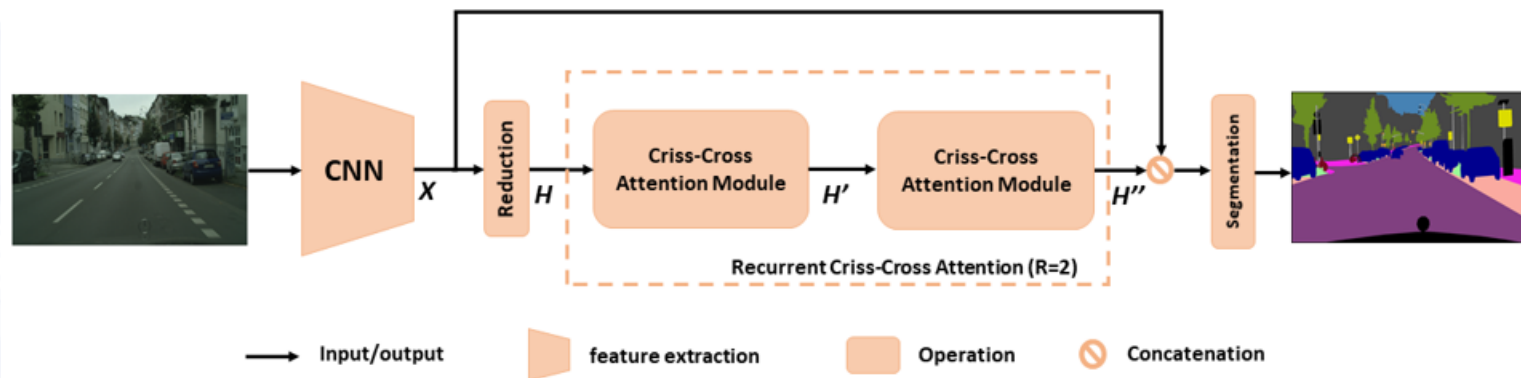
团队采用PixelAnnotationTool软件进行图片标记。2D语义分割对标签精度要求较高，工作量较大。



3. 2D语义分割

3.4 模型调研

团队采用mmsegmentation框架。该框架集成了40种2D语义分割模型，方便对比训练效果，从而找到表现效果最好的模型，提高实车识别精度。



CCnet模型示意图

3. 2D语义分割

3.5 模型适配

原始标签图片并不符合模型训练要求。首先，团队将标签图片转为**灰度图像**；其次，团队将每一个待训练模型的**配置文件**都进行了**修改**，确保模型能够读取到自定义数据集。



```
# Copyright (c) OpenMMLab. All rights reserved.
```

```
from mmdet.datasets import CocoDataset  
from mmdet.datasets import CocoDataset
```

对应标签类别

```
classes_exp = ('background', 'road', 'road marks', 'vegetation',  
'concrete', 'pedestrian', 'water',  
, 'glass')
```

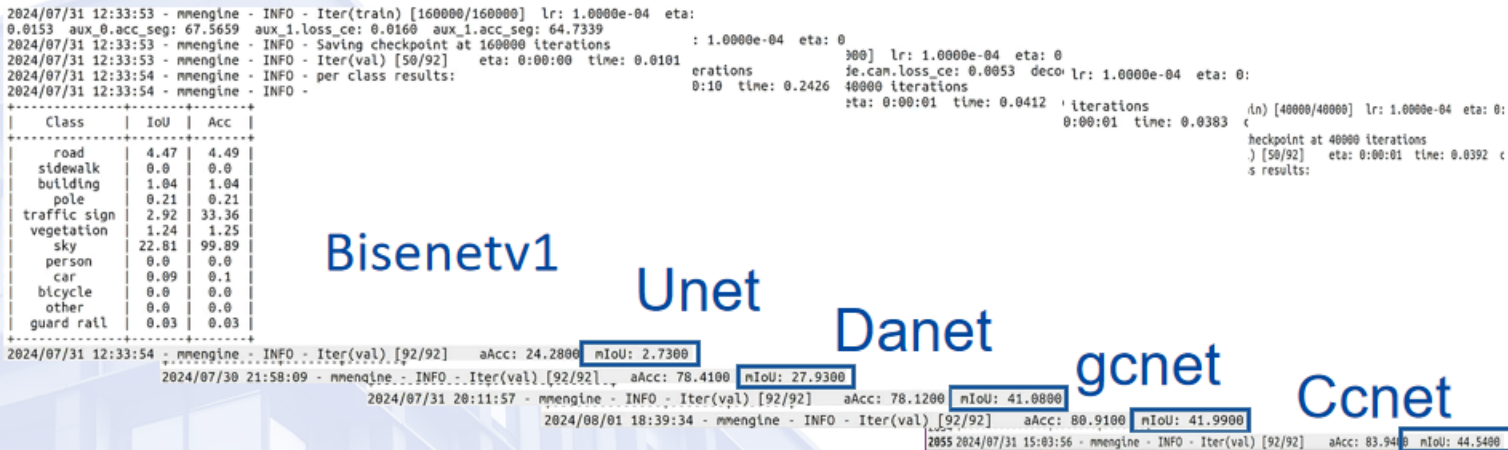
对应标签颜色

```
palette_exp = [[0, 0, 0], [77, 77, 77], [255, 255, 255], [0, 255, 0],  
[255, 0, 0], [0, 0, 255], [102, 51, 0], [255, 255, 0],  
[0, 207, 250], [255, 166, 0], [0, 204, 204]]
```

3. 2D语义分割

3.6 模型训练

在比赛时间内，团队一共训练了十余种2D语义分割模型，大部分模型都取得了不错的效果，IoU指标都达到了较好的预期。



3. 2D语义分割

3.6 模型训练

在这其中，表现最好的是Ccnet，IoU指标为44.54，在各单项指标上也几乎处于领先地位。

其他模型在某些指标也能有相对突出表现，如Ccnet在识别Building上就具有最好的效果。

Class \ Model	road	sidewalk	Building	pole	traffic sign	vegetation	sky	person	car	bicycle	other	guard rail	mIoU
Ann	71.78	25.64	61.22	16.58	30.09	65.88	83.63	34.27	48.83	0.0	1.81	63.28	40.18
Beit	43.62	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.64
Bisenetv1	4.47	0.0	1.04	0.21	2.92	1.24	22.81	0.0	0.09	0.0	0.0	0.03	2.73
Ccnet	76.95	27.03	56.83	16.64	25.68	68.94	87.25	41.84	60.56	0.0	2.59	70.2	44.54
Danet	66.1	20.5	58.28	16.07	20.24	59.78	85.38	41.59	56.62	0.0	1.32	67.04	41.08
Fcn	69.74	7.53	45.86	13.44	35.71	68.47	73.85	28.13	39.45	0.0	0.98	41.38	35.38
Gcnet	71.13	15.4	62.6	14.85	28.03	58.94	84.75	43.65	57.56	0.0	1.23	65.81	41.99
Isanet	72.6	23.25	50.17	11.16	31.25	64.67	84.13	6.53	45.23	0.0	0.0	55.86	37.07
Unet	64.7	15.06	40.68	0.03	0.0	51.83	84.36	0.0	6.34	0.0	0.0	55.44	26.54
Convnext	51.23	0.0	34.4	5.97	34.42	22.09	34.68	10.25	34.02	0.0	0.0	0.19	18.94
Deeplabv3	69.87	16.19	35.72	0.0	0.0	64.61	85.33	0.0	8.98	0.0	0.0	54.52	27.93

3. 2D语义分割

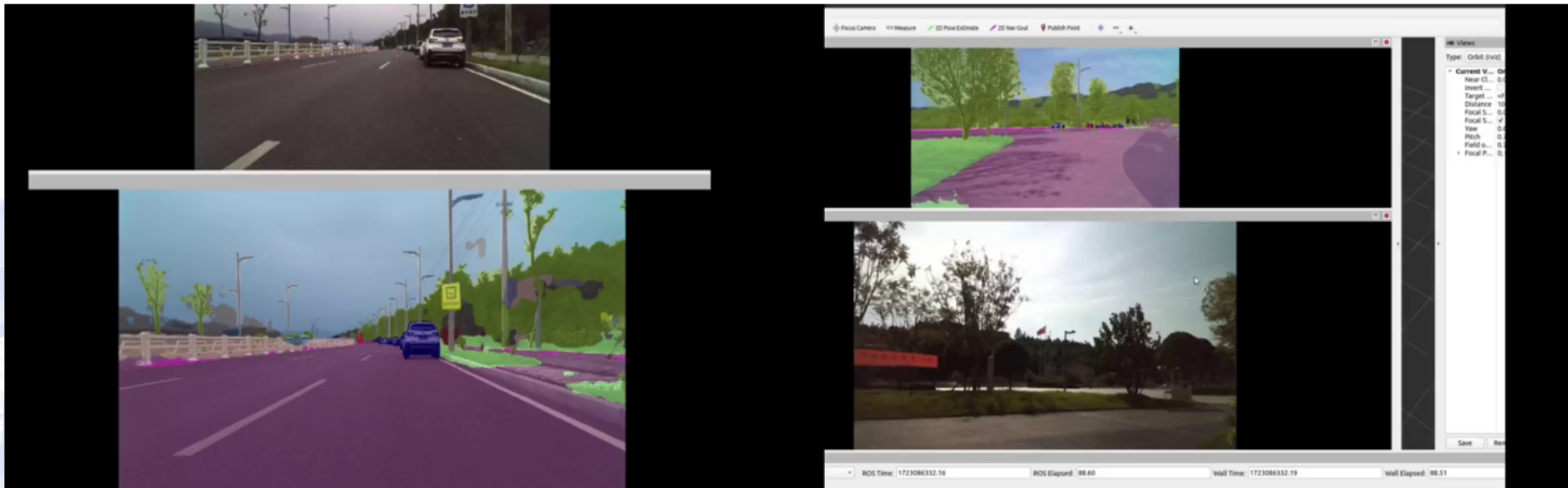
3.7.1 模型部署

团队将训练模型与ros进行交互，通过launch写入节点，采用多机通讯的方式，在rviz中进行可视化展示。



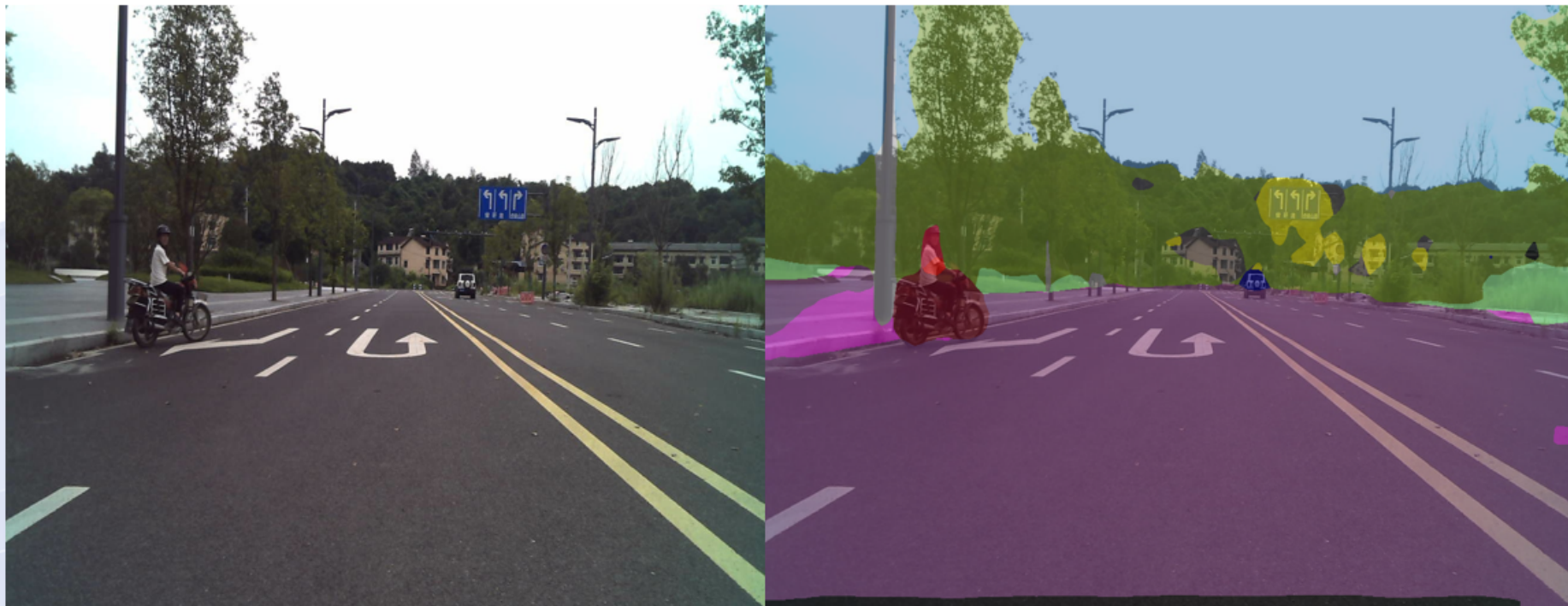
3. 2D语义分割

3.7.2 效果展示



3. 2D语义分割

3.8 反思



度。

3. 2D语义分割

3.9 优化改进——采用官方数据集

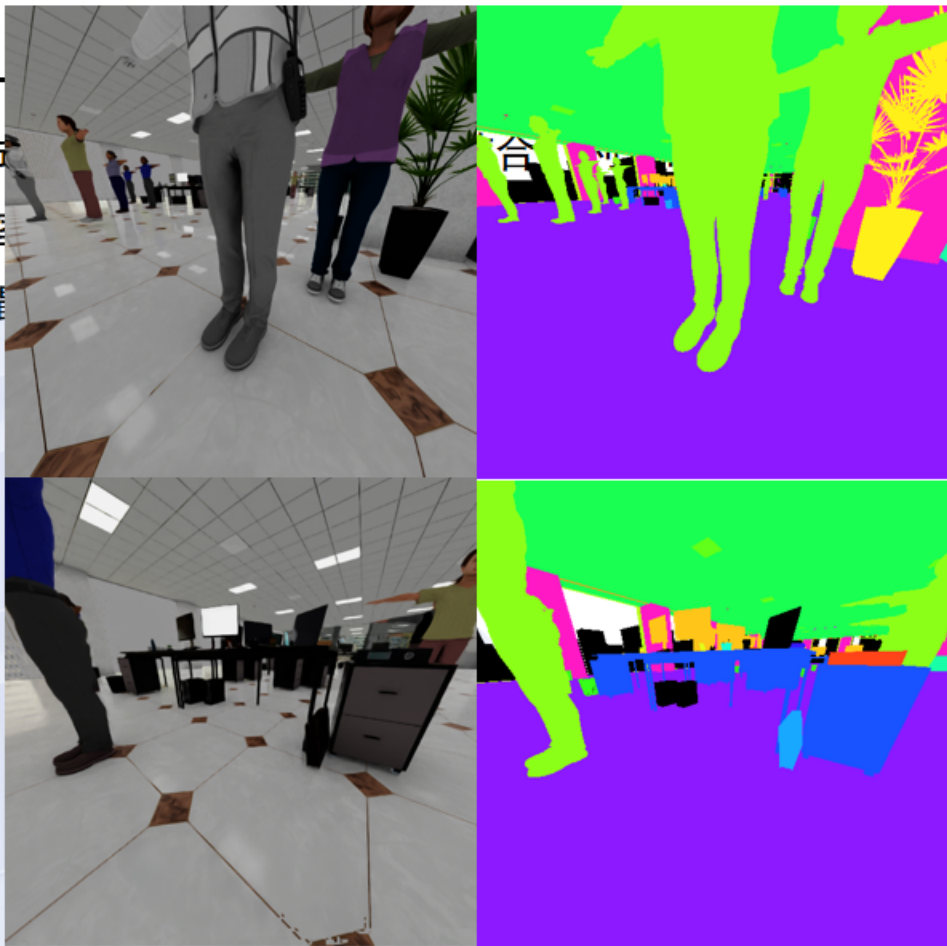
为了优化模型部署识别效果，团队尝试采用与自采数据集类似的**官方街景数据集Cityscapes**，并用其训练表现最好的五个模型。结果显示**IoU得到提升**，这说明自采数据集的精度问题一定程度上影响了模型训练效果。

Class Model	road	sidewalk	building	pole	traffic sign	vegetation	sky	person	car	bicycle	other	guard rail	mIoU
Ann	84.94	50.21	66.77	20.39	22.75	77.99	54.74	44.54	78.1	40.94	67.14	0.0	50.71
Ccnet	84.34	52.26	71.08	23.22	31.03	80.48	64.2	49.25	79.52	42.09	64.52	0.0	50.71
Danet	84.08	51.09	71.25	23.11	28.86	79.67	60.53	45.7	79.51	46.39	63.15	0.0	52.77
Gcnet	85.34	52.78	71.91	23.94	29.62	80.39	68.33	46.01	80.45	42.79	64.41	0.0	53.83
Isanet	85.26	48.6	70.07	21.57	25.79	80.55	69.44	48.32	77.02	41.93	70.99	0.0	53.30

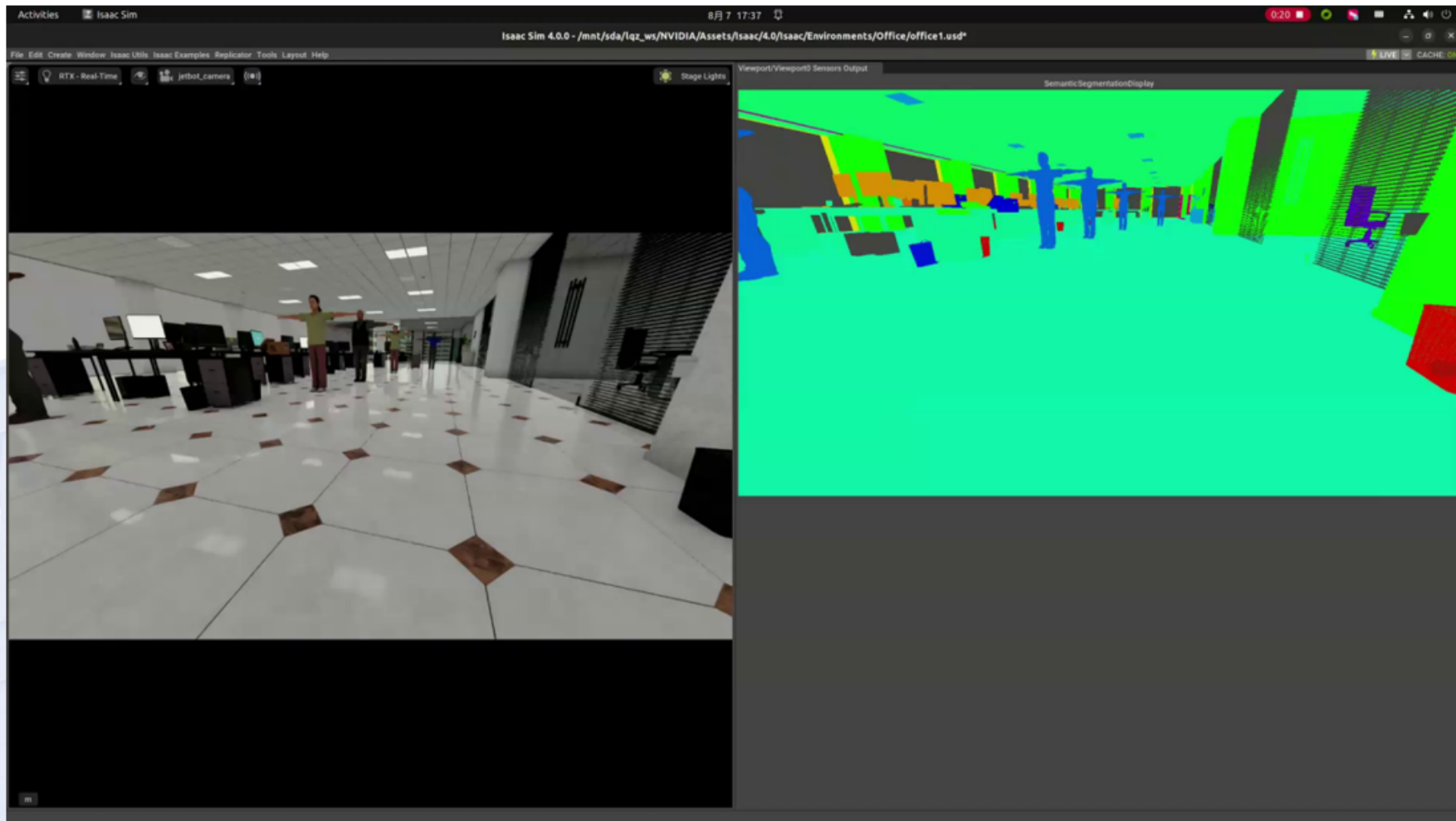
3. 2D语义分割

3.9 优化改进一

团队还尝试
了一个逼真的3D
数据集训练，最



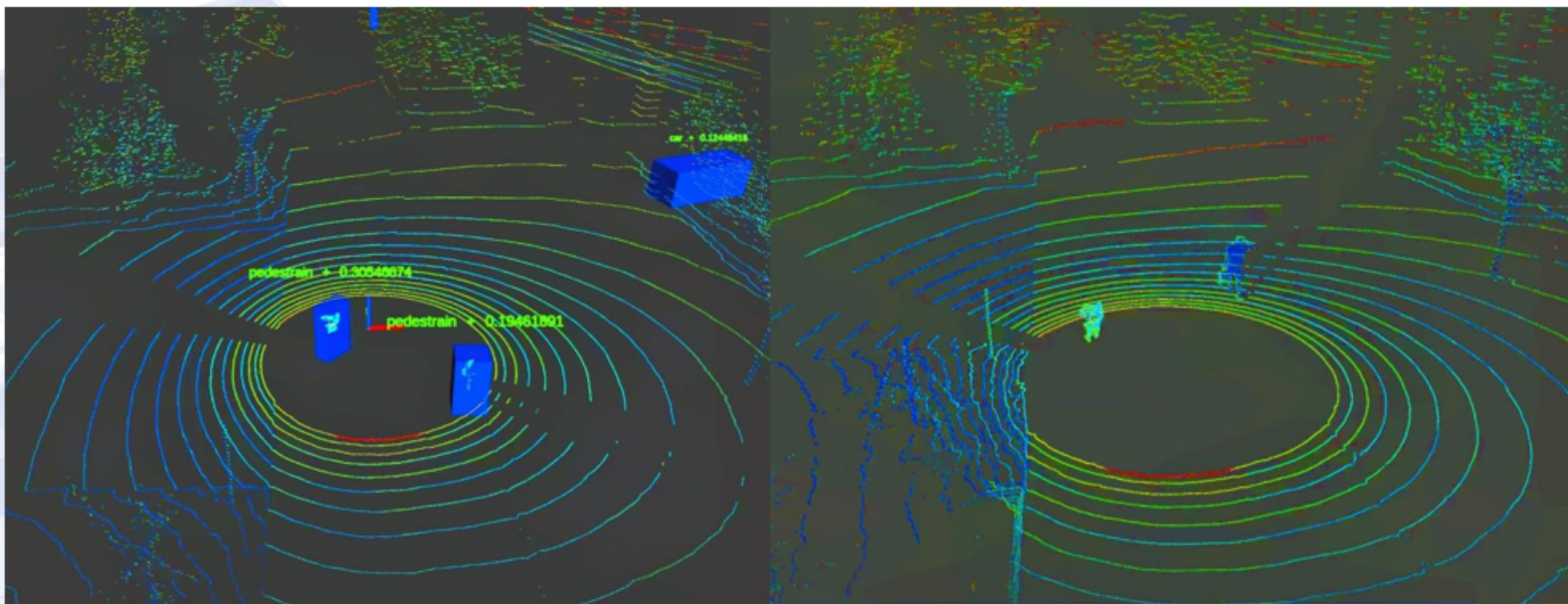
机器人模拟平台IsaacSim，团队搭建
的语义分割图像采集，整理生成



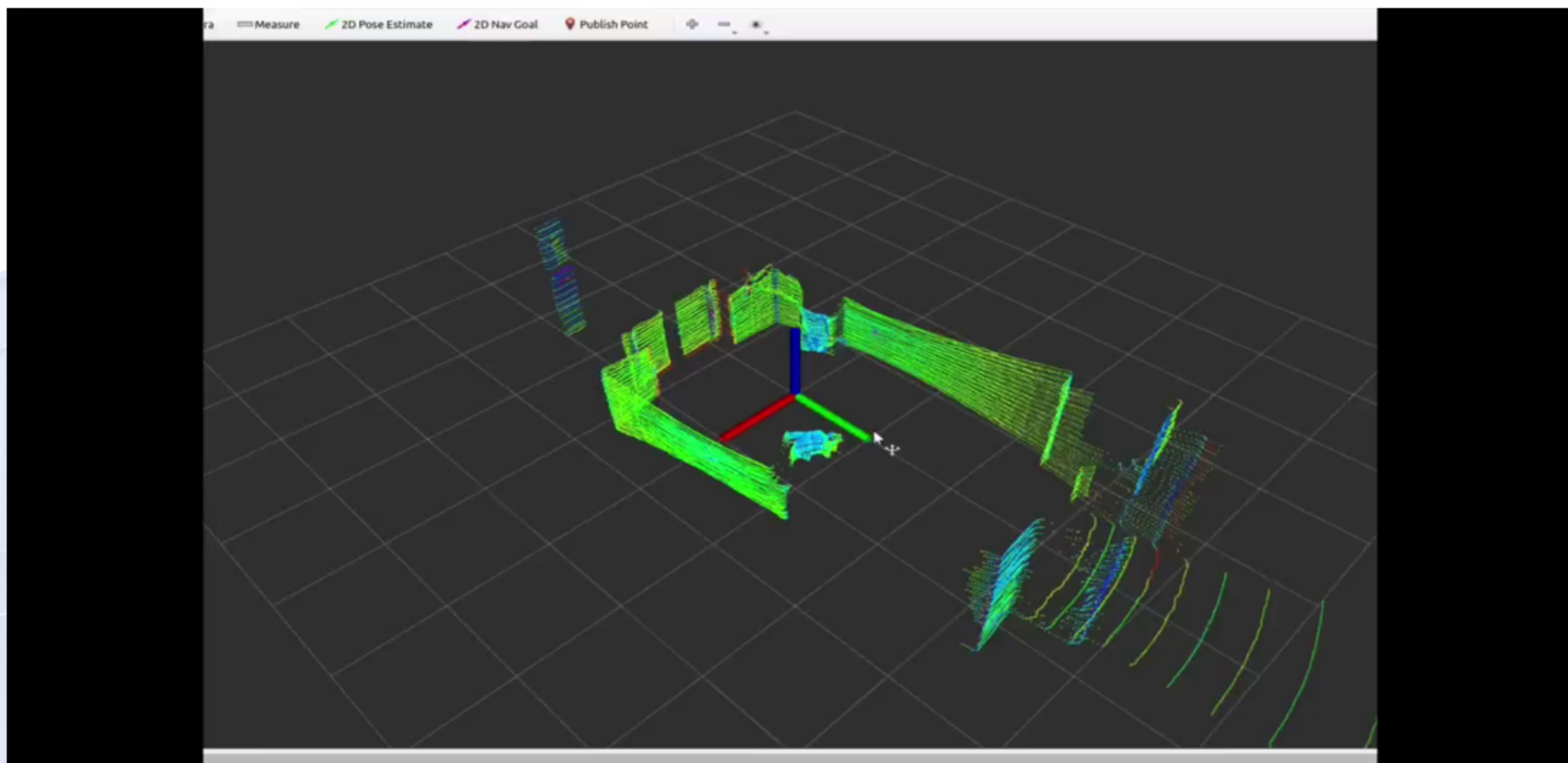
4. 3D 目标检测

4. 3D目标检测

团队还尝试突破二维的维度限制，实现更优的视觉识别效果。团队采用OpenPCDet框架下的3D目标检测模型，该模型通过雷达对物体进行立体建模，给实车提供了更多路况信息，有助于实车在应用过程中提高避障能力。



4. 3D目标检测



5. 总结与展望

5. 总结与展望

5.1 总结

团队在将近一个月的比赛时间内，基本完成了比赛要求的各项指标，包括但不限于：

- 采集图像，整理数据集
- 模型的学习、配置与训练
- 实车部署
- 模型的对比与优化

此外，团队还针对模型的改进与应用做了更进一步的探究：

- 采用官方数据集训练，尝试提升部署效果
- 采用IsaacSim合成数据集训练，通过仿真平台减小数据集收集难度，提高精度
- 尝试学习、弥合Sim2Real的gap

5.总结与展望

5.2 展望

展望未来，团队期望在如下方面进行改进：

- 了解语义分割的模型原理，学习相关算法如SAN、VPD等，跟进调研视觉模型的科研进展，正确选择baseline
- 多读、精读相关领域的论文，从而将对模型的理解从应用提升到本质
- 持续探究提升数据集质量的方式和手段，不断提升模型的应用效果
- 继续学习IsaacSim等仿真平台的使用，探究Sim2Real的gap

谢谢观看



重庆中科汽车软件创新中心
AUTOMOTIVE SOFTWARE INNOVATION CENTER CHONGQING

ISCAS 中国科学院软件研究所
Institute of Software Chinese Academy of Sciences

IISG 广州市智能软件产业研究院
Institute of Intelligent Software, Guangzhou



深信科创
SYNKROTRON



西南大学
计算机与信息科学学院 软件学院
COLLEGE OF COMPUTER AND INFORMATION SCIENCE, SOUTHWEST UNIVERSITY
SCHOOL OF SOFTWARE, SOUTHWEST UNIVERSITY