SYSU 2023
PRESENTATION SLIDE

中山大學
SUN YAT-SEN UNIVERSITY

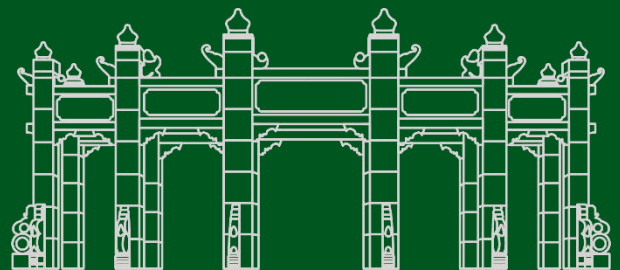
面向草坪环境的小车自动驾驶技术研究

基于扩散策略与BEV投影的自主探索与建图

答辩人：林泽

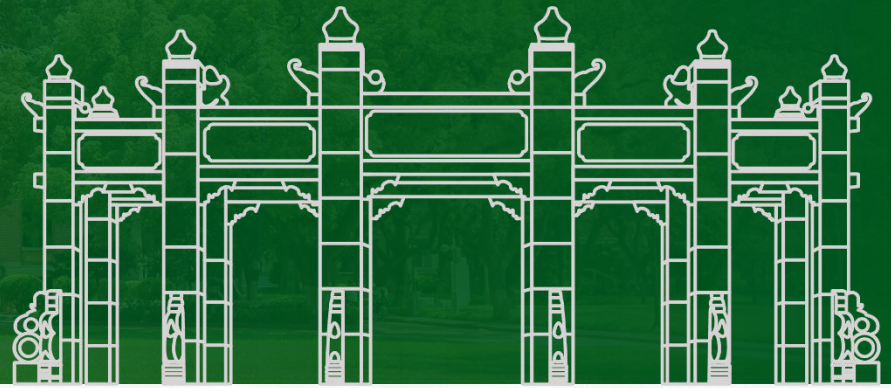
指导老师：廖思宇

项目成员：廖俪霜、陈锦宇、胡书琨、何申、林泽



中山大學 集成电路学院
School of Integrated Circuits, Sun Yat-sen University

目录 Contents



研究背景

市场潜力　随着智能家居和精准农业的快速发展，自动割草机器人作为智能园艺的关键应用，具有巨大的市场潜力

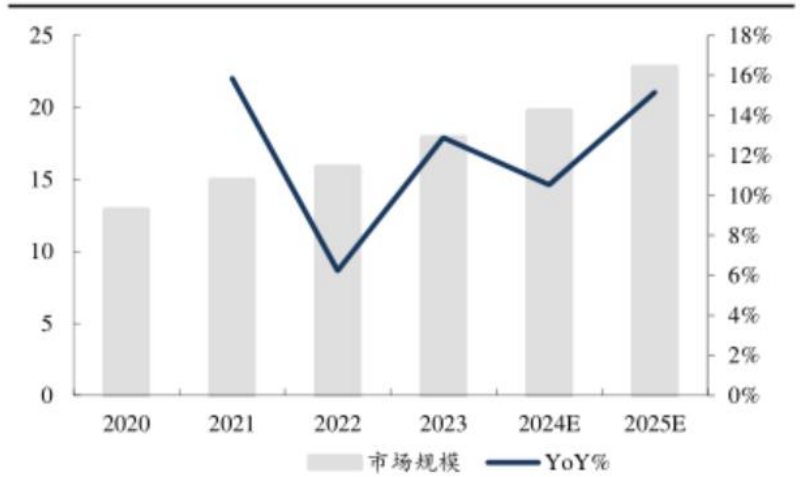
核心需求

全自动化作业
实现在无人干预下的边界识别与高效覆盖

低成本普及
将高精度导航技术部署到消费级、低成本的硬件平台

环境适应性
系统能处理草地、沙地、非铺装路面等非结构化环境的复杂性

图26：全球割草机器人市场规模（亿美元）



数据来源：华经产业研究院，东吴证券研究所

国内割草机器人围绕智能化破局出海。根据华经产业研究院，2023 年宝时得 Worx智能割草机器人在德国占有率达 54%，法国达到 53%，意大利 50%，销售额近 2 亿美元；富世华 2023 年割草机器人收入近 7.6 亿美元。追觅科技、九号公司等国内割草机器人公司围绕智能化提升产品力，解决用户前期布线费时费力、充电慢、效率低等痛点问题，份额有望迅速提升。

研究背景

现有技术的不足

尽管市场需求旺盛，但现有的大多数解决方案仍存在明显的局限性：

高成本依赖

主流高精度方案依赖LiDAR或RTK-GPS，显著推高硬件成本，难以普及。

环境限制

传统SLAM在草地等重复的环境中容易定位失败和建图漂移。

依赖人工布设的电子围栏，部署繁琐，缺乏动态感知能力。

实时性瓶颈

复杂的深度学习模型在嵌入式平台上直接运行时，推理时间较高，不满足机器人实时决策的要求。

边界规划僵硬

传统规划器在面对开阔草地和多样性边界（高障碍/平坦/混合边界）时，无法生成平滑、最优的轨迹，需要额外的条件性修正逻辑。

项目目标与拟解决问题

中山大學 SUN YAT-SEN UNIVERSITY

大创项目申请书回顾

核心目标

提出一种纯视觉、基于扩散策略和BEV投影的自主探索与建图框架

技术路径

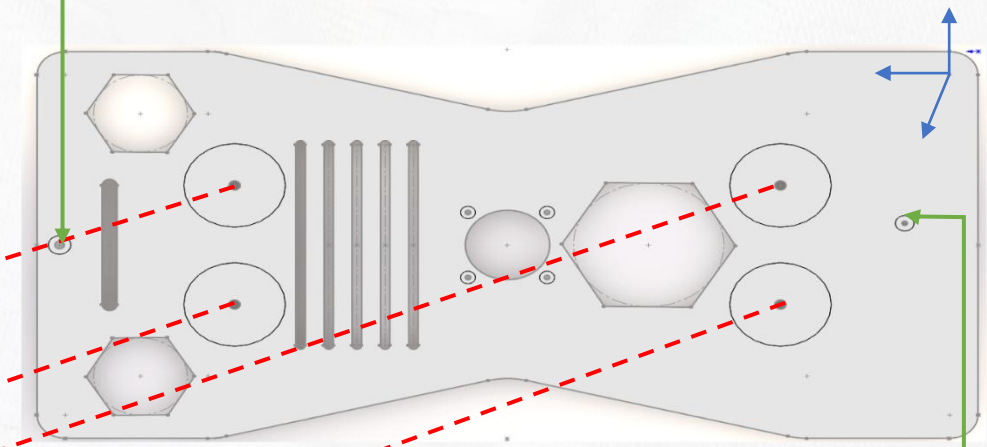
利用 YOLOv8 和 NoMaD 等先进算法，在保证低成本（仅使用RGB-D相机）的同时，实现高鲁棒性、高实时性的自主导航

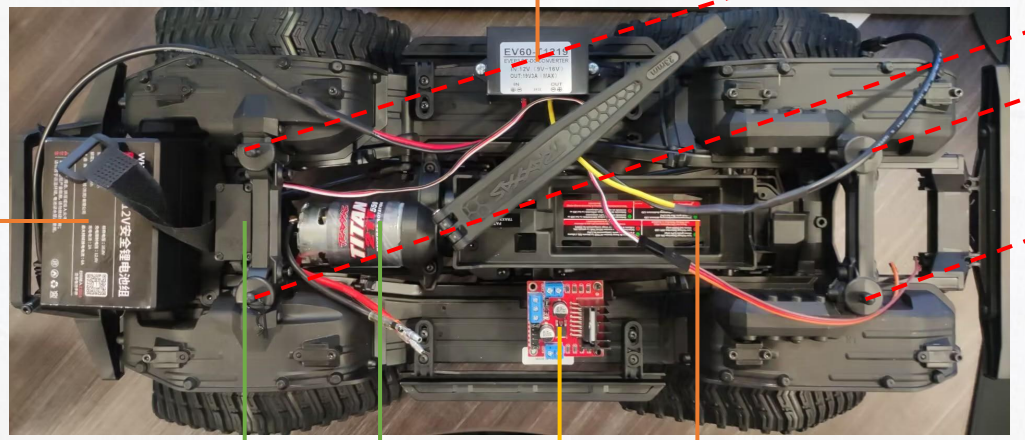
预期成果

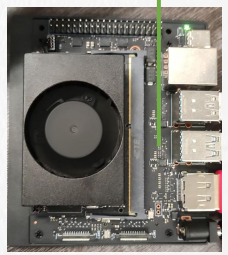
解决复杂环境视觉遮挡问题，小车系统在室内、复杂障碍环境中实现较高精度路径规划，初步适应光照变化与遮挡

小车设计与搭建

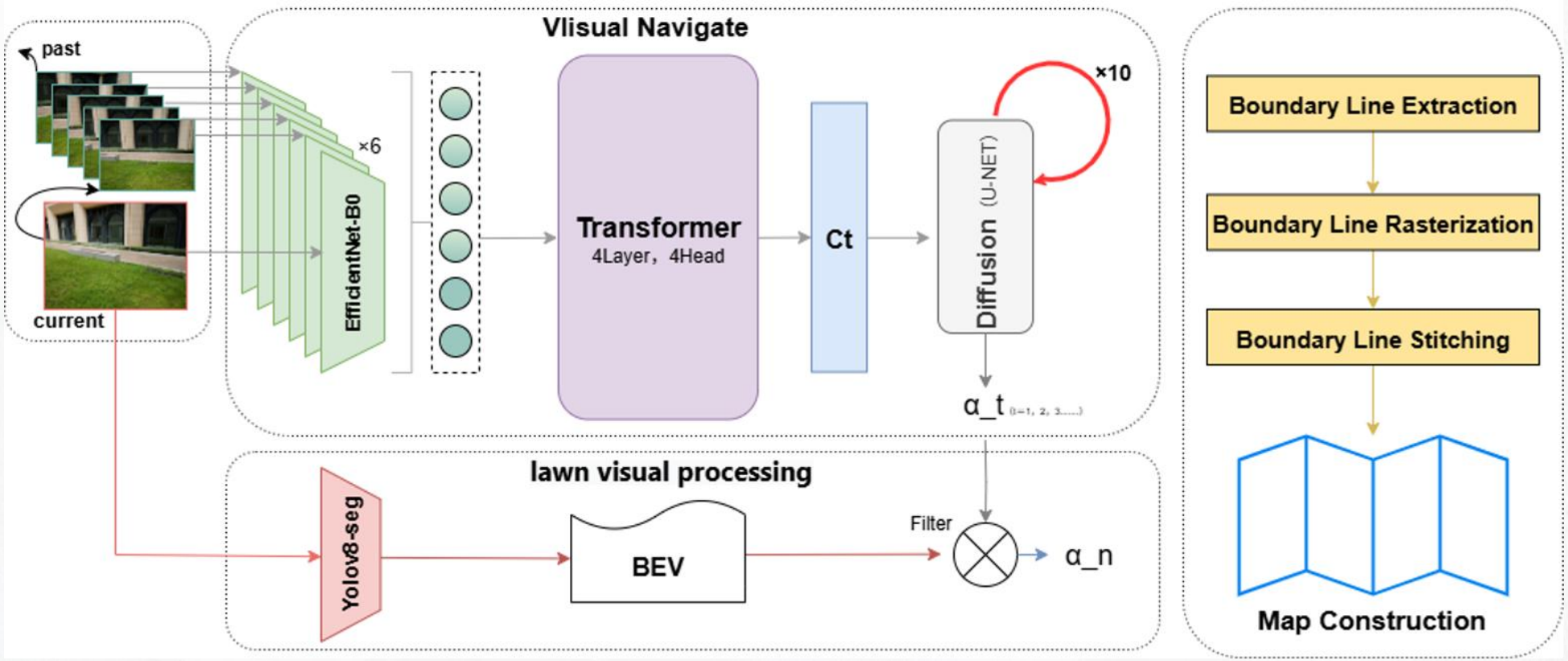
·底盘设计：改装市面上的 RC遥控车底盘，极大降低机械平台成本
·计算平台：采用 NVIDIA Jetson Orin Nano，运行YOLOv8分割模型和NoMaD扩散策略模型，实现边缘端实时推理
·感知系统：使用 Intel RealSense D435i RGB-D相机，提供彩色图像（用于分割）和深度信息（用于规划），契合“纯视觉”需求
·供电集成（工程难点）：设计多级供电系统，独立12V电池供电电机，19V变压器供电计算平台，解决电压不统一问题







总体技术框架



·环境感知：采用自训练 YOLOv8-seg 模型实现草坪/非草坪像素级分割

·路径规划：基于 NoMaD Diffusion 生成候选轨迹，结合 BEV投影 验证路径可行性

·地图构建：通过粗搜索 + ICP精修的两阶段配准实现BEV边界拼接，构建全局地图

关键技术1 - NoMaD扩散策略

模型简介

NoMaD (Goal Masked Diffusion Policies) 是一种用于导航与探索的统一深度学习模型

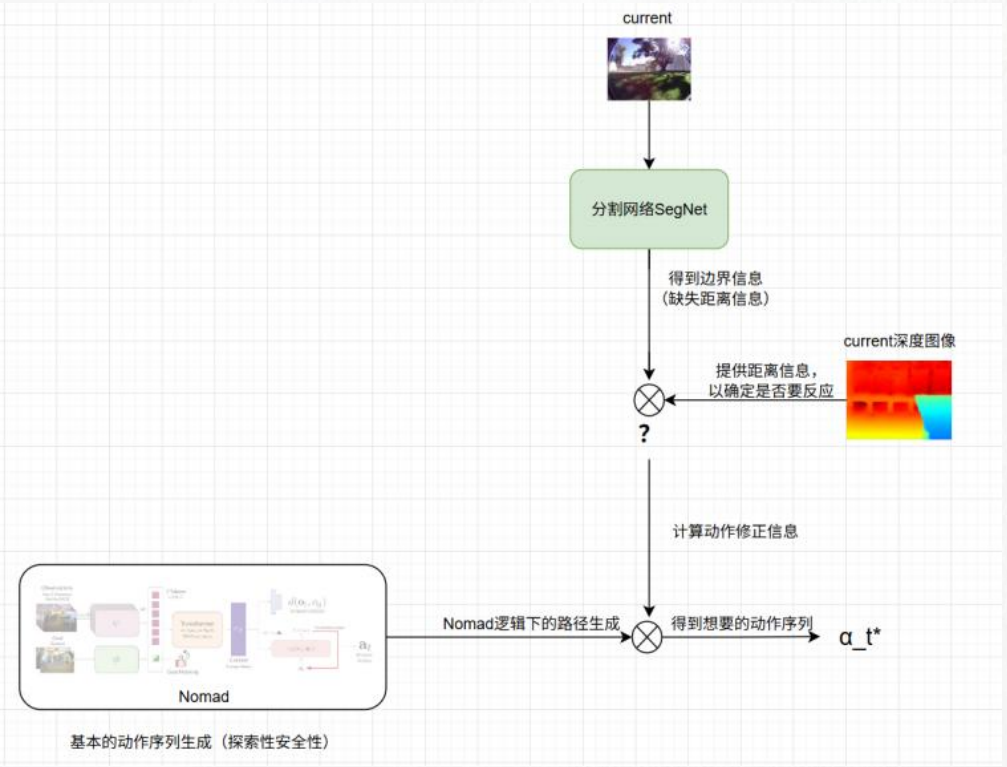
核心机制

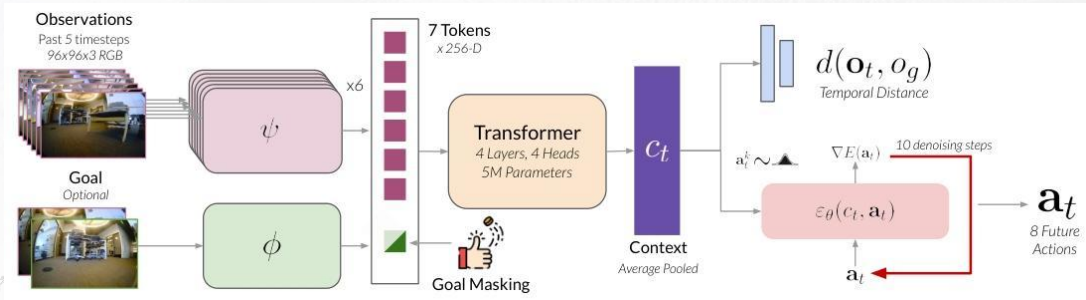
单模型双模式——通过目标掩码机制灵活切换任务导向导航与无任务探索模式。

动作生成——结合 Transformer 编码视觉序列与扩散模型（Diffusion），多次采样生成多条未来动作序列

优势

无需分离模型即可实现“先探索定位、后精准抵达”，输出碰撞规避动作序列[1]





[1]Sridhar A, Shah D, Glossop C, et al. Nomad: Goal masked diffusion policies for navigation and exploration[C]//2024 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2024: 63-70.

模型训练与数据集

阶段一 —— 避障预训练

使用 Tartan (越野场景) 和 Recon 公开数据集，让模型掌握安全避障与轨迹平滑的底层能力

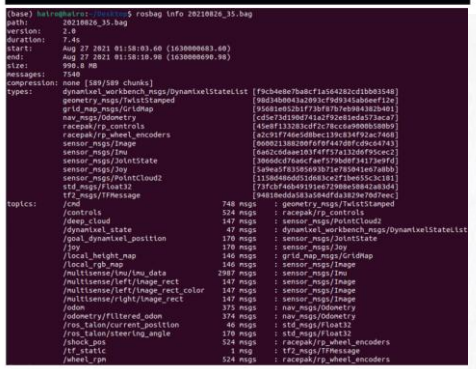
阶段二 —— 贴边联合微调

采用手持遥控录制的草坪边界示范数据，微调模型以适应贴边导航任务

整个流程中，避障作为底层能力、贴边作为目标偏好，两者解耦后统一在 Diffusion 阶段实现

Data Description

Modality	Dimension	Topic	Frequency
State	7	/odometry/filtered_odom	50Hz
Action	2	/cmd	100Hz
RGB Image	1024 x 512	/multisense/left/image_rect_color	20Hz
RGB Map	501 x 501	/local_rgb_map	20Hz
Heightmap	501 x 501	/local_height_map	20Hz
IMU	6	/multisense/imu/imu_data	200Hz
Shock Pos	4	/shock_pos	50Hz
Wheel RPM	4	/wheel_rpm	50Hz
Pedals	2	/controls	50Hz



Tartan公开数据集 01

贴边联合微调 02

	Dataset	Platform	Speed	Total Hrs.	Hrs. Used	Environment
1	GoStanford [28]	TurtleBot2	0.5m/s	17h	14h	office
2	RECON [36]	Jackal	1m/s	25h	25h	off-road
3	CoryHall [42]	RC Car	1.2m/s	2h	2h	hallways
4	Berkeley [29]	Jackal	2m/s	4h	4h	suburban
5	SCAND-S [53]	Spot	1.5m/s	8h	4h	sidewalks
6	SCAND-J [53]	Jackal	2m/s	1h	1h	sidewalks
7	Seattle [54]	Warthog	5m/s	1h	1h	off-road
8	TartanDrive [55]	ATV	10m/s	7h	5h	off-road
9	NeBula [56]	ATV	10m/s	10h	10h	off-road
10	SACSoN [57]	TurtleBot2	0.5m/s	75h	10h	office
11	BDD [13]	Car(s)	20m/s	10h	4h	on-road
	Total			160h	80h	

阶段一

避障预训练 03

阶段二

贴边联合微调 04

关键技术2 - NoMaD 路线生成与应用阶段筛选

生成阶段

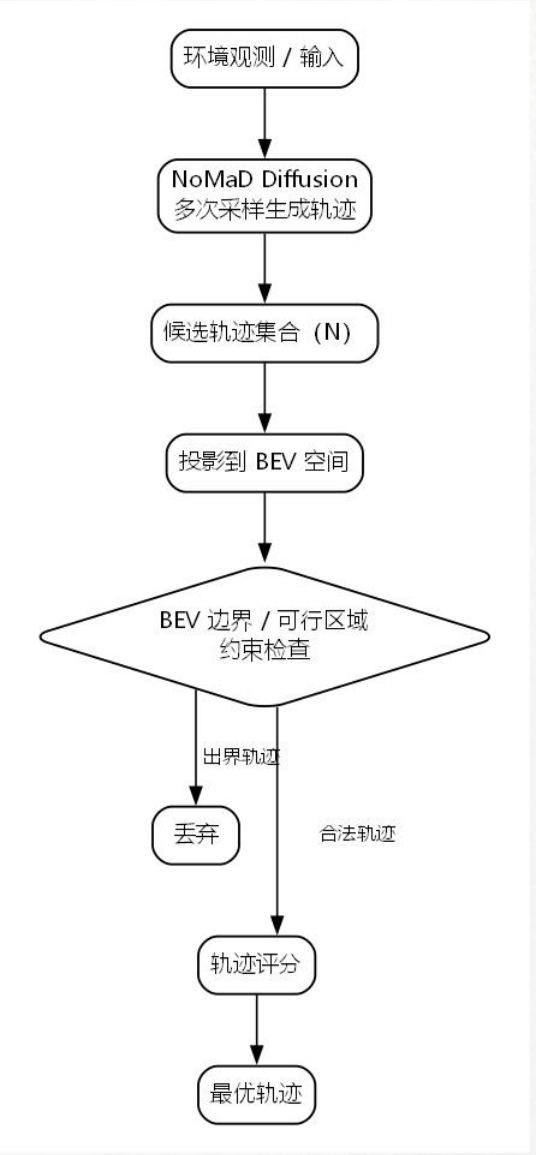
NoMaD模型一次生成 8-10 条候选轨迹，每条轨迹包含未来 10 步的动作序列，通过 TensorRT 加速扩散模型推理，将轨迹生成时间从350ms 降至150ms，结合 BEV 轨迹验证，使规划模块整体延迟控制在 200ms 内

BEV约束检查

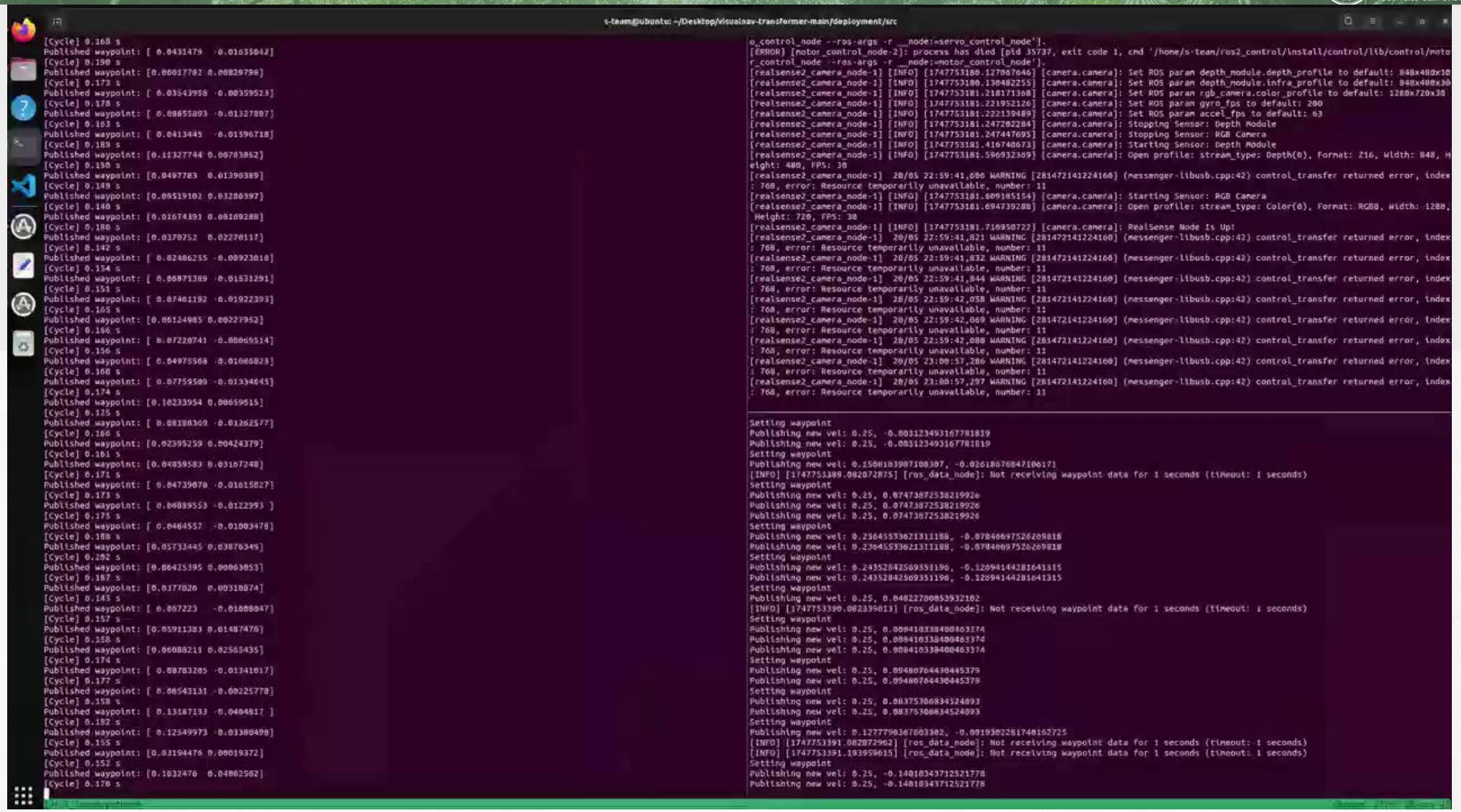
将候选轨迹映射到 BEV（鸟瞰）空间，利用 YOLO 分割得到的草坪 Mask 评估可行性，剔除不完全落在可行区域内的轨迹

最优决策

在合法轨迹集合中，选择最优（如贴边或指向未知区域）的轨迹执行，从而实现“条件性修正”，例如选“完全在边界内、且与边界近似相切/贴边”的那条作为最终路线。



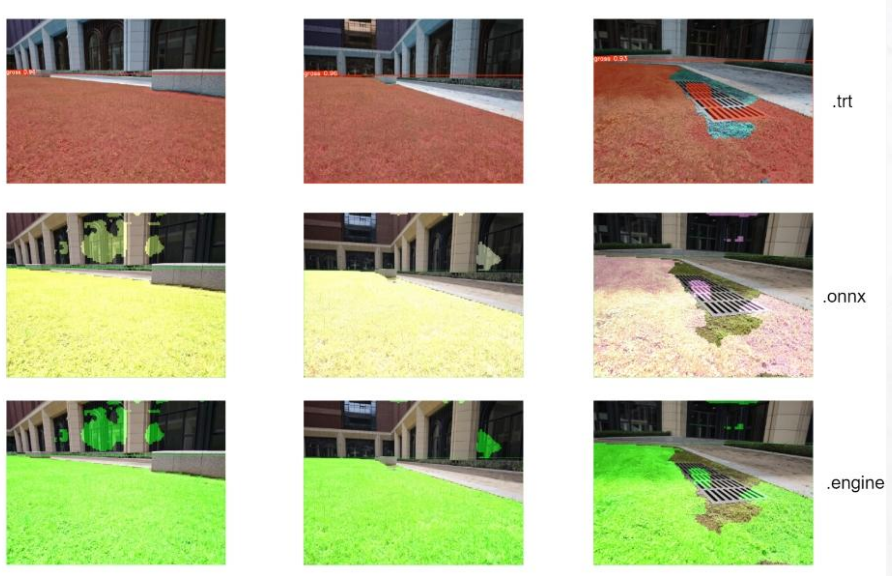
NoMaD运行日志


s-team@ubuntu: ~/Desktop/visualnav-transformer-main/deployment/src
[Cycle] 0.168 s
Published waypoint: [0.0431479 -0.01635042]
[Cycle] 0.190 s
Published waypoint: [0.06017781 0.00829798]
[Cycle] 0.173 s
Published waypoint: [0.03543958 -0.00359523]
[Cycle] 0.178 s
Published waypoint: [0.08855893 -0.01327807]
[Cycle] 0.163 s
Published waypoint: [0.0413445 -0.01596718]
[Cycle] 0.189 s
Published waypoint: [0.11327744 0.00783852]
[Cycle] 0.150 s
Published waypoint: [0.0497783 0.01390389]
[Cycle] 0.149 s
Published waypoint: [0.09519102 0.03286397]
[Cycle] 0.140 s
Published waypoint: [0.01674391 0.00169288]
[Cycle] 0.180 s
Published waypoint: [0.0370752 0.02270117]
[Cycle] 0.142 s
Published waypoint: [0.02486255 -0.00923018]
[Cycle] 0.154 s
Published waypoint: [0.06871389 -0.01531291]
[Cycle] 0.151 s
Published waypoint: [0.07401192 -0.01922393]
[Cycle] 0.165 s
Published waypoint: [0.06124985 0.00227952]
[Cycle] 0.166 s
Published waypoint: [0.07220741 -0.00869514]
[Cycle] 0.156 s
Published waypoint: [0.04975568 -0.01068823]
[Cycle] 0.160 s
Published waypoint: [0.07759589 -0.01334645]
[Cycle] 0.174 s
Published waypoint: [0.10233954 0.00659615]
[Cycle] 0.125 s
Published waypoint: [0.08188369 -0.01262577]
[Cycle] 0.166 s
Published waypoint: [0.02395259 0.00424379]
[Cycle] 0.161 s
Published waypoint: [0.04859583 0.03167248]
[Cycle] 0.171 s
Published waypoint: [0.04739078 -0.01615827]
[Cycle] 0.173 s
Published waypoint: [0.06089553 -0.0122993]
[Cycle] 0.175 s
Published waypoint: [0.0484557 -0.01003478]
[Cycle] 0.188 s
Published waypoint: [0.05733445 0.03876349]
[Cycle] 0.202 s
Published waypoint: [0.00425395 0.00063053]
[Cycle] 0.187 s
Published waypoint: [0.0377026 0.00318874]
[Cycle] 0.143 s
Published waypoint: [0.057223   -0.01088047]
[Cycle] 0.157 s
Published waypoint: [0.05911383 0.01487476]
[Cycle] 0.158 s
Published waypoint: [0.06088213 0.02565435]
[Cycle] 0.174 s
Published waypoint: [0.08783285 -0.01341017]
[Cycle] 0.177 s
Published waypoint: [0.08543131 -0.00225778]
[Cycle] 0.158 s
Published waypoint: [0.13187193 -0.0404817]
[Cycle] 0.182 s
Published waypoint: [0.12549973 -0.03388409]
[Cycle] 0.155 s
Published waypoint: [0.03194476 0.00019372]
[Cycle] 0.152 s
Published waypoint: [0.1032476 0.04862562]
[Cycle] 0.170 s
o_control_node --ros-args -r __node:=servo_control_node'].
[ERROR] [motor_control_node-2]: process has died [pid 35737, exit code 1, cmd '/home/s-team/ros2_control/install/control/lib/control/motor_control_node --ros-args -r __node:=motor_control_node'].
[realsense2_camera_node-1] [INFO] [1747753180.127867646] [camera.camera]: Set ROS param depth_module.depth_profile to default: 848x480x30
[realsense2_camera_node-1] [INFO] [1747753180.130482255] [camera.camera]: Set ROS param depth_module.infra_profile to default: 848x480x30
[realsense2_camera_node-1] [INFO] [1747753181.21817136B] [camera.camera]: Set ROS param rgb_camera.color_profile to default: 1280x720x30
[realsense2_camera_node-1] [INFO] [1747753181.221952126] [camera.camera]: Set ROS param gyro_fps to default: 200
[realsense2_camera_node-1] [INFO] [1747753181.222139489] [camera.camera]: Set ROS param accel_fps to default: 63
[realsense2_camera_node-1] [INFO] [1747753181.247202284] [camera.camera]: Stopping Sensor: Depth Module
[realsense2_camera_node-1] [INFO] [1747753181.247447695] [camera.camera]: Stopping Sensor: RGB Camera
[realsense2_camera_node-1] [INFO] [1747753181.416740673] [camera.camera]: Starting Sensor: Depth Module
[realsense2_camera_node-1] [INFO] [1747753181.596932369] [camera.camera]: Open profile: stream_type: Depth(0), Format: Z16, Width: 848, Height: 480, FPS: 30
[realsense2_camera_node-1] 20/05 22:59:41,606 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] [INFO] [1747753181.609185154] [camera.camera]: Starting Sensor: RGB Camera
[realsense2_camera_node-1] [INFO] [1747753181.6947392BB] [camera.camera]: Open profile: stream_type: Color(0), Format: RGB8, Width: 1280, Height: 720, FPS: 30
[realsense2_camera_node-1] [INFO] [1747753181.710950722] [camera.camera]: RealSense Node Is Up!
[realsense2_camera_node-1] 20/05 22:59:41,821 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 22:59:41,83Z WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 22:59:41,844 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 22:59:42,058 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 22:59:42,069 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 22:59:42,080 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 23:00:57,286 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
[realsense2_camera_node-1] 20/05 23:00:57,297 WARNING [281472141224160] (messenger-libusb.cpp:42) control_transfer returned error, index: 768, error: Resource temporarily unavailable, number: 11
Setting waypoint
Publishing new vel: 0.25, -0.003123493167781819
Publishing new vel: 0.25, -0.003123493167781819
Setting waypoint
Publishing new vel: 0.250#103907108307, -0.02618076847106171
[INFO] [1747753389.082072875] [ros_data_node]: Not receiving waypoint data for 1 seconds (timeout: 1 seconds)
Setting waypoint
Publishing new vel: 0.25, 0.07473872538219926
Publishing new vel: 0.25, 0.07473872538219926
Publishing new vel: 0.25, 0.07473872538219926
Setting waypoint
Publishing new vel: 0.23645533621311188, -0.07840697526269818
Publishing new vel: 0.23645533621311188, -0.07840697526269818
Setting waypoint
Publishing new vel: 0.24352842569351196, -0.12694144281641315
Publishing new vel: 0.24352842569351196, -0.12694144281641315
Setting waypoint
Publishing new vel: 0.25, 0.04822700853932102
[INFO] [1747753390.082339013] [ros_data_node]: Not receiving waypoint data for 1 seconds (timeout: 1 seconds)
Setting waypoint
Publishing new vel: 0.25, 0.08841033840463374
Publishing new vel: 0.25, 0.08841033840463374
Publishing new vel: 0.25, 0.08841033840463374
Setting waypoint
Publishing new vel: 0.25, 0.09480764430445379
Publishing new vel: 0.25, 0.09480764430445379
Setting waypoint
Publishing new vel: 0.25, 0.08375306834524893
Publishing new vel: 0.25, 0.08375306834524893
Setting waypoint
Publishing new vel: 0.1277790367603302, -0.09193022817481627 25
[INFO] [1747753391.082872962] [ros_data_node]: Not receiving waypoint data for 1 seconds (timeout: 1 seconds)
[INFO] [1747753391.193959615] [ros_data_node]: Not receiving waypoint data for 1 seconds (timeout: 1 seconds)
Setting waypoint
Publishing new vel: 0.25, -0.14810343712521778
Publishing new vel: 0.25, -0.14810343712521778

边缘端部署优化

特性	PyTorch 直接部署	TensorRT 部署
硬件兼容性	支持 CPU、多厂商 GPU、TPU 等	仅 NVIDIA GPU
开发灵活性	动态图调试方便，支持快速迭代	需固定模型结构，转换流程复杂
泛化能力	一次训练，多处部署	强依赖 NVIDIA 生态
极限性能	中等（依赖硬件驱动优化）	极高（针对 NVIDIA GPU 深度优化）

硬件针对优化后速度提升，效果基本不变



工程挑战

复杂的扩散模型在 Jetson Orin Nano 上直接推理延迟极高（>4000ms），无法满足实时性

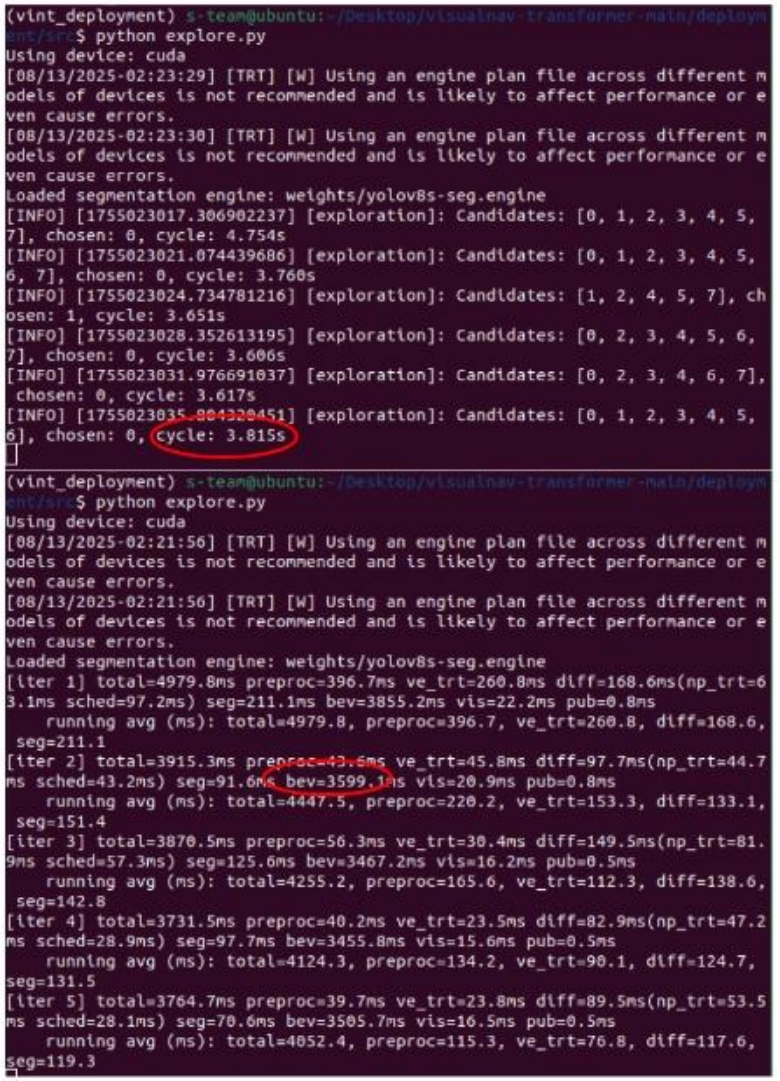
解决方案

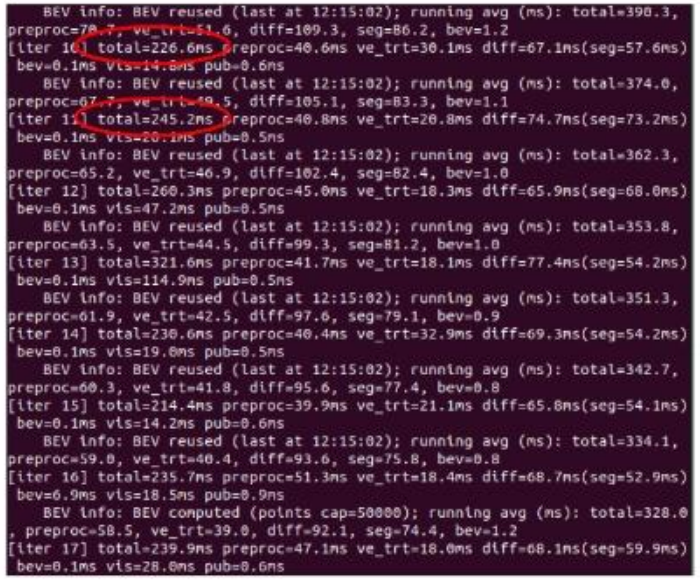
实现从 PyTorch 到 TensorRT 的迁移，通过算子融合与精度量化加速推理

优化结果

总体运行平均时间由 4000+ms 缩短到 200-300ms，结合BEV验证，整体规划延迟控制在 200ms 内，成功实现实时导航

边缘端部署优化





工程挑战

复杂的扩散模型在 Jetson Orin Nano 上直接推理延迟极高（>4000ms），无法满足实时性

解决方案

实现从 PyTorch 到 TensorRT 的迁移，通过算子融合与精度量化加速推理

优化结果

总体运行平均时间由 4000+ms 缩短到 200-300ms，结合BEV验证，整体规划延迟控制在 200ms 内，成功实现实时导航

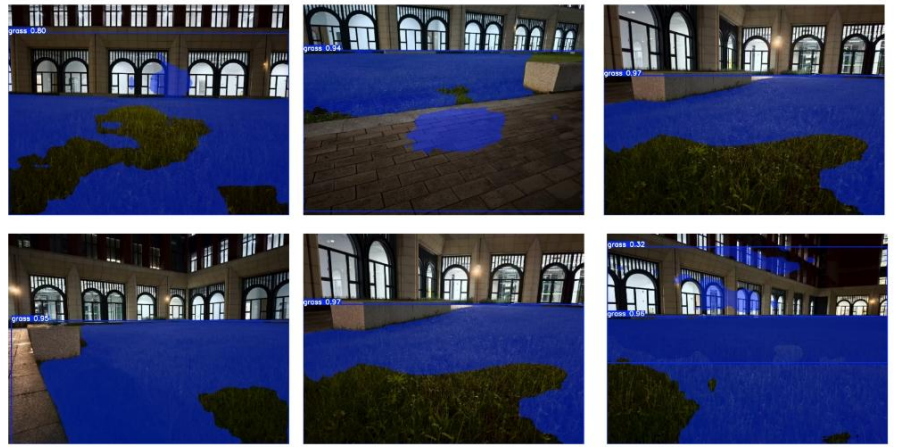
关键技术3 - 视觉感知与BEV建图

语义分割

使用通用+专用数据集训练 YOLOv8，在复杂光照下（如夜晚干草）仍保持高鲁棒性

其中“白天干草”mAP@0.5为0.872，“夜晚干草”为0.806

YOLOv8分割图



关键技术3 - 视觉感知与BEV建图

语义分割

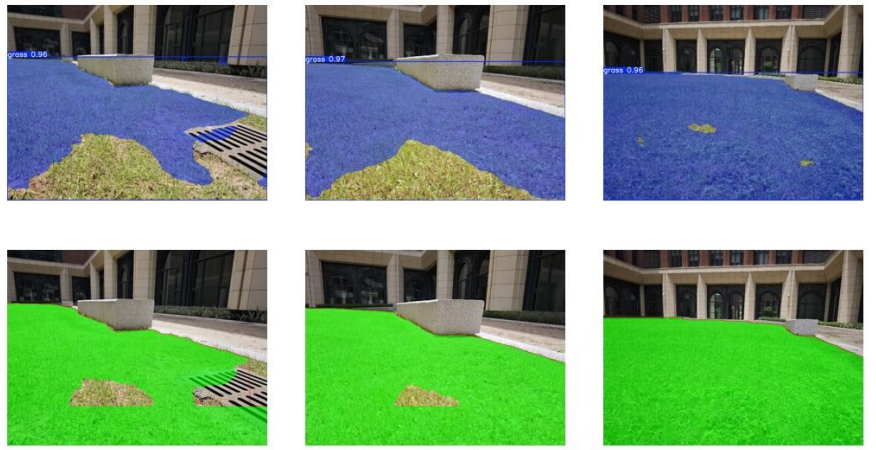
使用通用+专用数据集训练 YOLOv8，在复杂光照下（如夜晚干草）仍保持高鲁棒性

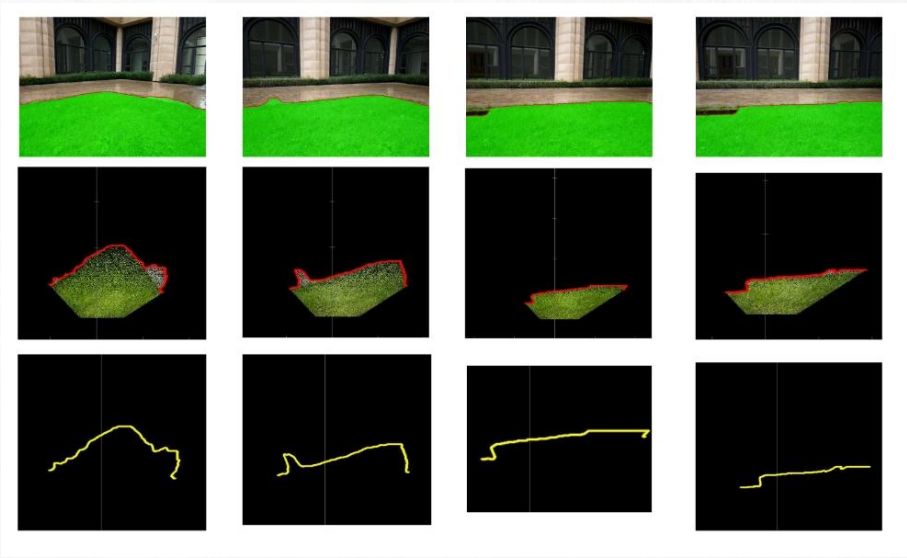
其中“白天干草”mAP@0.5为0.872，“夜晚干草”为0.806

逆透视变换(IPM)

利用相机内参及安装外参计算单应性矩阵，将YOLO 生成的草坪掩码精确投影到 BEV 栅格地图

BEV转换图





关键技术3 - 视觉感知与BEV建图

语义分割

使用通用+专用数据集训练 YOLOv8，在复杂光照下（如夜晚干草）仍保持高鲁棒性
其中“白天干草”mAP@0.5为0.872，“夜晚干草”为0.806

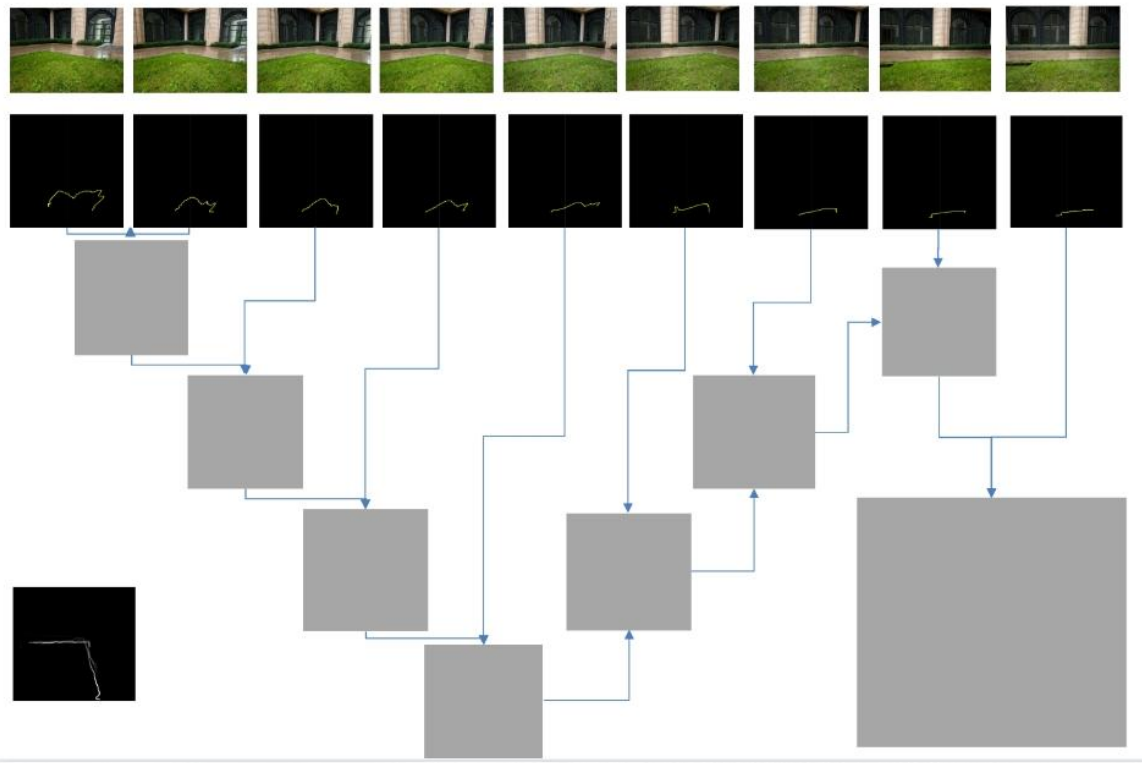
逆透视变换(IPM)

利用相机内参及安装外参计算单应性矩阵，将YOLO 生成的草坪掩码精确投影到 BEV 栅格地图

地图拼接

引入IMU陀螺仪数据辅助，使用 Pairwise Stitching 技术进行地图拼接，构建草坪俯视地图[1]

拼接图



[1]Sani E, Sgorbissa A, Carpin S. Improving the ros 2 navigation stack with real-time local costmap updates for agricultural applications[C]//2024 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2024: 17701-17707.

研究总结与未来计划

项目总结

1.实现了纯视觉端到端的低成本自主导航方案

2. 创新结合了 NoMaD Diffusion (生成多样性) 与 BEV技术 (安全约束)，解决了传统规划器僵硬的问题

3. 成功攻克了边缘设备上的实时推理难题

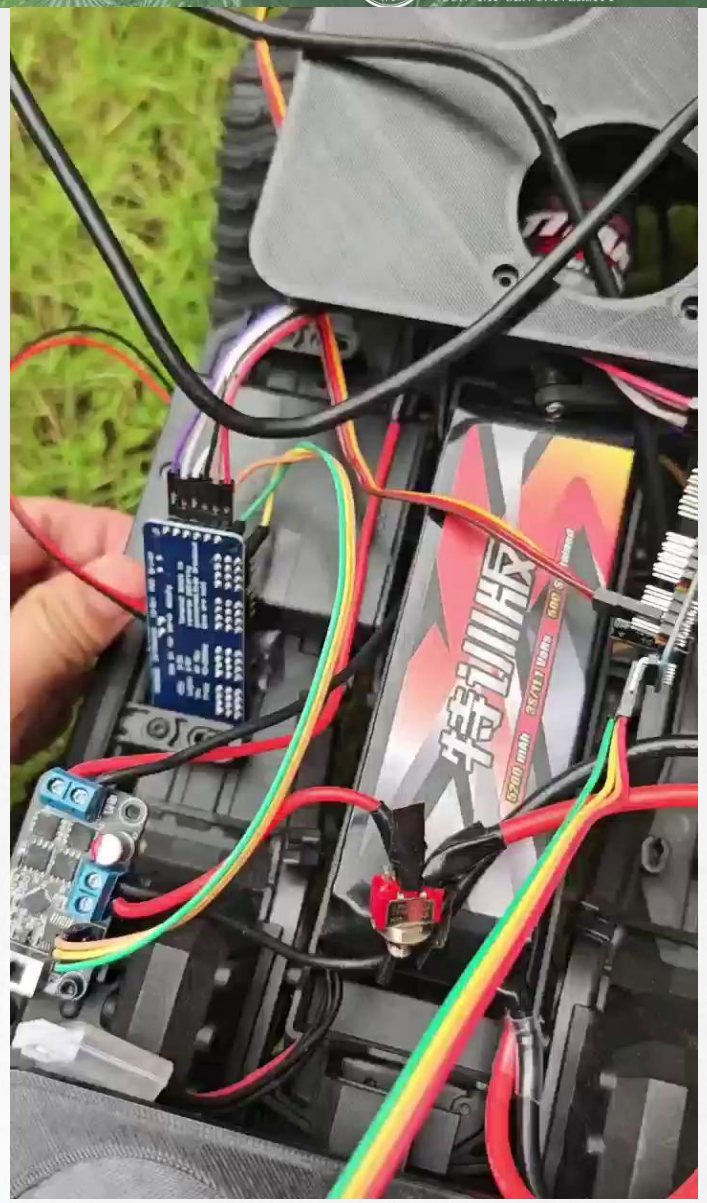
未来计划

1. 多模态融合
针对RGB视觉在极端环境的局限，计划引入IMU或低成本雷达提高建图准确度 。

2. 全尺度地图优化
改进地图拼接算法，解决大范围场景下的累积漂移问题

实机视频展示

Phase 1: Safety Testing of the Robot During Lawn Operation



中山大學
SUN YAT-SEN UNIVERSITY

谢谢观看

SUN YAT-SEN UNIVERSITY 2025