A Robot for Cleaning and Sorting Garbage in the Home Environment

Jiajun Long, Ximan Zhang, Yikai Zheng, Jie Luo, Vong Sinrithy

2023.6.1

Contents

Introduction

Detailed Introduction

Task:

A defined home environment
some household items carelessly placed
One or more trash cans

mapping the overall environment
Identify furniture and garbage

One or more garbage that appears randomly is thrown into the corresponding trash can

AncoraSIR.com

Related Work

[1] Jo˜ao Machado Santos, David Portugal, and Rui P. Rocha. An evaluation of 2d slam techniques available in robot operating system. In 2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), pages 1–6, 2013.
[2] Peiyuan Jiang, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. A review of yolo algorithm developments. Procedia Computer Science, 199:1066–1073, 2022. The 8th International Conference on Information Technology and Quantitative Management (ITQM 2020 & 2021): Developing Global Digital Economy after COVID-19.
[3] Yoonyoung Cho, Donghoon Shin, and Beomjoon Kim. ω2: Optimal hierarchical planner for object search in large environments via mobile manipulation. In 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 7888–7895, Oct 2022.
[4] Alexey Bochkovskiy, Chien-Yao Wang, and Hong- Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. CoRR, abs/2004.10934, 2020.
[5] Cihang Xie, Jianyu Wang, Zhishuai Zhang, Yuyin Zhou, Lingxi Xie, and Alan L. Yuille. Adversarial examples for semantic segmentation and object detection. CoRR, abs/1703.08603, 2017.
[6] Chien-Yao Wang, Hong-Yuan Mark Liao, I-Hau Yeh, Yueh-Hua Wu, Ping-Yang Chen, and Jun-Wei Hsieh. Csp- net: A new backbone that can enhance learning capability of CNN. CoRR, abs/1911.11929, 2019.
[7] Diganta Misra. Mish: A self regularized non-monotonic neural activation function. CoRR, abs/1908.08681, 2019.
[8] Golnaz Ghiasi, Tsung-Yi Lin, and Quoc V. Le. Drop- block: A regularization method for convolutional net- works. CoRR, abs/1810.12890, 2018.
[9] Giorgio Grisetti, Cyrill Stachniss, and Wolfram Bur- gard. Improved techniques for grid mapping with rao- blackwellized particle filters. IEEE Transactions on Robotics, 23(1):34–46, 2007.

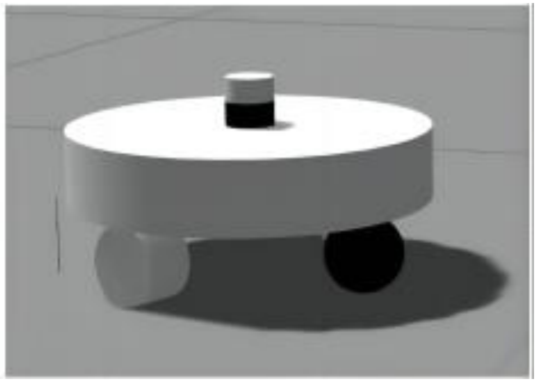
SUSTech
Southern University of Science and Technology

Dataset and Environment

AncoraSIR.com

Model and Environment

Model



Parameters: Maximum speed: 0.5m/s
Maximum acceleration: 1m/s2
IMU equipped
LiDAR:
particles: 500
detection radius: 0.1-6m
RGB-D Camera

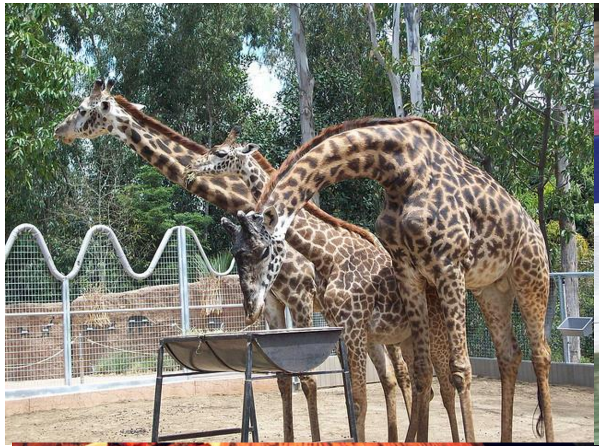
Environment

Living room: sofas, trash cans, tables
Kitchen: dining table, chairs
Bedroom: fitness equipment, desks

AncoraSIR.com

Dataset - COCO

Giraffe



Horse

Cat

Cup

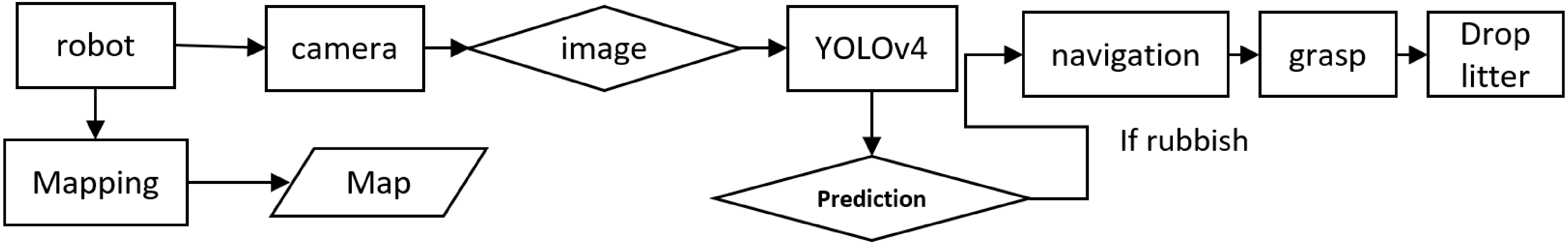
Bottle

Bowl

AncoraSIR.com

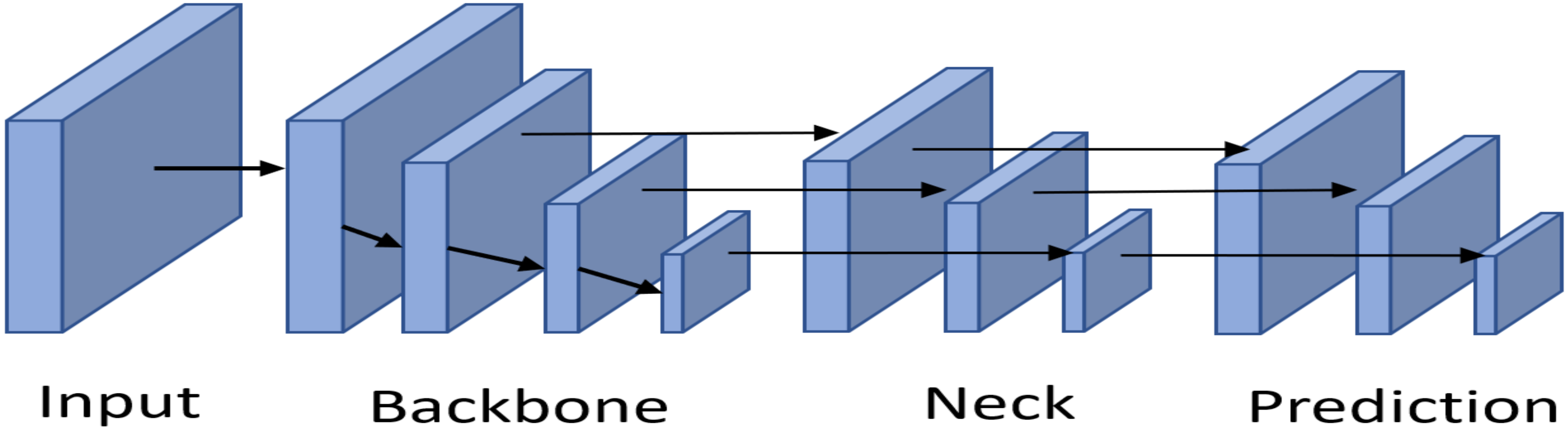
Methods

Work Flow



AncoraSIR.com

YOLOv4



- Mosaic data enhancement
- CmBN
- SAT(Self – Adversarial Training

- CSPDarknet53
- Mish activation function
- Dropblock regularization

- SPP
- FPN+PAN

- YOLOv3

AncoraSIR.com

SLAM

Prediction stage

Algorithm 1 Improved RBPF for Map Learning

Require:

$\mathcal{S}_{t-1}$, the sample set of the previous time step

z_t, the most recent laser scan

u_{t-1}, the most recent odometry measurement

Ensure:

$\mathcal{S}_t$, the new sample set

$\mathcal{S}_t = \{\}$

for all $s_{t-1}^{(i)} \in \mathcal{S}_{t-1}$ **do**

$\langle x_{t-1}^{(i)}, w_{t-1}^{(i)}, m_{t-1}^{(i)} \rangle = s_{t-1}^{(i)}$

Calibration stage

// scan-matching

$x_t'^{(i)} = x_{t-1}^{(i)} \oplus u_{t-1}$

$\hat{x}_t^{(i)} = \operatorname{argmax}_x p(x \mid m_{t-1}^{(i)}, z_t, x_t'^{(i)})$

if $\hat{x}_t^{(i)}$ = **failure then**

$x_t^{(i)} \sim p(x_t \mid x_{t-1}^{(i)}, u_{t-1})$

$w_t^{(i)} = w_{t-1}^{(i)} \cdot p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)})$

else

// sample around the mode

for $k = 1, \ldots, K$ **do**

$x_k \sim \{x_j \mid |x_j - \hat{x}^{(i)}| < \Delta\}$

end for

Resampling stage

// compute Gaussian proposal

$\mu_t^{(i)} = (0,0,0)^T$

$\eta^{(i)} = 0$

for all $x_j \in \{x_1, \ldots, x_K\}$ **do**

$\mu_t^{(i)} = \mu_t^{(i)} + x_j \cdot p(z_t \mid m_{t-1}^{(i)}, x_j) \cdot p(x_t \mid x_{t-1}^{(i)}, u_{t-1})$

$\eta^{(i)} = \eta^{(i)} + p(z_t \mid m_{t-1}^{(i)}, x_j) \cdot p(x_t \mid x_{t-1}^{(i)}, u_{t-1})$

end for

$\mu_t^{(i)} = \mu_t^{(i)} / \eta^{(i)}$

$\Sigma_t^{(i)} = \mathbf{0}$

for all $x_j \in \{x_1, \ldots, x_K\}$ **do**

$\Sigma_t^{(i)} = \Sigma_t^{(i)} + (x_j - \mu^{(i)})(x_j - \mu^{(i)})^T \cdot$

$p(z_t \mid m_{t-1}^{(i)}, x_j) \cdot p(x_j \mid x_{t-1}^{(i)}, u_{t-1})$

end for

$\Sigma_t^{(i)} = \Sigma_t^{(i)} / \eta^{(i)}$

// sample new pose

$x_t^{(i)} \sim \mathcal{N}(\mu_t^{(i)}, \Sigma_t^{(i)})$

// update importance weights

$w_t^{(i)} = w_{t-1}^{(i)} \cdot \eta^{(i)}$

Map estimation

end if

// update map

$m_t^{(i)} = \text{integrateScan}(m_{t-1}^{(i)}, x_t^{(i)}, z_t)$

// update sample set

$\mathcal{S}_t = \mathcal{S}_t \cup \{\langle x_t^{(i)}, w_t^{(i)}, m_t^{(i)} \rangle\}$

end for

$N_{\text{eff}} = 1 / \sum_{i=1}^{N} (\tilde{w}^{(i)})^2$

if $N_{\text{eff}} < T$ **then**

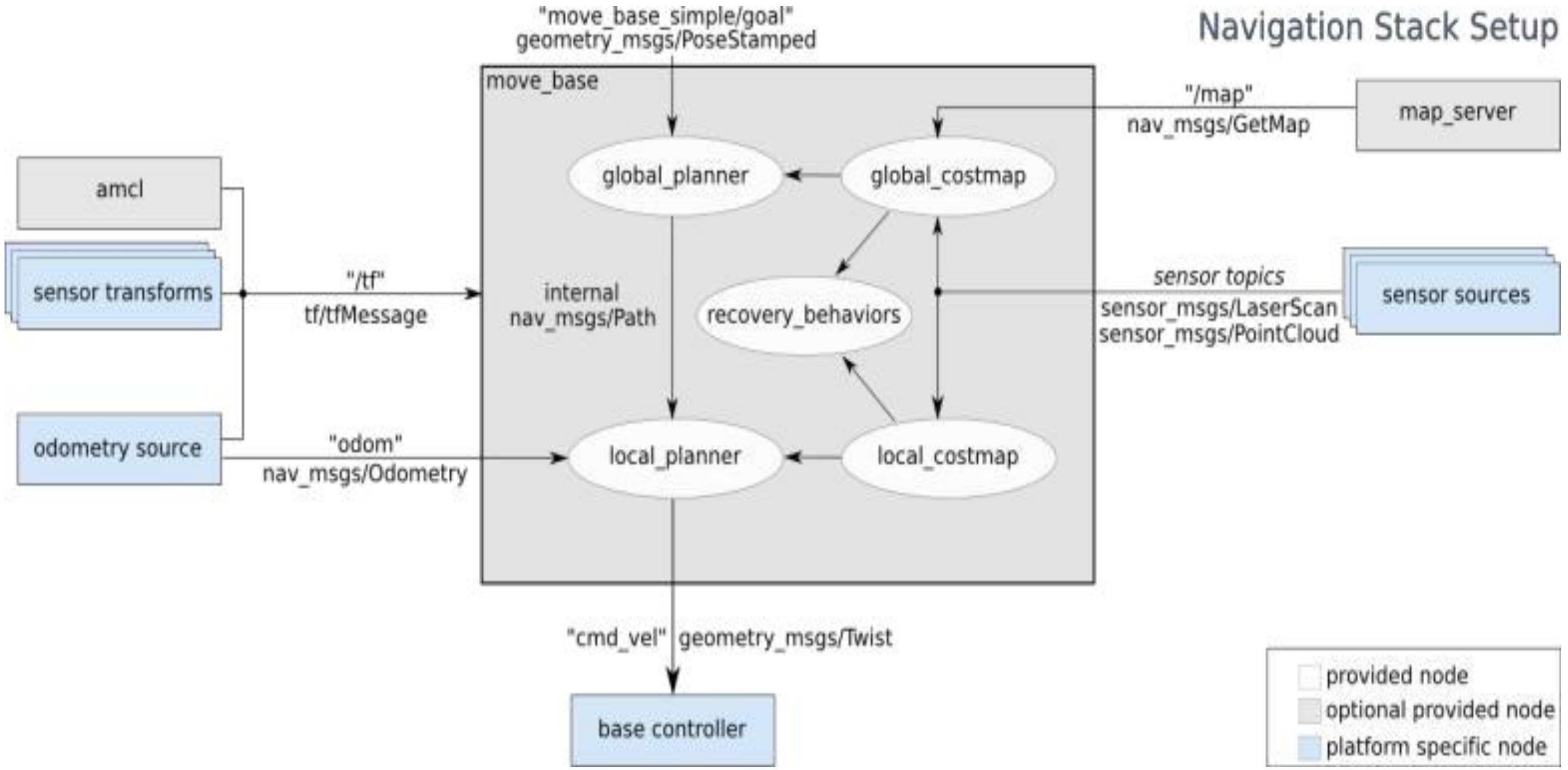
$\mathcal{S}_t = \text{resample}(\mathcal{S}_t)$

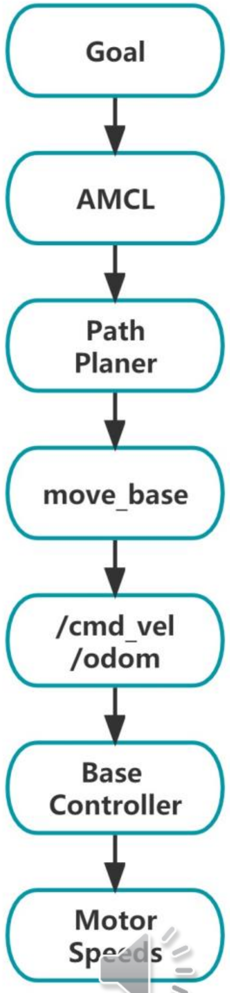
end if

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Navigation





AncoraSIR.com

Experiments

Experiment 1-1 Object Recognition

Tested the model accuracy by COCO test dataset

```
Class     Images  Instances          P          R      mAP50      mAP75   mAP50-95: 100%|██████████| 157/157 [13:16<00:00,  5.07s/it]
  all       5000      36335      0.672      0.519      0.566      0.401      0.371
```

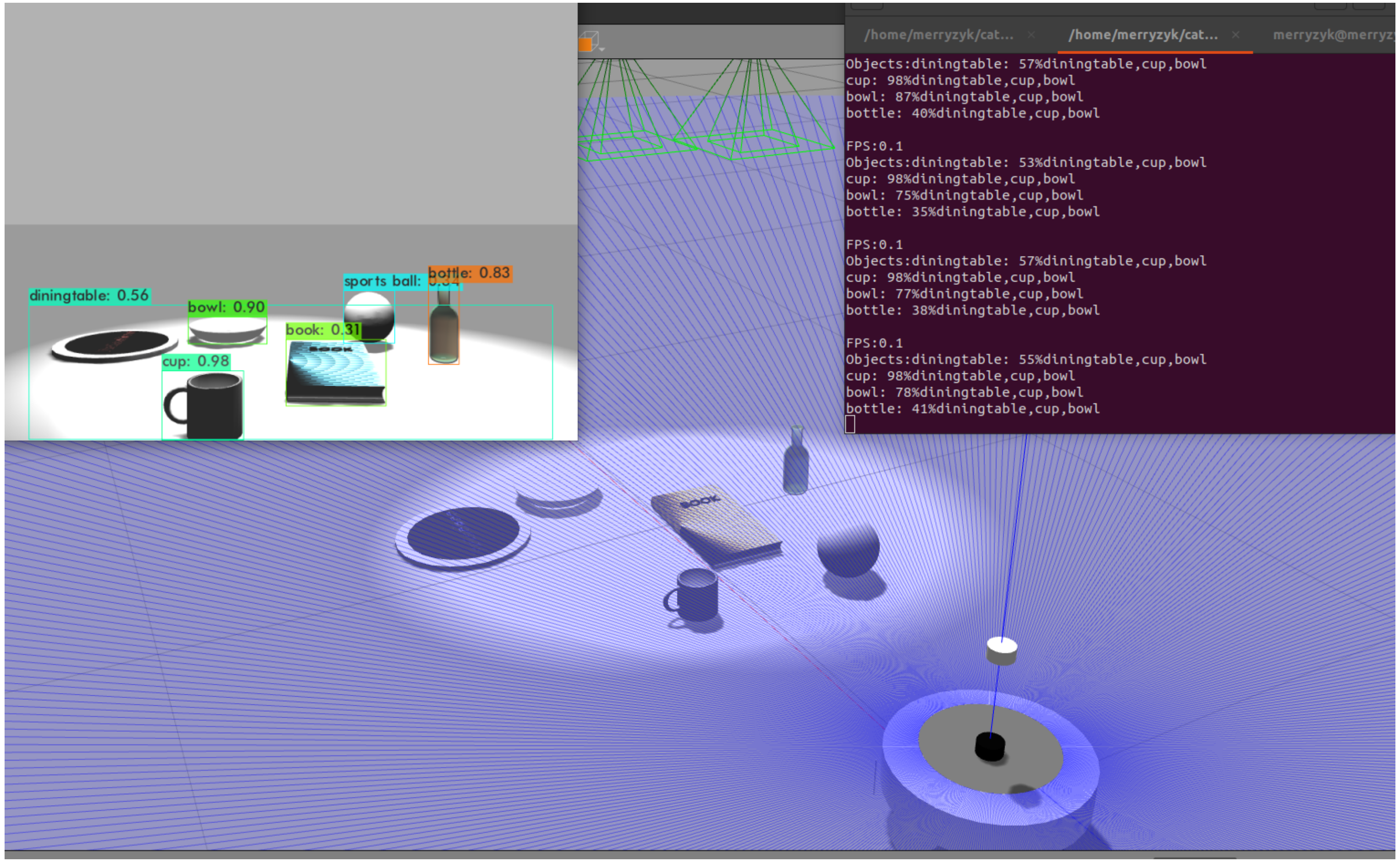
mAP50 = 0.566
mAP75 = 0.401
mAP50-95 = 0.371

Similar to the paper of YOLOv4

High accuracy

Experiment 1-2 Object Recognition

Tested the model accuracy on objects of simulation environment



Final choice：

Cup: 98%

Bowl: 90%

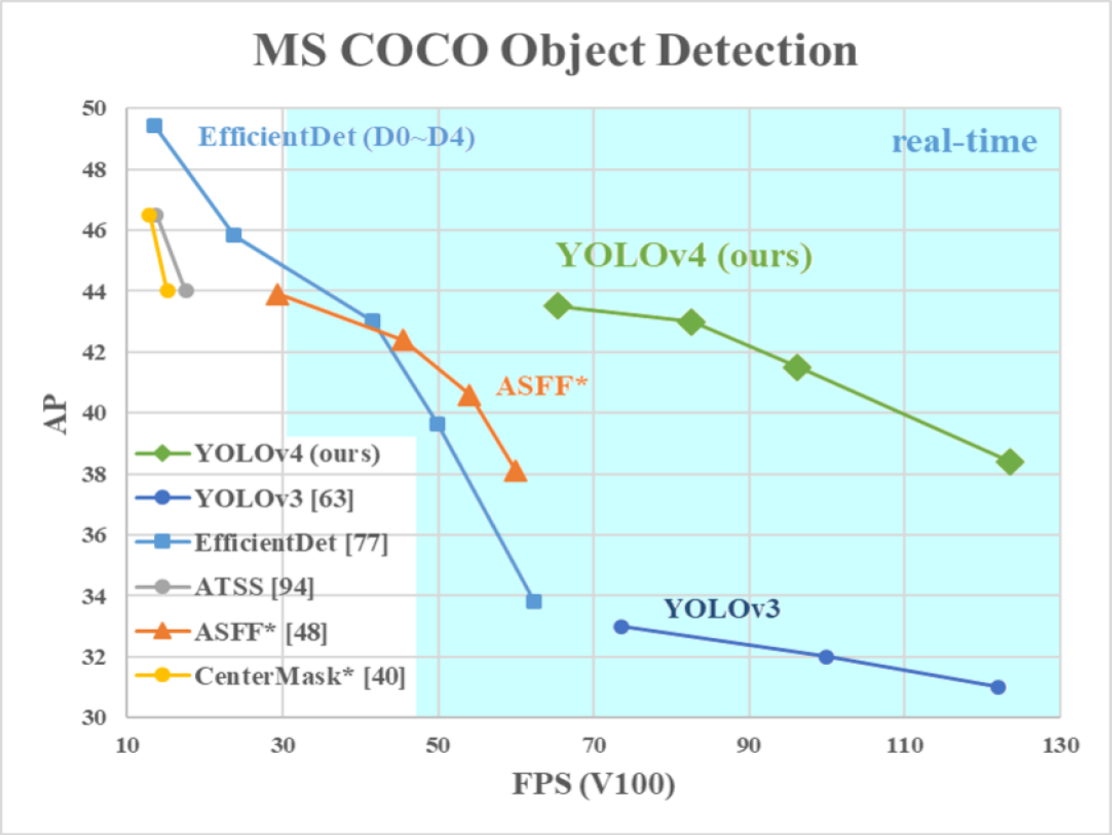
Bottle: 83%

Book: 31%

AncoraSIR.com

Experiment 1-3 Object Recognition

Compare with other algorithms and focus on error performance



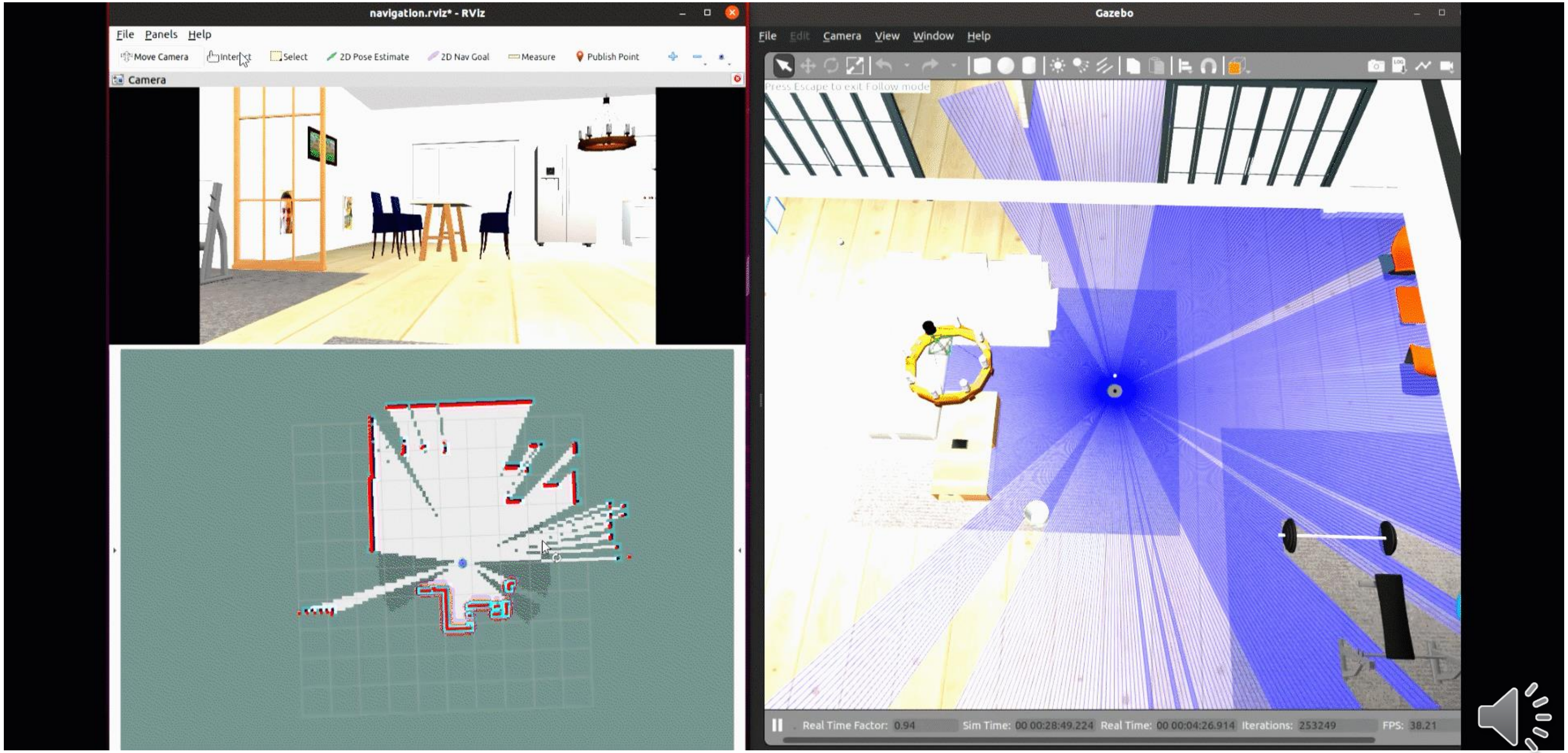
YOLOv4 three times faster than EfficientDet

YOLOv4 improve AP and FPS 10% and 12% from YOLOv3

AncoraSIR.com

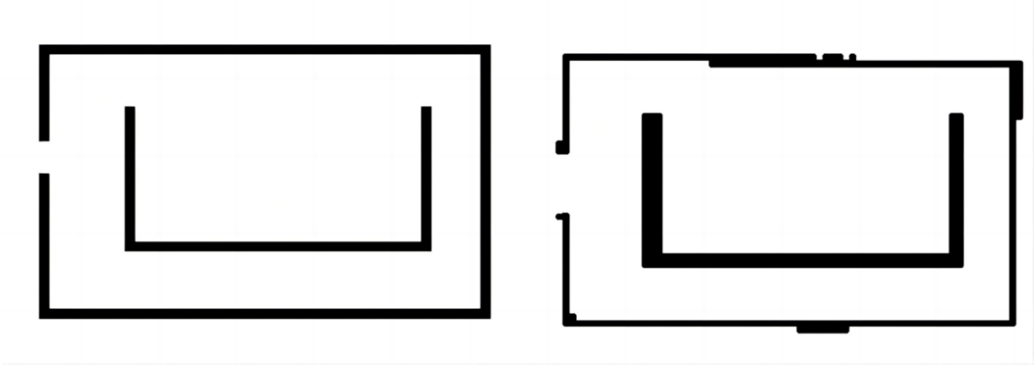
Experiment 2 - Slam and Navigation

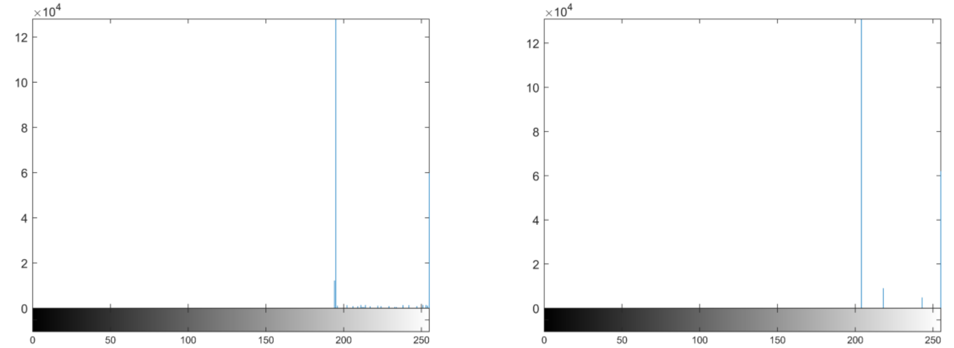
The process of self-mapping and navigation

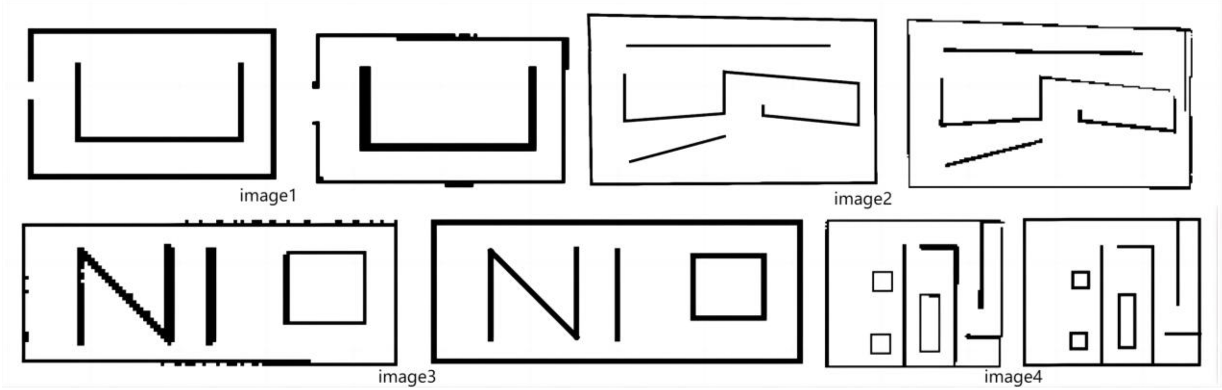


AncoraSIR.com

Experiment 2 - Slam and Navigation





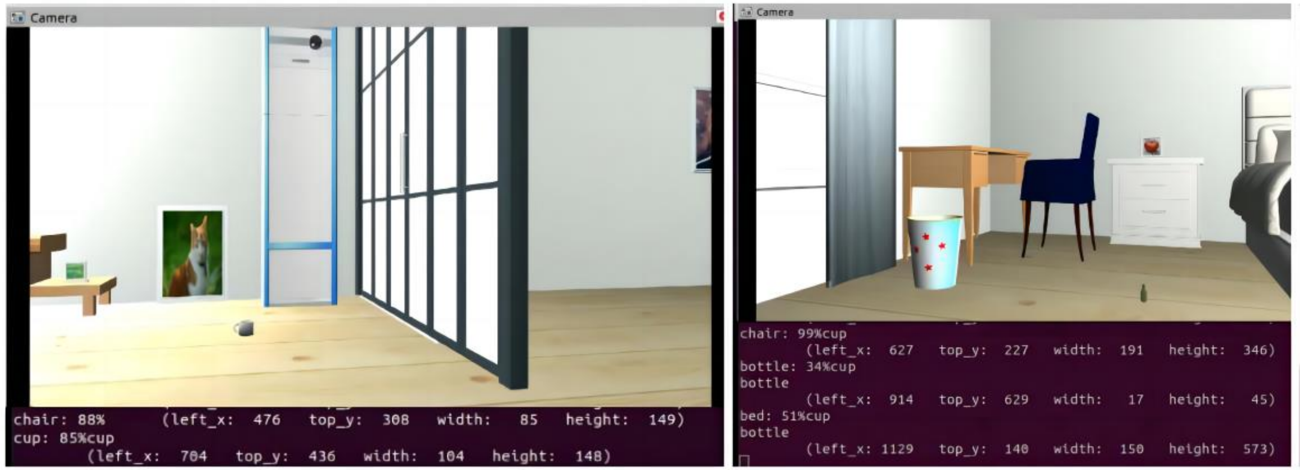


Simulations Experiments				
	image1	image2	image3	image4
RMSE	20.1239	91.4085	92.3797	94.5178
SSIM	0.8528	0.7035	0.6576	0.7171

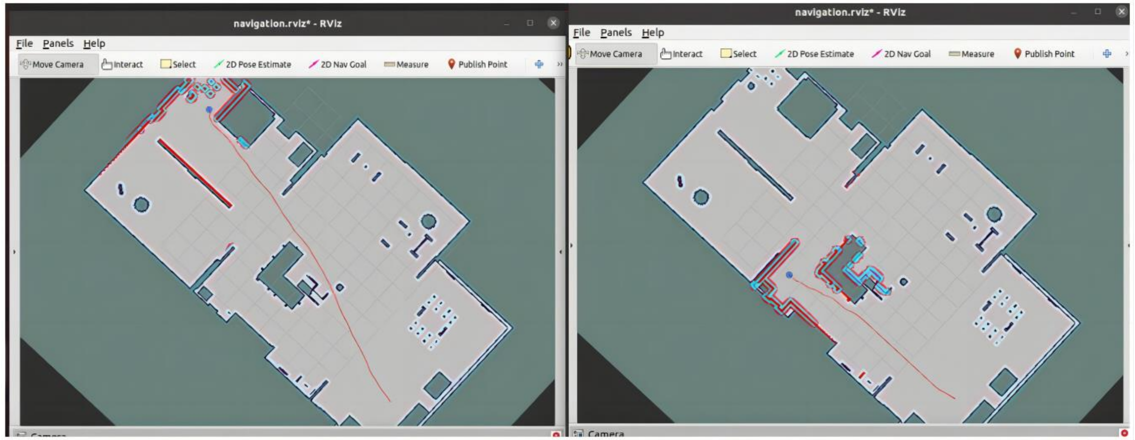
AncoraSIR.com

Experiment 3 - Integrated process

The garbage we need for object recognition can be classified.

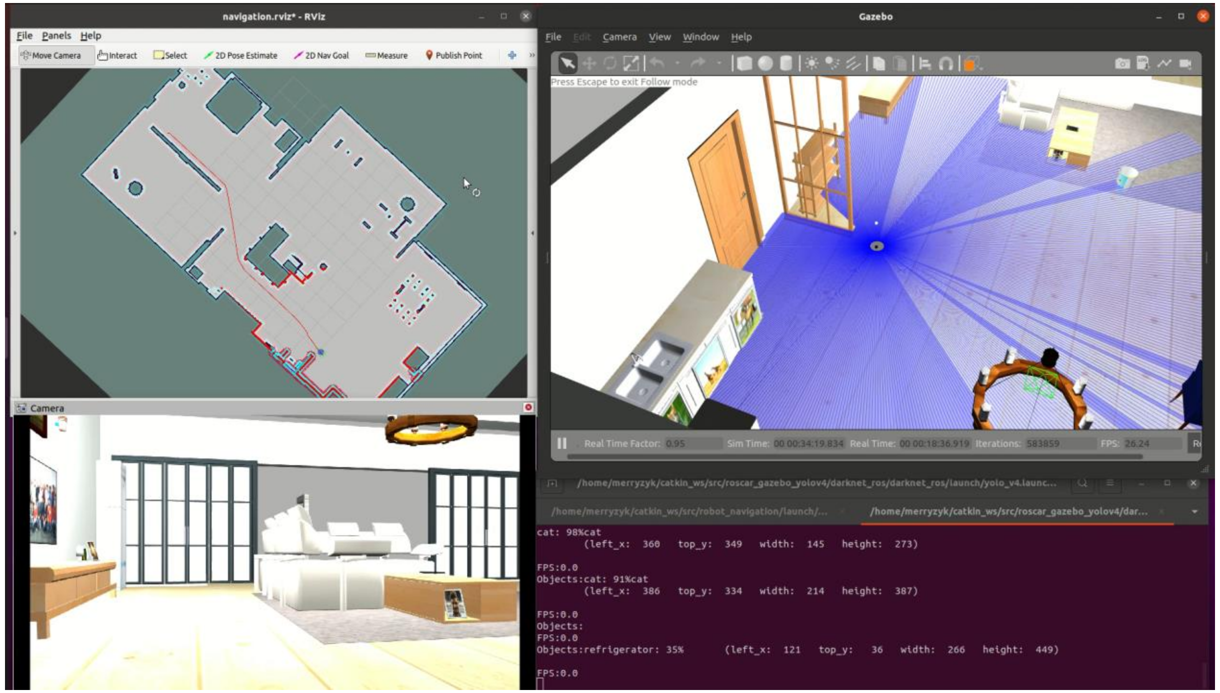


The map of the successful navigation path back to the trash bin area after picking up the trash

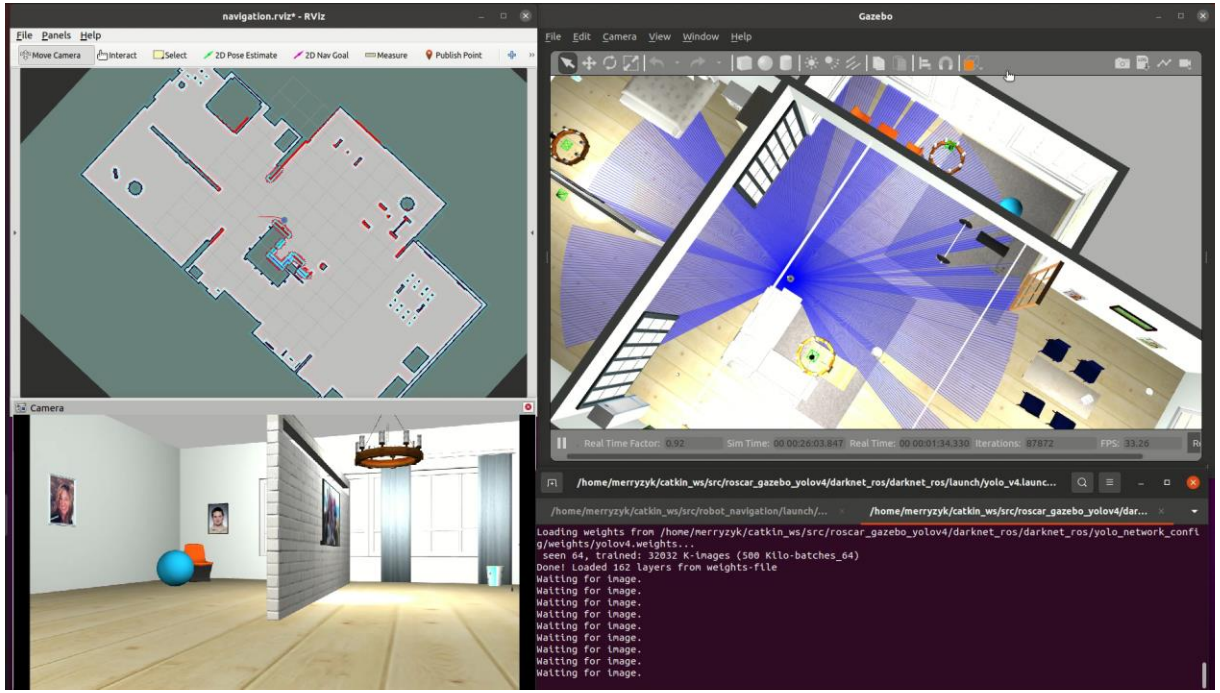


AncoraSIR.com

Experiment 3 - Integrated process



Bottle



Cup

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Experiment 3 - Problem

It is believed that it is caused by the error between the object in the simulation environment and the model trained by the real image

AncoraSIR.com

Conclusion

Conclusion

Task completion degree:

- Object Recognition
- SLAM and Navigation
- Integrated Process

Limitations:

- **model and environment** – more stable structures and diverse environment
- **object recognition** – used more advanced algorithms and more obvious features
- **navigation and mapping** – advanced mapping and more intelligent autonomous navigation

AncoraSIR.com