RoboCup@Home 2023: Stickler for the Rules Task Solutions

Tomoya Shiba
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Hakaru Tamukoh
Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Email: shiba.tomoya627@mail.kyutech.jp, tamukoh@brain.kyutech.ac.jp
https://www.brain.kyutech.ac.jp/~tamukoh/

Abstract

This paper proposes a robot vision system that detects illegal persons who violate house rules such as wearing shoes in a home environment. Such complex vision systems often require multiple AI systems including person detection, object detection, and more. Our approach simplifies this by leveraging the combined capabilities of Grounding DINO and SAM to detect rule violations effectively. The success of our method was proven at RoboCup@Home 2023, where it secured the highest score among all participating teams.

Keywords: Service robot, Object recognition, Mobile manipulator, RoboCup@Home,

1. Introduction

Home service robots are gaining increased attention due to the needs of aging societies with declining birth rates [1], [2], [3]. Their key functions include object recognition, picking and placing, identifying people and environments, and interacting with humans. Our research has involved using TOYOTA's home service robot HSR [4] and our Exi@ robot [5]. Various studies focusing on the social implementation of service robots for home use have been explored, often presented in competitions [6], [7], [8], [9]. These studies have mainly been proposed using dedicated recognition models. However, home environments vary greatly, and even within the same household, the objects to be recognized change over time and with the seasons. In RoboCup@Home competitions, the objects are usually announced in advance, leading most teams to develop recognition systems tailored to these specific objects [10]. However, some competitions require the detection of items not disclosed beforehand, such as people, their clothing, and common household objects. Therefore, it has been challenging to prepare recognition systems that can distinguish a wide variety of shoes to enforce a no-shoes rule or ensure adherence to a dress code at a party.

In this study, we propose a method for detecting rule violations in the competition task "Stickler for the Rules Task" using the detection models Grounding DINO [11] and Segment Anything [12], which employ prompt tuning for general object recognition. We implement these methods on TOYOTA's home service robot HSR. As a validation, we used them in RoboCup@Home [13], a competition for home service robots, to verify their effectiveness. (Fig. 1)

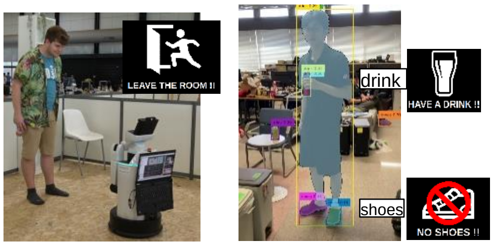

Fig. 1. Detection of rule violations

2. Related works

2.1. Object recognition

We have been researching methods for home service robots to detect the location of people and objects and then perform some action in response. Specifically, we have studied methods for tidying up and detecting people within a home. In our previous research, we developed a method for estimating the position of objects and available space, utilizing YolactEdge [14] and point cloud information [15].

2.2. *Problem*

The home environment contains many objects, including furniture such as desks and shelves. Additionally, the variety in human clothing is vast, requiring recognition of entirely different items depending on the season. Therefore, preparing a model that accounts for all these variations in a short period is challenging, and repurposing existing DNN models may lead to unstable recognition outcomes.

3. Proposed Method

3.1. *RoboCup@Home: Stickler for the Rules Task*

Here, I will explain the "Stickler for the Rules" task in RoboCup@Home. This task is designed for home service robots to identify party guests who have broken house rules and politely ask them to stop. The following four house rules are set:

1. All guests must take off their shoes at the entrance.
2. No guests are allowed in the Black Room.
3. Guests are not allowed to leave garbage on the floor.
4. All guests must always have a drink in hand.

3.2. *Proposed Method*

We propose a method for detecting violations using two Transformer-based models capable of prompt tuning. Grounding DINO is used for detecting bounding boxes, and Segment Anything is employed for segmentation tasks within the bounding box area. This allows for the identification of a wide variety of clothing and objects held in hands (Fig. 2). Additionally, to obtain the three-dimensional position, we use the pixel coordinates of the centroid of the bounding box and depth images.

Fig. 2. SAM (Segmentation) & Grounding DINO (BBox)

4. Experiments and Results

4.1. *Partial Recognition*

To detect rule violations within a home, high recognition accuracy, as well as ease of surveillance, are required. To evaluate these performances, we tested the recognition accuracy of partially visible individuals. In this experiment, we prepared data with humans partially visible in images and conducted a comparison between YOLO v8 [16], MMDetection [17], and our proposed method. The images used were from the open-source human dataset DensePose-COCO [18], and we created images divided into quarters, reducing the image size to half, to test recognition accuracy. The results are summarized in (Table 1)
As a result, our proposed method showed the highest recognition performance in the same task. However, it was also shown to be the most time-consuming method.

Table. 1. Results of Partial Recognition (Stickler for the Rules)

Method	Detection rate [%]	Speed[s]
Proposed Method	**43.2**	1.122
Yolo v8 (Detect)	23.1	**0.103**
MMDetection	24.2	0.239

4.2. *RoboCup@Home task*

At RoboCup@Home 2023, we tested the effectiveness of this task by featuring more than 5 individuals committing rule violations, arranged in 3 different patterns of person placement. The results are summarized in (Table 2). As a result, we were able to perfectly address two types of rule violations and achieve the highest score.

Table. 2. Results of RoboCup@Home2023 (Stickler for the Rules)

Team	Score
HMA (Our team)	**1000**
Tidyboy	700
TRAIL	**1000**
Tech United	800
eR@sers+ Pumas	300
RoboCanes-VISAGE	200

5. Discussion

The proposed method has demonstrated higher recognition accuracy for partial visibility compared to existing methods. However, since it is based on a Transformer model, it tends to have longer processed times. When implementing this method as a function of home service robots, it may be effective to use the existing models for fast recognition in some scenarios, and our method for its strong recognition accuracy in cases of partial visibility. Additionally, the recognition method using prompt tuning is often influenced by the skill of the prompt setter.
In the future, it will be necessary to establish objective criteria for prompt tuning.

6. Conclusion

In this study, we proposed a method for utilizing recognition models capable of prompt tuning to address the "Stickler for the Rules" task in RoboCup@Home, and conducted experiments in RoboCup tasks. The results showed that our method was effective, achieving the highest score in the same task. In the future, we aim to reclassify similar objects based on recognition scores and other information to reduce misrecognition.

References

1. T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara, and K. Murase, ROBOMECH Journal, 2019.
2. L. Iocchi, D. Holz, J. Ruiz-del-Solar, K. Sugiura, T. van der Zant, Artificial Intelligence, pp. 258-281, 2015.
3. H. Okada, T. Inamura, and K.Wada, Advanced Robotics, 2019.
4. Yamamoto, T., Terada, K., Ochiai, A., Saito, F., Asahara, Y., & Murase, K. (2019). Development of human support robot as the research platform of a domestic mobile manipulator. ROBOMECH journal, 6(1), 1-15.
5. Hori S, Yutaro I, Kiyama Y, et al. Hibikino-Musashi@Home 2017 team description paper. Preprint. 2017. Avail-able from: arXiv:1711.05457
6. Savage J, Rosenblueth DA, Matamoros M, et al. Semantic reasoning in service robots using expert systems. Robot Auton Syst.2019;114:77–92..
7. Ishida Y, Morie T, Tamukoh H. A hardware intelligent processing accelerator for domestic service robots. Adv Robot. 2020 June;34(14):947–957.
8. Yoshimoto Y, Tamukoh H. FPGA implementation of a binarized dual stream convolutional neural network for service robots. J Robot Mechatron. 2021;33(2):386–399
9. Taniguchi, A., Isobe, S., El Hafi, L., Hagiwara, Y., & Taniguchi, T. (2021). Autonomous planning based on spatial concepts to tidy up home environments with service robots. Advanced Robotics, 35(8), 471-489.
10. T. Shiba, A. Mizutani, Y. Yano, et al,. "Hibikino-Musashi@Home 2023 Team Description Paper," arXiv:2310.12650, 2023.
11. Liu, S., Zeng, Z., Ren, T., et al.: 'Grounding DINO: Marrying DINO with Grounded PreTraining for Open-Set Object Detection', 2023, arXiv:2303.05499
12. Kirillov, A., Mintun, E., Ravi, N., et al.: 'Segment Anything', 2023, arXiv:2304.02643
13. "RoboCup Federation website" (2023 12/15 accessed)
14. Liu, H., Soto, R. A. R., Xiao, F., & Lee, Y. J. (2021, May). Yolactedge: Real-time instance segmentation on the edge. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 9579-9585). IEEE.
15. T. Shiba, T. Ono, and H. Tamukoh, Object Search and Empty Space Detection System for Home Service Robot, The 2023 International Conference on Artificial Life and Robotics (ICAROB2023) OS17-2, Online, February 9-12(10), 2023.
16. Ultralytics.: 'Ultralytics YOLOv8 Docs', https://docs.ultralytics.com/, (Accessed on 12/14/2023).
17. Chen, K, et al. "MMDetection: Open mmlab detection toolbox and benchmark." arXiv preprint arXiv:1906.07155 (2019).
18. Güler, Rıza Alp, Natalia Neverova, and Iasonas Kokkinos. "Densepose: Dense human pose estimation in the wild." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

Authors Introduction

Mr. Tomoya Shiba

He received the B.Eng. degree from National Institute of Technology, Kagoshima College, Japan, in 2021. He received the M.Eng. from Kyushu Institute of Technology, Japan, in 2023. He is currently in a Ph.D. student in the graduate school of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interest includes image processing, motion planning, and domestic service robots.

Dr. Hakaru Tamukoh

He received the B.Eng. degree from Miyazaki University, Japan, in 2001. He received the M.Eng and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2003 and 2006, respectively. He was a postdoctoral research fellow of 21st century center of excellent program at Kyushu Institute of Technology, from April 2006 to September 2007. He was an assistant professor of Tokyo University of Agriculture and Technology, from October 2007 to January 2013. He is currently an associate professor in the graduate school of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research interest includes hardware/software complex system, digital hardware design, neural networks, soft-computing and home service robots. He is a member of IEICE, SOFT, JNNS, IEEE, JSAI and RSJ.