RoboFEI@Home Team Description Paper for the Brazilian Robotics Competition 2024

Cauan V. E. de Sousa, Gabriela Bassegio, Mateus Scarpelli,
Guilherme Morais Escudeiro, Gustavo Rosell Collado, Matheus Pereira Rodrigues,
Gustavo Garcia Bagio, Fagner de Assis Moura Pimentel and Plinio Thomaz Aquino-Junior
FEI University Center
São Bernardo do Campo, São Paulo, Brazil
plinio.aquino@fei.edu.br

***Abstract*—The RoboFEI@Home team, consisting of undergraduate and postgraduate students in mechanical, electronic, and computer engineering, has been developing a robotic platform for the Brazilian Robotics Competition since 2015 and RoboCup since 2016. This platform is designed to assist individuals with limited mobility and the elderly. The 2024 iteration introduces significant research advancements in object perception and recognition through motor effort and current analysis in the robotic manipulator. Additional enhancements include improved robot pose tracking and object segmentation using point clouds for refined object manipulation. Furthermore, an innovative approach to domestic robot programming incorporates gamification elements. This paper details the platform's technical evolution, emphasizing the collaborative nature of the project and its contributions to the field, with resources available in the team's repository https://github.com/robofei-home.**

***Index Terms*—Robotic Manipulation, Object Recognition, Assistive Robotics.**

I. Introduction

Since its inception in 2015, the RoboFEI team has been developing home assistance robots for the @Home league, beginning with the PeopleBot platform. The team has also been actively participating in the SSL and Humanoid Leagues since 2009 and 2013, respectively. The @Home robots, designed for home interaction, integrate safe human engagement with adaptive behaviors and specialized tasks [1]. Advanced academic projects have furthered research to enhance human-robot interaction, emphasizing behavior, design, and navigation [2]–[4]. Addressing the challenge of parts scarcity, the team has engineered an economical and sustainable robot design for a variety of domestic tasks. For the Brazilian Robotics Competition 2024, the team has improved the HERA robot's autonomy by incorporating new object recognition, pose tracking, and gamification techniques for programming, thus enhancing the robot's integration into household environments.

II. Research Focus and Interests

The RoboFEI@Home team, affiliated with the postgraduate program in Applied Artificial Intelligence and Robotics, is deeply invested in the dynamic field of human-machine interaction. This area of research, at the intersection of advanced technologies such as autonomous vehicles, robotic systems, and smart home automation, is pivotal in shaping the future of service robotics. Our university's strong foundation in Engineering and Computer Science drives our research and development efforts, dovetailing with the curriculum of both undergraduate and postgraduate courses that span mechanics, electronics, automation, robotics, materials science, and computer science.

Centro Universitário FEI - Fundação Educacional Inaciana

In addition to our core focus on human-machine synergy, the RoboFEI@Home group is dedicated to advancing methodologies, techniques, models, and algorithms across a spectrum of domains. These include adaptive interfaces, brain-computer interfaces, planning, intelligent automation for homes and buildings, autonomous systems, and the burgeoning field of the Internet of Things (IoT). Our collaboration extends into the realm of IoT with a particular emphasis on enhancing connectivity within robotic applications.

Our research interests encompass a broad scope but particularly focus on mechanical aspects, embedded electronics, sensor technology, battery management, and embedded computing. Additionally, our team has a constant interest in the creation of new robotic platforms. We are currently researching the redesign of our existing platform to make it more efficient in various aspects, including mechanical efficiency, electronic connectivity, and battery portability.

These interests are both theoretical and practical, aimed at addressing real-world challenges associated with designing and implementing intelligent, autonomous robotic systems. Through this multifaceted research approach, the RoboFEI@Home team aspires to contribute significantly to the field of robotics, forging new paths in the interplay of machines and the everyday lives of humans.

III. Team Milestones, Competitions, and Partnerships

The RoboFEI@Home team has garnered accolades and driven robotic innovation, particularly noted for its consecutive victories at the Brazilian Robotics Competition, a testament to our standing in South American robotics. Our success extended to the international stage at RoboCup 2022 in Thailand, where we secured first place in the Open Platform League.

Investments from our university into a dedicated research space compliant with RoboCup standards have significantly bolstered our experimental capabilities, allowing for refining techniques and strategies in real-world scenarios. This has led to substantial improvements in areas such as human interaction, object recognition, and advanced navigation.

Recent advancements leading up to RoboCup 2024 have seen our team achieve significant strides in object perception and recognition, employing sophisticated analyses of motor effort and current in our robotic manipulator. Our team has greatly enhanced our robots' manipulation skills, coupled with developments in robot pose tracking and point cloud-based object segmentation. Additionally, the introduction of gamification into robot programming has made our approach more compelling and improved its effectiveness.

In 2024, our team participated in RoboCup in the Netherlands and achieved sixth place, competing in various Level 1 and Level 2 challenges. This participation highlights our continuous commitment to excellence and innovation in the field of robotics.

The RoboFEI team also plays a pivotal role in organizing the Brazilian Robotics Olympiad, an event that engages elementary school students and fosters their learning in the field of robotics, encompassing all facets of service robotics. Hosted by the FEI University Center, the event saw the participation of approximately 120 teams from across the state of São Paulo, with a total of 814 competitors in the 2023 edition. In 2024 we will continue collaborating on this initiative, with increasing numbers of participants, enhancing robotics in schools. This initiative not only promotes educational development in robotics but also serves to inspire the next generation of engineers and innovators.

IV. Methodology for Tackling RoboCup@Home Challenges

The RoboCup@Home competition plays a crucial role in advancing autonomous service robots for domestic environments, evaluating them across a range of household tasks. Robots are assessed on their ability to interact and cooperate with humans, navigate and map dynamic spaces, excel in computer vision and object recognition under natural lighting conditions, and perform precise object manipulation. The competition also evaluates platform robustness, behavioral integration, and overall system intelligence and cohesiveness [5].

A. Robot Vision

Object detection is accomplished using the YOLOv8 architecture within the PyTorch framework, enhancing detection accuracy. Synthetic data generation has expedited dataset creation, improving image tagging and expanding dataset volume. A Deep Salient Object Detection algorithm has been implemented for precise background segregation, isolating pertinent objects and generating binary object masks. These masks enable the synthesis of images against variable backdrops, refining the detection process. Additionally, image segmentation via color extraction has been utilized for precise object manipulation without necessitating extensive model retraining for individual object recognition, thus optimizing training time. A people recognition system has been integrated for human interaction, employing the dlib library for facial and landmark recognition, supporting a broad spectrum of interactive tasks.

B. Voice Recognition

The team has transitioned to the Whisper API for voice recognition, replacing the previous Google Speech Recognition API. This integration is supported by a specialized ROS package, optimized for Ubuntu, to enhance performance. The Whisper API is particularly adept at adapting to diverse speech patterns and word choices, making it suitable for various environments. The updated system efficiently processes voice commands and seamlessly interfaces with the robot's core system, ensuring robust and accurate voice command recognition across different settings.

C. Manipulation

The robotic manipulator, engineered by the mechanical team, mimics the human arm's degrees of freedom, enhancing domestic task efficiency and human-robot interaction by applying the anthropomorphic principle. Studies of human upper limb anatomy and kinesiology, emphasizing extension and flexion movements, informed its design.

Material innovation has been integrated into the manipulator's design, with 3D printing utilized for complex shapes and carbon fiber for flatter components, achieving increased durability and reduced weight and size.

The Dynamixel Workbench package is utilized for direct kinematics in basic movements for the manipulation system. For advanced trajectory planning and precision, MoveIt with inverse kinematics is employed. OctoMap integration with the manipulation system ensures optimized and safe operations, incorporating environmental perception from vision systems into the robotic arm's trajectory planning, which enhances the success rate of ”pick and place” actions.

D. Robot Navigation and Social Navigation

Autonomous robot navigation necessitates mapping, spatial positioning, and optimal route determination capabilities, facilitated by sensors that convert environmental data into navigable paths through Simultaneous Localization and Mapping (SLAM). This process involves real-time corrections of path errors to avoid obstacles and select the most efficient route.

An ontological framework was developed to guide socially acceptable robot navigation and approaches to individuals and groups, integrating semantic mapping and social navigation principles [3]. The robot's reasoning on this ontology informs its navigation strategies, adjusting its trajectory and approach based on proxemics and interaction zones, considering social norms. The outcome is a computational model enabling mobile social robots to interact suitably within social contexts, aiming to minimize human discomfort [4].

V. Ongoing Research and Scientific Contributions

All the research topics discussed in this section represent functionalities currently under development and embedded in our robot, poised to be demonstrated at the upcoming competition.

A. Object Perception and Recognition Using Motor Effort and Current in Robotic Manipulators

The research project has advanced the manipulation module of the Hera robot, achieving synergistic development with the computer vision module for RoboCup 2024. The project's primary objective is to enable the robot to identify unknown objects and ascertain their positions solely through the generated Point Cloud and feedback from the motor effort exerted by the gripper. The motor's current during the grasping process indicates the object's category or even its specific identity. This inference is facilitated by the use of 3D printed flexible material in the robot's gripper, which molds to the contours of the object, allowing for enhanced interaction.

Experiments have been conducted to validate the system's object perception and recognition capabilities. The motor control system's response to varying degrees of effort and current was meticulously recorded in these trials as different objects were grasped. The results consistently correlate the motor's effort feedback and the object's type, size, and required grip strength. Objects ranging from soft textiles to rigid geometrical shapes were successfully categorized, underscoring the system's versatility (figure 1).

The practicality of this method was highlighted in experiments where the robot performed a series of object manipulations without prior knowledge of the items. The gripper's adaptive behavior, coupled with motor current analysis, reliably classified objects in real-time, showcasing the potential for this technology in scenarios where quick and accurate object identification is critical.

This research contributes to the field of robotic manipulation by demonstrating that motor effort and current can be effective indicators for object recognition. The findings suggest significant potential for applications in domestic robots, where the ability to adapt to various objects is paramount. The team's approach presents a step forward in robotic perception, potentially reducing the reliance on extensive object databases and complex vision systems for object identification tasks.

B. Point Cloud Segmentation for Autonomous Service Robots

Object segmentation with point clouds presents a significant challenge in robotics, essential for enabling robots to manipulate objects in unstructured environments. Point clouds, capturing three-dimensional information through sensors like 3D cameras and laser scanners, provide detailed environmental data, crucial for decision-making in tasks such as base positioning, object approach, and gripper actions.

The process of object segmentation involves extracting relevant information such as shape, position, and orientation from a point cloud dataset. Techniques for segmentation include region-based approaches, which divide the scene into distinct object regions, and edge-based approaches, which detect object edges to facilitate segmentation.

Fig. 1. Sensor interpretation using OctoMap in manipulation system.

Robust object segmentation is vital across various robotic applications, including industrial automation, hazardous environment cleaning, healthcare assistance, and domestic services. By accurately identifying and manipulating objects, robots can perform tasks such as picking, moving, or stacking items efficiently and safely.

This research advances the field by developing sophisticated object recognition and manipulation techniques through point cloud analysis. By enabling robots to identify and locate three-dimensional objects in unknown environments, we push the boundaries of autonomous operation in various contexts, from automated factories to residential settings.

The methodology involves creating a synthetic dataset using NVIDIA Omniverse Replicator, simulating high-fidelity environments with randomized elements to enhance data diversity. This dataset trains a convolutional neural network (CNN) to detect and segment objects in real-time, followed by motion planning to determine optimal manipulation trajectories.

Experiments validate the proposed method's effectiveness, using a combination of depth cameras like RealSense and the HERA robot platform [1], [6]. The integration of advanced simulation tools and real-world testing ensures that the system adapts to dynamic environments, improving the precision and efficiency of robotic tasks.

Preliminary results demonstrate the method's scalability and efficiency in generating labeled data, significantly reducing the need for manual annotations. This approach enhances training data quality and accelerates the development process, as shown in Figure 6.

Future work focuses on addressing the limitations of noise in natural environments by incorporating machine learning techniques for edge detection. This approach aims to improve segmentation accuracy and robustness, enabling effective autonomous object manipulation in real-world conditions. Comprehensive experiments will validate the new method's performance across diverse scenarios, ensuring reliability before practical deployment.

By overcoming current challenges and continuously refining our methodology, this research contributes to the advancement

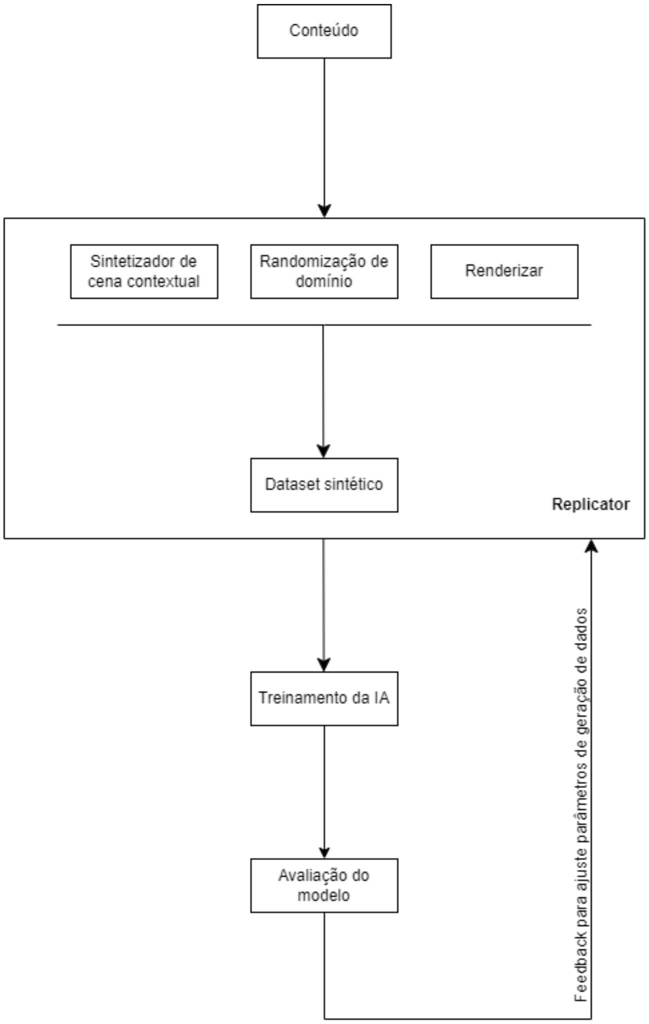

Fig. 2. Methodology Using Replicator

Fig. 3. The output of segmentation from step 1

of service robotics, setting new standards for robot interaction in complex, unstructured environments.

C. Detection and Analysis of Joints and Movements Applied in the Context of Robotics

In the domain of service robotics, the utilization of sophisticated computer vision systems, such as the OpenPose

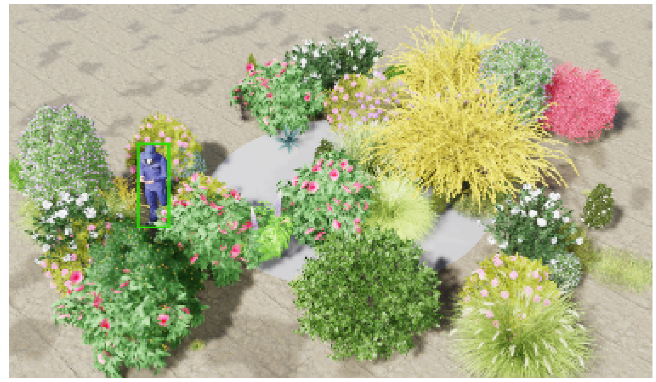

Fig. 4. The output of 2D marking from Step 1

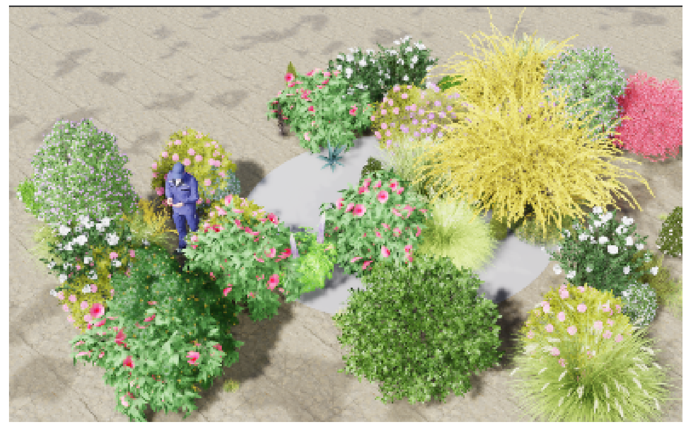

Fig. 5. The output of the camera from Step 1

algorithm for human pose tracking, has been instrumental. Originally aimed at the meticulous observation of physiotherapy exercises via joint detection, the scope of application extends significantly into the operational domain of assistive robotics.

Employing the OpenPose algorithm, robots have been endowed with the ability to accurately track human poses in real-time. This technology has been integrated with the robot's processing units to allow for immediate and responsive interaction with users, adapting to their movements and gestures. Such advancements have facilitated the robot's capacity to preemptively recognize potential hazards, such as fall risks or ergonomic discrepancies in elderly users' postures, offering swift assistance or alerting support systems.

In the home environment, robots equipped with pose tracking can now adjust their actions to the residents' body language, thereby personalizing task execution, whether it be household chores or engaging in interactive activities. This level of adaptability is achieved through continual feedback loops where the robot's sensors and motors adjust in response to the pose data processed by OpenPose.

Furthermore, the research has extended to telemedicine applications. Robots serve as data collection points, accurately capturing patient movements for remote diagnostics and treatment efficacy assessments in physiotherapy, with the

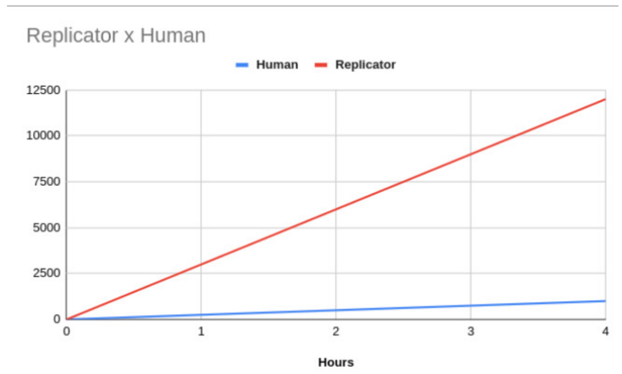

Fig. 6. Comparative between images labeled by human and replicator per hour

OpenPose algorithm ensuring high fidelity in data.

Institutional settings, such as nursing homes and hospitals, benefit from robots that continuously monitor patient safety, using pose detection to assist mobility and prevent accidents.

This research has been substantiated through testing. Experiments have been conducted where robots, utilizing the OpenPose algorithm, have successfully navigated complex human interactions. They have demonstrated proficiency in recognizing and mimicking human joint movements, highlighting the potential for these systems to contribute meaningfully to the field of assistive robotics.

Such technological integration represents significant progress in human-machine interaction, with the OpenPose algorithm at the forefront of this transformation. The ongoing research and development in this area promise to yield functional and empathetic robots that cater to the nuanced needs of human companionship and assistance.

VI. Advanced AI and Computer Vision in Robotics

In recent years, the intersection of artificial intelligence (AI) and robotics has transformed how robots perceive and interact with their surroundings. Key advancements in computer vision, driven by deep learning, have enabled robots to ”see” and understand their environment. Notably, the implementation of YOLO v8, a real-time object detection system, and GPT Vision, which integrates visual comprehension with natural language processing, represent significant milestones.

YOLO v8, developed by Ultralytics, enhances object recognition with its speed and accuracy, while GPT Vision extends cognitive abilities, enabling robots to identify and describe complex visual contexts. These technologies improve object manipulation, autonomous navigation, and human-robot interaction.

The fusion of AI and robotics defines a new era of service robots capable of learning, adapting, and making decisions based on intelligent environmental interpretation. Computer vision, a cornerstone of AI in robotics, allows robots to visually process and understand their surroundings, crucial for tasks ranging from navigation to object manipulation and safe human interaction. Deep learning advancements have enabled robots to perform tasks requiring object identification and contextual understanding.

Technologies like GPT Vision and YOLO v8 exemplify this synergy. GPT Vision provides deep scene analysis and detailed descriptions, while YOLO v8 excels in real-time image processing, crucial for rapid responses in dynamic environments. Other AI techniques, such as Convolutional Neural Networks (CNNs) and Reinforcement Learning, also enhance visual perception and adaptability in varied scenarios, expanding service robot applications in healthcare, hospitality, and more.

This research evaluates YOLO v8 and GPT Vision’s effectiveness in object detection and description. Five images were analyzed, representing varied scenarios challenging AI’s detection and description capabilities. YOLO v8 was used for real-time object recognition and classification, while GPT Vision provided detailed narrative descriptions of the same images, focusing on contextual depth.

Listing 1. Simple Use Example of YOLOv8

```python
from ultralytics import YOLO
model = YOLO('yolov8n.pt')
results = model(['im1.jpg', 'im2.jpg'])
for result in results:
    boxes = result.boxes
    masks = result.masks
    keypoints = result.keypoints
    probs = result.probs
    result.show()
    result.save(filename='result.jpg')
```

The results highlight YOLO v8’s quantitative precision in object identification against GPT Vision’s qualitative contextual richness. The methodologies’ strengths and limitations were assessed, demonstrating their impact on specific computer vision tasks in robotics.

YOLO v8	GPT Vision (detection):	
1 bed	1 bed	
1 painting	1 painting	1 plant
2 decorative letters	2 decorative letters	1 sculpture
1 nightstand	1 nightstand	

Fig. 7. Elements Identified in the Tested Images

Image 1:

Description: The image shows a modern, minimalist bedroom with a white textured wall featuring a black and white zig-zag patterned artwork and metal decorative letters ”5” and ”C”. The bed has a medium-toned wood frame and a large turquoise blanket. To the right, a dark wood vintage nightstand with a plant, and to the left, a piece of furniture with a white sculpture. A neutral-colored rug lies under the bed on a light wood floor. The room is illuminated by natural light from an unseen window, creating an airy and inviting atmosphere with a blend of modern and vintage elements.

The integration of advanced AI technologies like GPT Vision and YOLO v8 transforms the capabilities of artificial intelligence, setting new standards for autonomous and effective machine-environment interaction. As these technologies evolve, they promise to enhance robotic systems' ability to perform complex tasks with precision and accuracy.

Future research will focus on overcoming current challenges, particularly noise in natural environments, by incorporating edge detection and machine learning techniques to improve segmentation accuracy and robustness. Comprehensive experiments in real settings will validate the new method's performance, ensuring reliability before practical deployment.

This research contributes to advancing service robotics, setting new standards for robot interaction in complex, unstructured environments. The continuous evolution of AI and computer vision promises to enhance human-machine collaboration and open new frontiers for innovation.

VII. Conclusion

The journey of the RoboFEI@HOME team reflects a steadfast commitment to advancing the field of domestic service robotics. Participation in the RoboCup has been instrumental in this process, offering a valuable platform for showcasing our technological developments and engaging in knowledge exchange with leading teams worldwide. The challenges and competitions we have faced have tested our designs and served as catalysts for innovation, driving the team to develop cost-effective and sustainable robotic solutions tailored to the dynamic needs of home environments.

The advancements made in the HERA robot for RoboCup 2024, including enhanced object perception, sophisticated pose tracking, and the introduction of gamification in robot programming, underscore our pursuit of excellence and the continuous evolution of our robotic platforms. The collaboration and competition inherent in RoboCup have significantly contributed to our growth, allowing us to benchmark our research against the best in the world while fostering an environment of mutual learning and improvement.

As we prepare for the upcoming competition, we are reminded of the importance of such international forums in the broader context of robotics research. They are not merely contests but opportunities for collective advancement and shared triumphs in the field of robotics. The RoboFEI@HOME team is proud to be a part of this global community, contributing to the ongoing dialogue that shapes the future of robotics and the role these extraordinary machines will play in enriching human lives.

References

[1] P. T. Aquino Junior, B. F. V. Perez, D. R. Meneghetti, F. A. M. Pimentel, G. N. Marostica, J. G. R. Amorim, L. C. Neves, L. I. Gazignato, M. Y. Gonbata, R. M. Souza, R. C. Techi, T. S. B. Meyer, V. S. M. Schmiedl, and W. Y. Yaguiu. Hera: Home environment robot assistant. In Centro Universitario FEI, editor, *Proceedings of the II Brazilian Humanoid Robot Workshop (BRAHUR) and III Brazilian Workshop on Service Robotics (BRASERO)*, volume 1 of *1*, pages 68–73, Sao Bernardo do Campo, Sao Paulo, Brasil, 4 2019. IEEE South Brazil Robotics & Automation Society Chapter.

[2] Thiago Freitas dos Santos, Danilo Gouveia de Castro, Andrey Araujo Masiero, and Plinio Thomaz Aquino Junior. Behavioral persona for human-robot interaction: A study based on pet robot. In Masaaki Kurosu, editor, *Human-Computer Interaction. Advanced Interaction Modalities and Techniques*, pages 687–696, Cham, 2014. Springer International Publishing.

[3] Fagner Pimentel and Plinio Aquino. Performance evaluation of ros local trajectory planning algorithms to social navigation. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 156–161, 2019.

[4] Fagner de Assis Moura Pimentel and Plinio Thomaz Aquino. Proposal of a new model for social navigation based on extraction of social contexts from ontology in service robots. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 144–149, 2019.

[5] RoboCup@Home. http://www.robocupathome.org/. Accessed on: May, 10th 2024.

[6] Guilherme Nicolau Marostica, Nicolas Alan Grotti Meireles Aguiar, Fagner de Assis Moura Pimentel, and Plinio Thomaz Aquino-Junior. Robofei@home: Winning team of the robocup@home open platform league 2022. In Amy Eguchi, Nuno Lau, Maike Paetzel-Prüsmann, and Thanapat Wanichanon, editors, *RoboCup 2022: Robot World Cup XXV*, pages 325–336, Cham, 2023. Springer International Publishing.