FBOT@Home CBR 2024 RoboCup@Home Team Description Paper

João F. S. S. Lemos Jardel D. S. Dyonisio Luís F. M. Quadros Cris L. Froes Pedro E. Nunes Marina Z. Rocha Gabriel A. Dorneles Murilo C. Bicho Nicolas F. Dias André P. Maurell Cassio N. S. Teixeira Eduardo A. D. Evangelista Luis F. F. Silva Manoela R. Suita Richard J. C. Assis José D. G. Ramos Luis H. M. Mattiazzi Igor P. Maurell Rodrigo S. Guerra Paulo L. J. Drews Jr

August 23, 2024

***Abstract*—The following description paper (TDP) describes the FBOT@Home team and its Brazilian Open Robot for In-home Service (BORIS), created by members of the NAUTEC group at the Federal University of Rio Grande (FURG), in Brazil. This document also describes the team's customized mechanical designs and every code package developed, provided to the league as public contributions, including: dataset generation unity, speech, behavior, world, simulation and more. All of the developed modules are open-sourced and available at: https://github.com/butia-bots.**

***Index Terms*—RoboCup Brazil, BORIS, Animatronic Face, DIY Mobile Base.**

I. Introduction

As the demand for intelligent and autonomous systems continues to grow, our team is dedicated to pushing the boundaries of what can domestic robots achieve. This paper presents a comprehensive description of our work, highlighting our recent research and changes made to further progress in domestic robotics.

During the past year, our team has undergone a significant transition, switching from our established domestic robot, DoRIS, to a brand new and more modular model, the Brazilian Open Robot for In-home Service (BORIS). This shift was driven by the need to address the limitations of a commercial mobile base. For BORIS' mobility, we have developed a DIY low-cost base which is not only more adaptable, but also allows a quick assembling and disassembling, which facilitates testing process and makes it easier to transport for competition travels.

BORIS, shown in Fig. 1, is a service robot consisting of a mobile platform, a torso (equipped with CPU and GPU units), an animatronic face and a manipulator. The team FBOT@Home is part of the NAUTEC research group. The group's goal is to research and develop robotics and automation solutions applied to real world problems.

With DoRIS, the team FBOT@Home, previously known as BUTIABots, has already taken part in several RoboCup@Home competitions, securing notable achievements:

- 3rd Place: Brazilian Robotics Competition 2018 (Robocup Brazil);
- 3rd Place: Brazilian Robotics Competition 2019 (Robocup Brazil);
- 3rd Place: Brazilian Robotics Competition 2020 (Robocup Brazil);
- 2nd Place: Brazilian Robotics Competition 2021 (Robocup Brazil);
- 3rd Place: RoboCup Bangkok 2022;
- 1st Place: Brazilian Robotics Competition 2022 (Robocup Brazil);
- 9th Place: RoboCup Bordeaux 2023;
- 1st Place: Brazilian Robotics Competition 2023 (Robocup Brazil).
- 11th Place: RoboCup Eindhoven 2024;

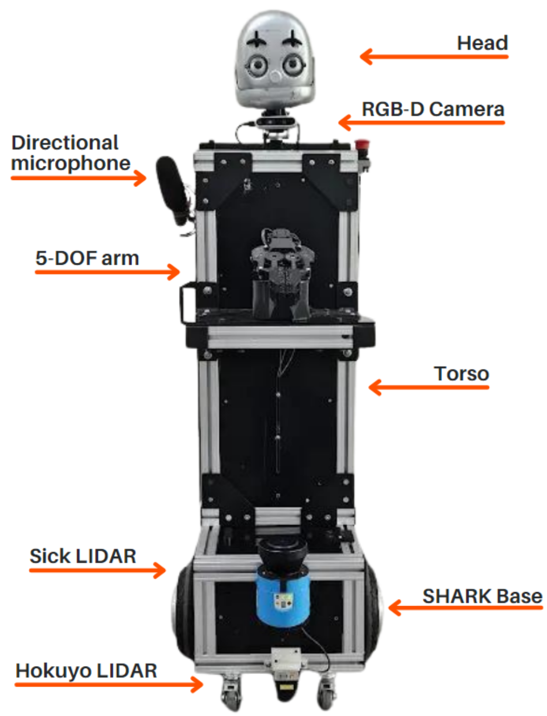


Fig. 1: BORIS' main differences compared to our previous model include a rectangular torso, a 5-DOF manipulator mounted on a platform, and the SHARK mobile base.

To introduce our team's work, this document is organized as follows: Sec. II presents our research, Sec. III presents our contributions to the league, and Sec. IV presents the conclusion.

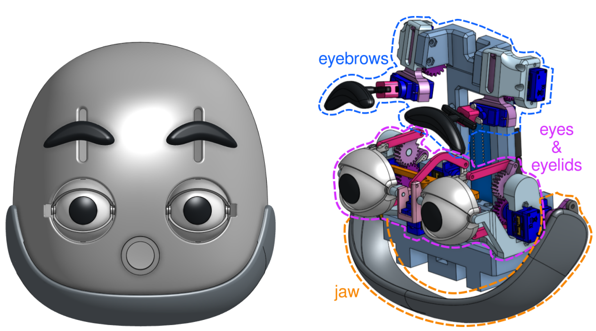
II. Research

The team maintains a strong commitment to research, constantly exploring new methods and edge technologies to push the boundaries of robotics and automation while also improving our members knowledge and abilities.

A. Head

One of the highlights of this robot design is its electro-mechanical 3D printed animatronic face. The head gestures are given by the pan and tilt motion of the neck, in addition to the movement of eyes, eyelids, eyebrows and jaw. Currently there are six pre-programmed standard expressions: happy, sad, neutral, angry, scared and sleepy. Additionally, there is a thread that triggers a routine of constant blinking motions, while the jaw is set to move accordingly to the speech process, both features being intended to make the robot appear more natural and approachable, what helps people to empathize with the robot, in order to facilitate human-robot interaction.

The face is used to associate emotions to the robot's progress in ongoing tasks. For example, if the robot succeeds in accomplishing a task, it makes a happy face, whereas if something is uncomprehended it changes to a sad expression. Over the years, mainly due to the several travels, the head had suffered irreparable damages and was not working properly. Consequently, it was decided to remodel the robot's face entirely, and this new version is already being tested and validated. Fig. 2 shows the new design elaborated with Onshape[1].



(a) External design (b) Internal systems

Fig. 2: Remodeled DoRIS' face

The structure is arranged in three major systems: eyes, eyebrows and jaw, as indicated in Fig. 2, the eyes system is subdivided in two: eyes pan and tilt motion and eyelids mechanism. This new model has been structured to be more robust than the existing version, guaranteeing more precise control. Additionally, it features a base support system that securely integrates all components, for easy maintenance access. It is important to note that this is an ongoing developmental project that has evolved beyond the scope of the robot itself, being part of a social behavior study. The objective of this research is to improve human-robot interaction, by assessing individuals' responses to an animatronic and chattering face totem present in the university lobbies.

[1]Official Website: https://www.onshape.com

B. Manipulation

The robot is equipped with the 5-DoF Interbotix WidowX-200 Robot Arm, which was relocated from a Interbotix LoCoBot. This arm employs Dynamixel XM430-W350 servo-actuators for the waist, shoulder, elbow and wrist angle joints and Dynamixel XL430-W250 servo-actuators for the wrist rotation and gripper joints.

In the previous version, with DoRIS, our team had developed a custom robotic arm. However, this arm caused numerous issues due to inaccuracies in the UDRF system, making it difficult to achieve high precision with its DIY components. Additionally, the deprecated Dynamixel servo-actuators used were outdated, lacked support, and had worn out, leading to movement loss. Consequently, with all the changes implemented for BORIS, the manipulation unit was also updated.

One of the disadvantages encountered with the WidowX-200 was the limited opening range of the gripper, which was insufficient for grabbing larger objects required in competition tasks. In order to solve this problem, it was decided to upgrade to a new gripper, 3D printed in PLA, with a scissor-like motion. This gripper was specifically designed to optimize both dimensions (width and range) and movement precision.

Another challenge that we have been facing is the gripper's weight limitation, as the WX-200 Robot Arm only supports a maximum working payload of 200 grams due to the servo's torque. To address this, a new upgrade is being formulated to accommodate the potential requirement of lifting heavier items in competitions.

We employ the Interbotix ROS Toolboxes[2] package, to enable the robot to execute arm movements in any Cartesian direction, unrestricted by specific axes. Additionally, our system incorporates the MoveIt[3] framework with OMPL[4] for motion planning. The integration of these tools ensures optimal arm movement, enabling the robot to use the most suitable tool for each specific situation.

Lately, our team has dedicated efforts to improve the safety of manipulation tasks. Leveraging our computer vision packages, we can precisely identify the midpoint of objects. This capability allows the robot arm to grasp items near their bases, ensuring secure handling. Upon release, the gripper can safely deposit the item on a flat surface, mitigating the risk of accidental falls.

C. Digital environment description and reconstruction using panoptic segmentation

We present a novel digitalization pipeline using panoptic segmentation to achieve faithful digital representations of real-world environments. Our pipeline employs posed images from a given scene and a method based on DM-NeRF for

[2]GitHub Repository: https://github.com/Interbotix/interbotix_ros_toolboxes
[3]Official Website: https://moveit.ros.org/
[4]Official Website: https://ompl.kavrakilab.org/

segmentation and reconstruction of the environment with no prior object detection or segmentation. The resulting three-dimensional meshes are then converted into XML files, categorized according to their panoptic classification, to produce an SDF file compatible with simulation software. A description of the proposed method can be seen in Fig. 3.

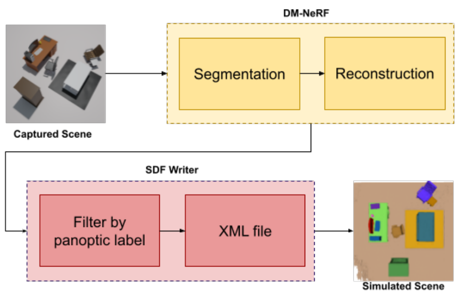


Fig. 3: The proposed pipeline: All individual objects are segmented and reconstructed with the appropriate labels, then transformed into the required format for simulation.

Quantitative evaluations of the scene reconstruction and segmentation processes yielded results with high fidelity and minimal noise. The pipeline was validated on the DM-SR dataset, which features eight typical household rooms. Our method produces accurate digital renditions of real-world environments and integrates them into the Gazebo simulation platform. This fact enables robots to navigate and interact within the virtual scene, Fig. 4, allowing system validation without the risks and costs associated with real-world experiments.

One of the applications for these features is the Robocup@Home competition, where the limited time allotted for each team in the arena makes creating a simulated environment both challenging and valuable, while the controlled nature of the environment serves as an ideal setting. Using the described method, posed images of the arena can be captured to quickly generate a simulated version of all objects in the scene.

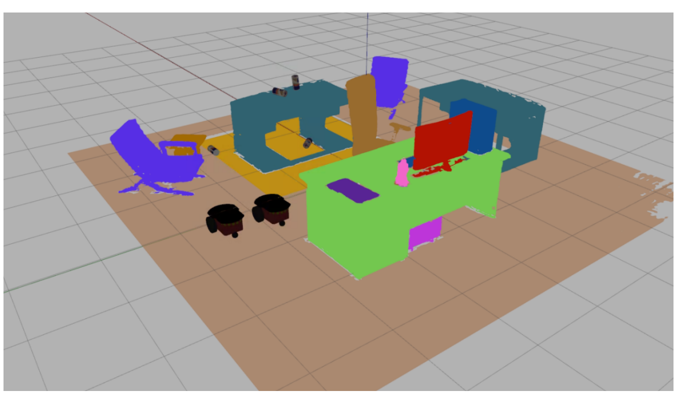


Fig. 4: Simulation of the reconstructed environment with newly added, removed, and replicated objects, alongside two robots positioned in the scene.

D. Pose detection and tracking using YoloV8-pose

YoloV8 [2] stands as the latest evolution in the Ultralytics convolutional neural network series, boasting an array of well-established variations, including their segmentation model. Notably, they have created a model that targets human joint detection, with high precision and the ability to handle joint occlusion, being the best choice for gesture detection, such as pointing to objects or raising hands. The package is also used to track people across multiple image frames, being extrapolated after by the team to the 3D space (as shown in figure 5), allowing the robot to follow someone even if there are many people present in the image.

(a) Front view (b) Side view

Fig. 5: YoloV8-Pose in 3D space

E. Open-source and open-hardware mobile base for service robotics

Currently, the team is actively involved in a binational research funded by the RoboCup Federation within the context of the RoboCup@Home Open Platform League. This collaborative project, conducted jointly between Brazil and Uruguay, aims to foster the development of service robotics in Latin America by providing a cost-effective solution for mobile bases, thereby enabling new researchers to kickstart their activities in this field.

Despite the availability of numerous mobile bases in the market, a prevailing challenge is their limited payload capacity or high cost. To address this, we are developing an affordable, high-capacity standard DIY robot base platform capable of carrying payloads up to 120kg, all while maintaining a cost below USD 2.000,00. Refer to Fig. 6 for a visual representation of the proposed platform.

In pursuit of this goal, we are repurposing the hardware found in self-balancing scooters, colloquially known as hoverboards. This process involves integrating new firmware to the motherboard [5] in order to allow communication with ROS. The adoption of this approach not only reduces costs but also ensures a robust and reliable system thanks to the well-established and proven components. The mechanical design is developed around widely available hardware parts such as the standard square cross-section extruded aluminum bars and connection accessories, industrial caster wheels, springs and other supporting structures.

During RoboCup Brazil 2023, as detailed in section III-C, we conducted a workshop in the Latin American Robotics Symposium, acquainting attendees with this project and providing comprehensive instructions on assembling a mobile base either akin to this version or tailored to their specific requirements.

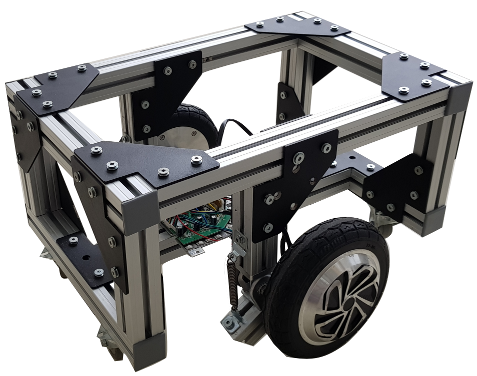


Fig. 6: SHARK mobile base.

F. Multi-Agent LLM Architecture for General Purpose Service Robots

In order to execute the General Purpose Service Robot task of the RoboCup@Home Competition, we propose a multi-agent architecture built on top of the open-source AutoGen [4] library. The architecture is composed of a set of large language model powered agents that interact with each other through chat, in order to accomplish goals given by a user. The decision about which agent to speak at a given time is made by a separate Group Chat Manager Agent.

The following functions are available for the agents to call, in order to interact with the control API of our robot.

1) Wait Door: Uses the LIDAR on the mobile base to wait until the entrance to the arena has been opened.

2) Navigate: Takes as input a named pose in the world model of the robot, and navigates to said pose.

3) Follow Person: Takes as input a named pose in the world model, navigates to it, and begin following the first person in front of the robot.

4) Answer Question: Listens for a question from a nearby human, and answers it by calling an LLM with retrieval augmented generation over a PDF file with provided domain knowledge about the competition and the questions of the GPSR task.

[5]Github Repository: https://github.com/EFeru/hoverboard-firmware-hack-FOC

5) Grasp: Takes as input a named pose to navigate to, and a list of object types to attempt to grasp at the given location. In case an object type that is not in the vocabulary of the custom trained YOLOv8 object detector is provided by the LLM agent, a feedback message with the available object types is provided as an answer. In the future, we intend to integrate a large multimodal model, such as the recently released GPT-4V [3] model API, in order to classify regions detected by a more generic object detector.

6) Place: Takes as input a named pose to navigate to, and places the object on the nearest place zone defined in the world model, after navigating to it.

7) Give: Takes as input a named pose to navigate to, and, after navigating to it, hands over the currently held object to a nearby person.

III. Contributions

The team has made significant contributions to the field of robotics through research, development, and the organization of events aimed at fostering and inspiring the study of robotics. Notable inclusions are the involvement our team had organizing the Brazilian Robotics Competition (CBR) in 2019, which was hosted at FURG, open source code and tutorials on Github, as well as the representation on the technical committee of the Brazilian league through one of the team's members.

A. Re-usability

All the ROS packages developed for our robot are open source and available on Github (https://github.com/butia-bots), complete with comprehensive documentation. Through our Butia Learning repository, we aim to provide a range of tutorials that cover the utilization of various systems commonly employed in domestic robotics. The primary objective of this initiative is to facilitate the training of new members of our team and enhance the seamless dissemination of information across RoboCup teams.

Our packages include all of the research described in the previous section. Notably, a significant portion of these packages can be adapted to development platforms beyond DoRIS, rendering them valuable for a diverse range of applications leveraging ROS.

B. An Open-Source Robot and Framework for Research in Human-Robot Social Interaction

Human-Robot Interaction (HRI) is essential to the widespread use of robots in daily life. Robots will eventually be able to carry out a variety of duties in human civilization through effective social interaction. Creating straightforward and understandable interfaces to engage with robots as they start to proliferate in the personal workspace is essential. Typically, interactions with simulated robots are displayed on screens. A more appealing alternative is virtual reality (VR), which gives visual cues more like those seen in the real world. In this study, it is introduced Jubileo, a robotic animatronic face with various tools for research and application development in human-robot social interaction field.

The physical segment of Jubileo project [1], is applied as DoRIS's robot face. Additionally, it gives a comprehensive framework to operate with a VR interface, enabling an immersive environment for HRI application tests and noticeably better deployment speed.

C. SHARK Mobile Base Workshop

In October 2023, supported by the RoboCup Federation, the team organized a workshop at the Latin American Robotics Symposium (LARS). The workshop's objective was to showcase our mobile base and offer comprehensive instructions to attendees for constructing a similar or modified plataform.

Throughout the workshop, our team members provided hands-on demonstrations, explaining the step-by-step process of repurposing all components from a hoverboard to create a cost-effective platform. This showcase not only highlighted the technical aspects of the process but also emphasized the economic benefits of adopting such an approach. Additionally, we provided guidance on integrating the system with ROS, offering an overview of its utility, particularly beneficial for about half of the attendees who declared having no prior experience with it.

We believe such initiatives are extremely important for the longevity of the league, as they actively foster the formation of new teams. Notably, 85% of the participants were students representing a diverse array of universities. Among these participants were members of aspiring @Home teams who explicitly expressed their interest in initiating their project using an affordable DIY platform.

IV. Conclusion

This paper outlines the approaches used by the FBOT@Home team for creating an intelligent system for home service robots, while presenting our new robot named BORIS. Our discussion covered the remodeled electro-mechanical 3D printed animatronic face, designed to enhance human-robot interaction by conveying emotions through expressions. Moreover, our research encompasses advanced pose detection and tracking using YoloV8, the development of a multi-agent LLM architecture for GPSR tasks, as well as a low-cost, high capacity mobile base developed with the aim of fostering advancements in home robotics, particularly in the context of Latin America.

References

[1] Jair A. Bottega, Victor A. Kich, Alisson H. Kolling, Jardel D. S. Dyonisio, Pedro L. Corçaque, Rodrigo da S. Guerra, and Daniel F. T. Gamarra. Jubileo: An open-source robot and framework for research in human-robot social interaction, 2022.

[2] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. YOLO by Ultralytics, January 2023.

[3] OpenAI. Gpt-4 technical report, 2023.

[4] Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Shaokun Zhang, Erkang Zhu, Beibin Li, Li Jiang, Xiaoyun Zhang, and Chi Wang. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. 2023.