

FBOT@Home Team Description Paper

RoboCup@Home Eindhoven 2024

João F. S. S. Lemos Jardel D. S. Dyonisio Emily G. Lamotte
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
Aline S. Marchiori Luis F. F. Silva Jardel J. P. Pardo
Eduardo A. D. Evangelista Igor P. Maurell Rodrigo S. Guerra
Paulo L. J. Drews Jr

November 27, 2023

Abstract. This team description paper (TDP) describes the team FBOT@Home and its Domestic Robotics Intelligent System (DoRIS), created by members of the NAUTEC group at the Universidade Federal do Rio Grande (FURG), Brazil. This document also describes our custom mechanical designs and public contributions we've made to the league, in the form of code packages developed by our team, such as: dataset generation unity, speech, behavior, world, simulation and more. All the developed modules are open-sourced and available at: <https://github.com/butia-bots>.

1 Introduction

Domestic robots have been an ambition of engineers for decades. Furthermore, the growing challenges and pressures of modern life lead to a future where the use of robots in a domestic environment will be common. Thanks to current advances in technology, this aspiration is finally viable.

Pursuing that goal, we present the *Domestic Robotic Intelligent System - DoRIS*, a domestic robot designed by the FBOT@Home team. *DoRIS* 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).

To introduce the focus of our team, this document is organized as follows: Sec. 2 presents our research, Sec. 3 presents our contributions to the league and Sec. 4 presents the conclusion. Finally, an overview finishes the paper in the last page.

2 Research

The team maintains a strong commitment to research, constantly exploring new methods and edge technologies to push the boundaries of robotics and automation.

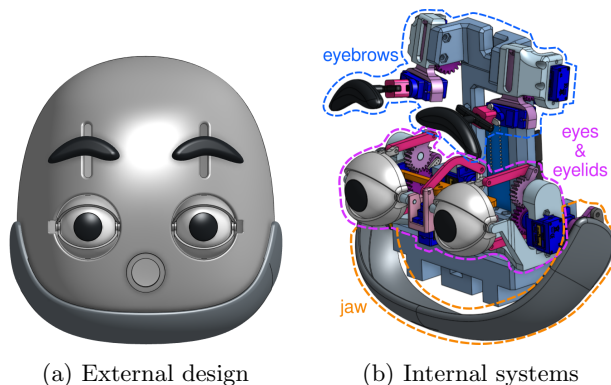
2.1 Head

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 five pre-programmed standard expressions: happy, sad, neutral, angry and scared. 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 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 has suffered irreparable damages and is not working properly. Consequently, it was decided to remodel the robot's face entirely, Fig. 1 shows the new design elaborated with Onshape¹.

The structure is arranged in three major systems: eyes, eyebrows and jaw, as indicated in Fig. 1, 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>



(a) External design (b) Internal systems

Fig. 1: Remodeled DoRIS' face

2.2 Manipulation

The robot includes a self-built 5DOF arm developed by the team, it employs Dynamixel MX-106T servo-actuators for the shoulder joints and Dynamixel MX-64T servo-actuators for the elbow, wrist and gripper joints. The structure of the links is mainly 3D printed in PETG polymer parts for the shoulder-elbow link and elbow-wrist, with some aluminum parts where a reinforced structure is required. The gripper is 3D printed in PLA and operates with a scissor-like motion, specifically designed to optimize both dimensions and movement precision.

We employ the Interbotix ROS Toolboxes² package, to enable the robot to execute arm movements in any Cartesian direction, unrestricted by specific axes. Additionally, our system incorporates the MoveIt³ framework with OMPL⁴ 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.

2.3 Streamlining object training with a low-cost 3D model generator

Achieving robust performance in object recognition tasks is of great importance. In past years, our team has consistently achieved exceptional results by meticulously capturing images and manually annotating objects. However, this traditional approach demands a lot of time from the team members, which is vital

² GitHub Repository: https://github.com/Interbotix/interbotix_ros_toolboxes

³ Official Website: <https://moveit.ros.org/>

⁴ Official Website: <https://ompl.kavrakilab.org/>

in a competition environment. With this in mind, we have developed a new approach to speed up this process. Our team previously employed Meshroom[1], an open-source tool that, while capable of generating reasonably high-quality meshes, presented challenges in terms of time efficiency. Thus, by utilizing the smartphone app Kiri Engine⁵ we were able to generate high-quality 3D models of many objects in an impressively short time frame.

While numerous commercial alternatives for 3D scanners exist, they often come with a substantial cost. The team, however, managed to seamlessly combine the impressive generation capabilities of Kiri Engine with the automation inherent in a scanner by adopting a cost-effective 3D printed design that incorporates smartphone support. Subsequently, the 3D models are employed in conjunction with our Dataset Generator package⁶ for the synthesis of data based on Unity Perception[2], a toolkit crafted around the principle of domain randomizations for sim-to-real transfer. Through our dataset generation tool, we can efficiently produce annotated ground truth data at scale, catering to tasks like object detection and instance segmentation.

2.4 Pose detection and tracking using YoloV8-pose

YoloV8[3] 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 2), allowing the robot to follow someone even if more people are present in the image.

2.5 Open-source and open-hardware mobile base for service robotics

Presently, we are 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

⁵ Official Website: <https://www.kiriengine.app/>

⁶ Github Repository: https://github.com/butia-bots/butia_dataset_generation_unity/

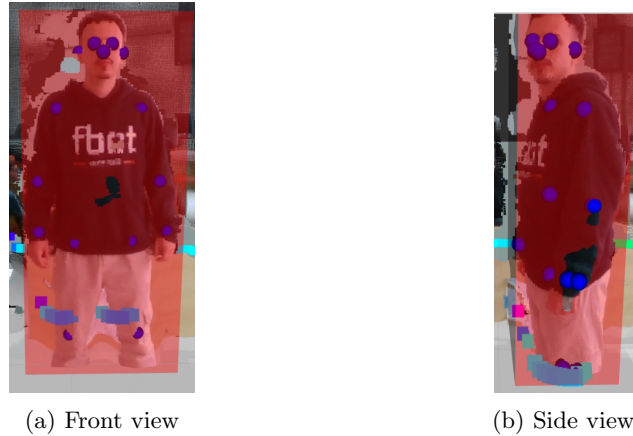


Fig. 2: YoloV8-Pose in 3D space

capable of carrying payloads up to 120kg, all while maintaining a cost below USD 2.000,00. Refer to Fig. 3 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 ⁷ 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 3.3, 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.

2.6 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.

⁷ Github Repository: <https://github.com/EFeru/hoverboard-firmware-hack-FOC>

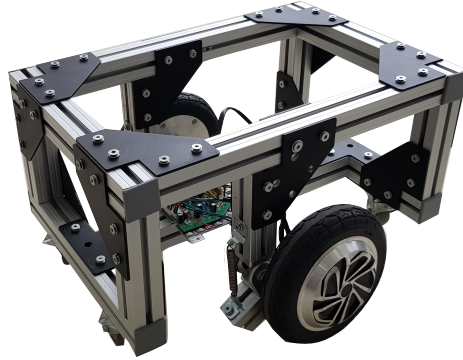


Fig. 3: SHARK mobile base.

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

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

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.

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.

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[5] model API, in order to classify regions detected by a more generic object detector.

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.

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.

3 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.

3.1 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.

3.2 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[6], 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.

3.3 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 platform.

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.

4 Conclusion

This paper outlines the approaches used by the FBOT@Home team for creating an intelligent system for home service robots. Our discussion covered the remodeled electro-mechanical 3D printed animatronic face, designed to enhance human-robot interaction by conveying emotions through expressions. In the realm of object recognition, we revolutionized our approach by employing a smartphone app together with a 3D printed scanner for fast and high-quality data training. 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. Carsten Griwodz, Simone Gasparini, Lilian Calvet, Pierre Gurdjos, Fabien Castan, Benoit Maujean, Gregoire De Lillo, and Yann Lanthony. Alicevision Meshroom: An open-source 3D reconstruction pipeline. In *Proceedings of the 12th ACM Multimedia Systems Conference - MMSys '21*. ACM Press, 2021.
2. Unity Technologies. Unity Perception package. <https://github.com/Unity-Technologies/com.unity.perception>, 2020.
3. Glenn Jocher, Ayush Chaurasia, and Jing Qiu. YOLO by Ultralytics, January 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.
5. OpenAI. Gpt-4 technical report, 2023.
6. 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.

Title: FBOT@Home Team Description Paper RoboCup@Home Eindhoven 2024

Members: João F. S. S. Lemos, Jardel D. S. Dyonisio, Emilly G. Lamotte, 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, Aline S. Marchiori, Luis F. F. Silva, Jardel J. P. Pardo, Eduardo A. D. Evangelista, Igor P. Maurell, Rodrigo S. Guerra, Paulo L. J. Drews Jr,

Robot Hardware Description

The robot is built based on performing household tasks. The specifications are described below:

- Base: PatrolBot (MobileRobots Inc) - Differential programmable autonomous general purpose service robot
- Torso: V shaped format, with 5 shelves in the inside, to support the hardware
- Arm: Mounted on torso. 5DOF. Maximum load: 1.5kg
- Neck: 2DOF
- Head: 12DOF, that are: jaw, eyelid, eyebrow and eyes
- Robot dimensions: height: 1.48m (max), width: 0.59m (max)
- Robot weight: 58kg

Also our robot incorporates the following devices:

- Body LIDAR: RSICK LMS-100 (Body LiDAR)
- Ground LIDAR: URG-04LX-UG01 (Ground LiDAR)
- Camera: RealSense D435i
- Microphone: Rode VideoMic
- CPU: Intel NUC NUC8i7HNK
- GPU: Nvidia Jetson TX2



Fig. 4: DoRIS robot

Robot's Software Description

For our robot we are using the following software:

- OS CPU unit: Ubuntu 20.04 LTS
- OS GPU unit: Ubuntu 18.04 LTS
- Platform: ROS - Robot Operating System
- Navigation: MoveBase
- Localization: AMCL
- Mapping: Gmapping
- Face recognition: OpenFace + OpenCV + Python's face-recognition library
- People Tracking: DeepSORT
- Object Recognition: YOLOv8
- Speech Recognition: SpeechRecognition library for Python
- Arm Control: MoveIt + Interbotix ROS Toolboxes
- Knowledge Storage: Redis
- Task Executor: SMACH