Autonomous Ground Navigation in Highly Constrained Spaces

Lessons Learned From the Second BARN Challenge at ICRA 2023

By (Competition Organizers) Xuesu Xiao, Zifan Xu, Garrett Warnell, and Peter Stone; (Winning Teams) Ferran Gebelli Guinjoan, Rômulo T. Rodrigues, Herman Bruyninckx, Hanjaya Mandala, Guilherme Christmann, Jose Luis Blanco-Claraco, and Shravan Somashekara Rai

The second Benchmark Autonomous Robot Navigation (BARN) Challenge took place at the 2023 IEEE International Conference on Robotics and Automation (ICRA 2023) in London, U.K., and continued to evaluate the performance of state-of-the-art autonomous ground navigation systems in highly constrained environments. Compared to the first BARN Challenge at ICRA 2022 in Philadelphia, the competition has grown significantly in size, doubling the numbers of participants in both the simulation qualifier and physical finals: 10 teams from all over the world participated in the qualifying simulation competition, six of which were invited to compete with each other in three physical obstacle courses at the conference center in London. Three teams won the challenge by navigating a Clearpath Jackal robot from a predefined start to a goal with the shortest amount of time without colliding with any obstacle. The competition results, compared to those of last year, suggest that the teams are making progress toward more robust and efficient ground navigation systems that work out of the box in many obstacle environments. However, a significant amount of fine-tuning is still needed on site to cater to different difficult navigation scenarios. Furthermore, challenges still remain for many teams when facing extremely cluttered obstacles and increasing navigation speed. In this article, we discuss the challenge, the approaches used by the three winning teams, and lessons learned to direct future research.

Digital Object Identifier 10.1109/MRA.2023.3322920
Date of current version: 12 December 2023

THE SECOND BARN CHALLENGE OVERVIEW

The second BARN Challenge took place as a conference competition at ICRA 2023 in London, U.K. As a continuation of the first BARN Challenge at ICRA 2022 in Philadelphia [1], the second challenge aimed to evaluate the capability of state-of-the-art navigation systems to move robots through static, highly constrained obstacle courses, an *ostensibly* simple problem even for many experienced robotics researchers, but in fact, as the results from the first competition suggested, a problem far from being solved.

Each team needed to develop an entire navigation software stack for a standardized and provided mobile robot, i.e., a Clearpath Jackal with a 2D 270° field-of-view Hokuyo lidar for perception and a differential drive system with 2m/s maximal speed for actuation. The developed navigation software stack needed to autonomously drive the robot from a given starting location through a dense obstacle field and to a given goal without any collisions with obstacles or any human interventions. The team whose system could best accomplish this task within the least amount of time would win the competition. The second BARN Challenge had two phases: a qualifying phase evaluated in simulation, and a final phase evaluated in three physical obstacle courses. The qualifying phase took place before the ICRA 2023 conference using the BARN dataset (with the recent addition of DynaBARN), which is composed of 300 obstacle courses in a Gazebo simulation randomly generated by cellular automata. The top six teams from the simulation phase were then invited to compete in three different physical obstacle courses set up by the organizers at ICRA 2023 in the ExCeL London conference center.

In this article, we report on the simulation qualifier and physical finals of the second BARN Challenge at ICRA 2023, present the approaches used by the top three teams, discuss lessons learned from the challenge compared against the first BARN Challenge at ICRA 2022, and point out future research directions to solve the problem of autonomous ground navigation in highly constrained spaces.

SIMULATION QUALIFIER

The simulation qualifier of the second BARN Challenge started on 1 January 2023. The qualifier used the BARN dataset, which consists of 300 5m × 5m obstacle environments randomly generated by cellular automata, each with a predefined start and goal. The teams' systems were evaluated in another 50 unseen environments not available to the public. In addition to the 300 BARN environments, six baseline approaches were also provided for the participants' reference, ranging from classical sampling-based and optimization-based navigation systems, to end-to-end machine learning methods and hybrid approaches. All baselines were implementations of different local planners used in conjunction with Dijkstra's search as the global planner in the Robotic Operating System (ROS) `move _ base` navigation stack.

RULES

Each participating team was required to submit its developed navigation system as a (collection of) launchable ROS node(s). The challenge utilized a standardized evaluation pipeline to run each team's navigation system and compute a standardized performance metric that considers navigation success rate (collision or not reaching the goal counts as failure), actual traversal time, and environment difficulty (measured by optimal traversal time). Specifically, the score s for navigating each environment i was computed as

$$s_i = 1_i^{\text{success}} \times \frac{\mathrm{OT}_i}{\mathrm{clip}(\mathrm{AT}_i, 4\mathrm{OT}_i, 8\mathrm{OT}_i)}$$

where the indicator function 1_{success} evaluates to one if the robot reaches the navigation goal without any collisions, and evaluates to zero otherwise. AT denotes the actual traversal time, while OT denotes the optimal traversal time, as an indicator of the environment difficulty and measured by the shortest traversal time assuming the robot always travels at its maximal speed (2m/s):

$$\mathrm{OT}_i = \frac{\text{Path Length}_i}{\text{Maximal Speed}}.$$

The Path Length is provided by the BARN dataset based on Dijkstra's search from the given start to goal. The clip function clips AT within 4OT and 8OT to ensure that navigating extremely quickly or slowly in easy or difficult environments, respectively, won't disproportionally scale the score. Note that the hyperparameters 4 and 8 for OT were manually selected before the first BARN Challenge, and the organizers found that the performance of the submitted navigation systems in the second BARN Challenge has reached the upper bound of this specific metric, 0.25. So the organizers plan to change these hyperparameters next year to increase the upper bound. The overall score of each team is the score averaged over all 50 unseen test BARN environments, with 10 trials in each environment. Higher scores indicate better navigation performance. The six baselines scored between 0.1627 and 0.2334.

RESULTS

The simulation qualifier started on 1 January 2023 and lasted through a soft submission deadline (20 April 2023) and a hard submission deadline (20 May 2023). Submitting by the soft deadline guaranteed an invitation to the final physical competition, given good navigation performance in the simulation, and left sufficient time for invited participants to make travel arrangements to London. The hard deadline was to encourage broader participation, but final physical competition eligibility depended on the available capacity and travel arrangements made beforehand. In total, 10 teams from all over the world submitted their navigation systems. The performance of each submission was evaluated by the standard evaluation pipeline. The results are shown in Table 1 with the baselines shown in the fourth column as a reference.

Compared to the simulation competition in the first BARN Challenge at ICRA 2022, in which only one team (Temple TRAIL) outperformed all baselines, four teams [KUL+FM (KU Leuven and Flanders Make), INVENTEC (Inventec Corporation), University of Almeria (University of Almeria and Teladoc Health), and UT AMRL (The University of Texas at Austin)] achieved better scores than the best baseline, Learning from Learned Hallucination (LfLH, 0.2334). The top six teams, KUL+FM, INVENTEC, University of Almeria, UT AMRL, Temple TRAIL (Temple University), and UVA AMR (University of Virginia) were invited to the physical finals at ICRA 2023. Note that Temple TRAIL's simulation score was decreased compared to last year since the team has learned from last year's experience and focused on the sim-to-real gap.

TABLE 1. SIMULATION RESULTS.

RANK	TEAM	SCORE	BASELINE
1	KUL+FM	0.249	
2	INVENTEC	0.2445	
3	University of Almeria	0.2439	
4	UT AMRL	0.2424	LfLH
5	Temple TRAIL	0.229	
6	UVA AMR	0.2237	
7	RIL	0.2203	E-Band, e2e
8	Staxel	0.2019	APPLR-DWA
9	The MECO Barners	0.1829	(Fast) DWA
10	Lumnonicity	N/A	

LfLH: Learning from Learned Hallucination; APPLR-DWA: adaptive planner parameter learning-with reinforcement-dynamic window approach.

PHYSICAL FINALS

The physical finals took place at ICRA 2023 in the ExCel London conference center on 30 and 31 May 2023 (Figure 1). Two physical Jackal robots with the same sensors and actuators were provided by the competition sponsor, Clearpath Robotics.

RULES

Physical obstacle courses were set up using 90 cardboard boxes in the conference center (Figure 1). The organizers used the same guidelines to set up three obstacle courses as in the first BARN Challenge, i.e., all courses aimed at testing a navigation system's local planning. Therefore, the courses had an obvious passage but with minimal clearance (a few centimeters around the robot) when the robot was traversing this passage.

The organizers also used the same competition rules agreed on by all of the physical competition participants. Although it was still impractical to run exactly the same navigation systems submitted by the teams in the simulation qualifier because of poor performance in the real world, the organizers reduced both the setup time and competition time from last year's 30 min to 20 min. After the 20-min setup time, each team had the opportunity to run five timed trials (after notifying the organizers to be timed) within another 20 min. The fastest three out of the five timed trials were counted, and the team that had the most successful trials (reaching the goal without any collision) was the winner. In the case of a tie, the team with the fastest average traversal time was declared the winner.

RESULTS

The navigation performances of the six teams are shown in Table 2. Note that, for travel-related reasons, Temple TRAIL could not attend in person, so the organizers ran Temple TRAIL's system submitted to the simulation qualifier for that team, and, unfortunately but also as expected, they could not finish a single trial without on-site fine-tuning. Compared to last year, many teams struggled less on the obstacle avoidance problem in the first two easier environments and therefore were able to shift their attention to increasing speed and were mostly navigating at a much higher speed (>1 m/s). The detailed results of all five timed trials (only the top three were counted in the final score) are listed in the last three columns of Table 2, where "X" indicates failure.

The winner, KUL+FM, is the very first team in the BARN Challenge history that has finished all nine counted physical trials without any collisions. In fact, it only failed three trials in all 15 timed trials in the three obstacle courses. The second place winner, INVENTEC, was able to quickly navigate all six counted trials in the first two environments, sometimes even faster than KUL+FM, but it didn't manage to finish the last most constrained obstacle course. University of Almeria also failed all three trials in the last course and one in the first one.

THE TOP THREE TEAMS AND APPROACHES

In this section, we report the approaches used by the three winning teams.

KUL+FM

The core algorithm of the KUL+FM team is an adaptive free-space motion tube. Consider that a robot's maneuver is defined by a curvature that the robot follows for a time horizon (T) at a constant forward velocity. The motion tube corresponds to the swept volume, that is, the area to be occupied by the vehicle when performing a maneuver. To deal with the discrete resolution of range sensors as well as measurement uncertainty, the footprint of the platform is inflated. Figure 2(a) shows the motion tube for a particular maneuver using the physical (blue polygon) and the inflated (green polygon) footprint of the vehicle. To evaluate whether a motion tube is within the free space, the edge of the inflated swept volume is sampled at an interval d_{sample} [Figure 2(b)]. For computational efficiency, samples are projected in the sensor space by associating them to beam indices. A candidate maneuver is said to be available if, for each sample, the corresponding measurement reported by the sensor is greater than the distance from the sensor to the sample. The reasoning for choosing the inflation and sampling interval values is discussed in the original research article from the team [2]. Finally, within the available motion tube, one can decide the control input of the vehicle based on a high-level application goal.

In essence, instead of inflating obstacles in a local map, the adaptive motion tube inflates the robot and its corresponding trajectory. Because motion tubes are computed with respect to the robot,

FIGURE 1. Final physical competition participants, sponsor Clearpath Robotics, and organizers at the second BARN Challenge in London.

TABLE 2. PHYSICAL RESULTS.

RANK.	TEAM	SUCCESS/TOTAL	AVERAGE TIME	COURSE 1	COURSE 2	COURSE 3
1	KUL+FM	9/9	34/56/91	36/X/33/34/34	63/64/52/47/54	86/79/X/79/X
2	INVENTEC	6/9	44/64/NA	47/42/45/45/41	58/66/X/67/X	X/X/X/X/X
3	University of Almeria	5/9	119/79/NA	X/103/134/X/X	85/86/93/X/53	X/X/X/X/X
4	UVA	4/9	68/NA/103	X/X/76/57/71	X/X/X/X/X	103/X/X/X/X
5	UT AMRL	3/9	84/NA/NA	91/X/79/X/81	X/X/X/X/X	X/X/X/X/X
6	Temple TRAIL	0/9	NA/NA/NA	X/X/X/X/X	X/X/X/X/X	X/X/X/X/X

there is no dependency between localization accuracy and free-space navigation. Another important characteristic of the method is the computational efficiency, thanks to the projection of Cartesian samples in the sensor space, which allows the computing of thousands of motion tubes in low-end computers. A potential disadvantage is that the method relies on the current sensor reading; therefore, it is not robust to outliers and measurement errors. Fortunately, both measurement outliers and errors are rather rare and negligible in the proximity of the sensor.

SOFTWARE STACK

KUL+FM's software stack for the BARN Challenge used an adaptive free-space motion tube for local navigation, the ROS global planner for global planning, and Hector SLAM for online mapping and tracking. Based on the most updated map, the global planner provides subgoals for the local navigation. These subgoals are used to assign costs to available motion tubes: the closer a maneuver takes the vehicle to a subgoal, the smaller its cost is. The control input sent to the vehicle corresponds to the weighted average of the available tubes.

PARAMETER TUNING

The parameters of the local navigation were kept constant during the simulation and the physical competition (2,000 motion tubes with different curvatures and forward velocities, $T = 1$ s, and $d_{\text{sample}} = 0.02$ m). Most of the parameter tuning took place in the global navigation: to avoid obstacles becoming too large and causing the global planner to fail in narrow passages, the team incrementally decreased the obstacle inflation and the distance of the obstacles to be considered in the planner.

INVENTEC

The INVENTEC team's approach was to extend the best-performing baseline, LfLH, with improved collision check and recovery behaviors via a finite-state machine (FSM). The main driving mode relies on a learning-based model that learns to drive a robot by collecting random trajectories and hallucinating obstacles. However, to address poor generalization of the learned model in out-of-distribution deployment scenarios, the team introduced two alternative modules for front safety checks: footprint inflation (FI) used in the simulation qualifier and model predictive control (MPC) used in the physical finals. Additionally, during backward movements, the team performed safety checks by extracting a region of interest (ROI) from the obstacle costmap in the robot memory. Details can be found in the comprehensive technical report by the INVENTEC team [3].

NAVIGATION FSM

The FSM consists of five states: *Initial*, *Heading*, *LfLH*, *Forward*, and *Backtrack*. In the *Initial* state, the navigation controller waits for the path computed by a Dijkstra's search in the `move_base` global planner with a *NavFn* plug-in. The state is switched to *Heading*, which aligns the robot to the target path within a tolerance of ±30°. The *LfLH* model produces velocity commands taking as input the current lidar scans and a local goal drawn from the global path 0.5 m ahead of the robot. At every step a safety check is performed. If a future collision is detected, the state changes to *Backtrack* recovery behavior and an alternative slow *Forward* state if the robot is stuck during backtrack.

RECOVERY BEHAVIORS

During the forward movement in *LfLH*, the robot's path is recorded [the green line shown in Figure 3(a)]. When *Backtrack* is first triggered, it samples a point 0.3 m behind the robot along the recorded path, aligns the heading to the target point, and performs a straight backward command. Moving backward means moving toward the lidar blind spot. Therefore, the team defined a rectangular ROI in the costmap directly behind the robot, as illustrated in Figure 3(b). At every step, the method checks for possible collisions in the costmap ROI, which contains information about past obstacles. If a potential collision is detected during the reverse movement, the state is switched to a slow-forward recovery behavior.

SIMULATION APPROACH

In the simulation stage, the INVENTEC team's strategy for improving the baseline LfLH model was threefold. It consisted in first utilizing FI for obstacle checking with the latest lidar data [shown in Figure 4(a)]; second, checking the costmap for history obstacles to assess the safety of the robot's rear side; and finally, clipping the maximum velocity to 0.7m/s, which provided a

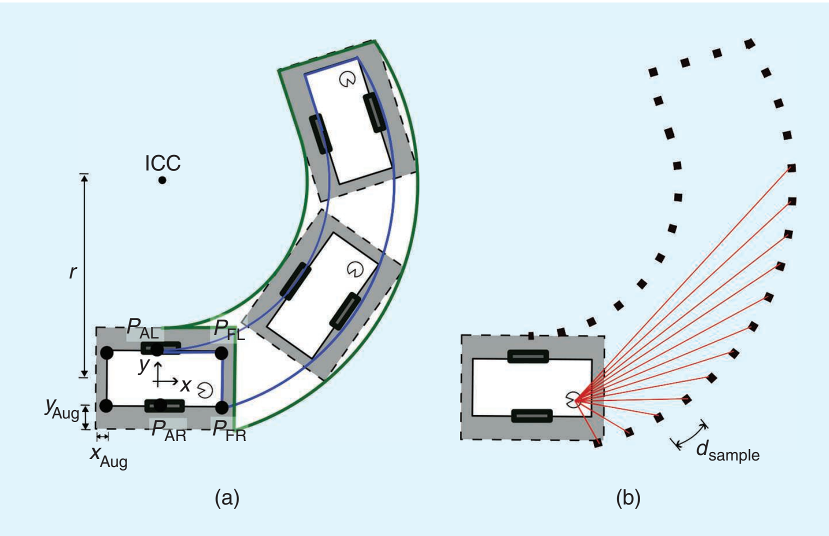

FIGURE 2. (KUL+FM) (a) A motion tube corresponds to the inflated trajectory of the vehicle. (b) For computational efficiency, samples are projected in the sensor space. ICC: instantaneous center of curvature.

significant performance improvement. The FI was 0.04 m. As a result, whenever the inflated footprint overlaps any detected obstacles, the robot transitions into recovery behaviors. Two illustrations of the inflated footprint are depicted in Figure 4(a), in which the green region and red region indicate safe and unsafe conditions, respectively.

REAL-WORLD APPROACH

The only difference between simulation and real-world methods was in the safety check when moving forward. An MPC approach was adopted for the safety check during forward movement in the real world. MPC not only takes into account the robot's current footprint but also integrates future position information, enabling the robot to proactively plan its movement and adjust its trajectory accordingly. Figure 4(b) depicts this process, with green indicating safe portions of the trajectory and red showing a detected future collision.

UNIVERSITY OF ALMERIA

The University of Almeria team's implementation was built upon the Mobile Robot Programming Toolkit (MRPT), an open source C++ framework specifically developed for robotics applications, including libraries for navigation. The MRPT ROS nodes were fine-tuned to align with the specifications of the Clearpath Robotics Jackal robot. A notable difference between the simulation qualifier and physical finals is that in simulation the goal was well known, with fixed coordinates in a map known a priori. By contrast, in the physical finals there were no absolute coordinates of the goal. In practice, this makes the real-world navigation close to pure *exploration*. Those caveats aside, the architecture of the system comprises the subsystems enumerated in the following paragraphs, each one implemented as an independent ROS node.

LOCALIZATION

A custom implementation of particle filter-based localization with an adaptive number of particles using the Kullback–Leibler distance approach was used by the University of Almeria team and is capable of using several metric maps at once for localization.

LOCAL OBSTACLES MAP

The purpose of this node is to perform real-time acquisition and processing of sensor data, in this case the lidar scans. These raw sensor data were subsequently transformed into a 2D point cloud representation, localized within the robot's own coordinate frame and decimated into lower resolution for faster processing.

PATH PLANNER (HIGHER LAYER)

At a relatively low rate (1 Hz or slower), the local obstacles were considered to find a kinematically feasible path using a path planning algorithm, which was then sent to a local path follower. The team's path planner used a custom algorithm based on A* on a discrete lattice of the state-vector space of the vehicle, i.e., the SE(2) pose plus velocities. Arcs between the lattice nodes were explored efficiently using parameterized trajectory generators (PTGs), a concept derived from past works, which defines families of paths to help explore the environment with kinematically and dynamically feasible paths.

LOCAL PLANNER (LOWER LAYER)

Once a path is found, the task of generating motor commands to follow it, including avoiding any new obstacles, is accomplished by a reactive navigation system, which is also based on the trajectory parameter space (TP-Space). In TP-Space, the robot, regardless of its physical shape and kinematic constraints, is transformed into a free-flying point within a newly formulated parameter space. This transformation incorporates the robot's shape and kinematic restrictions, thereby allowing for efficient navigation by taking into account the robot's specific physical characteristics and movement capabilities. On the other hand, PTGs define a set of potential trajectories for the robot, parameterized by variables such as path shape,

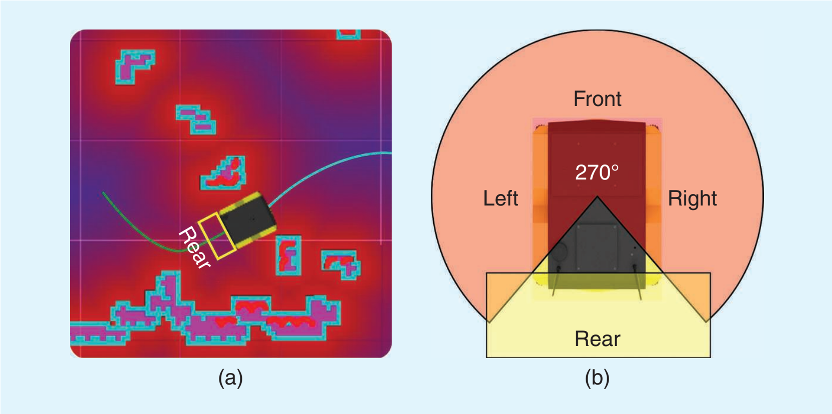

FIGURE 3. (INVENTEC) (a) Costmap with robot rear ROI (yellow rectangle). (b) Jackal lidar field of view.

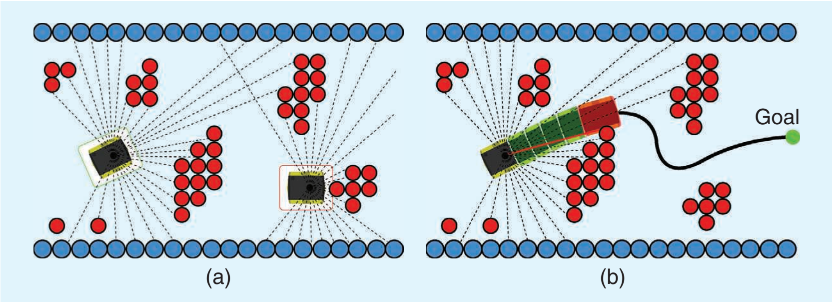

FIGURE 4. (INVENTEC) (a) FI for collision detection. (b) MPC footprint forward safety check.

robot speed, and turning radius. By dynamically adjusting these parameters based on real-time sensor data, the robot can select the optimal path to avoid obstacles and reach its destination. The reactive navigation system, through its integration with TP-Spaces and PTGs, acts as an advanced behavior planning algorithm. The implementation used the ROS node developed based on these principles. It processed the 2D point cloud data to dynamically generate an optimal path for the robot, while simultaneously accounting for the robot's kinematic constraints and potential obstacles within its environment. This local planner can run at a higher rate than the path planner, typically between 5 and 10 Hz.

DISCUSSION

While many discussion points from the first BARN Challenge are still valid this year, we discuss new findings and lessons from the second BARN Challenge and point out promising future research directions to push the boundaries of mobile robot navigation in highly constrained spaces.

MORE PARTICIPANTS FROM INDUSTRY

One interesting change in the second year of the BARN Challenge was the participation from the robotics industry. Compared to last year's competition, in which all participants were from universities worldwide, five out of the 10 teams who participated in the second BARN Challenge were from industry or hybrid academia–industry teams (KU Leuven with Flanders Make, INVENTEC, the University of Almeria with Teladoc Health, Staxel, and Lumnonicity from JIO AICOE). Notably, all three winning teams had industry ties.

The significant increase in industry participation indicates that the problem the BARN Challenge aims to solve is of significant interest to real-world robotics manufacturers, providers, and users. As discussed in last year's report, even very experienced robotics researchers in academia may have the impression that such an ostensibly simple problem has already been solved. However, the fact that the robotics industry is still working on such a problem and using the BARN Challenge as a testbed for its methods suggests that it still does not have a satisfactory solution for this problem in the real world. Therefore, the organizers suggest that academic researchers consider real-world problems the robotics industry faces to realize technology transfer from academia to industry. The first and third place winners, KUL+FM and University of Almeria (with Teladoc) are very good examples of applying state-of-the-art academic research to the real-world robotics industry and identifying real-world problems from industry to solve together with academia.

> "THE ORGANIZERS SUGGEST THAT ACADEMIC RESEARCHERS CONSIDER REAL-WORLD PROBLEMS THE ROBOTICS INDUSTRY FACES TO REALIZE TECHNOLOGY TRANSFER FROM ACADEMIA TO INDUSTRY."

SIMULATION PERFORMANCE APPROACHING THE CURRENT UPPER BOUND

Another observation from the simulation qualifier is that the simulation performance has approached the current metric's upper bound, i.e., 0.25. The upper and lower bounds are determined by the two hyperparameters 4 and 8 in the evaluation metric to ensure that navigating extremely quickly or slowly in easy or difficult environments, respectively, won't disproportionally scale the score. Approaching 0.25 means that most of the navigation trials in the BARN environments can be successfully finished within four times of the optimal traversal time, even in the most difficult ones with very dense obstacles. Therefore, the organizers will change these hyperparameters in next year's challenge to increase the upper and lower bounds to encourage the teams to score higher by achieving less traversal time than four times of the optimal time, without colliding with any obstacle.

THE FIRST TEAM THAT FINISHED ALL PHYSICAL COURSES

KUL+FM is the first team in The BARN Challenge history that was able to successfully navigate the Jackal through all three physical obstacle courses in all nine final counted trials. However, it did fail three out of the total 15 timed trials in obstacle courses 1 and 3, as shown in Table 2. While the failure trial in obstacle course 1 was likely due to the need of initial parameter fine-tuning to fit the physical environments, the two failure trials in the last obstacle course were both due to the desire to push on faster navigation speed. Considering that the physical obstacle courses constructed this year were qualitatively more difficult than last year's, all three winning teams' stable performances in the first two obstacle courses and KUL+FM's nine successful trials in all three courses suggest that the physical performance of the navigation stacks has been improved compared to last year.

FINAL RANKING STILL DECIDED BY SUCCESS RATE, NOT SPEED

Despite the improved performance and more confident deployment experience (e.g., less struggling to navigate to the goal and to avoid obstacles, but more focus on fine-tuning for robustness and speed) observed by the organizers, completely collision-free navigation is still out of reach for most teams regardless of speed. Therefore, the final ranking in the physical finals is still decided by success rate, not navigation speed. Although the second place winner, INVENTEC, had a chance to win the competition before it started the third obstacle course by pushing on reducing the traversal time, it eventually failed all five attempts in the last course. So the community is still waiting for the first BARN Challenge in which the final ranking is determined by traversal speed in highly constrained obstacle environments, after the success rate has been guaranteed to be 100%.

LESS OBVIOUS SIM-TO-REAL GAP

Compared to last year, when the winning approach in simulation suffered from significant collisions in the real world, the rankings in Tables 1 and 2 suggest a decreasing sim-to-real gap between the simulation qualifier and physical finals this year. Based on the approaches taken by the teams, it is possible that such a smaller sim-to-real gap is caused by the decreased usage on learning-based navigation methods. Despite the popularity of using machine learning to address visual inputs, off-road conditions, social contexts, and multirobot navigation, only one team, INVENTEC (except Temple TRAIL, who did not compete in the physical finals), used a learning-based method to navigate in highly constrained obstacle environments in the second BARN Challenge. INVENTEC's approach is mostly based on the Learning from Hallucination paradigm, especially the latest LfLH approach (one of the baselines provided by the competition organizers). Furthermore, INVENTEC designed sophisticated recovery behaviors to address real-world scenarios where the learning approach did not work well (see the following detailed discussion), which also helped to reduce the sim-to-real gap.

IMPORTANCE OF GOOD RECOVERY BEHAVIORS TO DEPLOY END-TO-END LEARNING-BASED SYSTEMS IN THE REAL WORLD

During the first BARN Challenge last year, the end-to-end learning approach trained by deep reinforcement learning by Temple TRAIL experienced a significant sim-to-real gap because the training was conducted in simulation on a different, smaller robot platform (Turtlebot2), and there was no recovery behavior. This year, Temple TRAIL was not able to compete in person, and the system deployed by the organizers on behalf of the team did not perform well without fine-tuning. As the other team that adopted machine learning for its navigation system, INVENTEC used the LfLH policy trained using simulated data but also devised a set of recovery behaviors to address real-world scenarios where the learning approach did not work well. By developing good recovery behaviors to complement an end-to-end learned motion policy, INVENTEC's approach outperformed the original LfLH approach, which is assisted only by a set of very simple recovery behaviors, and achieved very good sim-to-real transfer. This observation suggests the potential of end-to-end learning approaches for navigation when augmented by a sophisticated mechanism to complement learning during out-of-distribution real-world scenarios.

> "THIS OBSERVATION SUGGESTS THE POTENTIAL OF END-TO-END LEARNING APPROACHES FOR NAVIGATION WHEN AUGMENTED BY A SOPHISTICATED MECHANISM TO COMPLEMENT LEARNING DURING OUT-OF-DISTRIBUTION REAL-WORLD SCENARIOS."

TUNING STILL NECESSARY FOR ALL CLASSICAL SYSTEMS FACING DIFFERENT OBSTACLE ENVIRONMENTS

Similar to last year's observation, the original intention of "out-of-the-box" deployment of the navigation systems submitted to the simulation qualifier directly in the physical finals was still impossible for all of the teams: all teams that used classical systems had to extensively fine-tune their navigation stacks, while INVENTEC, the only team that used a learning-based approach, fine-tuned its recovery behaviors, the classical part of its system. Although slight fine-tuning to adapt a system from simulation to the real world is reasonable, all teams needed to fine-tune their navigation systems before competing in all three obstacle courses. It is unclear whether a single navigation system configuration that works for all obstacle courses exists or not. The organizers suggest that the community keep such an intention in mind when developing navigation systems because the option of fine-tuning for every deployment scenario is not possible, especially when the goal is to deploy autonomous mobile robots at scale in the wild. One single navigation system configuration or parameterization to address all possible scenarios is certainly ideal, but another promising approach is autonomous parameter tuning to adapt to different obstacle configurations around the robot.

INTEREST FROM OUTSIDE THE ROBOTICS COMMUNITY

The second BARN Challenge has even raised interest from outside the robotics community. For example, Harald Carlens, a machine learning researcher, Malte Schwerin, an artificial intelligence student researcher from RWTH Aachen University, and Dr. Gerald Steinbauer-Wagner, a professor in robotics and psychology from Graz University of Technology in Austria, have reached out to use the BARN Challenge for ML Contests (https://mlcontests.com) to study ranking methods for algorithm competitions and to assess robotic capabilities using methods from psychology, respectively.

REFERENCES

[1] X. Xiao et al., "Autonomous ground navigation in highly constrained spaces: Lessons learned from the benchmark autonomous robot navigation challenge at ICRA 2022 [Competitions]," *IEEE Robot. Autom. Mag.*, vol. 29, no. 4, pp. 148–156, Dec. 2022, doi: 10.1109/MRA.2022.3213466.

[2] R. T. Rodrigues, N. Tsiogkas, N. Huebel, and H. Bruyninckx, "Clutter-resilient autonomous mobile robot navigation with computationally efficient free-space features," in *Proc. Int. Symp.* Robot. Res., Cham, Switzerland: Springer Nature Switzerland, 2022, pp. 522–537.

[3] H. Mandala and G. Christmann, "The BARN challenge 2023–Autonomous navigation in highly constrained spaces–Inventec team," 2023, *arXiv: 2307.14580.*