The AI Driving Olympics: An Accessible Robot Learning Benchmark

Andrea Censi[*] **Liam Paull**[†] **Jacopo Tani**[*] **Matthew R. Walter**[‡]

Abstract

Despite recent breakthroughs, the ability of deep learning and reinforcement learning to outperform traditional approaches to control physically embodied robotic agents remains largely unproven. To help bridge this gap, we have developed the "AI Driving Olympics" (AI-DO), a competition with the objective of evaluating the state-of-the-art in machine learning and artificial intelligence for mobile robotics. Based on the simple and well specified autonomous driving and navigation environment called "Duckietown," AI-DO includes a series of tasks of increasing complexity—from simple lane-following to fleet management. For each task, we provide tools for competitors to use in the form of simulators, data logs, code templates, baseline implementations, and low-cost access to robotic hardware. We evaluate submissions in simulation online, on standardized hardware environments, and finally at the competition events. We have held successful AI-DO competitions at NeurIPS 2018 and ICRA 2019, and will be holding AI-DO 3 at NeurIPS 2020. Together, these competitions highlight the need for better benchmarks, which are lacking in robotics, as well as improved mechanisms to bridge the gap between simulation and reality.

1 Background

Machine learning has made it possible to tackle problems in perception and decision making, where classic robotics approaches have traditionally failed [Kemp et al., 2007], including image-based grasping of diverse objects [Levine et al., 2018] and navigation based on semantic clues [Zhu et al., 2017]. Robotic domains such as self-driving vehicles could greatly benefit from advancements in learning. However, it is not clear how well ML-based solutions would transfer from the simplified environments in which they are often developed to the real-world, where decisions must be made in real-time based on noisy sensor measurements, the world is uncertain and difficult to model, computation is limited, and communication is subject to bandwidth constraints and latencies.

Solutions to embodied robotic tasks are notoriously difficult to compare, making it difficult to evaluate the state-of-the-art [Anderson et al., 2011]. By making solutions comparable and results reproducible, competitions provide an effective means to understand current progress and open problems. With this in mind, we conceived the AI Driving Olympics (AI-DO) as a benchmark for evaluating learning-based solutions on accessible physically embodied systems. Specifically, AI-DO provides standardized simulation and hardware environments for assessing the state-of-the-art in autonomous driving, including tasks related to multisensor perception, autonomous mobility on demand (AMoD), and embodied AI. The first AI-DO live competition took place at NeurIPS 2018, and the second was held at ICRA 2019. Between the two, AI-DO has attracted over 4000 submissions from more than 200 participants.

[*]ETH Zürich, Zürich, Switzerland
[†]Mila, Université de Montréal, Montréal, Canada
[‡]Toyota Technological Institute at Chicago, Chicago, IL, USA

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

2 The AI Driving Olympics

The AI-DO competition was designed with a particular set of characteristics in mind. The competition must be accessible to people with diverse backgrounds, with no up-front costs and hardware that is affordable, but optional. The competition should involve automatic, reproducible experiments with physical hardware that is subject to common power, computation, memory, and network constraints. Additionally, the framework should support data-intensive learning algorithms through access to a high-fidelity simulation environment as well as large amounts of real-world data.

AI-DO involves a set of diverse tasks for the domain of autonomous driving, including both interactive embodied and simulated tasks, and non-interactive inference tasks. Thus, it presents a balanced view of a very complex domain, without trivializing it to one particular task or environment. AI-DO involves four leagues: urban driving, racing, autonomous mobility on demand, and advanced sensing.

Urban Driving The AI-DO *urban driving* league consists of competitions that require Duckiebot robots to drive on a closed road within a formalized Duckietown environment. The overall goal is to travel as far as possible in the right lane, while avoiding static and dynamic objects that may be present on the course. We have developed several resources to support participants, including an affordable robot ($250 USD) and a rigorously defined environment [Paull et al., 2017], implementations of fully functional baseline algorithms, large datasets from diverse environments, a cloud-based simulation and training environment, remotely accessible Duckietowns ("Robotariums") with Duckiebots.

Racing The *racing* league tasks participants with autonomously navigating at least one lap of a course as fast as possible. Unlike urban driving, racing focuses more on speed and agility, which presents a different set of challenges for sim-to-real transfer. The ability to correctly model the vehicle's dynamics becomes increasingly important at higher speeds. Joint perception and dynamics are required for predictive path planning in order to achieve racing-level performance. These are exciting research problems. Participants can deploy their trained models in a cloud-based simulator as well as onto the AWS DeepRacer, a $1/18$-scale Ackerman-drive vehicle.

Autonomous Mobility on Demand The *AMoD* league considers the question of how manage a fleet of autonomous vehicles in order to best serve the needs of customers. Participants implement a centralized dispatcher that commands individual agents within a fleet in simulation in order to meet customer demands. The environment involves thousands of vehicles and customer requests that must be handled simultaneously in a large, complex network with varying congestion effects.

Advanced Sensing The *advanced sensing* league considers challenging, multisensor tasks including object detection, object tracking, and trajectory estimation. The league uses nuScenes, a large-scale publicly available dataset for autonomous driving. Unlike other datasets and challenges, nuScenes is multimodal, with data from six cameras, five radars, and one lidar from 1000 driving scenes that include 1.4M images, 400k lidar scans, and 1.3M radar scans.

References

J. Anderson, J. Baltes, and C. T. Cheng. Robotics competitions as benchmarks for AI research. *The Knowledge Engineering Review*, 26(1):11–17, 2011.

C. C. Kemp, A. Edsinger, and E. Torres-Jara. Challenges for robot manipulation in human environments [grand challenges of robotics]. *IEEE Robotics & Automation Magazine*, 14(1):20–29, 2007.

S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research*, 37(4-5): 421–436, 2018.

L. Paull, J. Tani, H. Ahn, J. Alonso-Mora, L. Carlone, M. Cap, Y. F. Chen, C. Choi, J. Dusek, Y. Fang, and others. Duckietown: an open, inexpensive and flexible platform for autonomy education and research. In *Proc. Int'l Conf. on Robotics and Automation (ICRA)*, 2017.

Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi. Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *Proc. Int'l Conf. on Robotics and Automation (ICRA)*, 2017.