Continual Learning on Incremental Simulations for Real-World Robotic Manipulation Tasks

Josip Josifovski[1] *josip.josifovski@tum.de*
Mohammadhossein Malmir[1] *hossein.malmir@tum.de*
Noah Klarmann[1] *noah.klarmann@tum.de*
Alois Knoll[1] *knoll@mytum.de*

***Abstract*—Current state-of-the-art approaches for transferring deep-learning models trained in simulation either rely on highly realistic simulations or employ randomization techniques to bridge the reality gap. However, such strategies do not scale well for complex robotic tasks; highly-realistic simulations are computationally expensive and hard to implement, while randomization techniques become sample-inefficient as the complexity of the task increases. In this paper, we propose a procedure for training on incremental simulations in a continual learning setup. We analyze whether such setup can help to reduce the training time for complex tasks and improve the sim2real transfer. For the experimental analysis, we develop a simulation platform that can serve as a training environment and as a benchmark for continual and reinforcement learning sim2real approaches.**

***Index Terms*—simulation, sim2real, reality gap, continual learning, robotics**

I. Introduction

Current deep-learning-based models show impressive results in simulations, as they win in video games [1], outperform the world champion in the game of Go [2], or develop creative strategies for multi-agent hide-and-seek environments [3].

However, the success in applying such models to control real robots is limited. For example, real-world training of a data-driven model, as in [4], is impractical or impossible in many cases where environments are dynamic or resources are limited. Training in simulation, on the other hand, faces with the reality gap problem when the model is transferred to the real world. To bridge the reality gap, current state-of-the-art methods use either highly realistic simulations or randomization strategies. As the tasks get more complex, such strategies do not scale well. For instance, highly-realistic simulations of a complex task become computationally expensive and harder to develop, while randomization approaches need exponentially more training samples with the linear increase of randomization parameters due to increasing task complexity. On the other hand, different continual learning approaches [5] allow incremental learning and reusing of skills from learned to novel tasks and adaptation to dynamic environments. In this paper, we analyze whether incremental simulations can be combined in a continual learning manner to address the problem of scalability and facilitate the real-world transfer of the simulation-trained model.

[1] *Chair of Robotics, Artificial Intelligence and Real-time Systems, Department of Informatics, Technical University of Munich*, Munich, Germany

This work has been financially supported by the ECSEL Joint Undertaking under the H2020 AI4DI project (grant agreement 826060).

Supplementary video is available at: https://youtu.be/UoJolTE6SMg

II. Background and Related Work

Training in simulation for robotic control and the reality gap problem have been addressed in many earlier works [6]–[9]. In this section we focus only on recent approaches that employ deep learning architectures.

The approaches that are introduced in [10], [11] aim at precise reproduction of the relevant real-world phenomena in simulation. For instance, in [10] the authors trained a 7-DOF robotic arm in simulation for locating and grasping a cube by adjusting the simulation to resemble the real world as closely as possible. However, matching and correctly simulating all relevant real-world phenomena is often complicated to achieve. Instead, other approaches [12]–[15], rely on randomization of the important simulation parameters during model training to increase model robustness and minimize the effect of the mismatch between simulation and real world. For example, in [12] the authors trained a robotic arm to perform a pushing task by randomizing the physical properties of the robot and the dynamics of the simulation. Similarly, in [13] the authors employed domain randomization to create millions of unrealistically-appearing training images for robust object localization in a robotic grasping task. In the same direction, James et al. [15] combined the domain randomization technique with a generative model to facilitate the learning of a more complex task, like grasping of an unknown object with a robotic arm.

Most of the above-mentioned approaches employ domain adaptation, where the model trained in simulation is fine-tuned on the real robot as a final step in order to improve its performance. This two-stage training can be framed as a continual learning problem by considering the same robotic task in simulation and in the real world as two separate tasks to solve. For example, in [16] the authors first trained a model for a reaching task in simulation, and then reused the features from the first model to train a second model in the real world. They reported that due to the reusability of the skills learned in simulation, the additional training steps in the real world are drastically decreased.

III. Methodology and Experiments

Considering domain adaptation as a continual learning problem, we propose incremental simulations of several manipulation tasks to facilitate the learning of a more complex task and smoothly transfer the learned policy in simulation to the real robot. In this setup, a model like [17] can be trained using

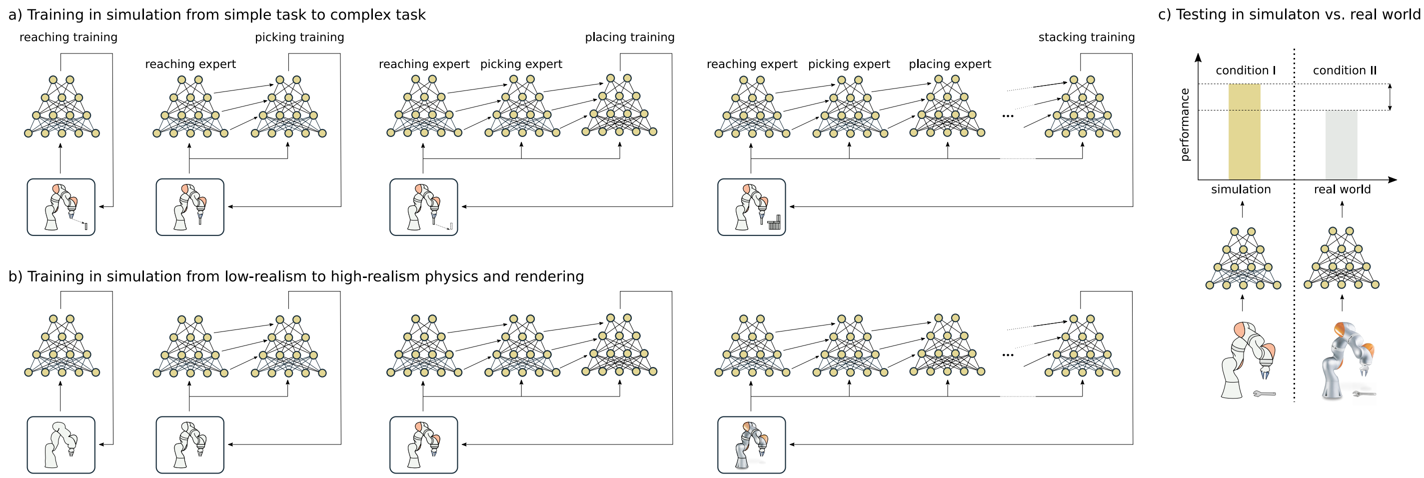

Fig. 1. Experimental setup for training and testing.

curriculum or continual learning. For example, starting from a simple task like reaching, the manipulator is trained on tasks with increasing difficulty, while reusing the features of the previously learned tasks, as shown in Figure 1a). Alternatively, the model can be trained on simulations with increasing level of realism, starting from simplified meshes and unrealistic renderings and enhancing the simulation towards more precise physics and photo-realistic rendering, as shown in Figure 1b). The two incremental approaches can also be combined to provide speed-realism tradeoffs during training.

Different training strategies will be evaluated in order to gather data for exploratory analysis. The results from the analysis will be used to propose model changes or more efficient simulation training strategies for complex manipulation tasks in the real world. To objectively evaluate the transfer to the real robot, an experiment like the one described in Figure 1c is planned. Here, the model is trained to perform a complex task on incremental simulations. Afterwards, it is tested in simulation first, on previously unseen scenarios and the performance of the model is measured (condition I) to serve as a baseline. The same simulation test scenarios are carefully recreated in the real world, while taking special care to minimize confounding variables, like different initial conditions. The performance in the real world is also measured (condition II) and compared to the performance in the simulation. This way, the difference in performance which is due to the reality gap can be quantified to see whether and to what extent the incremental simulations help in bridging the reality gap. The incremental simulations approach will also be compared with existing techniques, like domain randomization, to see whether it can provide improvements for complex manipulation tasks.

To implement the experimental setup, we are developing a simulation platform that can support different robotic manipulation tasks, like reaching, pushing, picking, placing, or stacking with a Kuka™ robotic manipulator [18], which is commonly used in current research. The simulation platform is developed using the Unity™ game engine [19], in order to provide GPU-accelerated physics calculations and highly-realistic rendering. To ensure a high degree of interconnectivity with the environment, we employ same interfaces as OpenAI Gym [20]. In addition, as the simulation platform focuses on real-world transfer, it supports techniques like domain and dynamics randomization and interfaces to the robot for easily switching from simulation to real world. The simulation platform and the real robots used in the experiments are presented in Figure 2.

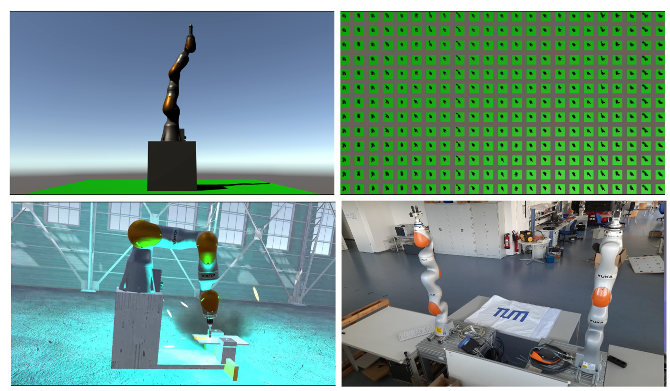

Fig. 2. The simulation platform and the real robots. Upper left: Simple simulation environment for a reaching task. Upper right: Top-down view of matrix of simulation environments running in parallel. Bottom Left: Simulation of highly-realistic rendering and physics. Bottom right: The Kuka™ iiwa manipulators [18] used in the experiments.

IV. Conclusion and Future Work

The experimental results of this paper should contribute to better understanding of whether and to what extent continual learning on incremental simulations can improve the training process and the sim2real transfer for complex robotic manipulation tasks. Furthermore, even though there are many different simulation platforms available, very few tackle real-world transfer and allow objective evaluation and experiment reproducability of sim2real approaches. In order to facilitate research, we plan to provide the simulation platform as a benchmark for comparison of continual and (hierarchical) reinforcement learning models in sim2real transfer.

REFERENCES

[1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, *Playing atari with deep reinforcement learning*,

[2] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016. DOI: 10.1038/nature16961.

[3] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, and I. Mordatch, *Emergent tool use from multi-agent autocurricula*,

[4] 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*, vol. 37, no. 4-5, pp. 421–436, 2018, ISSN: 0278-3649. DOI: 10.1177/0278364917710318.

[5] G. I. Parisi, R. Kemker, J. L. Part, C. Kanan, and S. Wermter, "Continual lifelong learning with neural networks: A review," *Neural networks : the official journal of the International Neural Network Society*, vol. 113, pp. 54–71, 2019. DOI: 10.1016/j.neunet.2019.01.012.

[6] N. Jakobi, Phil Husbands, and Inman Harvey, "Noise and the reality gap: The use of simulation in evolutionary robotics," *European Conference on Artificial Life*, 1995.

[7] N. Jakobi, "Running across the reality gap: Octopod locomotion evolved in a minimal simulation," *European Workshop on Evolutionary Robotics*, 1998.

[8] J. C. Zagal, J. Ruiz-del-Solar, and P. Vallejos, "Back to reality: Crossing the reality gap in evolutionary robotics," *IFAC Proceedings Volumes*, vol. 37, no. 8, pp. 834–839, 2004, ISSN: 14746670. DOI: 10.1016/S1474-6670(17)32084-0.

[9] S. Koos, J.-B. Mouret, and S. Doncieux, "The transferability approach: Crossing the reality gap in evolutionary robotics," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 1, pp. 122–145, 2013, ISSN: 1089-778X. DOI: 10.1109/TEVC.2012.2185849.

[10] S. James and E. Johns, *3d simulation for robot arm control with deep q-learning*, 2016.

[11] J. Mahler, M. Matl, X. Liu, A. Li, D. Gealy, and K. Goldberg, "Dex-net 3.0: Computing robust vacuum suction grasp targets in point clouds using a new analytic model and deep learning," pp. 5620–5627, 2018. DOI: 10.1109/ICRA.2018.8460887.

[12] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, "Sim-to-real transfer of robotic control with dynamics randomization," pp. 3803–3810, 2018. DOI: 10.1109/ICRA.2018.8460528.

[13] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, *Domain randomization for transferring deep neural networks from simulation to the real world*,

[14] OpenAI, I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, J. Schneider, N. Tezak, J. Tworek, P. Welinder, L. Weng, Q. Yuan, W. Zaremba, and L. Zhang, *Solving rubik's cube with a robot hand*, 2019.

[15] S. James and Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, Konstantinos Bousmalis, "Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 12 627–12 637, 2019.

[16] A. A. Rusu, M. Vecerik, T. Rothörl, N. Heess, R. Pascanu, and R. Hadsell, *Sim-to-real robot learning from pixels with progressive nets*,

[17] A. A. Rusu, N. C. Rabinowitz, G. Desjardins, H. Soyer, J. Kirkpatrick, K. Kavukcuoglu, R. Pascanu, and R. Hadsell, *Progressive neural networks*, 2016.

[18] *Kuka lbr-iiwa*, https://www.kuka.com/products/robot-systems/industrial-robots/lbr-iiwa, Accessed: 2020-6-26.

[19] *Unity 3d*, https://unity.com/, Accessed: 2020-6-26.

[20] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *arXiv preprint arXiv:1606.01540*, 2016.