ES-sim2real optimization

Cesare Maria Dalbagno, spring internship 2024

AI for Robotics Lab (AIR-Lab), Department of Cognitive Science and Artificial Intelligence
Tilburg University, Tilburg, The Netherlands

Introduction

Reinforcement learning (RL) trains agents in simulations which are different from the real world: "**sim-to-real gap**"

Domain randomization enhances adaptability by changing the parameters of the simulations

By introducing an adaptation phase the agent could meta-learn a general policy for real world deployement

Research question

To what extent the performance of an agent trained in simulation using meta-deep reinforcement learning (mDRL) with evolutionary strategies (ES) and domain randomization (DR) will be able to generalize to the real world?

Proof of concept: test generalizability of different training paradigms in MuJoCo simulator

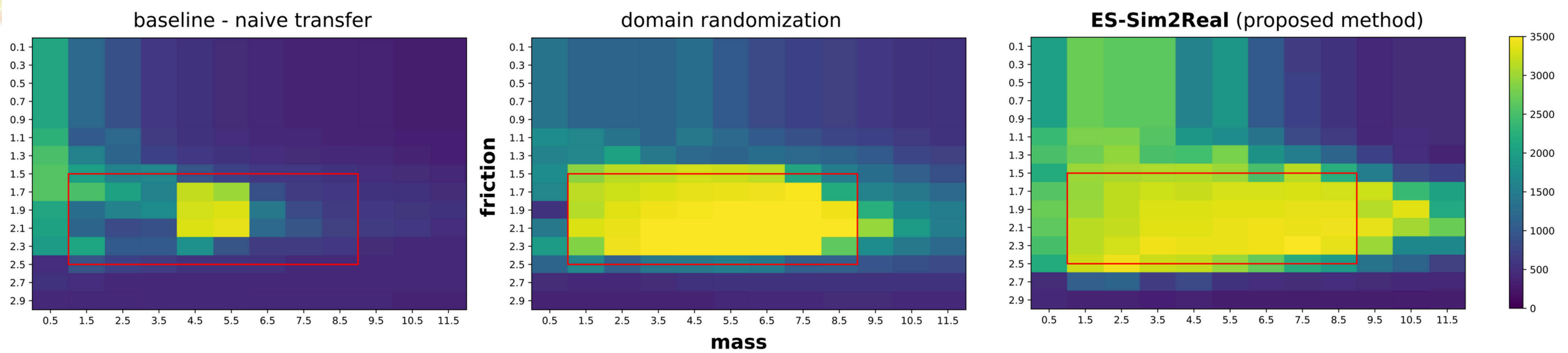
Methods

Trained 6 agents in simulation (4 MuJoCo envs, 2 Gymnasium robotics envs) with different setups:

- set parameters (with PPO)
- domain randomization with LSTM
- ES domain randomization
- ES domain randomization + meta learning adaptation

Testing their performances on a range of tasks wider than the one they were trained on (out of randomization boundaries)

Results


baseline - naive transfer
domain randomization
ES-Sim2Real (proposed method)
friction
mass
0.1 0.3 0.5 0.7 0.9 1.1 1.3 1.5 1.7 1.9 2.1 2.3 2.5 2.7 2.9
0.5 1.5 2.5 3.5 4.5 5.5 6.5 7.5 8.5 9.5 10.5 11.5
0 500 1000 1500 2000 2500 3000 3500

Conclusion and discussion

Still have to try on real robot, results seem so far convincing, even though ES alone could already constitute a competitive method.