Hessisches Kompetenzzentrum
für Hochleistungsrechnen

End-to-end Learning for Multi-Embodiement Locomotion

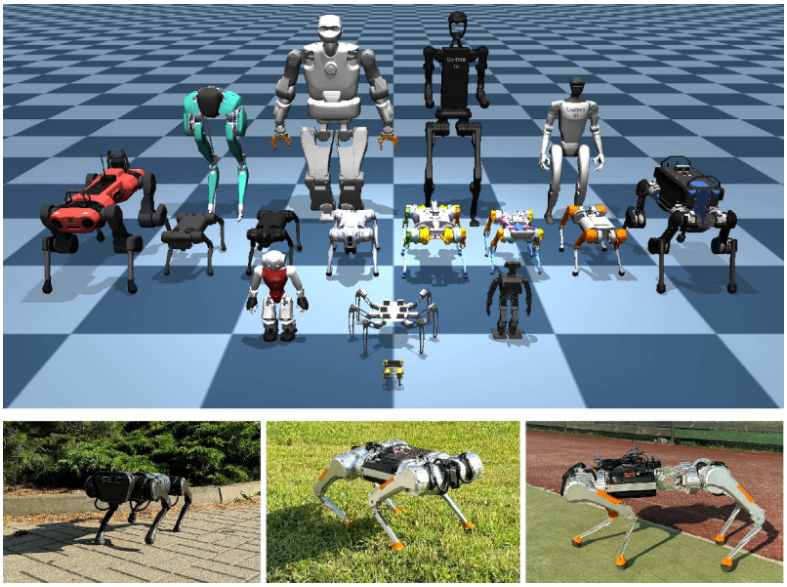


Figure 1: Top – We train a single locomotion policy for multiple robot embodiments in simulation. Bottom – We can transfer and deploy the policy on three real-world platforms by randomizing the embodiments and environment dynamics during training.

Project Manager
Dr. Davide Tateo

Researchers
Nico Bohlinger, Dr. Piotr Kicki, Grzegorz Czechmanowski, Maciej Krupka and Dr. Krzysztof Walas

Principal Investigator
Prof. Jan Peters (PhD)

Project Term
2023 - 2024

Clusters
Lichtenberg II Cluster Darmstadt

Software
Python, PyTorch

Additional Software
JAX, MuJoCo

Institute
Intelligent Autonomous Systems

University
Poznan University of Technology, Poland, Technische Universität Darmstadt

Introduction

The rapid development of robotics has enabled legged robots capable of performing highly dynamic tasks such as walking on uneven terrain, jumping, or even completing challenging parkour courses. However, training robots of different morphologies, such as quadrupeds, bipeds, and hexapods, still is a significant challenge for current training paradigms. To overcome these challenges, scalable Deep Reinforcement Learning algorithms and parallelized simulations significantly improve the learning for such robots when coupled with clever neural network policy architectures.

Methods

In this research, a novel framework called URMA (Unified Robot Morphology Architecture) is used to create a single neural network architecture capable of learning control policies for various robot morphologies. The approach involves training the robots in a simulation environment, where multiple robots are trained simultaneously. Our method uses a combination of DRL and a morphology-aware neural network architecture to allow robots to share knowledge between their embodiements. The framework is designed to be morphology-agnostic, meaning that it can be applied to any robot with any number of joints. Key to our approach is the use of special encoders and decoders that

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

translate the robot’s specific joint and sensor information into a shared representation space. This enables the model to generalize its learning and control a wide range of robots.

Results

In the first phase of the project, we successfully trained a locomotion policy on 16 different robot platforms, including quadrupeds, hexapods, and humanoids. The trained policy was able to control these robots both in simulation and also on selected real-world environments. For example, zero-shot transfer of the trained policy to the MAB Silver Badger quadruped robot in the real world was achieved without any further fine-tuning, demonstrating the framework’s robustness and adaptability. A significant achievement in the later stages was the successful zero-shot and few-shot transfer of the trained policy to robots not seen during trainingt. The framework also demonstrated an ability to handle missing sensor data and adapt to changes in robot morphology. These results highlight the potential of the URMA framework to serve as a foundation model for legged locomotion, enabling rapid deployment across a wide variety of robot platforms.

Discussion

The results of our project demonstrate the effectiveness of the URMA framework in learning robust control policies for different types of robots. The use of HPC was crucial in achieving these results, as it enabled the parallel simulation and training of multiple robot platforms, reducing the overall training time. The ability of the framework to transfer learned policies to new robots without additional training underscores its potential for practical applications in robotics.
However, challenges remain, particularly in generalizing to highly diverse or completely new robot morphologies. While the framework shows strong performance on quadrupeds and humanoids, more work is needed to handle extreme variations in robot design or to improve the fine-tuning process for entirely new platforms. Future research will focus on extending the framework’s capabilities to more complex locomotion tasks, such as highly dynamic locomotion in challenging terrain and integrating advanced sensing capabilities.

Publications

Bohlinger, N.; Czechmanowski, G.; Krupka, M.; Kicki, P.; Walas, K.; Peters, J.; Tateo, D.; One Policy to Run Them All: an End-to-end Learning Approach to Multi-Embodiment Locomotion; Conference on Robot Learning (2024) https://doi.org/10.48550/arXiv.2409.06366

Bohlinger, N.; Tateo, D.; Kicki, P.; Walas, K.; Peters, J. Benefits of an Actuated Spine in Agile Quadruped Locomotion. ICRA2024 Workshop on Bio-inspired Robotics and Robotics for Biology (2024)

Bohlinger, N.; Czechmanowski, G.; Krupka, M.; Kicki, P.; Walas, K.; Peters, J.; Tateo, D.; One Policy to Run Them All: an End-to-end Learning Approach to Multi-Embodiment Locomotion;Workshop on Embodiment-Aware Robot Learning, Robotics: Science and Systems (2024)

Last Update: 2025-01-10 16:28