Optimal Gait Control for a Tendon-driven Soft Quadruped Robot by Model-based Reinforcement Learning

Xuezhi Niu[†1], Kaige Tan[†2], Didem Gürdür Broo[1] & Lei Feng[2]

[1] Cyber-Physical Systems Lab, Department of Information Technology, Uppsala University
[2] Unit of Mechatronics, KTH Royal Institute of Technology

May 21, 2025

IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION

UPPSALA UNIVERSITET

KTH VETENSKAP OCH KONST

Key challenges:

- High simulation cost
- Inefficient RL training
- Reality gap

Our Contributions:

- A fast DNN Surrogate dynamics model
- MBRL + post-training
- Action-space reduction
- Hardware validation

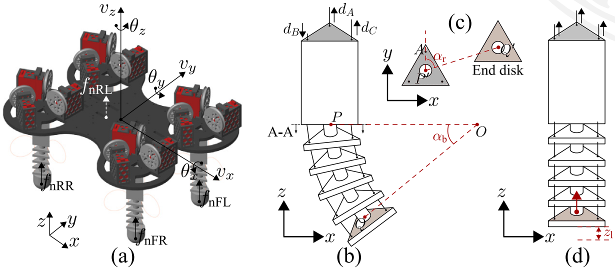


Figure 1: Overview of SoftQ and CTSA: (a) Rendered robot with key states. (b) CTSA bending angle α_b. (c) CTSA rotational angle α_r. (d) CTSA compression length z_l.

Method Overview

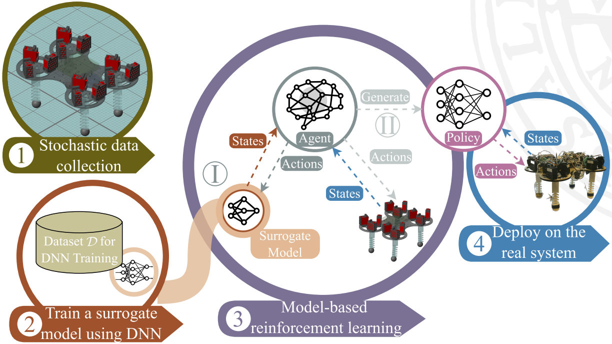


Figure 2: Gait control policy generation workflow.

- Surrogate model for fast training
- Parametric model for reduced action space
- Post-training in high-fidelity SimScape

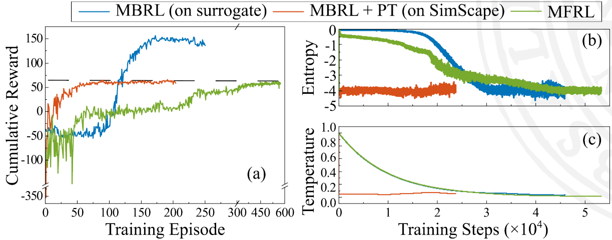


Figure 3: Training results at 0.2 m/s reference speed, comparing MBRL, MBRL+PT, and MFRL [1]. (a) Cumulative reward over episodes, (b) entropy, and (c) temperature evolution during training.

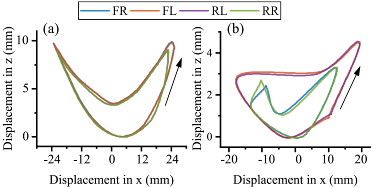


Figure 4: Foot trajectories for SoftQ by using (a) expert gait [2] and (b) MBRL+PT.

Table 1: Quantitative comparison of the performance

		MBRL+PT	MFRL	Expert
stability	[min,max]	[-0.04, 0.8]	[0.02, 0.745]	-
	avg	0.65	0.03	0.73
COT ($J/kg/m$)	[min,max]	[41, 146]	[40, 1375]	-
	avg	68	501	81
training time (h)	[min,max]	[7.9, 19.8]	[19.7, 61.8]	-
	avg	10.8	22.6	-
speed (m/s)	[min,max]	[0.04, 0.47]	[0.002, 0.22]	-
	avg	0.26	0.003	0.09

Key numbers:

- Training time: 10.8 h vs 22.6 h (MBRL+PT vs MFRL)
- Speed: 0.26 m/s vs 0.09 m/s (MBRL+PT vs expert)

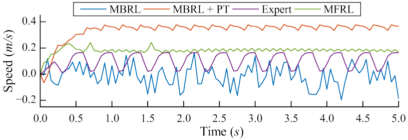


Figure 5: Resultant forward walking speed in simulation for expert gait, MBRL, MBRL+PT, and MFRL.

Real-World Validation

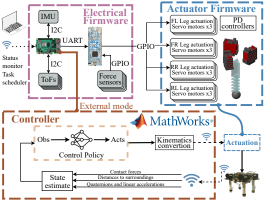
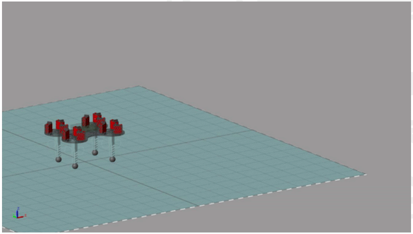


Figure 6: Control architecture.



Result: 0.15 m/s achieved in reality

Conclusions & Future Works

Takeaways:

- MBRL+PT enables efficient and stable gait learning with less training time and lower energy cost compared to MFRL.
- Validated on hardware: Achieves 0.15 m/s ($\approx$1 BL/s), three times faster than prior model-free RL.

Future Work

- Real-time adaptation to varying terrains via online learning.
- Robust generalization using domain randomization and sensor fusion.

References

Back to start

[1] Q. Ji, S. Fu, K. Tan, S. T. Muralidharan, K. Lagrelius, D. Danelia, G. Andrikopoulos, X. V. Wang, L. Wang, and L. Feng, "Synthesizing the optimal gait of a quadruped robot with soft actuators using deep reinforcement learning," *Robotics and Computer-Integrated Manufacturing*, vol. 78, p. 102382, 2022.

[2] Q. Ji, S. Fu, L. Feng, G. Andrikopoulos, X. V. Wang, and L. Wang, "Omnidirectional walking of a quadruped robot enabled by compressible tendon-driven soft actuators," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 11 015–11 022.