

Increasing Interaction Fidelity: Training Routines for Biomechanical Models in HCI

Michał Patryk Miazga
miazga@uni-leipzig.de
ScaDS.AI, Leipzig University
Leipzig, Germany

Patrick Ebel
ebel@uni-leipzig.de
ScaDS.AI, Leipzig University
Leipzig, Germany

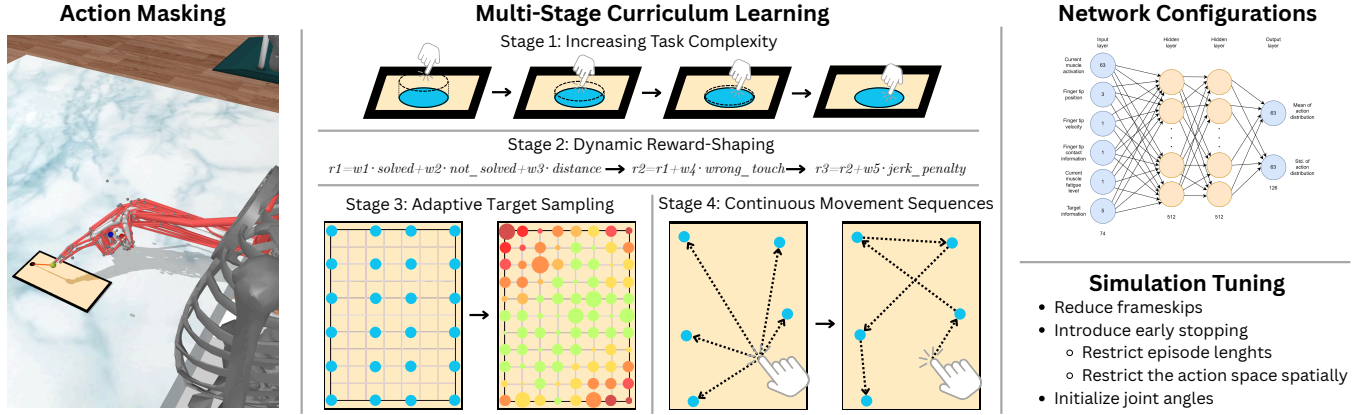


Figure 1: Overview of training routines to improve RL-based biomechanical forward simulations.

Abstract

Biomechanical forward simulation holds great potential for HCI, enabling the generation of human-like movements in interactive tasks. However, training biomechanical models with reinforcement learning is challenging, particularly for precise and dexterous movements like those required for touchscreen interactions on mobile devices. Current approaches are limited in their interaction fidelity, require restricting the underlying biomechanical model to reduce complexity, and do not generalize well. In this work, we propose practical improvements to training routines that reduce training time, increase interaction fidelity beyond existing methods, and enable the use of more complex biomechanical models. Using a touchscreen pointing task, we demonstrate that curriculum learning, action masking, more complex network configurations, and simple adjustments to the simulation environment can significantly improve the agent’s ability to learn accurate touch behavior. Our work provides HCI researchers with practical tips and training routines for developing better biomechanical models of human-like interaction fidelity.

CCS Concepts

• **Human-centered computing** → **User models; Pointing; Systems and tools for interaction design;** • **Computing methodologies** → **Reinforcement learning.**

Keywords

Biomechanical Modeling; Simulated Agents; Deep Reinforcement Learning; Machine Learning

ACM Reference Format:

Michał Patryk Miazga and Patrick Ebel. 2025. Increasing Interaction Fidelity: Training Routines for Biomechanical Models in HCI. In *The 38th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '25)*, September 28–October 1, 2025, Busan, Republic of Korea. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3746058.3758385>

1 Introduction

Simulated users that model human-like, dexterous interactions with technology hold great potential to make user-centered design more efficient, rigorous, and predictable [13, 16]. At the core of these systems are biomechanical models that capture anatomical and physiological detail. Advances in computational power have made it possible to use increasingly complex models, such as those in *MyoSuite* [2, 25] and *User-in-the-Box* [11, 12], which can simulate muscle fatigue or neuromuscular diseases.

However, training Reinforcement Learning (RL) agents to learn muscle-level control in physically simulated environments remains a highly complex task. Training is often brittle, slow, and difficult to reproduce. This limits interaction fidelity and accordingly the broader adoption of biomechanical modeling in the Human-Computer Interaction (HCI) community for technology design and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UIST Adjunct '25, Busan, Republic of Korea

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2036-9/2025/09

<https://doi.org/10.1145/3746058.3758385>

Table 1: Evaluation results represent the mean values of 100 episodes, each with a maximum duration of 10 seconds.

Network Size	Max Timesteps	Action Masking	Curriculum Learning	Stage 2 Dynamic Reward	Stage 2 Early Reward	Button Radius (mm)	Success Rate	Avg. Errors / Success Episode	Avg. Time / Success Episode (s)
128 × 128	200M	✓	✓	✗	✓	1.5	82%	176.08	3.39
256 × 256	200M	✓	✓	✓	✗	1.5	99%	156.81	3.35
256 × 256	800M	✗	✓	✓	✗	1.5	5%	3.6	1.6
512 × 512	200M	✓	✓	✓	✗	1.5	100%	8.07	1.40
512 × 512	200M	✓	✗	✗	✗	6	0%	0	✗
512 × 512	200M	✓	✗	✗	✓	1.5	64%	179.56	3.62

evaluation. For instance, in *User-in-the-Box* [6, 11], the pointing targets had a radius of 5–15 cm. Given that the iPhone 16 has a diagonal of 15.5 cm this is large, especially when such models should be applied to mobile use cases. Furthermore, the action space for arm movements was simplified by disabling wrist flexion and various muscles, effectively reducing the model to five degrees of freedom and 26 active muscles. In *Sim2VR*, Fischer et al. [7] also reduced the number of muscles and replaced some with joint torque actuators, which are easier to control. Their model also focuses primarily on coarse movements in VR. Overall, current approaches demonstrate the potential of musculoskeletal forward simulation for HCI, but they fall short of enabling the precise control required for most interactive tasks. We argue that these shortcomings are primarily due to ineffective training routines. Although reward shaping has been studied [3, 19, 21], the influence of other training routines on the learning of fine motor skills remains underexplored.

In this work, we present and evaluate performance improvements achieved through action masking, curriculum learning, alternative policy network configurations, and task-specific improvements to the simulation environment. Based on our findings, we propose actionable training routines that lower technical barriers and support the development of high-fidelity, human-like control in physics-based biomechanical simulations for HCI.

2 Training Routines for Biomechanical Models

Our primary goal is to learn stable muscle-level control policies that enable precise and dexterous movements. All evaluations are conducted using Proximal Policy Optimization (PPO) [20], and their performance is assessed based on the success rate of touch interactions. We consider an interaction as successful if the target element is touched within 10 seconds after the start of the episode. We further count unintended interactions outside the target area (e.g., due to overshooting) as errors, indicating suboptimal control and a lack of fine motor skills. In the following, we introduce all the routines evaluated in this work.

Action Masking. As shown in prior work [7, 10, 11, 15, 22], action masking is an effective technique to reduce the dimensionality of the action space, limit unnecessary exploration, and accelerate learning. In our implementation, we apply action masking by disabling all fingers except the index finger. Our results (see Table 1) show that action masking highly improves training efficiency.

Multi-Stage Curriculum Learning. In RL, curriculum learning improves learning efficiency and final performance by training agents on tasks of gradually increasing difficulty [14, 17, 18]. To achieve the fine motor control required for touchscreen interactions,

we introduce a structured, multi-stage curriculum. Each curriculum stage increases complexity via sub-stages. The agent advances to the next (harder) stage only once it meets a performance threshold in the final sub-stage of the current curriculum stage. As shown in Table 1 and illustrated in Figure 1, the following curriculum stages significantly improve training efficiency:

Stage 1: Increasing Task Complexity. To simplify early training, the agent begins from a fixed position and must reach a large, button-shaped 3D target placed just above the surface with its fingertip. As training progresses, the 3D target gradually flattens into a 2D surface (see Figure 1), shifting the task from interacting with a volumetric object to a more demanding surface-based interaction that requires greater precision.

Stage 2: Dynamic Reward-Shaping. Reward shaping involves modifying the reward function to guide the agent toward desired behaviors more effectively [4, 5, 8, 9, 24]. Our experiments show that starting with a simple reward to then gradually increase reward complexity, based on the training progress, significantly improves performance. Early rewards penalize incorrect button presses, while later stages discourage jerky movements and excessive muscle effort. Without this step-wise shaping, the agent fails to learn meaningful behavior.

Stage 3: Adaptive Target Sampling. Once the agent achieves satisfactory performance in the early stages of the curriculum, adaptive target sampling is used to improve generalization. The interaction surface is first discretized, and target sizes are randomized by uniformly sampling radii between 1.5 mm and 7 mm—encouraging the agent to learn the speed-accuracy trade-off observed in human interaction [26]. We also found that success rates vary significantly across target locations. To address this, we sample low-performing target locations more frequently, promoting balanced performance across the task environment.

Stage 4: Continuous Movement Sequences. Human interaction with mobile devices such as smartphones can be seen as a sequence of movement primitives. Thus the agent must learn to perform any interaction movement from any starting position. To achieve this, each new target is sampled relative to the agent’s current position and velocity.

Network Configurations. Most current approaches [7, 11] are limited by the use of small policy networks, which struggle to capture the complexity of fine-motor tasks. While smaller networks are easier to train, they lack the capacity for precise control. In contrast, our training routines, combined with tuned optimizer settings, enable the use of significantly larger networks, leading to improved performance. We set the learning rate and clip range to 6×10^{-4} and 0.2, respectively, and reduce both linearly over

time. At timestep t , each parameter p is computed as: $p(t) = p_{t=0} \times r(t)$, where $r(t)$ linearly decays from 1 to 0 as training progresses. This schedule encourages exploration early on and gradually shifts toward more stable convergence.

Our results (Table 1) further show that a larger policy network with a 512×512 architecture significantly outperforms smaller networks, enabling more accurate and stable fine-motor control.

Simulation Tuning. We propose multiple small tweaks to the simulation environment, such as (1) reducing the frameskip parameter, which controls how frequently the agent receives observations and issues actions in MuJoCo [1, 23]. We found that a low value (e.g., 3) allows fine-grained control, while higher values (e.g., 10) cause overshooting due to prolonged muscle activations. (2) We introduce early stopping criteria that terminate episodes whenever the agent exceeds the episode length restriction or leaves the permitted 3D action space constraint, thereby enforcing realistic temporal and spatial boundaries. (3) We initialize joint positions within the valid 3D action space, setting the index finger to point forward, to ensure consistent, task-relevant initial conditions.

3 Conclusion

This work demonstrates that tailored training routines and hyperparameter settings are key to achieving accurate and dexterous control in RL-based musculoskeletal simulations. By evaluating task success and errors, we show that reliable fine-motor behavior is achievable, overcoming limitations of prior approaches. Our methods improve performance while lowering technical barriers, making biomechanical simulations more accessible for HCI research and applications. Implementations of our routines are available at: <https://github.com/ciao-group/RL-training-routines-for-Biomechanical-Models>.

Acknowledgments

The authors acknowledge the financial support by the Federal Ministry of Research, Technology and Space of Germany and by Sächsische Staatsministerium für Wissenschaft, Kultur und Tourismus in the programme Center of Excellence for AI-research „Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig“, project identification number: ScaDS.AI

References

- [1] Alex Braylan, Mark Hollenbeck, Elliot Meyerson, and Risto Miikkilainen. 2000. Frame skip is a powerful parameter for learning to play atari. *Space* 1600 (2000), 1800.
- [2] Vittorio Caggiano, Huawei Wang, Guillaume Durandau, Massimo Sartori, and Vikash Kumar. 2022. MyoSuite-A contact-rich simulation suite for musculoskeletal motor control. *arXiv preprint arXiv:2205.13600* (2022).
- [3] Wenqian Chen, Yaru Chen, Yongxuan Wang, and Rong Liu. 2023. Static standing balance with musculoskeletal models using ppo with reward shaping. *Procedia Computer Science* 226 (2023), 78–84.
- [4] Daniel Dewey. 2014. Reinforcement Learning and the Reward Engineering Principle. In *AAAI spring symposia*.
- [5] Jonas Eschmann. 2021. Reward function design in reinforcement learning. *Reinforcement learning algorithms: Analysis and Applications* (2021), 25–33.
- [6] Florian Fischer, Miroslav Bachinski, Markus Klar, Arthur Fleig, and Jörg Müller. 2021. Reinforcement learning control of a biomechanical model of the upper extremity. *Scientific Reports* 11, 1 (2021), 14445.
- [7] Florian Fischer, Aleks Ikkala, Markus Klar, Arthur Fleig, Miroslav Bachinski, Roderick Murray-Smith, Perttu Hämmäläinen, Antti Oulasvirta, and Jörg Müller. 2024. SIM2VR: Towards Automated Biomechanical Testing in VR. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–15.
- [8] Marek Grzes. 2017. Reward shaping in episodic reinforcement learning. (2017).
- [9] Yujing Hu, Weixun Wang, Hangtian Jia, Yixiang Wang, Yingfeng Chen, Jianye Hao, Feng Wu, and Changjie Fan. 2020. Learning to utilize shaping rewards: A new approach of reward shaping. *Advances in Neural Information Processing Systems* 33 (2020), 15931–15941.
- [10] Shengyi Huang and Santiago Ontañón. 2020. A closer look at invalid action masking in policy gradient algorithms. *arXiv preprint arXiv:2006.14171* (2020).
- [11] Aleks Ikkala, Florian Fischer, Markus Klar, Miroslav Bachinski, Arthur Fleig, Andrew Howes, Perttu Hämmäläinen, Jörg Müller, Roderick Murray-Smith, and Antti Oulasvirta. 2022. Breathing life into biomechanical user models. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–14.
- [12] Aleks Ikkala and Perttu Hämmäläinen. 2020. Converting biomechanical models from opensim to mujoco. In *International Conference on NeuroRehabilitation*. Springer, 277–281.
- [13] Martin Lorenz, Tiago Amorim, Debargha Dey, Mersedeh Sadeghi, and Patrick Ebel. 2024. Computational Models for In-Vehicle User Interface Design: A Systematic Literature Review. In *Proceedings of the 16th AutomotiveUI Conference*. ACM, Stanford CA USA, 204–215. <https://doi.org/10.1145/3640792.3675735>
- [14] Sha Luo, Hamidreza Kasaei, and Lambert Schomaker. 2020. Accelerating reinforcement learning for reaching using continuous curriculum learning. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.
- [15] Michał Patryk Miazga, Daniel Abitz, Matthias Täschner, and Erhard Rahm. 2025. Automated Configuration of Schema Matching Tools: A Reinforcement Learning Approach. *Datenbanksysteme für Business, Technologie und Web (BTW 2025)* (2025), 331.
- [16] Roderick Murray-Smith, Antti Oulasvirta, Andrew Howes, Jörg Müller, Aleks Ikkala, Miroslav Bachinski, Arthur Fleig, Florian Fischer, and Markus Klar. 2022. What simulation can do for HCI research. *Interactions* 29, 6 (Nov. 2022), 48–53. <https://doi.org/10.1145/3564038>
- [17] Sanmit Narvekar. 2017. Curriculum Learning in Reinforcement Learning. In *IJCAI*. 5195–5196.
- [18] Sanmit Narvekar, Bei Peng, Matteo Leonetti, Jivko Sinapov, Matthew E Taylor, and Peter Stone. 2020. Curriculum learning for reinforcement learning domains: A framework and survey. *Journal of Machine Learning Research* 21, 181 (2020), 1–50.
- [19] Katharine Nowakowski, Philippe Carvalho, Jean-Baptiste Six, Yann Maillet, Anh Tu Nguyen, Ismail Seghiri, Loick M’pemb, Theo Marcille, Sy Toan Ngo, and Tien-Tuan Dao. 2021. Human locomotion with reinforcement learning using bioinspired reward reshaping strategies. *Medical & biological engineering & computing* 59 (2021), 243–256.
- [20] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- [21] Hannah Selder, Florian Fischer, Per Ola Kristensson, and Arthur Fleig. 2025. What Makes a Model Breathe? Understanding Reinforcement Learning Reward Function Design in Biomechanical User Simulation. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–10.
- [22] Roland Stolz, Hanna Krasowski, Jakob Thumm, Michael Eichelbeck, Philipp Gassert, and Matthias Althoff. 2024. Excluding the irrelevant: Focusing reinforcement learning through continuous action masking. *Advances in Neural Information Processing Systems* 37 (2024), 95067–95094.
- [23] Emanuel Todorov, Tom Erez, and Yuval Tassa. 2012. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 5026–5033.
- [24] Peter Vamplew, Benjamin J Smith, Johan Källström, Gabriel Ramos, Roxana Rădulescu, Diederik M Roijers, Conor F Hayes, Fredrik Heintz, Patrick Mannion, Pieter JK Libin, and others. 2022. Scalar reward is not enough: A response to Silver, Singh, Precup and Sutton (2021). *Autonomous Agents and Multi-Agent Systems* 36, 2 (2022), 41. <https://doi.org/10.1007/s10458-022-09575-5> Publisher: Springer.
- [25] Huawei Wang, Vittorio Caggiano, Guillaume Durandau, Massimo Sartori, and Vikash Kumar. 2022. MyoSim: Fast and physiologically realistic MuJoCo models for musculoskeletal and exoskeletal studies. In *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 8104–8111.
- [26] Jacob O. Wobbrock, Edward Cutrell, Susumu Harada, and I. Scott MacKenzie. 2008. An error model for pointing based on Fitts’ law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy) (CHI ’08)*. Association for Computing Machinery, New York, NY, USA, 1613–1622. <https://doi.org/10.1145/1357054.1357306>