roto 2.0: The Robot Tactile Olympiad

Elle Miller[1], Jayaram Reddy[2], Ayush Deshmukh[1], Trevor McInroe[1], David Abel[1], Oisin Mac Aodha[1], Sethu Vijayakumar[1]

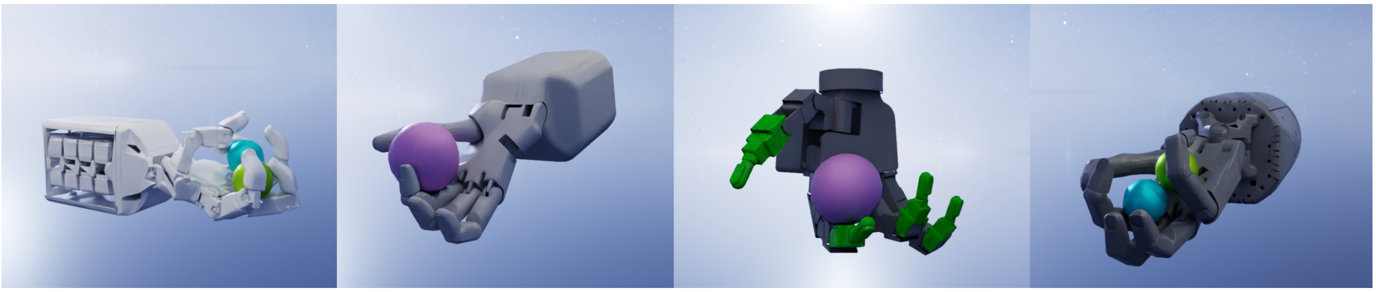

Fig. 1. **The roto 2.0 Benchmark Suite:** A standardised RL framework across four distinct dexterous morphologies (from L to R): ORCA Hand, Shadow Lite, Allegro Hand, and the Shadow Dexterous Hand. The suite facilitates "blind" tactile manipulation tasks, such as Baoding ball rotation and ball bouncing.

***Abstract*—Tactile-based reinforcement learning (RL) is currently hindered by fragmented research and a focus on over-saturated orientation tasks. We introduce v2 of the Robot Tactile Olympiad (roto 2.0), a GPU-parallelised benchmark designed to standardise tactile-based RL across four distinct robotic morphologies (16-DOF to 24-DOF). Unlike prior benchmarks, roto focuses on end-to-end "blind" manipulation, utilising only proprioception and tactile sensing without state information or distillation. We demonstrate a significant performance leap, with our blind agents achieving 13 Baoding ball rotations in 10 seconds, an order of magnitude faster than current state-of-the-art speeds. By open-sourcing our environments and robustly tuned baselines, we reduce the barrier to entry and enable researchers to prioritise fundamental algorithmic challenges over tedious RL tuning. Website: https://elle-miller.github.io/roto/**

I. Introduction

Real-world robotic manipulation requires the ability to interact robustly in unstructured environments where visual line-of-sight is frequently obstructed. To achieve this, robots must learn to "feel." While Reinforcement Learning (RL) has revolutionised locomotion across complex terrains [1], tactile-based manipulation lags significantly behind. Progress is currently hindered by a fragmented landscape: most labs work in isolation using unique combinations of sensors and robots, making cross-validation difficult. Furthermore, the community has largely over-saturated the task of in-hand orientation, e.g. [2], [3], [4], [5], [6], [7], [8], [9]. While impressive, a single task fails to capture the broader spectrum of challenges that tactile observations pose, leaving the true utility of tactile feedback an open debate. The difficulty of tactile-based RL stems from a trifecta of complexity: manipulation is inherently hard [10], on-policy RL is notoriously difficult to tune, and thus effectively combining the two with sparse and discontinuous tactile observations is a significant undertaking. This difficulty is exacerbated by a lack of standardised tactile-based RL benchmark environments across complex tasks and morphologies. *Tactile-Gym 2.0* [11] is limited to a 4-DOF arm and single-point contact tasks, while *VTDexManip* [12] focuses on pre-training policies with pre-collected datasets. To fill this gap, we introduce v2 of the Robot Tactile Olympiad (**roto 2.0**), an RL benchmark built upon GPU-parallelised Isaac Lab [13]. Originally introduced in [14] for the Shadow Hand, we expand the suite here to include four distinct dexterous morphologies: the anthropomorphic Shadow Dexterous Hand (24-DOF), Shadow Dexterous Hand Lite (16-DOF), Allegro Hand (16-DOF) and ORCA Hand (17-DOF). Crucially, we demonstrate that with our RL training pipeline, "blind" policies (utilising only proprioception and tactile data) can master sophisticated manipulation without the need for the teacher-student distillation or explicit pose estimators common in prior work [3], [15], [16]. While the current state-of-the-art for Baoding ball rotation achieves a maximum of 3 rotations in 10 seconds [17], our blind agents exploit the raw potential of tactile-proprioceptive loops to achieve significantly higher throughput of up to 15 rotations. As demonstrated in [14], integrating self-supervised forward dynamics to aid representation learning can further push this boundary to 25 rotations per 10 seconds, nearly closing the gap between blind policies and state-based agents. From these results, we argue that establishing robust haptic foundations is a prerequisite to integrating visual modalities, rather than a secondary addition to vision-centric systems. By open-sourcing **roto**, we aim to reduce the barrier-to-entry for researchers interested in tactile-based RL and help focus community efforts on high-impact research directions instead of RL tuning. Contributions:

- **The roto benchmark:** A set of tactile-based environments with integrated hyperparameter optimisation and robustly tuned baselines.
- **Morphological diversity:** A cross-platform evaluation of four robot hands to study how hardware complexity influences tactile policy convergence.
- **Performance breakthroughs:** We provide "blind" policies that achieve state-of-the-art speeds in complex manipulation, providing a new ceiling for tactile intelligence.

[1]University of Edinburgh, UK. Email: elle.miller@ed.ac.uk
[2]National University Singapore, Singapore

II. Methodology

RL. We use a customised implementation of Proximal Policy Optimisation (PPO) [18] from SKRL [19] to incorporate observation stacking, self-supervision, separated environments for continuous evaluation, and various training tricks [20]. We use $8,092$ parallelised environments for training and 100 for evaluation. For each combination of robot ($n = 3$), task ($n = 2$), and observation setting (blind vs. state-based), we perform a hyperparameter sweep across seven PPO parameters using 40 trials (8 warm-up runs) to ensure robust baselines.

- *Bounce:* The agent must bounce a ball as many times as possible in 10 seconds (600 timesteps). A bounce is defined as a contact event after a period of at least 5 timesteps ($\sim$ 83ms) without contact.
- *Baoding:* The agent must rotate two balls (55g) around each other in-hand as many times as possible within 10 seconds (600 timesteps). We use 1.5 inch diameter for Shadow Hand and ORCA, 2 inches for Allegro and 1.2 inches for Shadow Lite.

MDP. The blind agents receive a history length $k = 4$ of proprioceptive and binary tactile observations, and are joint-position controlled, see Table I for details. We define the each task-relevant robot link as a tactile sensor. The state-based agents additionally receive the object position(s) and linear velocity. For *Bounce*, the agent is rewarded with $r_{bounce} = 10$ for every successful bounce. For *Baoding*, we specify two static target positions and define the reward as $r_{dist_1} + r_{dist_2} + r_{rotation}$. The dense distance rewards r_{dist_1}, r_{dist_2} encourage the balls to the targets. When the centers of both balls are within 1.0 cm of the targets, the targets switch and the agent receives a bonus reward $r_{rotation} = 10$. The episode terminates if any object is out of reach or the maximum episode length $T = 600$ is reached. The physics simulation runs at 240 Hz, the control policy at 60 Hz.

TABLE I
OBSERVATION AND ACTION SPACES

Type	Description	*Shadow*	*Shadow Lite*	*Allegro*	*ORCA*
Tactile obs.	binary contacts	17	14	20	17
Proprio. obs.	joint positions	20	16	16	17
	joint velocities	20	16	16	17
	joint command error	20	13	16	17
	last action	20	13	16	17
Total obs.	single timestep	97	72	84	85
	$k = 4$ timesteps	388	288	336	340
Actions	joint positions	20	13	10	17

III. Experimental results

We evaluate the learning efficiency and asymptotic performance of four distinct morphologies for state-based and blind agents. The mean evaluation returns are summarised in Figure 2; we refer the reader to the project page for the policy videos. In the Bounce environment, state-based agents approach the theoretical maximum return ($1,000$ reward, corresponding to 100 successful bounces). While the hands converge to similar success rates, they exhibit hardware-specific strategies, e.g. the ORCA hand adopts an "outstretched" starfish-like pose. Our blind agents demonstrate high sample efficiency, approaching 80 bounces by 200M steps. Despite the vast differences in hardware, we find that performance trends remain consistent across the full-hand morphologies, excepting the Shadow Lite. The Baoding task reveals a more significant performance gap between state-based and blind agents. State-based agents achieve a throughput of up to 35 rotations in 10 seconds. Blind agents exhibit much lower performance with high variance; while a top-performing Shadow Hand seed achieved 13 rotations, others failed to converge. This stochasticity highlights a core challenge in tactile-based RL: efficient feature extraction [14].

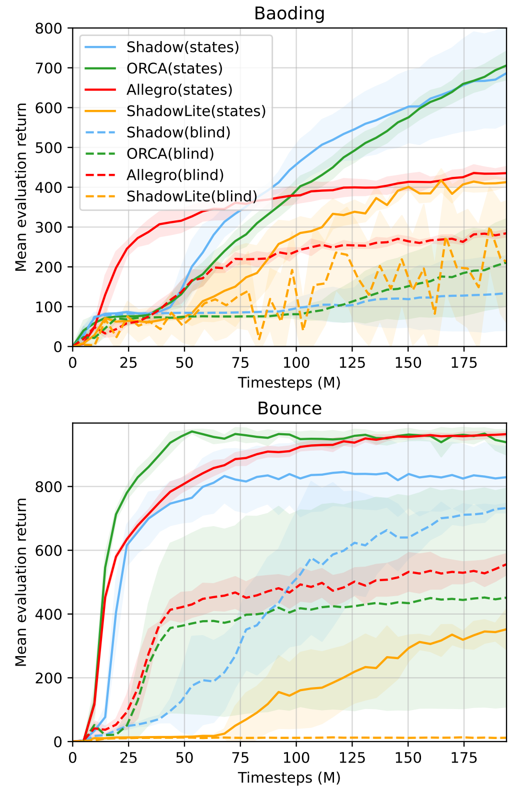

Fig. 2. Mean evaluation returns across 5 seeds for state-based and blind agents in the Baoding and Bounce tasks.

IV. Discussion & Conclusion

We introduce **`roto 2.0`**, a multi-morphology benchmark for blind dexterous manipulation. While our high-speed simulated policies represent a "performance ceiling" that exceeds current real-world hardware limits, they provide a benchmark for developing the next generation of robust RL pipelines. Our results indicate that while blind policies can approach privileged performance in simple tasks like bouncing, complex or multi-object manipulation (Baoding) remains an open challenge. We identify high-priority research directions for the community: beyond sparse binary contacts to richer forms of tactile information, ML methodologies with inductive biases for tactile data, and expanding tasks beyond hands e.g. whole-body humanoid manipulation [21]. We are actively investigating the sim-to-real transferability of these policies and welcome community contributions to the **roto** suite to accelerate the arrival of the "locomotion moment" for robotic touch.

References

[1] N. Rudin, J. He, J. Aurand, and M. Hutter, "Parkour in the wild: Learning a general and extensible agile locomotion policy using multi-expert distillation and rl fine-tuning," 2025. [Online]. Available: https://arxiv.org/abs/2505.11164

[2] O. M. Andrychowicz, B. Baker, M. Chociej, R. Józefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, and W. Zaremba, "Learning dexterous in-hand manipulation," *The International Journal of Robotics Research*, vol. 39, no. 1, Jan. 2020. [Online]. Available: https://doi.org/10.1177/0278364919887447

[3] L. Sievers, J. Pitz, and B. Bäuml, "Learning Purely Tactile In-Hand Manipulation with a Torque-Controlled Hand," in *2022 International Conference on Robotics and Automation (ICRA)*, May 2022, pp. 2745–2751. [Online]. Available: https://ieeexplore.ieee.org/document/9812093

[4] T. Chen, J. Xu, and P. Agrawal, "A System for General In-Hand Object Re-Orientation."

[5] H. Qi, A. Kumar, R. Calandra, Y. Ma, and J. Malik, "In-Hand Object Rotation via Rapid Motor Adaptation," in *Conference on Robot Learning (CoRL)*, 2022.

[6] H. Qi, B. Yi, S. Suresh, M. Lambeta, Y. Ma, R. Calandra, and J. Malik, "General In-hand Object Rotation with Vision and Touch," in *Proceedings of The 7th Conference on Robot Learning*. PMLR, Dec. 2023, pp. 2549–2564, iSSN: 2640-3498. [Online]. Available: https://proceedings.mlr.press/v229/qi23a.html

[7] L. Röstel, J. Pitz, L. Sievers, and B. Bäuml, "Estimator-Coupled Reinforcement Learning for Robust Purely Tactile In-Hand Manipulation," in *2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*. Austin, TX, USA: IEEE, Dec. 2023, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/10375194/

[8] M. Yang, C. Lu, A. Church, Y. Lin, C. Ford, H. Li, E. Psomopoulou, D. A. W. Barton, and N. F. Lepora, "Anyrotate: Gravity-invariant in-hand object rotation with sim-to-real touch," in *Conference on Robot Learning (CoRL)*, 2024.

[9] S. Suresh, H. Qi, T. Wu, T. Fan, L. Pineda, M. Lambeta, J. Malik, M. Kalakrishnan, R. Calandra, M. Kaess, J. Ortiz, and M. Mukadam, "Neural feels with neural fields: Visuo-tactile perception for in-hand manipulation," *Science Robotics*, p. adl0628, 2024.

[10] M. T. Mason, "Toward Robotic Manipulation," *Annual Review of Control, Robotics, and Autonomous Systems*, no. 1, 2018.

[11] Y. Lin, J. Lloyd, A. Church, and N. F. Lepora, "Tactile gym 2.0: Sim-to-real deep reinforcement learning for comparing low-cost high-resolution robot touch," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10 754–10 761, 2022.

[12] Q. Liu, Y. Cui, Z. Sun, G. Li, J. Chen, and Q. Ye, "Vtdexmanip: A dataset and benchmark for visual-tactile pretraining and dexterous manipulation with reinforcement learning," in *The Thirteenth International Conference on Learning Representations*, 2025. [Online]. Available: https://openreview.net/forum?id=jf7C7EGw21

[13] M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, J. L. Yuan, R. Singh, Y. Guo, H. Mazhar, A. Mandlekar, B. Babich, G. State, M. Hutter, and A. Garg, "Orbit: A unified simulation framework for interactive robot learning environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3740–3747, 2023.

[14] E. Miller, T. McInroe, D. Abel, O. Mac Aodha, and S. Vijayakumar, "Enhancing tactile-based reinforcement learning for robotic control," in *NeurIPS*, 2025.

[15] Z.-H. Yin, B. Huang, Y. Qin, Q. Chen, and X. Wang, "Rotating without Seeing: Towards In-hand Dexterity through Touch," Mar. 2023, arXiv:2303.10880 [cs]. [Online]. Available: http://arxiv.org/abs/2303.10880

[16] L. Yang, B. Huang, Q. Li, Y.-Y. Tsai, W. W. Lee, C. Song, and J. Pan, "TacGNN: Learning Tactile-Based In-Hand Manipulation With a Blind Robot Using Hierarchical Graph Neural Network," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3605–3612, Jun. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10093019/

[17] Y. Yuan, H. Che, Y. Qin, B. Huang, Z.-H. Yin, K.-W. Lee, Y. Wu, S.-C. Lim, and X. Wang, "Robot synesthesia: In-hand manipulation with visuotactile sensing," in *ICRA*, 2024.

[18] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[19] A. Serrano-Muñoz, D. Chrysostomou, S. Bøgh, and N. Arana-Arexolaleiba, "skrl: Modular and flexible library for reinforcement learning," *Journal of Machine Learning Research*, vol. 24, no. 254, pp. 1–9, 2023. [Online]. Available: http://jmlr.org/papers/v24/23-0112.html

[20] E. Miller, "The art of robot reinforcement learning," *parallelles.substack.com*, 2026. [Online]. Available: https://parallelles.substack.com/p/the-art-of-robot-reinforcement-learning

[21] C. Sferrazza, D.-M. Huang, X. Lin, Y. Lee, and P. Abbeel, "Humanoid-bench: Simulated humanoid benchmark for whole-body locomotion and manipulation," 2024.