Enhancing Tactile-based RL for Robotic Control

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

1 THE UNIVERSITY of EDINBURGH 2 DeepMind

NEURAL INFORMATION PROCESSING SYSTEMS 2025

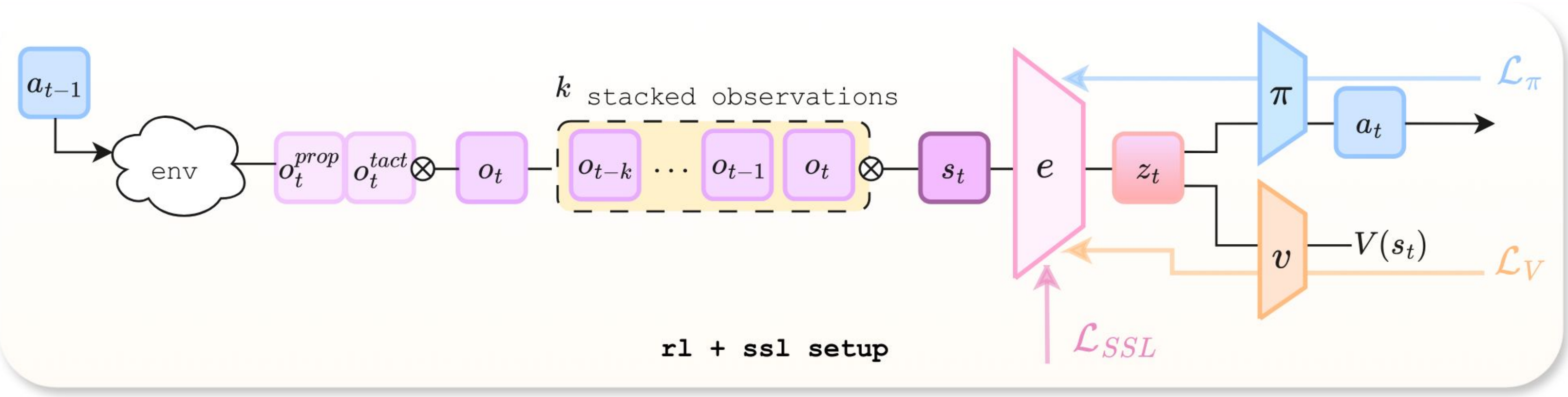
Stay up to date:

ellemiller101

elle-miller.github.io/tactile_rl

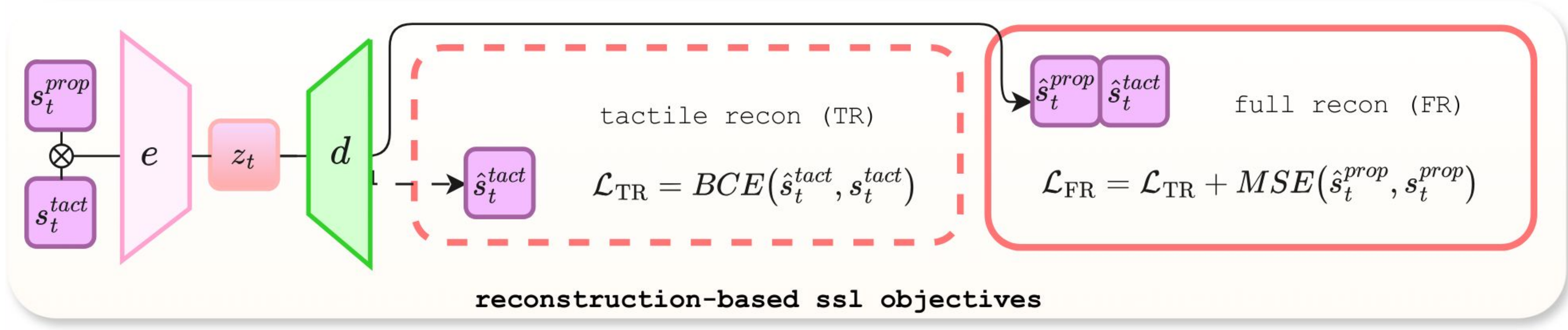
What's the problem?

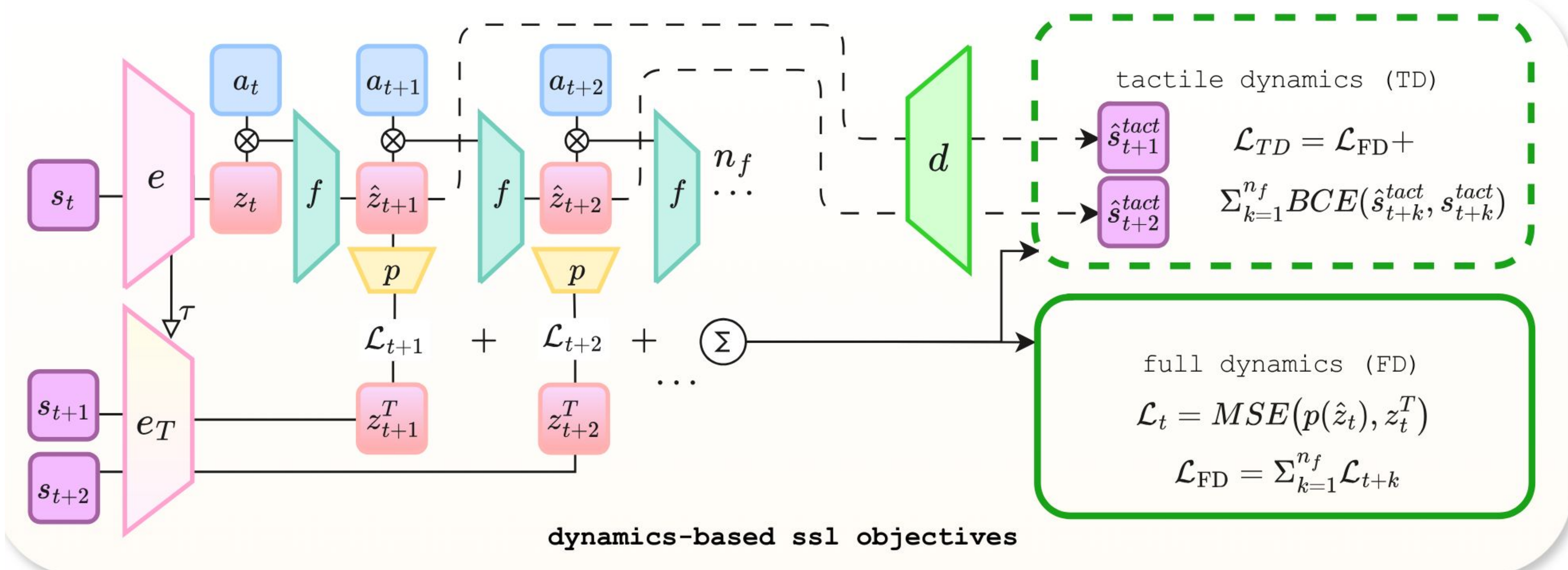
RL agents must learn **both** a useful observation representation & policy optimisation with **one** supervisory signal: reward



Self-supervised learning (SSL) provides an additional signal for converting complex observations into useful representations

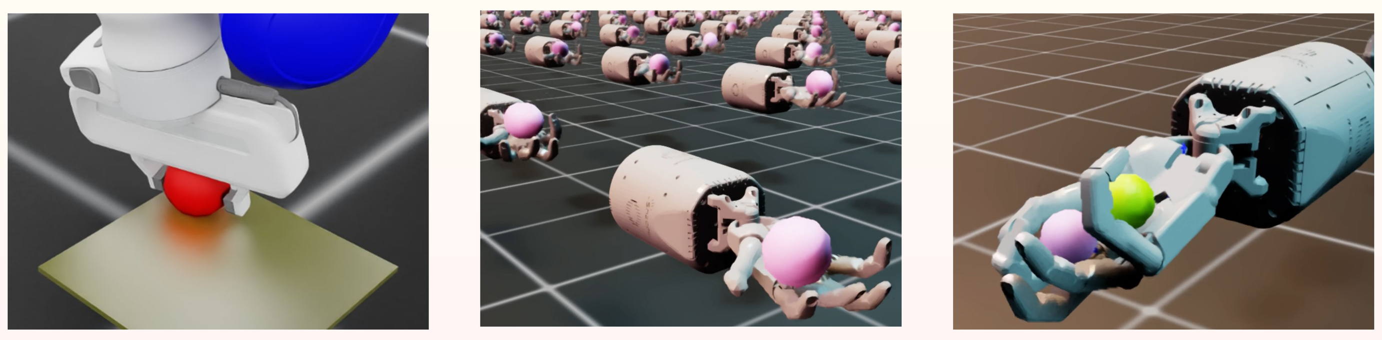
Proposition: self-supervision





RoTO: Robot Tactile Olympiad

To inspire progress in tactile-based dexterity, we are releasing our novel Isaac Lab environments as a new benchmark! We include in-built hyperparameter optimisation, robustly tuned baselines, self-supervision integration, and more.

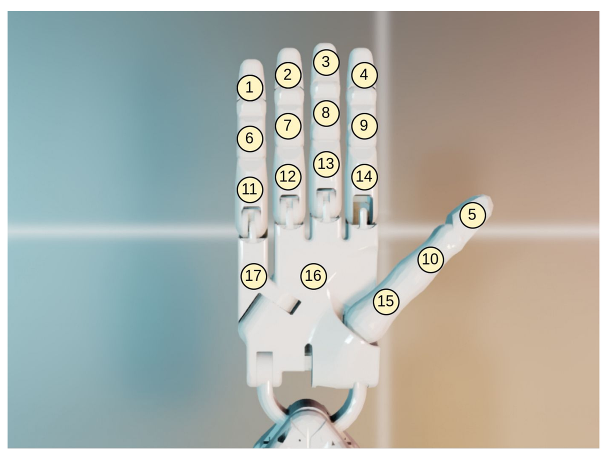


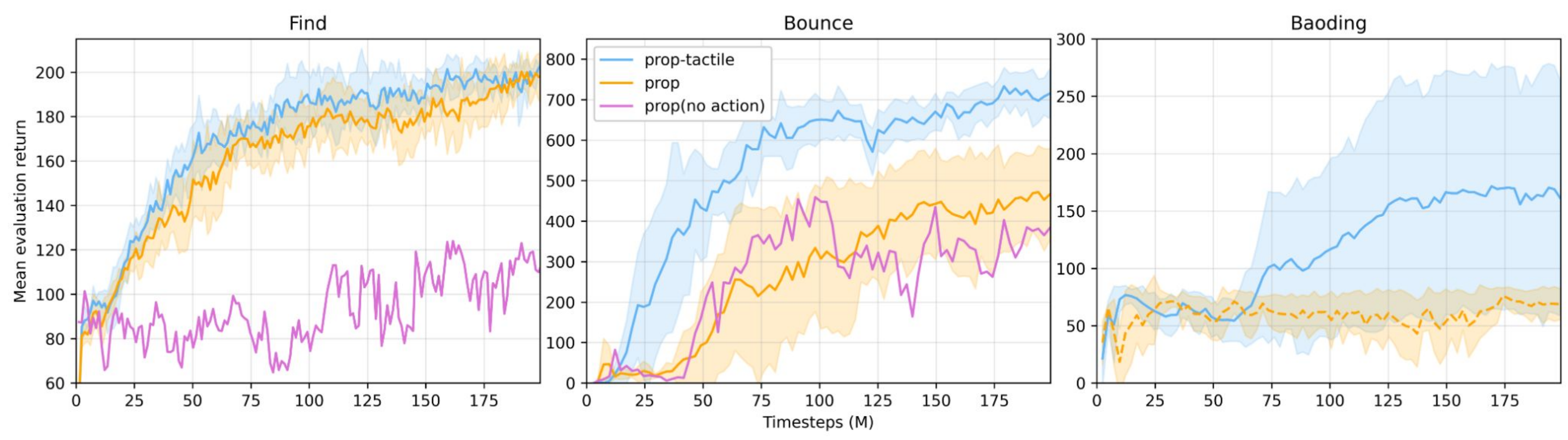
1. **Find** an object
2. **Bounce** a ball without dropping
3. **Baoding** ball rotation

Finding 1. Sparse binary contacts + proprioception sufficient for superhuman dexterity

Following the intuition that truly blind dexterity should be possible, we studied a setup with no vision or privileged information.

With our methodology we achieve super-human dexterity with only a history of proprioception and 17 binary tactile activations as observations.



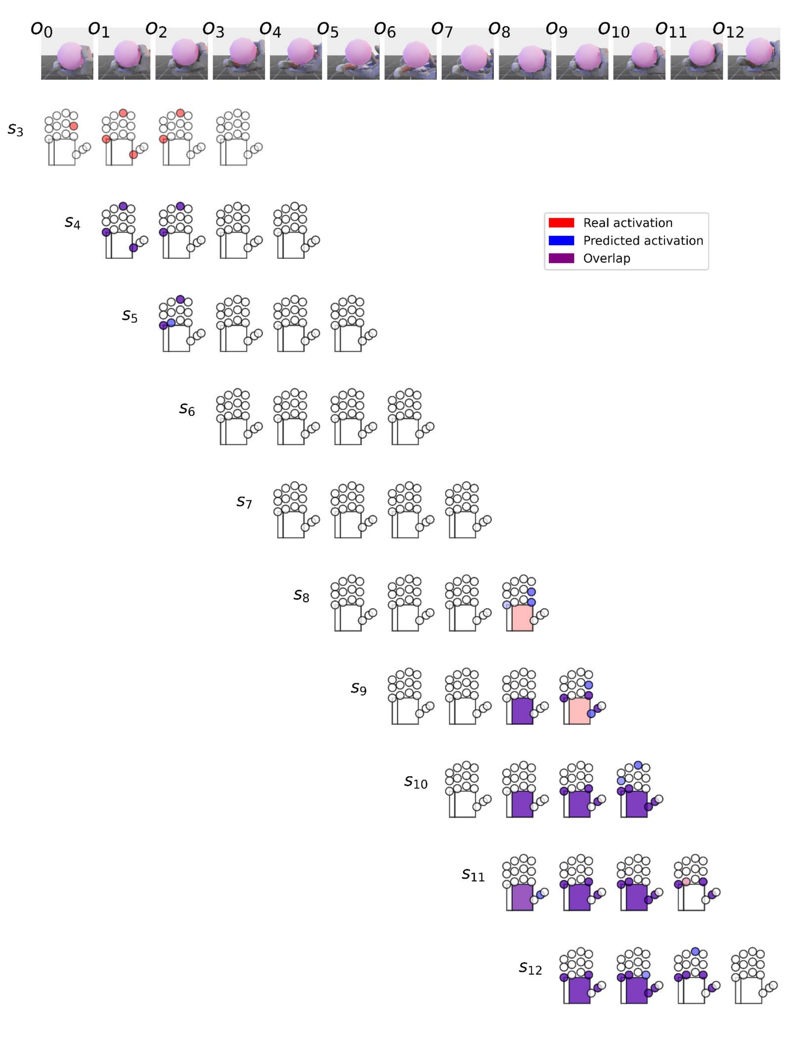


Finding 2. Self-supervision greatly improves performance of tactile-based agents

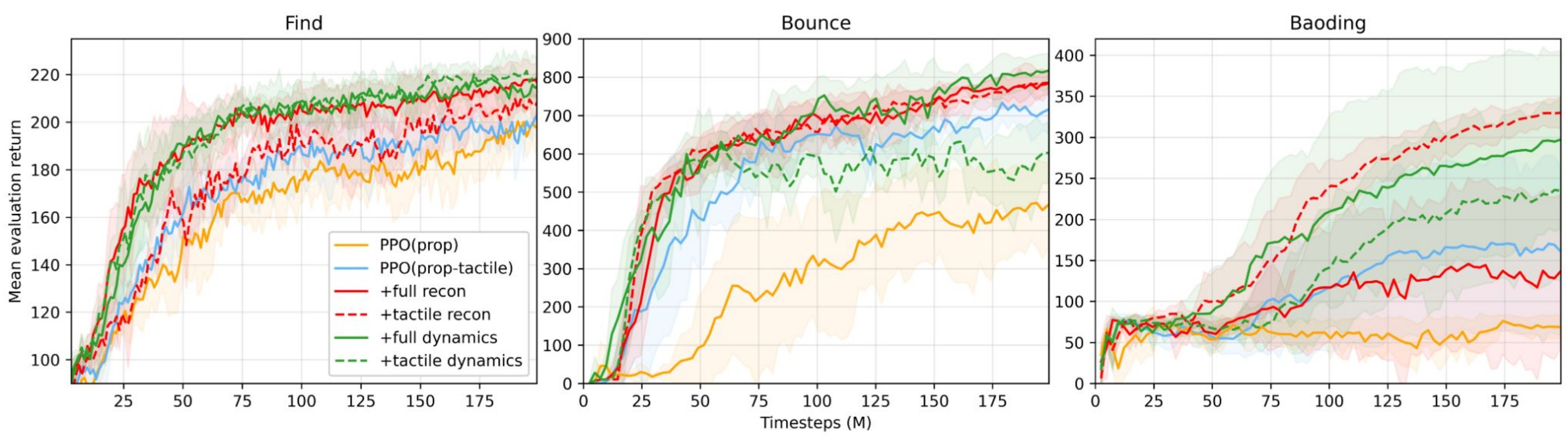
Compared to RL-only agents the best SSL agents on average find an object 36% faster (1.4 vs 1.9 seconds), bounce a ball 8 more times in 10 seconds (79 vs 71), and complete 17 Baoding rotations compared to 5 in 10 seconds.

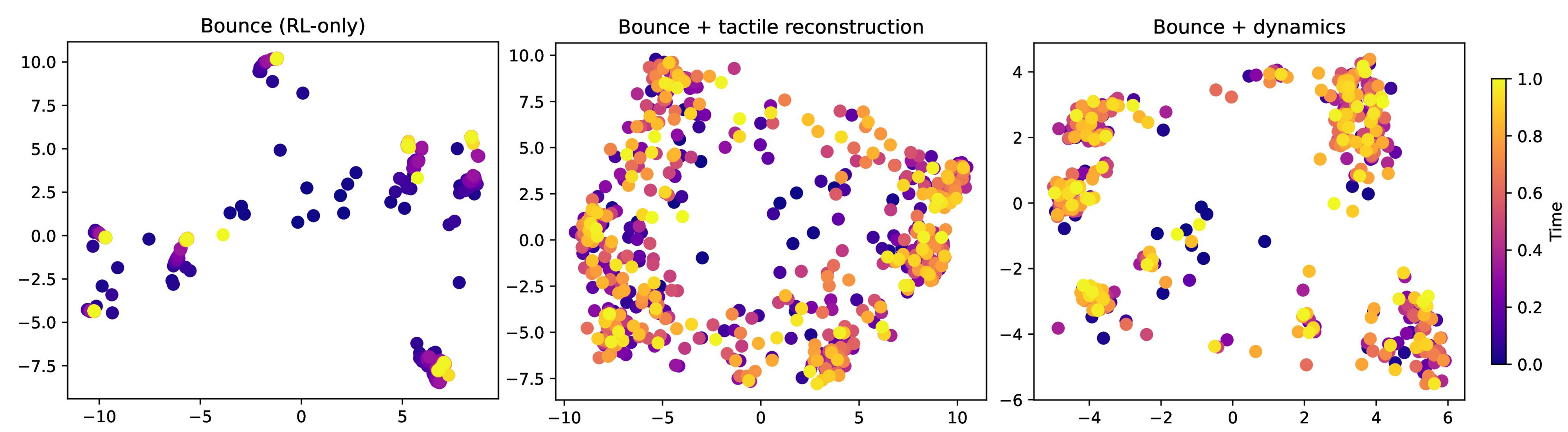
Dynamics-based objectives had the strongest performance, and led to agents that could predict states multiple steps into the future with high precision (e.g. when and where the ball would land)

We also find that **on-policy SSL agents can benefit from off-policy data**, by separating the auxiliary memory from the RL rollout memory and increasing its size.



Learned forward model could predict future contact states to high precision





PCA of state representation z for one episode for Bounce agents