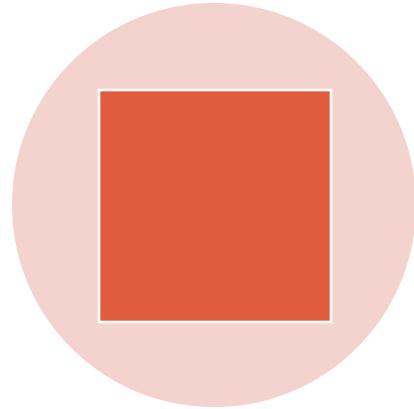
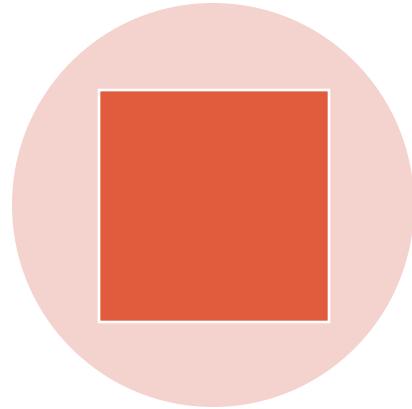


DayDreamer: World Models for Physical Robot Learning, Wu et al. (2022)



GÖRKEM BASAR, SINA
KHUDIDA, JOSEPHINE STÖHR



SEMINAR: MACHINE
LEARNING IN ROBOTICS



PROF. DR. VERENA HAFNER

Topic

Learn online on the real robot (no simulator, no demos)

Replay buffer: store (observation, action, reward) from real runs

World model (predictor): encode obs → latent; predict next latent + reward

Imagined rollouts: practice inside the model's latent space

Actor–Critic: actor = chooses actions; critic = scores future reward (λ -returns)

Deploy policy → collect more data → repeat

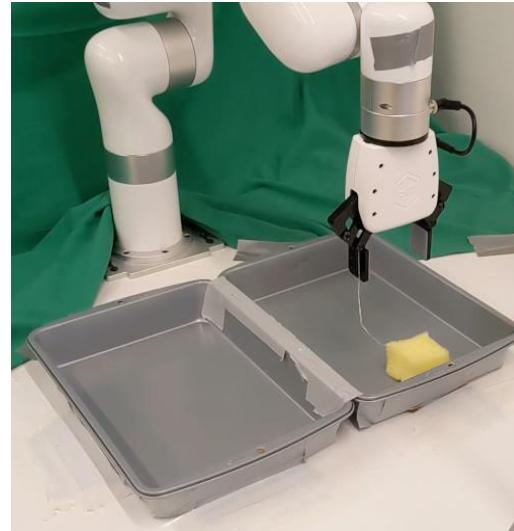
Experiments in the paper



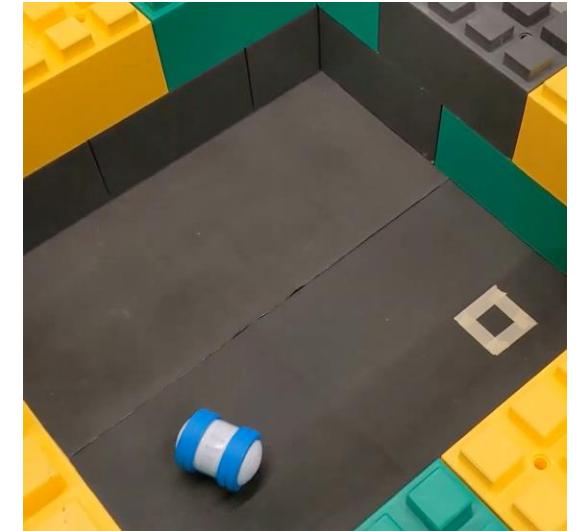
A1: quadruped Walking



UR5: Visual Pick and Place



XArm: Visual Pick and Place



Sphero: navigation

[1]

Stress-Testing Perception: The Ambiguity Gap

Side View: "Fully Observable"



Top-Down View: "Ambiguous"



Mechanism Comparison: Why Dreamer Wins on Ambiguity

Feature	DrQ-v2 (Model-Free)	Dreamer (Model-Based)
Input Data	Stack of 3 Images (Frame Stacking)	Single Image + Internal Memory
Motion Detection	Infers from pixel difference in stack	Infers from Memory + Action History
Action Awareness	Implicit (Actions are outputs only)	Explicit (Previous Action is an Input to the next state)
Top-Down Failure	Fails because Z-axis motion is invisible in pixel stack	Succeeds because it "remembers" the action caused the motion

Planned Experiments

Experiment	Algorithm	Perspective	Expected outcome
1	DrQ-v2 (Model-Free)	Side view	Success (Easy mapping)
2	Dreamer (Model-Based)	Side view	Success (Easy mapping)
3	DrQ-v2 (Model-Free)	Top-Down	Failure (Can't see height and joints)
4	Dreamer (Model-Based)	Top-Down	Success (Infers height and joints)

References

- [1] P. Wu, A. Escontrela, D. Hafner, K. Goldberg, and P. Abbeel, “DayDreamer: World Models for Physical Robot Learning,” *arXiv.org*, 2022. <https://arxiv.org/abs/2206.14176v1> (accessed Dec. 16, 2025).