Make the Pertinent Salient: Task-Relevant Reconstruction for Visual Control with Distraction

Kyungmin Kim, Charless Fowlkes, and Roy Fox
University of California, Irvine

1. Visual Control with Distraction

- **Visual control task**: Control actions based on visual information.
  - ➢ e.g. DeepMind Control suite (DMC)
- Add **distractions** for a more **challenging** and **realistic** setup.

2. Model-Based Reinforcement Learning (MBRL)

- **Cooperation** between a **world model** and **behavior learning**.
- Promising with **great sample efficiency** in visual control tasks.
- Often **struggles** in **distracting** environments.

Representation learning	Examples	Drawbacks
Reconstruction-based	Dreamer [1], etc.	Unnecessary information included
Reconstruction-free	TD-MPC [2], DreamerPro [3], etc.	Sample inefficient

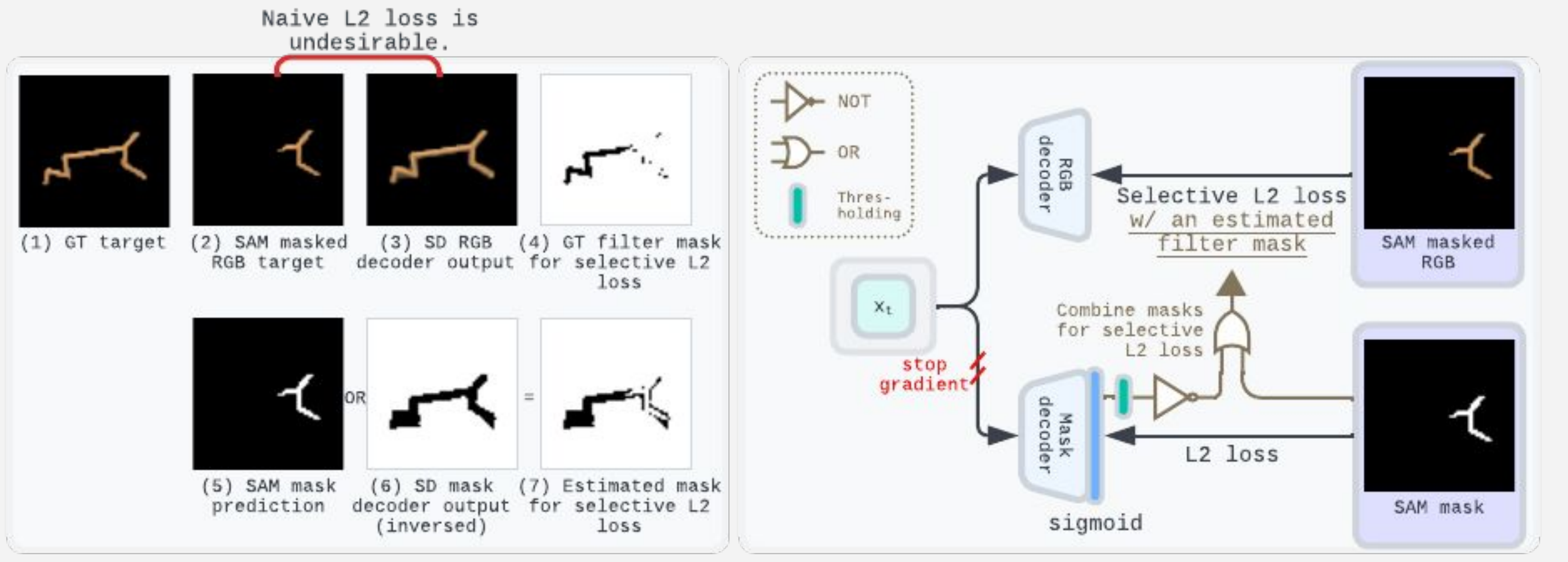
3. Method

- **Assumption**: Task-relevant components are **straightforward to identify** within images, given available prior knowledge.
- Use **prior knowledge** with **segmentation foundation models**.
- **SD**: Reconstruct **task-relevant** components **only**.
- $\mathbf{SD_{GT}}$: When **GT mask** for task-relevant components are **available**.
- $\mathbf{SD_{approx.}}$: Use **PerSAM** [4] fine-tuned with **a single data point**.

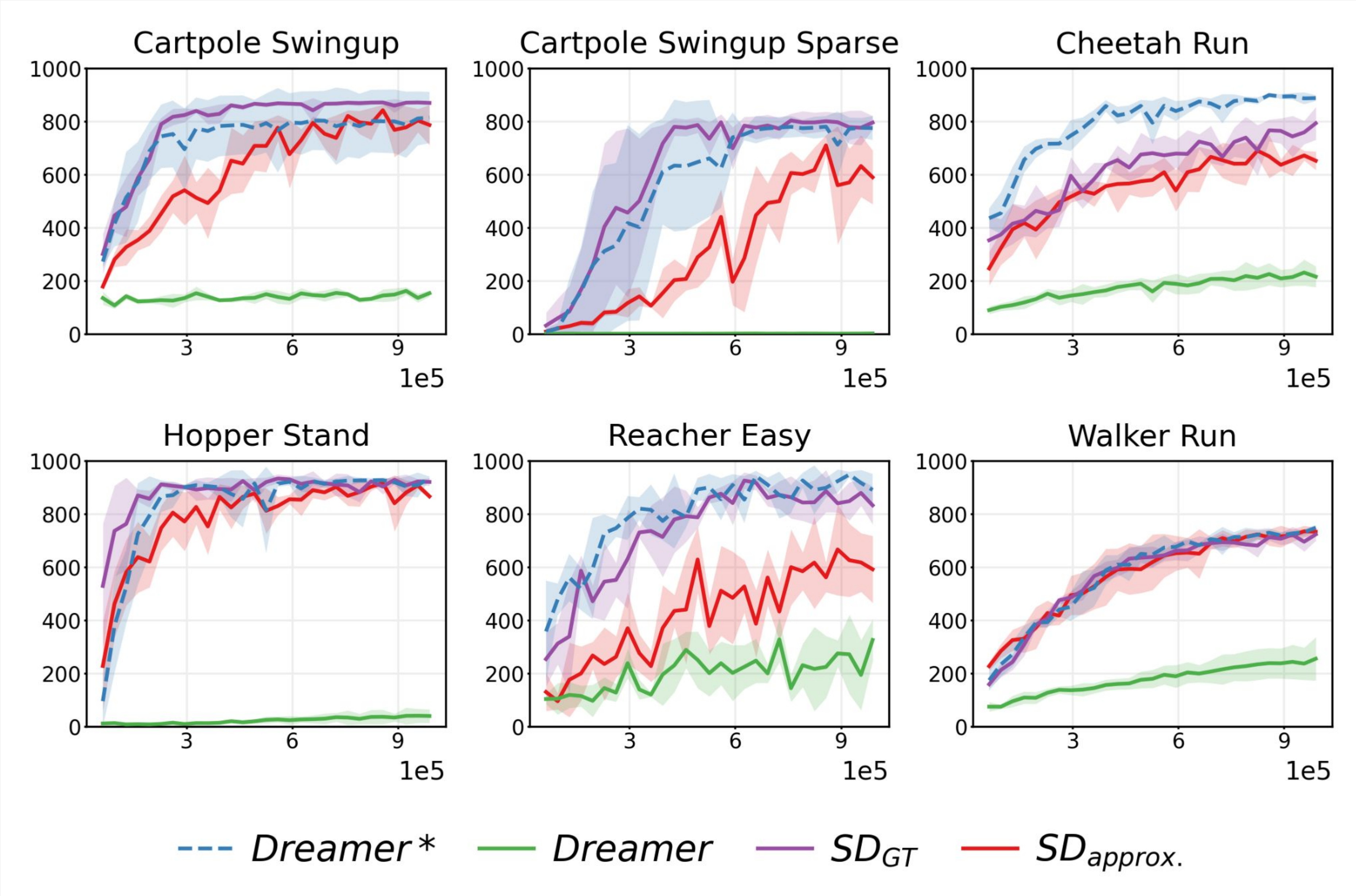
(a) Dreamer target (b) $\mathrm{SD_{GT}}$ target (c) $\mathrm{SD_{approx.}}$ target

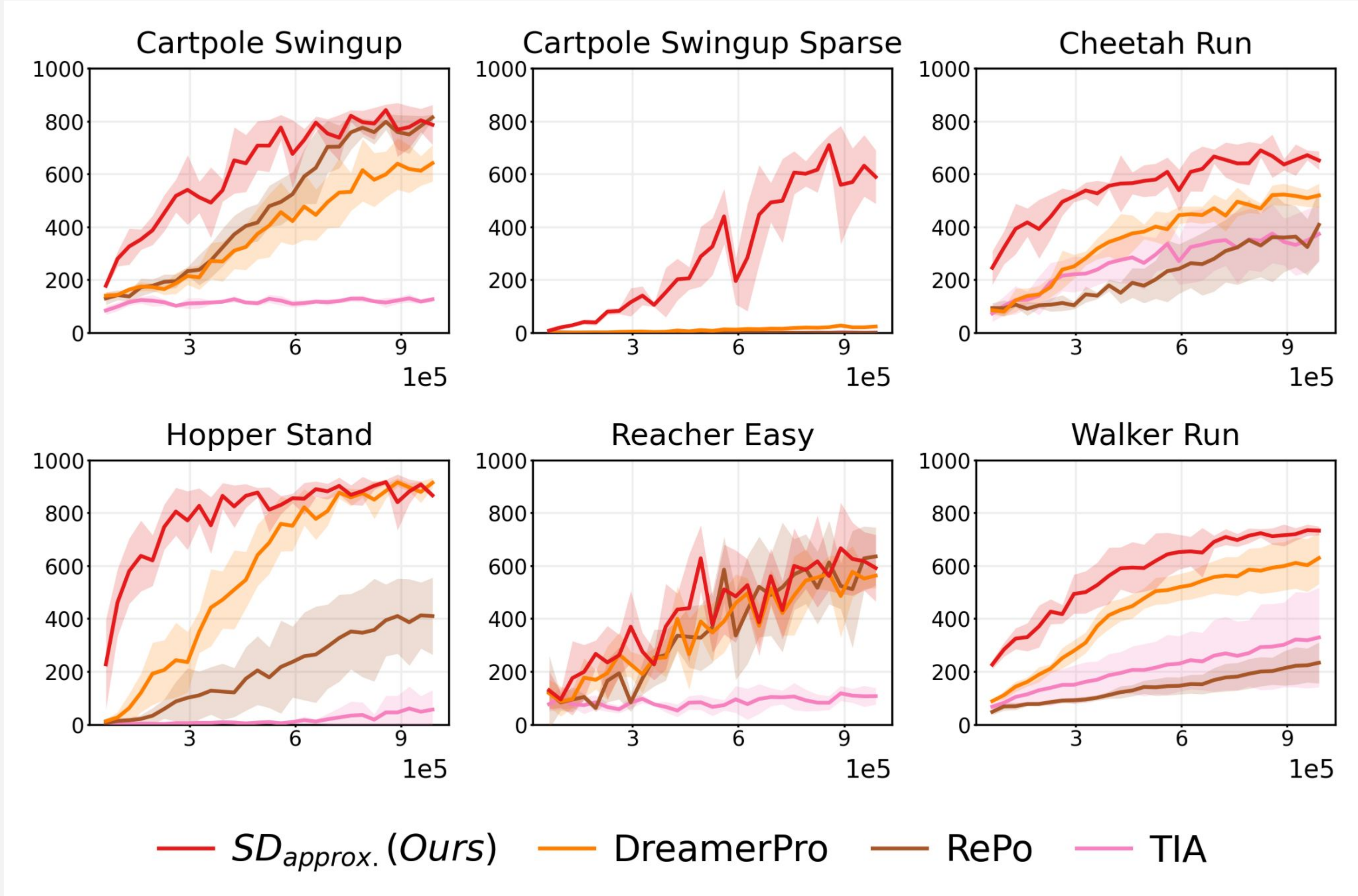
- To make $\mathbf{SD_{approx.}}$ more **robust** to **noisy targets**, we devise **selective $\mathbf{L_2}$ loss**.
- **Identify** pixels where **predicted labels** may be **wrong** but the world model **decoder** is **correct**, **ignoring $\mathbf{L_2}$ loss** for such pixels to avoid providing wrong signals.

$$\mathrm{mask}_{estimate} = \mathrm{mask}_{\mathrm{SAM}} \vee \neg\mathrm{mask}_{\mathrm{SD}}$$



4. Experimental Results





- **Dreamer*** is trained in a **clean** environment whereas the other methods are trained in a **distracting** environment.
- Overall, $\mathrm{SD_{GT}}$ matches Dreamer*, and $\mathrm{SD_{approx.}}$ eventually reaches $\mathrm{SD_{GT}}$ while Dreamer falls short.
- **Reconstruction-free methods** take **lots of samples** to converge.
- In **sparse reward** scenario, **only our method** achieves compelling performance.

5. References

[1] Hafner et al. Mastering diverse domains through world models. arXiv preprint, 2023.
[2] Hansen et al. Td-mpc2: Scalable, robust world models for continuous control. ICLR, 2024.
[3] Deng et al. Dreamerpro: Reconstruction-free model-based reinforcement learning with prototypical representations. ICML, 2022.
[4] Zhang et al. Personalize segment anything model with one shot. arXiv preprint, 2023.