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

Kyungmin Kim, JB Lanier, 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 RL

- **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.	Irrelevant information included
Reconstruction-free	TD-MPC [2], DreamerPro [3], etc.	Sample inefficient

- Our method, **SD**, fixes this with **segmentation-guided** reconstruction.

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.

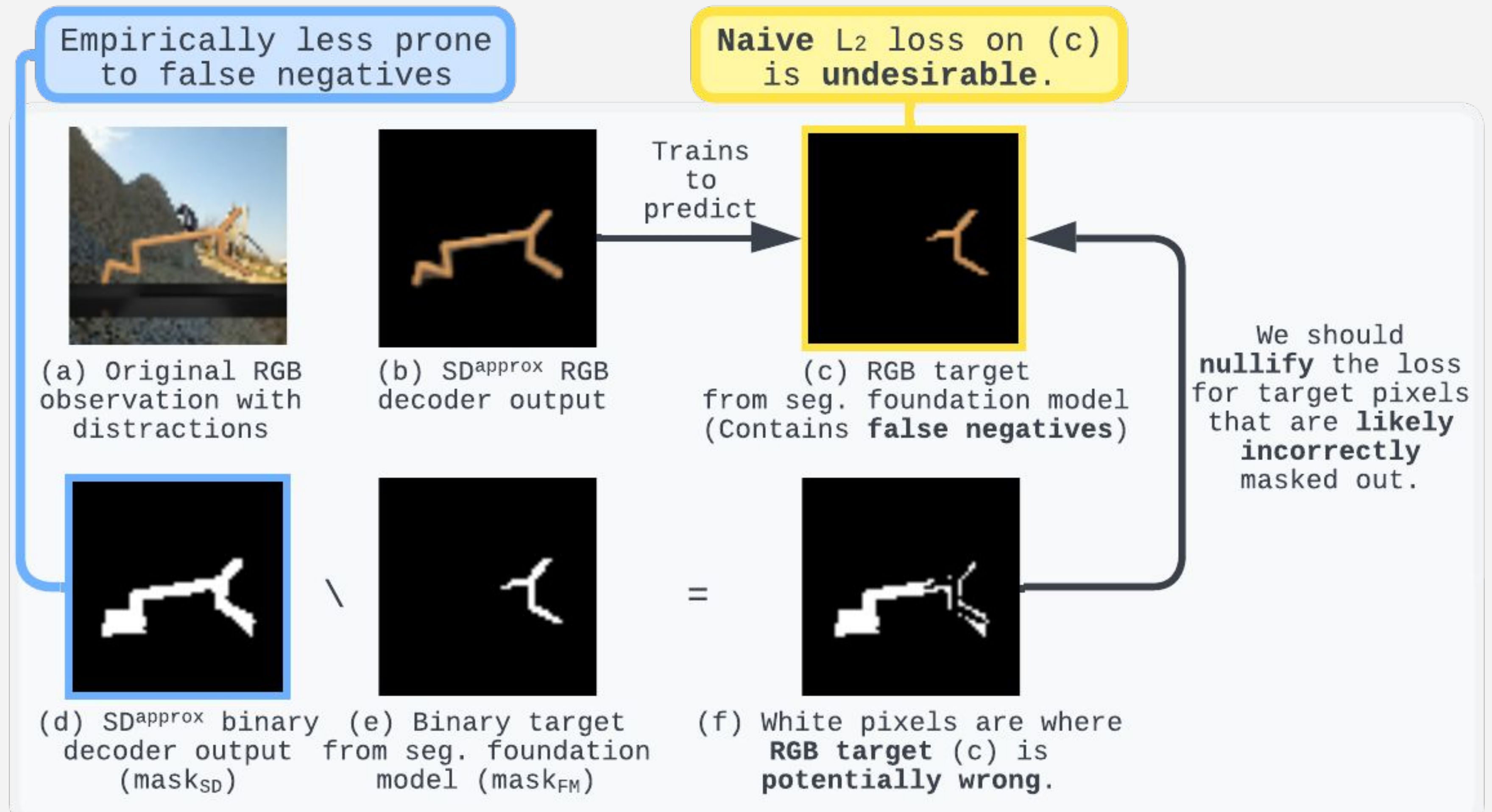
3. Method

- **Assumption**: Task-relevant components in the image are **easily identifiable** using available prior knowledge.
- Use the **prior knowledge** of pre-trained **segmentation foundation models**.
- **SD**: Reconstruct **only task-relevant** components.
- **SD^{GT}**: Uses **ground-truth masks** for task-relevant components when available (e.g., in simulation).
- **SD$^{approx.}$**: Uses **a segmentation model** fine-tuned with as few as one annotated example.

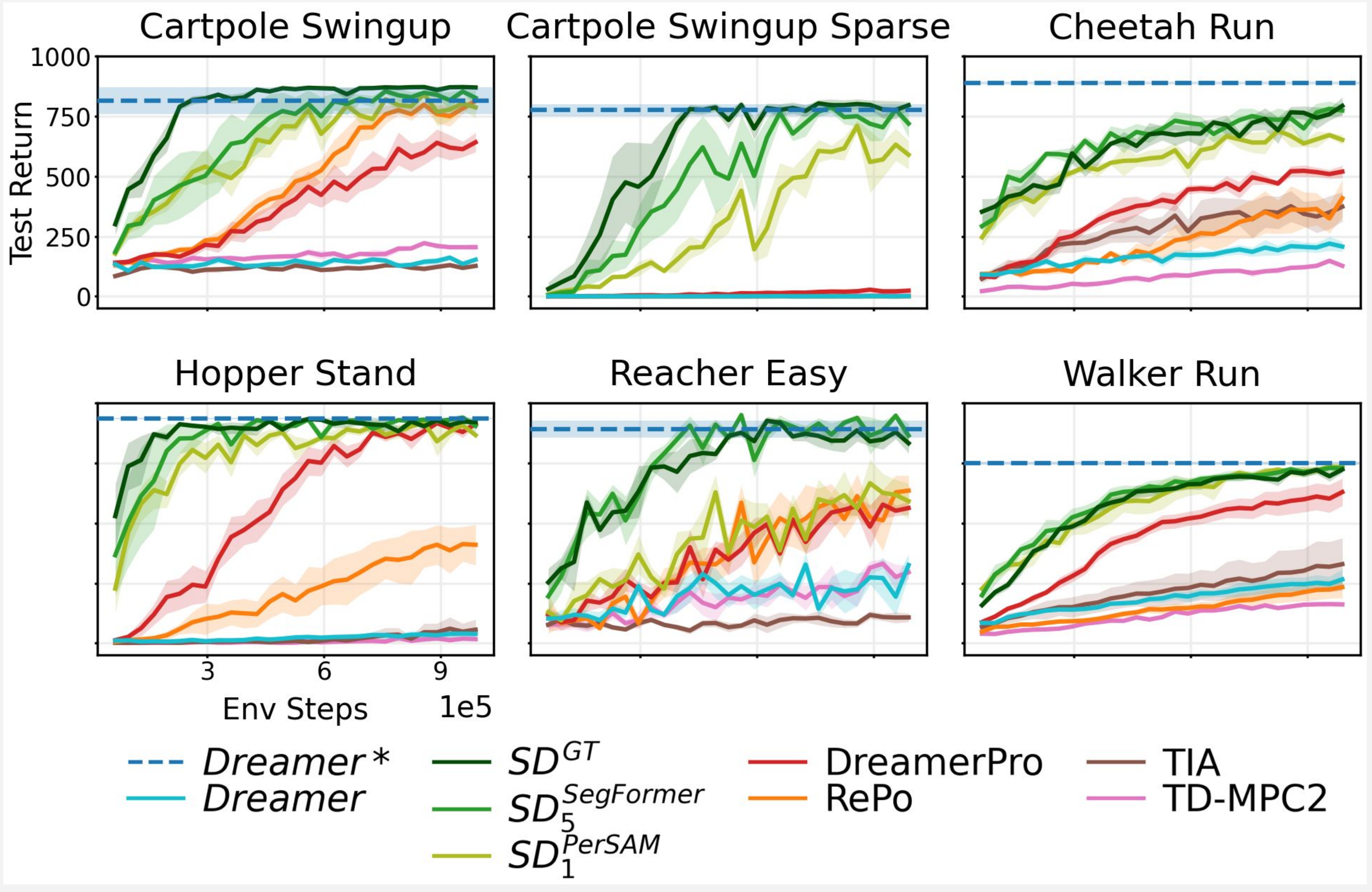
(a) Dreamer target (b) SD^{GT} target (c) SD$^{approx.}$ target

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

$$\mathrm{pixel}_{\mathrm{nullify}} = \mathrm{pixel}_{\mathrm{SD}} \setminus \mathrm{pixel}_{\mathrm{FM}}$$



4. Experimental Results



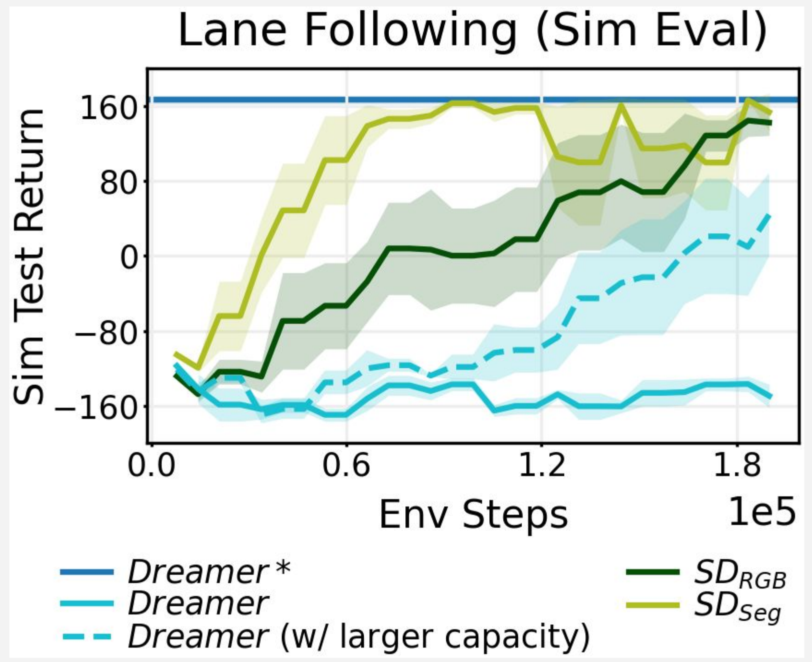
- SD^{GT} matches Dreamer*; also, SD$^{approx.}$ eventually reaches SD^{GT} while Dreamer falls short.
- **Reconstruction-free methods** take **lots of samples** to converge.

Sim-to-Real Experiments on DuckieTown (Real-World)

(a) **Training**-time observations

(b) **Test**-time observations (**sim**)

(c) **Test**-time observations (**real**)



(d) **Sim** test-time performance

Method	Real-world Return
DREAMER*	-172.2 ± 14.4
DREAMER	-119.7 ± 10.7
DREAMER (large)	3.9 ± 23.1
SD$_{RGB}$	106.2 ± 4.4
SD$_{Seg}$	**116.2 ± 5.1**

(e) **Real-world** evaluation

- SD variants outperform Dreamer, highlighting their effectiveness for **sim-to-real transfer** by **reducing variance** during training.