ReCoRe
Regularized Contrastive Representation Learning of World Model




Rudra P.K. Poudel Harit Pandya Stephan Liwicki Roberto Cipolla

Cambridge Research Laboratory
Toshiba Europe Ltd.

@RudraPoudel

TOSHIBA UNIVERSITY OF CAMBRIDGE

World Model

Learns the environment dynamic in a (self-)supervised manner, which

- **enables planning in the imaginations**
- **improves the sample efficiency**

Current Problems of World Models

Works well on in-distribution tests

but

poor out-of-distribution (OoD) generalization!

iGibson OoD Success %	**100k**	**500K**
RAD	0.5	36.3
CURL	4.7	31.4
DreamerV2	1.3	1.5
Masked WM	1.7	2.5

Image-based Control: visual point-goal navigation task.

Inspiration

Works well on in-distribution tests

but

poor out-of-distribution (OoD) generalization!

iGibson OoD Success %	**100k**	**500K**
RAD	0.5	36.3
CURL	4.7	31.4
DreamerV2	1.3	1.5
Masked WM	1.7	2.5
DreamerV2 + Grounding DINO	**37.1**	**50.0**

Image-based Control: visual point-goal navigation task.

ReCoRe

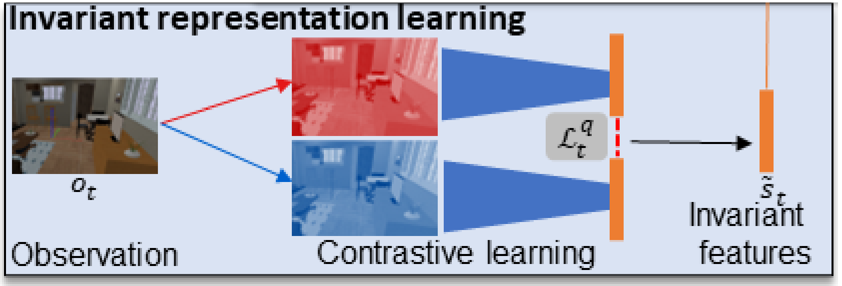
Can we learn generalized representation with

- less training data?
- smaller model?
- efficient adaptation to downstream tasks?

Ideas,

- **World model:** separates representation and policy learning
- **Contrastive learning:** invariant global context representations
- **Regularization:** (causality) intervention invariant auxiliary task to explicitly enforce the invariance and preserve the local details

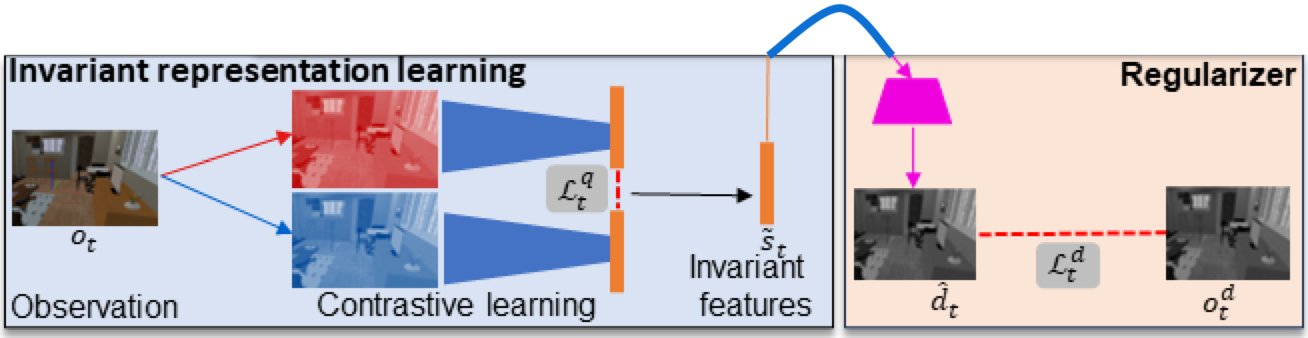
Contrastive Learning: Improves Invariant Global Context Representations



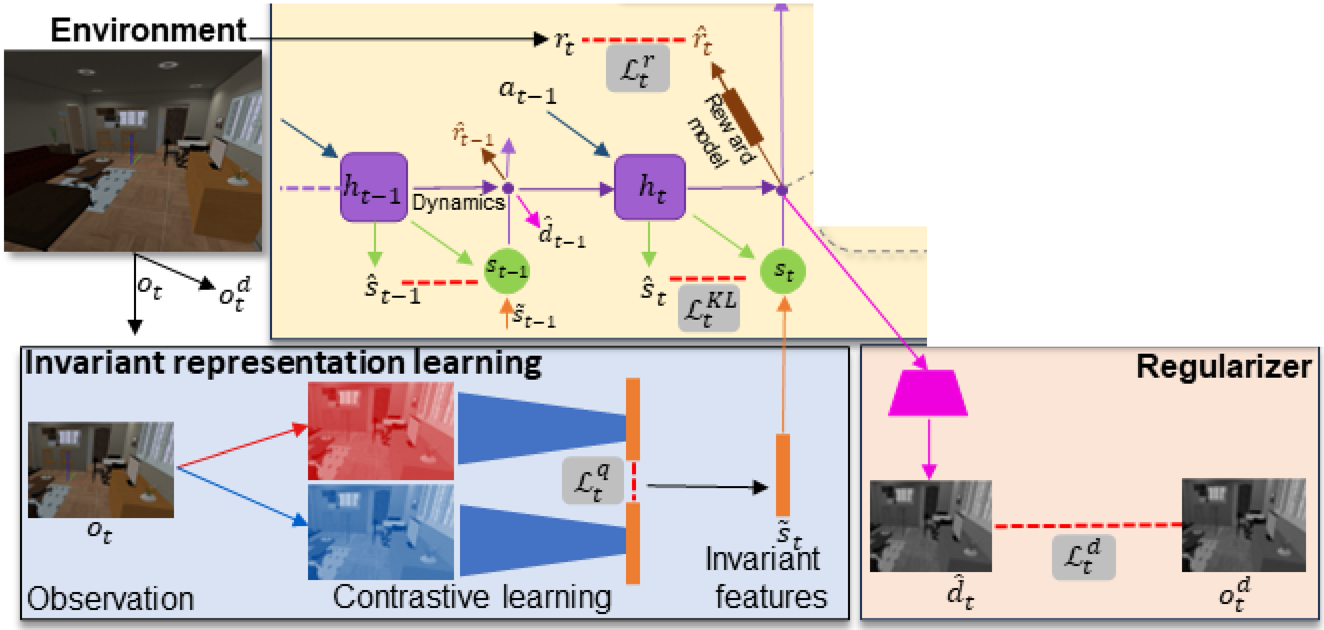
Intervention Invariant Auxiliary Tasks:
Explicitly Forcing the Invariance and Preserving Details

1. Image denoising
2. Depth reconstruction
3. Segmentation mask
4. Object detections
5. etc.

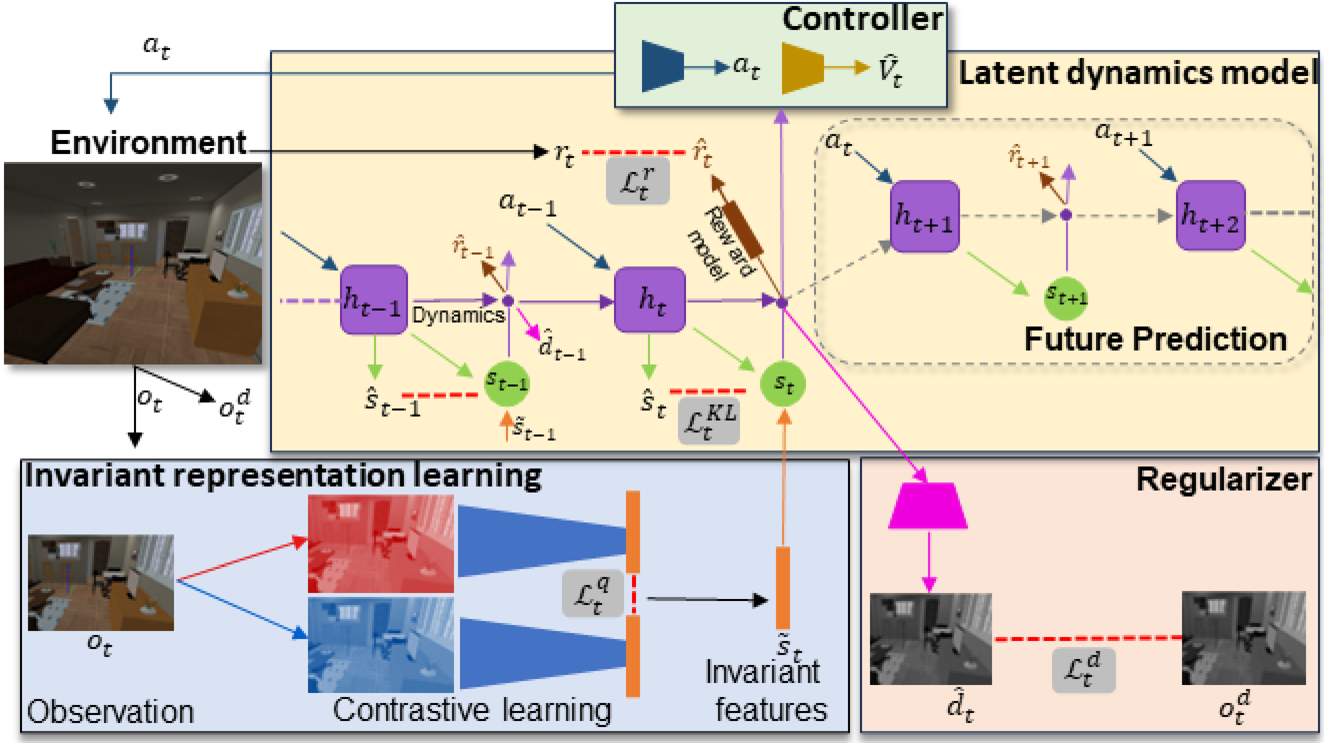
Auxiliary Task(s)



Learning World Model Independent of the Policy, Better Generalization on Downstream Tasks



Learning the Policy in the Imaginations, Improves Sample Efficiency



Out-of-Distribution Results

ReCoRe matches SoTA Grounding DINO on 100k, outperforms on 500k.

iGibson OoD Success %	100k	500K
RAD	0.5	36.3
CURL	4.7	31.4
DreamerV2	1.3	1.5
Masked WM	1.7	2.5
DreamerV2 + Grounding DINO	**37.1**	50.0
ReCoRe	36.0	**59.7**

Image-based Control: visual point-goal navigation task.

Sim-to-Real Results

ReCoRe outperforms Grounding DINO.

iGibson-to-Gibson Success %	**100k**	**500K**
RAD	0.0	30.2
CURL	6.7	36.7
DreamerV2 + Grounding DINO	40.8	60.3
ReCoRe	**41.6**	**71.9**

Image-based Control: visual point-goal navigation task.

In-Distribution Results with DMC Control

ReCoRe also works with **image-denoising as an auxiliary task**.
Better intervention invariant auxiliary tasks, better results.

100k Steps Total Rewards	**ReCoRe + Image**	**ReCoRe + Seg**	**CURL**	**Dreamer**
Finger, spin	486	474	**767**	341
Cartpole, swingup	472	449	**582**	326
Reacher, easy	327	**982**	538	314
Cheetah, run	321	**400**	299	235
Walker, walk	654	**739**	403	277
Ball in cup, catch	830	**859**	769	246

Similar results hold for 500k.

Ablation of ReCoRe

Ablation of proposed hypothesis,

- **Contrastive learning:** good for global context representations
- **Regularization:** (causality) intervention invariant auxiliary task to explicitly enforce the invariance and preserve the local details

iGibson OoD Success %	**500K**
ReCoRe	59.7
ReCoRe - CL (Contrastive Learning)	5.0
ReCoRe - CL (Contrastive Learning) + DA (Data Augmentation)	22.1
ReCoRe - D (invariant aux. task)	0.8
ReCoRe - D (invariant auxiliary task) + I (RGB reconstruction)	19.2

Image-based Control: visual point-goal navigation task.

ReCoRe
Regularized Contrastive Representation Learning of World Model

- Intervention invariant auxiliary tasks,
  - improves OoD and sim-to-real generalization of world model
  - stronger invariance auxiliary tasks, improves results

Project Website and Code