Unsupervised state representation learning with robotic priors: a robustness benchmark

Timothée Lesort, Mathieu Seurin, Xinrui Li, Natalia Díaz-Rodríguez* and David Filliat,
U2IS, ENSTA ParisTech, Inria FLOWERS team, Université Paris Saclay, Palaiseau, France
{timothee.lesort, mathieu.seurin, xinrui.li, natalia.diaz,
david.filliat}@ensta-paristech.fr

(1) First two authors contributed equally

Motivation

Our goal is to learn a relevant state from images, actions and rewards. We train a neural network in unsupervised manner using prior knowledge about the physical world in form of robotic priors.

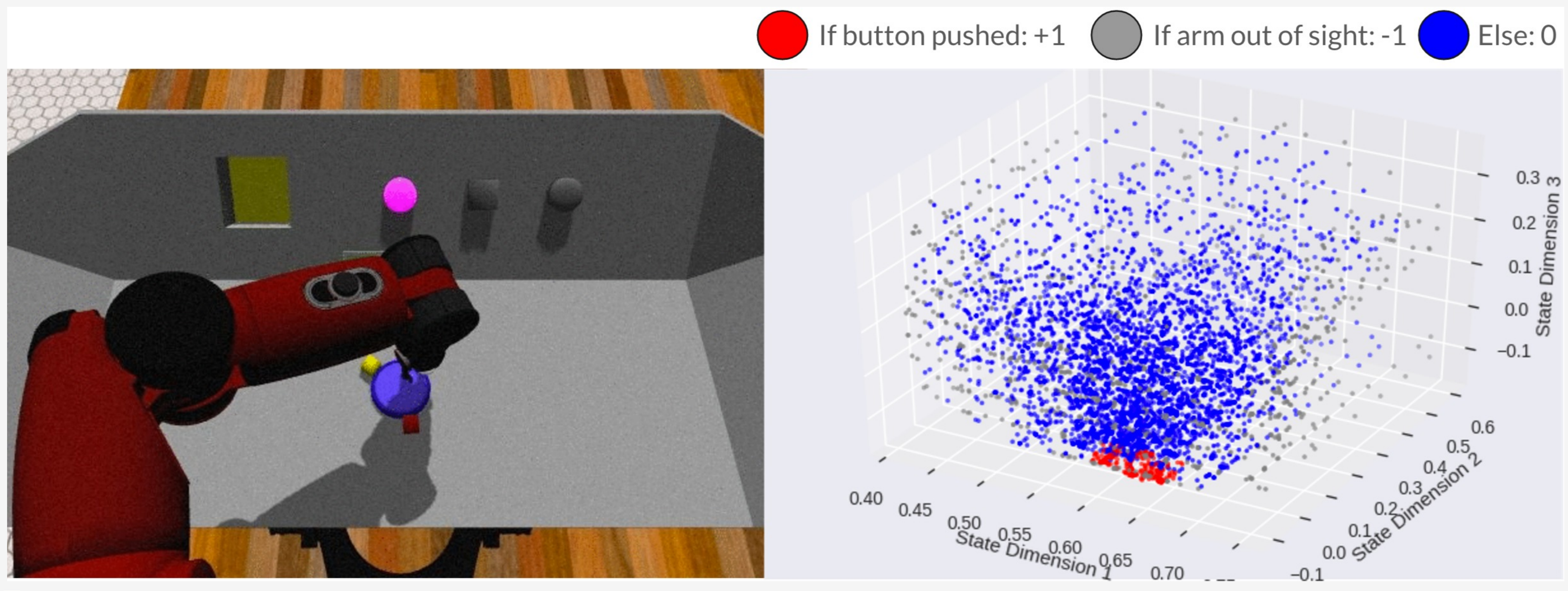

Figure 1: Left: Baxter's camera view for Static-Button-Distractors dataset 2. Right: Baxter's left hand ground truth position and its coded reward

Method: Robotic Priors

Robotic Priors [1] provide the model with basic knowledge about the environment dynamical features. Each prior is encoded as a loss function:

- **Temporal coherence Prior:** *Two states close to each other in time are also close to each other in the state representation space.*

$$L_{Temp}(D, \hat{\phi}) = \mathrm{E}[\| \Delta \hat{s}_t \|^2] \ , \quad (1)$$

- **Proportionality Prior:** *Two identical actions should result in two proportional magnitude state variations.*

$$L_{Prop}(D, \hat{\phi}) = \mathrm{E}[(\| \Delta \hat{s}_{t_2} \| - \| \Delta \hat{s}_{t_1} \|)^2 | a_{t_1} = a_{t_2}] \ , \quad (2)$$

- **Repeatability Prior:** *Two identical actions applied at similar states should provide similar state variations, not only in magnitude but also in direction.*

$$L_{Rep}(D, \hat{\phi}) = \mathrm{E}[e^{-\|\hat{s}_{t_2} - \hat{s}_{t_1}\|^2} \| \Delta \hat{s}_{t_2} - \Delta \hat{s}_{t_1} \|^2 | \ a_{t_1} = a_{t_2}] \ , \quad (3)$$

- **Causality Prior:** *If two states on which the same action is applied give two different rewards, they should not be close to each other in the state representation space.*

$$L_{Caus}(D, \hat{\phi}) = \mathrm{E}[e^{-\|\hat{s}_{t_2} - \hat{s}_{t_1}\|^2} \ | \ a_{t_1} = a_{t_2}, r_{t_1+1} \neq r_{t_2+1}] \ , \quad (4)$$

- **Reference point Prior:** *Two states corresponding to the same reference point should be close to each other*

$$L_{Ref}(D, \hat{\phi}) = \mathrm{E}[\| \hat{s}_{t_i} - \hat{s}_{t_j} \|^2 | \ s_{t_i} = s_{t_j} = s_{Ref}] \quad (5)$$

where s_{Ref} is the embedded state of a fixed reference point.

Contributed Benchmark Dataset

The goal is to discover a state space allowing Baxter robot to push a button (in which case he gets a positive reward) [2]. We created environments with moving and static distractors and domain randomization.

Figure 2: A sample of each dataset (1-4), created for our benchmark with increasing complexity

Experiments on task: *Baxter pushing button* (4 datasets)

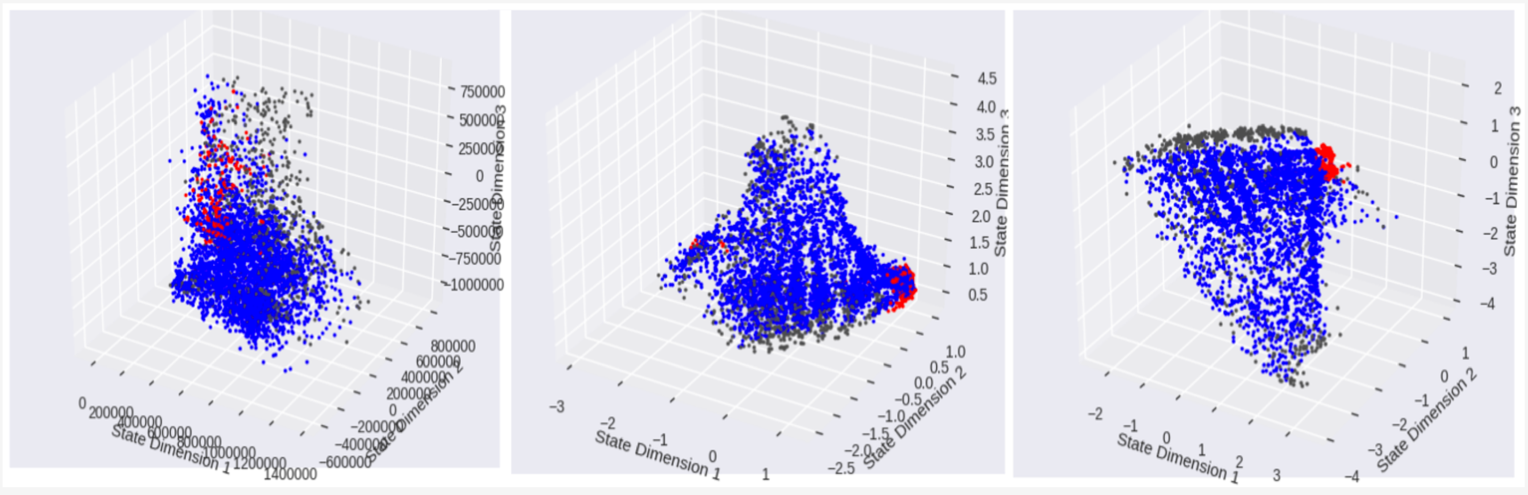

Figure 3: Learned state space on Static-Button-Distractors (dataset 2): Left: Denoising Autoencoder. Middle: 4 Priors. Right: 5 Priors. A red reward (value +1) state means the button is being pushed, gray (-1) if the hand is out of sight, and blue (reward 0) if hand is elsewhere.

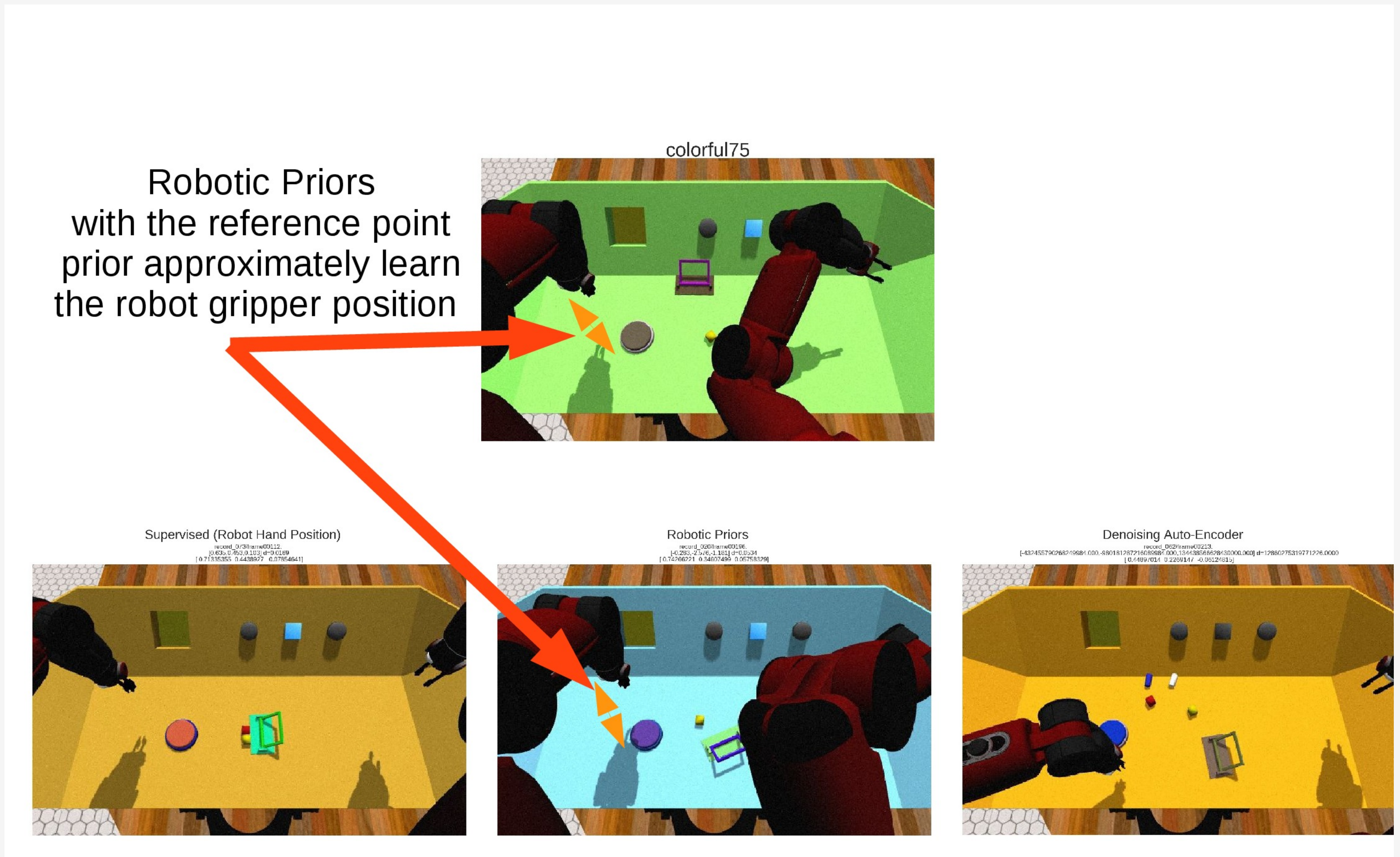

Figure 4: Nearest neighbor evaluation of robotic priors against supervised learning and Autoencoders on (the domain randomization) *Colorful75* dataset using the fixed ref. point prior. The auto-encoder-learnt nearest state's observation is far from the ground truth position.

Criterion	GT	Superv	4 Priors	5 Priors	AE	5 Priors 0f
KNN-MSE	0.024	0.03	0.079	0.053	0.099	0.047
NIEQA local	0	0.239	0.66	0.50	0.599	0.52
NIEQA global	0	0.048	0.41	0.20	0.465	0.21

Table 1: *Static-Button-Distractor 3D* (dataset 2) Results. *xf* means x ResNet frozen layers, GT: (hand position) Ground Truth

Conclusions

Robotic Priors ignore distractors, learn representations more relevant to the task than autoencoders, and provide performance close to the one learnt with supervision. We contributed[a]:

- A dataset benchmark on Baxter robot on the *button pushing* task
- A quantitative and qualitative evaluation metric based on KNN-MSE testing the limits of the robotic priors
- A 5^{th} fixed reference point prior to overcome the vulnerabilities found on the original priors

Future work should focus on transfer learning from simulation to reality.

[a] DREAMProject: www.robotsthatdream.eu

References

[1] R. Jonschkowski and O. Brock.
Learning state representations with robotic priors.
Autonomous Robots, 39(3):407–428, 2015.

[2] T. Lesort, M. Seurin, X. Li, N. Díaz-Rodríguez, and D. Filliat.
Unsupervised state representation learning with robotic priors: a robustness benchmark.
CoRR, abs/1709.05185, 2017.