



# Efficient Focus for Autonomous Agents

15 MAR 2024

Bram Grooten, PhD Candidate

Supervisors: Decebal Mocanu (University of Luxembourg), Mykola Pechenizkiy (TU/e)  
Email: [bj.grooten@tue.nl](mailto:bj.grooten@tue.nl) Twitter/X: @BramGrooten



# Bram Grooten



**PhD** candidate

Started: Nov 2021

**Internship** at Sony AI  
**Research visit** in Alberta



**Master** Applied Mathematics  
**Master** Science Education



**Bachelor** Applied Mathematics (Boston, US)



# Finding Focus in RL

## Automatic Noise Filtering with Dynamic Sparse Training in Deep Reinforcement Learning

Bram Grooten  
Eindhoven University of Technology  
b.j.grooten@tue.nl

Glada Sokar  
Eindhoven University of Technology  
g.a.z.n.sokar@tue.nl

Shubhansh Dohare  
University of Alberta  
sdohare@ualberta.ca

Elena Mocanu  
University of Twente  
e.mocanu@utwente.nl

Matthew Taylor  
University of Alberta  
matthew.taylor@ualberta.ca

Mykola Pechenizkiy  
Eindhoven University of Technology  
m.pechenizkiy@tue.nl

Decebal Constantin Mocanu  
University of Twente  
d.c.mocanu@utwente.nl

### ABSTRACT

Tomorrow's robots will need to distinguish useful information from noise when performing different tasks. A household robot for instance may continuously receive a plethora of information about the house, but needs to focus on just a small subset to successfully execute its current chore.

Filtering distracting inputs that contain irrelevant data has received little attention in the reinforcement learning literature. To meet learning in the extremely noisy environment (ENE) where up to 90% of the input features are pure noise. Agents need to detect which features actually provide task-relevant information about the state of the environment.

Consequently, we propose a new method termed Automatic Noise Filtering (ANF) which uses the principles of dynamic sparse training in synergy with various deep reinforcement learning algorithms. The sparse input layer learns to focus its connectivity on task-relevant features, such that ANF-SAC and ANF-TD outperform standard SAC and TD by a large margin, while using up to 95% fewer weights.

Furthermore, we devise a transfer learning setting for ENEs, by permitting all features of the environment after IM iterates relevant to the world evolves. Again, ANF surpasses the baselines in final performance and sample complexity. Our code is in the supplementary material and will be put online at a later stage.

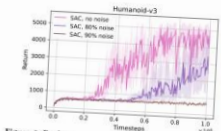
### KEYWORDS

deep reinforcement learning, noise filtering, sparse training

### ACM Reference Format

Bram Grooten, Glada Sokar, Shubhansh Dohare, Elena Mocanu, Matthew Taylor, Mykola Pechenizkiy, and Decebal Constantin Mocanu. 2023. Automatic Noise Filtering with Dynamic Sparse Training in Deep Reinforcement Learning. In *Proc. of the 29th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '23)*, London, United Kingdom, May 29– June 2, 2023. IFAAMAS, 12 pages.

*Proc. of the 29th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '23)*, London, United Kingdom, May 29– June 2, 2023. Copyright © 2023, IFAAMAS, an imprint of International Foundation for Autonomous Agents and Multiagent Systems. All rights reserved.



**Figure 3: Performance of SAC on Humanoid-v3 environments expanded with a different number of pure noise features. Once the environment contains too much noise, SAC struggles to learn a decent policy. On this problem standard dense networks cannot filter through the noise well enough.**

### 1 INTRODUCTION

Future robots will likely perceive a plethora of information about the state of the world, but only parts of it are going to be relevant to their current task. For example, imagine a household robot that house. For its current task, e.g. taking groceries, only a small subset of these information sources, or features, are relevant. We want agents to automatically detect which features are task-relevant, a learning aid distinguishing between noise and lateral noise, a magical robot receiving all possible information about the patient, or a self-driving car that needs to ignore distracting billboards.

We simulate the noisy real-world environment by adding synthetic noise features to an existing state space. This allows us to understand where we stand and what can be done. To illustrate the current situation, Soft Actor-Critic (SAC) [21] fails to learn a decent policy on an environment with 90% added noise features, see Figure 1. We need to invent methods that can effectively filter

The example: Conditions of focus, location, neighborhood, kitchen stimuli, CO, CO levels and temperature in each room up to the last block of 1000 states and final frames in the help and/or basement, novel, neighborhood, and health of all individuals etc.

## MaDi: Learning to Mask Distractions for Generalization in Visual Deep Reinforcement Learning

Bram Grooten<sup>1</sup>, Tristan Tomilin<sup>1</sup>, Gautham Vasan<sup>2</sup>, Matthew E. Taylor<sup>3,4</sup>,  
A. Ripan Mahamood<sup>5</sup>, Meng Fang<sup>1</sup>, Mykola Pechenizkiy<sup>1</sup>, Decebal Constantin Mocanu<sup>6,1</sup>  
<sup>1</sup>Eindhoven University of Technology <sup>2</sup>University of Alberta <sup>3</sup>Alberta Machine Intelligence Institute (Amii)  
<sup>4</sup>University of Liverpool <sup>5</sup>University of Liverpool <sup>6</sup>University of Luxembourg

### ABSTRACT

The visual world provides an abundance of information, but many input pixels received by agents often contain distracting stimuli. Autonomous agents need the ability to distinguish useful information from task-irrelevant perceptions, enabling them to generalize to unseen environments with new distractions. Existing work approaches this problem using data augmentation or large auxiliary networks with additional loss functions. We introduce MaDi, a signal only. In MaDi, the conventional actor-critic structure of the reward reinforcement learning agents is complemented by a small third mask to determine what the actor and critic will receive, such that they can focus on learning the task. The mask are created dynamically, depending on the current input. We run experiments on the DeepMind Control Generalization Benchmark, the Distracting Cart-pole's focus with useful masks, while its efficient Masker network only adds 0.2% more parameters to the original structure, in one results better than or competitive to state-of-the-art methods.<sup>1</sup>

### KEYWORDS

Deep Reinforcement Learning, Generalization, Robotics

### ACM Reference Format

Bram Grooten<sup>1</sup>, Tristan Tomilin<sup>1</sup>, Gautham Vasan<sup>2</sup>, Matthew E. Taylor<sup>3,4</sup>, A. Ripan Mahamood<sup>5</sup>, Meng Fang<sup>1</sup>, Mykola Pechenizkiy<sup>1</sup>, Decebal Constantin Mocanu<sup>6,1</sup>. 2024. MaDi: Learning to Mask Distractions for Generalization in Visual Deep Reinforcement Learning. In *Proc. of the 29th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '23)*, Auckland, New Zealand, May 29– June 2, 2023. IFAAMAS, 12 pages.

<sup>1</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

*Proc. of the 29th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '23)*, Auckland, New Zealand, May 29– June 2, 2023. Copyright © 2023, IFAAMAS, an imprint of International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). This is a licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

<sup>2</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>3</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>4</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>5</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>6</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>7</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>8</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>9</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>10</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>11</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>12</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>13</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>14</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>15</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>16</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>17</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>18</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>19</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

<sup>20</sup>Code and videos are in the Supplementary Material and will be available online. Corresponding author: b.j.grooten@tue.nl

degradation in the performance of deep RL agents, thereby hindering their applicability in the real world. To address this, we propose a novel algorithm, Masking Distractions, which learns to filter out the key ideas behind MaDi is its generalization capabilities: actor-critic architecture with a small lightweight component, the Masker (see Figure 1). This small neural network generates a mask that dims the irrelevant pixels, allowing the actor and critic to focus on learning the task at hand without printing too distracted. Unlike previous approaches that have attempted to address this while introducing minimal overhead in terms of model parameters, thus preserving the efficiency of the original architecture.

Furthermore, no additional loss function is necessary for the Masker to optimize its parameters. To ensure that the Masker maintains the viability of the task-relevant pixels, it is trained on the critic's objective, as pixels that are essential to determine the value of an observation should not be hidden. Figure 1 shows an example of corresponding to the current input frame. The Masker is able to bounding boxes, or other annotations. The reward alone is enough to evaluate the effectiveness of MaDi. We conduct experiments on multiple environments from three benchmarks: the DeepMind Control Generalization Benchmark [22], the Distracting Cart-pole [19], and a real UR5 Robotic Arm for which we design a novel generalization experiment with visual distractions. Our results by to focus on relevant visual information by generating helpful masks, leading to enhanced generalization performance. Furthermore, MaDi achieves state-of-the-art performance on many environments, surpassing well-known methods in vision-based reinforcement learning [4, 18, 21, 22, 28, 29].

Our main contributions are:

- We introduce a novel algorithm, MaDi, which supplements the standard actor-critic architecture of deep RL agents with a lightweight Masker. This network learns to focus on the task-relevant pixels solely from the reward signal.
- We present a comprehensive set of experiments on the DeepMind Control Generalization Benchmark and the Distracting Cart-pole. MaDi consistently achieves state-of-the-art or near state-of-the-art performance.
- We test MaDi on a physical robot, demonstrating that our algorithm increases the performance of the UR5 Robotic Arm videos are playing in the background.

Oral at AAMAS'24

Spotlight at SNN workshop of ICLR'23  
Full-paper at AAMAS'23



# Why this research?



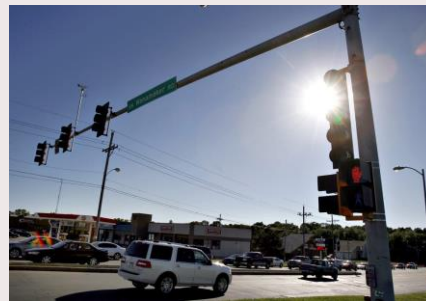
<https://www.istockphoto.com/nl/foto/robotmeisje-dat-een-dienblad-houdt-en-voedsel-en-drank-in-moderne-binnenlandse-keuken-gm1309871457-399451605>



<https://www.pexels.com/photo/people-walking-on-times-square-during-night-time-12729170/>

# Types of noise

1. Uncertainty in perception  
measurement errors
2. Task-irrelevant percepts  
distracting information



<https://eu.cjonline.com/story/news/local/2010/09/25/sun-poses-driving-challenges/16488302007/>



<https://www.pexels.com/photo/people-walking-on-times-square-during-night-time-12729170/>

# Contents

Motivation & Introduction

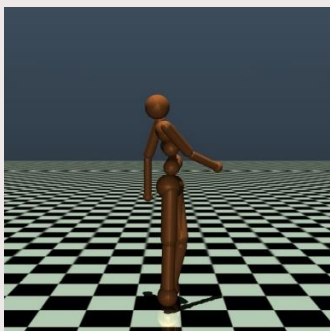
**Automatic Noise Filtering (ANF)**

Masking Distractions (MaDi)

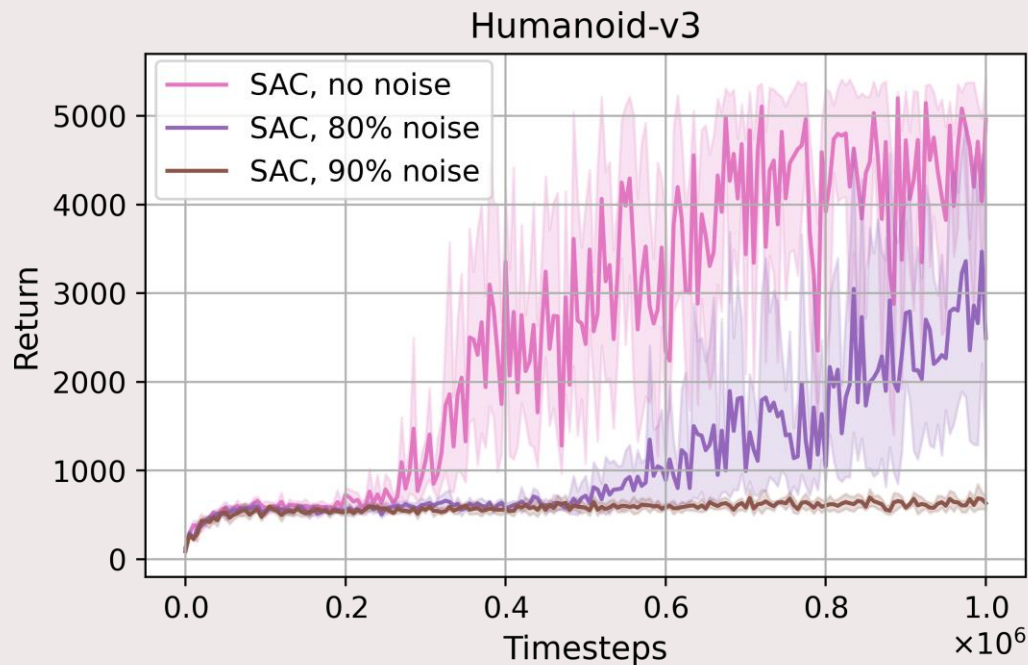
Potential future work

# Automatic Noise Filtering

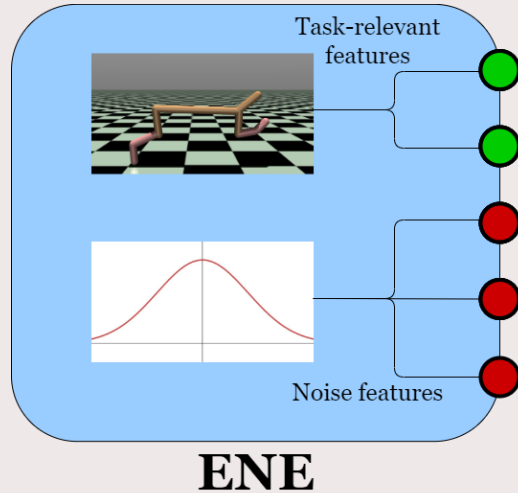
Current situation:



Humanoid-v3



# How does ANF work?

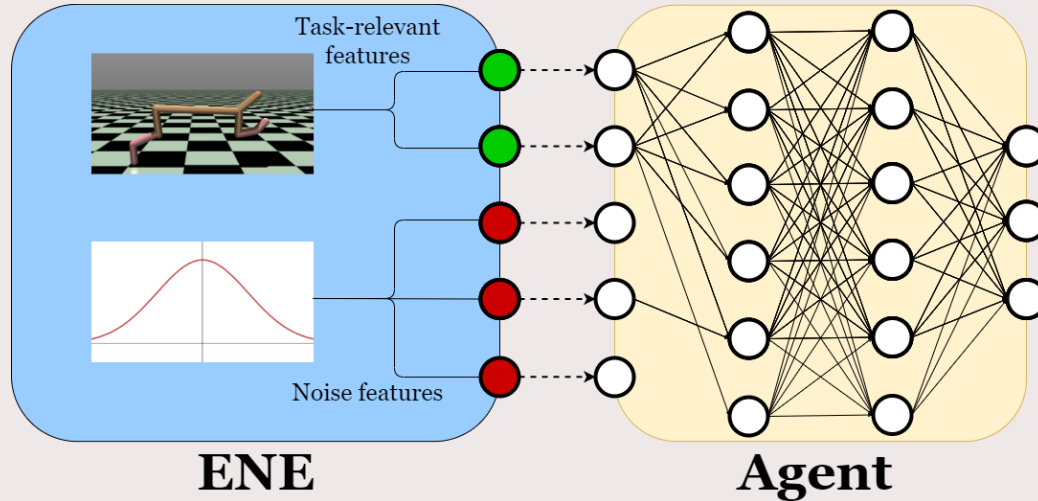


Extremely noisy environment:

simulate irrelevant information by adding many *noise features*

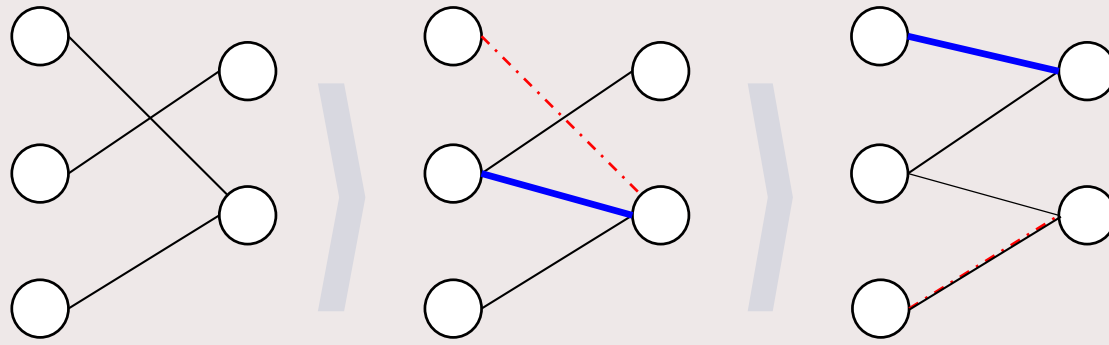


# How does ANF work?



Automatic Noise Filtering:  
adapt the input layer's connectivity to focus on relevant features

# How does ANF work?

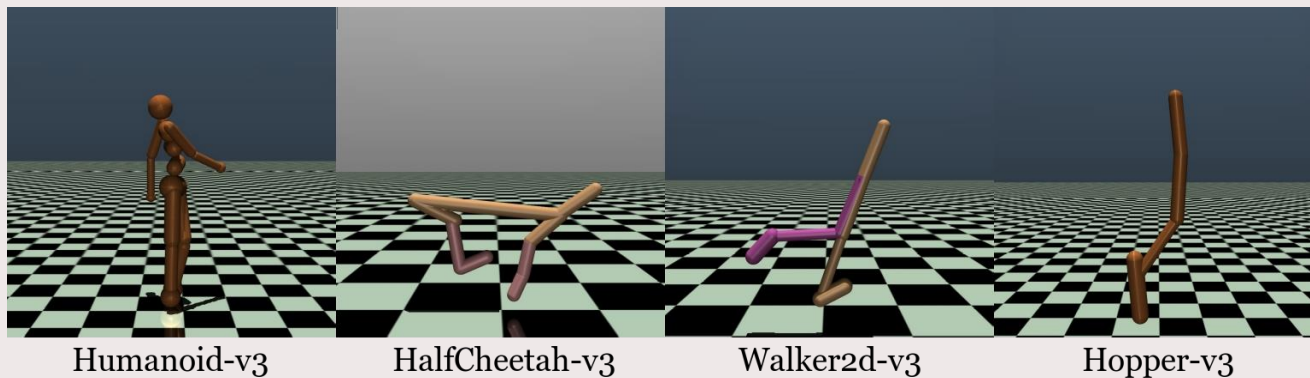


Dynamic Sparse Training (Drop - - - - , Grow — )

Sokar, G., et al. *Dynamic Sparse Training for Deep Reinforcement Learning*. IJCAI (2022).

Mocanu, D.C., et al. *Scalable training of artificial NNs with adaptive sparse connectivity inspired by network science*. Nature Comms. (2018).

# Experiments

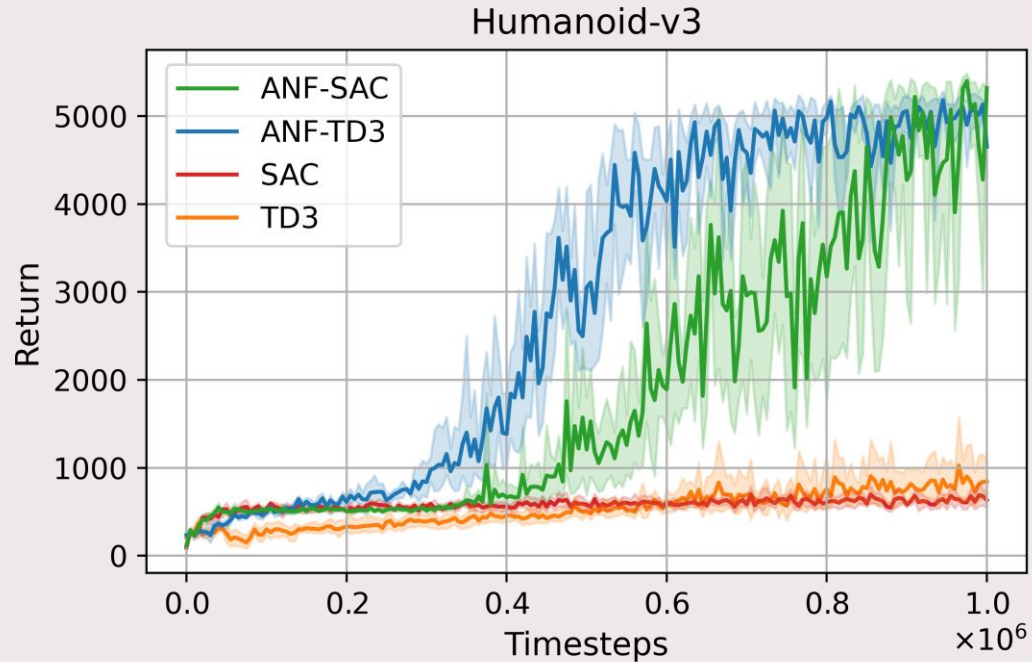


**Table 1: State and action space dimensions.**

| Environment    | State dim.      | Action dim.     | State dim.                         | State dim.                          |
|----------------|-----------------|-----------------|------------------------------------|-------------------------------------|
|                | <i>Original</i> | <i>Original</i> | <i>ENE (<math>n_f = .8</math>)</i> | <i>ENE (<math>n_f = .99</math>)</i> |
| Humanoid-v3    | 376             | 17              | 1880                               | 37600                               |
| HalfCheetah-v3 | 17              | 6               | 85                                 | 1700                                |
| Walker2d-v3    | 17              | 6               | 85                                 | 1700                                |
| Hopper-v3      | 11              | 3               | 55                                 | 1100                                |

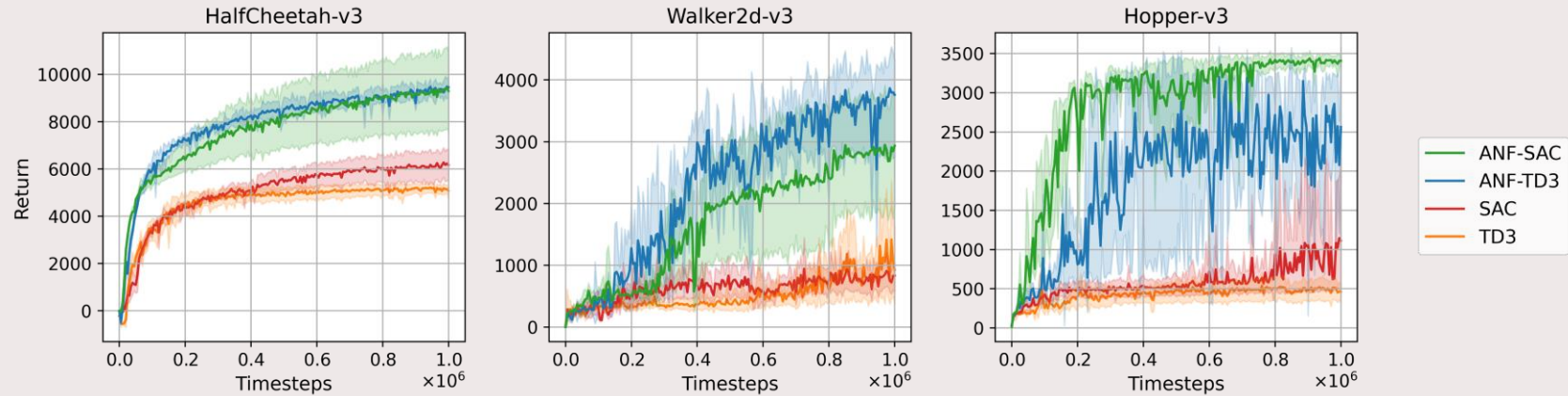
# Results

90% noise features:



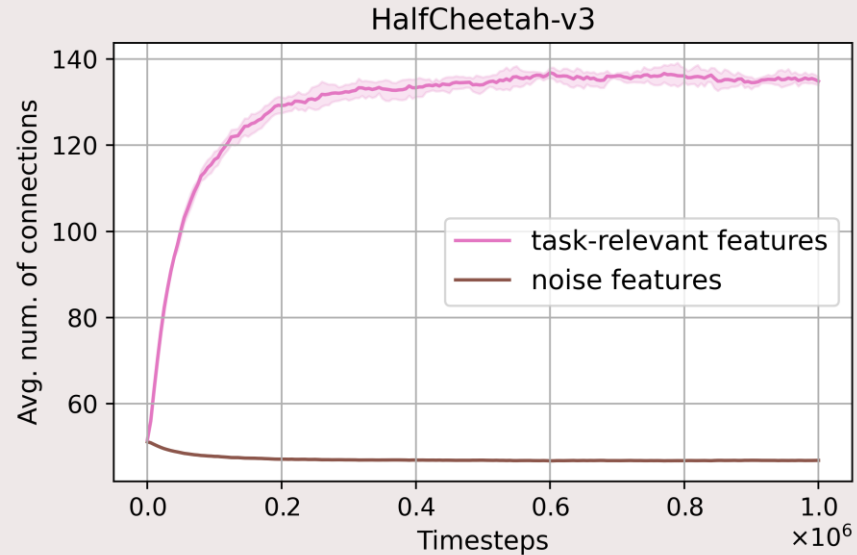
# Results

98% noise features:





# Results



ANF adapts the connectivity to focus on task-relevant features

# Louder noise

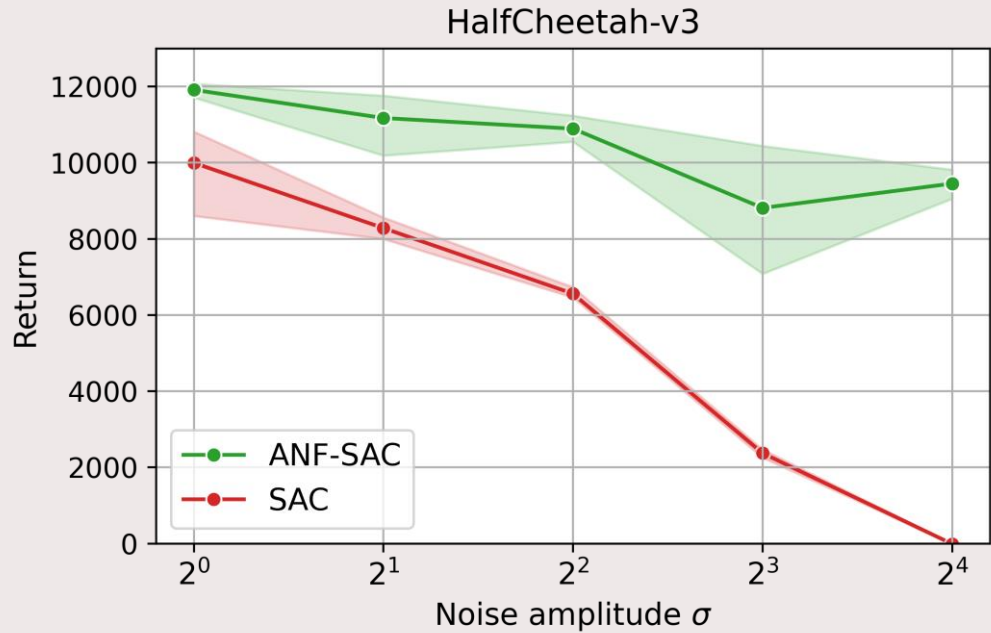
Previously:

Noise sampled from  $N(0, 1)$

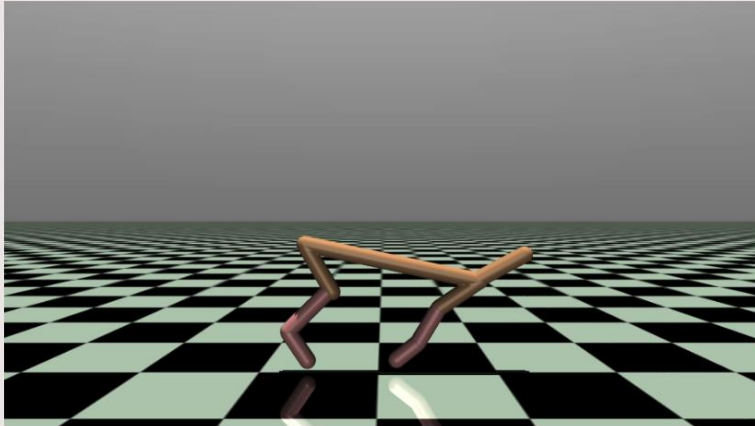
Now:

Sample from:  $N(0, \sigma^2)$

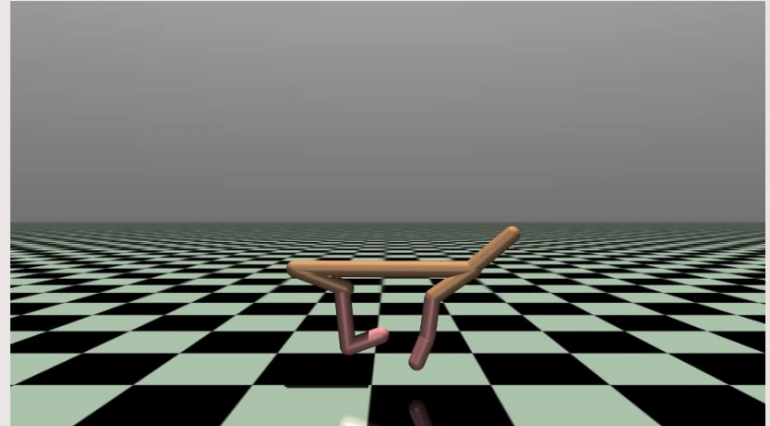
Varying the *noise amplitude*  $\sigma$



# Louder noise



SAC

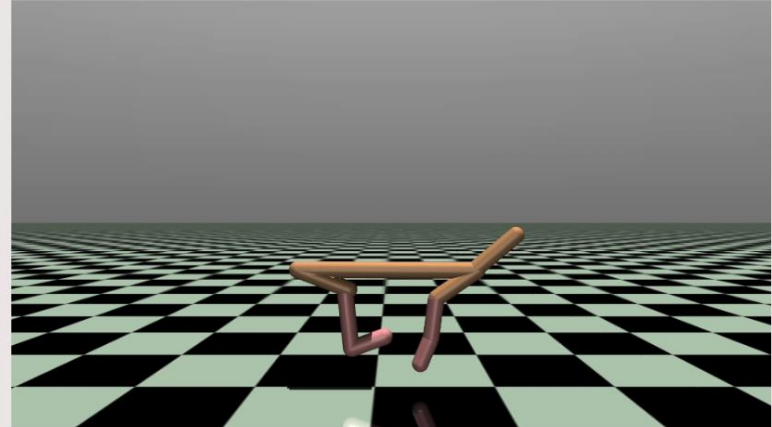


ANF-SAC

# Louder noise

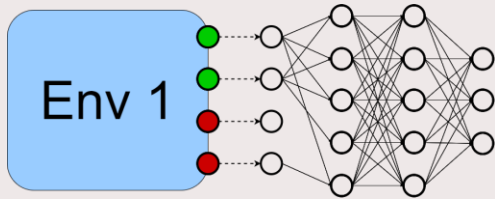


Source: [twitter.com/wonderofscience/status/1542146051658104832](https://twitter.com/wonderofscience/status/1542146051658104832)



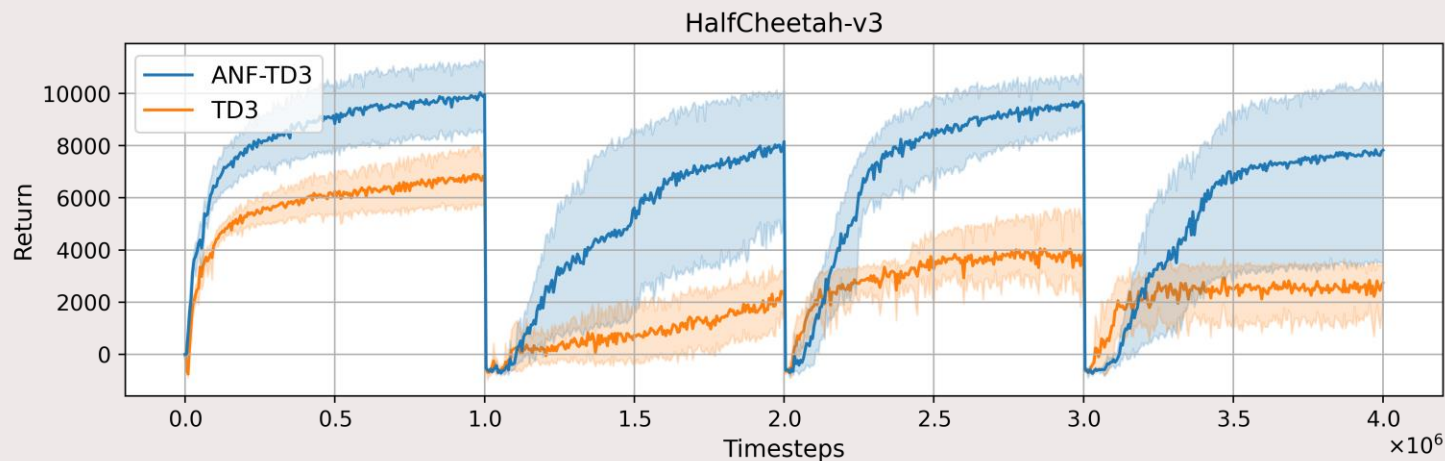
ANF-SAC

# Transfer learning



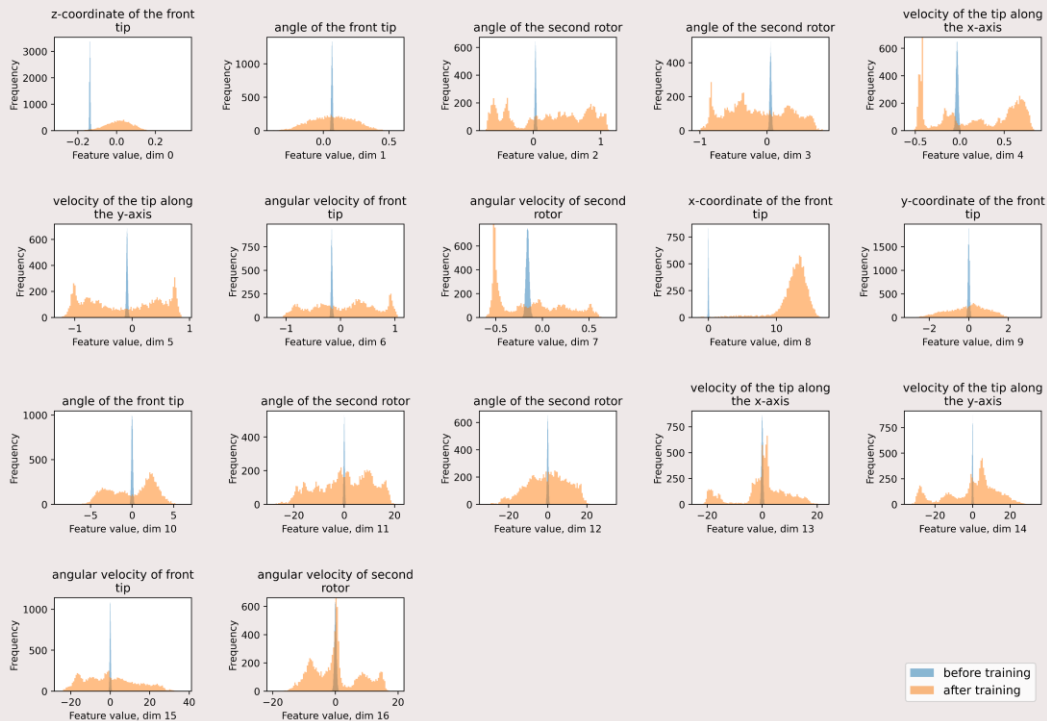
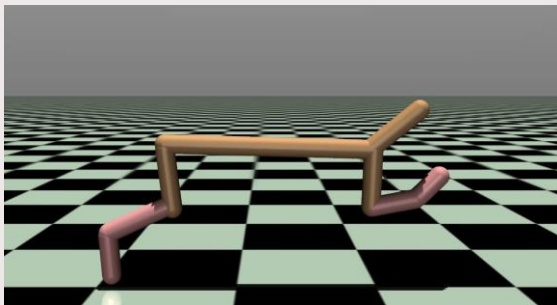


# Transfer learning

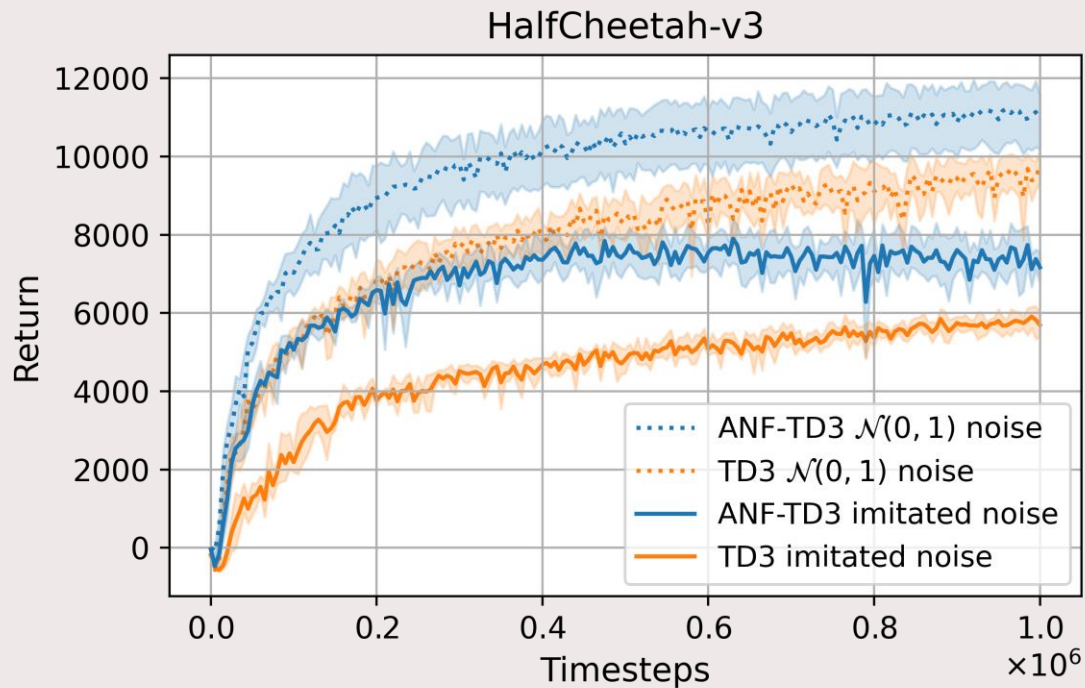


Every 1M timesteps the environment changes

# Imitating real features



# Imitating real features



Motivation & Introduction

Automatic Noise Filtering (ANF)

**Masking Distractions (MaDi)**

Potential future work

# Masking Distractions (MaDi)



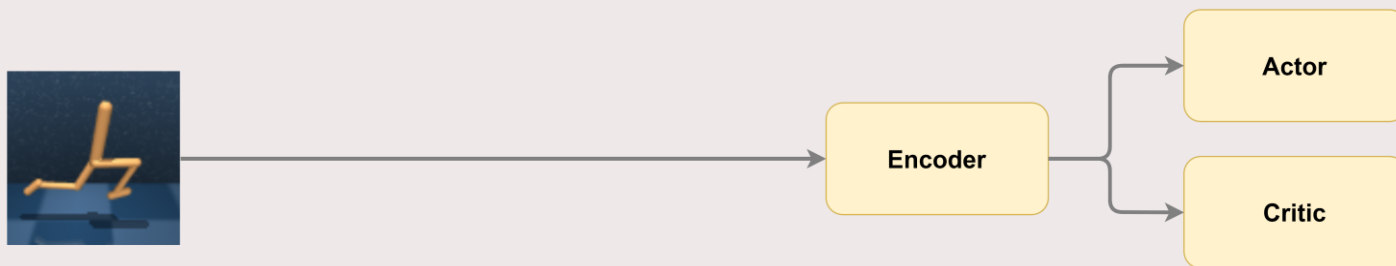
Training environments

Testing

Stone et al. *The Distracting Control Suite – A Challenging Benchmark for RL from Pixels.* (2021)

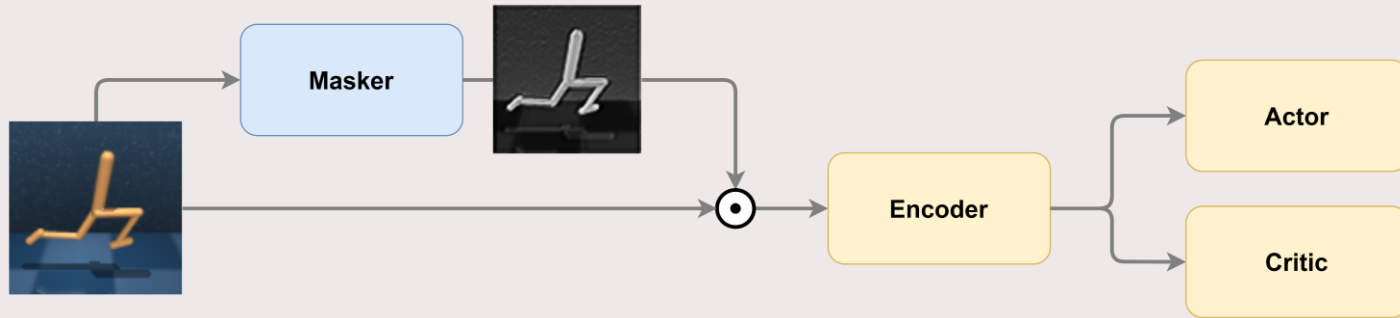


# How does MaDi work?

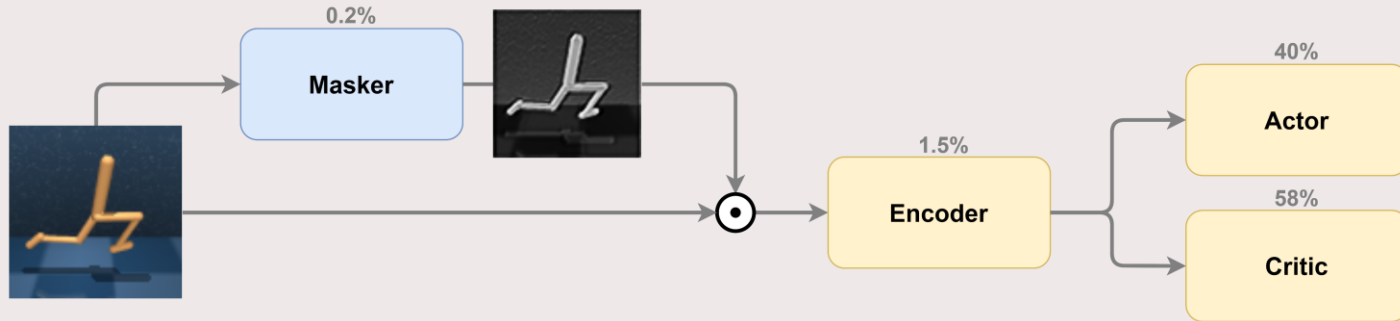


Hansen et al. *Stabilizing Deep Q-Learning with ConvNets and Vision Transformers under Data Augmentation*. (2021)

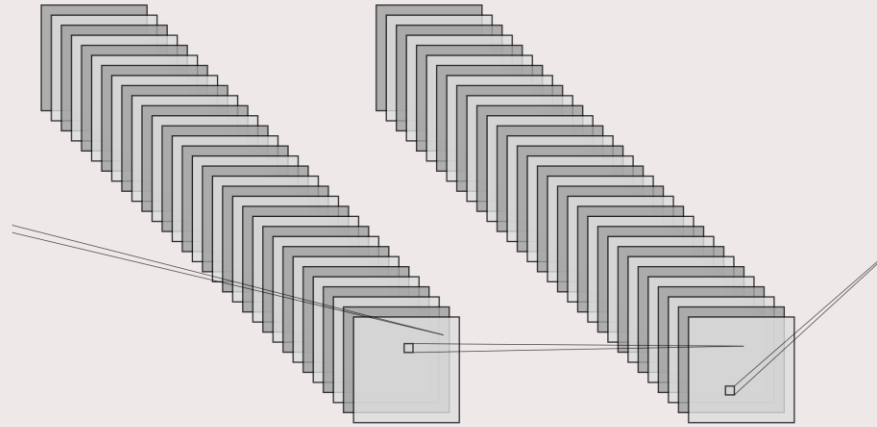
# How does MaDi work?



# How does MaDi work?

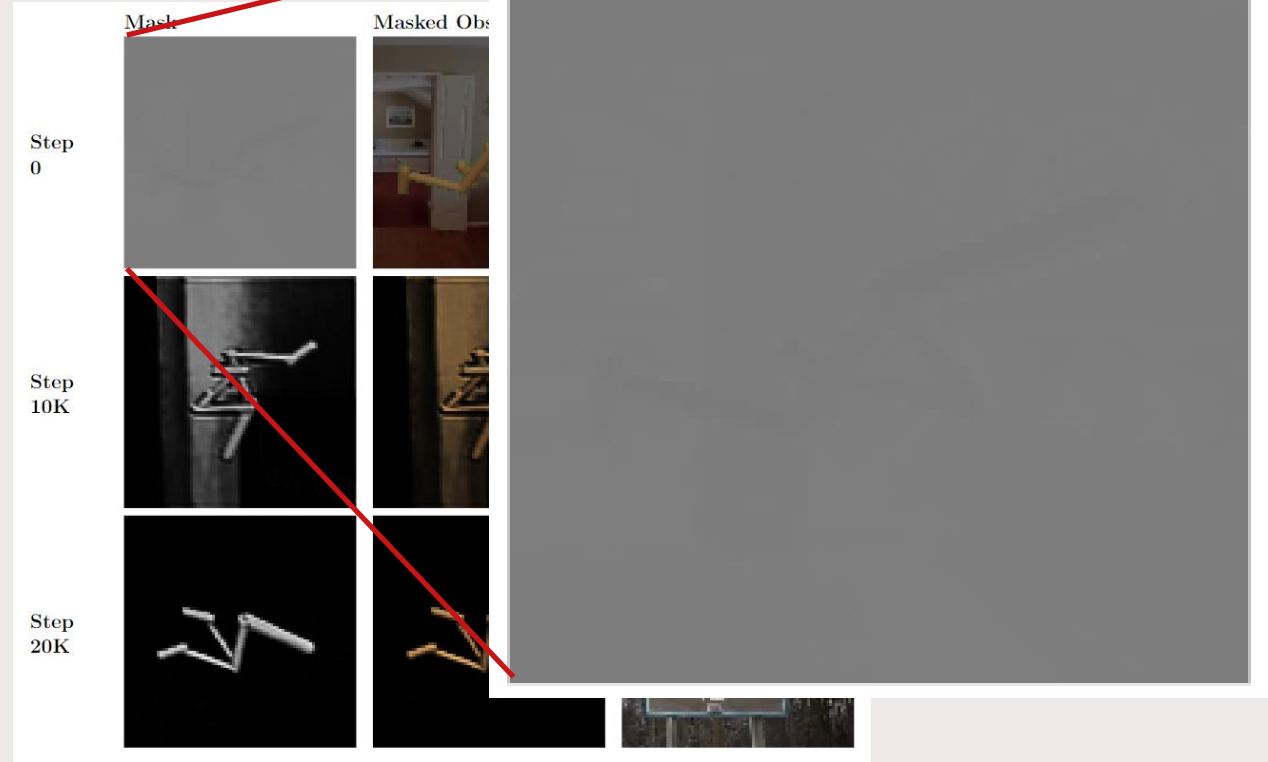


# How does MaDi work?

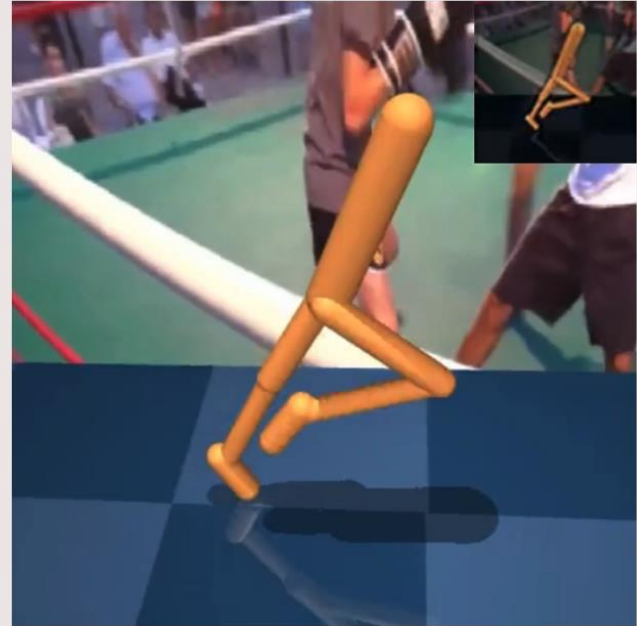
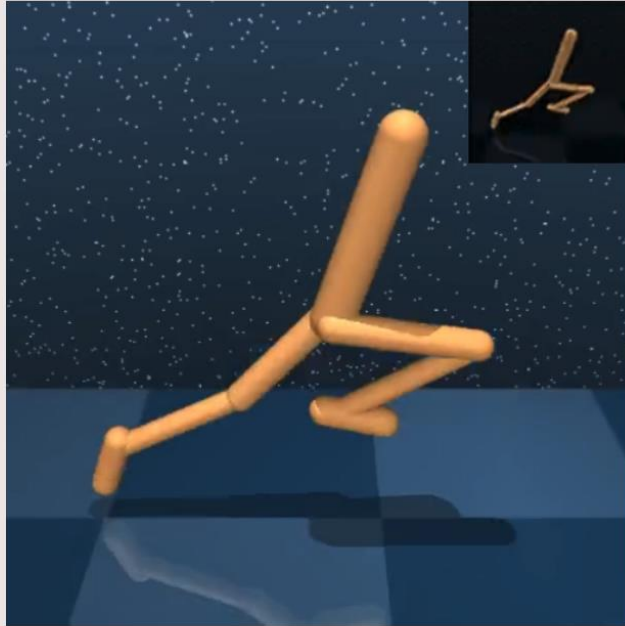


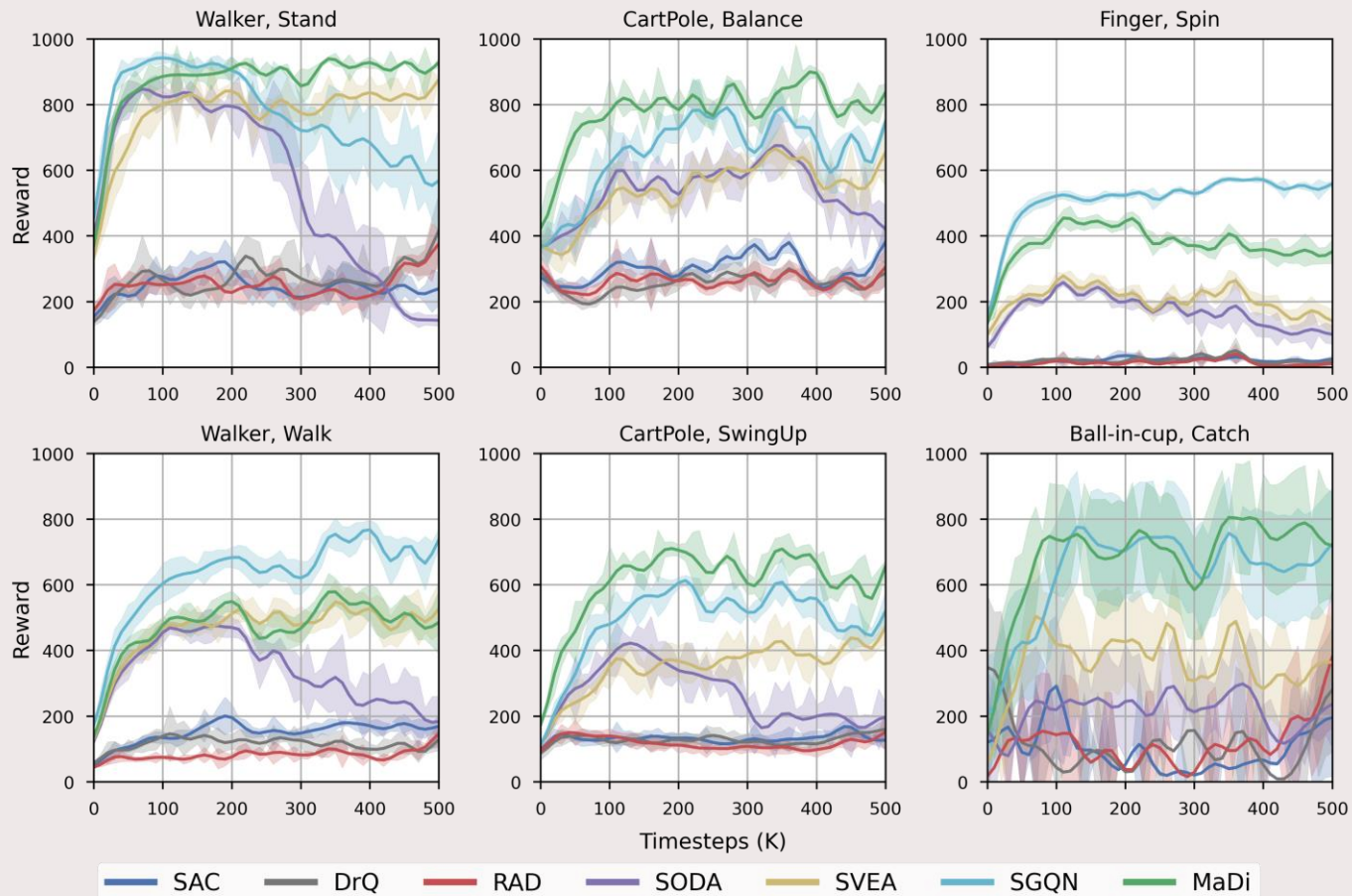
# How does MaDi work?

Initialization:



# Results







(a) training, walker



(b) video\_hard, walker



(c) distracting\_cs, walker



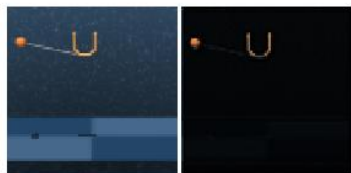
(d) training, cartpole



(e) video\_hard, cartpole



(f) distracting\_cs, cartpole



(g) training, ball\_in\_cup



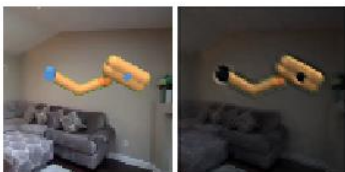
(h) video\_hard, ball\_in\_cup



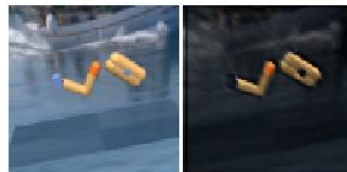
(i) distracting\_cs, ball\_in\_cup



(j) training, finger



(k) video\_hard, finger



(l) distracting\_cs, finger



# Robotic experiments

UR5 Robotic Arm

VisualReacher task

Training asynchronously

Inputs:

- webcam view
- proprioception

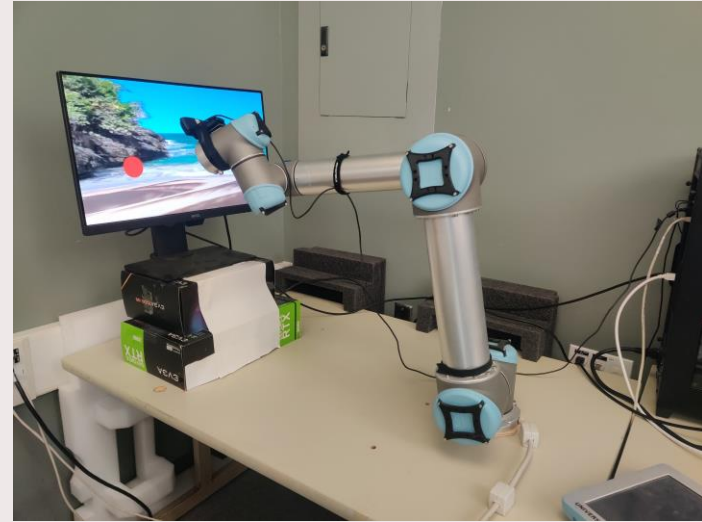


Wang et al. Real-Time Reinforcement Learning for Vision-Based Robotics Utilizing Local and Remote Computers. (2023)

# Robotic experiments

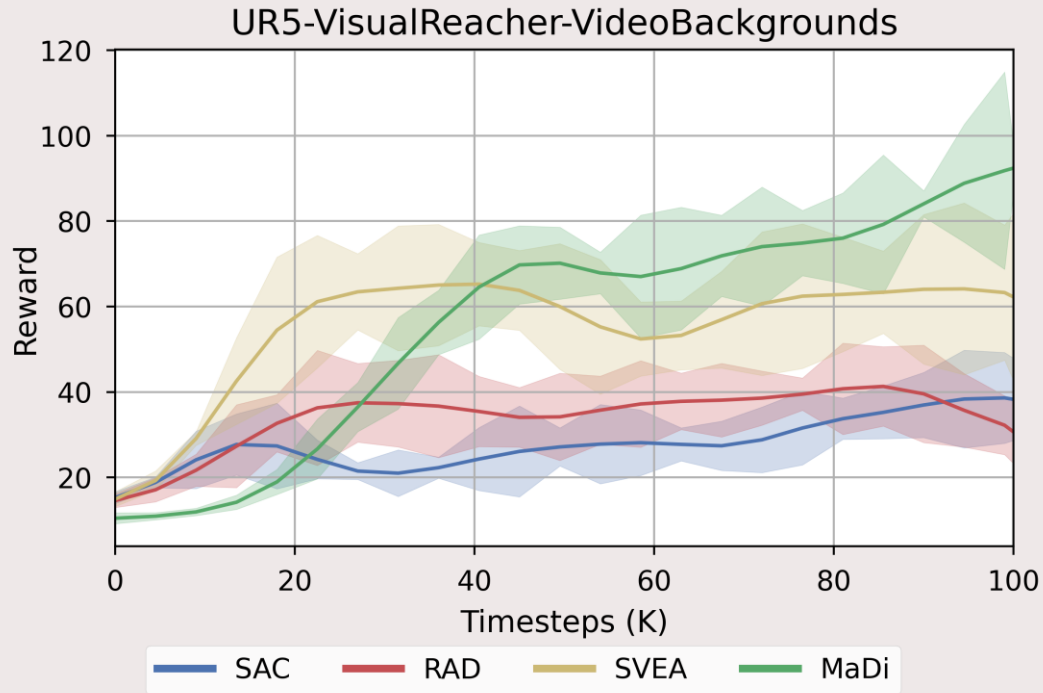


Training



Testing

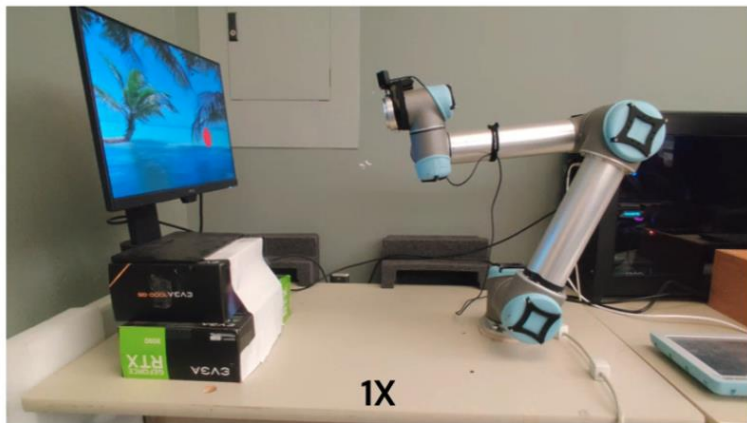
# Robotic experiments



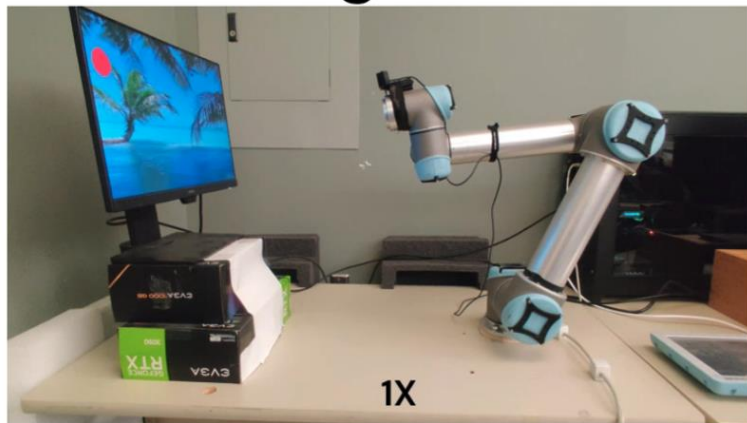


# UR5-VisualReacher-VideoBackgrounds

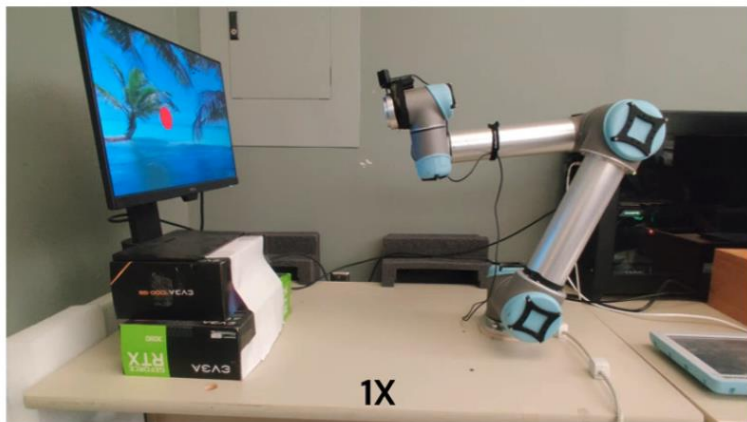
MaDi



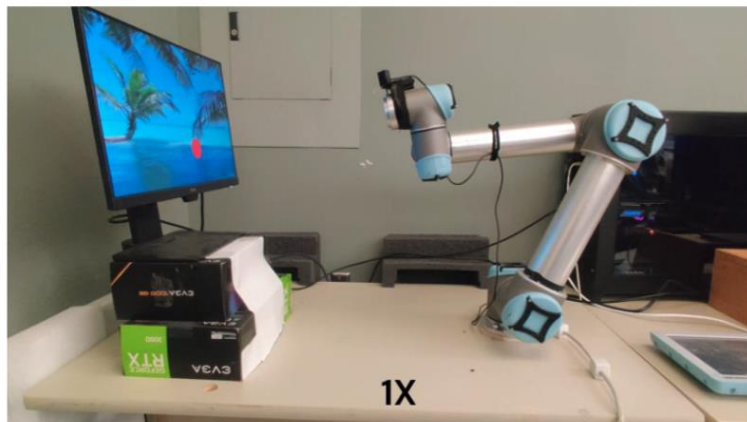
SVEA



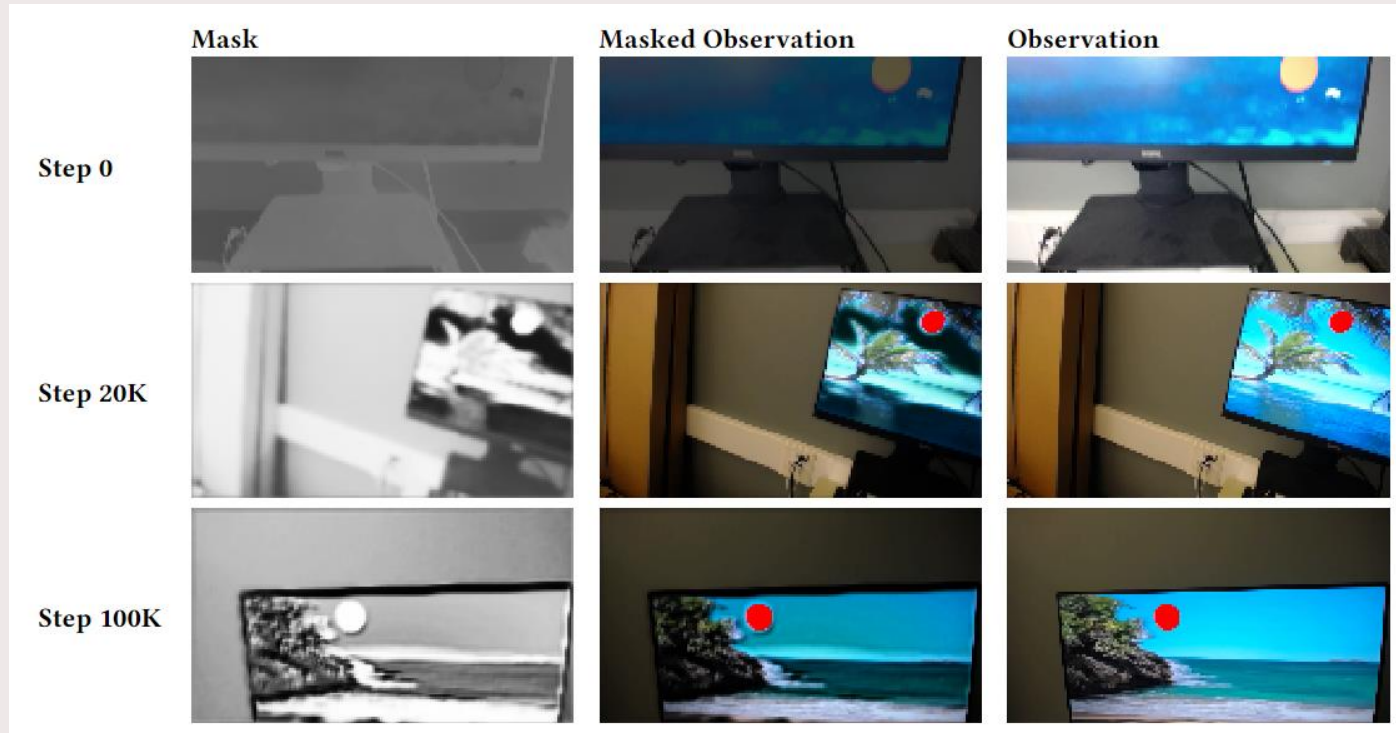
RAD



SAC



# Masks



Motivation & Introduction

Automatic Noise Filtering (ANF)

Masking Distractions (MaDi)

**Potential future work**

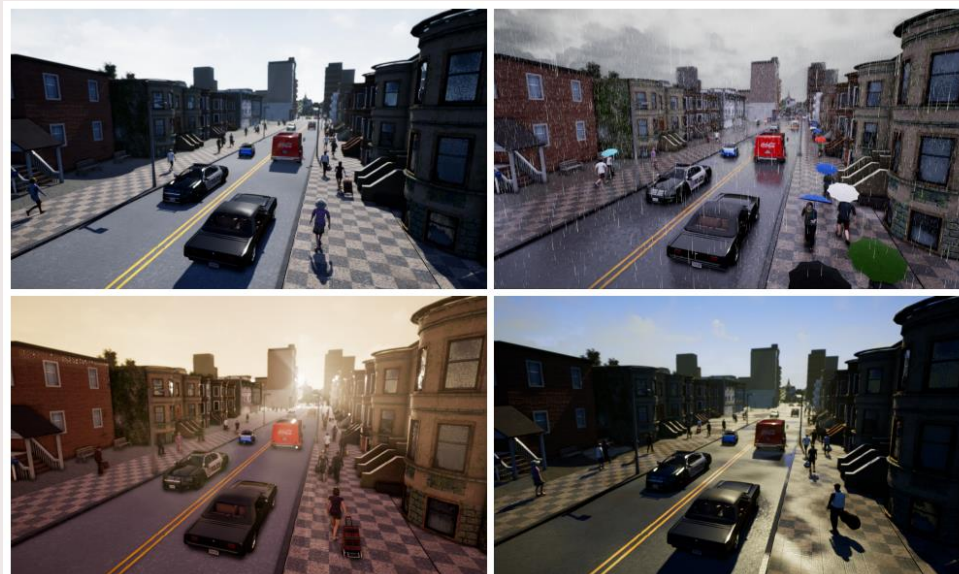


# Potential future work

- Other benchmarks or robotics applications
- Sparsify later in the network
- Transfer Learning with MaDi
- Supervised Learning



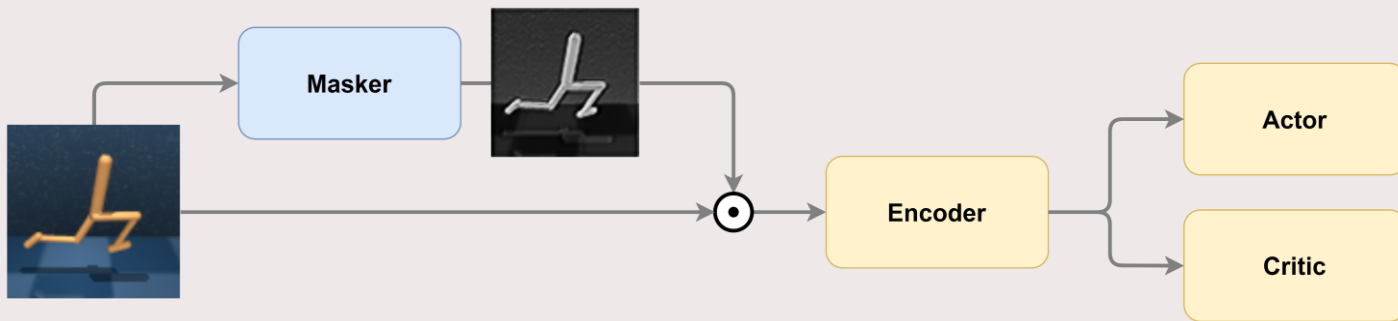
# MaDi on other benchmarks



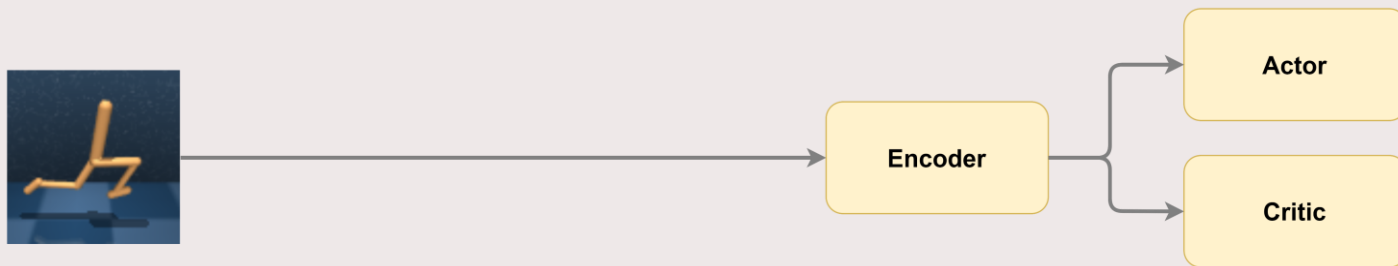
CARLA (Car Learning to Act)

Dosovitskiy et al. *CARLA: An Open Urban Driving Simulator*. CoRL (2017).

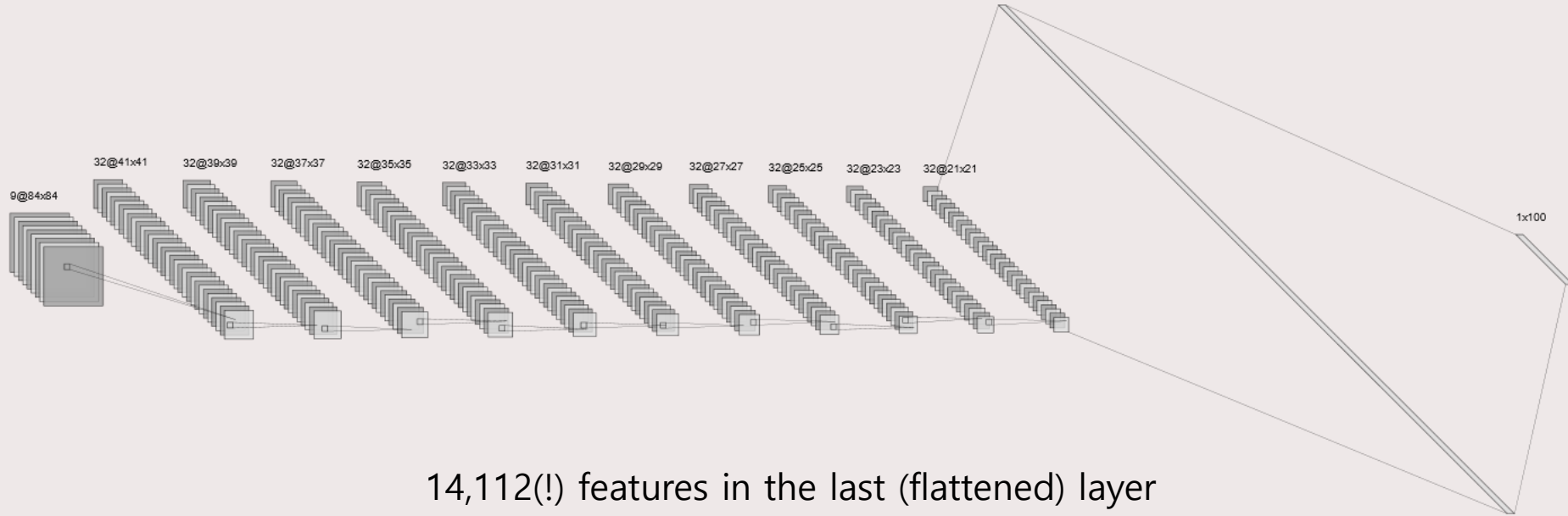
# Sparsify later in the network



# Sparsify later in the network



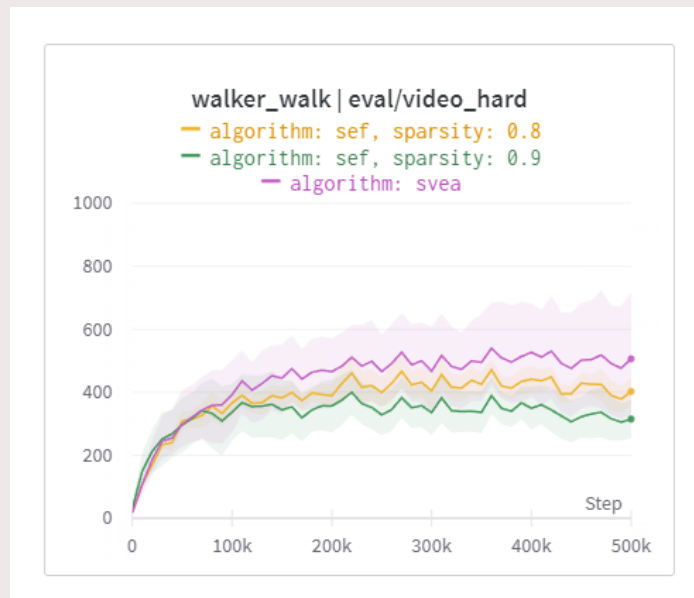
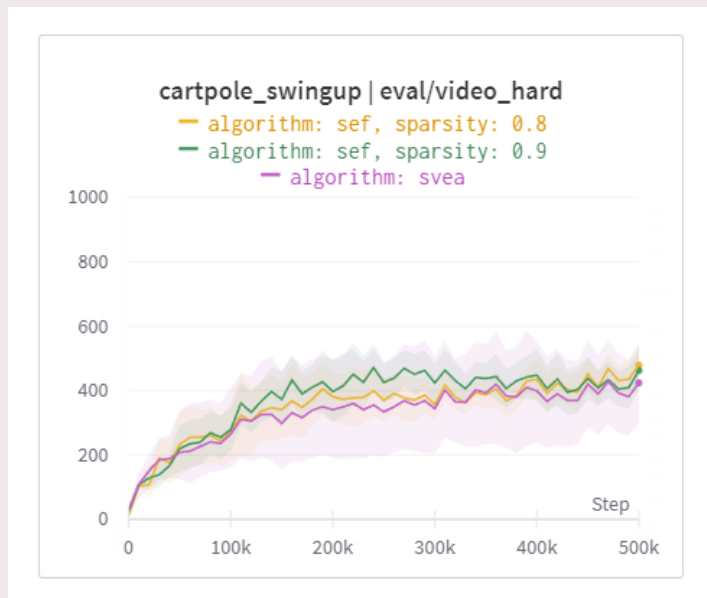
# Sparsify later in the network



14,112(!) features in the last (flattened) layer

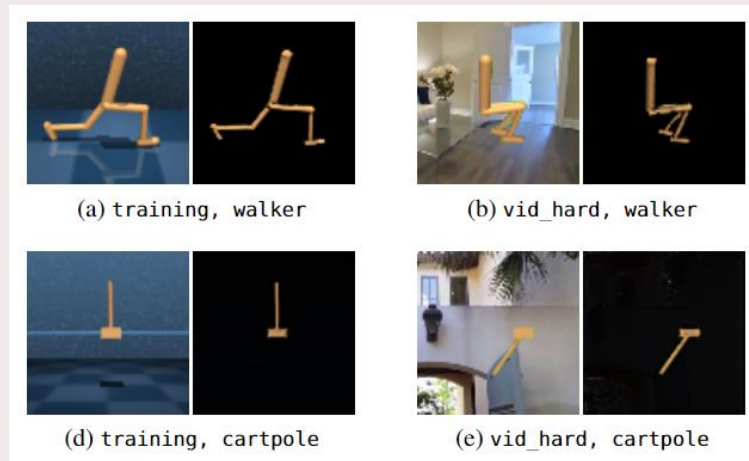
Made with: [alexlenail.me/NN-SVG](https://alexlenail.me/NN-SVG)

# Sparsify later in the network



# Transfer Learning with MaDi

walker\_stand  $\rightarrow$  walker\_walk

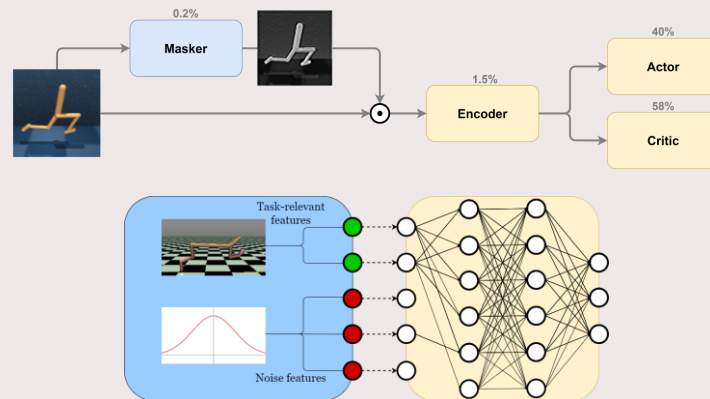


cartpole\_balance  $\rightarrow$  cartpole\_swingup

walker\_stand  $\rightarrow$  cartpole\_swingup (?)

## Main take-away

ANF and MaDi outperform deep RL methods by learning to focus on task-relevant features.




[arxiv.org/abs/2302.06548](https://arxiv.org/abs/2302.06548)



[arxiv.org/abs/2312.15339](https://arxiv.org/abs/2312.15339)

[github.com/bramgrooten/  
automatic-noise-filtering](https://github.com/bramgrooten/automatic-noise-filtering)



[github.com/bramgrooten/  
mask-distractions](https://github.com/bramgrooten/mask-distractions) 



@BramGrooten

# Tutorial

[github.com/bramgrooten/automatic-noise-filtering/blob/main/tutorial.md](https://github.com/bramgrooten/automatic-noise-filtering/blob/main/tutorial.md)

## Sparse Training in Deep RL - Tutorial

This guide is designed for anyone interested in using sparse neural networks in deep reinforcement learning. It was part of the [UCAI 2023 tutorial T27: Sparse Training for Supervised, Unsupervised, Continual, and Deep Reinforcement Learning with Deep Neural Networks](#). In the following we will play around with some of the settings in this repository. I hope this will give you a feeling for the types of problems we are researching in this field.

Author: [Bram Grooten](#).

### Install

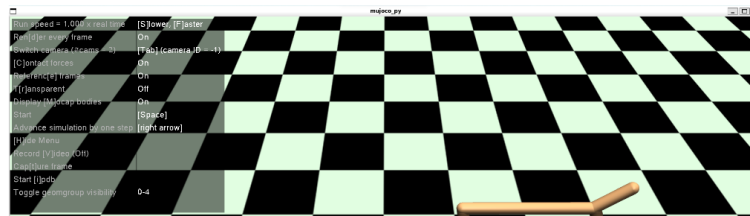
First `git clone` and install this repository, by following the instructions in the [README](#).

### Visualizing a trained agent

We have included a trained agent in the repository already, so let's view what this actually looks like! Go to this repository's main folder in your terminal (the directory with the `view_muJoco.py` file in it), make sure your venv is activated, and run:

```
python view_muJoco.py
```

A window should pop up showing you one episode of a running HalfCheetah! The camera doesn't track the agent by default, press `TAB` to do that.





# References

- Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. *CARLA: An Open Urban Driving Simulator*. Conference on Robot Learning (2017).
- Grooten, B., Sokar, G., Dohare, S., Mocanu, E., Taylor, M., Pechenizkiy, M., and Mocanu, D.C. *Automatic Noise Filtering with Dynamic Sparse Training in Deep Reinforcement Learning*. AAMAS (2023).
- Hansen, N., Su, H., Wang, X. *Stabilizing Deep Q-Learning with ConvNets and Vision Transformers under Data Augmentation*. NeurIPS (2021).
- Mocanu, D.C., Mocanu, E., Stone, P., Nguyen, P., Gibescu, M., and Liotta, A. *Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science*. Nature communications (2018).
- Sokar, G., Mocanu, E., Mocanu, D.C., Pechenizkiy, M., and Stone, P. *Dynamic Sparse Training for Deep Reinforcement Learning*. IJCAI (2022).
- Stone, A., Ramirez, O., Konolige, K., and Jonschkowski, R. *The Distracting Control Suite – A Challenging Benchmark for Reinforcement Learning from Pixels*. arXiv preprint (2021).
- Wang, Y., Vasan, G., and Mahmood, A.R. *Real-Time Reinforcement Learning for Vision-Based Robotics Utilizing Local and Remote Computers*. ICRA (2023).