Interactive learning for Neurosciences: Between Simulation and Reality

Audrey Durand

Canada CIFAR AI Chair

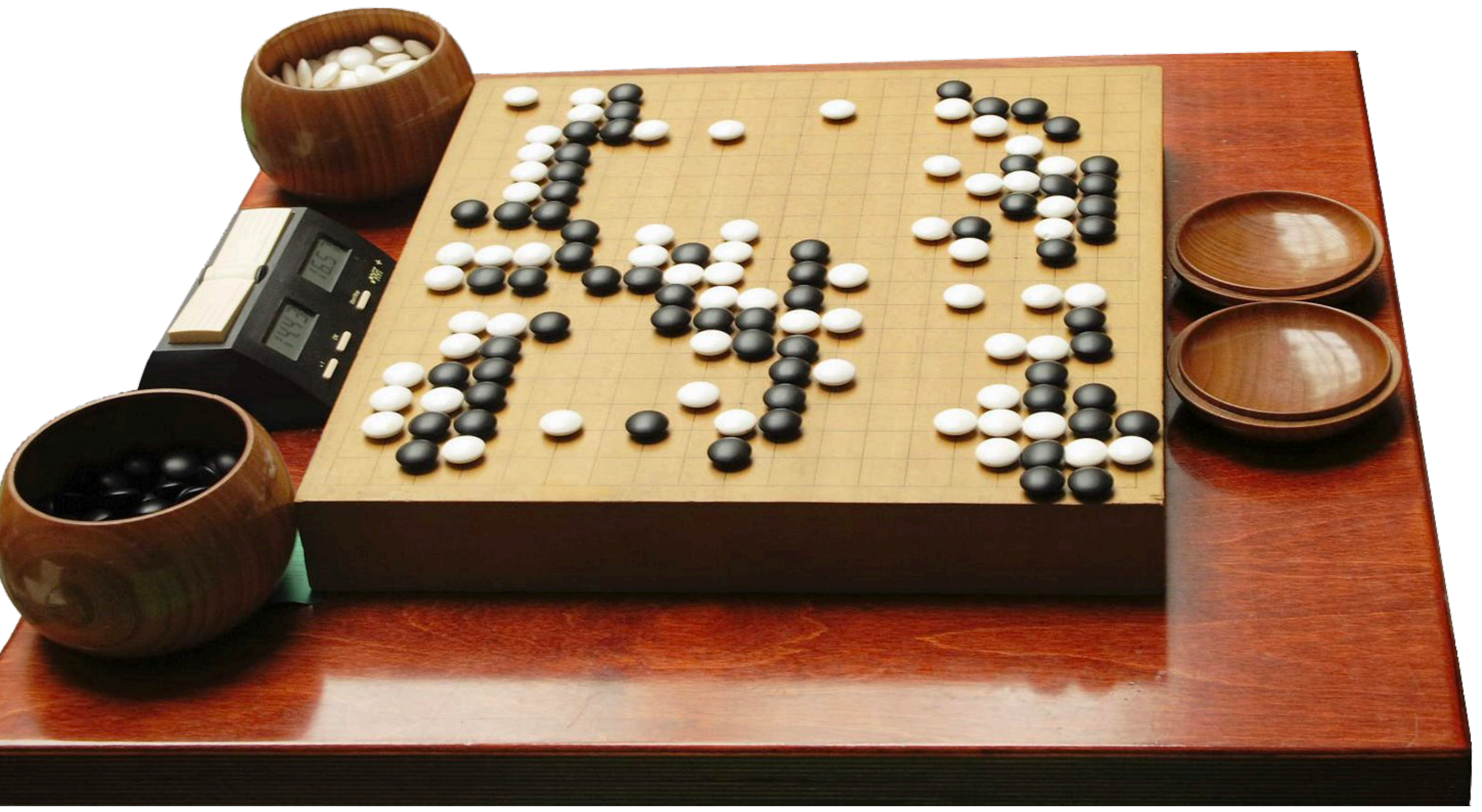
UNIVERSITÉ LAVAL

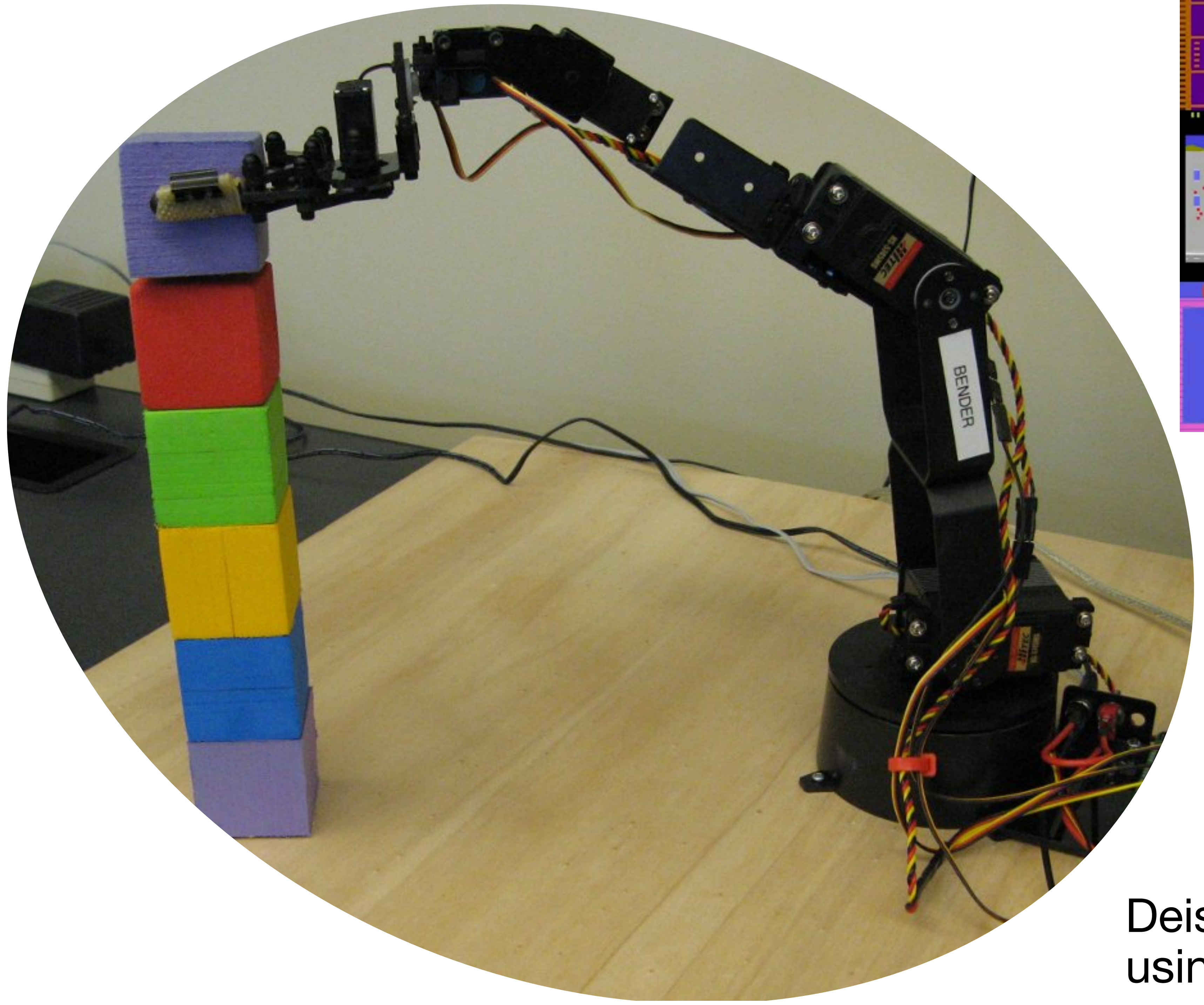
Institut intelligence et données

CIFAR

Mila

A popular topic

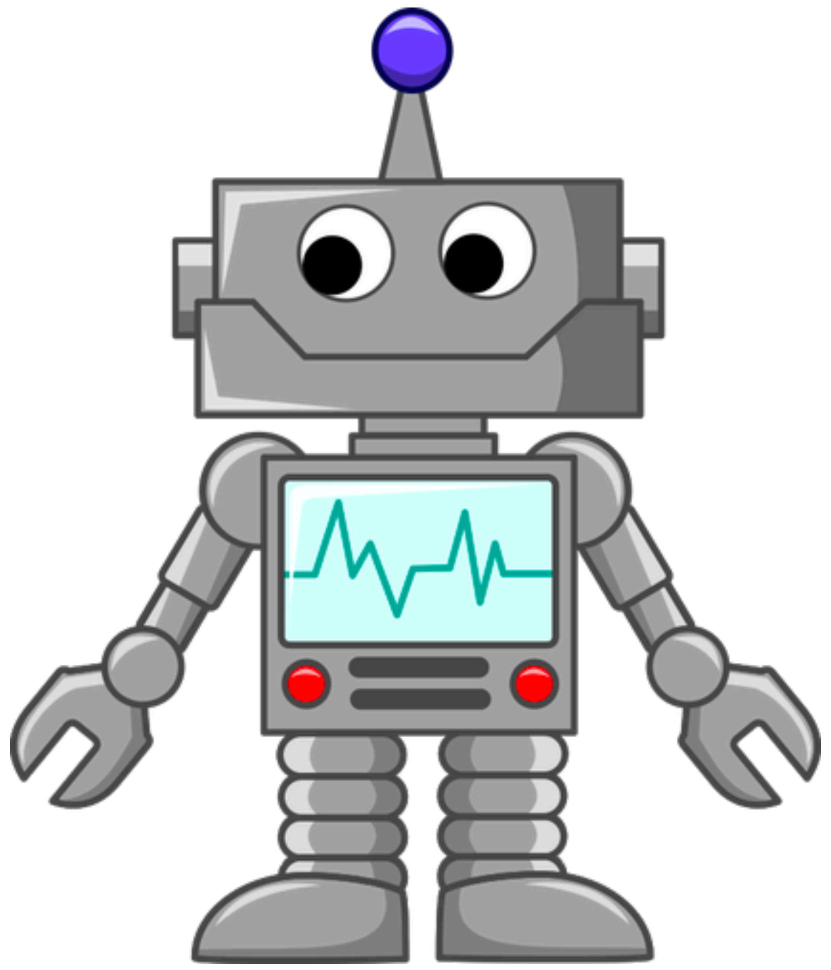




Deisenroth, M. P., Rasmussen, C. E., & Fox, D. (2011). Learning to control a low-cost manipulator using data-efficient reinforcement learning. Robotics: Science and Systems VII, 7, 57-64.

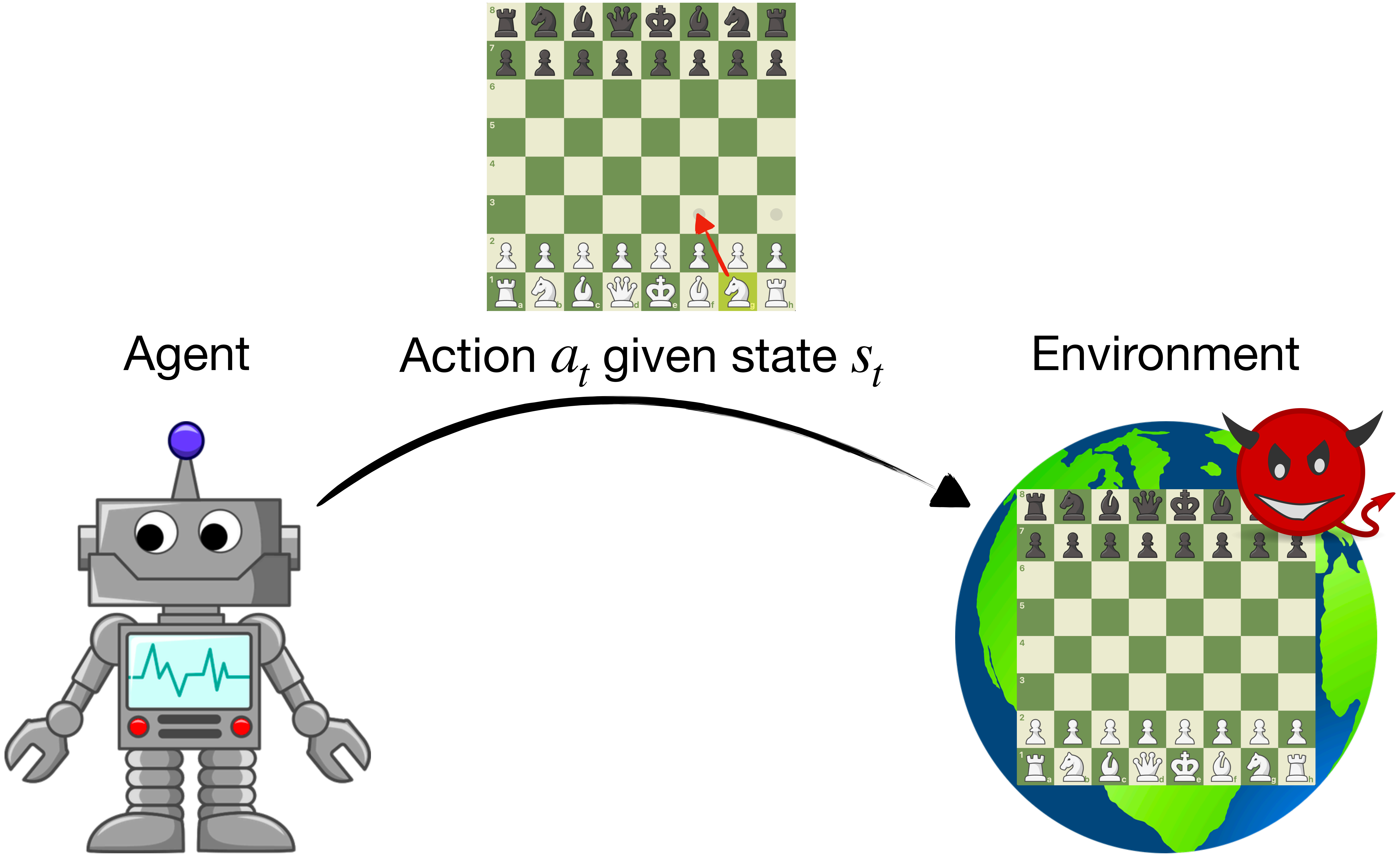
Reinforcement Learning (RL)

Agent

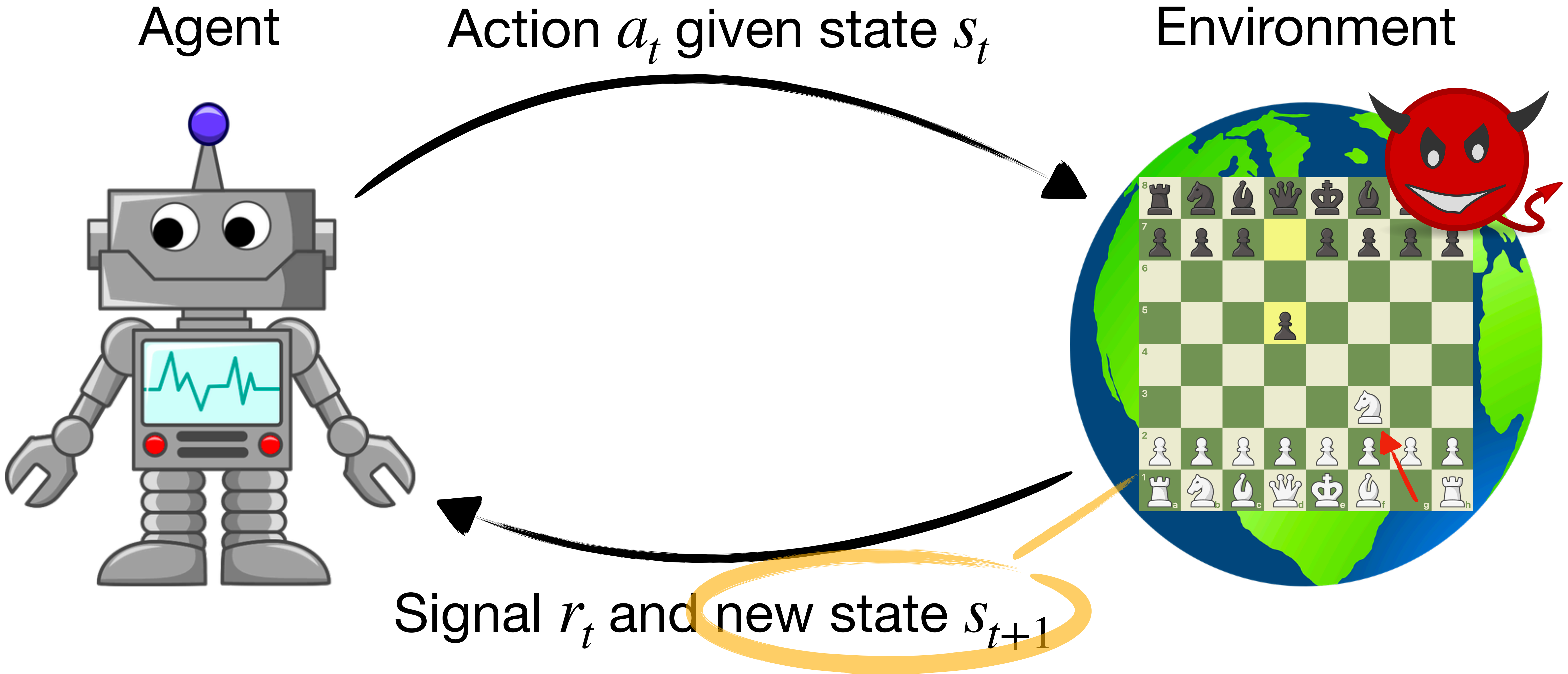


Environment

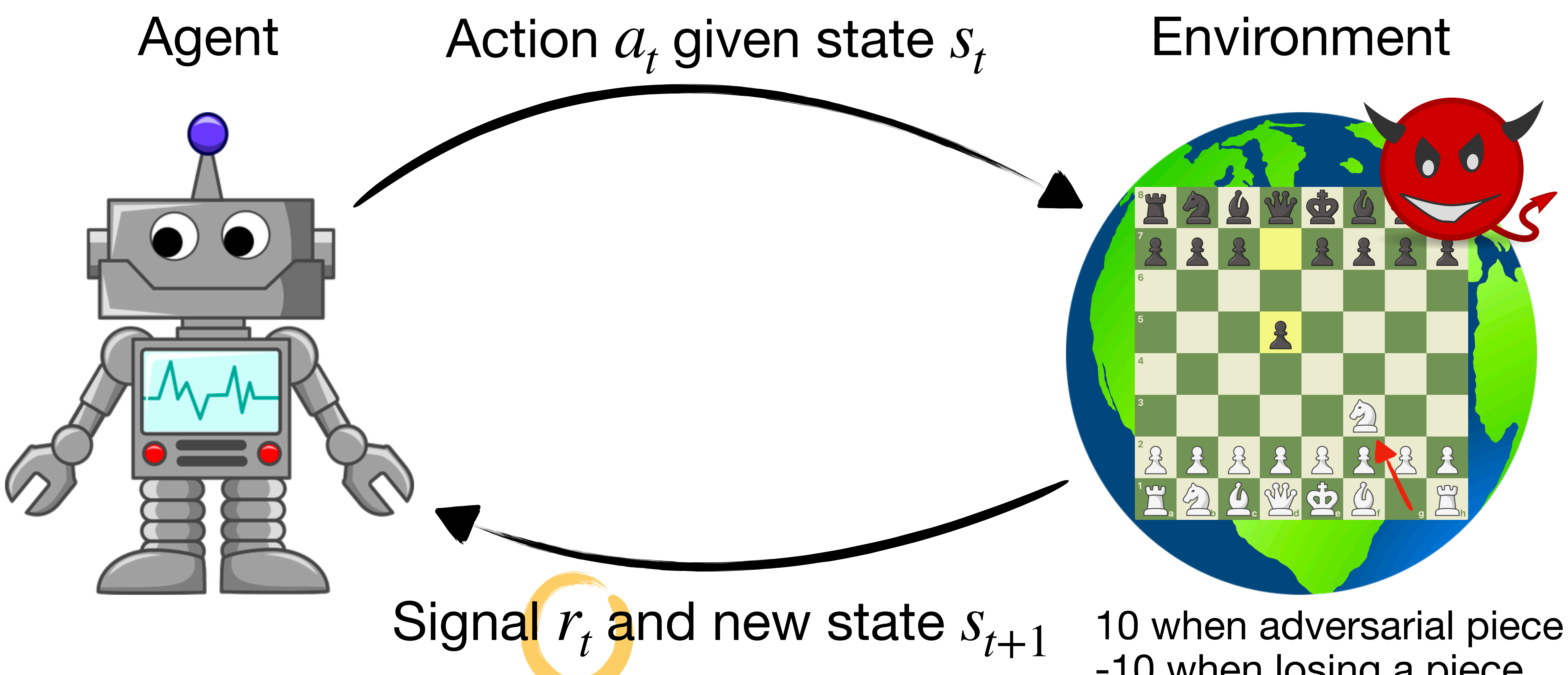
Reinforcement Learning (RL)



Reinforcement Learning (RL)



Reinforcement Learning (RL)

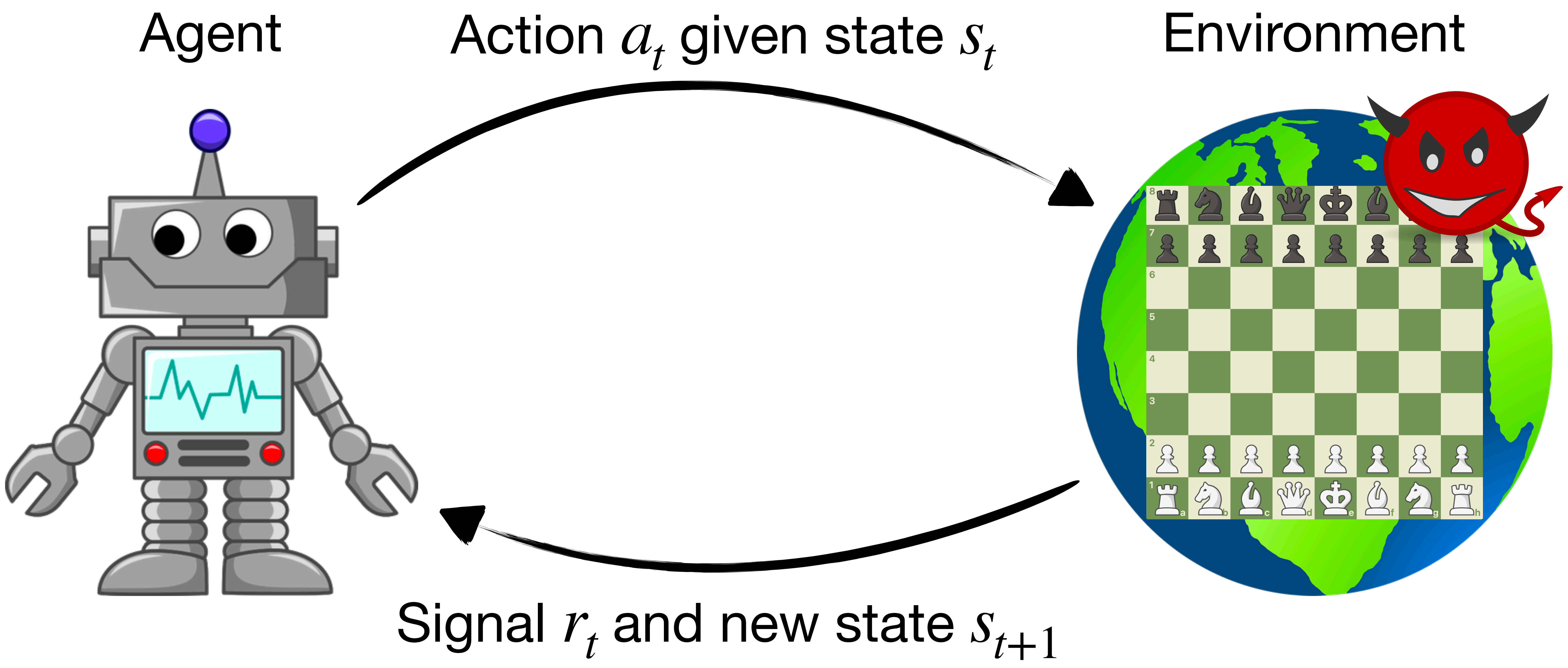


10 when adversarial piece eliminated
-10 when losing a piece
100 when winning game
-100 when losing game
-1 otherwise

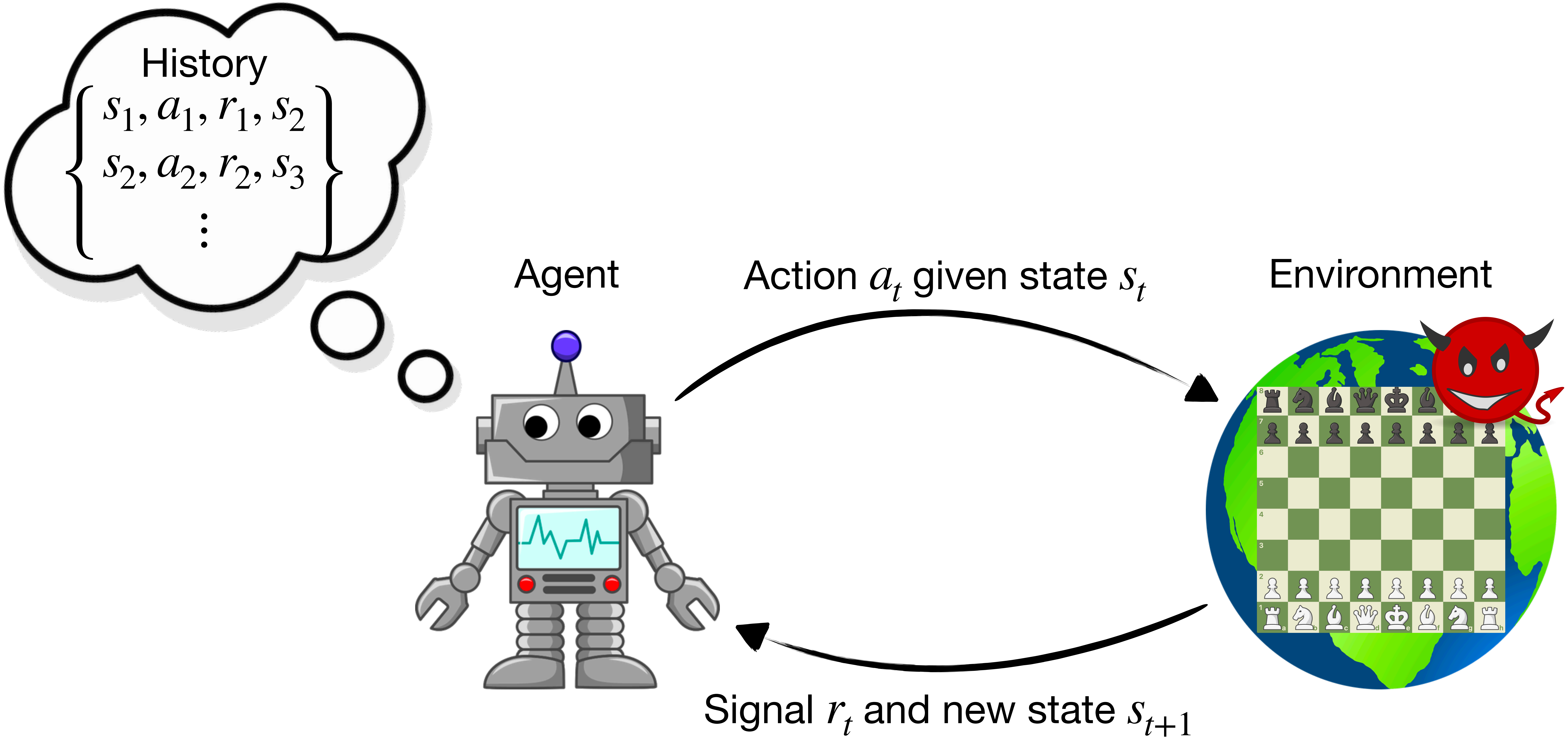
Episodic learning

1 episode = 1 game

- Game 1:  … Defeat!
- Game 2:  … Defeat!
- Game N:  … Victory!



Episodic learning

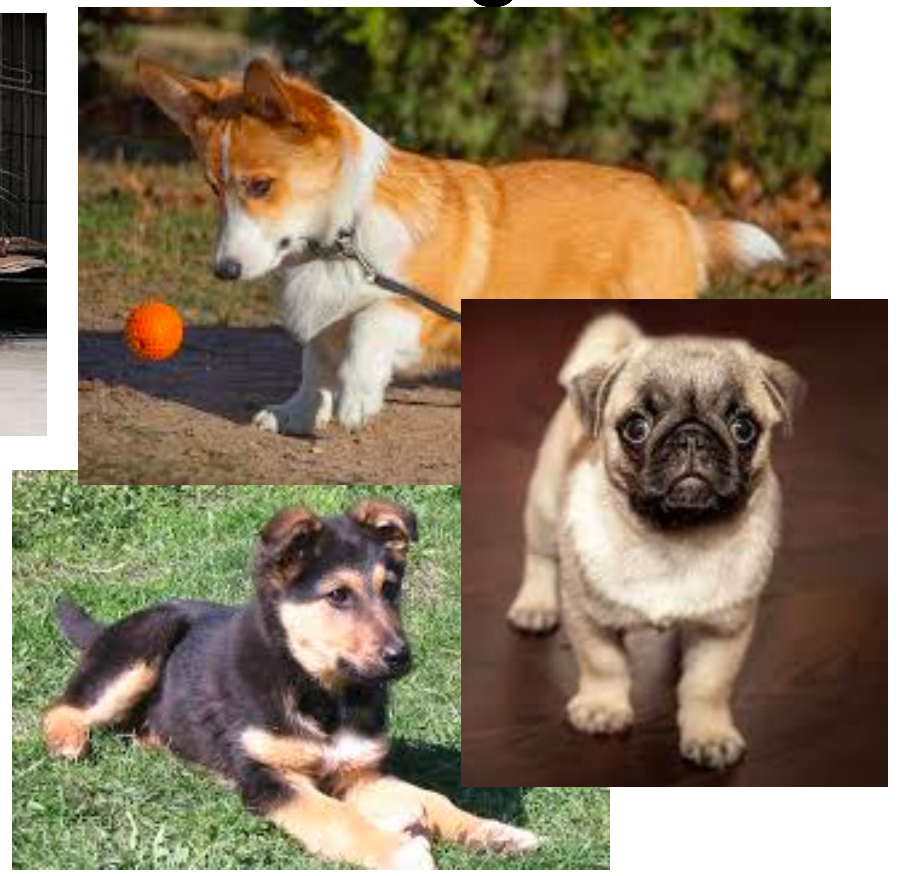


Recall Supervised Learning

Dataset

Cats:

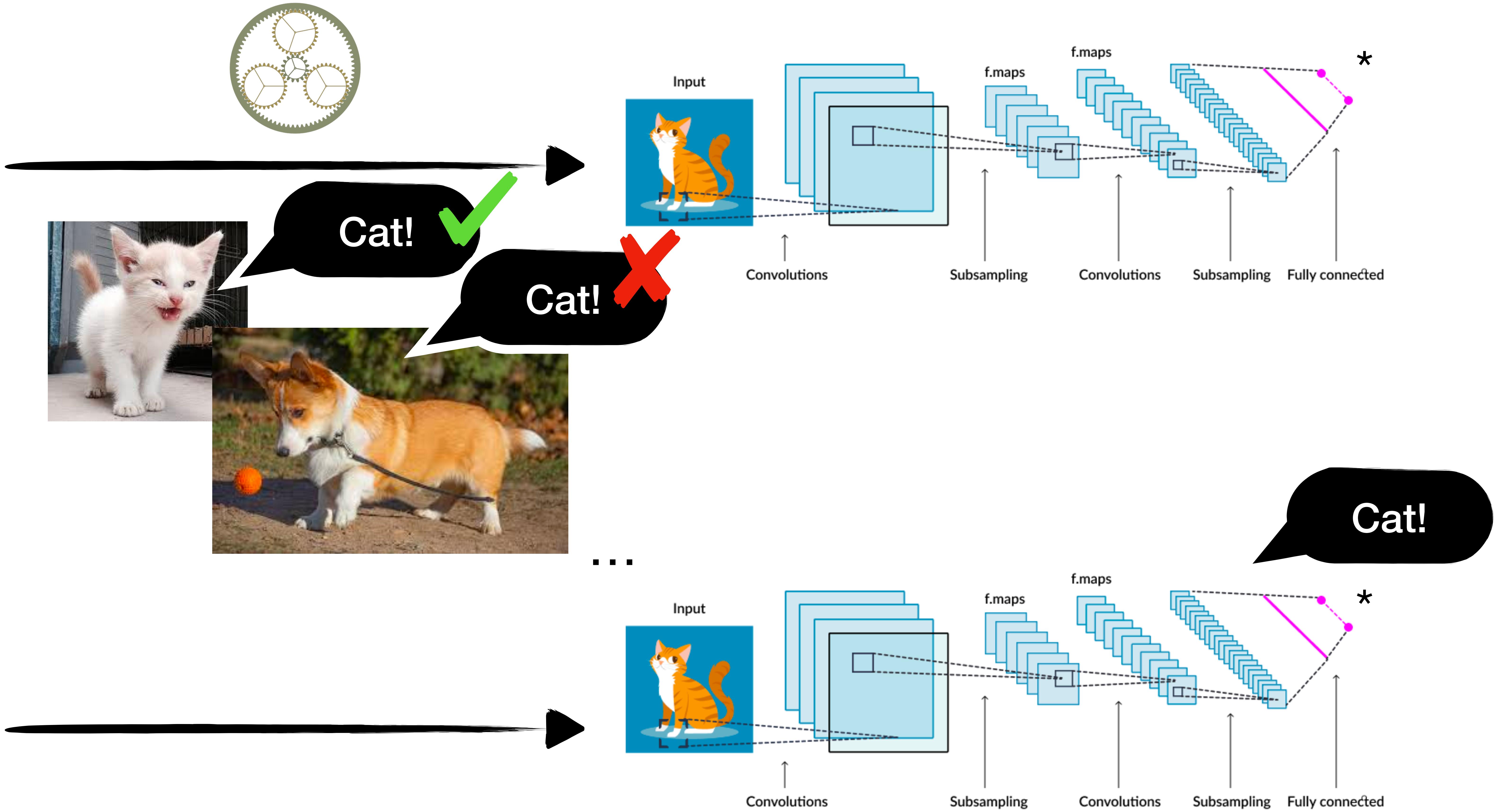
Dogs:



?

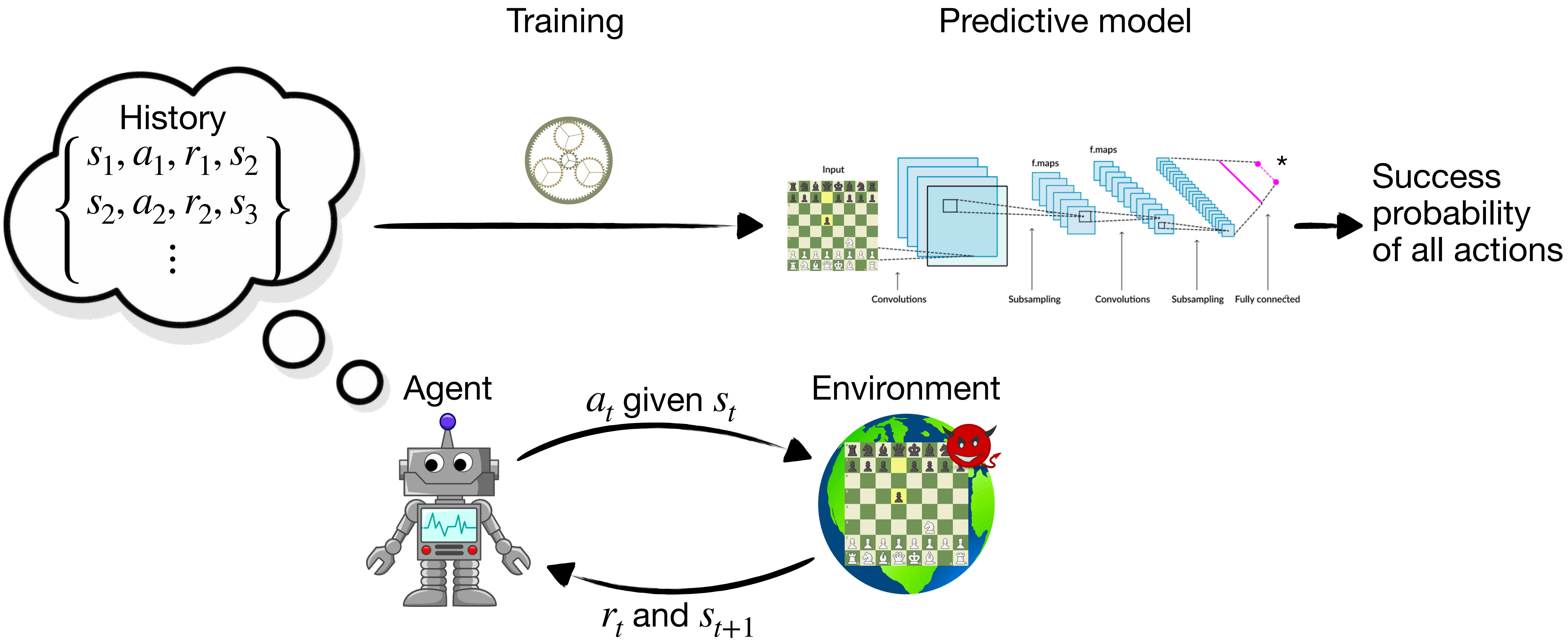
Training

Predictive model



* https://sandeep-bhuiya01.medium.com/disadvantages-of-cnn-models-95395fe9ae40

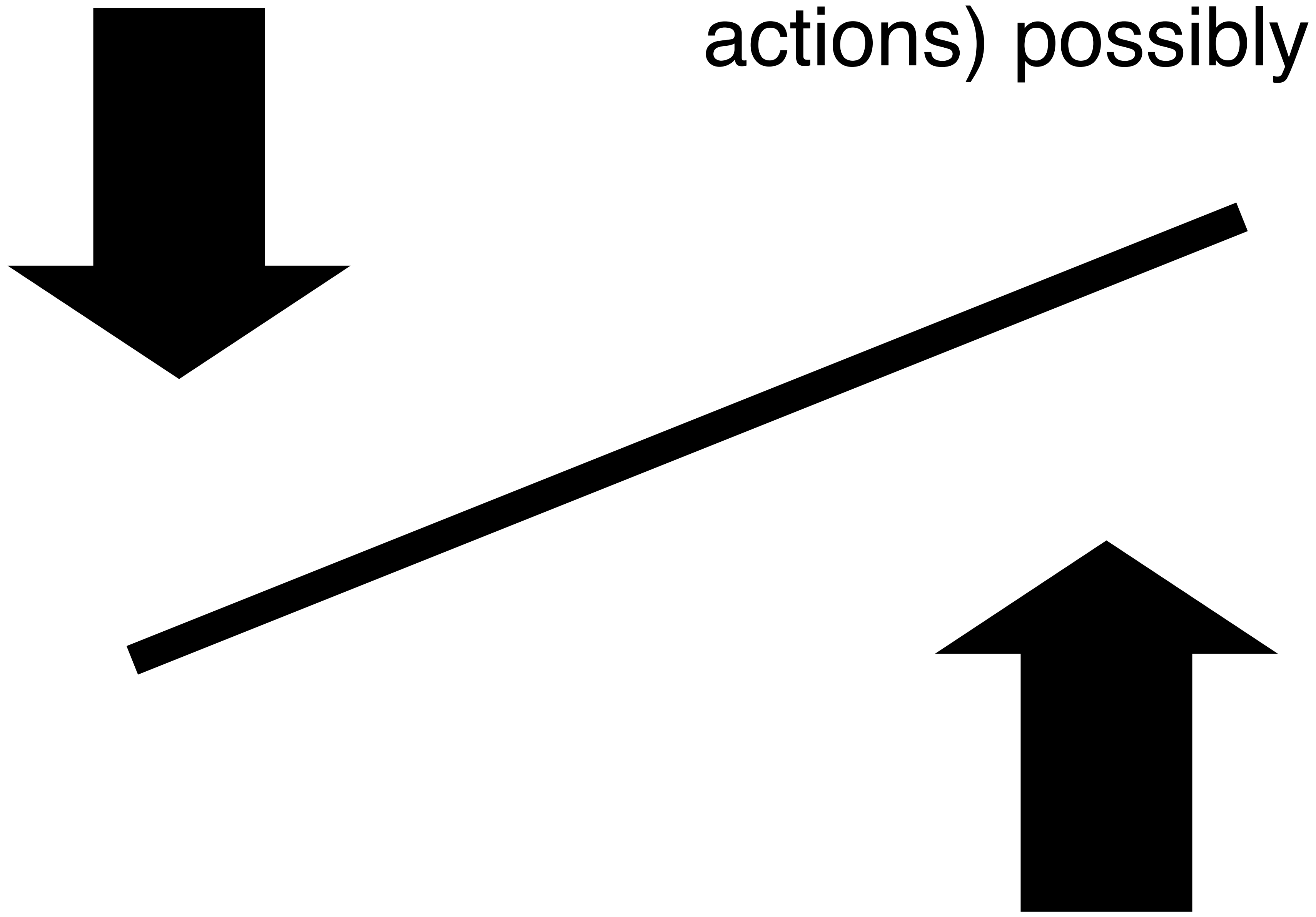
RL: History is the dataset!



* https://sandeep-bhuiya01.medium.com/disadvantages-of-cnn-models-95395fe9ae40

Exploration / Exploitation

Exploration: Improving knowledge about the problem (better understanding the impacts of actions) possibly at the cost of rewards



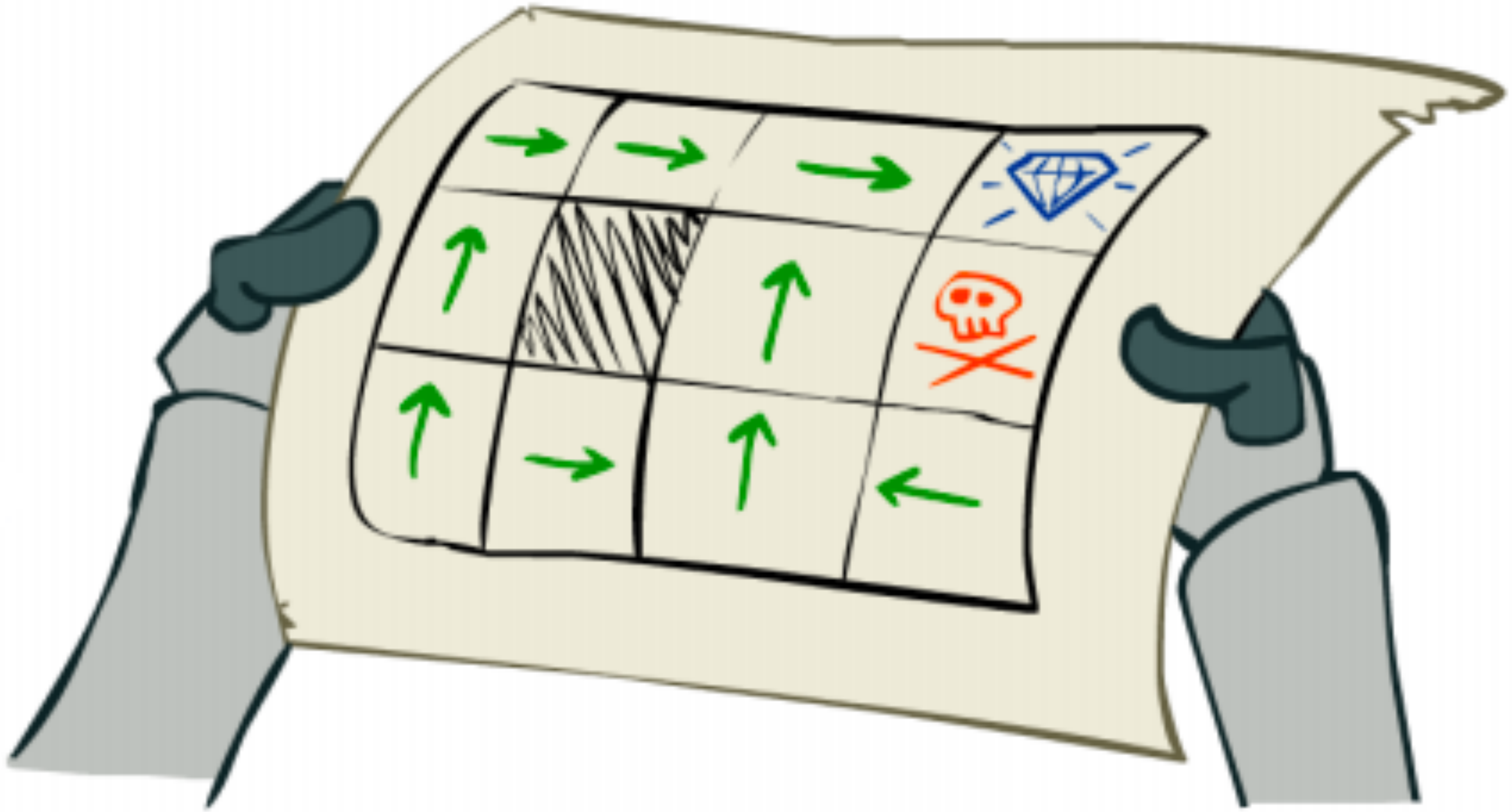
Exploitation: Using knowledge to maximize collected rewards

Example: Learning the optimal path

- Goal: Get from S to T as fast as possible
- Optimal path (unknown): Path 2
- Shortest path currently known: Path 5
- Exploitation: Follow Path 5
- Exploration: Deviate from Path 5
  → Could allow to discover Path 2

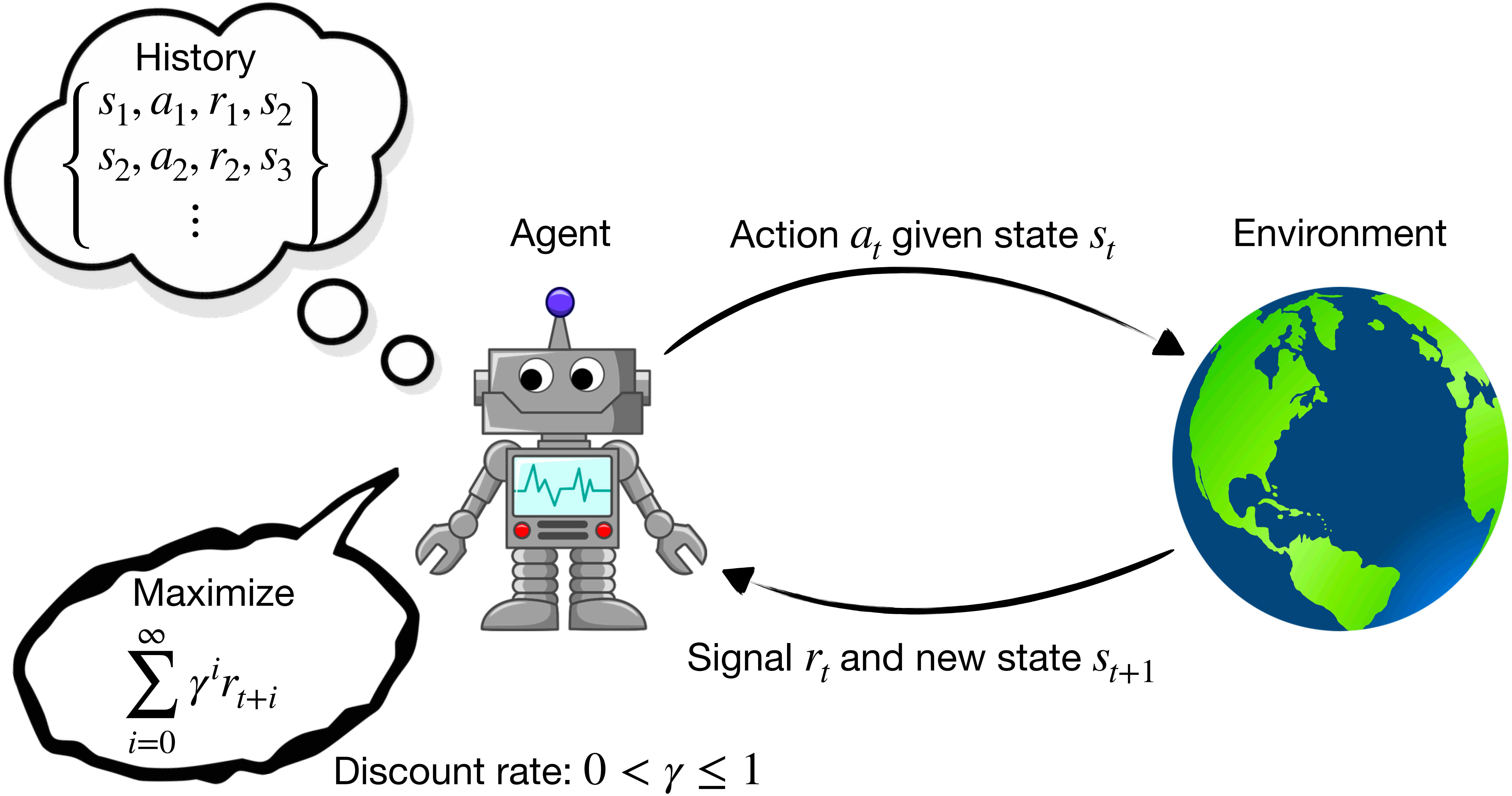
Galbrun, E., Pelechrinis, K., & Terzi, E. (2016). Urban navigation beyond shortest route: The case of safe paths. Information Systems, 57, 160-171.

How to achieve this with RL?

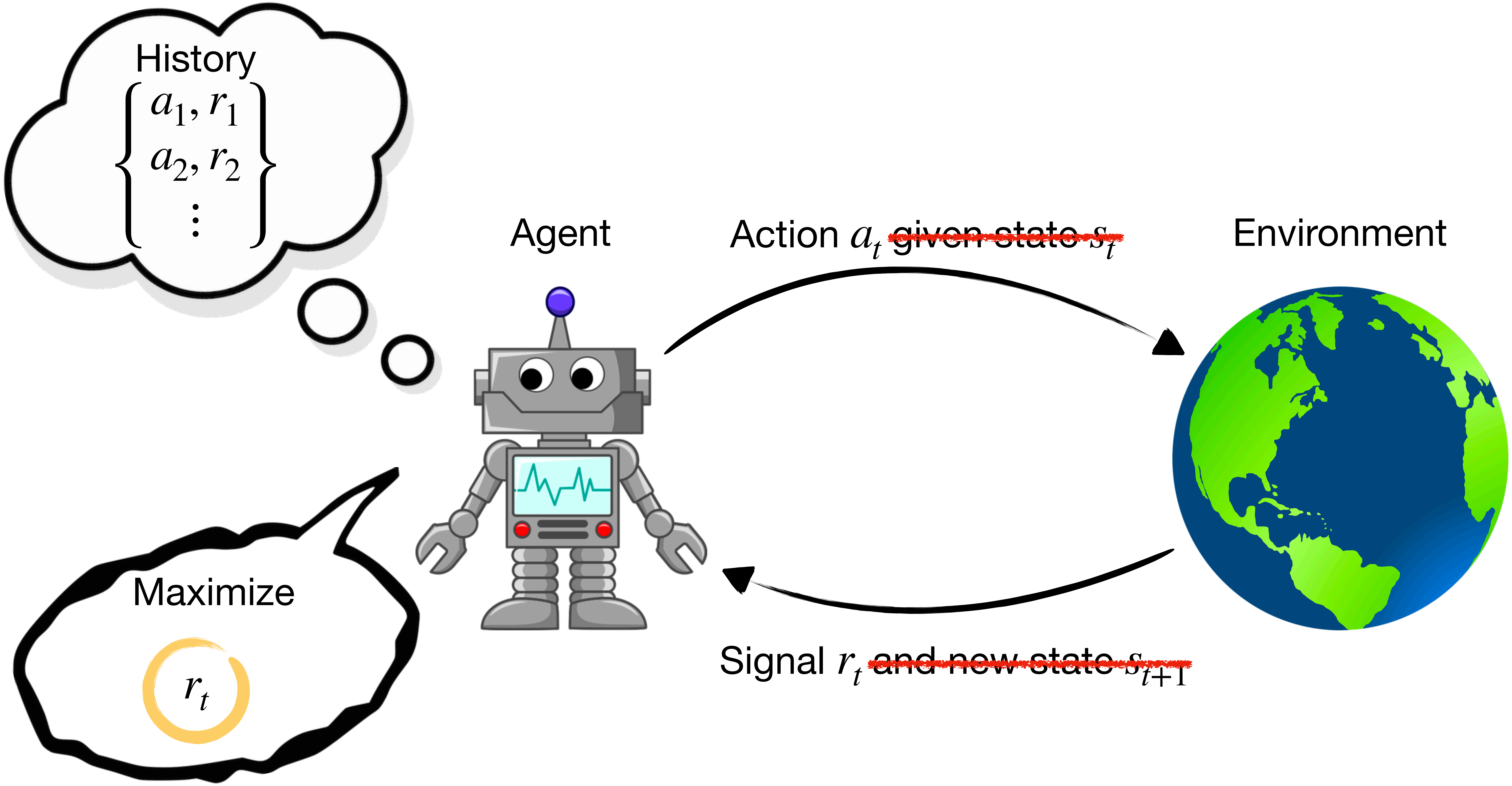


https://medium.com/@xaviergeerinck/the-markov-property-chain-reward-process-and-decision-process-4f63f7922401

Planning for the future

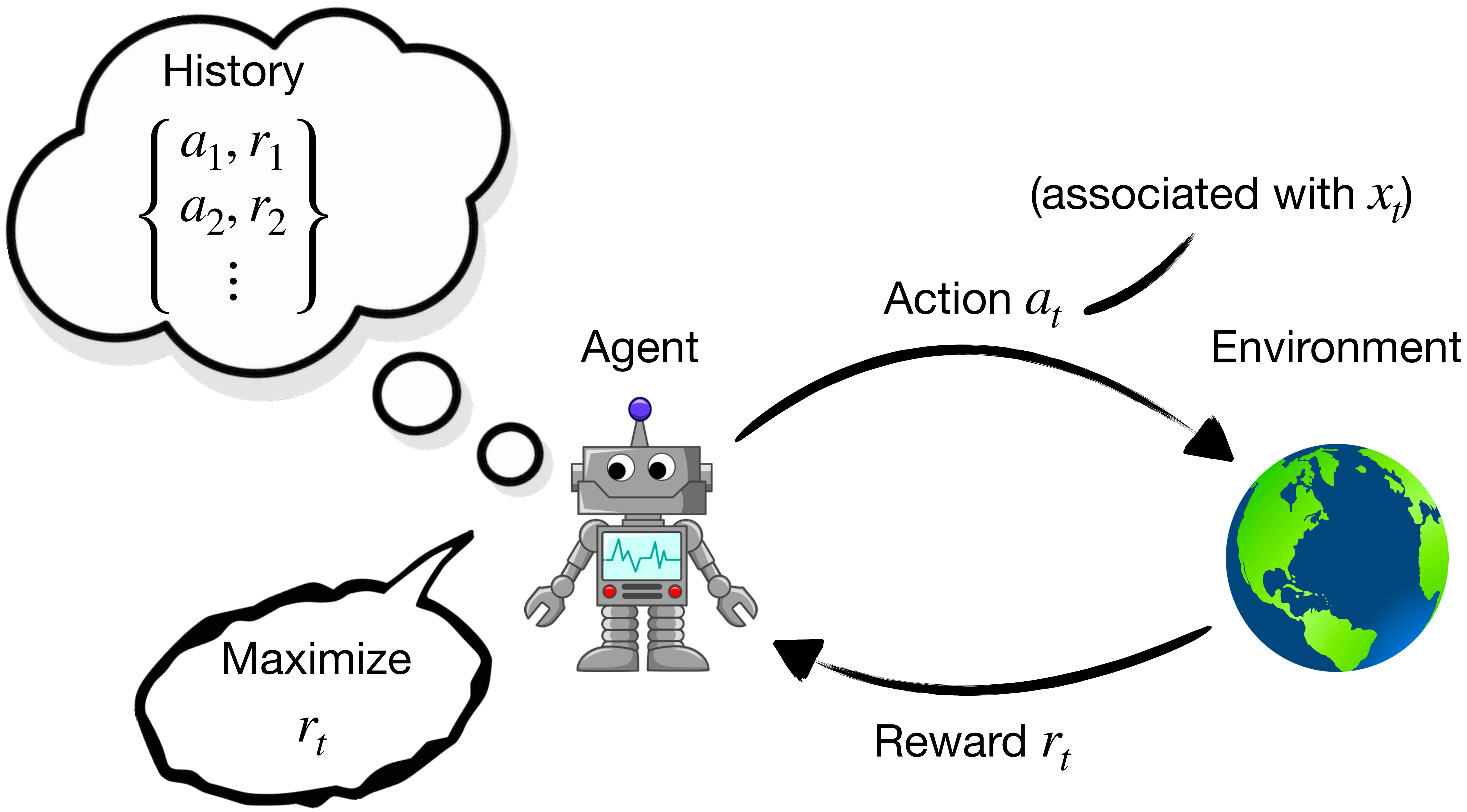


Bandits


History
$\left\{ \begin{matrix} a_1, r_1 \\ a_2, r_2 \\ \vdots \end{matrix} \right\}$
Maximize
r_t
Agent
Action a_t ~~given state s_t~~
Environment
Signal r_t ~~and new state s_{t+1}~~

Example: Structured bandits

- Action $a \in \mathscr{A}$ is associated with features $x \in \mathscr{X}$
- Simple case: Discrete set $\mathscr{X} \subset \mathbb{R}$
- Expected reward function: $f : \mathscr{X} \mapsto \mathbb{R}$
- Reward $r_t = f(x_t) + \varepsilon_t$ with $\varepsilon \sim \mathscr{N}(0,\sigma^2)$
- Assumption: Actions nearby in $\mathscr{X}$ have similar expected reward
- Assumption: $f(x) = \langle \phi(x), \theta \rangle$ with unknown θ and known $\phi(\cdot)$



Kernel regression

- Kernel $k(x, x') = \langle \phi(x), \phi(x') \rangle$
- Gaussian prior $\theta \sim \mathcal{N}_d(0,\Sigma)$ with $\Sigma = \dfrac{\sigma^2}{\lambda} I$ for $\lambda > 0$

$\mathbf{K}_N = [k(x_i, x_j)]_{1 \leq i,j \leq N}$ and $\mathbf{k}_N(x) = \big(k(x, x_i)\big)_{1 \leq i,j \leq N}$

$$\mathbb{P}[f \,|\, x_1, \ldots, x_N, y_1, \ldots, y_N] \sim \mathcal{N}\left(\big(f(x)_{x \in \mathcal{X}}\big), [k_N(x, x')]_{x,x' \in \mathcal{X}}\right)$$

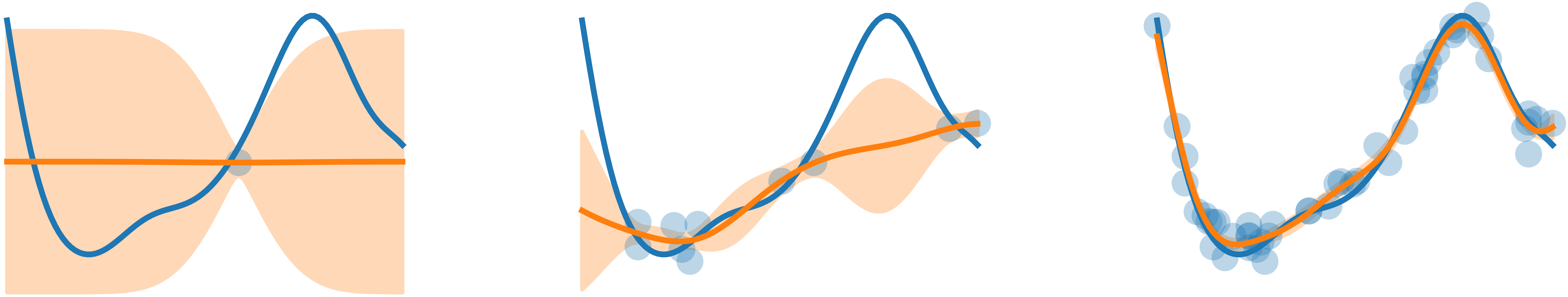
Posterior mean: $f_N(x) = \mathbf{k}_N(x)^\top (\mathbf{K}_N + \lambda I)^{-1} \mathbf{y}_N$

Posterior covariance: $k_N(x, x') = k(x, x') - \mathbf{k}_N(x)^\top (\mathbf{K}_N + \lambda I)^{-1} \mathbf{k}_N(x')$

Kernel regression

- Kernel $k(x, x') = \langle \phi(x), \phi(x') \rangle$
- Gaussian prior $\theta \sim \mathcal{N}_d(0,\Sigma)$ with $\Sigma = \dfrac{\sigma^2}{\lambda} I$ for $\lambda > 0$

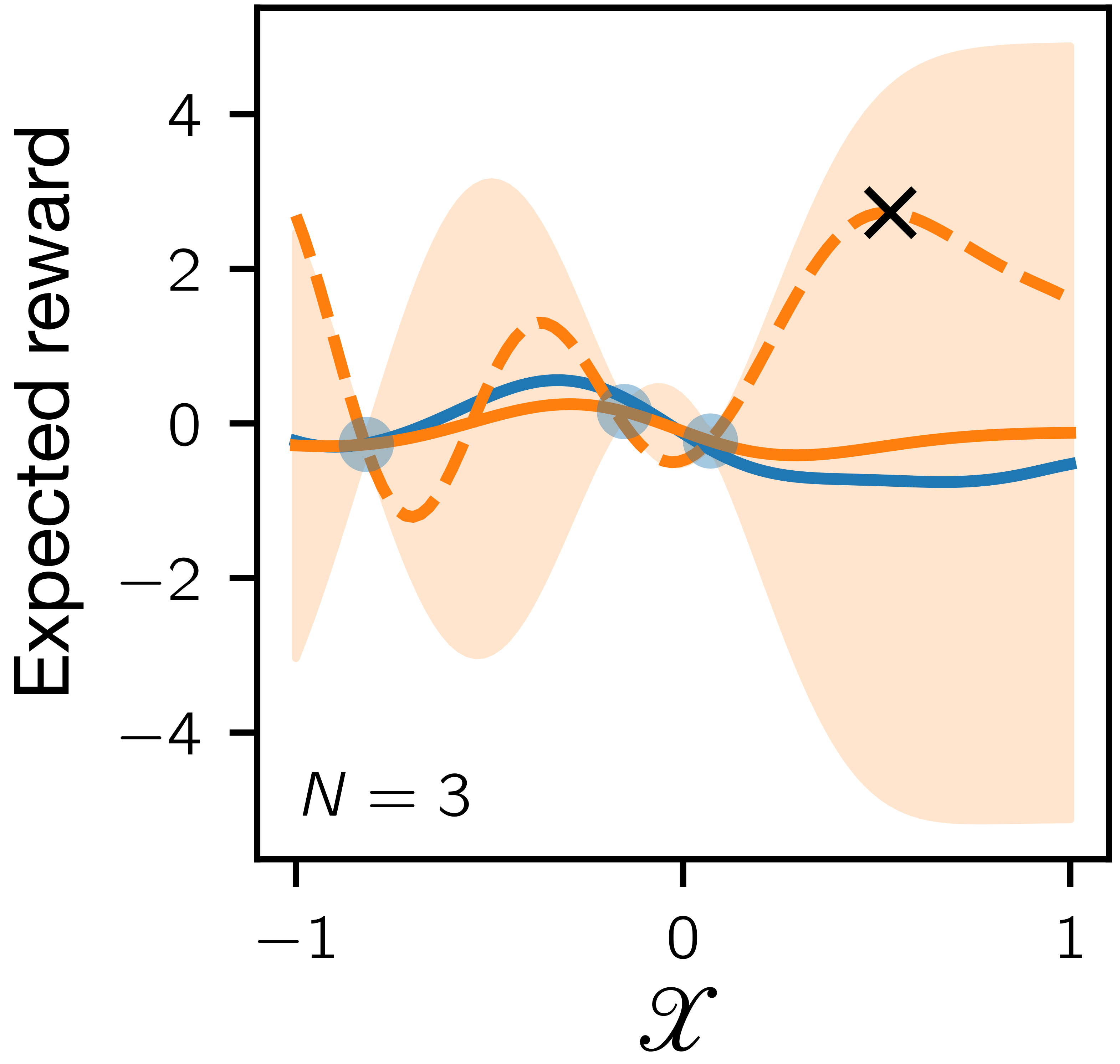
Example: Posterior distributions after 1, 10, and 100 observations (x_t, r_t)



Kernel regression + Thompson Sampling (TS)

Example: Decision making at $t = 4$

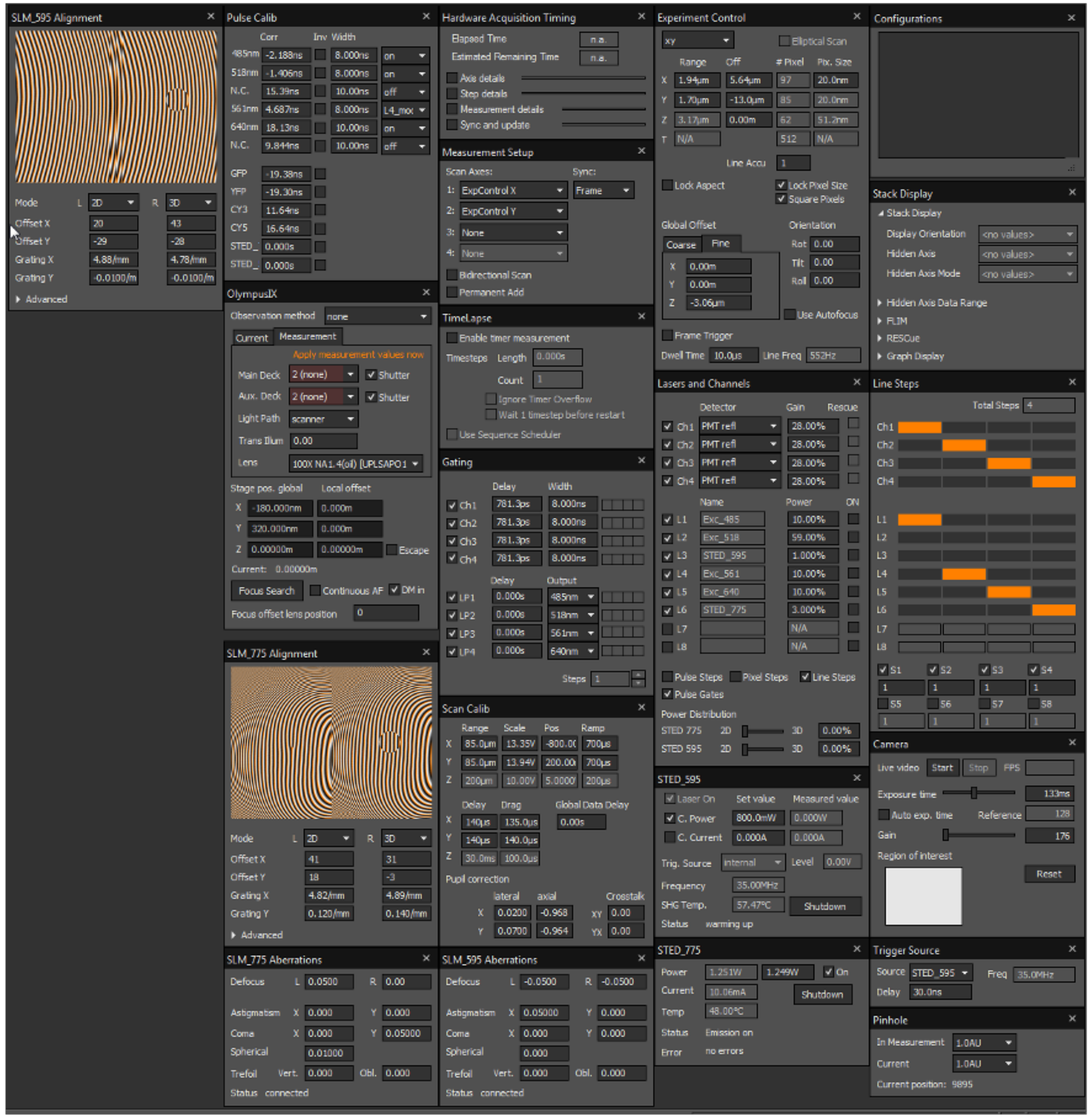
- Compute posterior distribution given previous observations
- Sample function $\tilde{f}$ from the posterior
- Select $x_t = \arg\max_{x \in \mathcal{X}} \tilde{f}(x)$



STED Microscopy Optimization

Flavie Lavoie-Cardinal et Theresa Wiesner

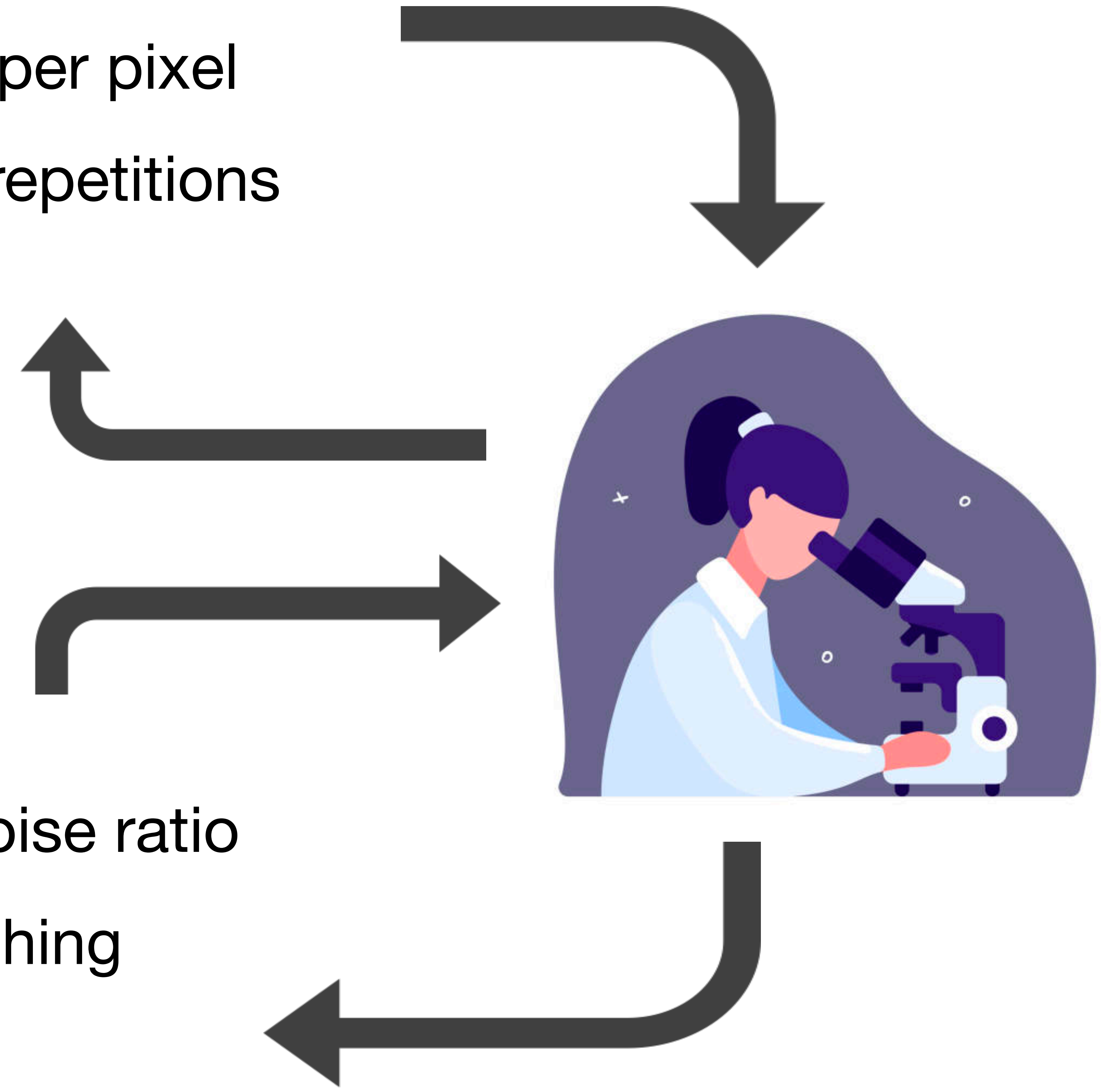
Difficult to configure to acquire good images



Imspector Software – Abberior Instruments

Parameters:

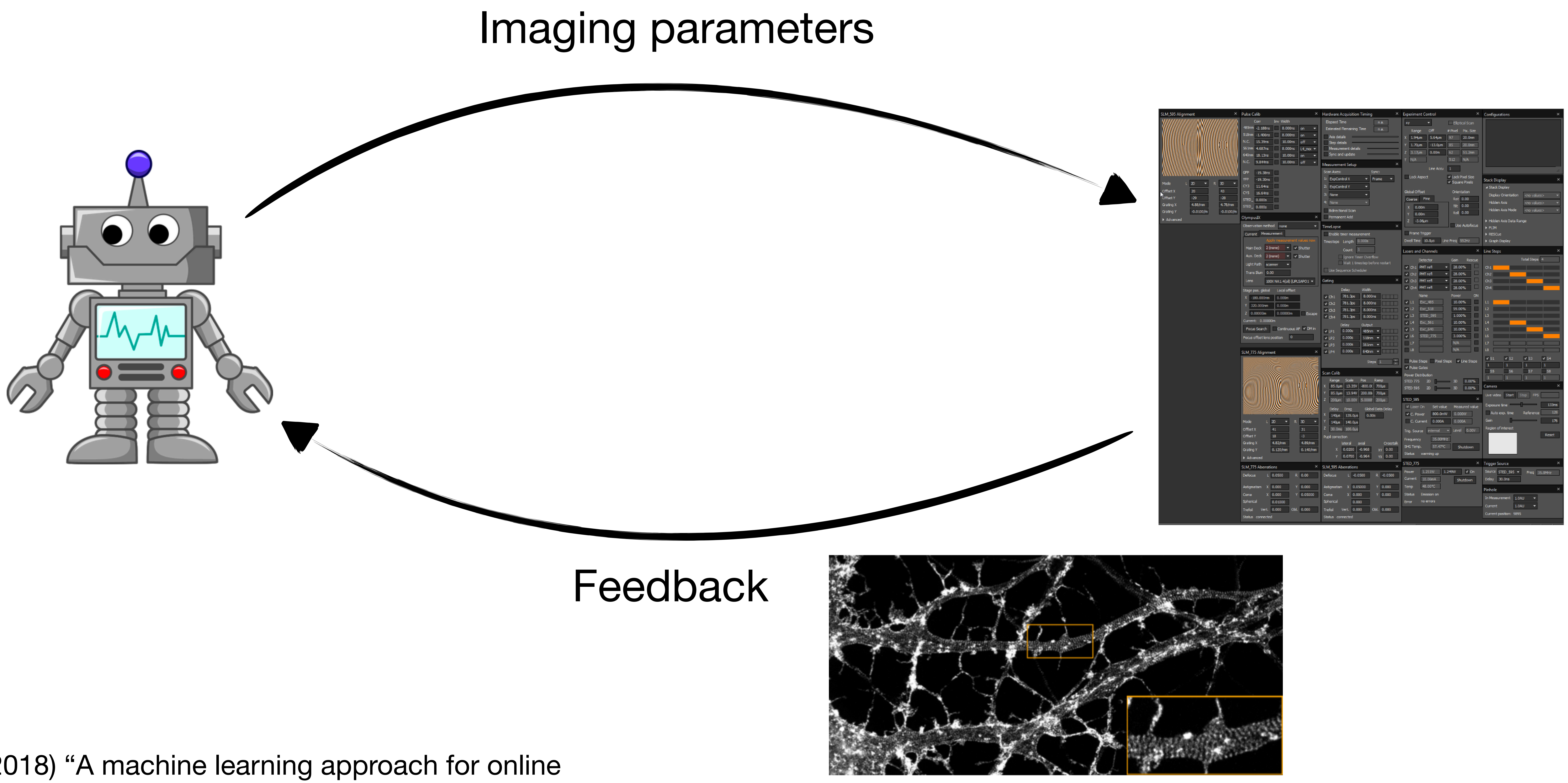
- Power of lasers
- Time spent per pixel
- Number or repetitions
- …



Objectives:

- Signal to noise ratio
- Photobleaching
- Resolution
- Quality
- …

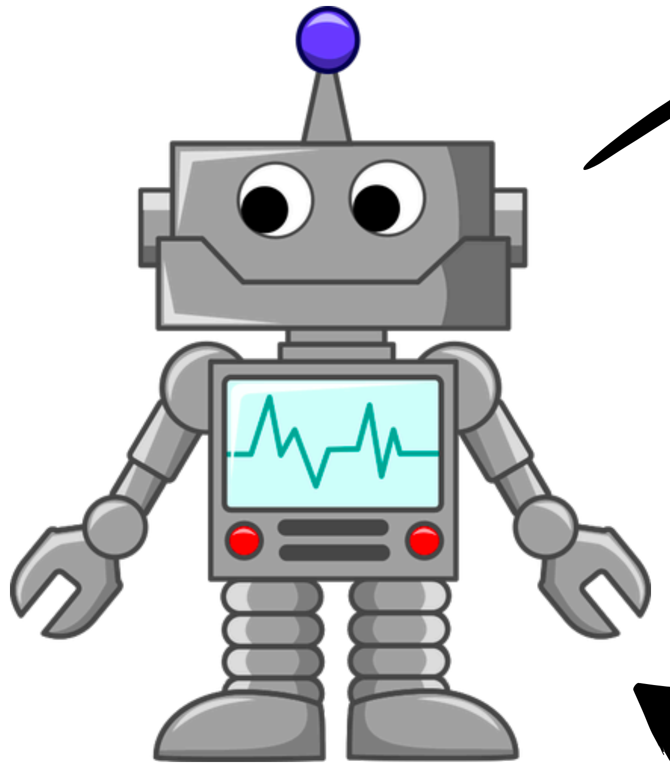
Bandit problem!

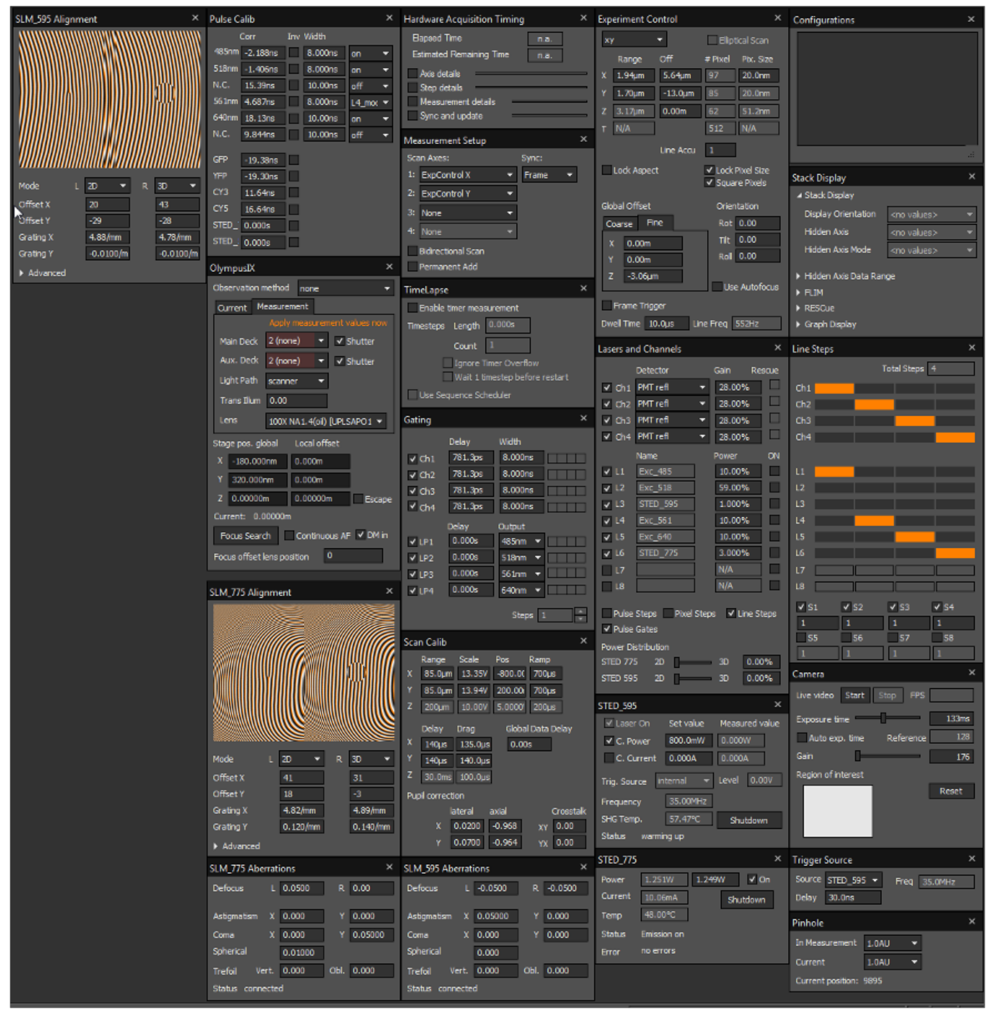


Durand et al. (2018) “A machine learning approach for online automated optimization of super-resolution optical microscopy”. *Nature Communications.*

Bandit problem!

Imaging parameters





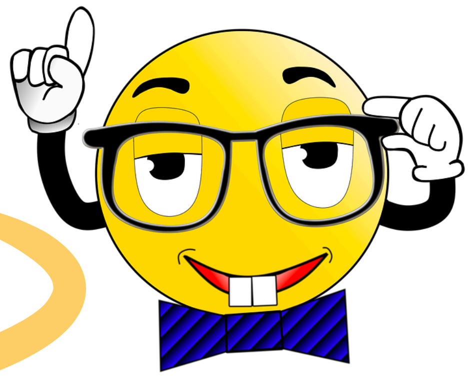
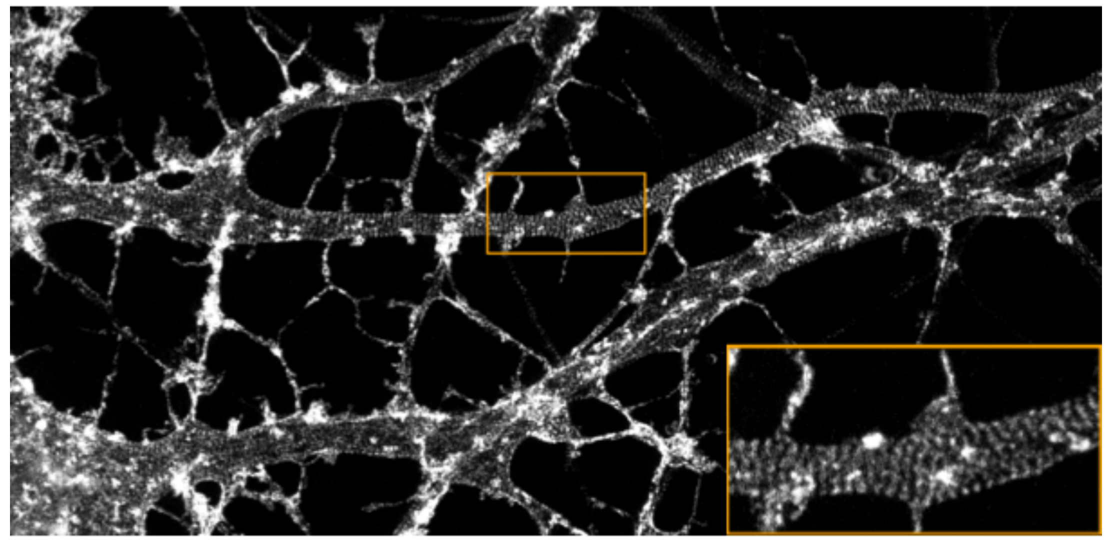


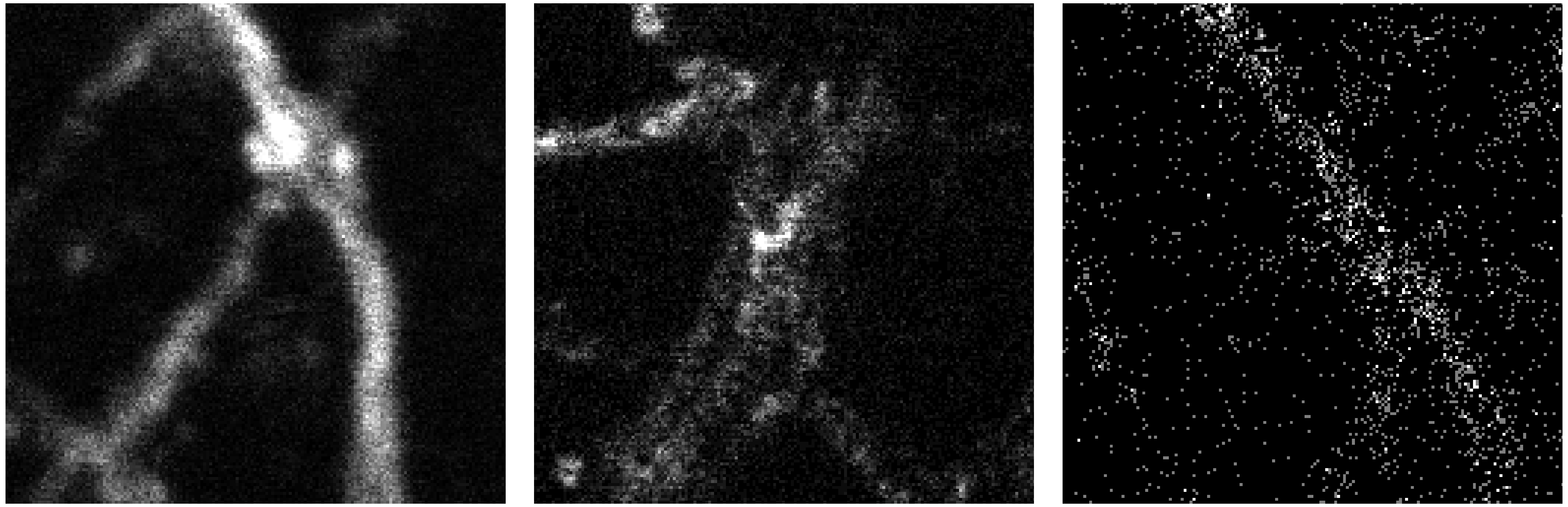
Image quality score



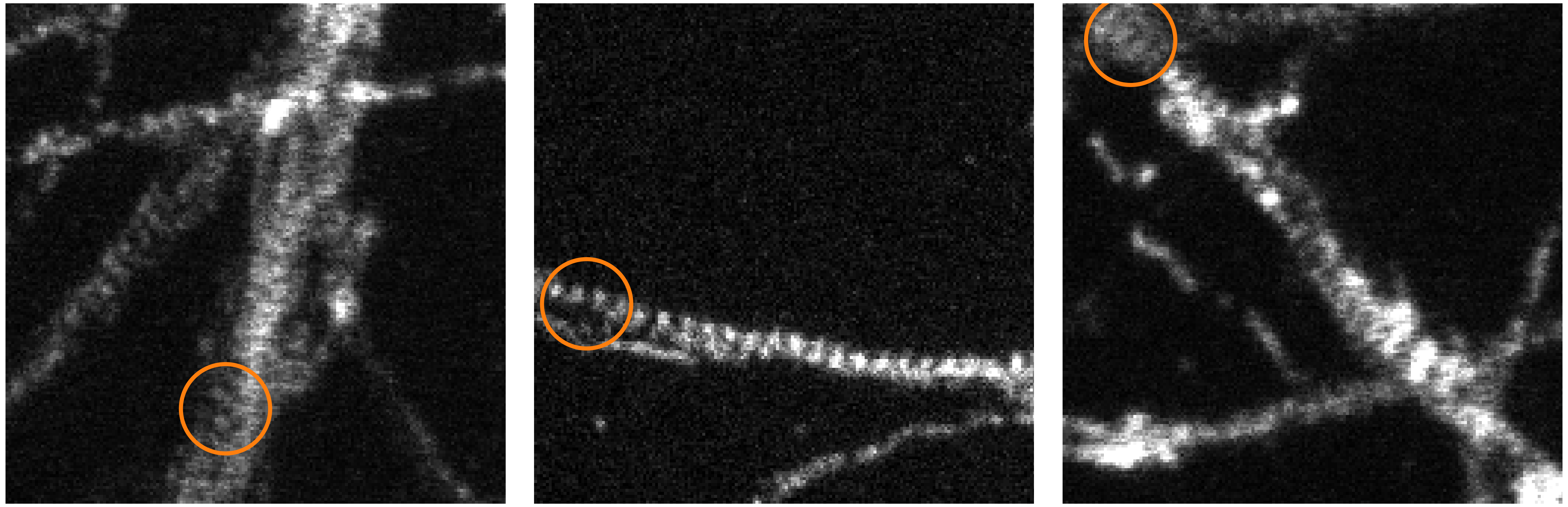
Durand et al. (2018) "A machine learning approach for online automated optimization of super-resolution optical microscopy". *Nature Communications.*

What is *good* image quality?

Avoiding images likes these:

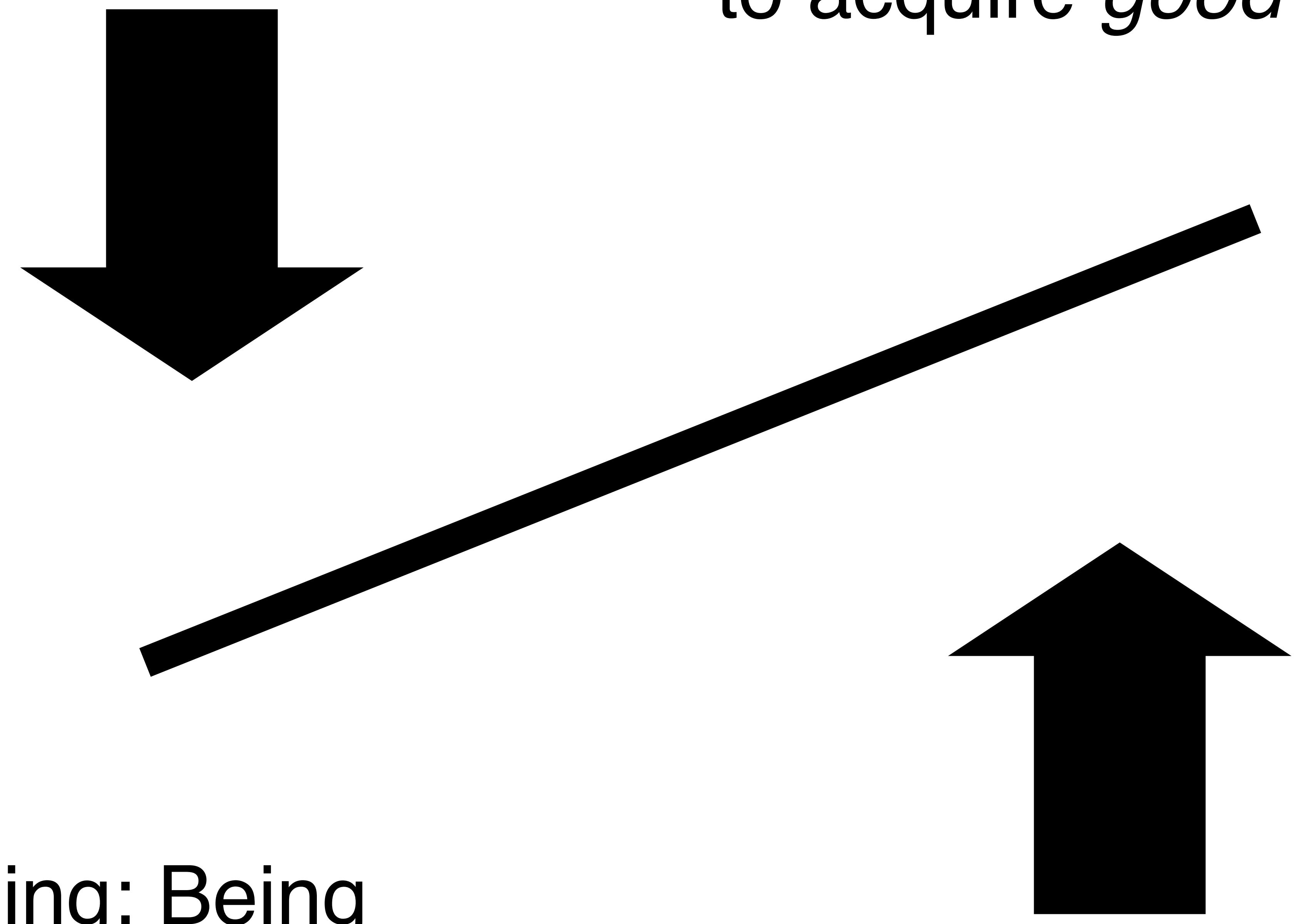


Acquiring more images likes these:



Multiple objectives to consider

Maximize image quality: Being able to acquire *good* images



Minimize photobleaching: Being able to acquire *multiple* images

Optimizing multiple objectives

Imaging parameters

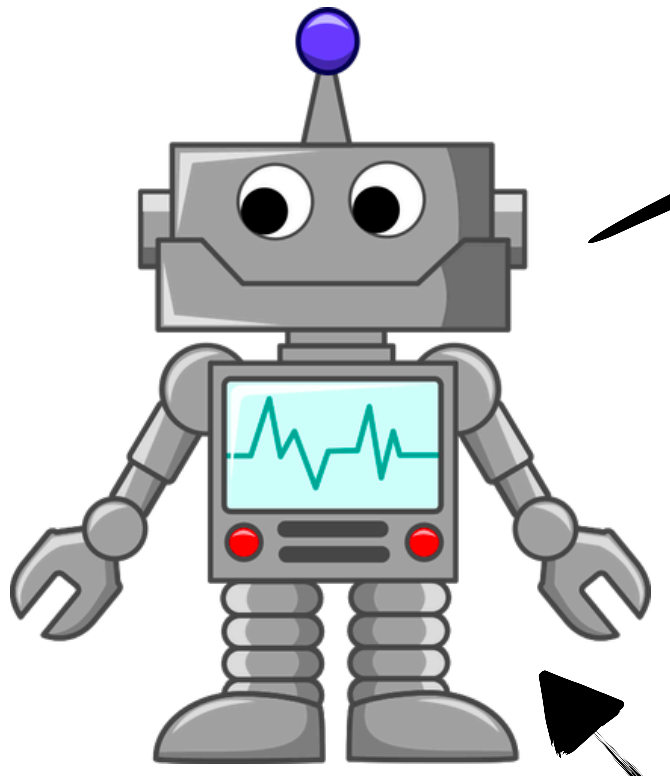
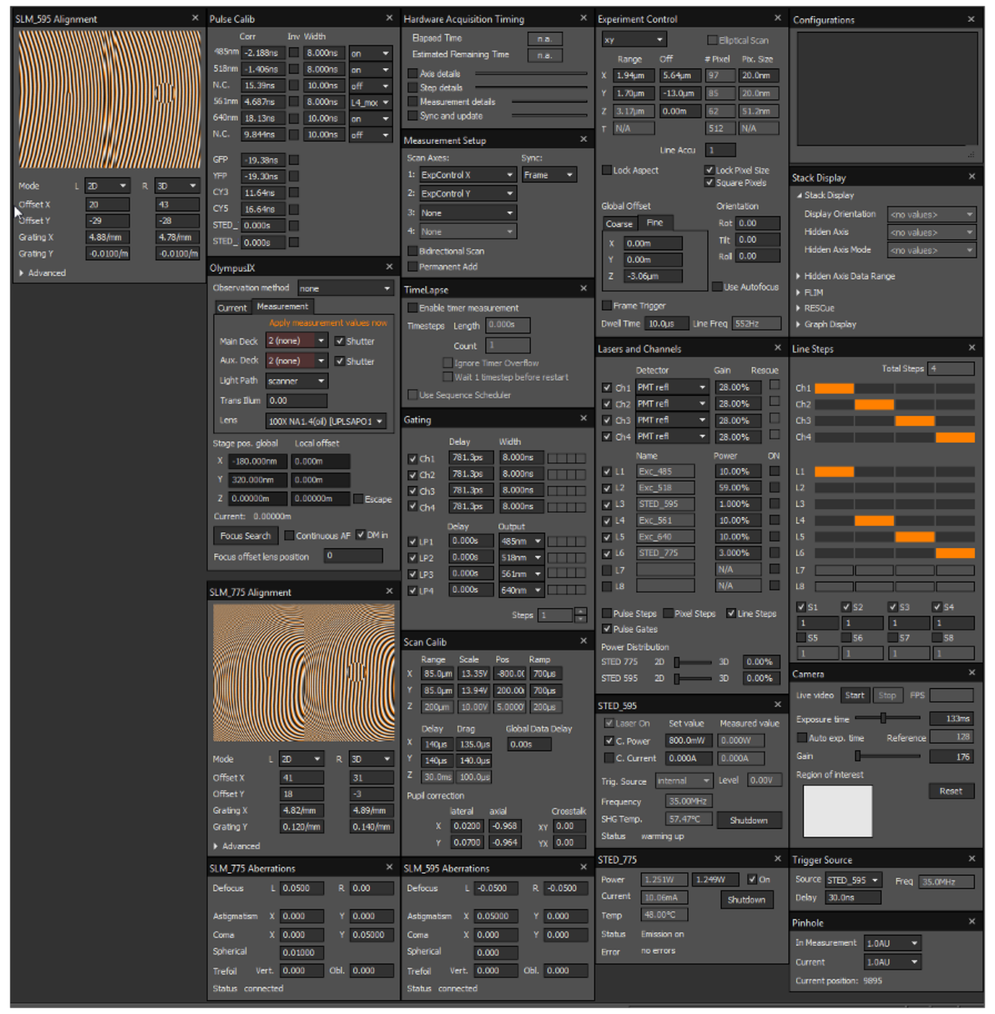
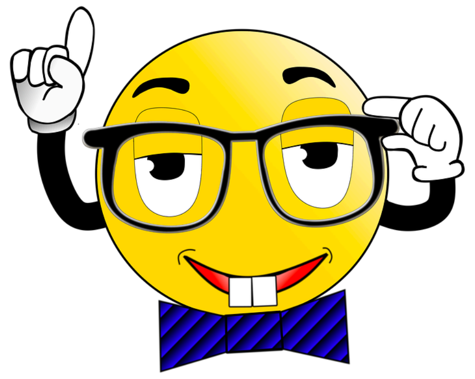
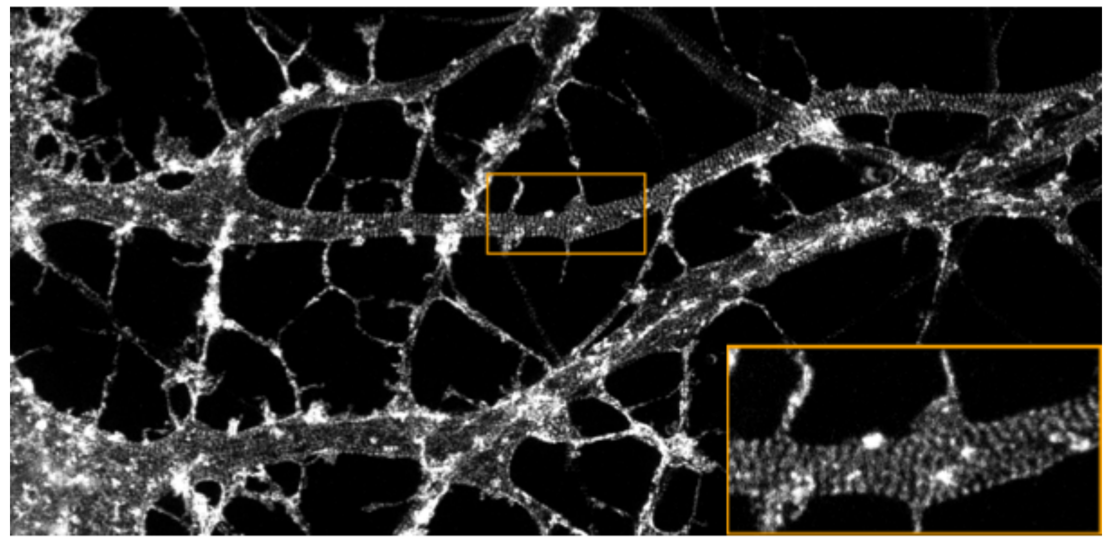


Image quality score $r_t^{(1)}$





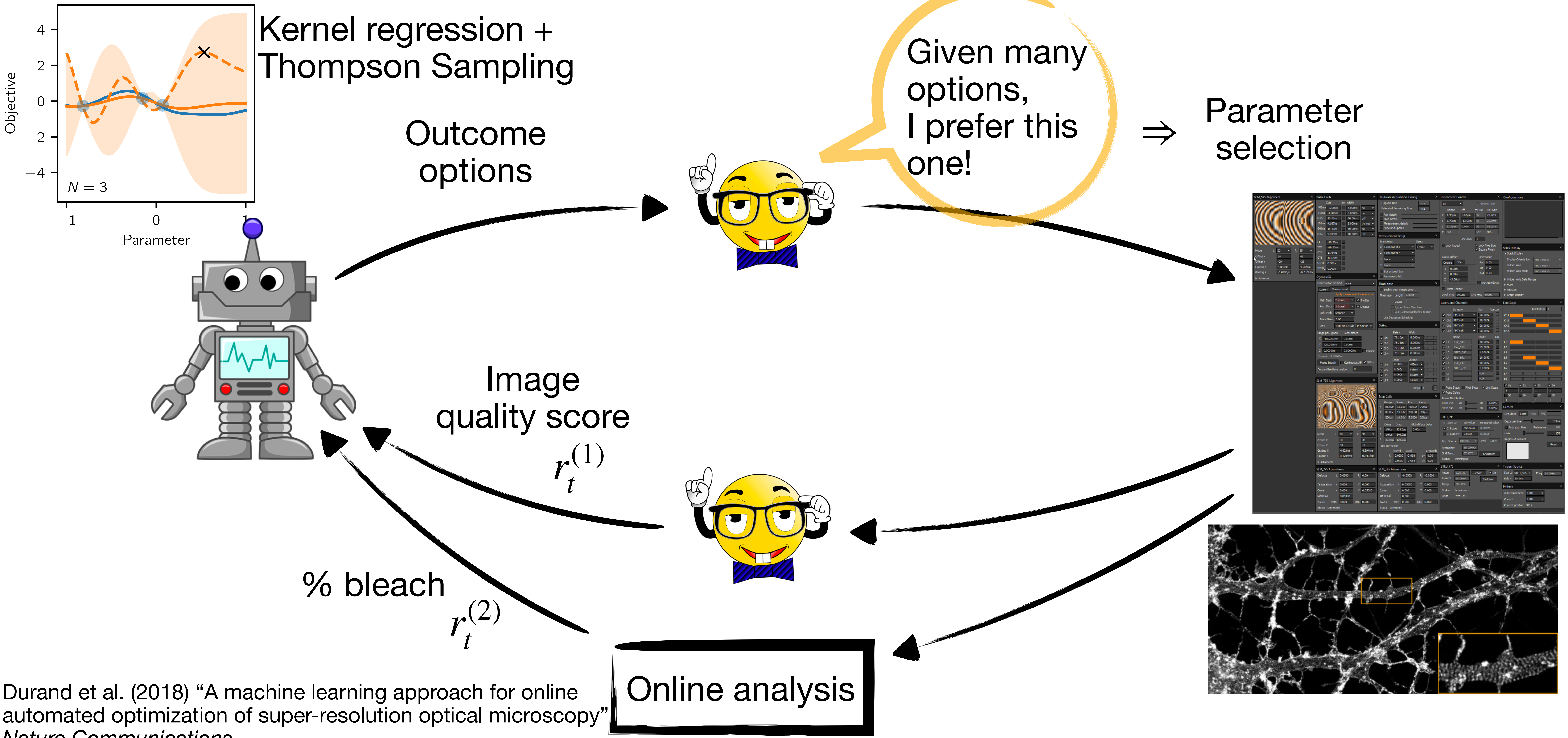


% bleach $r_t^{(2)}$

Online analysis

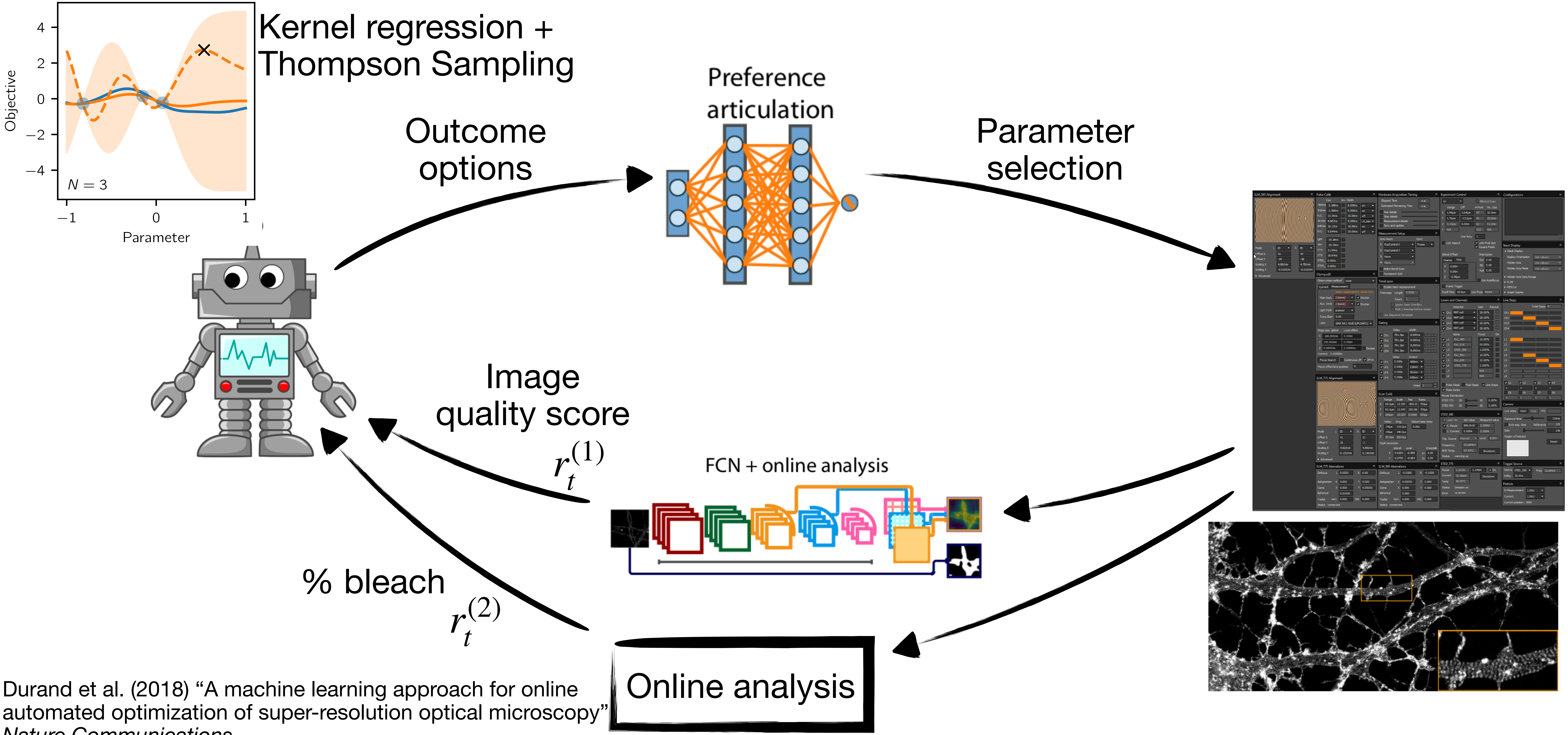
Durand et al. (2018) “A machine learning approach for online automated optimization of super-resolution optical microscopy” *Nature Communications.*

Optimizing multiple objectives



Durand et al. (2018) “A machine learning approach for online automated optimization of super-resolution optical microscopy” *Nature Communications.*

Automated multi-objective optimization

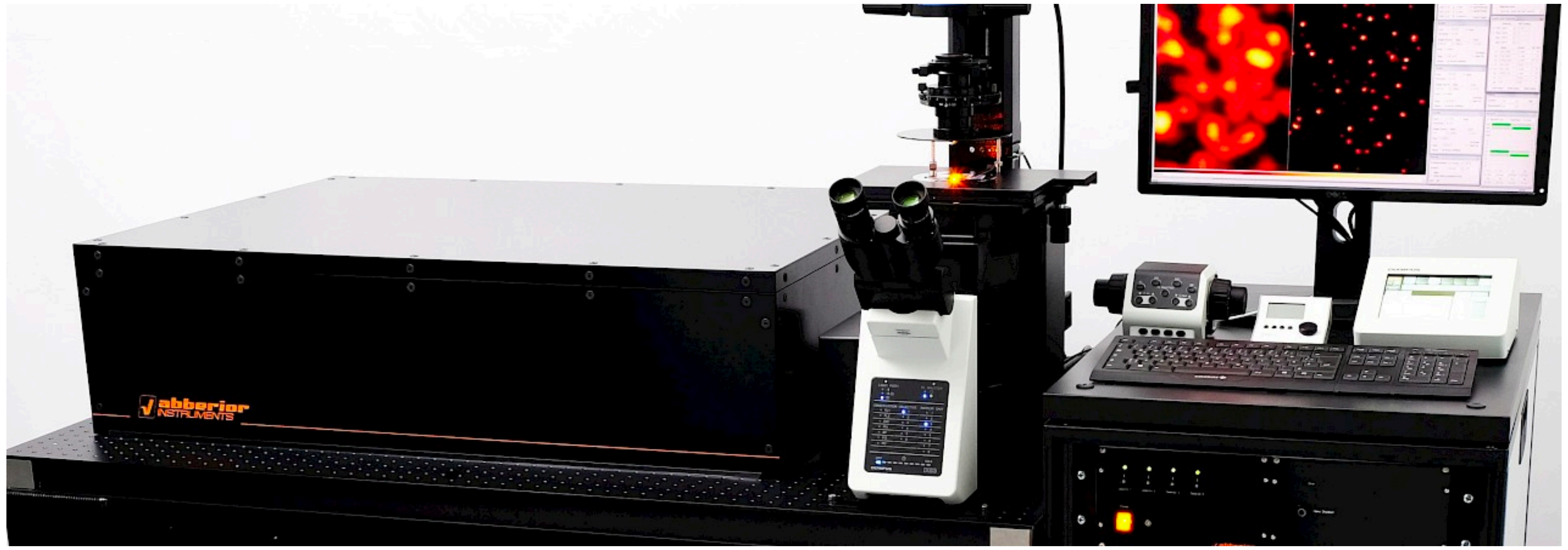


Durand et al. (2018) “A machine learning approach for online automated optimization of super-resolution optical microscopy” *Nature Communications.*

Experiments

Three parameters (1000 configurations):

- Excitation laser power
- Depletion laser power
- Duration of imaging per pixel



From abberior-instruments.com

Acquire:

- Confocal (low resolution image)
- First STED
- Second STED

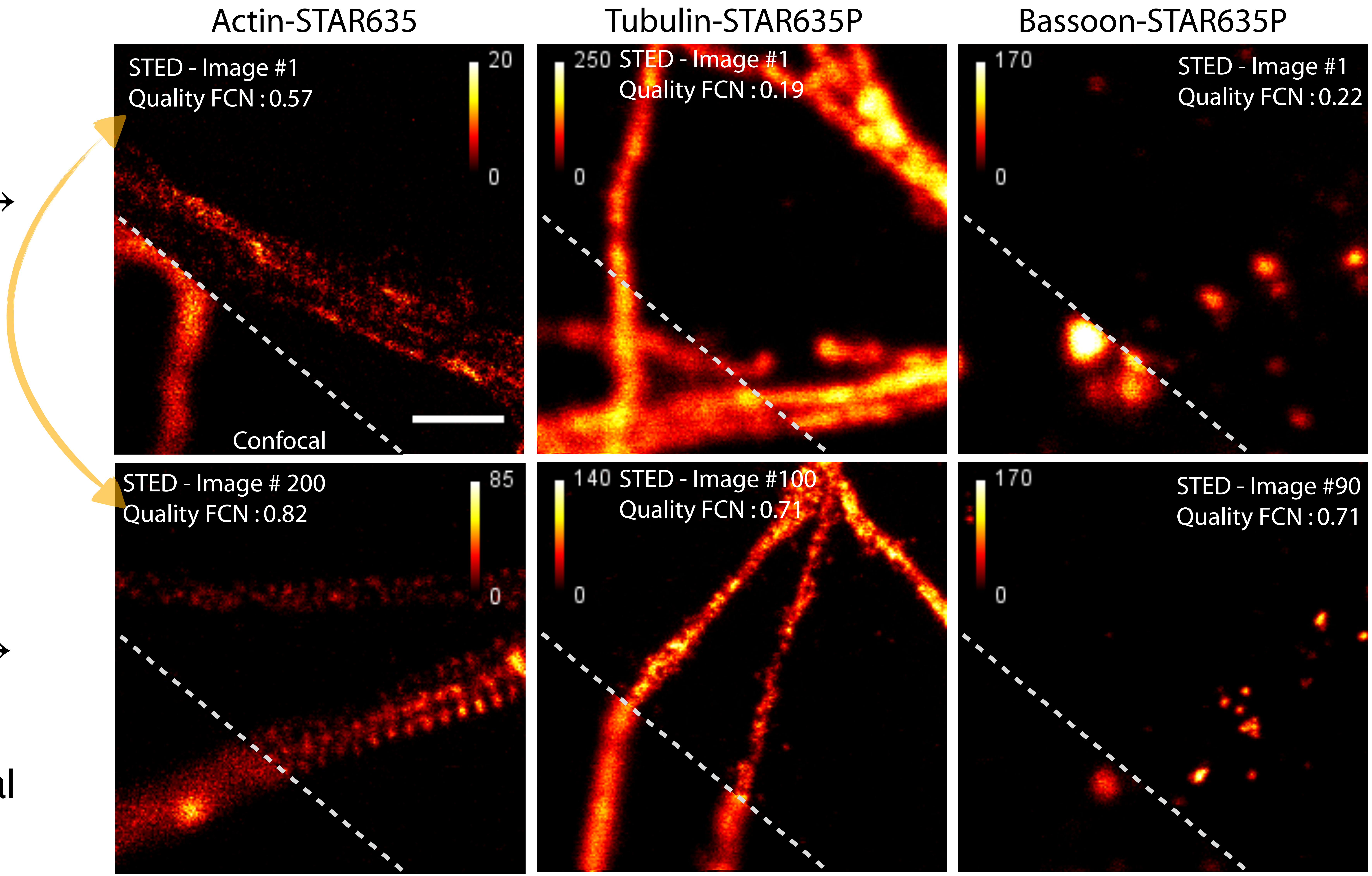
Goal: ↑ 1st STED quality and ↓ photobleaching

Fully automated imaging

Beginning of optimization →

End of optimization →

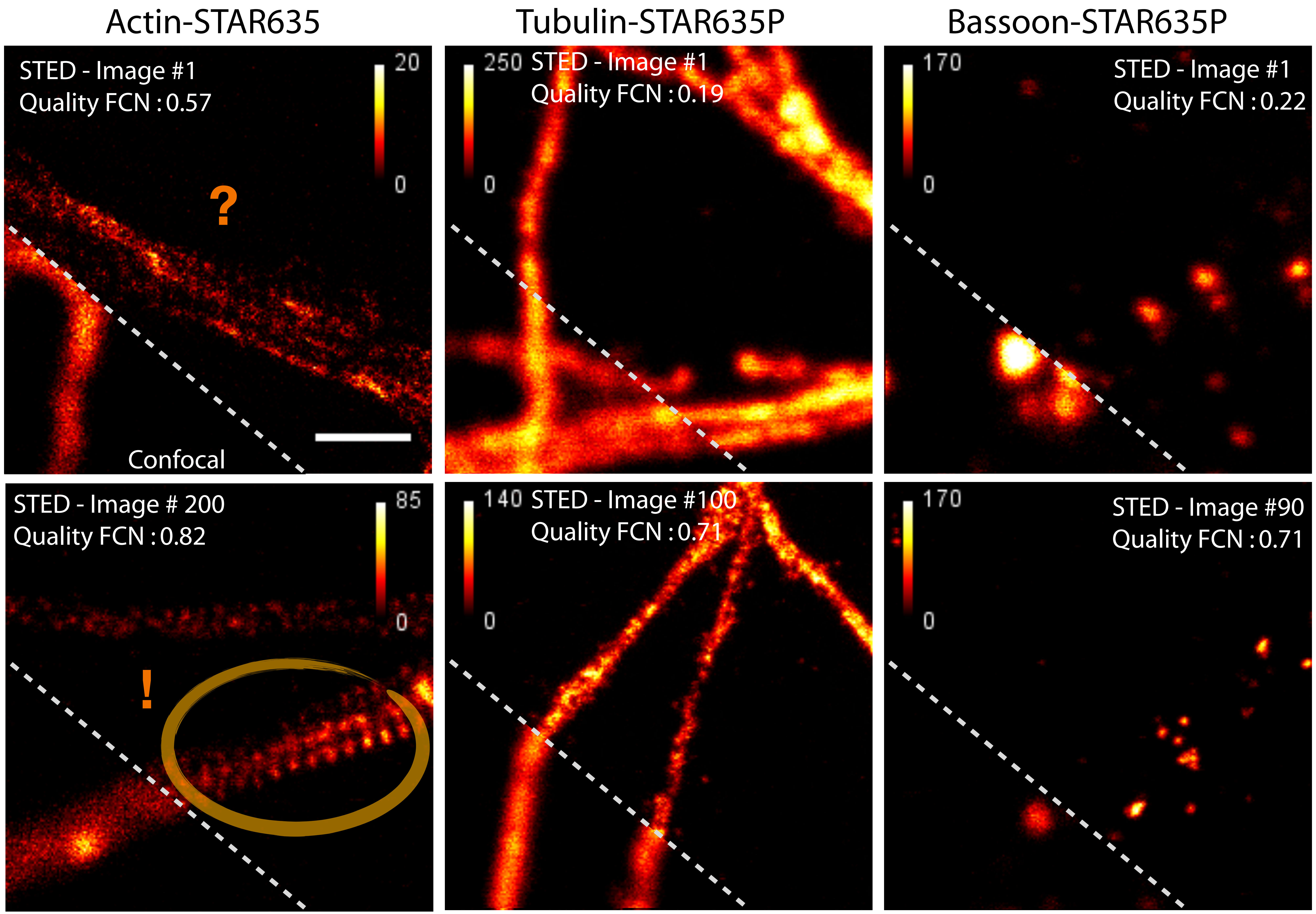
Bottom left corner: Confocal
<u>Not</u> super-resolution



Fully automated imaging

Beginning of optimization →

End of optimization →



Bottom left corner: Confocal

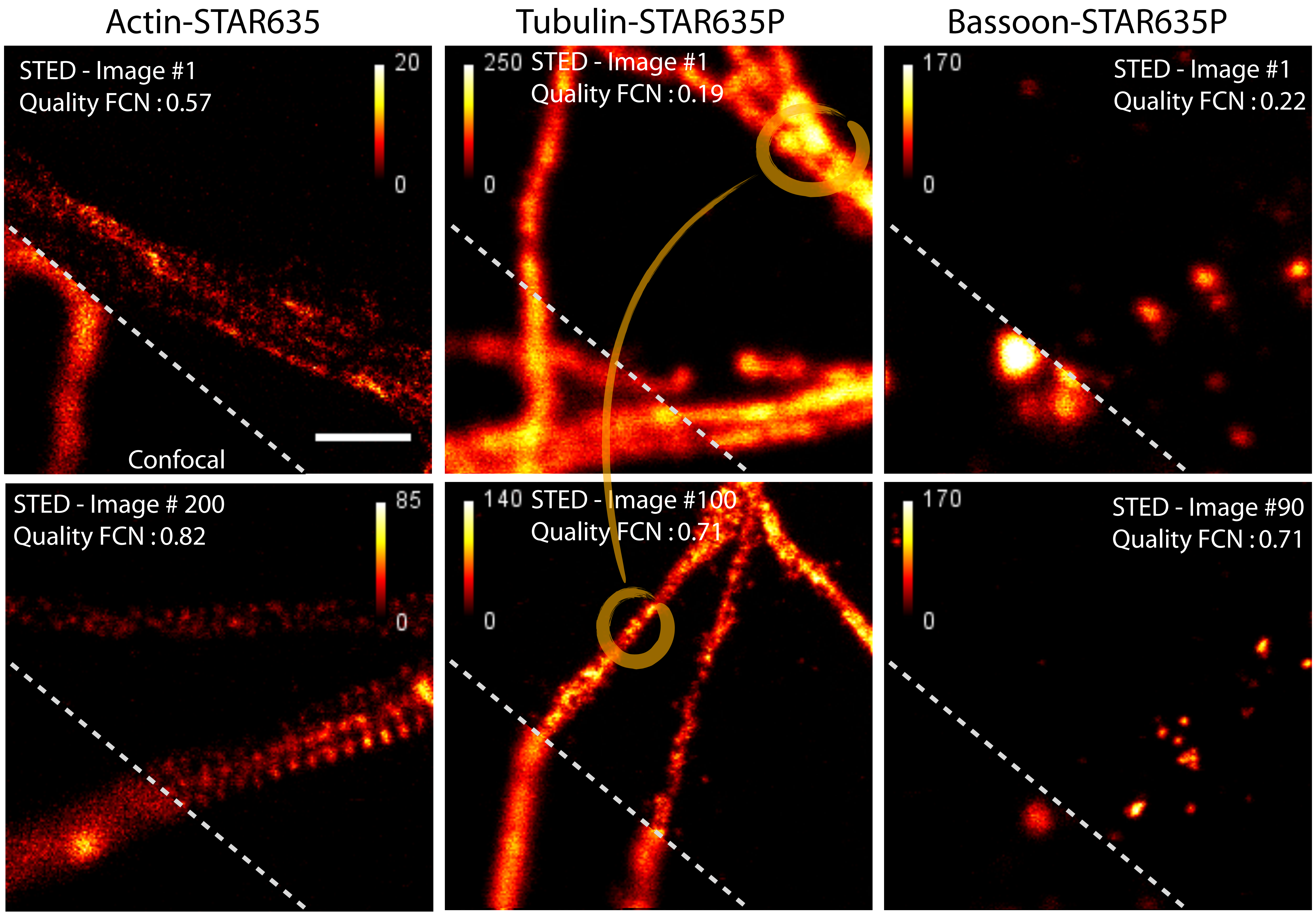
Not super-resolution

Fully automated imaging

Beginning of optimization →

End of optimization →

Bottom left corner: Confocal
<u>Not</u> super-resolution

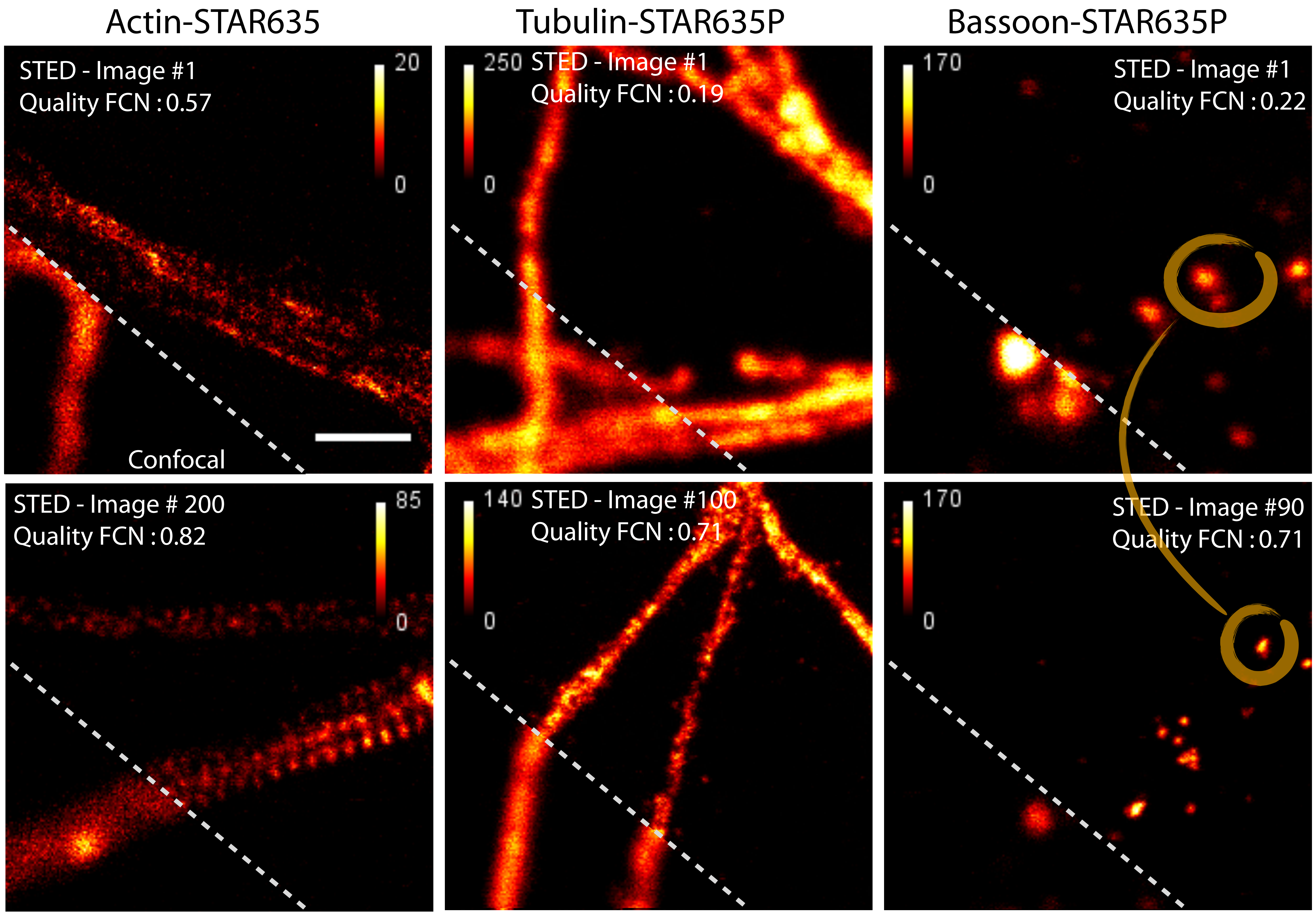


Fully automated imaging

Beginning of optimization →

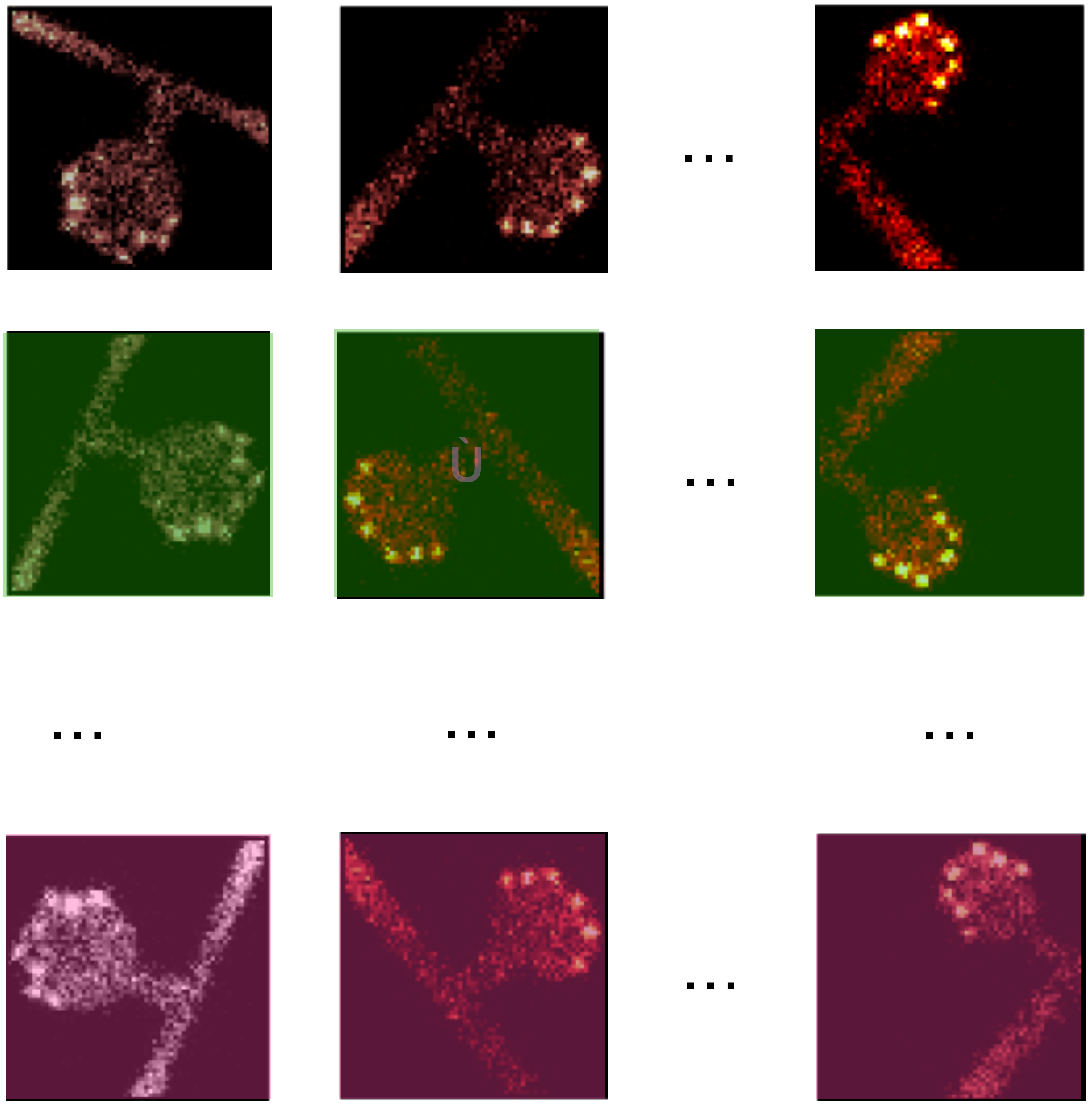
End of optimization →

Bottom left corner: Confocal
<u>Not</u> super-resolution



Bandits formulation

Steps $t = 1,2,\ldots$



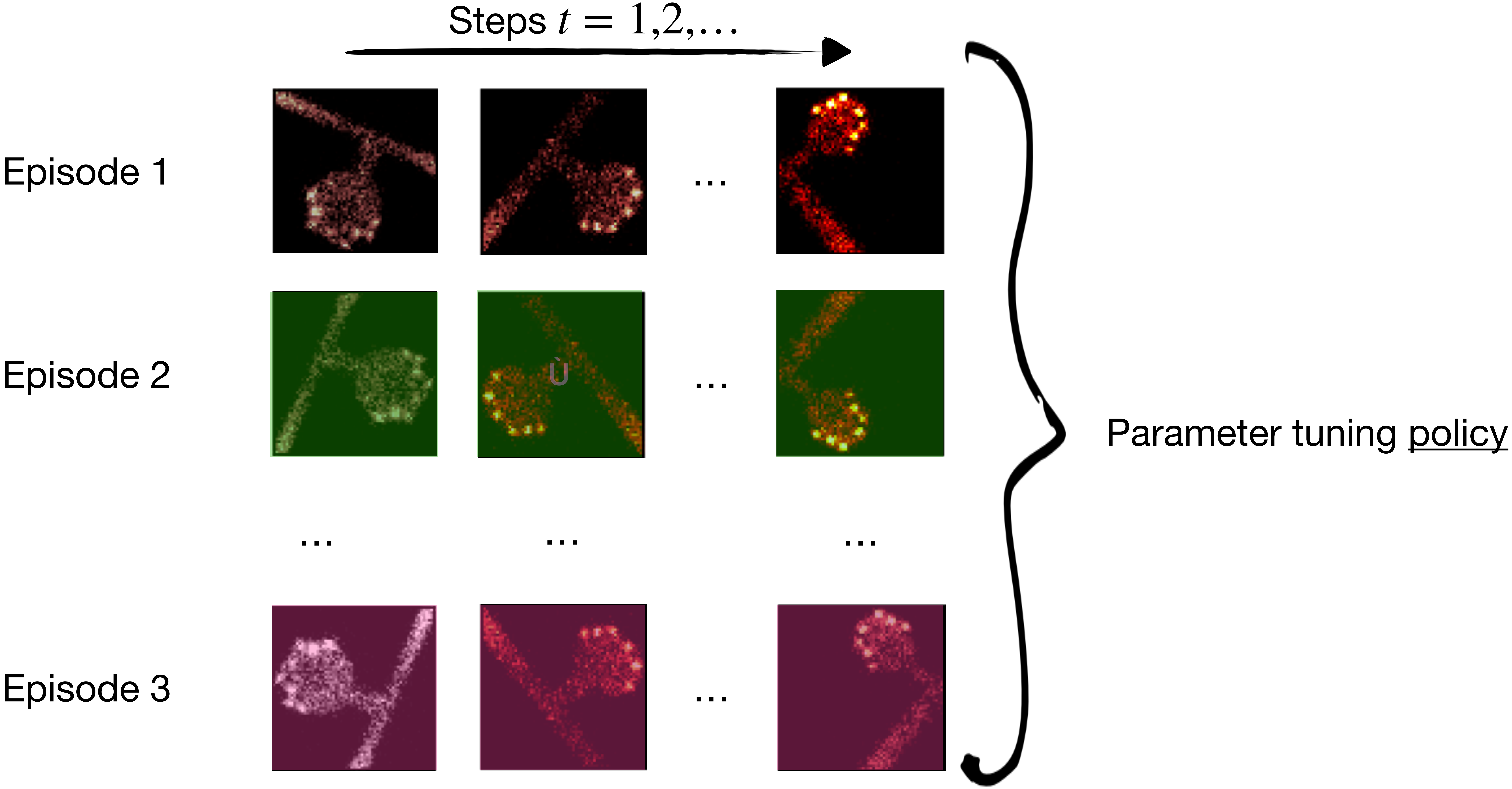
New problem: … → Good parameters for *this* sample

New problem: … → Good parameters for *this* sample

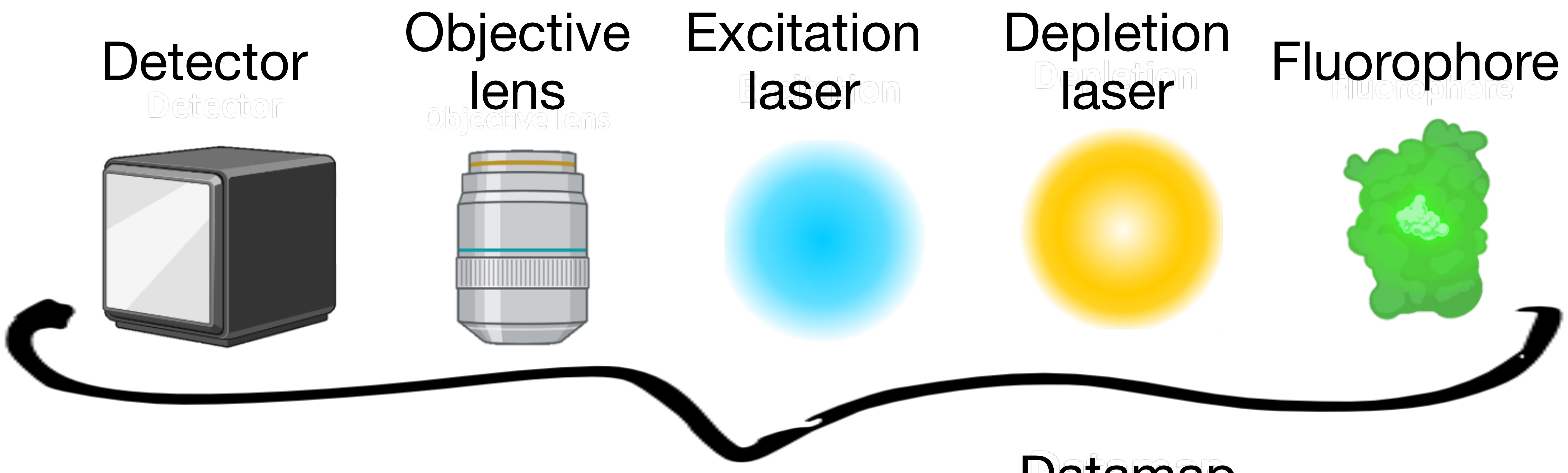
… … …

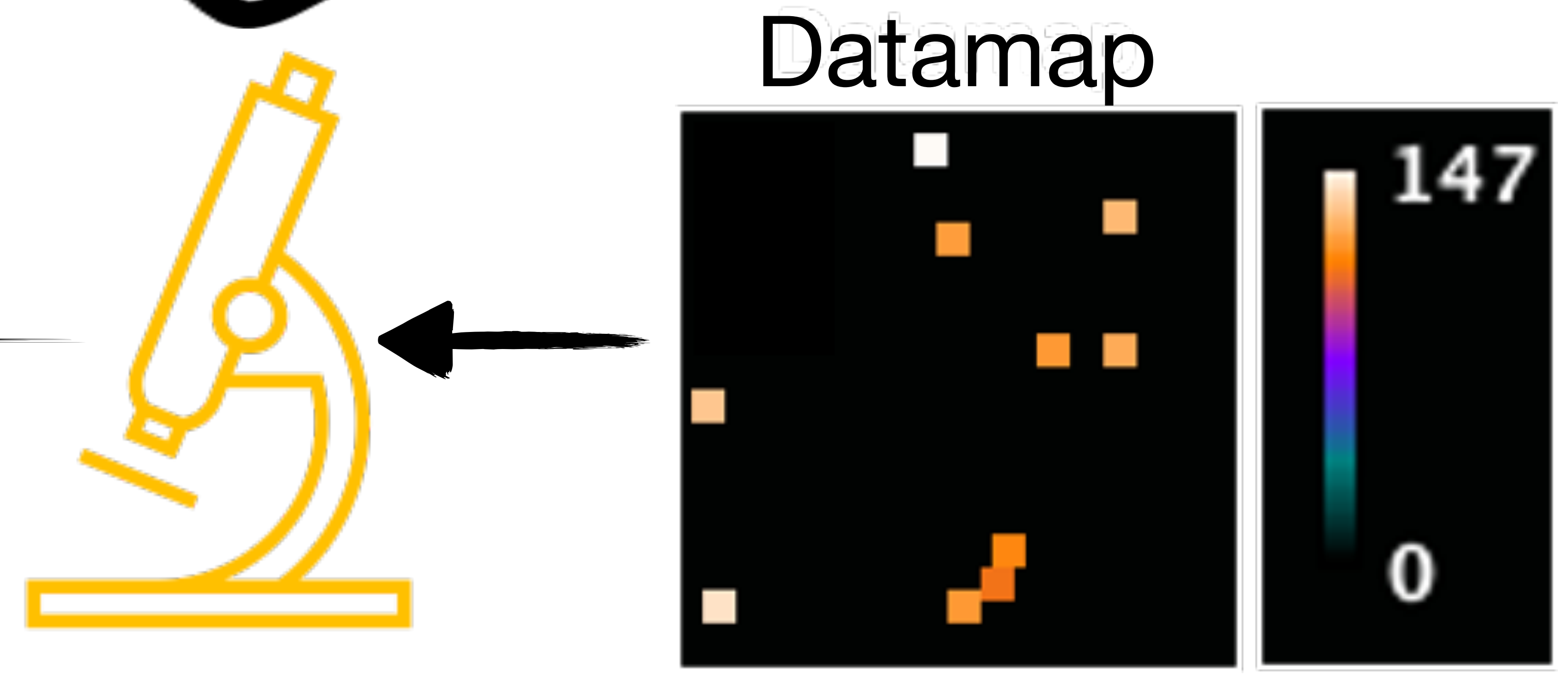
New problem: … → Good parameters for *this* sample

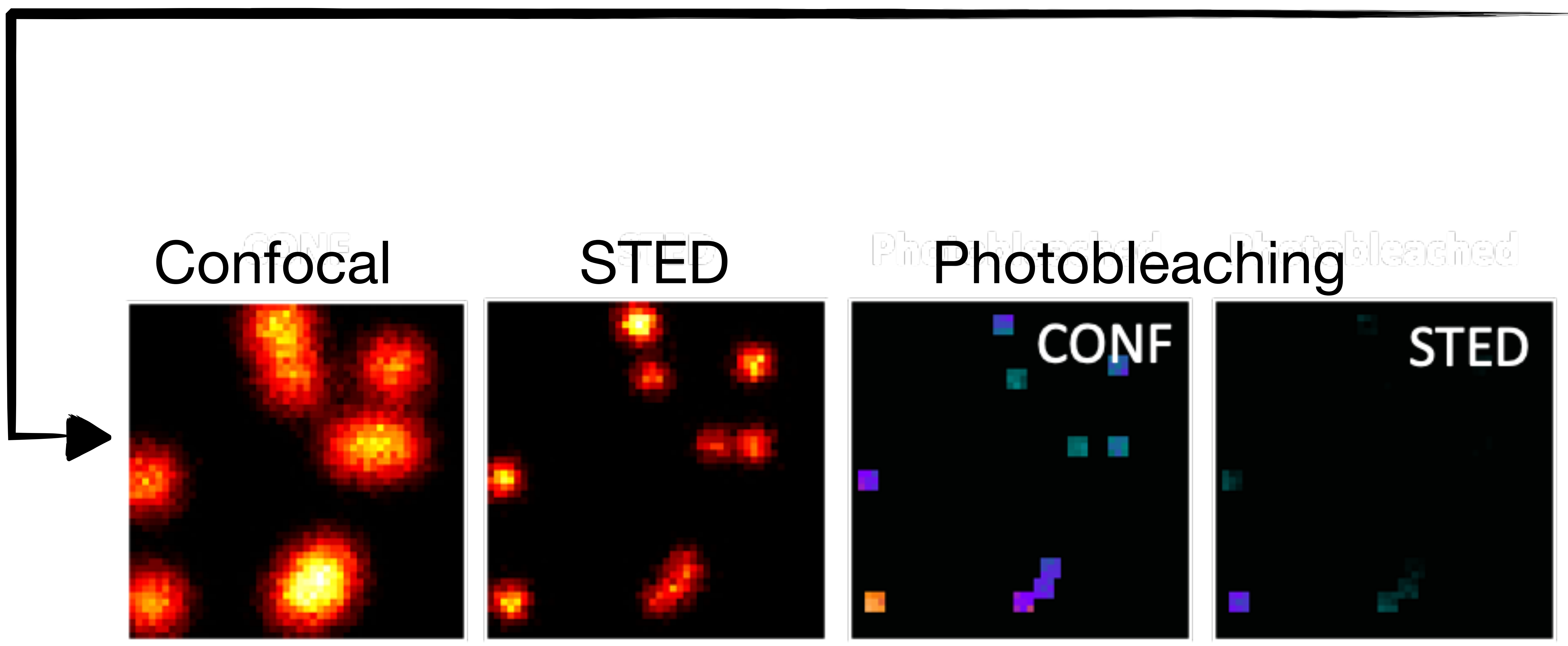
… vs RL formulation



pySTED

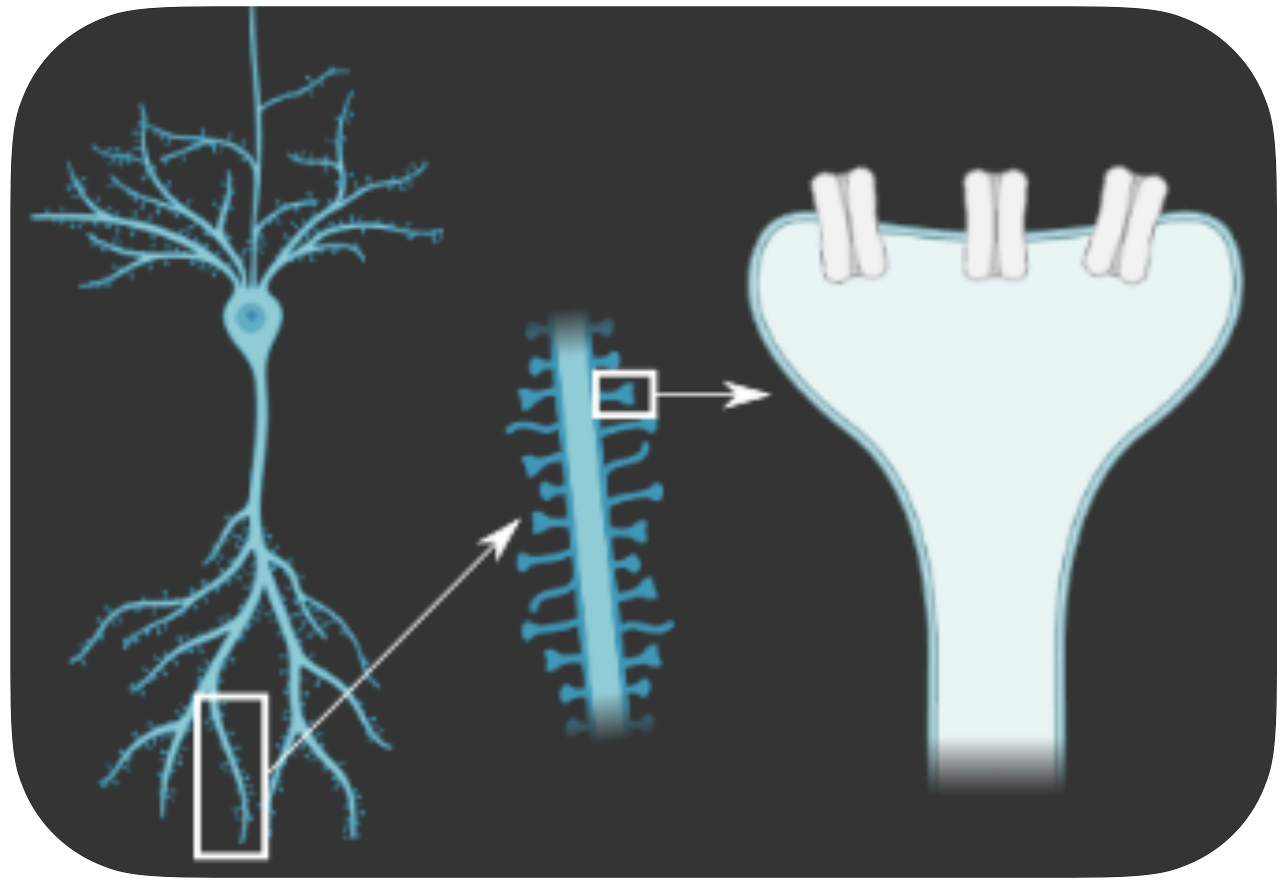




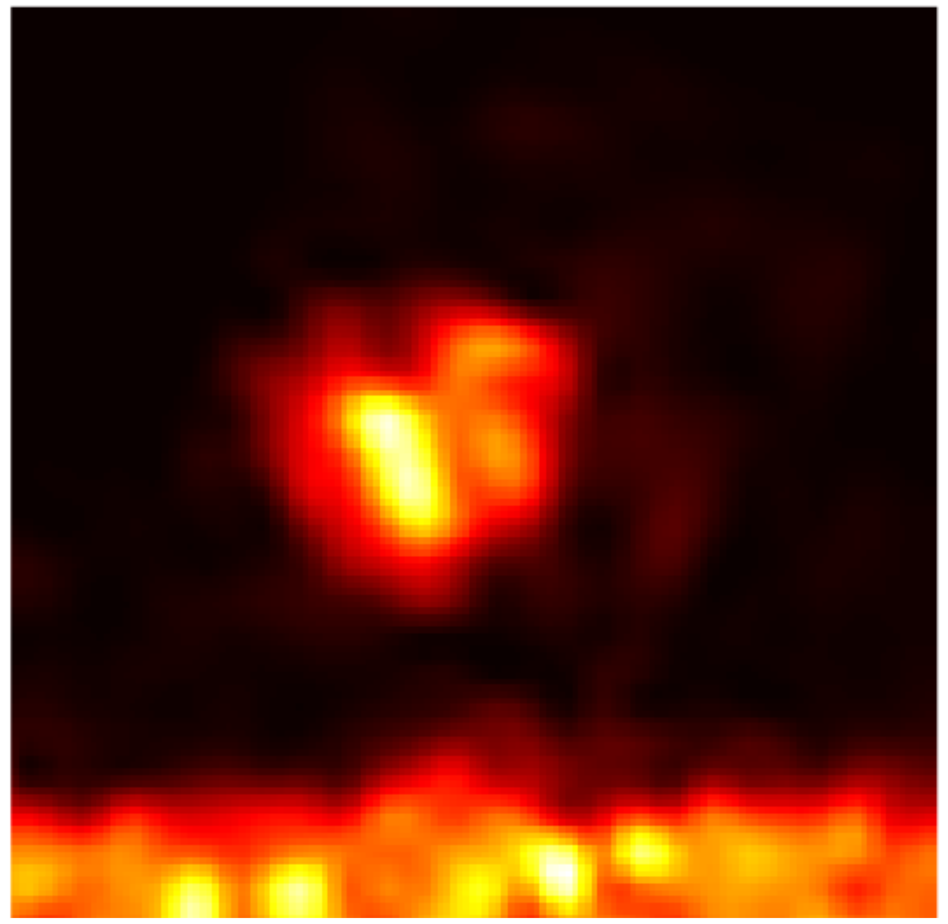


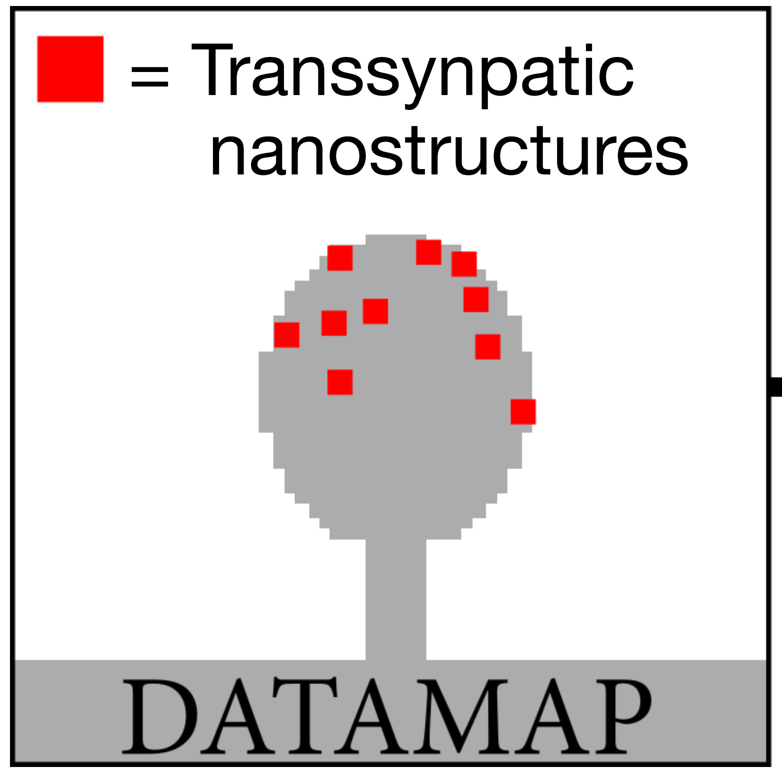
Turcotte et al. (2022) “pySTED : A STED Microscopy Simulation Tool for Machine Learning Training”. *AAAI workshop on AI to Accelerate Science and Engineering.*

Simulated imaging of a dendritic spine

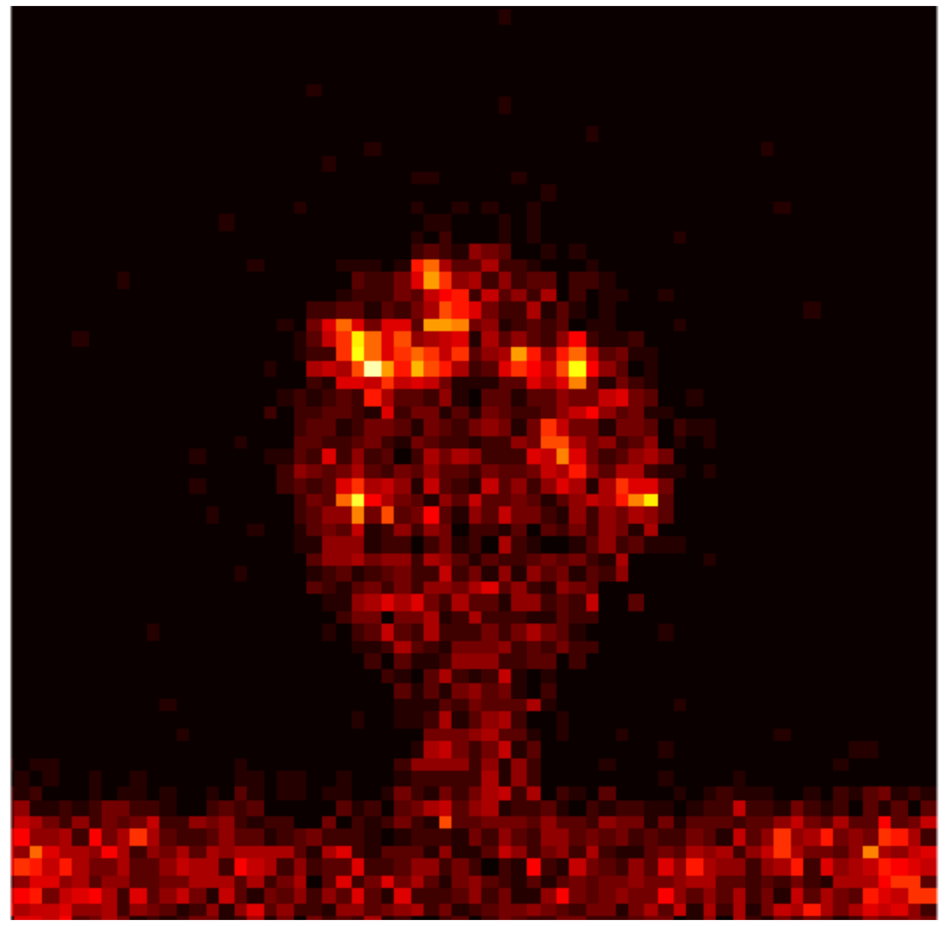


Real image

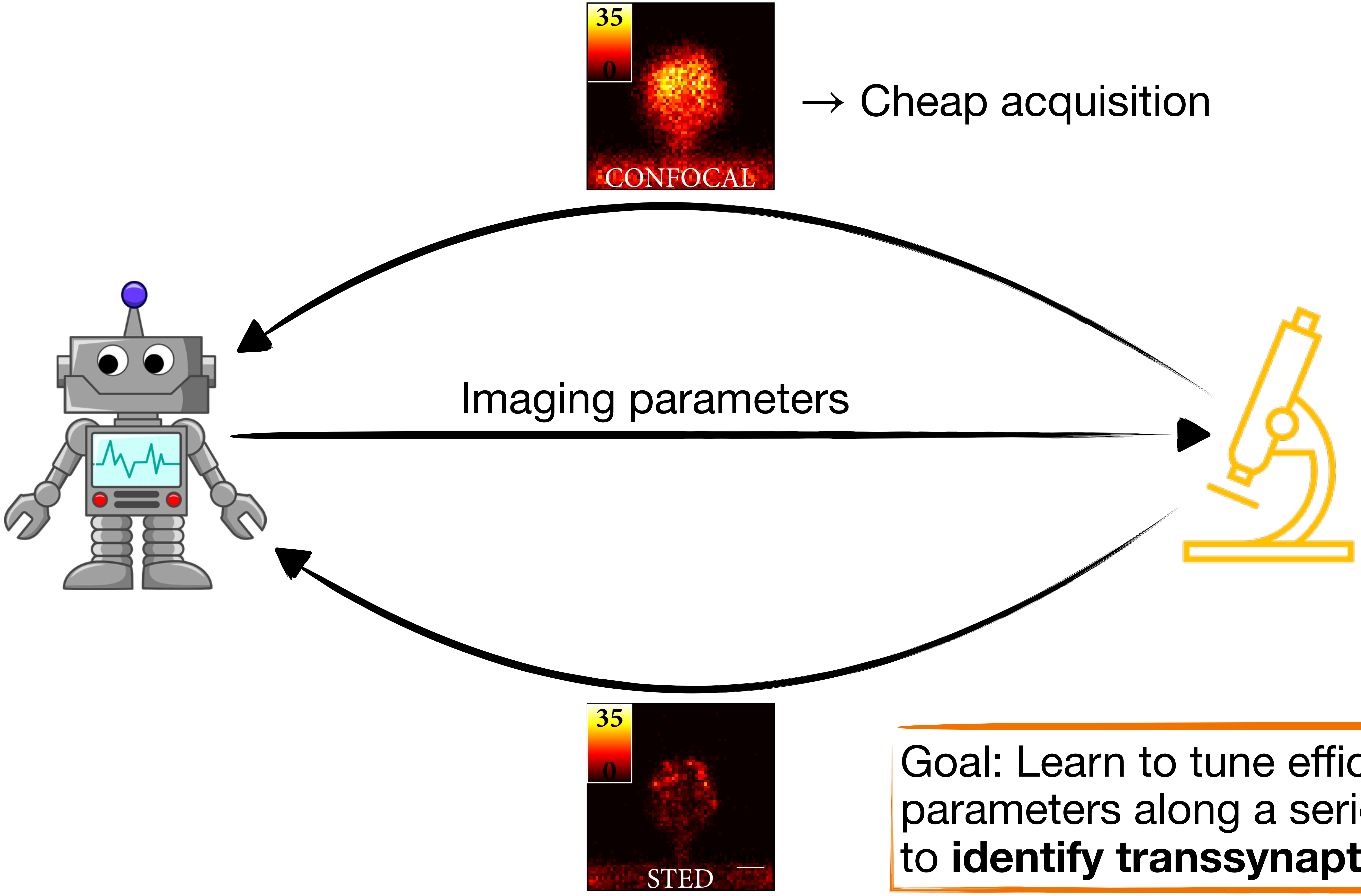




Simulated image



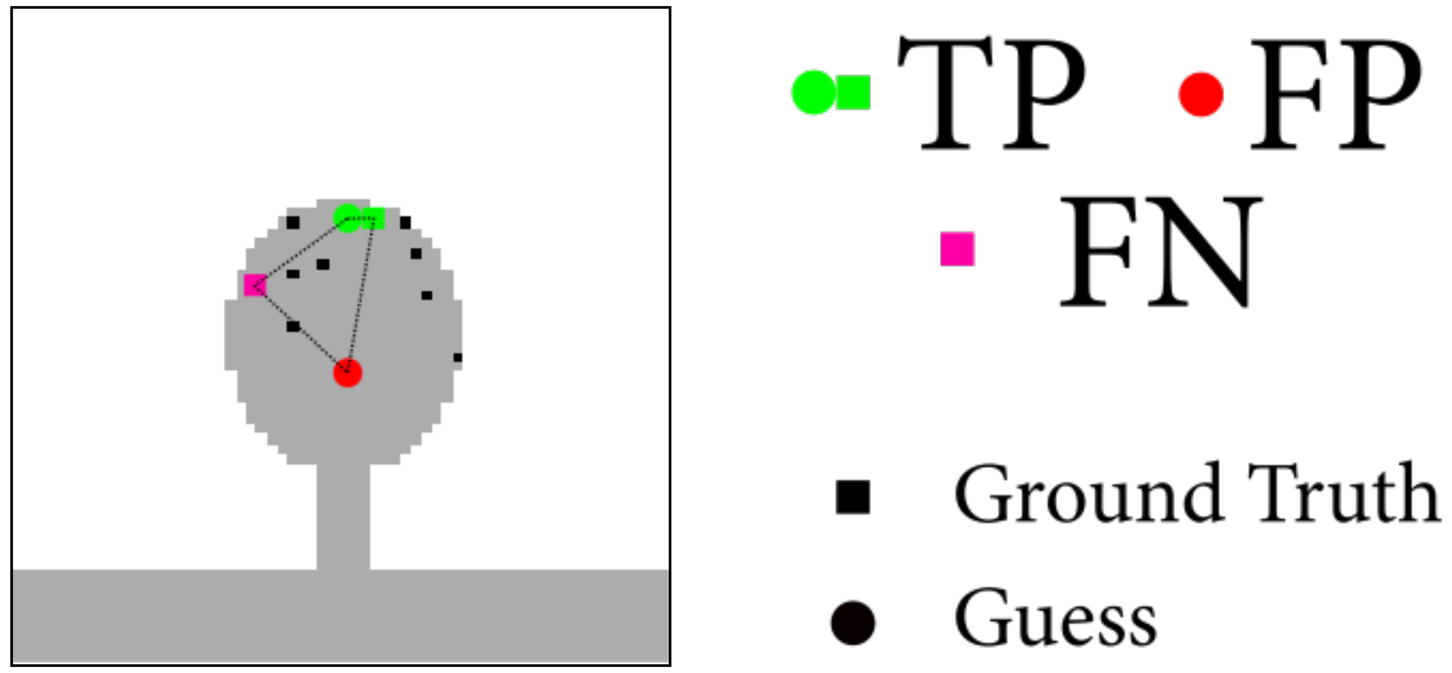
Learning an imaging policy



Goal: Learn to tune efficiently the imaging parameters along a series of images in order to **identify transsynaptic nanostructures**

Evaluating an acquisition

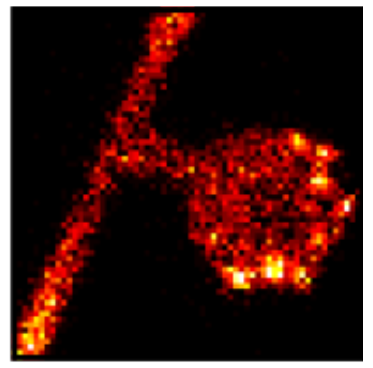
- Identifying nanostructures on a STED image:
  - Identify local intensity maxima + 2D Gaussian fit
  - Gaussian standard deviation (2 directions) < 250 nm (diffraction limit) → nanostructure
- Associating identified locations with true locations → Hungarian algorithm
- Comparing with true locations (ground truth) :

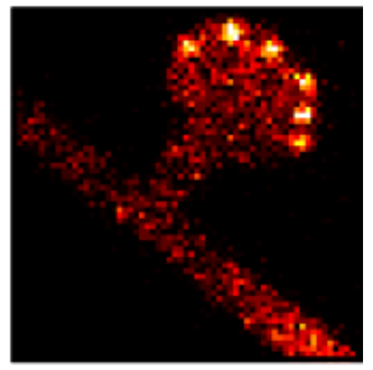


$$F_1\text{-score} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})}$$

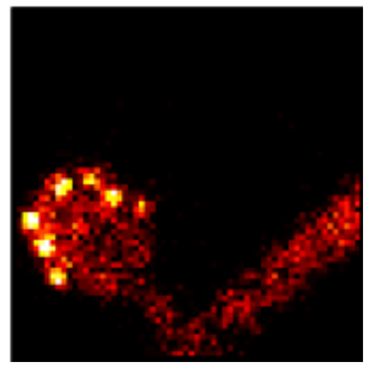
RL formulation

Episode:





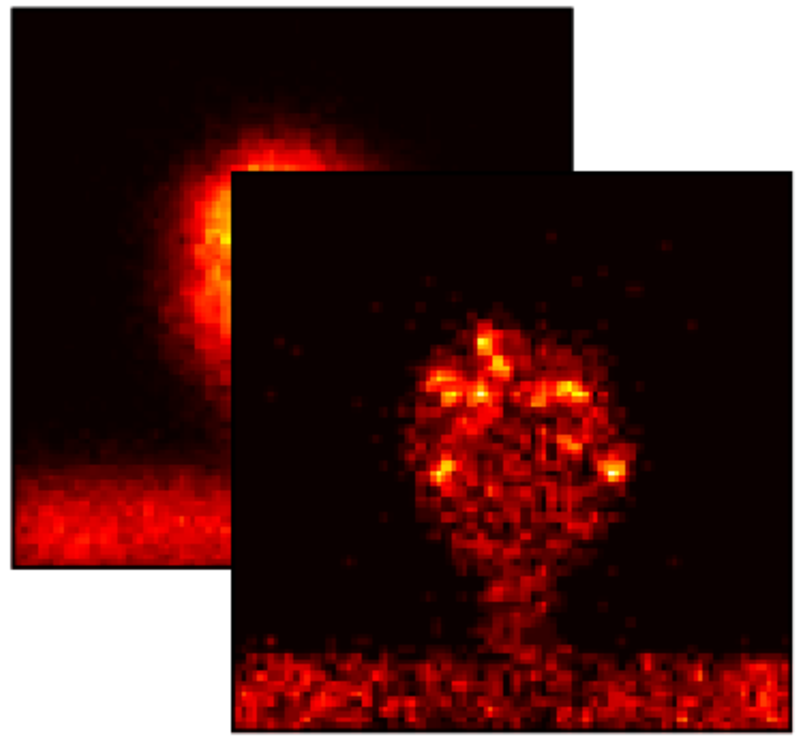
[. . .]



T images

State s_t:

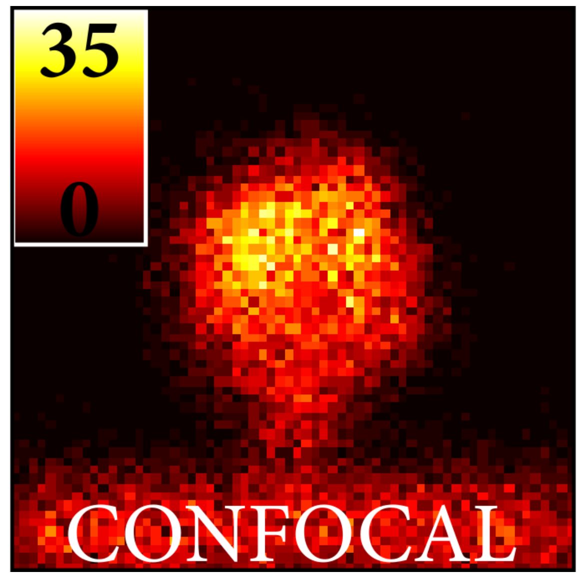
Last confocal-STED pair
and associated objective values



- Signal to noise ratio
- Resolution
- Photobleaching

+

Current confocal

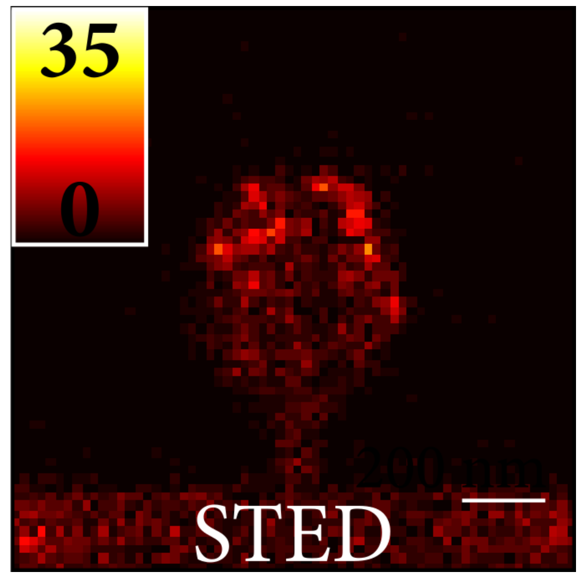


Action a_t: Imaging parameters

- Excitation laser power
- STED laser power
- Time spent per pixel

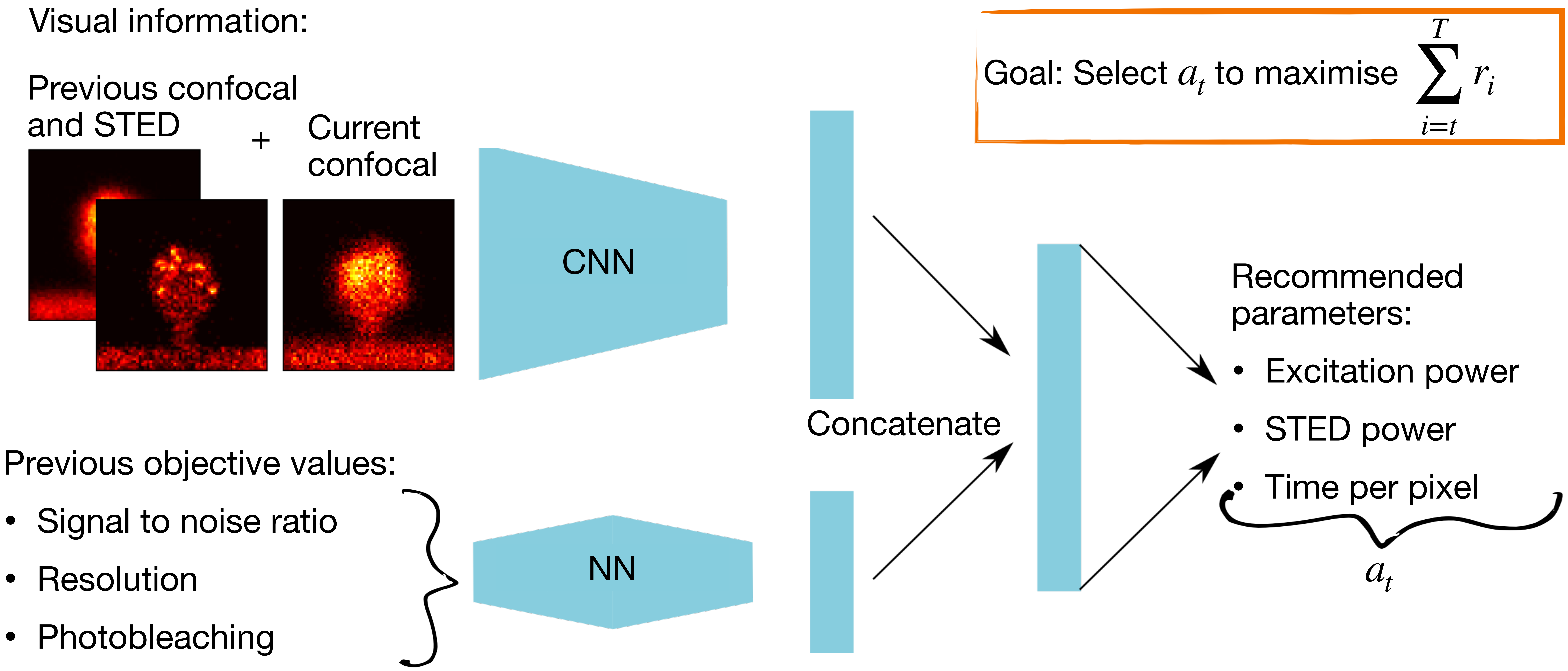
Reward r_t:

F_1-score (



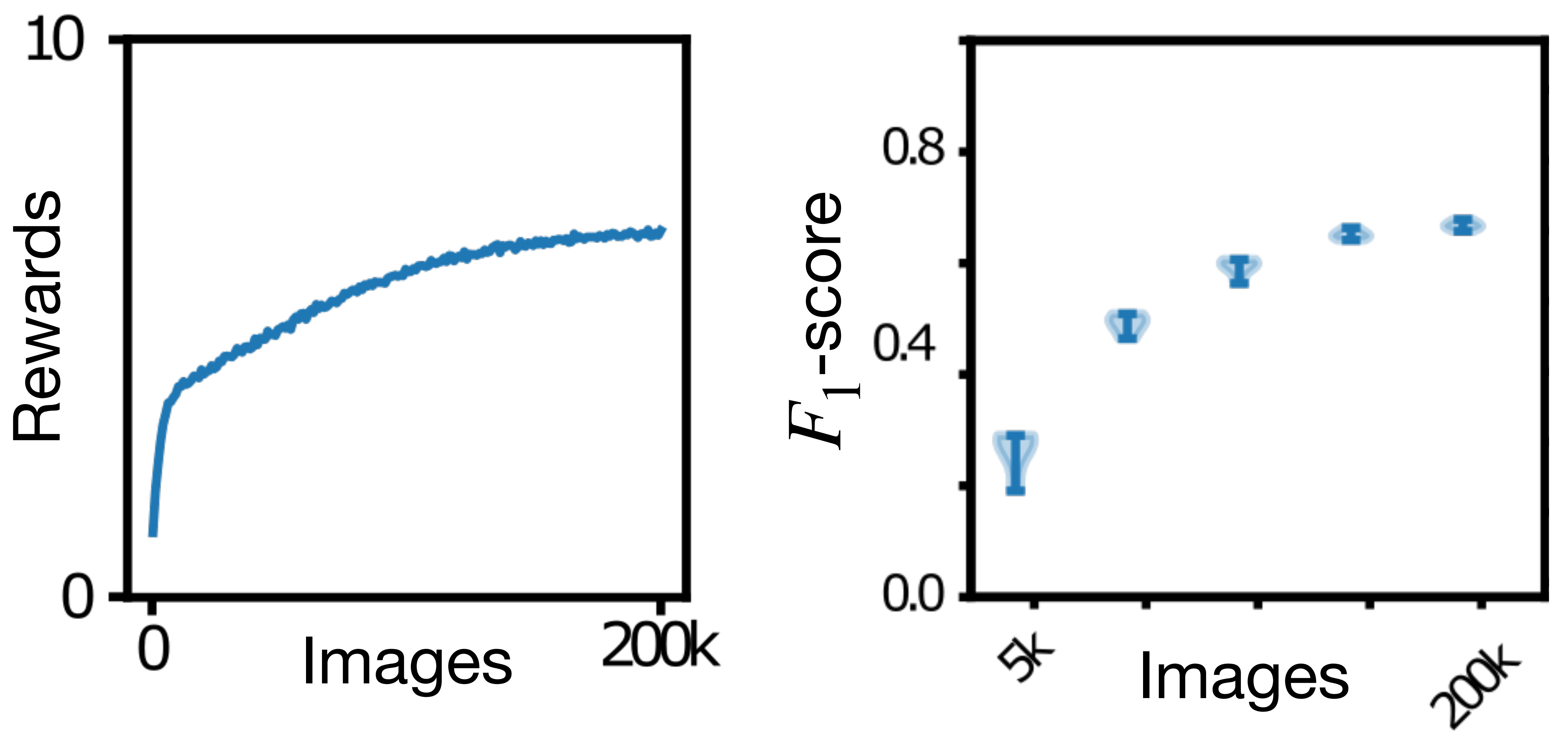
)

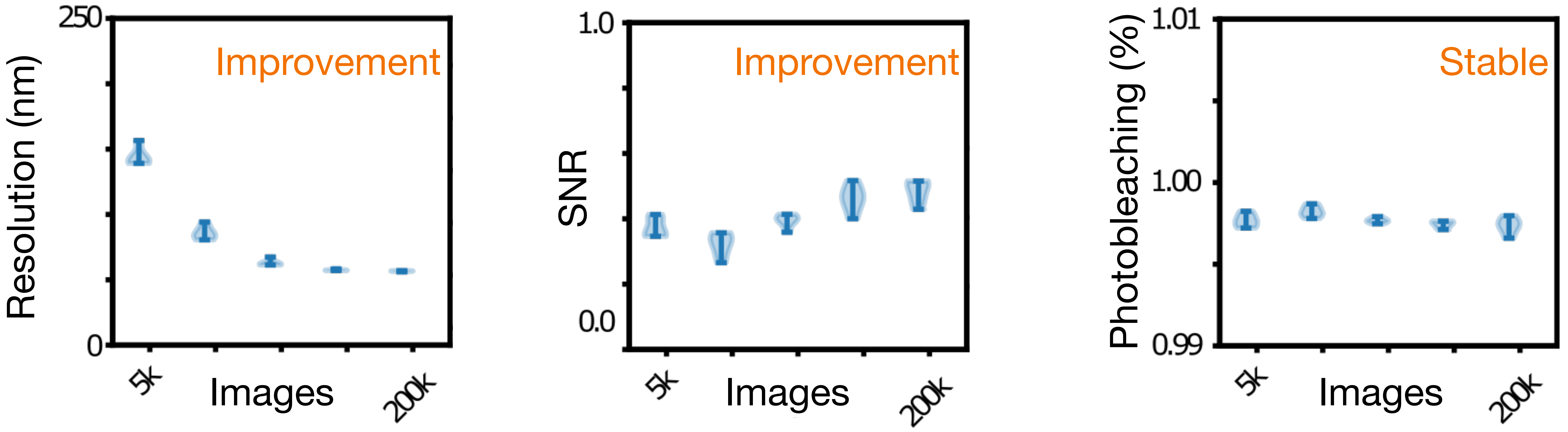
PPO (Proximal Policy Optimization)



Experiments

- Episodes of $T = 10$ images
- 20 000 episodes

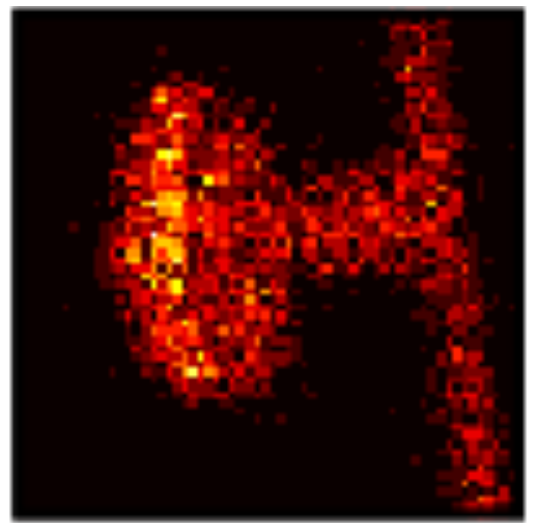




Rewards that do not depend on ground truth

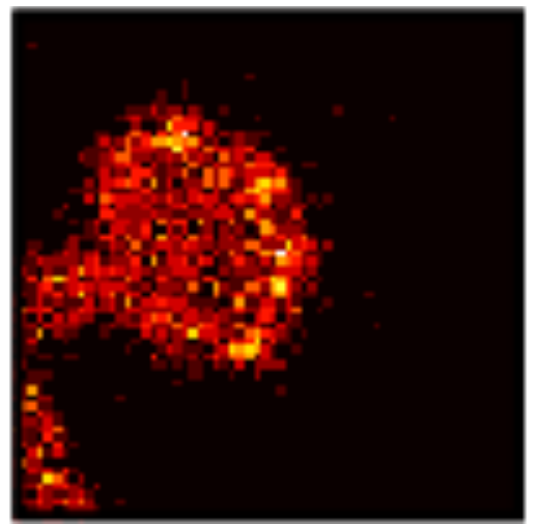
Expert image databank:

Best

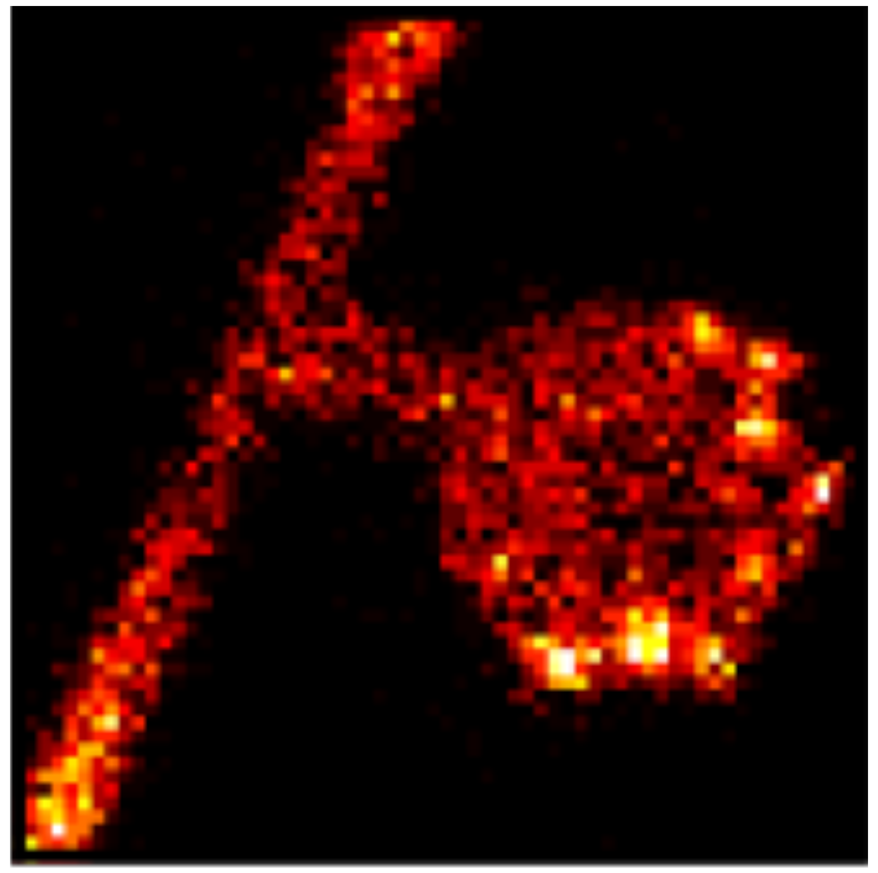


...

Worst



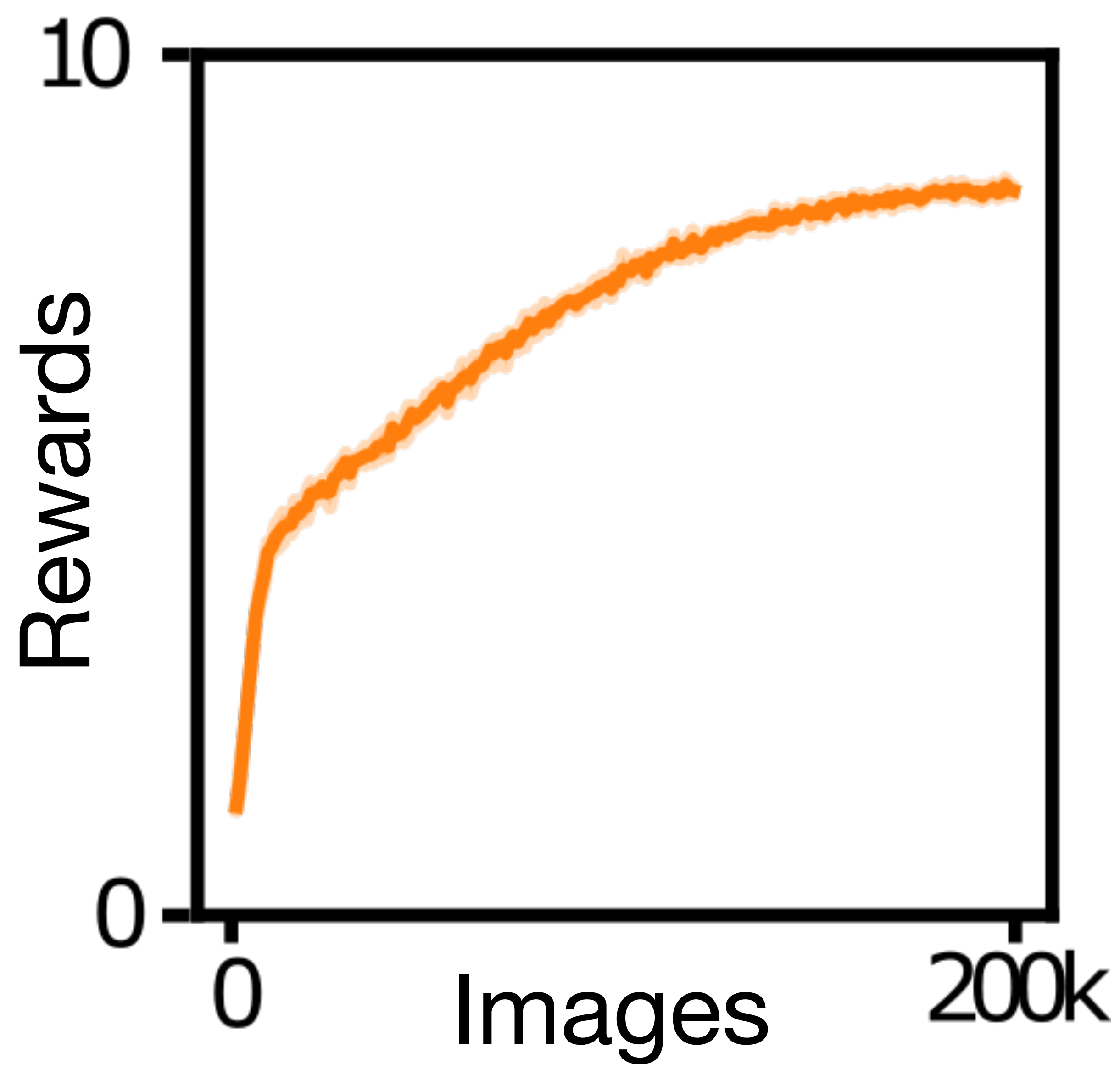
Acquired image:

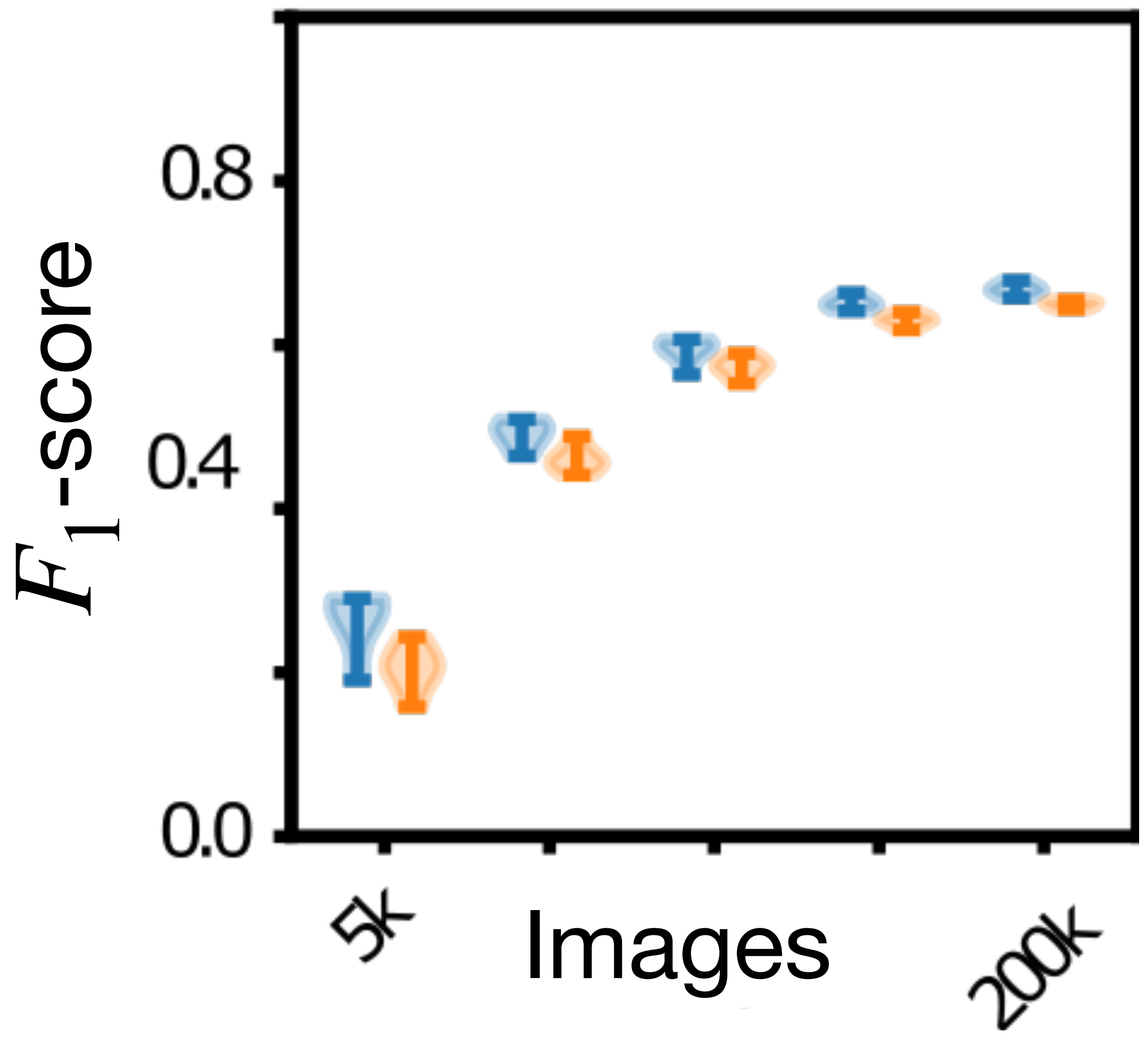


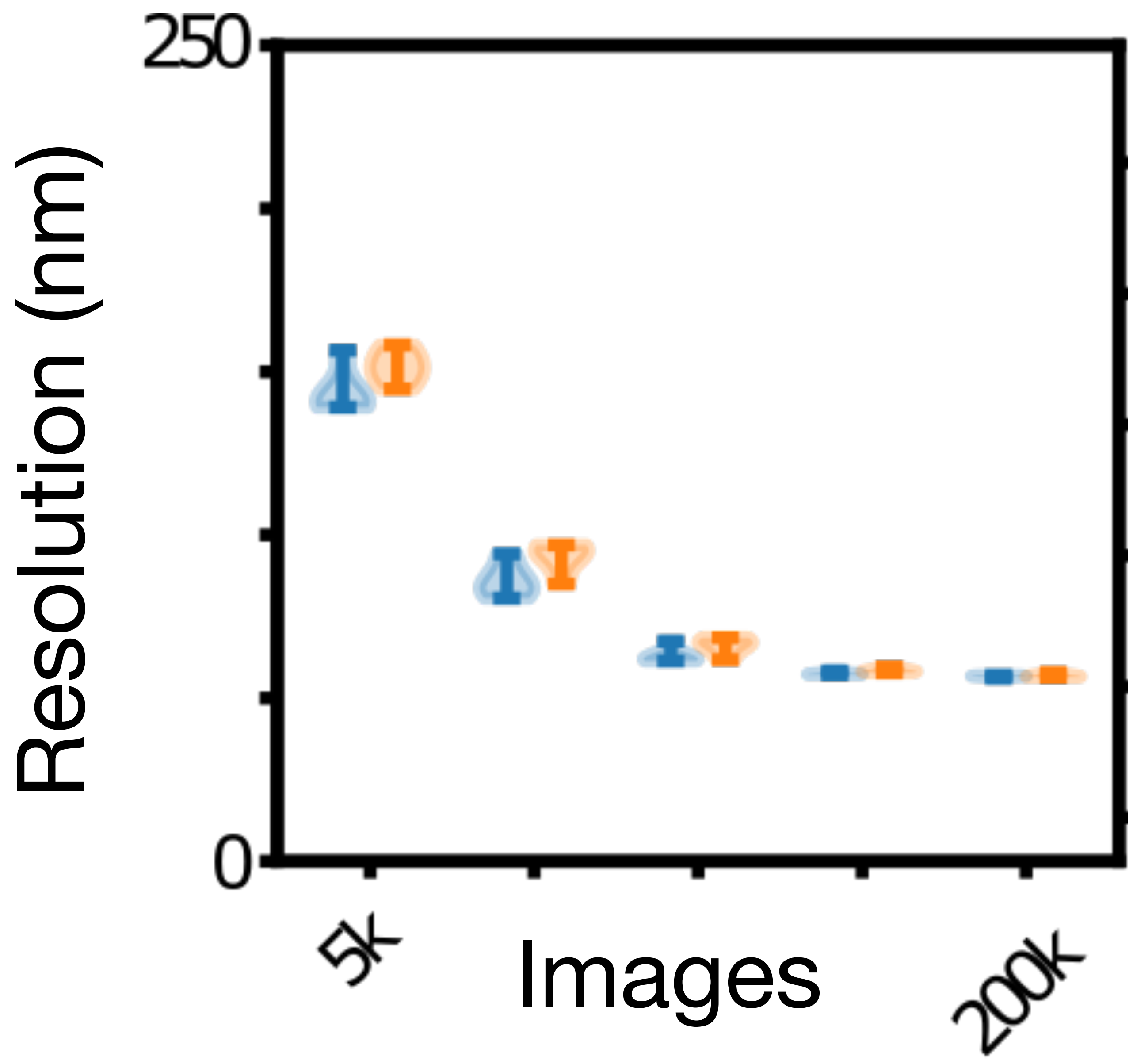
Reward r_t: Ranking between 0 (worst) and 1 (best)

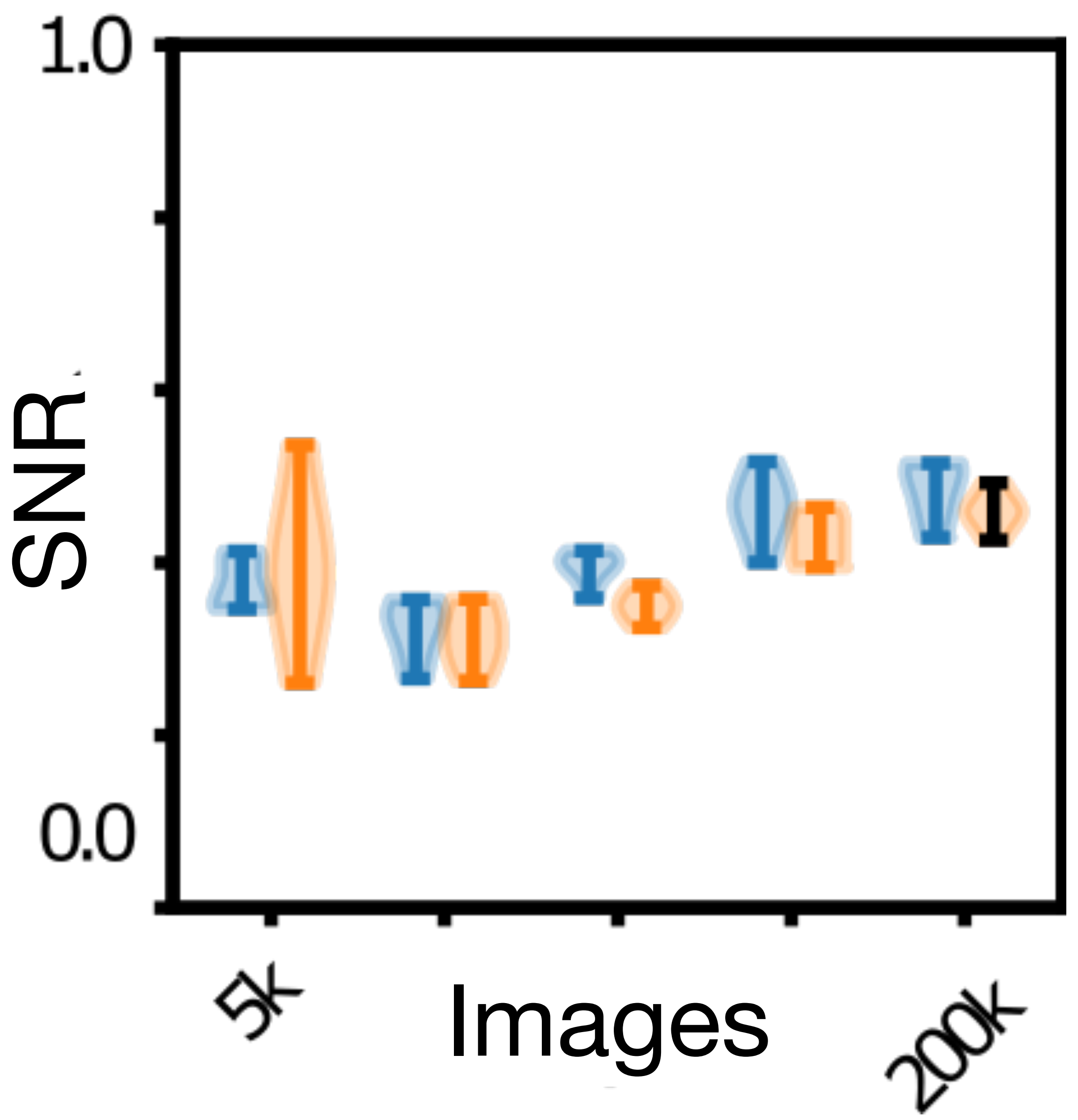
Experiments

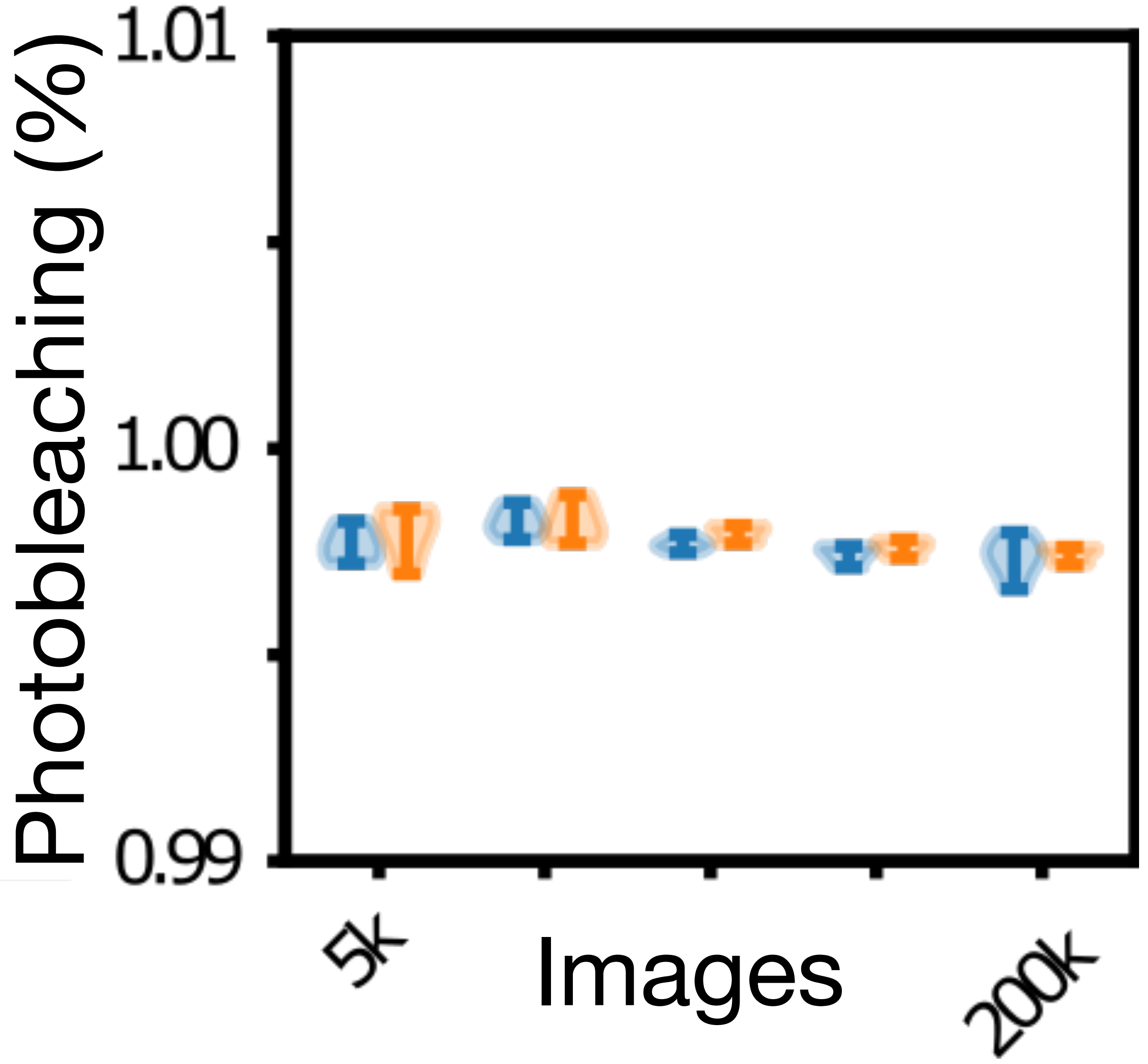
- Episodes of $T = 10$ images
- 20 000 episodes











A fully automated parameter tuning strategy!

Thanks!

UNIVERSITÉ LAVAL

Institut intelligence et données

CENTRE DE RECHERCHE CERVO BRAIN RESEARCH CENTRE



Questions?