

DeepMind

BYOL-Explore

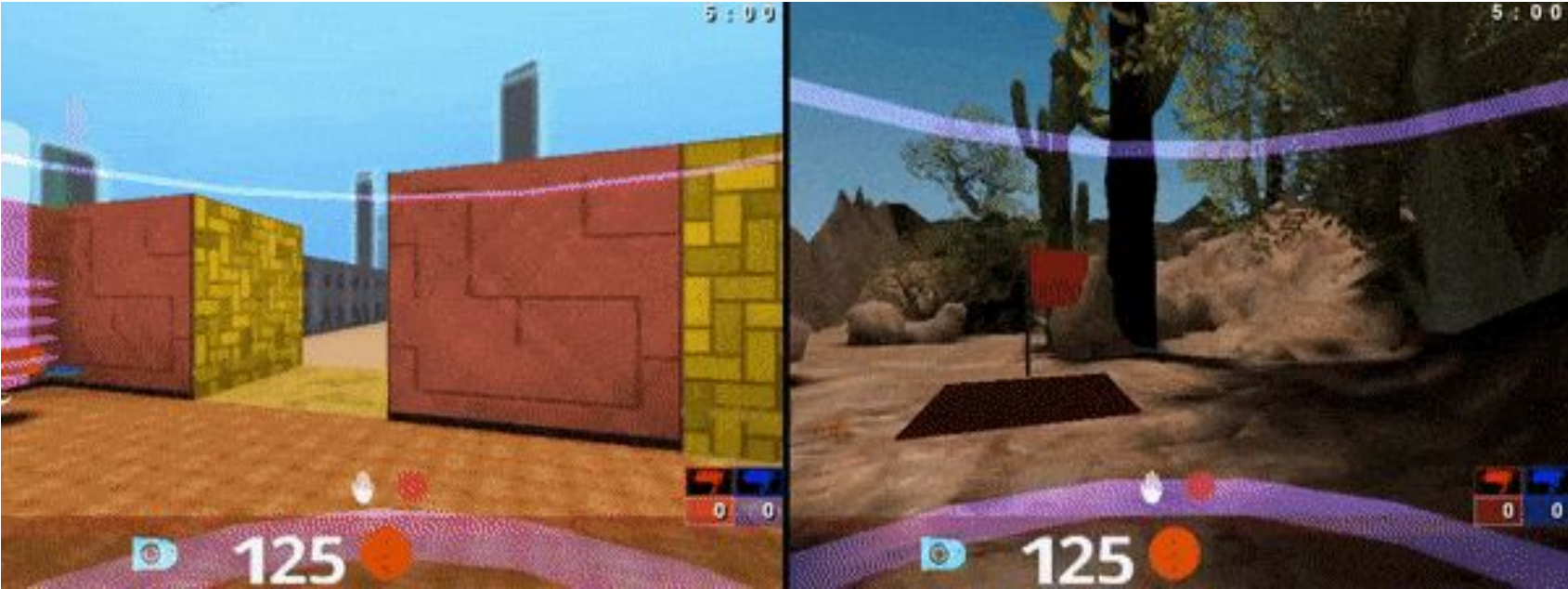
Exploration by Bootstrapped Prediction

Michal Valko

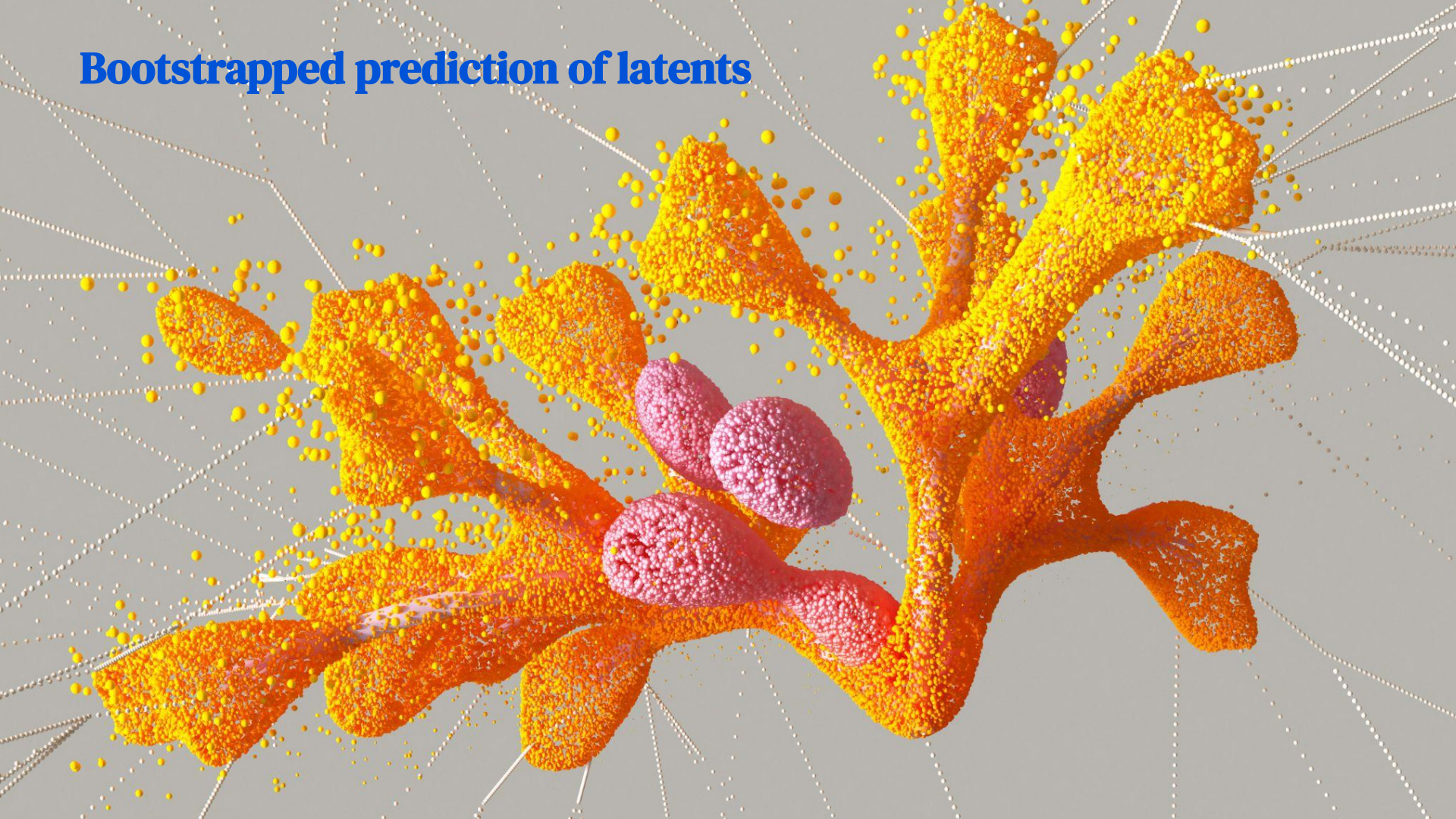
Join work with *Zhaohan Daniel Guo, Shantanu Thakoor, Miruna Pislari, Bernardo Avila Pires, Florent Altché, Corentin Tallec, Alaa Saade, Daniele Calandriello, Jean-Bastien Grill, Yunhao Tang, Rémi Munos, Mohammad Gheshlaghi Azar, Bilal Piot*



Visually-complex environments



Bootstrapped prediction of latents



Exploration in reinforcement learning

Sparse Reward Environment



Exploration is important for sparse reward environments

BUT

It is **infeasible** to try to explore everything in **large, complex environments**

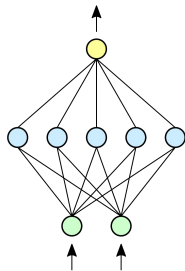
THUS

It is **crucial** to pick and choose what to explore.



Curiosity-Driven Exploration

Build a **world model**



Explore the **mistakes** of world model to better refine it

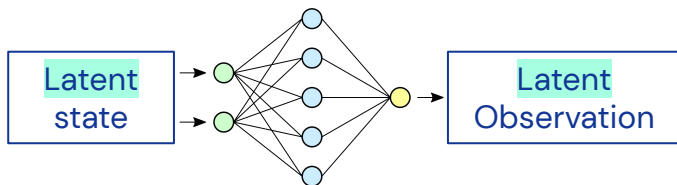
$$\max_{\pi} \text{WorldModelLoss}(\pi)$$

The **world model** determines what is **interesting** to explore and what to ignore



BYOL-Explore Latent World Model

Extend BYOL[1] to learn a self-supervised representation and latent dynamics model



Explore the latent mistakes of world model to better refine it

$$\underbrace{\max_{\pi}}_{\text{Exploration}} \min_{\theta} \text{BYOLLoss}_{\pi}(\theta)$$

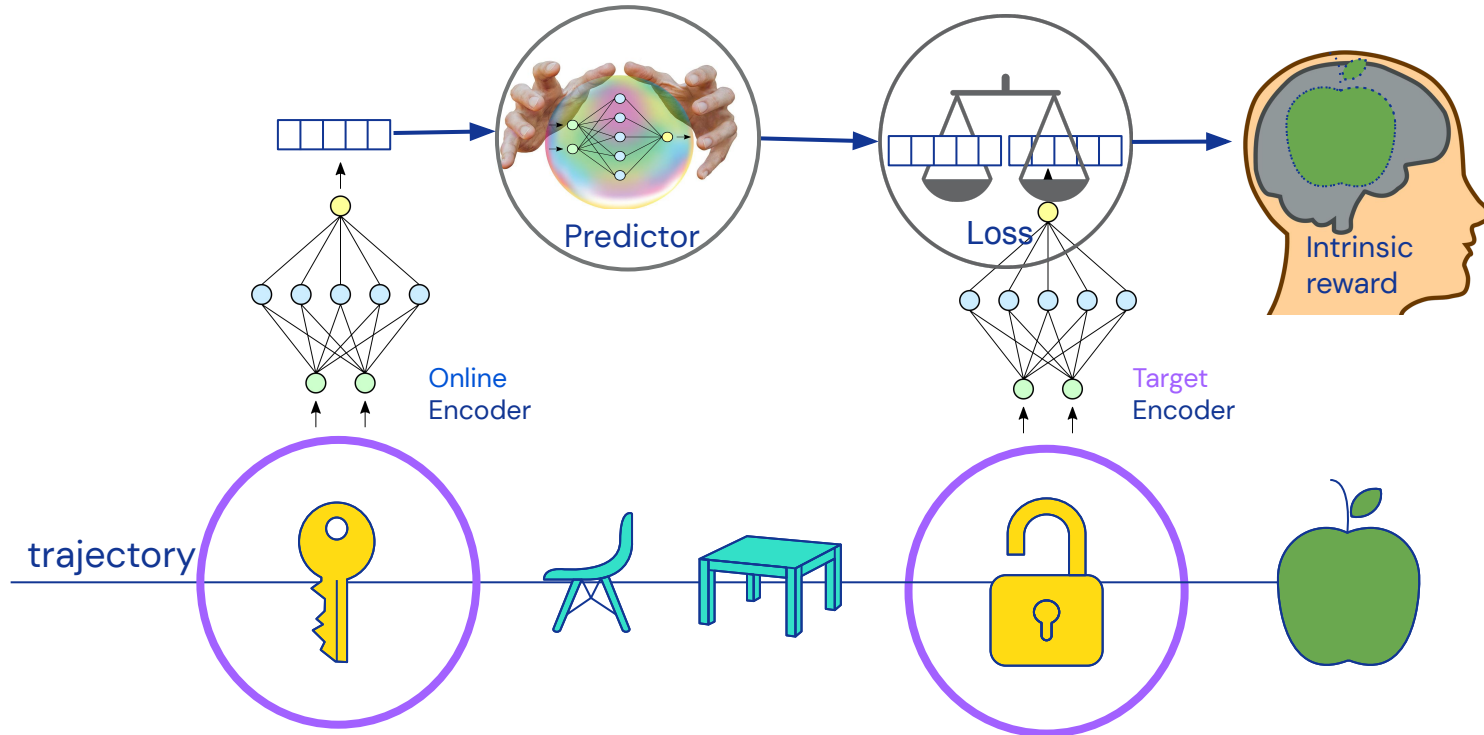
Dynamics Model Learning

The mistakes are dynamics-aware and structured, since they are in latent space

One unified objective for representation learning, dynamics modelling, and exploration



BYOL-Explore algorithm overview



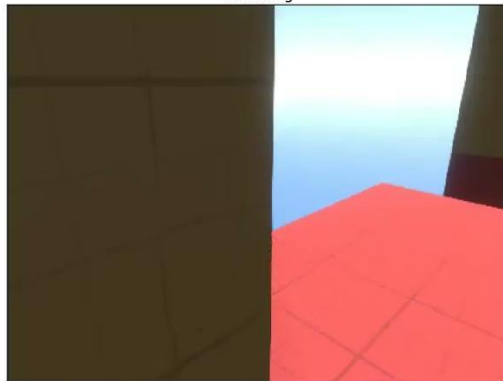
DeepMind

Experimental results

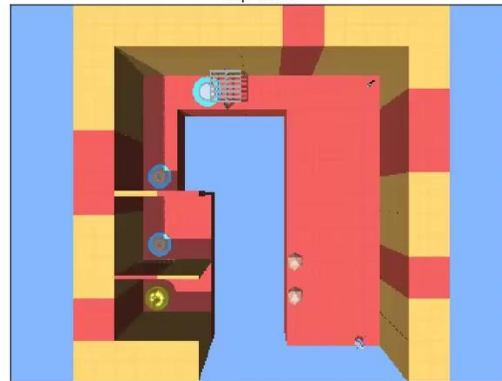
First Person



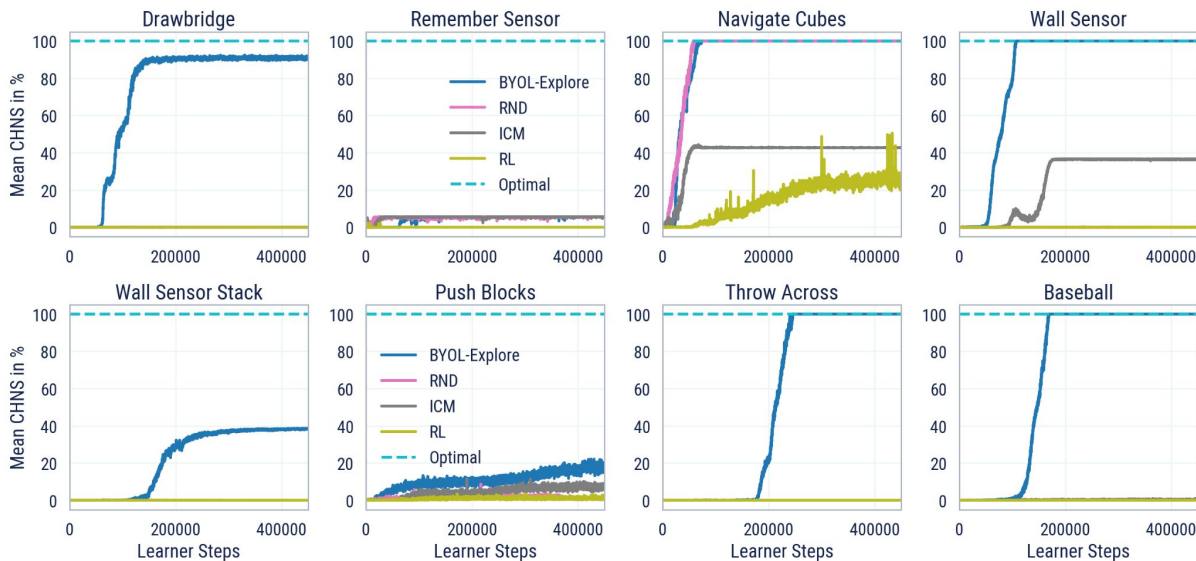
Following



Top-down



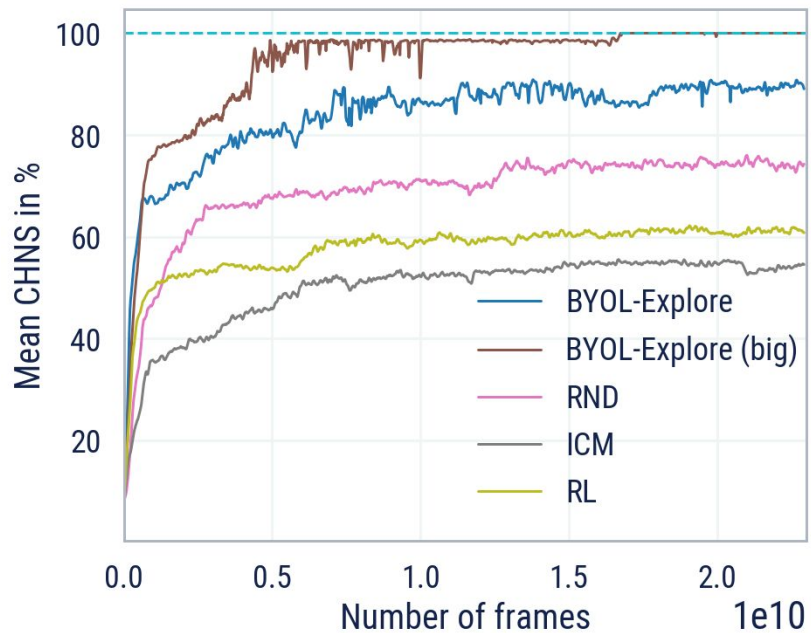
DM-Hard-8



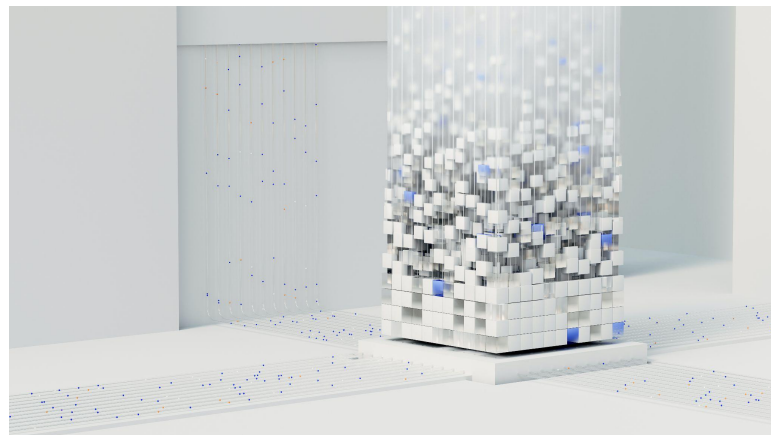
BYOL-Explore solves 5.5/8 tasks in DM-Hard-8, where **previously SOTA results used demonstrations**



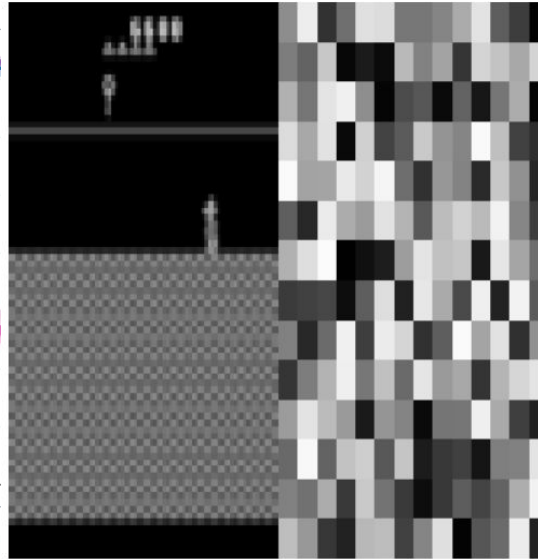
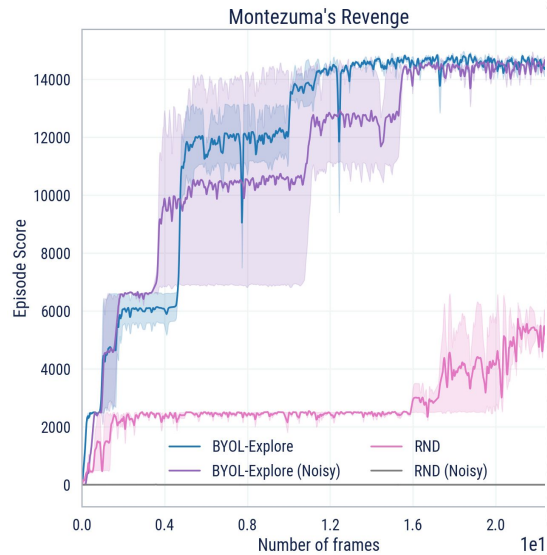
Atari - Hard exploration



Achieves **near-superhuman** performance on the 10 hardest exploration Atari games.



Atari - Hard exploration - Noise



Superhuman on Montezuma's Revenge and is robust to 'noisy-tv' because it is in **latent space**



DeepMind

BYOL-Explore

A single, simple curiosity-driven algorithm for

- **Representation learning**
- **Dynamics modelling**
- **Exploration**

See paper and poster for more details!

