Image Augmentation Is All You Need: Regularizing Deep Reinforcement Learning from Pixels

Denis Yarats*
NYU & FAIR
@denisyarats

Ilya Kostrikov*
NYU
@ikostrikov

Rob Fergus
NYU
@rob_fergus

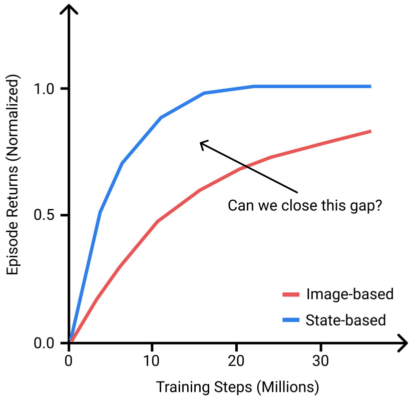
ICLR 2021

*Equal contribution. Author ordering determined by coin flip.

Deep Reinforcement Learning (DRL) from Pixels

- Sample-efficient Deep RL algorithms from images would enable many real-world applications.
- Learning from pixels requires a order of magnitude more samples than states.
- **Goal:** Make pixel-based RL as efficient as state-based.

State-based vs Image-based DRL methods

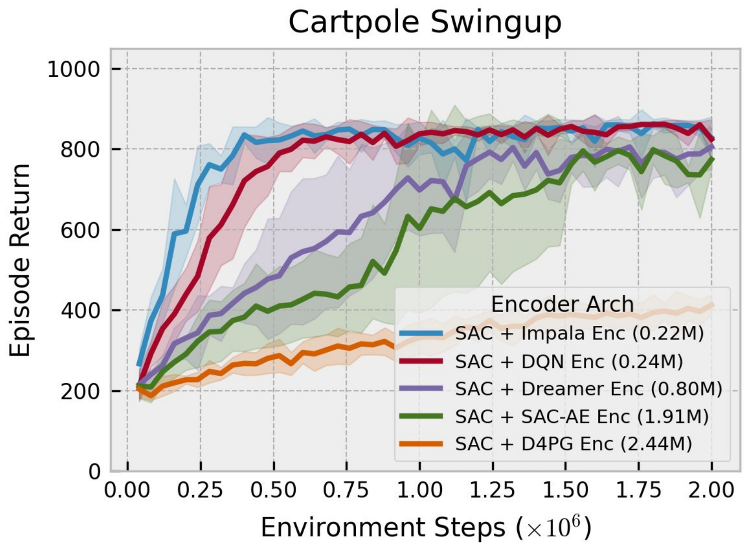


How to learn efficiently from Pixels?

- Why can’t we learn directly from pixels efficiently?
- Does overfitting prevent efficient learning?

Overfitting in Off-Policy DRL from Pixels

- (SAC) Soft Actor-Critic (Haarnoja et al., 2018) + Image Encoders of different capacity.
- Overfitting: more parameters leads to worse performance!



Diagnosing Off-policy DRL from Pixels

- Usually learning starts from several thousand transitions in replay buffer collected by a suboptimal policy.
- Highly correlated data:
  - Samples within the same trajectories are correlated.
  - Q-targets are correlated.
- This leads to overfitting!
- How do we break spurious correlations?

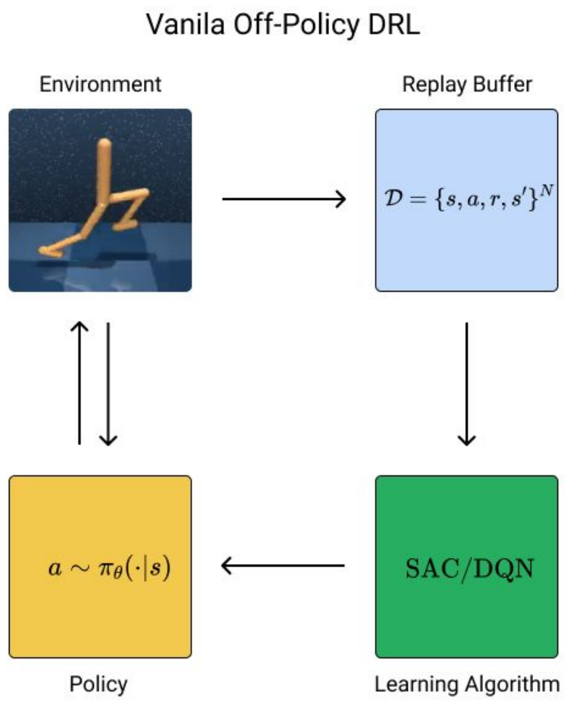


Image Augmentation in Off-Policy DRL from Pixels

- Attempt to reduce overfitting:
  - Take inspiration from Computer Vision (L2, dropout, data augmentation).
- Image augmentation with random shifts shows promise!
  - After sampling from the replay buffer and before using this data in SAC/DQN, we augment images.

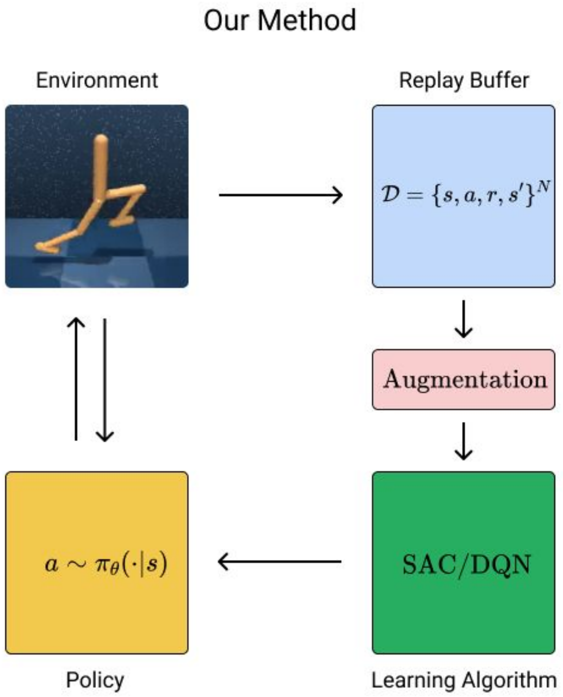
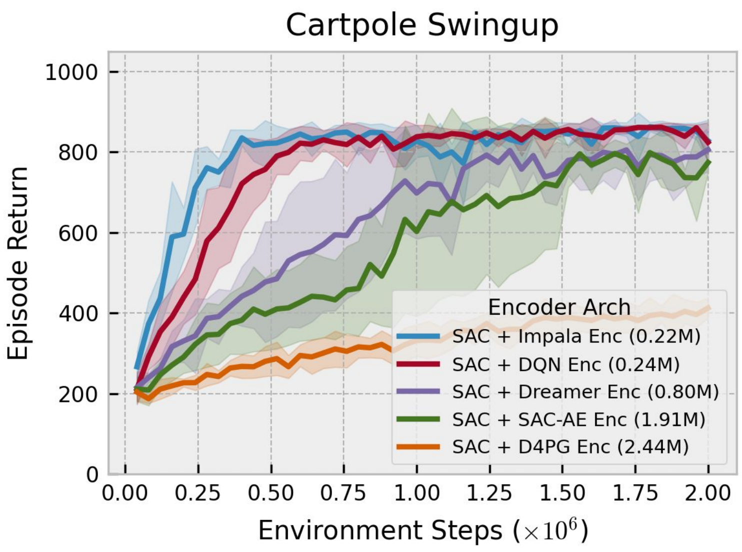
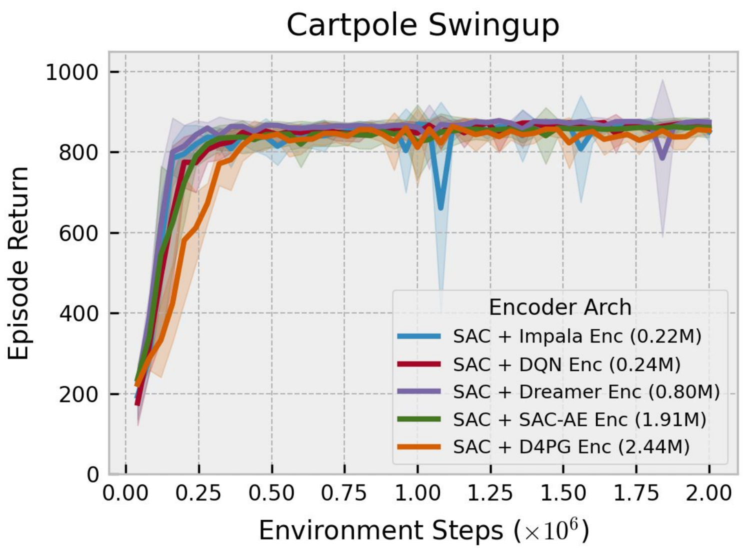


Image Augmentation in Off-Policy DRL from Pixels

- SAC+Image Encoder.
- Overfitting.

- SAC+Image Encoder+Regularization.
- Similar Performance.



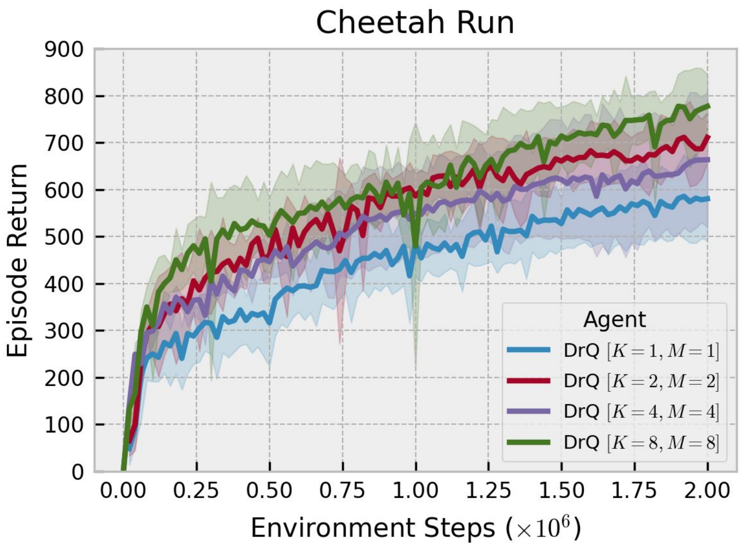


Further Augmentation strategies: Data-regularized Q

1. Regularize targets with several augmentations (**K**):

$$y_i = r_i + \gamma \frac{1}{K} \sum_{k=1}^{K} Q_\theta(f(s'_i, \nu'_{i,k}), a'_{i,k})$$

where $a'_{i,k} \sim \pi(\cdot | f(s'_i, \nu'_{i,k}))$



2. Regularize Q with several augmentations (**M**):

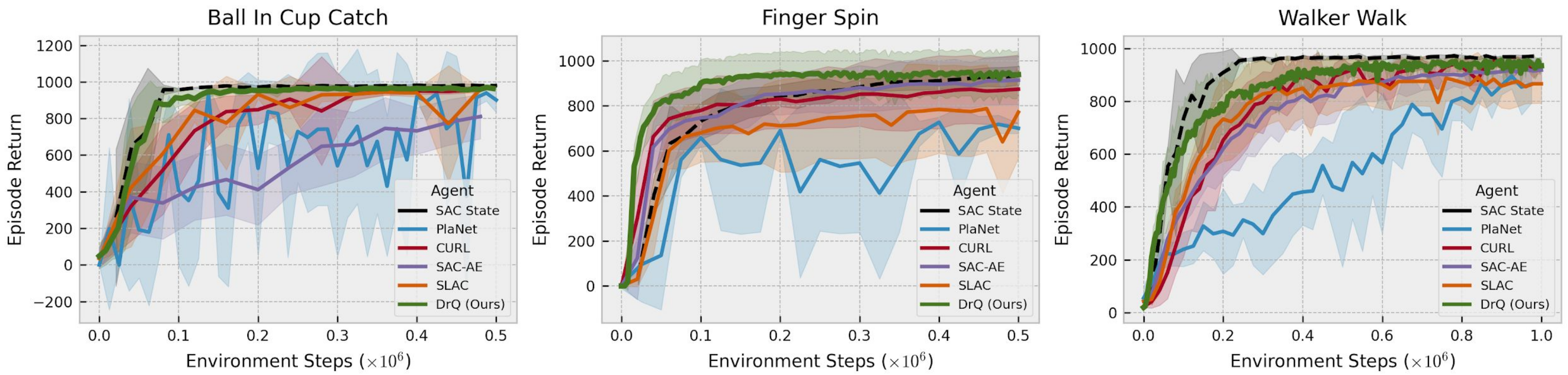
$$\theta \leftarrow \theta - \lambda_\theta \nabla_\theta \frac{1}{NM} \sum_{i=1,m=1}^{N,M} (Q_\theta(f(s_i, \nu_{i,m}), a_i) - y_i)^2$$

Prior work in Image-based DRL

- Model-Based methods:
  - PlaNet (Hafner et al., 2018), SLAC (Lee et al., 2019), Dreamer (Hafner et al., 2019).
  - Require learning a complex world-model which are hard to train.
- Auxiliary tasks methods:
  - UNREAL (Jaderberg et al., 2016), SAC-AE (Yarats et al., 2019), CURL (Srinivas et al., 2020).
  - Hard to design an axillary that correlates well with the downstream task.
  - Common unsupervised methods (e.g. autoencoder, contrastive loss) are not ideal, as they tend to compete with the downstream objective.
- Contemporaneous work on image augmentation:
  - RAD (Laskin et al., 2020).

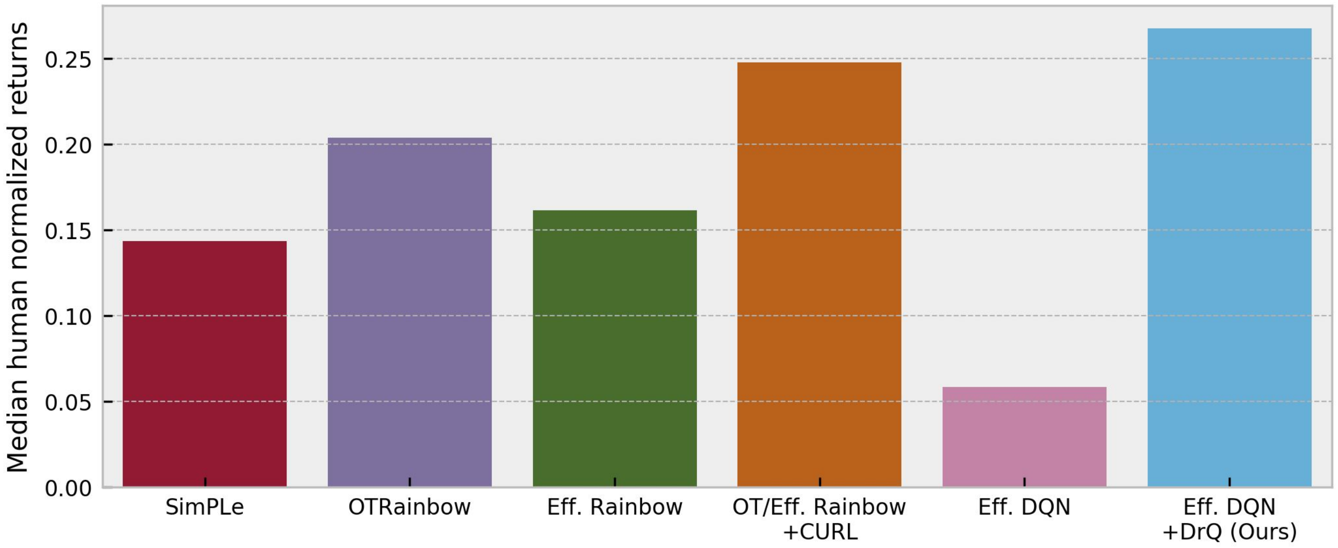
DeepMind Control Suite

- **Model-free**: SAC-AE (Yarats et al., 2019), CURL (Srinivas et al., 2020).
- **Model-based**: PlaNet (Hafner et al., 2018), SLAC (Lee et al., 2019).



Atari 100k

- **Model-free**: OTRainbow (Kielak et al., 2020), Efficient Rainbow (van Hasselt et al., 2019), CURL (Srinivas et al., 2020).
- **Model-based**: SimPLe (Kaiser et al., 2018).



Different Image Augmentations

None

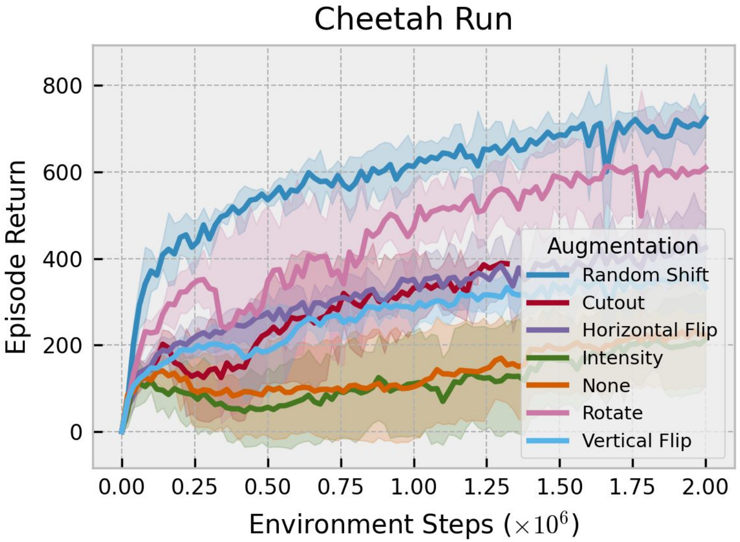
Cutout

Flip

Random Shift

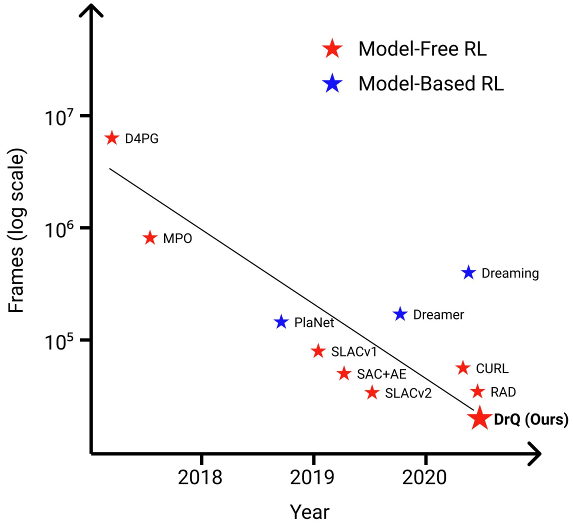
Intensity

Rotate



Evolution of Image-based DRL Algorithms

- Impressive progress in DRL from Pixels.
- ~2 orders of magnitude improvements for both model free/based methods.



Based on performance on DM Control Suite

Thank you for attention!

Q/A at poster session 2:
May 3, 2021, 9 AM PDT
May 3, 2021, 11 AM PDT