Understanding the Mechanism behind Data Augmentation's Success on Image-based RL

David Klee
Khoury College of Computer Sciences
Northeastern University
Boston, MA 02115
dmklee@ccs.neu.edu

Robin Walters
Khoury College of Computer Sciences
Northeastern University
Boston, MA 02115

Robert Platt
Khoury College of Computer Sciences
Northeastern University
Boston, MA 02115

Abstract

Reinforcement learning for continuous control tasks is challenging with image observations, due to the representation learning problem. A series of recent work has shown that augmenting the observations via random shifts during training significantly improves performance, even matching state-based methods. However, it is not well-understood why augmentation is so beneficial; since the method uses a nearly-shift equivariant convolutional encoder, shifting the input should have little impact on what features are learned. In this work, we investigate why random shifts are useful augmentations for image-based RL and show that it increases both the shift-equivariance and shift-invariance of the encoder. In other words, the visual features learned exhibit spatial continuity, which we show can be partially achieved using dropout. We hypothesize that the spatial continuity of the visual encoding simplifies learning for the subsequent linear layers in the actor-critic networks.

Keywords: Data Augmentation, Equivariance, Representation Learning

Acknowledgements

We are grateful to Kevin Esslinger who provided valuable feedback on the writing.

1 Introduction

Solving continuous control tasks with reinforcement is challenging when using image observations. The reinforcement learning loss is insufficient to solve the representation learning problem of encoding relevant information from the images. Many approaches [1, 2, 3, 4, 5] have added an auxiliary loss to encourage the encoder to form better representations. Other works [6, 7, 8] have achieved good performance by applying data augmentation, in the form of random pixel shifts, to the images during training. Data augmentation is an appealing technique for image-based RL due to its simplicity and effectiveness.

While data augmentation may be useful, it is not clear why it works so well. Specifically, why do random shifts significantly affect learning, given that the encoder is a nearly shift-equivariant convolutional network? To unpack this question further, we must consider two facts about [7, 6]: (1) the entities are centered in the image observations (so there is no benefit to generalizing to novel pixel translations); (2) the encoder used to process the image observations is made of convolutional layers which are known to exhibit shift-equivariance.

In this work, we investigate the representations learned when using data augmentations for image-based RL. We find that data-augmentation results in higher shift-equivariance in the convolutional layers compared to no augmentation. Further, we show that data augmentation results in qualitatively more robust feature maps, which cannot be achieved using traditional regularization techniques like dropout. This work is in progress, and we hope it will lead to insights that can further improve the sample efficiency of image-based RL methods.

2 Related Work

Reinforcement Learning from Images Many papers have focused on overcoming the representation learning problem for image-based RL. [1] proposes applying reconstruction loss to the convolutional encoder during training. CURL [2] uses a contrastive objective as an auxiliary loss during training, generating positive keys using random crop augmentations. Stooke et al. [9] train the encoder offline using a contrastive loss, such that the encodings are similar for nearby timesteps in a trajectory; they show that sub-pixel random shifts on the encoder's output can boost RL performance. Other works have applied auxiliary losses based on predictive information [3], forward modeling [4], and mutual information between augmented states [5]. DrQ [7] proposes averaging q-value estimates over multiple data augmentations during training, and claim that the augmentation serves as a regularizer to avoid overfitting. In concurrent work, RAD [6] compares multiple forms of data augmentation and find that random translations achieve best performance; they posit that augmentations result in more robust representations. [10] shows that excessive data-augmentation leads to higher q-target variance and over-regularization.

Shift-Equivariance of Convolutional Networks It is well known that the weight sharing scheme of convolutional networks leads to shift-equivariance. Azulay and Weiss [11] highlight that downsampling operations (stride or pooling) violate the classical sampling theorem and reduces shift-equivariance; performing data augmentation restores equivariance only for samples nearby the training distribution. Zhang [12] proposes a method for anti-aliased downsampling in convolutional networks, called BlurPool, that exhibits higher shift-equivariance. Cohen and Welling [13] generalize convolutional networks to exhibit the equivariance properties of arbitrary symmetry groups. To our knowledge, there is no work describing a mechanism, other than downsampling, that affects shift-equivariance of convolutional architectures.

3 Experiments

We are interested in uncovering the mechanisms underlying the effectiveness of data-augmentation for image-based RL. To narrow the scope of the investigation, we look at how modifications to the approach outlined in [6] affect the performance and representations learned. In this section, we briefly describe the learning algorithm for continuous-control, image-based RL, how to incorporate data augmentation during training, and how to quantitatively measure the equivariance of the encoder.

3.1 Method

Traditional continuous-control reinforcement learning algorithms can be applied to image-based tasks using a convolutional encoder. It is common to use a single convolutional encoder whose output is used for both the actor and critic networks. For this work, we use soft-actor critic (SAC) [14], where the convolutional encoder is made of four convolutional layers (see Fig. 1). To achieve higher performance, SAC can be improved by performing data augmentation during training. Following [6], we perform random translations on the state and next state images when calculating the actor and critic losses; we will refer to this approach as RAD-SAC in the figures below. For additional details on actor and critic networks or training details, see [6].

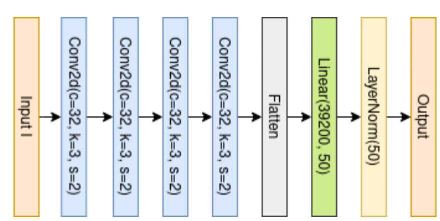

Figure 1: Encoder network for SAC

3.2 Domains

We run experiments on several tasks from the DeepMind Control Suite [15]. To perform data augmentation in the form of random translations, the environment images are rendered at 100×100 pixels and crops are randomly sampled at a size of 84×84 pixels. To provide full information to the agent, the agent recieves three stacked frames as an observation; augmentations are consistent across the frames. For additional details about the environments, see [6].

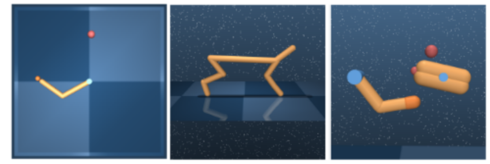

Figure 2: Tasks from DMControl Suite: reacher-easy, cheetah-run, finger-spin.

3.3 Metric

The most obvious explanation for why random-shifts impact learning of a convolutional encoder is that the augmentations lead to better shift-equivariance. Thus, we will measure how shift-equivariance is affected by augmentation. A network is shift-equivariant if shifting the input image results in a similarly shifted output feature map. From [12], the metric of shift-equivariance of a convolutional neural network, f_θ is calculated as:

$$\mathcal{M}_{equ}(f_\theta) = \mathbb{E}_{x \sim X}\left[\mathtt{sim}\big(f_\theta(\mathcal{T}(x)),\, \mathcal{T}'(f_\theta(x))\big)\right] \tag{1}$$

where $\mathtt{sim}$ is the cosine similarity function, $x \sim X$ is an image input sampled from training distribution, and $\mathcal{T}$ is an operator that translates images in pixel space, and $\mathcal{T}'$ is an equivalent operator in output pixel space (i.e. if f_θ downsamples the image by a factor of 2, then $\mathcal{T}'$ will translate by half the magnitude of $\mathcal{T}$). The shift-invariance of a neural network can be calculated by setting $\mathcal{T}'$ to the identity operator in Eqn. 1.

3.4 RAD-SAC vs. SAC

Conventional wisdom would suggest that the convolutional encoder of RAD-SAC is shift-equivariant so random-translations should have little impact on the convolutional features learned. As pointed out in Section 2, downsampling operations can lower shift-equivariance. We investigate the possible influence of the downsampling (which occurs in first convolutional layer via striding) in Figure 3. We compare the performance of RAD-SAC to a method where random translation occurs after the strided convolutional layer ('RAD-SAC after conv1') and to a method where SAC uses an anti-aliased downsampling operation from [12] instead of striding ('SAC + blurpool'). We observe minimal difference in performance when using an anti-aliasing downsampler, suggesting that achieving better shift-equivariance is not the driving mechanism behind data-augmentation. This is supported by the fact that applying random translations before or after the strided convolution layer has minimal impact on performance. This form of 'representational robustness' was noted in [6], however it was also observed for other forms of augmentation.

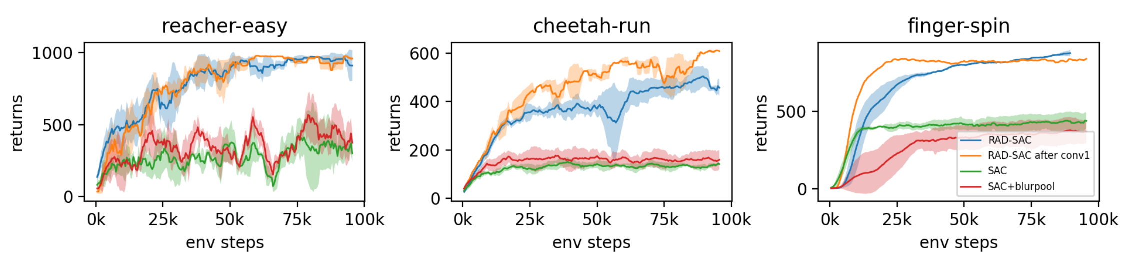

Figure 3: Learning curves on several DMControl tasks.

We show the measures of equivariance and invariance of the encoders after training on reacher-easy environment (Fig. 4). At the final convolutional layer (conv4), we observe a marked difference in shift-equivariance between the methods that used augmentation and those that did not. This suggests that augmentation does impact shift-equivariance, although the equivariance does not seem to be introduced by the downsampling operation. Perhaps more importantly, we observe that the methods with data-augmentation have much higher shift-invariance at the final convolutional layer. Invariance to shifts is desirable because the output of the encoder is flattened and processed by a linear layer.

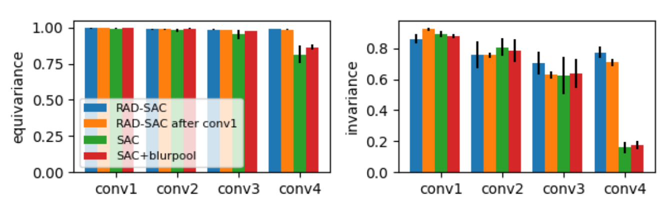

Figure 4: Measures computed using Eqn. 1 on random translations of 1 to 4 pixels. The network is measured after each layer so the value of 'conv3' means the equivariance up to, and including, the third convolutional layer.

3.5 RAD as a Regularizer

The results so far suggest that augmenting with random translations causes the learned feature maps to be more shift-equivariant *and* shift-invariant. This may seem counter-intuitive; however, it is possible when the output of the network is spatially-continuous. By inspecting the learned feature maps after the final convolution (Figure 5), we find that RAD-SAC generates feature maps with higher spatial continuity compared to when data augmentation is not used. Next, we test if this spatial continuity can be achieved via other means: applying dropout ($p = 0.9$) after final convolutional layer ('SAC+dropout'), applying gaussian blur ($\sigma = 2.$) to final feature map ('SAC+blur'), and adding regularization loss to enforce local spatial continuity ('SAC+regularizer'). The results (Fig. 6) show that dropout and blurring can achieve similar to RAD-SAC for some, but not all, environments. Thus, forming robust representations does not necessarily lead to success on the reinforcement learning task; increasing spatial continuity will lower how precise the network is about the spatial position of features.

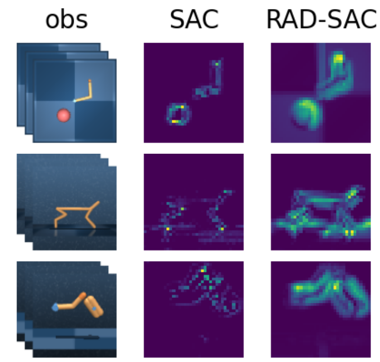

Figure 5: Spatial attention maps of final convolutional layer.

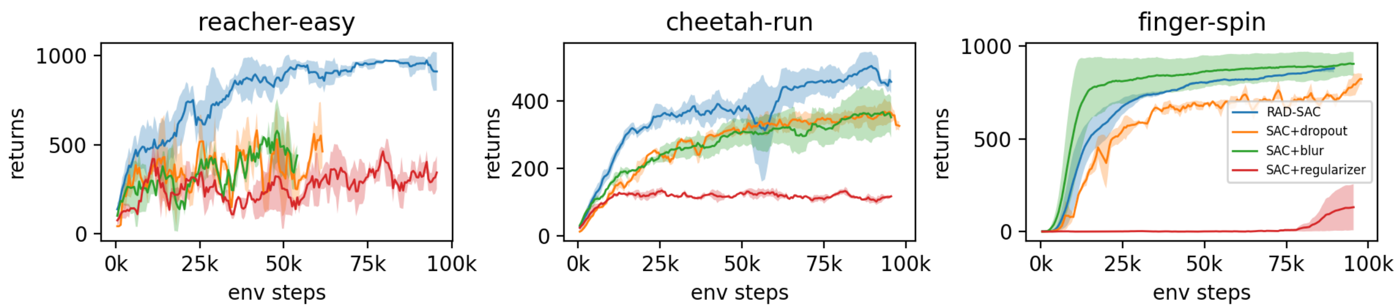

Figure 6: Comparing RAD-SAC to using traditional regularization techniques.

4 Conclusion

In this work, we investigated the effects of random translation augmentations on image-based RL for continuous control tasks. We find that augmentation of this sort results in encoders that have higher shift-equivariance and shift-invariance (i.e. have spatial consistency). We are unable to reliably achieve performance of RAD-SAC using regularization techniques alone. This suggests that we do not completely understand how data augmentation impacts the learning process. One possible direction to explore is inspired by the data from Fig. 7. It shows the shift-equivariance of two SAC agents trained on reacher-easy, one with 4 convolutional layers and the other with 5 convolutional layers. The shift-equivariance decreases based how far the layer is from the output. We hypothesize that the encoder learns to form non-equivariant representations that are more interpretable to the downstream linear layers. Designing a better interface between convolutional and linear layers may result in faster learning with less need for data augmentation.

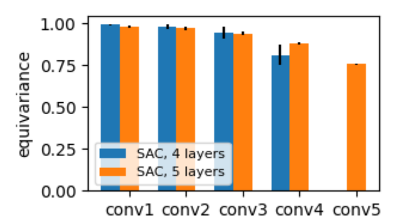

Figure 7: Shift-equivariance for encoders with different number of layers

References

[1] Denis Yarats, Amy Zhang, Ilya Kostrikov, Brandon Amos, Joelle Pineau, and Rob Fergus. Improving sample efficiency in model-free reinforcement learning from images. *arXiv preprint arXiv:1910.01741*, 2019.

[2] Michael Laskin, Aravind Srinivas, and Pieter Abbeel. Curl: Contrastive unsupervised representations for reinforcement learning. In *International Conference on Machine Learning*, pages 5639–5650. PMLR, 2020.

[3] Kuang-Huei Lee, Ian Fischer, Anthony Liu, Yijie Guo, Honglak Lee, John Canny, and Sergio Guadarrama. Predictive information accelerates learning in rl. *arXiv preprint arXiv:2007.12401*, 2020.

[4] Max Schwarzer, Ankesh Anand, Rishab Goel, R Devon Hjelm, Aaron Courville, and Philip Bachman. Data-efficient reinforcement learning with self-predictive representations. *arXiv preprint arXiv:2007.05929*, 2020.

[5] Nicklas Hansen and Xiaolong Wang. Generalization in reinforcement learning by soft data augmentation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13611–13617. IEEE, 2021.

[6] Michael Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srinivas. Reinforcement learning with augmented data. *arXiv preprint arXiv:2004.14990*, 2020.

[7] Ilya Kostrikov, Denis Yarats, and Rob Fergus. Image augmentation is all you need: Regularizing deep reinforcement learning from pixels. *arXiv preprint arXiv:2004.13649*, 2020.

[8] Denis Yarats, Rob Fergus, Alessandro Lazaric, and Lerrel Pinto. Mastering visual continuous control: Improved data-augmented reinforcement learning. *arXiv preprint arXiv:2107.09645*, 2021.

[9] Adam Stooke, Kimin Lee, Pieter Abbeel, and Michael Laskin. Decoupling representation learning from reinforcement learning. In *International Conference on Machine Learning*, pages 9870–9879. PMLR, 2021.

[10] Nicklas Hansen, Hao Su, and Xiaolong Wang. Stabilizing deep q-learning with convnets and vision transformers under data augmentation. In *Thirty-Fifth Conference on Neural Information Processing Systems*, 2021.

[11] Aharon Azulay and Yair Weiss. Why do deep convolutional networks generalize so poorly to small image transformations? *arXiv preprint arXiv:1805.12177*, 2018.

[12] Richard Zhang. Making convolutional networks shift-invariant again. In *International conference on machine learning*, pages 7324–7334. PMLR, 2019.

[13] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016.

[14] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*, pages 1861–1870. PMLR, 2018.

[15] Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, Nicolas Heess, and Yuval Tassa. dm_control: Software and tasks for continuous control. *Software Impacts*, 6:100022, 2020.