END-TO-END REINFORCEMENT LEARNING ON REACHER

A PREPRINT

Zihan Ding

May 1, 2019

1 Basic Settings

The reinforcement learning (RL) directly uses the screenshot images as inputs, and the task is to reach a fixed target goal position on Reacher environment. The algorithm used in experiments is Soft Actor-Critic (SAC). With screenshot RGB image of size 200*200*3 and grey-scale image of size 200*200*1 (single channel) as inputs, a pre-trained convolutional auto-encoder (AE) is used for compressing the images. And the encoded low-dimensional vectors are used as encoding of states for RL. The reason of applying an AE is because directly appling an image of size 200*200*3=120000 as input state could be quite large and therefore hard for RL to learn, which is already testified with experiments.

As the original screen size is 1000*1000, a down-sampling process is actually taken with a factor of 5. In order to make the objects in the down-sampled images still easy to be recognized, I actually enlarge and widen the objects in the original images, and the original screenshot of Reacher environment looks like in Fig. 1.

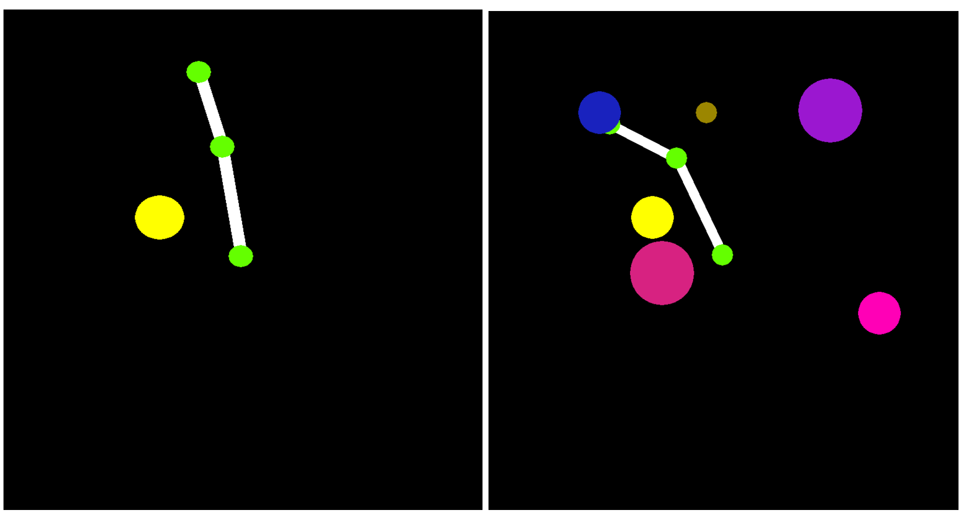

Figure 1: Original screenshot for 3 channels RGB images, the only one goal position is the yellow circle in each of two kinds of environments . The right environment is with lots of colorful objects (for the purpose of confusing the agent only) compared with the left one.

2 Results

2.1 Auto-encoder

The results of AE are shown in Fig. 2-6, with different channels of images and different latent encoding dimensions. The Fig. 2-4 are using environment on the left of Fig. 1 and Fig. 5-6 are with the right one in Fig. 1. One point to be noticed is that **the target goal position can sometimes be reconstructed imprecisely**.

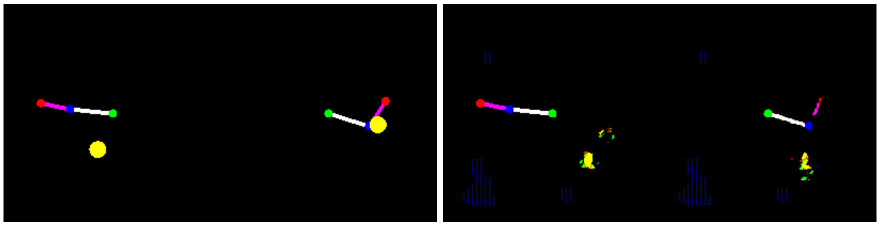

Figure 2: Original screenshot for 3 channels RGB images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 64.

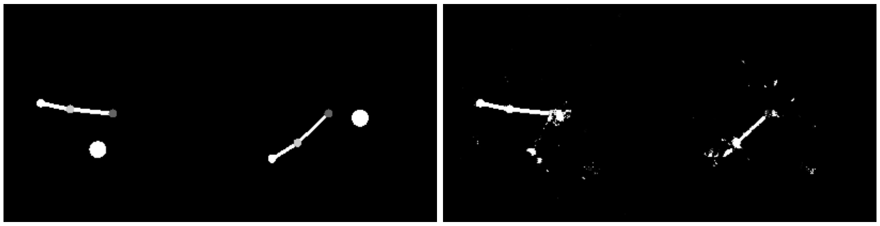

Figure 3: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 64.

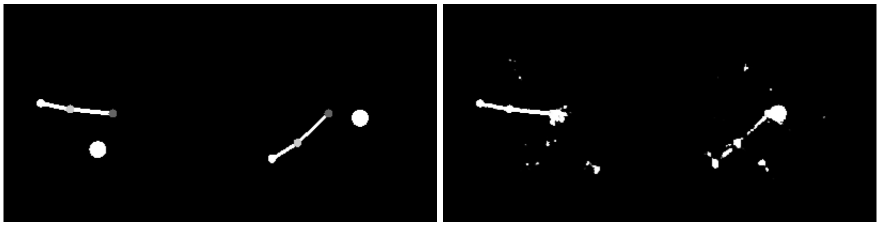

Figure 4: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 16.

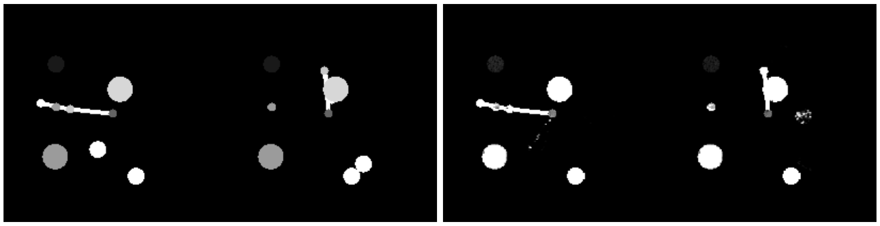

Figure 5: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 16.

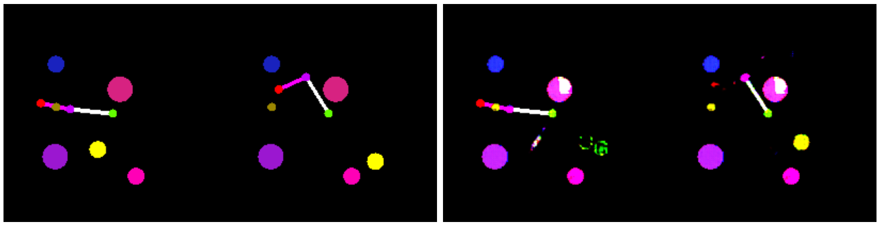

Figure 6: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 16.

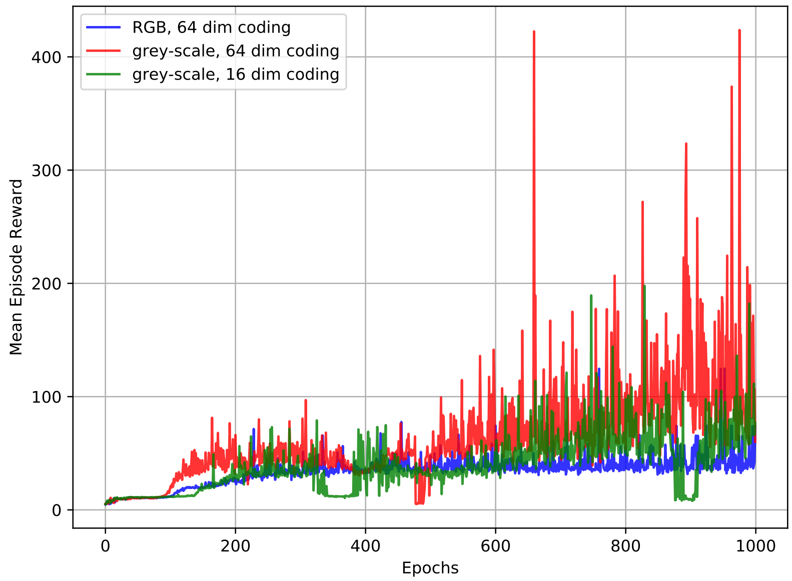

Figure 7: Reinforcement learning curves after encoding with auto-encoder, for RGB or grey-scale images with different encoding dimensions as state representation.

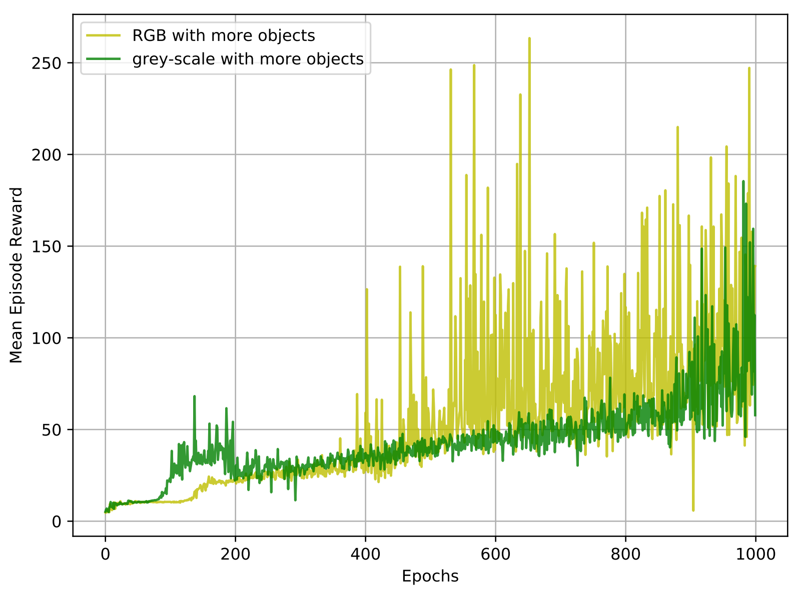

Figure 8: Reinforcement learning curves after encoding with auto-encoder (64-dimension coding), for RGB or grey-scale images as state representation.

2.2 Reinforcement Learning

The learning performances with end-to-end settings using AE is shown in Fig. 7, 8. Fig. 7 is using the left environment of Fig. 1, and Fig. 8 is using the right environment of Fig. 1 with more colorful objects.

From Fig. 7 we can derive two results:

(1). **the grey-scale images as state representation shows better learning performances compared with the RGB images**. This can be interpreted as the grey-scale image is a more compressed representation with only useful information, compared with the RGB one.

(2). **lower dimensional encoding will hurt the learning performances**, because of the more information loss in the coding layers of AE. And it also makes the reinforcement learning process unstable.

From Fig. 8, we can derive a result that:

when there are more colorful objects in the environment, which could be similar to the goal object except from the color, **the grey-scale images do not show predominance over RGB images as state representation for reinforcement learning process anymore**. On the contrary, as the distinctive color feature of important goal object is lost in grey-scale images, it may confuse the RL agent and therefore make it harder to learn.

2.3 Conclusions

* The RL is relatively robust for the AE coding results. This can be testified with that the target goal position can sometimes be reconstructed imprecisely with AE as shown in the reconstructed images, but this does not hurt the RL learning process much. And the RL works even with different size of objects in images using the AE.

* When the RGB-to-grey-scale transformation does not lose any important features for screenshots of the environment, the grey-scale images as state representation will help with subsequent reinforcement learning process. But when colors contain useful information for the agent to learn, the grey-scale images as state representation will confuse the agent more than the RGB images, and therefore hurt the reinforcement learning performances.

* Low-dimension encoding in AE, like 16-dimension encoding of grey-scale images, could hurt the learning performances of RL, making it unstable and slower than a high-dimension encoding (but still better than the 64 dimension RGB images).

References