
END-TO-END REINFORCEMENT LEARNING ON REACHER

A PREPRINT

Zihan Ding

April 30, 2019

1 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 applying 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.

2 Results

2.1 Auto-encoder

The results of AE are shown in Fig. 1,2,3, with different channels of images and different latent encoding dimensions. One point to be noticed is that **the target goal position can sometimes be reconstructed imprecisely**.

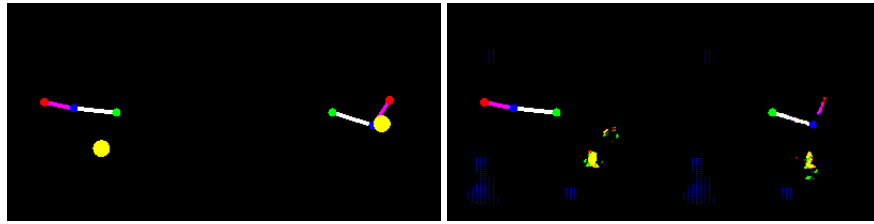


Figure 1: Original screenshot for 3 channels RGB images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 64.

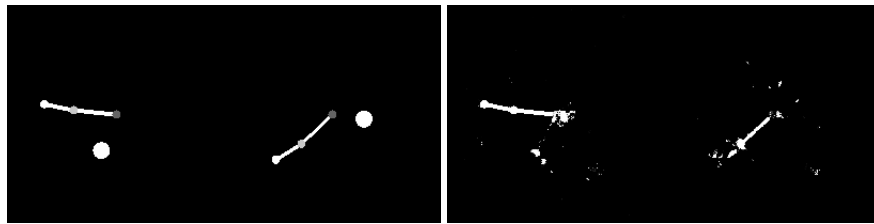


Figure 2: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 64.

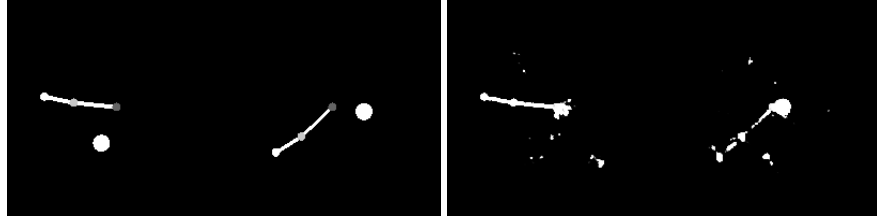


Figure 3: Original screenshot for 1 channel grey-scale images (2 random sampled images) and reconstructed images after training the AE for 300 epochs with encoding dimension of 16.

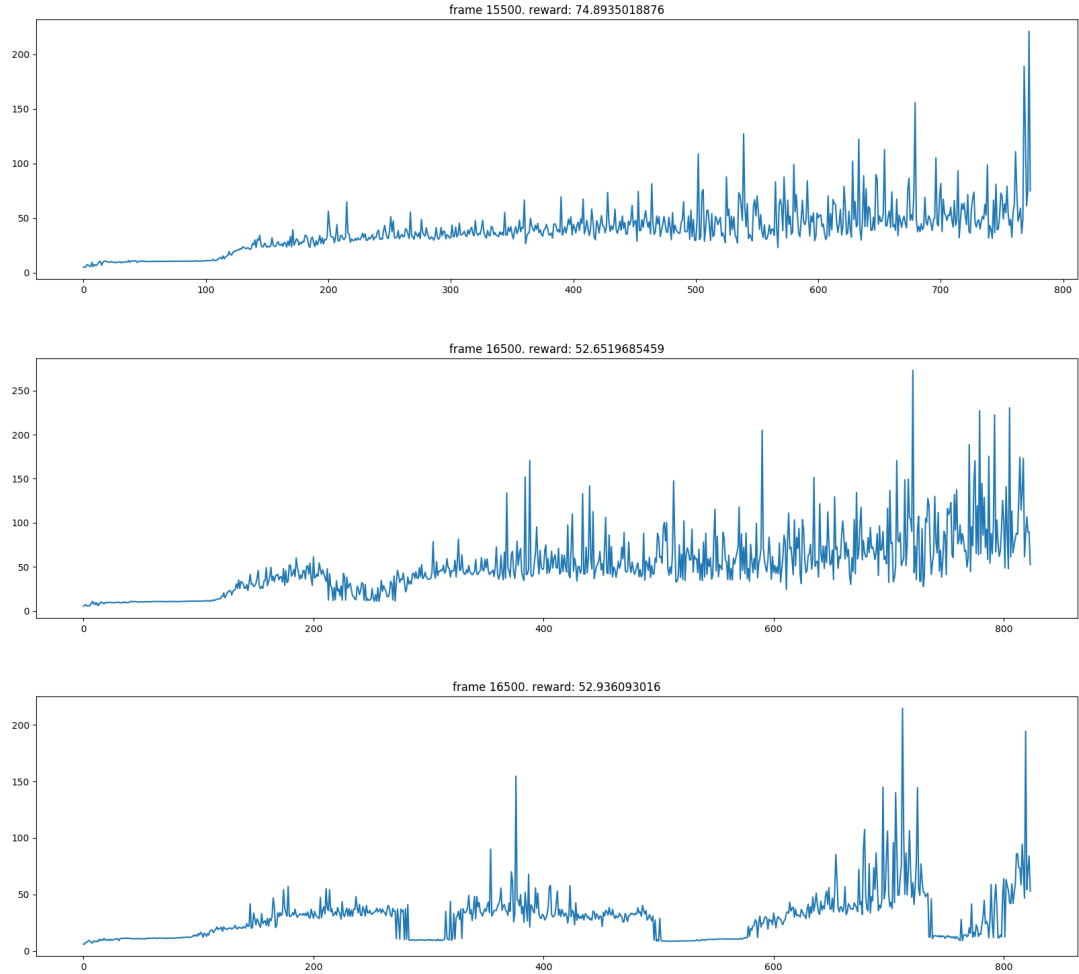


Figure 4: From top to bottom are using: 64-dimension coding for 3-channel RGB images, 64-dimension coding for 1-channel grey-scale images and 16- dimension coding for 1-channel grey-scale images, respectively. The learning curves do not show great difference for different settings.

2.2 Reinforcement Learning

The learning performances with end-to-end settings using AE is shown in Fig. 4. It seems using grey-value images do not show better learning performances compared with using RGB images, for both 64-dimension and 16-dimension encoding. And for 16-dimension encoding of grey-scale case, the learning process could be unstable.

2.3 Conclusions

- * 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.
- * It seems using grey-value images do not show better learning performances compared with using RGB images, for both 64-dimension and 16-dimension encoding.
- * Low-dimension encoding like 16-dimension of grey-scale images could even hurt the learning process of RL, making it unstable.

References