Paper summary: Improving sample efficiency in model-free reinforcement learning from images

May 22, 2022

1 Idea in few sentances

2 Explanation of the central concept

3 Methodology

4 Initial rambly notes

4.1 Abstract

Fitting a high-capacity encoder to extract features (state information) from images with only the reward signal leads to poor performance. One option is to incorporate reconstruction loss into an off-policy algorithm, but that often leads to training instability. Investigtion into why shows variational autoencoders to be a problem.

4.2 Introduction

Some solutions to low sample efficiency are:

1. use an off-policy algorithm
2. add an auxiliary task with an unsupervised objective

The simplest auxiliary task is an autoencoder with a pixel reconstruction objective. Prior works uses a two-step training procedure, but this often leads to lower final performance.

> We confirm that a pixel reconstruction loss is vital for learning a good representation, specifically when trained jointly, but requires careful design choices to succeed.

4.3 Method

4.4 Other stuff