Paper summary: Reinforcement learning with latent flow

May 6, 2022

1 Idea in few sentances

2 Explanation of the central concept

3 Methodology

4 Initial rambly notes

4.1 Abstract

Temporal information is essential for learning effictive policies with RL. Current state-of-the-art RL algorithms either assume that this information is present in the state space, or give it via a heuristic like frame stacking. However, in current video classification, explicit encodings of temporal information like optical flow or two-stream architectures are used in state-of-the-art methods. Inspired by this, the paper introduces Flow of LAtents for REinforcement learning (Flare) — a network architecture for RL which explicitly encodes temporal information through latent vector differences. Flare achieves the same performance as state-based RL (but without access to the state velocity, only positional state information) and is the most sample-efficient model-free pixel-based RL algorithm on 500k and 1M step benchmarks.

4.2 Introduction

RL holds a promise (amen brother)... Passing a stack of most recent frames as an input to the CNN can be intepreted as a form of early fusion. In contrast, modern video recognition systems employ optical flow and late fusion, where individual frames are individually processed with CNN layers before fusion and downstream processing. However, integrating such an approach to RL is not trivial. Flare is the proposed solution and it can be intepreted as a *structured late fusion* architecture. Along with the achievements listed in the abstract,

Rainbow + Flare outperforms the baseline on 5/8 challenging Atari games at the 100M step benchmark.

CNN with frame stacking was the way until 2016. Then CNN for embedding followed by an LSTM to aggregate temporal information became the best thing.

4.3 Method

4.4 Other stuff