

DQN-Atari Project Overview

Introduction	1
Project Structure	1
Notebook Contents	1
Memory Optimization	1
AtariPreprocessing Adjustment	2
Training Considerations	2
Results	2
Future Plans	2

Introduction

This project, inspired by DeepMind's paper on "Human-level control through deep reinforcement learning," is part of the Reinforcement Learning course at EPITA in 2023. The primary objective is to implement and train a Deep Q-Network (DQN) on the Atari Breakout game using the gym and torch modules.

Project Structure

The implementation resides in the ``torch_DQL.ipynb`` notebook. The project adheres to the parameters outlined in the original paper, with the exception of reducing memory size to accommodate limited RAM. Additionally, a preprocessing step has been introduced to use luminance instead of grayscale for image representation.

Notebook Contents

The notebook provides a straightforward usage of functions, featuring the DQN network in ``DQN_torch.py``, memory management in ``ReplayBuffer.py``, and a video creation function in ``create_video.py``.

Memory Optimization

To address memory constraints, the memory size was reduced to fit in limited RAM. The use of ``uint8`` variables in the ``ReplayBuffer`` instead of ``float32`` was adopted for quantization, significantly reducing memory requirements. Consequently, no image scaling (normalization) is performed.

AtariPreprocessing Adjustment

An adjustment in ``preprocess.py`` introduces a ``luminance_obs`` boolean parameter in the AtariPreprocessing configuration, enabling the use of luminance instead of grayscale.

Training Considerations

Training duration can be extensive, dependent on hardware and the chosen number of episodes. Despite using a GPU, stability issues in training may arise, impacting reward outcomes.

Results

On a system with limited RAM (memory size: 15,000), the agent was trained on thousands of episodes, achieving a mean reward of 15. On a higher-RAM system (memory size: 200,000), approximately 6,000 episodes (>2M frame_count) resulted in a mean reward of 27.

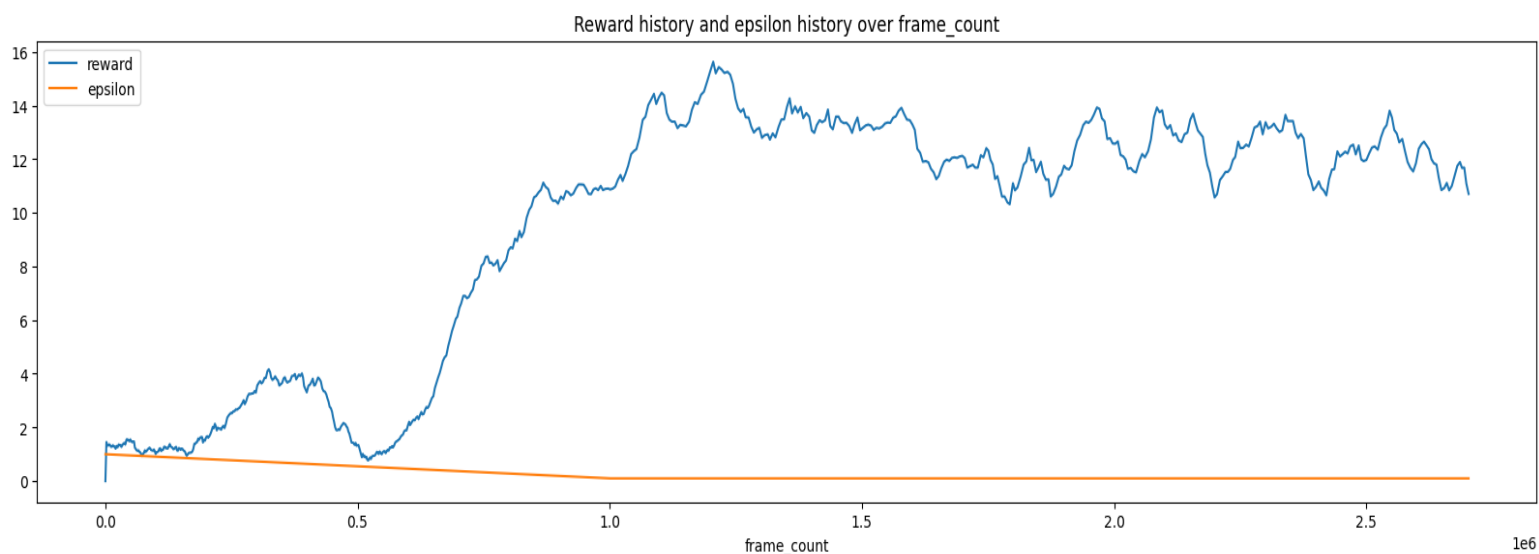


Figure 1: History of reward over frame_count using a memory size of 15 000

Future Plans

The project aims to explore advanced approaches by implementing a model playing Breakout based on a Decision Transformer, paving the way for further experimentation and improvement.