DQN-Atari Project from DeepMind paper on Breakout game

Hugo Deplagne[1]

Abstract: Reinforcement Learning Course, EPITA.

Keywords: DQN; Breakout

1 Intro

This research initiative, drawing inspiration from DeepMind's seminal work on "Human-level control through deep reinforcement learning,"constitutes a pivotal component of the academic curriculum within the context of the Reinforcement Learning course at EPITA in 2023. The overarching goal is to intricately design and execute a Deep Q-Network (DQN) to engage with the intricacies of the Atari Breakout game, using gym and torch modules as foundational frameworks.

1.1 Références

Original paper [Vo15].
Github code [H 23].

2 EPITA, Le Kremlin-Bicêtre France, hugo.deplagne@epita.fr,

2 Project Structure

The architectural framework of this endeavor is encapsulated within the `torch_DQL.ipynb` notebook, meticulously structured to align with the parameters elucidated in the original paper. An exception to this adherence lies in the judicious reduction of memory size, an imperative step necessitated by the constraints imposed by limited RAM. A deliberate preprocessing step has been introduced, leveraging luminance in lieu of grayscale for enhanced fidelity in image representation.

2.1 Notebook Contents

The intricacies of the implementation are delineated within the notebook, offering a coherent and comprehensible exposition of the utilized functions. Notable components include the DQN network encapsulated in `DQN_torch.py`, memory management exemplified in `ReplayBuffer.py`, and a video creation function embodied in `create_video.py`.

2.2 Memory Optimization

Addressing the inherent constraints of memory capacity, an astute reduction in memory size was implemented to accommodate the limitations of the available RAM. The strategic implementation of `uint8` variables in the `ReplayBuffer` in lieu of `float32` facilitated quantization, resulting in a substantial reduction in memory requisites. Consequently, the absence of image scaling or normalization is a deliberate and consequential choice.

2.3 AtariPreprocessing Adjustment

A nuanced modification introduced in `preprocess.py` incorporates a `luminance_obs` boolean parameter within the AtariPreprocessing configuration. This strategic adjustment empowers the utilization of luminance as opposed to grayscale, contributing to an enriched and nuanced image representation.

2.4 Training Considerations

The temporal dynamics of training are contingent upon the intricacies of the hardware configuration and the chosen number of training episodes. Despite the utilization of a Graphics Processing Unit (GPU), the stability of the training process may exhibit occasional fluctuations, thereby influencing the outcomes of reward metrics.

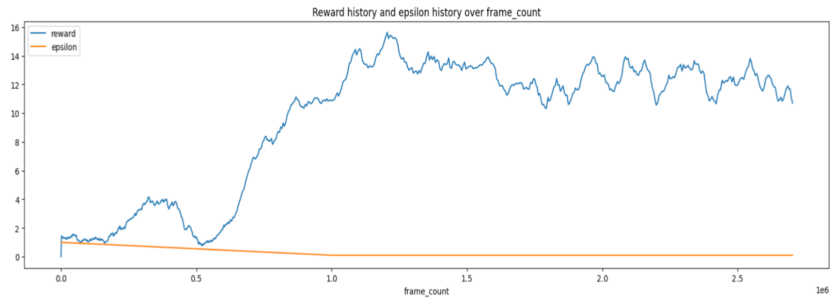

Fig. 1: History of reward over frame_count using a memory size of 15.000

2.5 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. Consequently, the best agent reaches a score of 46.
See Fig. 2.

2.6 Future Plans

This academic pursuit is poised to advance into uncharted territories by delving into sophisticated methodologies. The envisioned trajectory includes the implementation of a model proficient in playing Breakout, grounded in the

Fig. 2: Best agent end of play

principles of a Decision Transformer. This forward-looking initiative is poised to catalyze further experimentation and iterative refinement in the pursuit of enhanced performance and understanding.

Bibliography

[H 23] H. Deplagne: Torch model for DQN Breakout, 2023, URL: https://github.com/Coartix/ReinforcementLearningCourse/tree/master/DQN-Atari.

[Vo15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Hassabis: Human-level control through deep reinforcement learning.

Nature 345/, pp. 603–604, 2015, URL: https://www.nature.com/articles/nature14236.