

Rapport d'Entraînement DQN sur Breakout

COLLEMARE LUCAS

1 Objectif

Ce projet consiste à entraîner un agent Deep Q-Network (DQN) sur le jeu **Atari Breakout**.

2 Méthodologie

Le réseau utilisé suit l'architecture de DQN : trois convolutions successives (32 filtres en 8×8 stride 4, puis 64 filtres en 4×4 stride 2, puis 64 filtres en 3×3 stride 1), suivies d'une couche fully connected de 512 neurones qui produit les valeurs Q pour chaque action. L'environnement ALE/Breakout-v5 est prétraité avec un redimensionnement en 84×84 , une conversion en niveaux de gris et l'empilement des quatre dernières frames, auxquels s'ajoutent la désactivation des sticky actions, une courte phase de warm-up, ainsi qu'un clipping des récompenses où la valeur retournée est $-1, 0, +1$.

L'apprentissage repose sur un replay buffer, un réseau cible mis à jour régulièrement, une Huber Loss combinée à un gradient clipping, une mise à jour (Double DQN) et une politique ϵ -greedy décroissante de 1.0 à 0.1 au fil des frames.

3 Résultats

	Valeur
Reward maximal	341
Reward moyen (5k frames)	142

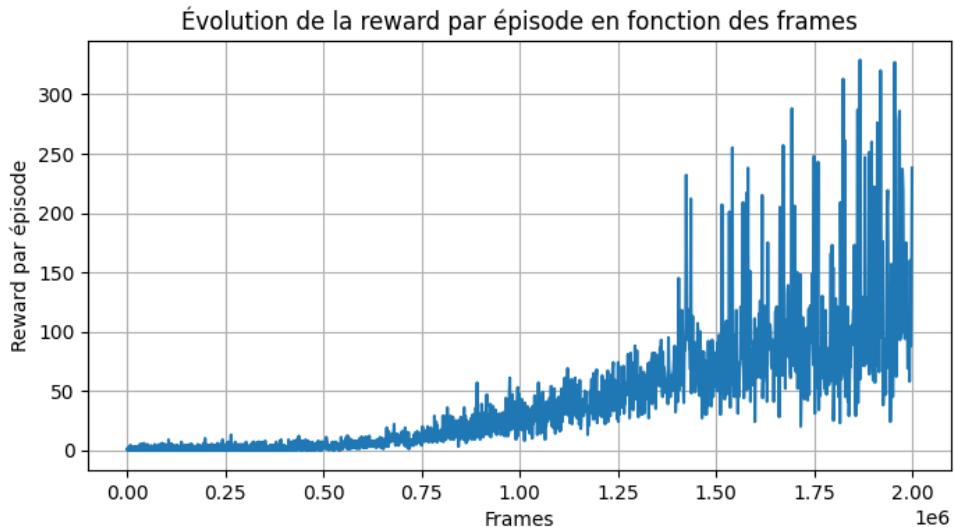


Figure 1: Évolution de la récompense en fonction des frames.

4 Conclusion & Analyse

Un entraînement plus long, un buffer plus large ou une modification de la fréquence de mise à jour du réseau cible pourraient améliorer les résultats, mais augmenteraient significativement le temps nécessaire et les ressources mobilisées pour l'apprentissage.