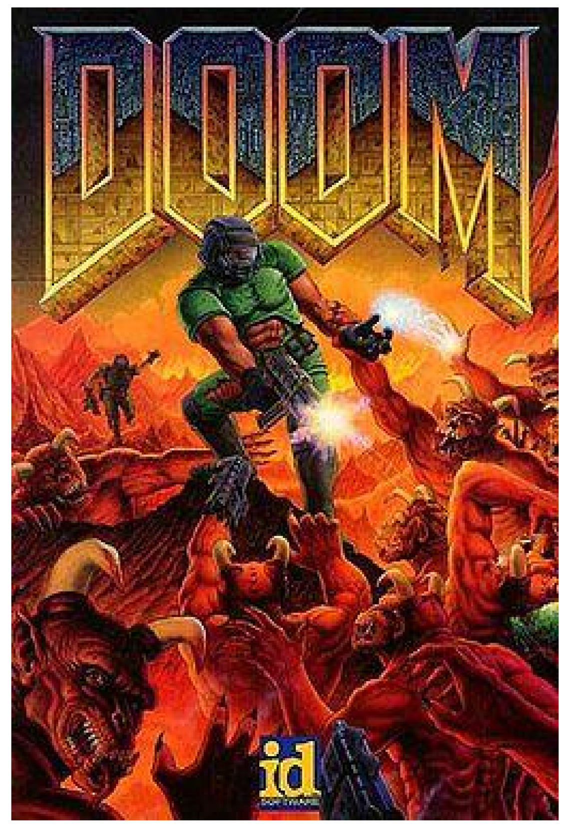
Prioritized Experience Replay Kills on Doom

Presenters: Brian Su*, Cem Koc*, Can Koc*

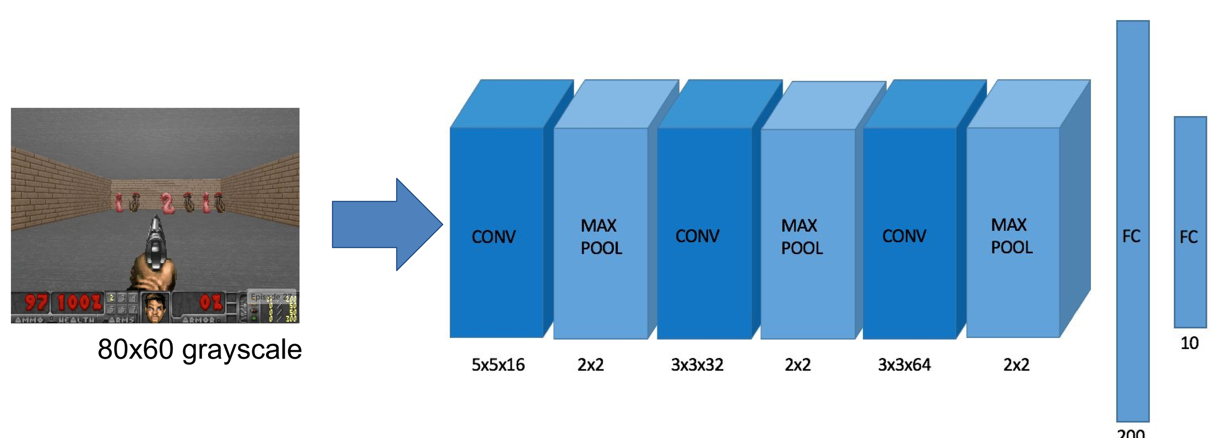
Dept of EECS, University of California Berkeley

Introduction

Since the publication of Deep Q-learning Networks (DQN) in *Playing Atari with Deep Reinforcement Learning*, many improvements have been made to improve the performance and reduce the training time of DQNs. In this work, we apply Prioritized Experience Replay from Schaul et al. 2016 on Doom, a first-person shooter game, and observe substantial performance improvements and training time reduction compared to the baseline DQN model on the Doom game modes "basic" and "defend the line."



Data Source and Model Architecture



Prioritized Experience Replay Architecture

Data is obtained from ViZDoom which is an implementation of the game to do deep learning. Preprocessing is done to reduce the image size and grayscale before inputting to the model.

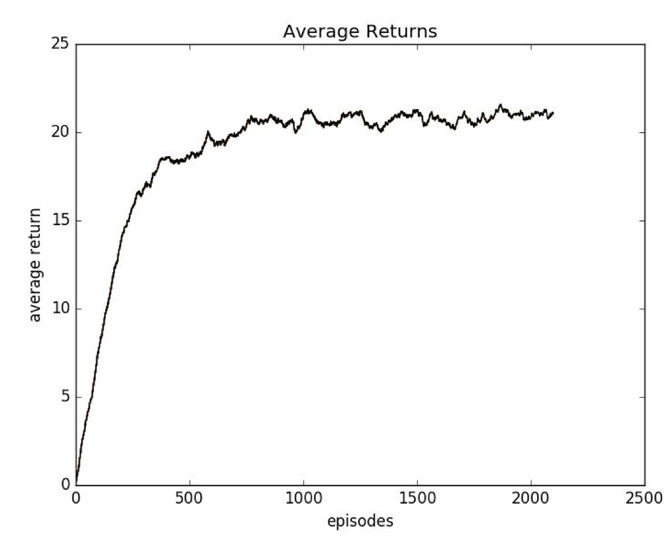
Used the DDQN model architecture from [2] with an additional convolutional layer. As it is suggested in [6] implemented a priority queue on top of the replay memory to rank the experiences by TD error.

Results on Game: Defend the Line

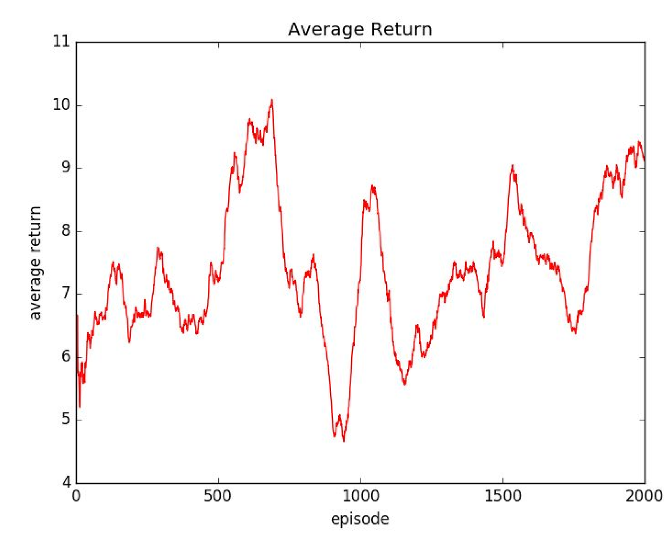
Performance metrics:

- Kill 16 monsters and get 15 points.
- +1 for killing a monster
- -1 for dying.

We trained a game of 2500 episodes with 300 steps/episode. Prioritized experience replay achieved much better rewards and saturated after 1000 episodes.

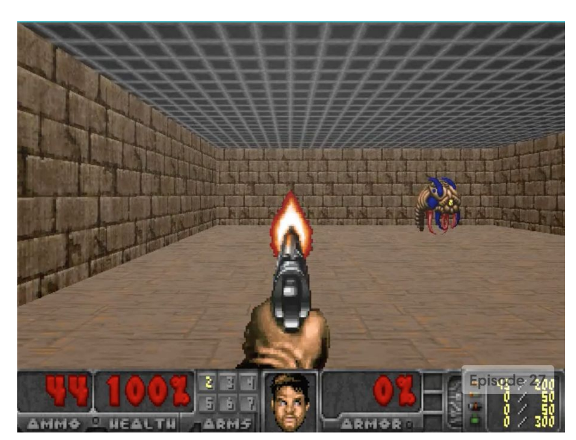


Prioritized Experience Replay from [6]



Baseline DQN from [1]

Results on Game: Basic



Performance metrics:

- Kill one monster in 3 seconds with 1 shot.
- +101 pts for killing monster
- -5 for missing a shot.
- -1 for every 0.028 second.

We trained a game of 10,000 episodes with max 400 steps/episode. Prioritized exp. replay achieved better rewards than baseline.

Prioritized Experience Replay from [6]

Conclusion

Our results from training a Prioritized Replay Experience DDQN on "defend the line" and "basic" game modes show that we obtain higher rewards, faster training times, and less stochastic average return performance compared to the baseline DQN model. Given the complicated nature of Doom, with a variety of maps, items to be picken up such as health kits, different monsters, we expect prioritized experience replay to have a noticeable impact when training on complex game modes such as "deathmatch."

Future Work

Prioritized Experience Replay is an example of current work to discover ways to reward an agent for exploring its environment. Intuitively, by encouraging exploration, the RL agent will stumble onto novel kinds of environments that allow for greater rewards. Using PER, an agent trains itself more frequently on experiences that yielded unexpected rewards. Our results from training a PER DDQN on Doom show that this direction of research is promising.

We would like to evaluate prioritized replay experience on more complicated game modes, such as "deathmatch." In addition, we would like to evaluate the recently released current state of the art, asynchronous advantage actor critic (A3C) on Doom. Based on the paper, A3C should allow us to reduce our reliance on GPUs and use CPUs instead to train.

References

1. Mnih et.al.Playing Atari with Deep Reinforcement Learning, 2013.
2. Hasselt et.al.Deep Reinforcement Learning with Double Q-Learning, 2015.
3. Wang et.al.Dueling Network Architectures for Deep Reinforcement Learning,2015
4. Dosovitskiy Alexey, Koltun Vladlen. Learning to Act by Predicting the Future, ICLR, 2017
5. Mnih et al., Asynchronous Methods for Deep Reinforcement Learning. in ICML 2016
6. Schaul et al, Prioritized Experience Replay. in ICLR 2016

*: contributed equally