

# Asynchronous Advantage Actor-Critic Methods on **ViZDOOM**

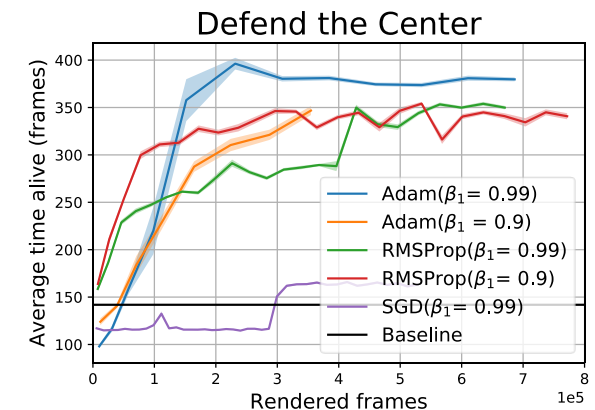
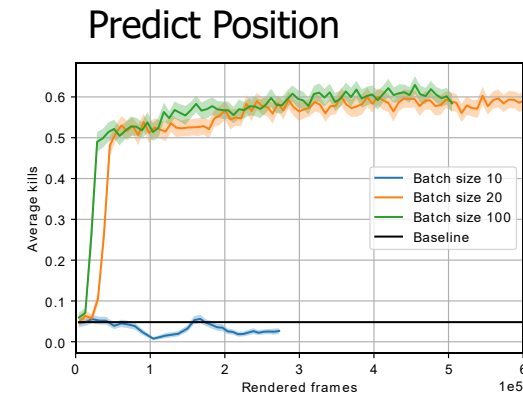
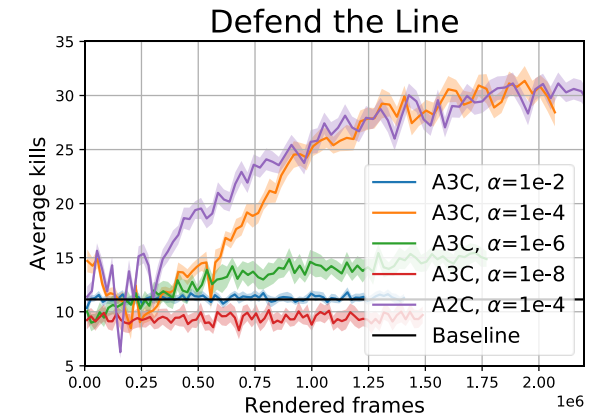
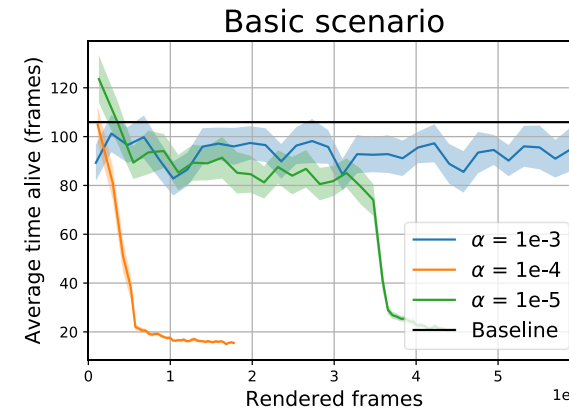
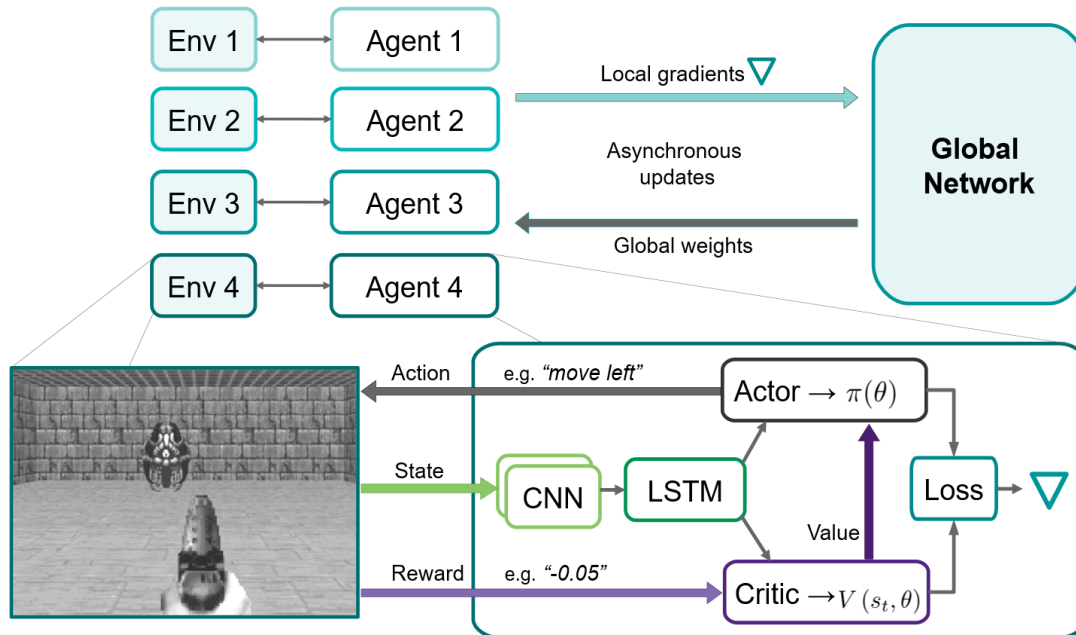
Aksel Obdrup (s144089), Elin Dijkstra (s182774), Louis Hein (s181573) & Tobias Konradsen (s144077) – DTU Compute – 02460 Advanced Machine Learning

Reinforcement learning algorithms learn behavior in various environments based on rewards for actions.

ViZDoom: Partially observable environment with sparse or delayed rewards

Other implementations based on experience replay are computationally heavy and potentially unstable

→ Asynchronous approach proposed



Learning is sensitive to experiment settings

- Learning rate: Optimal found  $10^{-4}$
- A2C/A3C: Similar learning curve
- Batch size: Scenario dependent, larger than 10
- Optimizers: Adam outperforms RMSProp and SGD

Non trivial behavior on all presented scenarios