THE BIOROBOTICS INSTITUTE

Not-so-Catastrophic Forgetting in Deep Reinforcement Learning on Sequential Atari

Giacomo Spigler
The BioRobotics Institute, Scuola Superiore Sant'Anna

Abstract

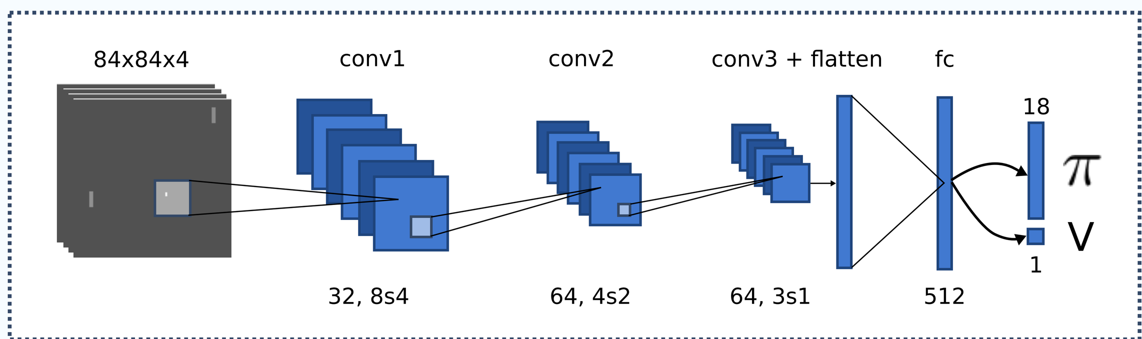
In this work we analyze the amount of forgetting on sequential Atari games learning using Deep Reinforcement Learning. Specifically, we show that while catastrophic forgetting is found to be highly disruptive in terms of the behavior of an agent, the changes to the parameters of its neural network are not as drastic as it could seem from the large drop in performance. Indeed, it is found that relatively short periods of retraining on previously learnt tasks are often sufficient to quickly recover and improve the lost performance.

Methods

- Training on pairs of Atari games from OpenAI Gym [2] and ALE [3].

Game A (10M frames) → Game B (10M frames) → Game A (10M frames)

- GA3C/A3C algorithm [4,5].



- Elu activations (except linear for the critic V and softmax for the actor π); training loss augmented with entropy loss on policy (beta=0.01); RMSProp optimizer with lr=0.0003, decay=0.95, momentum=0.0, eps=0.01; gradients clipped to [-1,1] and then globally with clip norm=50; rewards clipped to [-1,1]; lambda=0.99.

Motivation

- Sequential learning in neural networks is known to disrupt performance on previously learned tasks (catastrophic forgetting [1]).
- Here we set to analyze the amount of forgetting in Deep Reinforcement Learning on the sequential Atari benchmark.

- **Specifically**: are the changes to the parameters of the agent's neural network *major* or can previous performance be recovered *without re-training it from scratch.*

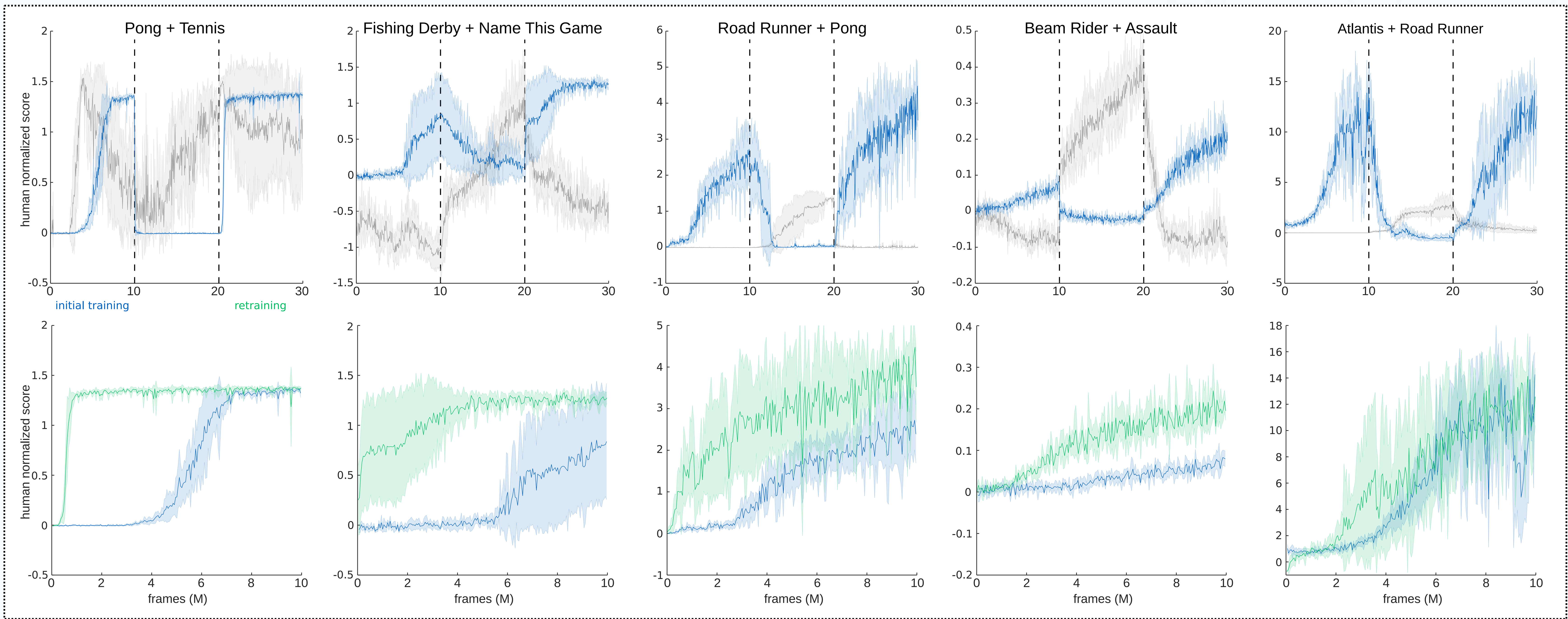
- **Idea**: measure the degree of catastrophic forgetting due to sequential training by the *amount of retraining required for the agent to recover the performance lost due to interference.*

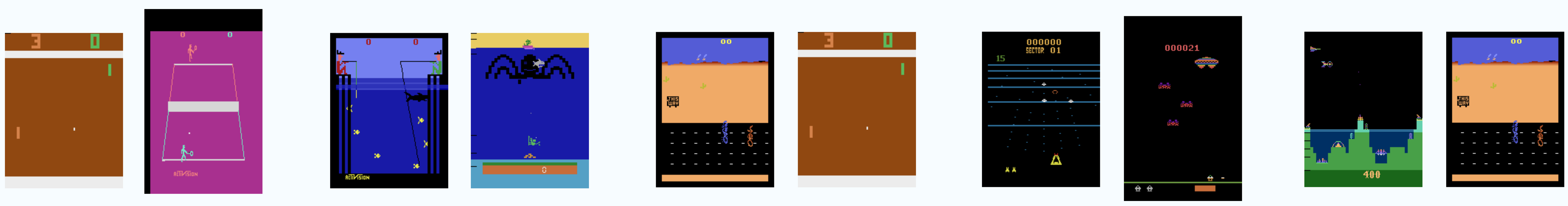
Results

- Training on the second game leads to quick disruption of the acquired performance on the first game (catastrophic forgetting).

- However, retraining on the first game was found to quickly recover and improve the performance achieved after first training on it.
- Shorter training required to recover the lost performance.

- The amount of interference and recovery was found to be specific to the game pairs used.





Conclusions

- The amount of forgetting in the network is lower than what could be thought from the drop in performance.
- Forgetting was highly disruptive in terms of the behavior of the agent, but the actual changes to its network's parameters were probably limited.
- This is in line with the hypothesis that in very high-dimensional weight spaces it may not be difficult to find parameters optimal for a task that are close to the parameters that are optimal for other tasks, as suggested in [6].

Acknowledgments

A sincere thanks to Microsoft for providing cloud computing resources via a *Microsoft Azure for Research award* (2017/2018).

Thanks to the organizers of *TMLSS 2018* for the great work in organizing the summer school!

References:

[1] M. McCloskey and N. J. Cohen, "Catastrophic interference in connectionist networks: The sequential learning problem," Psychology of learning and motivation, 1989
[2] G. Brockman et al., "Openai gym," 2016
[3] M. G. Bellemare et al., "The arcade learning environment: An evaluation platform for general agents," JAIR, 2013
[4] M. Babaeizadeh et al., "Reinforcement learning thorugh asynchronous advantage actor-critic on a gpu," ICLR, 2017
[5] V. Mnih et al., "Asynchronous methods for deep reinforcement learning," ICML, 2016
[6] J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," PNAS, 2017

Contact: Giacomo Spigler
giacomo.spigler@santannapisa.it
Scuola Superiore Sant'Anna, The BioRobotics Institute
Viale Rinaldo Piaggio, 34, 56025, Pontedera PI