RE: DATA-EFFICIENT REINFORCEMENT LEARNING WITH SELF-PREDICTIVE REPRESENTATIONS

Reinforcment Learning Project

Florian Langenhagen, Tri Hiep Bui and Nicolas Steiner

Leibniz University Hannover

March 4, 2022

Motivation

Goal: Learn on limited interactions $\Rightarrow$ efficient RL

- Combine successful approaches like
  - Self-supervised representations by constrastive learning
  - Data augmentation
  - Rainbow: Combining Improvements in Deep Reinforcement Learning [1]
- Self-predicitve representations $\Rightarrow$ Improve sample efficiency
  - Dynamic model predicts future latent representations

Architecture: Self-predictive Representations [2]

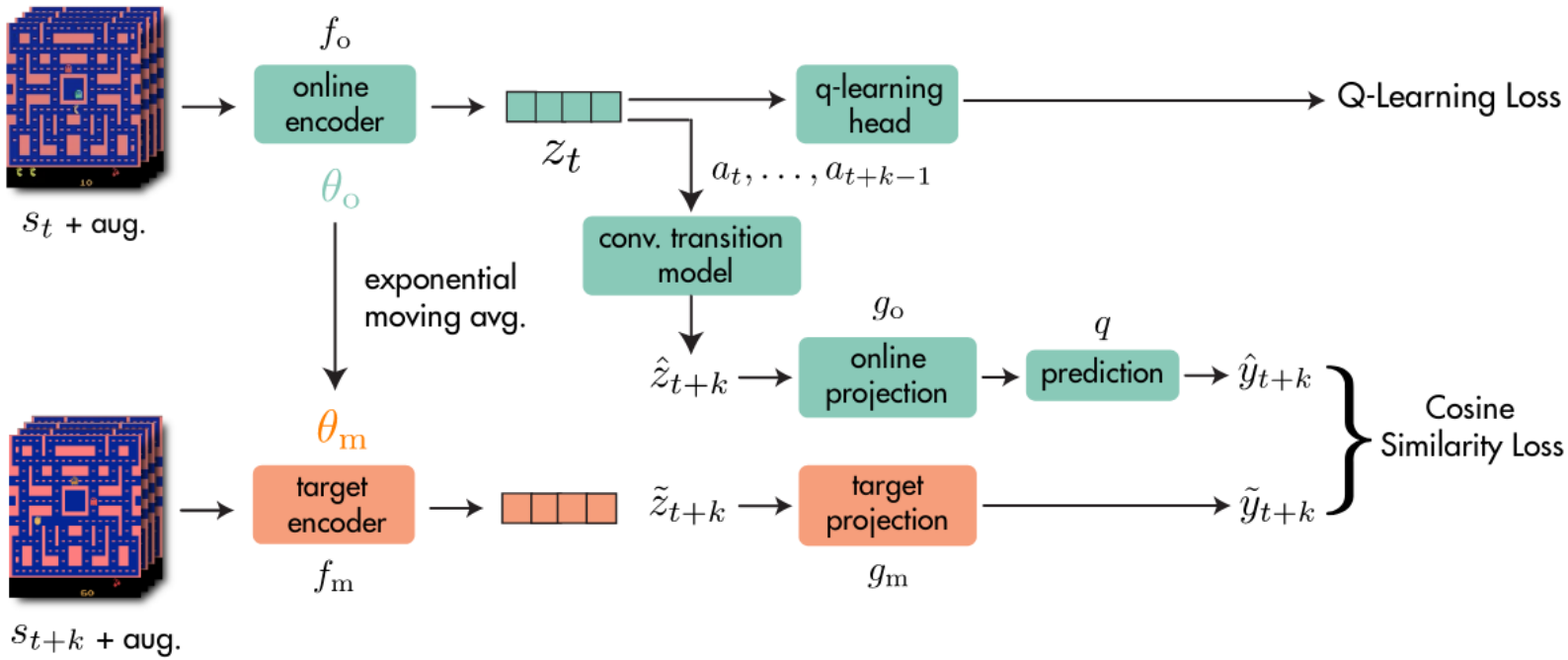
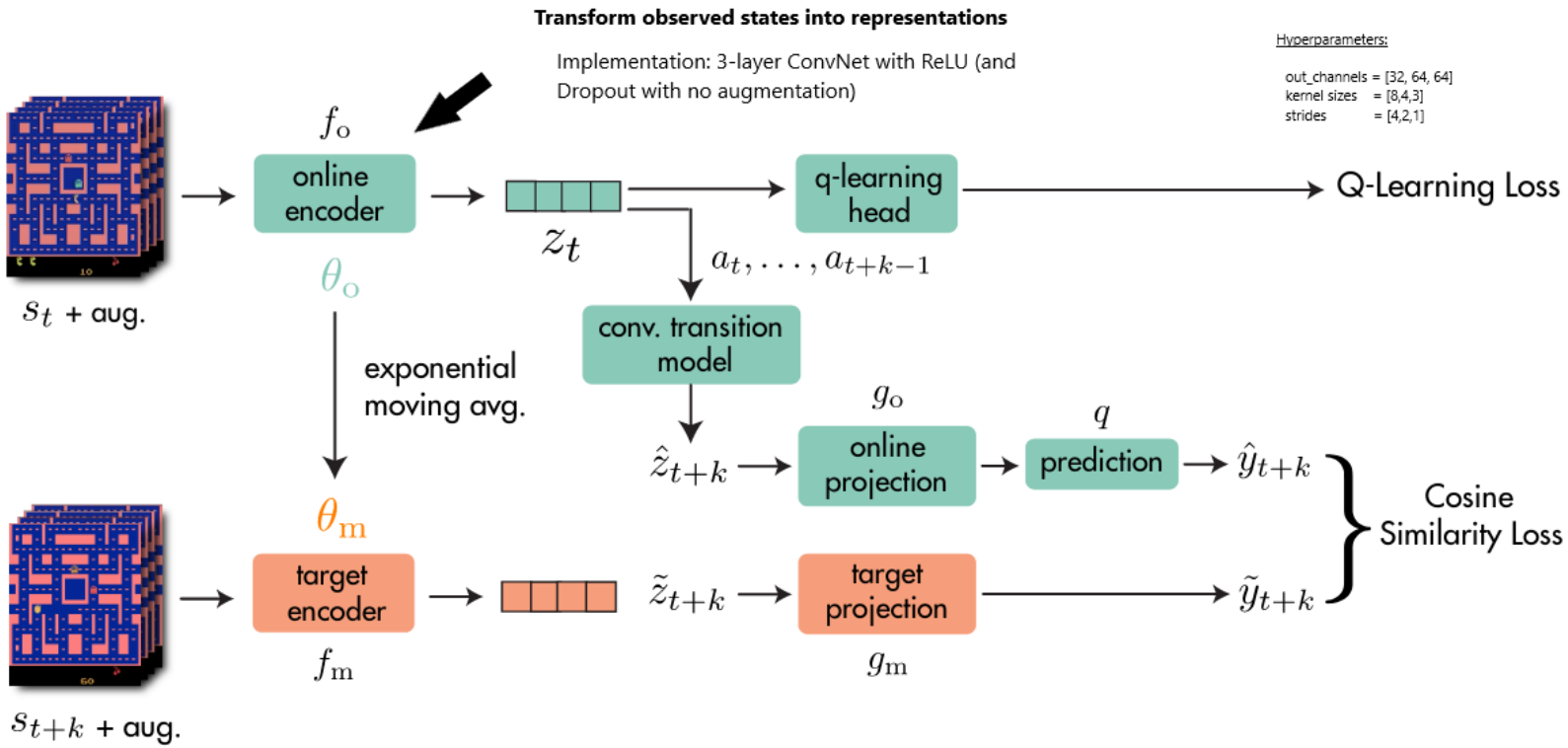
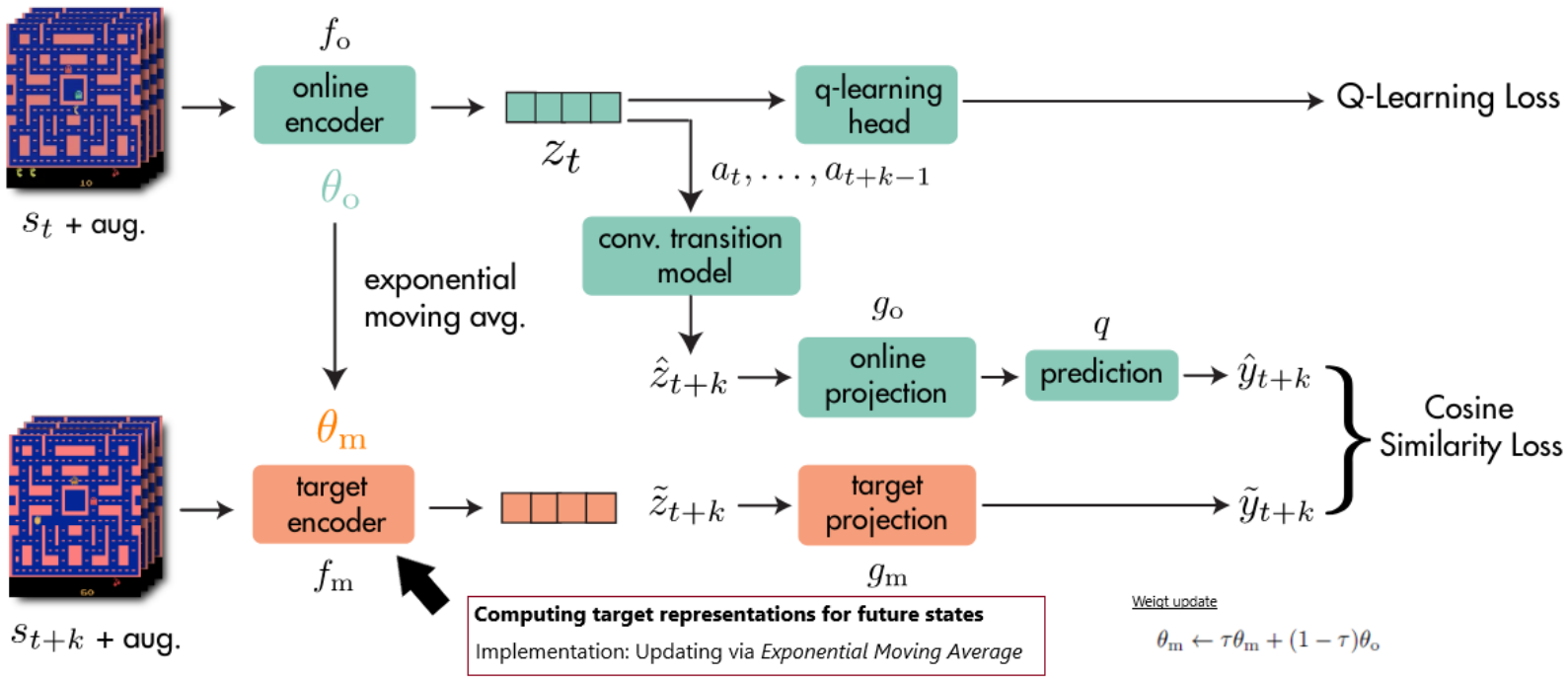


Figure: Illustration of the SPR-method

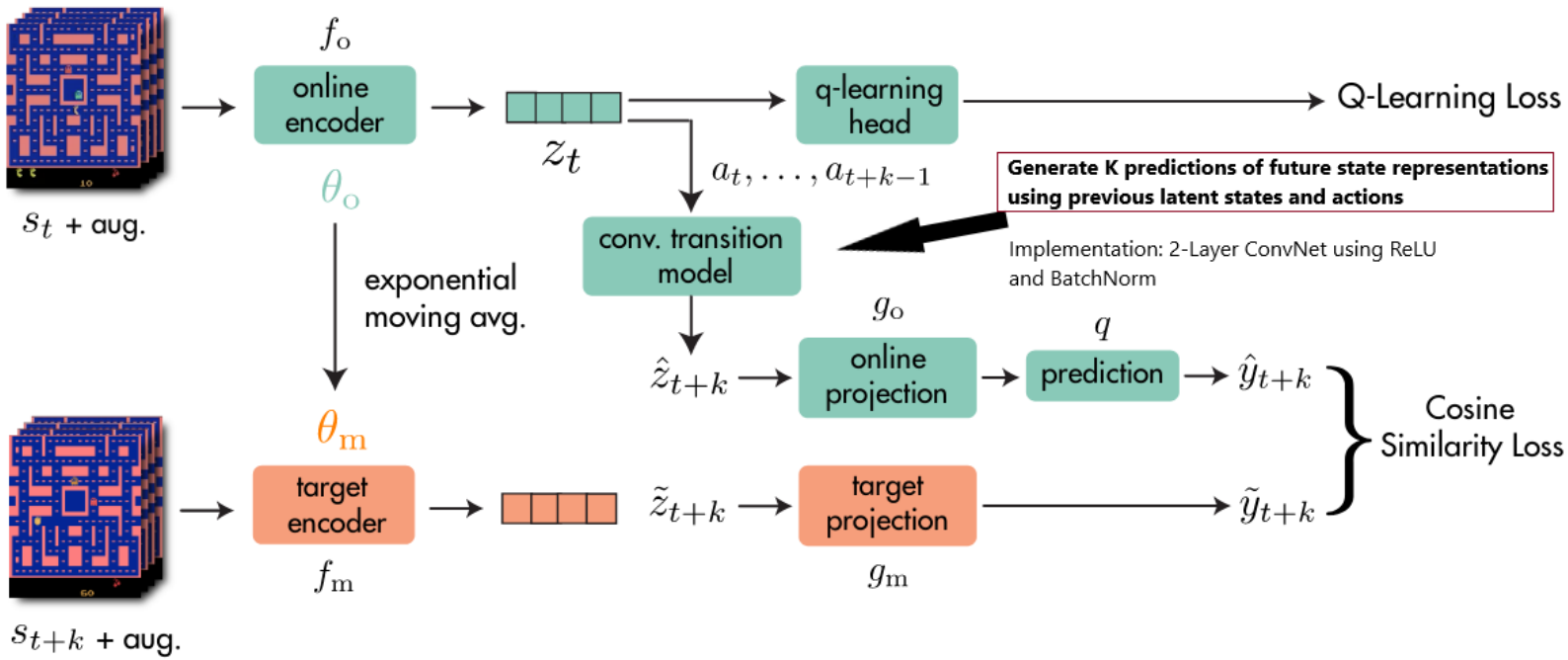
Architecture: Online-Encoder



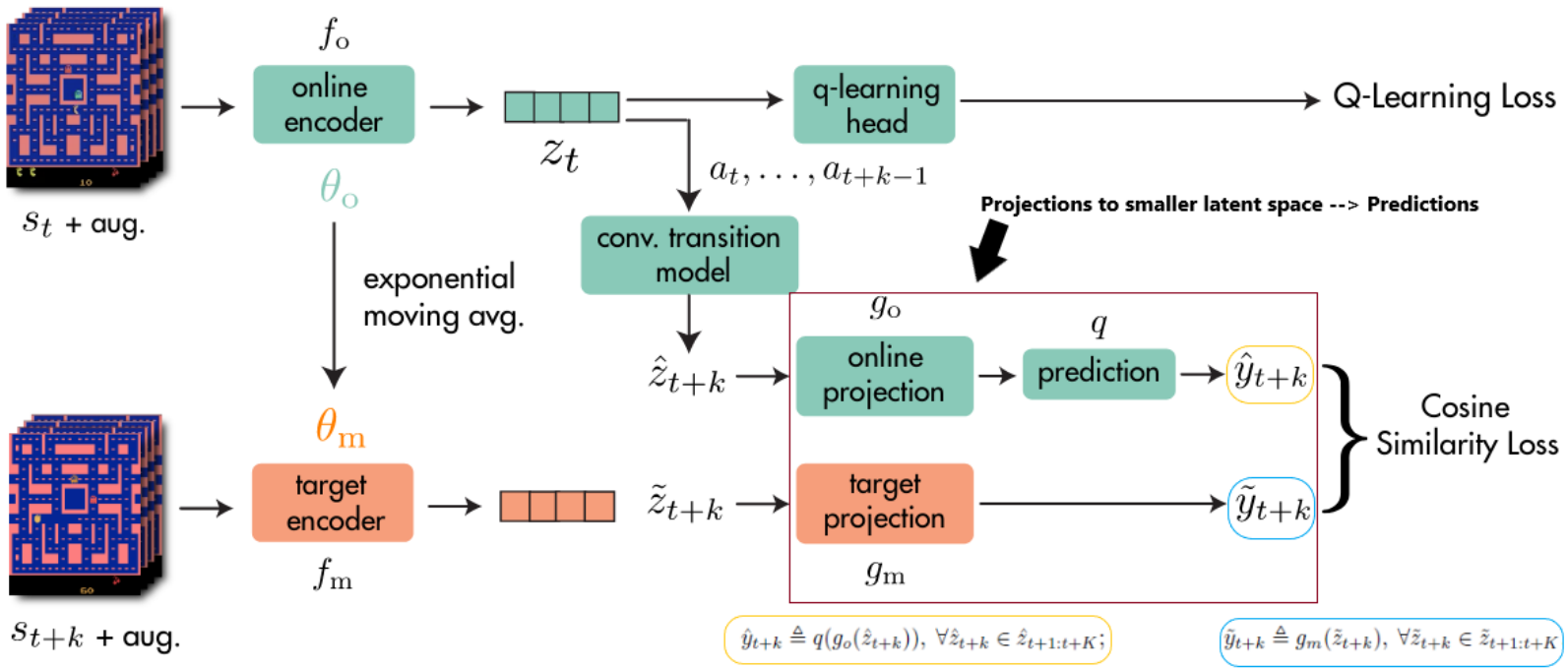
Architecture: Target-Encoder



Architecture: Transition model



Architecture: Projection heads



Architecture: Loss

Two different losses

- L_θ^{RL} : Q-Learning loss
- L_{SPR}^θ : SPR (cosine similarity) loss affecting f_o, g_o, q and h

Total loss: $L_\theta = L_\theta^{RL} + \lambda L_{SPR}^\theta$

Dueling Q-Network [3]

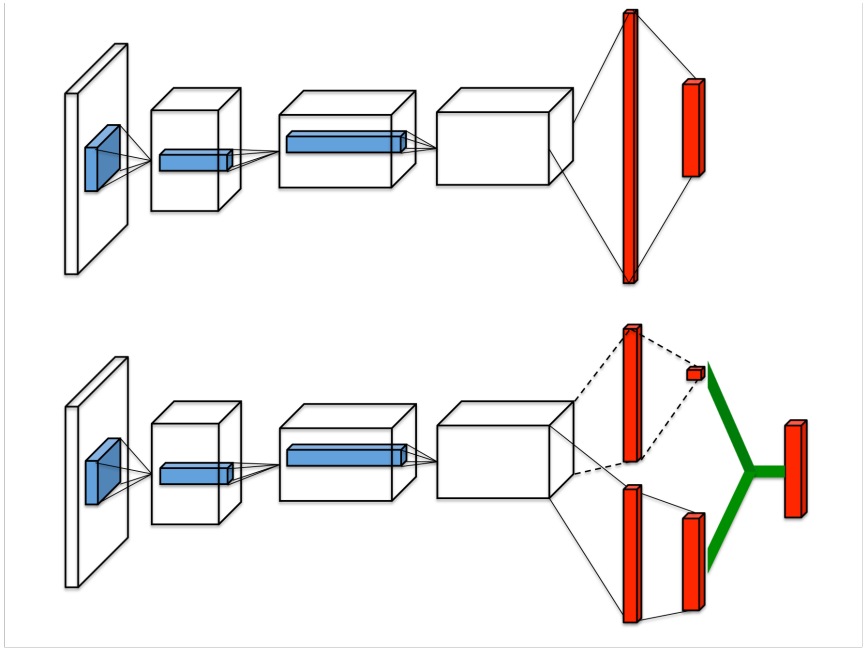


Figure 1. A popular single stream Q-network (**top**) and the dueling Q-network (**bottom**). The dueling network has two streams to separately estimate (scalar) state-value and the advantages for each action; the green output module implements equation (9) to combine them. Both networks output Q-values for each action.

OpenAi Gym

- Frameskipping
- Repeat Action Probability

Problems

- Nested original code
- Information flow hard to follow
- Much preknowledge needed ⇒ many references like
  - `(Data-Efficient-)Rainbow`[1]
  - `Double DQN`[2]
  - `Contrastive Learning`[3]
  - ...
- Compensate `rlpyt`-Library

[1] Van Hasselt et al. https://arxiv.org/abs/1710.02298
[2] Van Hasselt et al. https://arxiv.org/abs/1509.06461
[3] Chen et al. https://arxiv.org/abs/2002.05709

Resources

[allowframebreaks]

 Matteo Hessel et al. "Rainbow: Combining Improvements in Deep Reinforcement Learning". In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*. Ed. by Sheila A. McIlraith and Kilian Q. Weinberger. AAAI Press, 2018, pp. 3215–3222. URL: https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17204.

 Max Schwarzer et al. "Data-Efficient Reinforcement Learning with Self-Predictive Representations". In: *9th International Conference on Learning Representations, ICLR 2021, Virtual*