Paper:1

1. Title: "Good Robot!": Efficient Reinforcement Learning for Multi-Step Visual Tasks with Sim to Real Transfer
2. Authors: Andrew Hundt, Benjamin Killeen, Nicholas Greene, Hongtao Wu, Heeyeon Kwon, Chris Paxton, and Gregory D. Hager
3. Affiliation: Andrew Hundt, Benjamin Killeen, Nicholas Greene, Hongtao Wu, Heeyeon Kwon, and Gregory D. Hager are with The Johns Hopkins University, Baltimore, MD 21218 USA; Chris Paxton is with NVIDIA, Seattle, WA, 98105 USA.
4. Keywords: reinforcement learning, sim to real transfer, multi-step visual tasks, computer vision, grasping and manipulation.
5. Urls: Paper url: https://ieeexplore.ieee.org/document/9132494 ; Github code url: https://github.com/jhulcsr/good
6. Summary:

- (1): The research background of this paper is on the difficulty of learning long-term multi-step robotic tasks in real-world settings. Such tasks require the robot to learn the immediate physical consequences of its actions as well as progress towards the overall goal.
- (2): The past methods in reinforcement learning struggle with long-horizon tasks where time can be wasted exploring dead ends and task progress may be easily reversed. The approach proposed in this paper is motivated by the observation that reinforcement learning wastes significant time exploring unproductive actions. The Schedule for Positive Task (SPOT) framework is proposed to address this by incorporating common sense constraints in the learning process.
- (3): The research methodology proposed in this paper includes developing the SPOT framework, which explores within action safety zones, learns about unsafe regions without exploring them, and prioritizes experiences that reverse earlier progress to learn with remarkable efficiency. The framework successfully completes simulated trials of multi-step visual tasks and demonstrates direct sim to real transfer.
- (4): The methods in this paper are evaluated on a variety of tasks, including block stacking, creating rows of cubes, and clearing toys arranged in adversarial patterns. The SPOT framework significantly improves trial success rates and efficiency with respect to actions per trial, while training takes just 1-20 k actions depending on the task. The performance achieved can support the goals of efficient reinforcement learning for multi-step visual tasks with sim to real transfer.

7. Methods:

- (1): The research methodology proposed in this paper includes developing the Schedule for Positive Task (SPOT) framework, which incorporates common sense constraints in the learning process to address the challenge of learning long-term multi-step robotic tasks in real-world settings.

- (2): The SPOT framework explores within action safety zones, learns about unsafe regions without exploring them, and prioritizes experiences that reverse earlier progress to learn with remarkable efficiency.
- (3): The methods in this paper are evaluated on a variety of tasks, including block stacking, creating rows of cubes, and clearing toys arranged in adversarial patterns. The SPOT framework significantly improves trial success rates and efficiency with respect to actions per trial, while training takes just 1-20 k actions depending on the task.
- (4): The approach also includes reward shaping, which optimizes a reward to train policies efficiently, and dynamic action spaces, which assume the existence of an oracle and introduce a new target value function called SPOT-Q Learning to reduce unproductive attempts and accelerate training.
- (5): Algorithm 1 describes how the SPOT-Q Learning approach is continuously trained with Prioritized Experienced Replay (PER) as the current policy is rolled out.
- (6): The method was evaluated through a series of simulation experiments to understand the contribution of each element of the approach to their overall performance, achieving up to 100% trial success on the simulated stacking and row tasks, with models successfully transferring to the real world.

8. Conclusion:

- (1): The research presented in this article proposes an efficient reinforcement learning framework for multi-step visual tasks with sim to real transfer. The proposed Schedule for Positive Task (SPOT) framework explores within action safety zones, learns about unsafe regions without exploring them, and prioritizes experiences that reverse earlier progress to learn with remarkable efficiency.
- (2): Innovation point: The SPOT framework is a novel approach that addresses the challenge of learning long-term multi-step robotic tasks in real-world settings by incorporating common sense constraints in the learning process. Additionally, the article proposes reward shaping and dynamic action spaces to optimize policies efficiently and accelerate training.

Performance: The proposed framework significantly improves trial success rates and efficiency with respect to actions per trial, achieving up to 100% trial success on simulated block stacking and row-making tasks with successful sim to real transfer.

Workload: The training workload takes just 1-20 k actions depending on the specific task, demonstrating remarkable efficiency for reinforcement learning in multi-step visual tasks. However, SPOT's main limitation is that intermediate rewards need to be present, and future research should explore ways to overcome this limitation while also applying the approach to more complex tasks.