COMP579 Research Proposal

Anson Ng, Yinhan Lu, Yuze Liu

March 2025

1 Introduction & Motivation

While Atari games serve as standard benchmarks in reinforcement learning research, the neural network architectures used to process these environments still impose significant computational demands.
Double A3C (Zhong et al., 2023) enhances the A3C algorithm by incorporating dual value networks to reduce correlation during updates, but this additional complexity further increases parameter count and training costs. As deep RL models continue to grow in sophistication, their deployability becomes constrained by resource limitations.
Recent advances in Dynamic Sparse Training with Neural Pathways (Arnob et al., 2024), and Policy Pruning and Shrinking (PoPS) (Liu et al., 2021) offer promising avenues to maintain performance while drastically reducing resource requirements. This research addresses the critical need for more efficient reinforcement learning algorithms in resource-constrained environments.

1.1 Research Question & Approach

We propose exploring how parameter-efficient training techniques affect Double A3C performance on Atari games. Specifically, we ask:

1. Can Dynamic Sparse Training with discovered Neural Pathways (Arnob et al., 2024) maintain Double A3C performance while reducing parameter count?

2. How does the PoPS approach impact convergence speed and stability compared to dense networks?

3. What are the optimal trade-offs between model compression rates and performance across different game difficulty levels?

1.2 Proposed Approach

We propose to implement Double A3C proposed in Zhong et al. (2023) with two parameter-efficient variations: (1) Neural Pathway Discovery that progressively identifies and reinforces critical pathways during training as demonstrated by

Arnob et al. (2024), and (2) PoPS-based pruning that incrementally shrinks the policy network. Our approach combines these techniques in a novel way for reinforcement learning. We will implement this using PyTorch and OpenAI Gym framework, leveraging GPUs available through Google Cloud Computing resources or GPUs provided by the school laboratory.

1.3 Experimental Design

We will evaluate our approach on three Atari games (Breakout, Pong, and Ice Hockey) across multiple difficulty levels, comparing against vanilla A3C and original Double A3C. Performance metrics include average game score, training time, parameter count, and GPU/memory usage. We will conduct a systematic analysis of performance across different compression rates (25%, 50%, 75%) and integrate visualization tools including saliency maps to understand feature importance. This comprehensive evaluation will be completed within 10 weeks using our allocated Google Cloud GPU instances or resources given by school.

1.4 Expected Contribution

Our research aims to contribute parameter-efficient implementations of Double A3C that maintain performance while reducing computational requirements by up to 75%. We will provide an open-source code framework that enables direct comparison between different efficiency techniques, facilitating future research in this area. Additionally, we will deliver a comprehensive report with detailed instructions for reproducing both the original Double A3C results and our enhanced implementations. If successful, this work will provide practical techniques for deploying deep RL in resource-constrained environments while quantifying how different compression techniques affect performance across varying difficulty levels, extending understanding of the relationship between model capacity and task complexity in deep reinforcement learning.

References

Samin Yeasar Arnob, Riyasat Ohib, Sergey M. Plis, Amy Zhang, Alessandro Sordoni, and Doina Precup. Efficient reinforcement learning by discovering neural pathways. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL `https://openreview.net/forum?id=WEoOreP0n5`.

Xin Liu, Hsin-Pai Cheng, Mingyi Hong, and Meisam Razaviyayn. Learning to prune deep neural networks via layer-wise optimal brain surgeon, 2021. URL `https://arxiv.org/abs/1911.04252`.

Yangxin Zhong, Jiajie He, and Lingjie Kong. Double a3c: Deep reinforcement learning on openai gym games, 2023. URL `https://arxiv.org/abs/2303.02271`.