ARIA

APPLIED RESEARCH IN ACTION

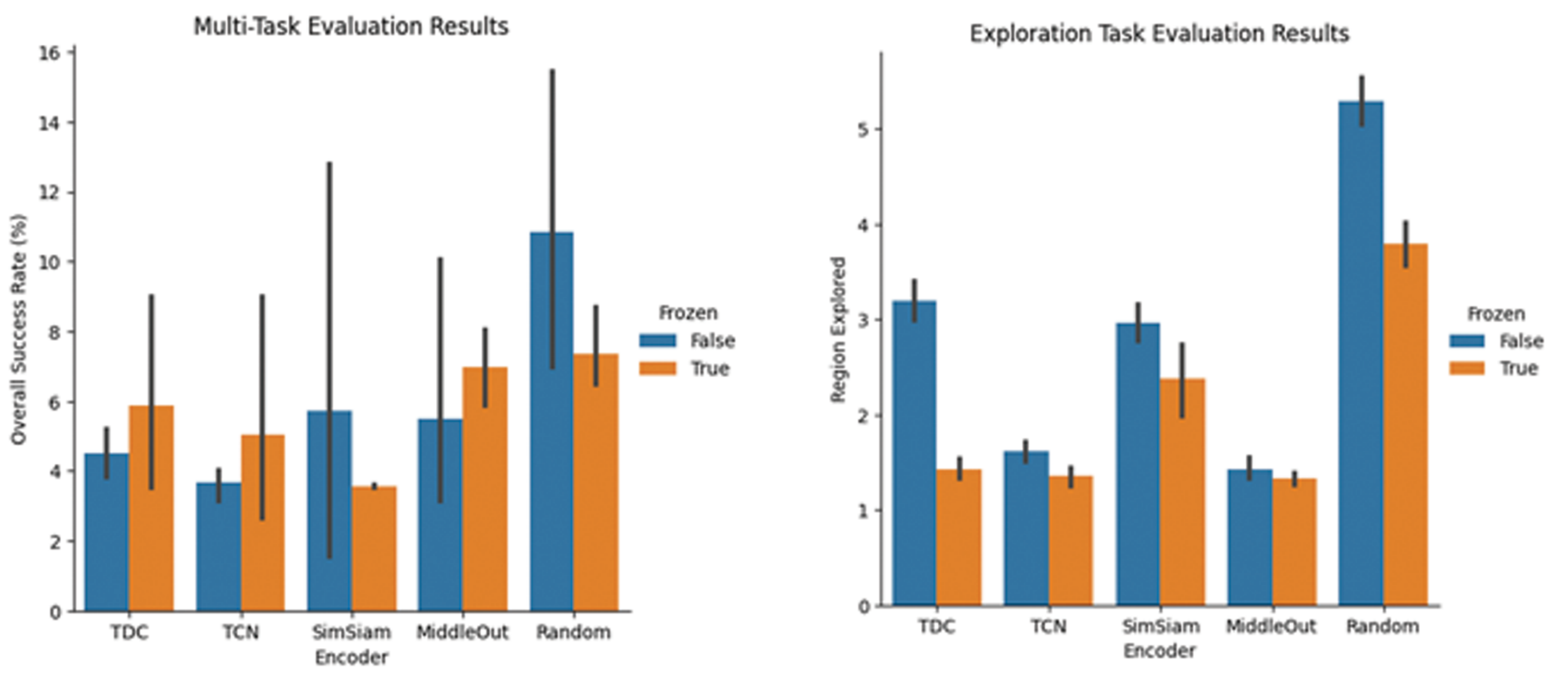
3D Image Representation for Gameplay Automation

Self-Supervised image representation encoding methods do not generalize to novel exploration and other Reinforcement Learning tasks.

Yi Cheng Zhu

Steve Engels
ACADEMIC SUPERVISOR

Yonas Bedasso
INDUSTRY SUPERVISOR



Performance of trained agents. Randomly initialized image encoder is label as "Random" in graphs. (Left) Percentage of tasks succeeded on multi-task environment on 1024 evaluation worlds after 4,000,000 training steps. (Right) Number of unique regions reached by agent in exploration task. Evaluated on 128 worlds after 500,000 training steps.

PROJECT SUMMARY

In this project, we explore the use of transfer learning for training Reinforcement Learning agents on raw pixel-based visual inputs. Existing scripting-based methods for gameplay automation has been very successful in automating linear, story driven games but fail at ensuring coverage of large open-world titles. Previous works have found that Reinforcement Learning methods significantly outperform random exploration but is difficult to train and suffers from catastrophic forgetting. We hypothesize that the image encoder is the most difficult component to train in a Reinforcement Learning scheme and it would benefit significantly from a pre-trained image encoder.

We focus on transferring image encoders trained to build meaningful embeddings of the complex 3D environment in modern AAA games through various self-supervised tasks such as Time Contrastive Networks (TCN) [1], Temporal Distance Classification (TDC) [2], Simple Siamese (SimSiam) [3], and a modified version of the Middle Out [4] methods. We evaluate the performance of our models on multi-task reinforcement learning and on a specific exploration task.

We found that all four pre-training methods resulted in significant performance degradation on both the multi-task and the exploration task. In the multi-task case, when the image encoder is frozen, the Middle Out method and the TDC methods showed similar performance to the randomly initialized image encoder. When the encoder is allowed to change, all four methods significantly underperformed compared to the randomly initialized. We conclude that the features learned by such pre-training methods are not applicable to the tasks, possibly due the absence of samples of the target task in the training dataset.

REFERENCES

1. Sermanet, Pierre, et al. "Time-contrastive networks: Self-supervised learning from video." 2018 IEEE international conference on robotics and automation (ICRA). IEEE, 2018.
2. Aytar, Yusuf, et al. "Playing hard exploration games by watching youtube." Advances in neural information processing systems 31 (2018).
3. Chen, Xinlei, and Kaiming He. "Exploring simple siamese representation learning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.
4. Li, Junnan, et al. "Self-supervised representation learning using 360 data." Proceedings of the 27th ACM International Conference on Multimedia. 2019.

AMD

Computer Science
UNIVERSITY OF TORONTO

Master of Science in Applied Computing