Self-Supervised Attention-Aware Reinforcement Learning

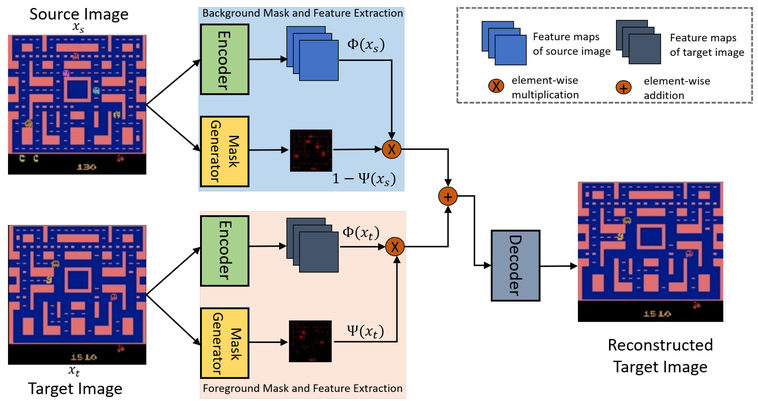
Haiping Wu Khimya Khetarpal Doina Precup

McGill University Mila

Introduction

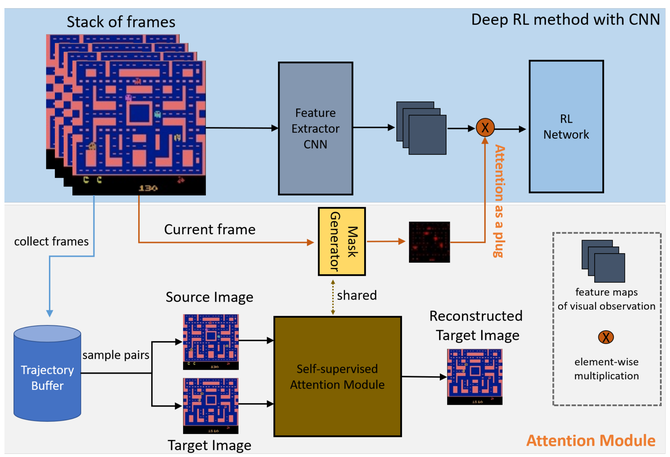
- Recent works try to obtain object keypoints in an unsupervised manner. However, current unsupervised keypoints detection methods including the Transporter are limited in that they do not deal with variable number of objects, scale, and classes of objects.
- Furthermore, the use of object-oriented representation for deep RL has not been highly explored.
- We propose to learn attention masks to understand the scene in a self-supervised manner and show it is easy to plug in the attention module for existing deep RL methods for policy learning.

Self-Supervised Attention Module



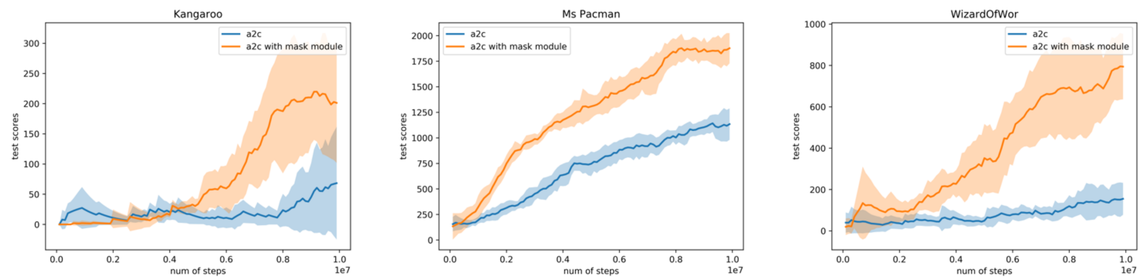
The core idea is to employ a self-supervised loss through an auto-encoder architecture with a bottleneck. The module tries to reconstruct the target image x_t by using minimal information (features of foreground regions) from the target image x_t, and other needed information from source image x_s.

Attention-aware Reinforcement Learning

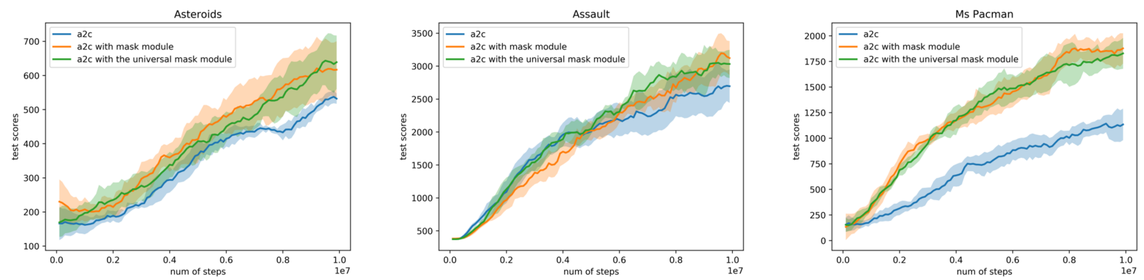


Experiments & Results

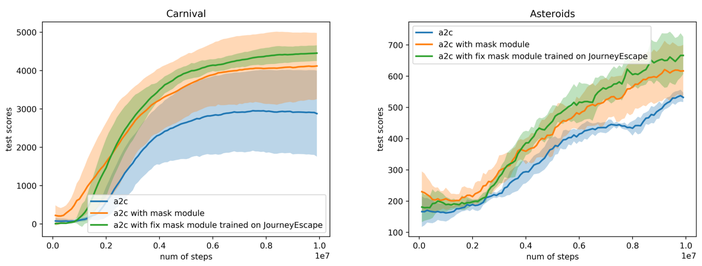
Single-task Learning. Average (over 5 random seeds) test scores during learning of A2C with/without the our self-supervised attention mask.



Multi-task Learning. Comparison between the baseline method A2C, A2C with the self-supervised attention module, A2C with the the universal attention module jointly trained on three games.

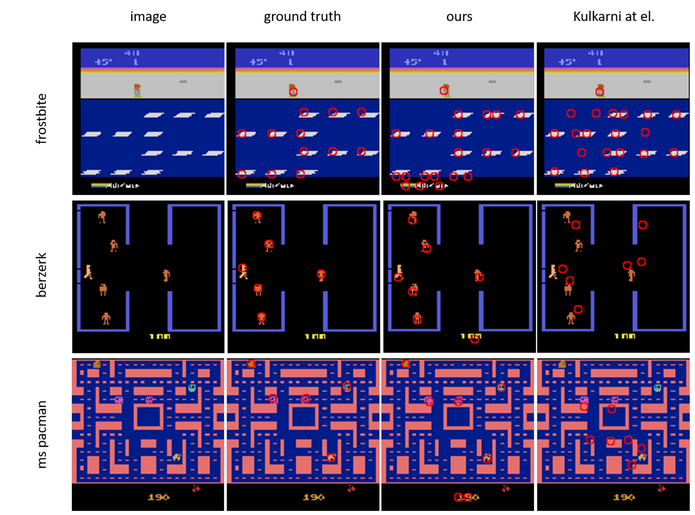


Transfer Learning.



Comparison between the baseline method A2C and A2C with the fix attention module trained on JourneyEscape, showing that the attention module has the ability to transfer across games.

Qualitative results of extracted object keypoints from learned mask.



Use Non-Maximum Suppression (NMS) to get object keypoints from learned attention masks.