Unsupervised Video Object Segmentation for Deep Reinforcement Learning

Vik Goel, Jameson Weng, Pascal Poupart
University of Waterloo, Waterloo AI Institute, Vector Institute

Introduction
- Current deep RL techniques require large amounts of data to find a good policy
- Once found, the policy remains a black box to practitioners
- Practitioners cannot verify that the policy is making decisions based on reasonable information
- MOREL (Motion-Oriented REinforcement Learning) automatically detects moving objects and uses the relevant information for action selection

Key Ideas
- Within the first few seconds of playing a game, humans are able to pick out all the important objects
- One of the most important reasons for this is that humans have strong priors
- Motion is a strong indicator for identifying important objects in games
- MOREL splits the training procedure into two phases
  - The first phase learns object segmentation in an unsupervised manner
  - The second phase uses the learned representation to optimize reward
- Practitioners can look at the segmented objects to diagnose model strengths and weaknesses

Learning to Segment Moving Objects
- We gather a dataset using a uniform random policy
- Train a network without supervision to capture a structured representation of motion between frames
- Network predicts object masks, object motion, and camera motion to warp one frame into the next

Transfer to RL Agent
- We add an extra path to track static objects
- RL agent is jointly optimized with object segmentation
  - This allows the agent to continue learning to segment objects as it encounters novel states
- Our method can be composed with any deep RL method, such as A2C and PPO
  - Practitioners can look at the segmented objects to diagnose model strengths and weaknesses

Experiments
- We gather a dataset using a uniform random policy
- Train a network without supervision to capture a structured representation of motion between frames
- Network predicts object masks, object motion, and camera motion to warp one frame into the next

Visualizations
- We visualize our model's object segmentations to allow greater interpretability
- For example, our method has trouble on Beam Rider, where the object masks focus on capturing animations unimportant to the game

Figure 1: Evaluation of sample complexity on all 59 Atari games after composing RL algorithms with MOREL

Figure 2: Ablation study of MOREL vs. vanilla A2C

Figure 3: Ablation study of modifications to our object segmentation network

Figure 4: Unsupervised video object segmentation results. The first and second frames are the inputs to the network. The masks are overlaid in green, where intensity indicates model confidence.