Paper review for "Zhan, Zhao et al.: A Framework for Efficient Robotic Manipulation"

EECS 598 Paper Review – Week 12 - Changyuan Qiu

Real-robot reinforcement learning is challenging for several reasons, and data efficiency is chief among them. While deep reinforcement learning algorithms have shown success learning policies from visual observations, they still require an impractical number of real-world data samples to learn effective policies.

Here the author proposed a Framework for Efficient Robotic Manipulation (FERM) building upon recent significant advances in unsupervised representation learning and data augmentation to tackle this sample inefficiency problem of reinforcement learning. The key idea is: first collect 10 human demonstrations and store in a replay buffer; then pre-train a CNN encoder on the demonstration data with a contrastive loss; finally use an offline data-augmented reinforcement learning algorithm to train an reinforcement learning agent with both the encoder (augmented images on data collected online during training) and the replay buffer (the initial demonstrations).

While prior works typically utilized workaround strategies relying on transferring policies like training in simulation and transferring the learned policy to the real robot (Sim2Real) or parallelizing training with robot farms (QT-Opt), this work focused on making pixel-based RL data-efficient itself. The key achievement of the work is FERM's amazing data efficiency – experiments showed that using only 10 demonstrations, FERM can learn optimal policies on 6 diverse manipulation tasks such as reaching, pushing, moving, pulling a large object, flipping a switch, drawer opening in 15-50 minutes of training. The final learned policies of FERM achieved an average success rate of 96.7% across 30 episodes of evaluation on each of the 6 tasks. Besides, FERM is also quite general (FERM use the same hyperparameters to solve all 6 tasks) and lightweight (FERM only requires a robot, two cameras and a handful of demonstrations, and the contrastive pre-training only takes ~ 40 seconds on a single GPU) compared with prior methods like Sim2Real, motion capture and engineering dense rewards.

Regarding limitations of this paper, it is worth noting that authors only evaluate FERM experiments on 6 basic robot manipulation tasks like reach, move and pull, since these basic manipulation tasks could be explicitly programmed and achieved competitive performance. And I highly suspected whether FERM could generalize to more complex tasks like motion planning, and I believed there are at least 2 bottlenecks for FERM to work on complex tasks: (1) Demonstrations are harder to collect and more demonstrations might be needed due to the complex problem settings (2) Contrastive pre-training could fail on complex problem settings. But still, FERM seems promising as a general backbone framework for real-robot reinforcement learning and presents exciting avenues for applying RL to real robots in a quick and efficient manner.