Adversarial Feature Training for Generalizable Robotic Visuomotor Control

Xi Chen[1], Ali Ghadirzadeh[1,2], Mårten Björkman[1], and Patric Jensfelt[1]

[1] RPL, KTH Royal Institute of Technology, Sweden
{xi8,algh,celle,patric}@kth.se
[2] Intelligent Robotics research group, Aalto University, Finland

1 Introduction

Deep reinforcement learning (RL) has become a popular framework to train visuomotor action-selection policies that directly map raw image pixels to motor actions, eliminating the need for large-scale image labeling. However, lack of generality is still a common problem when training a visuomotor action-selection policy. As an example, a policy trained to pour the liquid into a mug would fail if the mug is replaced by a glass that has not been seen during the training. In this work, we introduce an end-to-end training method based on adversarial training to extract visual features that generalize well to other instances of task objects, assuming that weakly labeled still images of such objects are provided. Here, weak labels mean binary values specifying whether a given image contains a task object or not. In our experiments, we show that we can train a policy for only a template object, an object with the simplest possible geometry and texture, and then selectively transfer the learned policy to other task objects surrounded by task-irrelevant visual clutter. We evaluate our method on two real robotic tasks, picking and pouring using an ABB YuMi robot. Our empirical analysis demonstrates that our method outperforms prior work with a good margin. Videos of our experiments can be found at https://youtu.be/BRk7ctg7ftI.

2 Method

The problem we address is how to train a deep feed-forward policy network which maps an input image to a sequence of motor actions. Similar to our earlier work [2], we split a deep policy network into three sub-networks, which are (1) the perception, (2) the low-dimensional policy, and (3) the motor trajectory networks. The perception network is responsible for finding a suitable feature representation of the input image, the low-dimensional policy network maps visual features to the latent space of the generative model, and the motor trajectory network maps a latent variable input into a trajectory of motor actions.

To avoid the expensive process of training deep visuomotor policies based on real robot data and at the same time allow policies to generalize to novel variations of the task, we first train the policy using reinforcement learning given interactive data collected by playing with a template object only, as described in our earlier works [2, 1, 3], and then transfer the gained manipulation knowledge to the rest of the task objects by further training the perception network and the policy network, assuming that still images of the task objects are provided. The main idea is to extract a set of visual features as the output of the perception model which are suitable for the RL task while not being distinguishable whether the features comes from the template object or the task objects.

In the second phase, we construct two sets of training data: *Task dataset (source domain data $\mathcal{D}_S$)*, which includes images of the template object and the output latent encoding of the pre-trained low-dimensional policy; and *Task object dataset (target domain data $\mathcal{D}_T$)*, which includes images accompanied by weak labels that specify whether an image contains a task object or not. We introduce two auxiliary networks, a discriminator and a classifier, which are not part of the final model but will only be used during training. The discriminator network receives the visual features generated by the perception model and decides whether the input is from the template object or the task objects. Similarly, the classifier network receives the visual features as input and outputs a classification of whether there is a task object or not. We train the perception, the policy, as well as the auxiliary networks, the discriminator and the classifier, jointly by optimizing the three loss functions introduced in the following:

$$\mathcal{L} = \mathcal{L}_{task}(\mathcal{D}_S) + \mathcal{L}_c(\mathcal{D}_T) + \mathcal{L}_D(\mathcal{D}_S \cup \mathcal{D}_T), \quad (1)$$

where $\mathcal{L}_{task}$ is a mean squared error loss applied to the source domain data to learn the mapping from the template object images to the corresponding latent encoding; $\mathcal{L}_c$ is a binary cross-entropy classification loss, which is used to speeds up the training by implicitly guiding the network where to attend in the input image; and $\mathcal{L}_D$ is a discriminator loss that forces the extracted features to be distributed similarly, whether the features come from the source domain data or the target domain data.

3 Results

We evaluate our method on two real-world robotic manipulation tasks: a pouring task and a picking task, using an ABB Yumi robot. In the pouring task, the robot needs to pour the content of a small cup into the desired cup, and in the picking task, the robot has to pick up a cuboid object. The cups and the cuboid objects are placed on a table and other task-irrelevant objects that are regarded as visual clutter. For each task, we first train a policy using a template and then transfer the learned policy to other task objects, as explained in the method section. We evaluate the policy on 3 task cups for the pouring task and 20 cuboid objects, including 5 novel objects, which is unseen during policy training, for the picking task. We perform 10 tests for each task object, and compare the results to two baseline methods: ADDA [5] and GPLAC [4]. Our method achieves a success rate of 100% on the pouring task, 86% on the picking task for objects with training data, and 82% for unseen objects. While the best result of the baseline methods achieves a success rate of 53% on the pouring task, 67%, and 56% on the picking task.

References

1. Ghadirzadeh, A.: Sensorimotor Robot Policy Training using Reinforcement Learning. Ph.D. thesis, KTH Royal Institute of Technology (2018)
2. Ghadirzadeh, A., Maki, A., Kragic, D., Björkman, M.: Deep predictive policy training using reinforcement learning. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 2351–2358. IEEE (2017)
3. Hämäläinen, A., Arndt, K., Ghadirzadeh, A., Kyrki, V.: Affordance learning for end-to-end visuomotor robot control. arXiv preprint arXiv:1903.04053 (2019)
4. Singh, A., Yang, L., Levine, S.: Gplac: Generalizing vision-based robotic skills using weakly labeled images. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 5851–5860 (2017)
5. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7167–7176 (2017)