

Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey

<https://arxiv.org/abs/2009.13303>

0. Introduction

- 강화학습(RL)은 로보틱스 분야에서 큰 성공을 거두었지만, 실제 환경에서의 데이터 수집은 비용과 시간 면에서 제한적입니다.
- 이를 해결하기 위해 시뮬레이션 환경에서 학습을 진행하지만, 시뮬레이션과 실제 환경 간의 차이로 인해 정책의 성능 저하가 발생합니다.
- 이러한 Sim-to-Real Gap을 극복하기 위한 다양한 연구가 진행되고 있으며, 본 논문은 이를 위한 주요 방법론들을 종합적으로 다룹니다.

1. Overview

- Sim-to-Real Transfer는 시뮬레이션에서 학습한 정책을 실제 환경에 적용하는 과정입니다.
- 이 과정에서 발생하는 Sim-to-Real Gap은 시뮬레이션과 실제 환경 간의 물리적, 시각적 차이로 인해 발생합니다.
- 주요 접근법으로는 Domain Randomization, Domain Adaptation, Imitation Learning, Meta-Learning, Knowledge Distillation 등이 있습니다.

2. Challenges

- Sim-to-Real Gap: 시뮬레이션과 실제 환경 간의 물리적, 시각적 차이로 인해 정책의 성능이 저하됩니다.
- 샘플 효율성: 실제 환경에서의 데이터 수집은 비용과 시간이 많이 소요되므로, 샘플 효율성이 중요한 문제로 대두됩니다.

- 안전성 문제: 실제 로봇을 사용한 학습은 실패 시 물리적 손상이 발생할 수 있어 안전성이 중요한 고려사항입니다.

3. Taxonomy of Methods

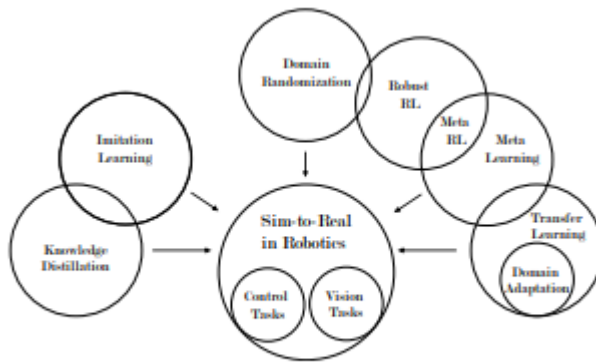


Fig. 2: Illustration of the different methods related to sim-to-real transfer in deep reinforcement learning and their relationships.

1. Domain Randomization

- 시뮬레이션 환경의 다양한 파라미터(예: 조명, 질감, 마찰 계수 등)를 무작위로 변화시켜 모델의 일반화 능력을 향상시킵니다.

2. Domain Adaptation

- 시뮬레이션과 실제 환경 간의 분포 차이를 줄이기 위해 Generative Adversarial Networks(GANs) 등을 활용하여 시뮬레이션 데이터를 실제 환경에 맞게 변환합니다.

3. Imitation Learning

- 전문가의 행동을 모방하여 학습하는 방식으로, 실제 환경에서의 데이터 수집을 최소화할 수 있습니다.

4. Meta-Learning

- 다양한 환경에서 빠르게 적응할 수 있는 모델을 학습하여, 새로운 환경에서도 높은 성능을 유지하도록 합니다.

5. Knowledge Distillation

- 복잡한 모델의 지식을 간단한 모델로 전달하여, 실제 환경에서도 효율적으로 동작할 수 있도록 합니다.

4. Applications / Case Studies

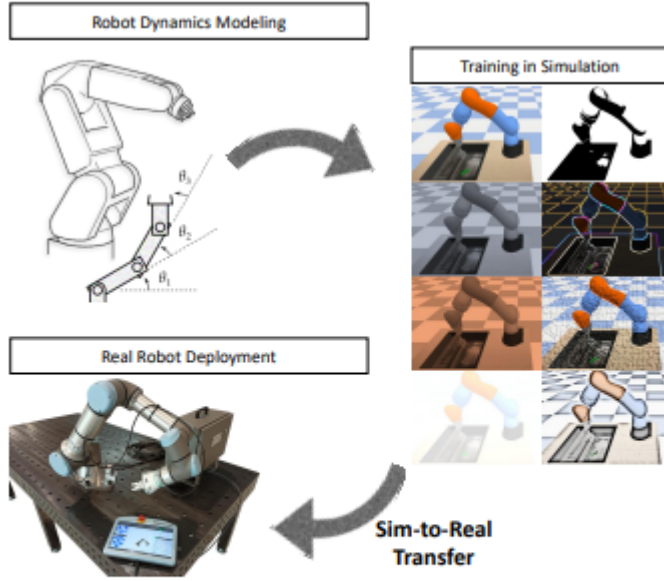


Fig. 1: Conceptual view of a simulation-to-reality transfer process. One of the most common methods is domain randomization, through which different parameters of the simulator (e.g, colors, textures, dynamics) are randomized to produce more robust policies.

- 로봇 조작: 시뮬레이션에서 학습한 조작 정책을 실제 로봇에 적용하여 다양한 물체를 조작하는 데 성공하였습니다.
- 로봇 팔 학습: 시뮬레이션에서 학습한 로봇 팔의 움직임을 실제 환경에 적용하여 높은 정확도로 목표를 달성하였습니다.
- 자율 주행: 시뮬레이션 환경에서 학습한 주행 정책을 실제 도로 환경에 적용하여 안전하게 주행하는 데 성공하였습니다.

5. Results & Comparisons

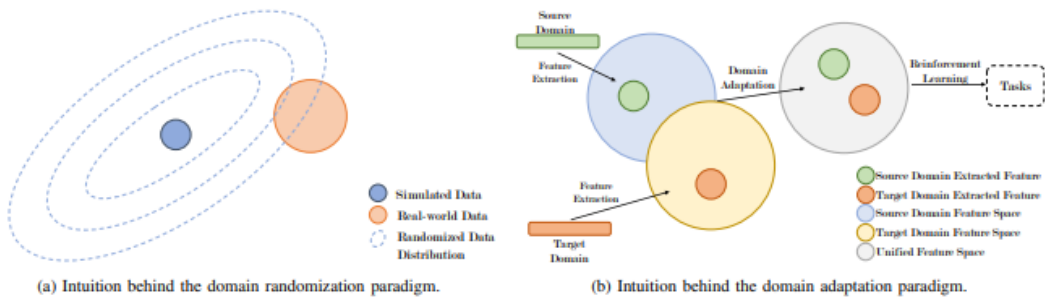


Fig. 3: Illustration of two of the most widely used methods for sim-to-real transfer in DRL. Domain randomization and domain adaptation are often applied as separate techniques, but they can also be applied together.

- Domain Randomization은 시뮬레이션과 실제 환경 간의 차이를 줄이는 데 효과적이며, 다양한 환경에서의 일반화 능력을 향상시킵니다.
- Domain Adaptation은 실제 환경에 맞게 시뮬레이션 데이터를 변환하여 성능을 향상시키지만, 추가적인 계산 자원이 필요합니다.
- Imitation Learning은 실제 환경에서의 데이터 수집을 최소화할 수 있지만, 전문가의 행동 데이터가 필요합니다.
- Meta-Learning은 다양한 환경에서 빠르게 적응할 수 있는 모델을 학습하여 새로운 환경에서도 높은 성능을 유지합니다.
- Knowledge Distillation은 복잡한 모델의 지식을 간단한 모델로 전달하여 실제 환경에서도 효율적으로 동작할 수 있도록 합니다.

6. Insight / Future Directions

- 다양한 환경에 대한 일반화 능력 향상: 다양한 환경에서의 학습을 통해 모델의 일반화 능력을 향상시킬 필요가 있습니다.
- 샘플 효율성 개선: 실제 환경에서의 데이터 수집을 최소화하고, 샘플 효율성을 개선하는 방법론의 개발이 필요합니다.
- 안전성 고려: 실제 환경에서의 학습 시 안전성을 고려한 방법론의 개발이 중요합니다.
- 다양한 방법론의 통합: 여러 방법론을 통합하여 더욱 효과적인 Sim-to-Real Transfer를 구현할 필요가 있습니다.