Alternative Distillation Methods in Sim-to-Real Transfer

Li, Kevin (School: Montgomery Bell Academy)
Cook, Brash (School: Montgomery Bell Academy)

Due to the need for unsafe or costly data collection to train a policy in the real world, sim-to-real transfer is the most common approach: learning a policy in simulation and deploying it in the real environment. However, these policies often fail to transfer reliably to the real world because of the sim-to-real gap. Domain randomization exposes the policy to a wide range of simulated environments with different environmental variables (friction, backlash, gravity, etc.) during training to improve the policy's generalization to the real world, but basic training methods often leave the robot in a local maximum. To enable a policy to adapt better to DR and thus improve generalization to the real world, model distillation is commonly employed. This project investigates how different distillation methods affect policy robustness under domain randomization. We designed and built a custom quadruped robot and evaluated multiple different training methods (PPO, Teacher-Student Distillation, P2PDRL, and CPD) in a simulated locomotion task in IsaacLab. Policies were compared via their reward performance in simulation under domain randomization and deployed onto real hardware. We found that CPD produced a significantly higher reward under DR (50% increase over standard PPO) and even managed to outperform teacher-student distillation. Reinforcement learning has become increasingly more common and more powerful for complex robot tasks. However, when deployed onto low-cost physical hardware, the improvements were negligible, which shows that while CPD increases generalization in simulation, its benefits may be limited in systems with hardware constraints.

Awards Won: