Context-Based Meta Reinforcement Learning for Robust and Adaptable Peg-in-Hole Assembly Tasks

LAMARR INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Authors: Ahmed Shokry, Walid Gomaa, Tobias Zaenker, Murad Dawood, Rohit Menon, Shady A. Maged, Mohammed I. Awad, Maren Bennewitz

Humanoid Robots Lab, University of Bonn

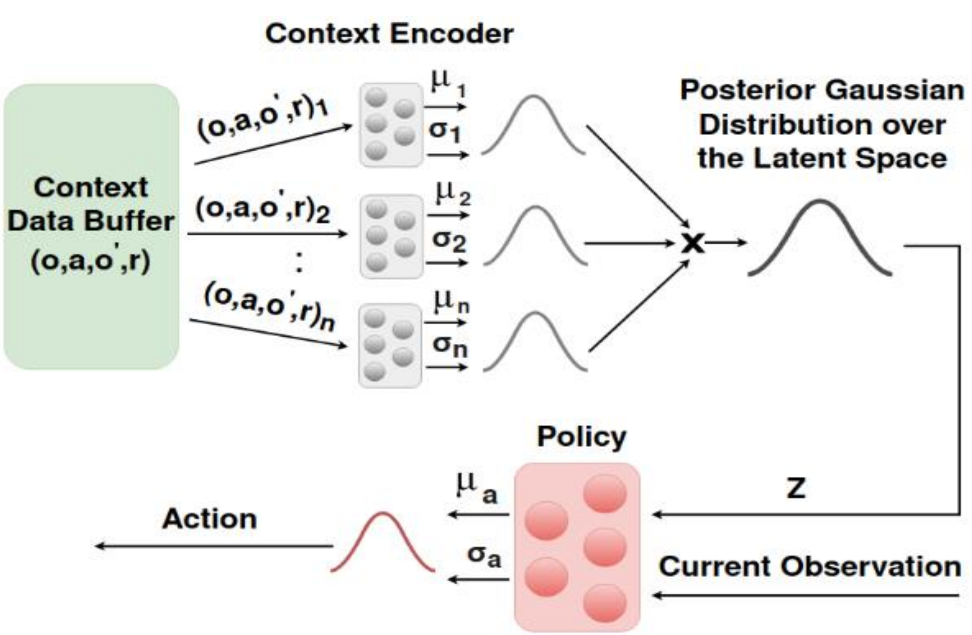
Motivation

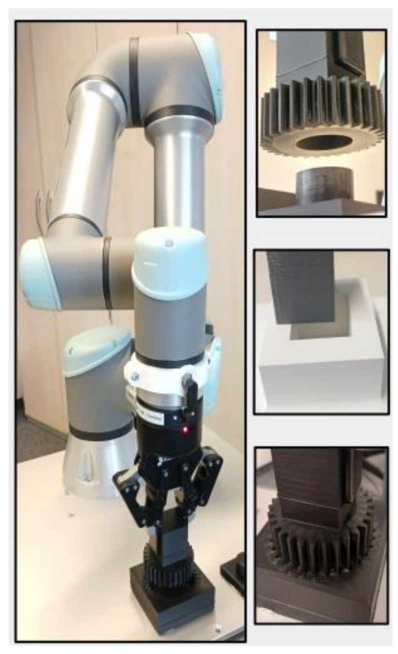
Peg-in-Hole Assembly

- **Peg-in-hole** (PiH) assembly is a **core** task for industrial and service robots
- **Uncertainty** in the hole position and orientation due to **sensor errors** represents a challenge for PiH assembly tasks in **unknown** environments

Meta Reinforcement Learning

- **Meta** reinforcement learning (RL) can infer and adapt to **unknown** task parameters and different rewards. However, it is **sample-inefficient**
- **PEARL** meta RL agent uses the **context data** to infer the unknown task parameters
- Observation **o** is the distance between the peg and the **estimated noisy** hole position
- Reward **r** is the distance between the peg and the **actual hole** position, which is **unknown** during test time





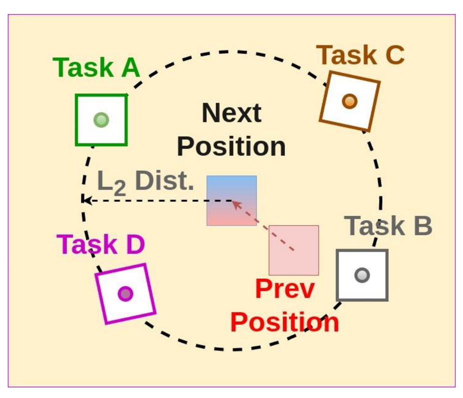
Challenges and Contributions

Challenges

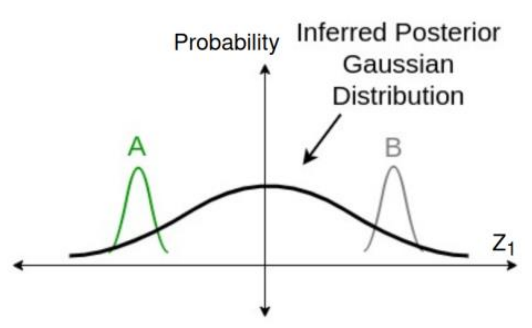
- immeasurable **reward,** due to the unknown hole position, results in a sample-inefficient adaptation during **test** time
- PEARL has a very limited **generalization** to out-of-distribution (OOD) tasks with more **uncertainty** in the hole position than that of the training tasks

Contributions

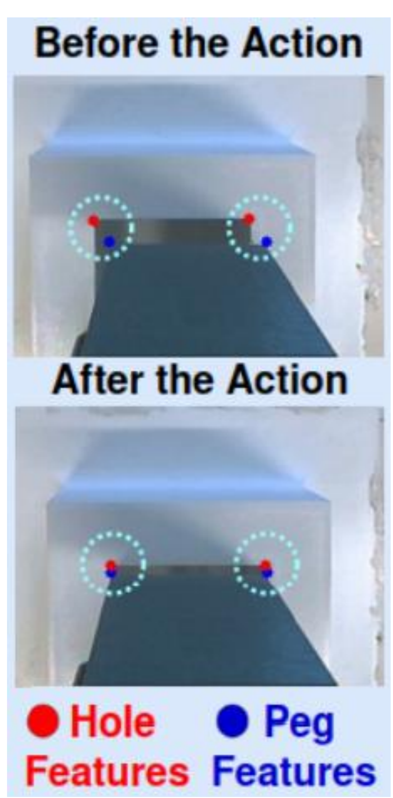
- Replace the reward in the context data with the **motion** towards the actual hole **m** due to the last action, **enhancing** the agent's ability to infer the task
- The motion **m** is calculated using the robot forward kinematics and hole features in **2D images** without requiring depth information, enhancing real-world **applicability**
- Replace the **meta-trained** context encoder with a **new** encoder that uses **force/torque** sensor readings and trained using a **limited** number of real-world data
- Propose a **fine-tunning** method that **pushes** the posterior distribution towards latent variables with **high** motion towards the actual hole **m** to safely and gradually adapt to **OOD** tasks

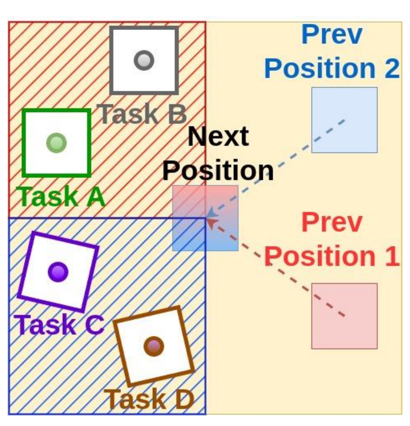


Tasks with the same reward require opposite actions

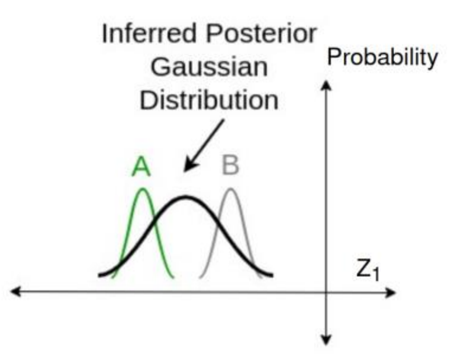


Uncertain and wide posterior Gaussian distribution





Tasks with the same motion towards the actual hole **m** require slightly different actions

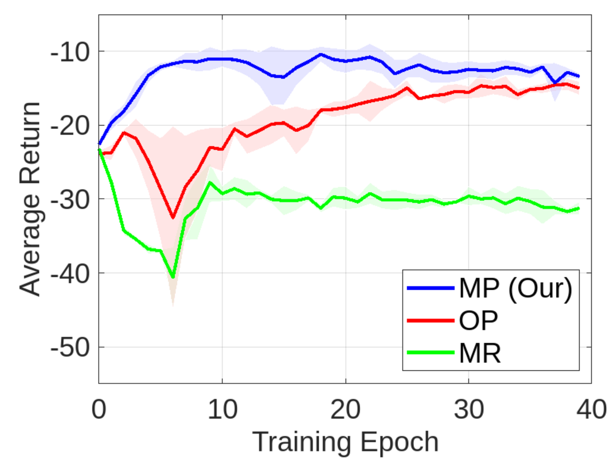


More certain posterior Gaussian distribution

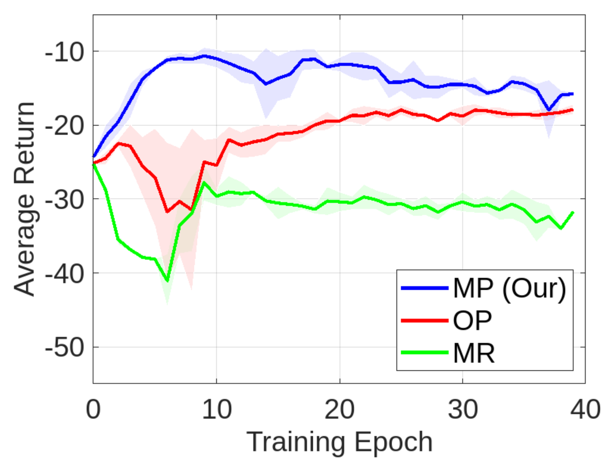
Results

Effect of Context Data Modification on the Training Efficiency

- Our modified agent achieves **higher return** in a **smaller** number of epochs



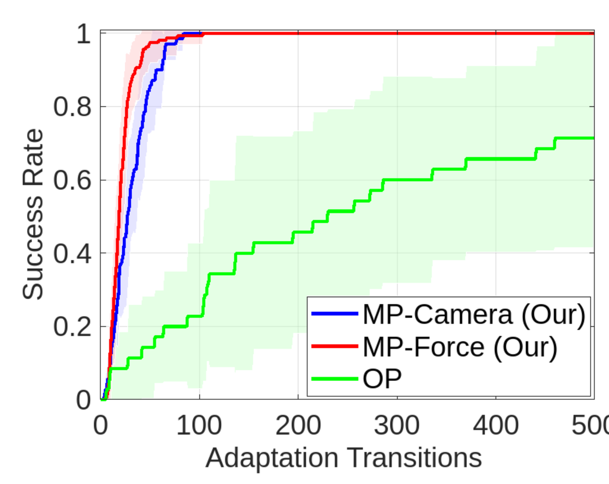
Training tasks



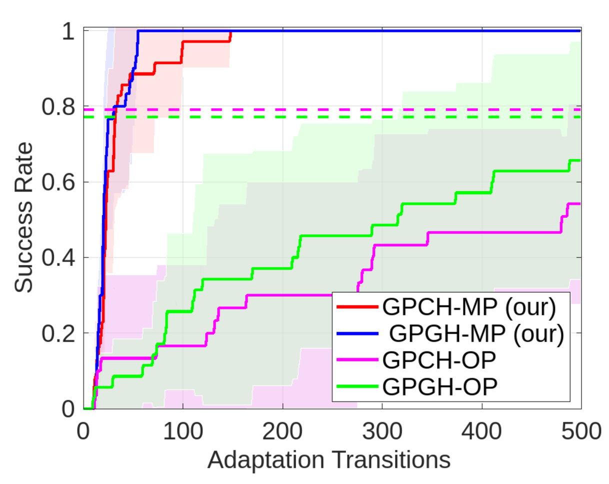
Test tasks

Real-World Adaptation Performance

- Our modified agents show **higher success** rate and more **consistent** performance



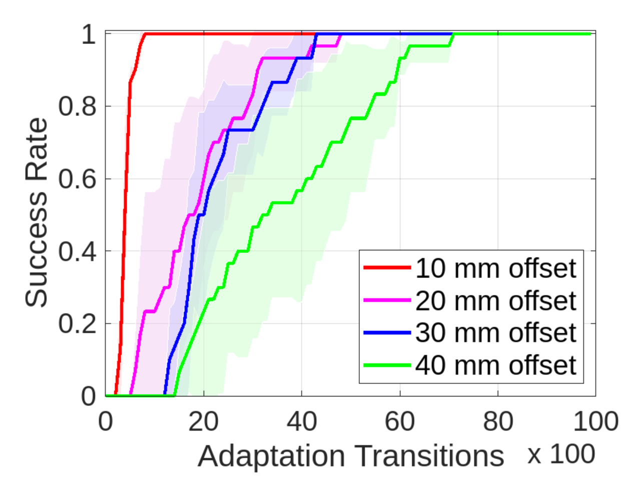
Simple shapes (Cuboid)

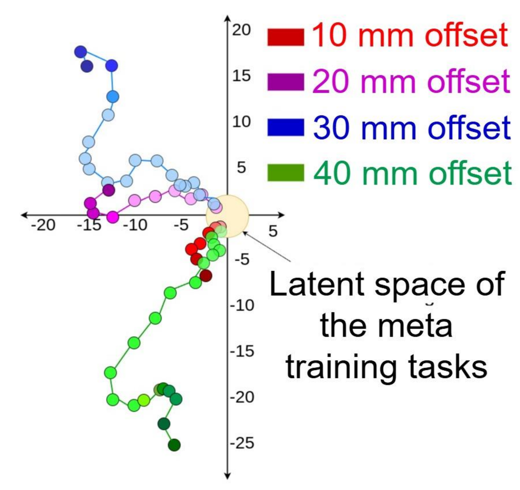


Complex shapes (Gear)

Out-of-Distribution Adaptation

- Our proposed method gradually and consistently **explores** the **latent space** and adapts to tasks with uncertainty up to **10x** more than that of the training tasks





Summary

- Our approach **enhances** the performance of the context-based **meta RL** agent in **PiH** assembly tasks
- We enhance the **sample efficiency** of the meta RL agent by **modifying** the context data used to infer the unknown task parameters
- The proposed modified context can be measured in the **real world** using an **uncalibrated** camera or a force/torque sensor enhancing the real-world **applicability**
- Context-based meta RL is known for its **limited** generalization to **OOD** tasks, that is why a safe and gradual latent space **exploration** method is proposed
- Experiments in simulation and in the real world prove the effect of the proposed methods in enhancing **training efficiency**, real-world **adaptation** performance, and out-of-distribution **generalization** capabilities of the meta RL agent

Contact

Ahmed Shokry
shokry@cs.uni-bonn.de
Humanoid Robots Lab
University of Bonn
Germany

HRL

Partner institutions:

technische universität dortmund
Fraunhofer IAIS
UNIVERSITÄT BONN
Fraunhofer IML

Institutionally funded by:

Bundesministerium für Forschung, Technologie und Raumfahrt
Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen