# Master's Thesis (m/w/d)
Multi-view Robotic Manipulation Policy with Foundation Models

Problem formulation

Developing robotic agents capable of executing diverse manipulation tasks from visual observations in unstructured real-world environments is a challenging problem. Beyond recognizing objects, the robot must understand their affordances—what actions they can perform with the objects. This includes grasping points, potential tool use, or interaction with other objects. Another major goal of introducing learning into robotic manipulation is to enable the robot to effectively handle unseen objects and successfully tackle various tasks in new environments.
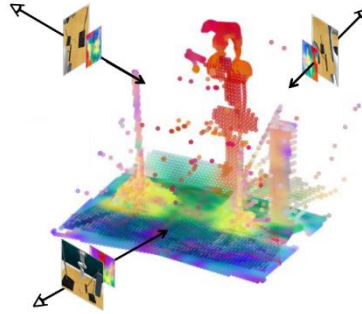
Task definition

In this thesis, the primary objective is to have a comprehensive survey and evaluate the current state-of-the-art robotic manipulation policy that leverages multi-view visual inputs and foundation models to perform complex manipulation tasks in unstructured environments. We'll integrate und evaluate pre-trained foundation models (e.g., DINO, CLIP) on their ability to enhance the robot's understanding of the scene. Finally, experiments on real robots is necessary for policy generalization tests.
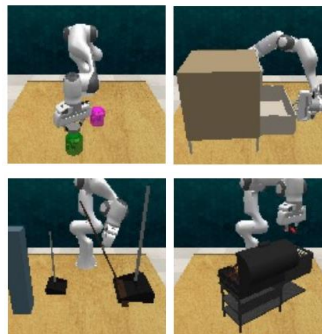
You shall offer
- Solid knowledge and experience in computer vision, deep learning.
- Coding skills in Python and Linux.
- Experience in simulation is a plus.

We will offer
- Powerful robot for experiments
- Powerful GPU server for training your AI.

References
- Ze, Yanjie, et al. "Gnfactor: Multi-task real robot learning with generalizable neural feature fields." Conference on Robot Learning. PMLR, 2023.

- Ke, Tsung-Wei, Nikolaos Gkanatsios, and Katerina Fragkiadaki. "3d diffuser actor: Policy diffusion with 3d scene representations." arXiv preprint arXiv:2402.10885 (2024).

Research area:
Computer Vision,
Deep Learning,
Imitation Learning

Requirement:
☒ Experimental
☒ Theoretical
☒ Practical
☐ Simulation
☐ Construction (CAD)

Studiengang:
☒ Mechanical Engineering
☒ Mechatronics
☒ Electronics
☒ Info-Tech
☐ Informationswirtschaft
☐ Wirtschaftsingenieurwesen

Begin: From now on

If you are interested, please send us an e-mail with your **curriculum vitae** and a current **transcript of records**.

Contact person:
Yitian Shi
Geb. 50.38; Raum 1.15
Phone: +49 721 608 48612
yitain.shi@kit.edu

Please note that your data will be treated in accordance with the applicable data protection regulations as part of the application process.