Foundational Vision-Trajectory Transformer for Industrial Robot Skills

ArtiMinds Robotics & Autonomous Learning Robots (ALR)
Prof. Gerhard Neumann

Project Type

- [X] Master Thesis
- [ ] Bachelor Thesis
- [X] Research Project

Supervisors

Claudius Kienle
claudius.kienle@artiminds.com
Philipp Becker
philipp.becker@kit.edu

Difficulty

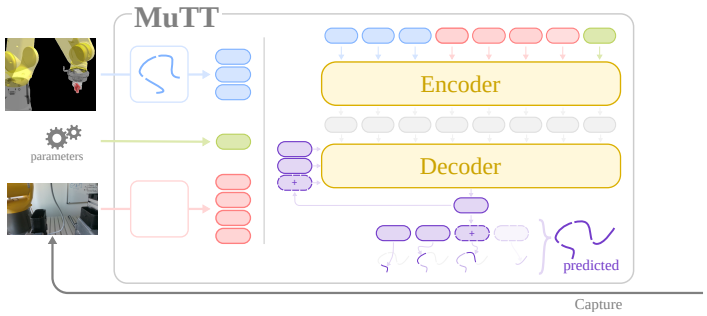
Algorithmic

Math

Application

Description

The use of neural networks for robot skill learning or optimization reduces the labor-intensive programming of robots in industry for the task at hand. However, the current models used for predicting robot skill executions [1, 3] rely heavily on large datasets recorded for specific environments. When the environment changes, the task undergoes slight modifications, or the robot itself is altered, a new dataset must be recorded to fine-tune the models for the new context. Consequently, these models struggle to generalize to new settings, posing a significant limitation on the widespread adoption of innovative machine learning solutions [1, 2] in industry.
This project aims to pioneer the development of a first-of-its-kind multimodal foundational model for industrial robot skills. A recently developed model [3] for robot skills can be used as starting point to be further developed into a foundational model.



Multimodal transformer for prediction of robot skill executions [3].

Tasks

The project involves following tasks:

- Literature Review: Conduct a thorough review of foundational model architectures used in related applications, explore dataset generation techniques for training foundational models, and devise strategies to minimize the required dataset size for efficient training.
- Model Refinement: Optimize the existing model architecture of [3] to ensure its suitability as a foundational model for industrial robot skills.
- Simulated Dataset: Establish a dataset generation pipeline for collecting data within a simulated environment, facilitating the development and training of the foundational model.
- Real-World Validation: Evaluate the developed model on various real-world industrial robots, including KUKA, Fanuc, or UR, available at ArtiMinds Robotics GmbH, to assess its generalization capabilities and suitability for diverse industry applications.

References

[1] Benjamin Alt, Darko Katic, Rainer Jäkel, Asil Kaan Bozcuoglu, and Michael Beetz. Robot program parameter inference via differentiable shadow program inversion. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, page 4672–4678, May 2021.

[2] Lars Johannsmeier, Malkin Gerchow, and Sami Haddadin. A framework for robot manipulation: Skill formalism, meta learning and adaptive control. In *2019 International Conference on Robotics and Automation (ICRA)*, page 5844–5850, May 2019.

[3] Claudius Kienle, Benjamin Alt, Onur Celik, Philipp Becker, Darko Katic, Rainer Jaekel, and Gerhard Neumann. Mutt: A multimodal trajectory transformer for robot skills. in review for iROS 2024.