End-to-end Robotic Manipulation from Verbal Commands

Student Project Proposal
Supervisors: Dr. Emmanuel Senft & Dr. Jean-Marc Odobez
Advisors and point of contact: Dr. Emmanuel Senft
Place: Idiap Research Institute, Martigny (part-time work at EPFL possible)
Keywords: Human-robot interaction, machien learning, transformers, robot manipulation

Project Description

As robots leave factories to enter human spaces, they will need to be easily instructed by people about tasks they should do. An intuitive way for humans is to use natural language to specify actions from the robot, such as "bring me the cup on the table". However, to deliver on such a command, a robot needs to understand the voice command and combine with its perception of the world to know what it should do. Then, it should activate its motors to deliver on the requested command.

Historically, the solution to this problem was using symbolic representations of object poses and then design trajectories specific to each actions. However, with the recent advances in machine learning, new end-to-end network can do the whole process from natural language understanding to command grounding in visual inputs to robot motion. This project aims to build on recent advances to develop new learning architecture to control assistive robot from verbal commands.

In this project (semester or master), you will deep learning methods to control robot from visual and voice inputs. This work builds on recent advances in the field (e.g., https://cliport.github.io/). You will reimplement state-of-the-art ML systems and extend them to control robots both in simulation and with real plateforms.

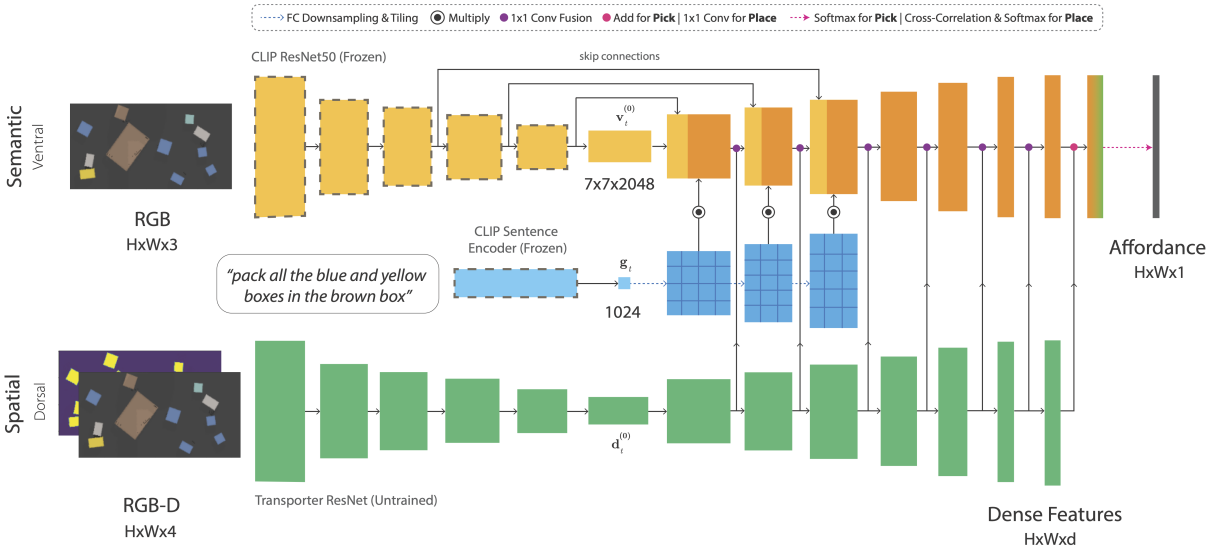

Figure 2: Example of architecture used in CLIport (image from https://cliport.github.io/).

Main tasks and goals

The exact project goals will vary depending of whether the project is a semester project or a master project.

1. Analyze the state of the art of end-to-end methods for robot manipulation.
2. Deploy existing methods (e.g., https://cliport.github.io/) in simulation.
3. Test the algorithm on the real hardware.
4. Extend the method to tackle more complexe environments, for example using mobile robots.

Practical Information

Prerequisites: Good command of Python, good experience with deep learning systems (e.g., pytorch), basics of Linux. Experience in robotics (inverse kinematics, control, or system architecture) would be a plus.

Tools you'll use: PyTorch, Simulators (e.g., Nvidia's Isaac Sim), robot (e.g., Panda, Lio).

Dates: Available immediately.

Contacts: If you are interested or you have any questions, please contact **Emmanuel Senft** (esenft@idiap.ch).