Imitation Learning-Based Control Framework For Vision-Guided Robotic Manipulation

IVR Team

Team Details

Chakrapani (Leader)
Roll No: 240021005
GitHub ID: `chakri2007`
LinkedIn: chakrapani-burra-086878352
Discord: chakrapani2006
Skills: ros2, px4, python

Bhav Makhija
Roll No: 240021004
GitHub ID: `Bhavdex`
LinkedIn: bhav-makhija
Discord: Bhavj
Skills: Algorithm Development

Srujana Patil
Roll No: 240021018
GitHub ID: `Sruj-08`
LinkedIn: srujana-patil
Discord: Sruj08
Skills: python

Tejasri
Roll No: 240041033
GitHub ID: `Tejasri676`
LinkedIn: tejasri-lakshmi-sana
Discord: Sana676
Skills: python

Sathvik
Roll No: 240021001
GitHub ID: `sathvik557`
LinkedIn: amadala-sathvik
Discord: Sathvik0586

Skills: px4, CAD

Project Details

Project Name: Imitation Learning-Based Control Framework For Vision-Guided Robotic Manipulation
Preference No: 3
Domain: IVR

1 Introduction

Imitation learning is transforming robotics by enabling robots to learn complex manipulation tasks from expert demonstrations. The LeRobot framework, with its suite of pre-trained policies (VQ-BeT, Pi0, Pi0Fast, Diffusion), provides a robust foundation for high-fidelity manipulation. This project aims to integrate these models into a real-time ROS2-Gazebo simulation, focusing on the design, analysis, and control of a custom 6-DoF robotic arm for vision-guided pick-and-place tasks. The broader objective is to bridge advanced learning-based control with practical robotic design and simulation.

2 Approach

Model Selection

We propose to use the VQ-BeT model from LeRobot. VQ-BeT is selected for its strong performance in vision-based manipulation, efficient policy inference, and proven robustness in simulation environments. Its discrete latent space and transformer-based architecture make it suitable for mapping visual observations to precise control commands, which is critical for real-time pick-and-place tasks.

Robotic Arm Design Plan

The design will be inspired by industrial 6-DoF arms, adapted for a 20 kg payload.

- **CAD & Simulation:** Modeled in SolidWorks/Fusion 360, exported to URDF/XACRO, and converted to SDF using the `gz` CLI tool.
- **Gripper:** A 2-finger adaptive gripper with a tendon-driven actuation mechanism will be designed.
- **FEA:** Static load and deformation analysis in ANSYS, targeting a safety factor of 1.6.

System Integration and Simulation Environment

- **Simulation Tools:** ROS2 and Gazebo will be used. MoveIt2 will handle motion planning and control.

- **Vision Input:** Camera plugin in Gazebo, publishing image topics compatible with LeRobot.
- **Control Pipeline:** Robot state and object pose fed to VQ-BeT. Outputs converted to ROS2 control commands.
- **Fallback:** A pre-integrated arm will be used for early validation before switching to the custom arm.

3 Timeline

Model Selection: 1 Week
Robotic Arm CAD Design: 2 Weeks
Simulation Setup and Integration: 2 Weeks
Control Pipeline and Model Development: 2 Weeks
Evaluation and Report: 1 Week

4 Conclusion

This project will demonstrate the integration of advanced imitation learning models with custom robotic hardware in a simulated environment, paving the way for robust, vision-guided manipulation systems. The proposed approach emphasizes both innovative control and rigorous mechanical design, ensuring practical feasibility and research value.