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

IVR Team

Team Details

Name	Chakrapani (Leader)
Roll No	240021005
Github Id	`chakri2007`
LinkedIn Id	linkedin.com/in/chakrapani-burra-086878352
Discord	chakrapani2006
Skills	ros2, px4, python
Name	Bhav Makhija
Roll No	240021004
Github Id	`Bhavdex`
LinkedIn Id	linkedin.com/in/bhav-makhija
Discord	Bhavj
Skills	Algorithm Development
Name	Srujana Patil
Roll No	240021018
Github Id	`Sruj-08`
LinkedIn Id	linkedin.com/in/srujana-patil-620677324
Discord	Sruj08
Skills	python
Name	Tejasri
Roll No	240041033
Github Id	`Tejasri676`
LinkedIn Id	linkedin.com/in/tejasri-lakshmi-sana-b6951b321
Discord	Sana676
Skills	python
Name	Sathvik
Roll No	240021001
Github Id	`sathvik557`
LinkedIn Id	linkedin.com/in/amadala-sathvik-333b7a323
Discord	Sathvik0586
Skills	px4, CAD

Project Details

Project Name: Imitation Learning-Based Control Framework For Vision-Guided Robotic Manipulation
Preference Number: 2
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 Reference

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

- **CAD & Simulation:** The arm will be modeled in SolidWorks/Fusion 360, exported to URDF/XACRO, and converted to SDF for Gazebo using the gz CLI tool.
- **Gripper:** A 2-finger adaptive gripper with a tendon-driven actuation mechanism will be designed for dexterous handling of small objects.
- **FEA:** Static load and deformation analysis will be performed in ANSYS, targeting a minimum safety factor of 1.6. Mesh refinement and material selection (e.g., aluminum alloy) will be justified based on simulation results.

System Integration and Simulation Environment

- **Tools:** ROS2 and Gazebo will be used for physics-based simulation. MoveIt2 will handle motion planning and actuation.
- **Vision Input:** A camera plugin will be added in Gazebo, publishing ROS2-compatible image topics for LeRobot inference.

- **Control Pipeline:** Robot state and object pose will be fed to the VQ-BeT model. Model outputs will be mapped to ROS2 control commands for real-time actuation.
- **Fallback:** If integration with the custom arm is challenging, the pre-integrated arm model will be used for initial validation, with continued development of the custom design.

3 Timeline

Task	**Duration**
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.