HARDWARE-II

System Disription

This project explores a vision-based robotic interaction system usng two camera inputs, computer vision, and ROS to control a UR10 robotic arm.

One camera observes the workspace from above and uses OpenCV with ArUco markers to detect and localize objects (cubes) and predefined target positions within a calibrated coordinate system.

A second camera captures human hand gestures, which are interpreted using MediaPipe to extract user intent, such as selecting an object or defining a target position.

Both vision streams are processed independently and communicated through ROS nodes, allowing modular data exchange between perception, decision-making, and robot control.

Based on the detected workspace state and the interpreted gesture input, a task planner generates a pick-and-place command for the robot.

The UR10 robot then executes the movement, autonomously picking up a cube from its current location and placing it at the specified target position.