Vision-Based Lego Detection, Localization and Assembly Using UR5 Robot Arm

Undergraduate Thesis Project by Anh Tu Duong
Supervised by Professor Niculae Sebe

Introduction

The "Vision-Based Lego Detection, Localization, and Assembly Using UR5 Robot Arm" project endeavors to achieve autonomous assembly of Lego blocks through a UR5 robot arm equipped with a ZED camera for perception. The project involves three key components: Vision for object detection and localization, Motion for robot arm manipulation, and the Planning component, currently in development.

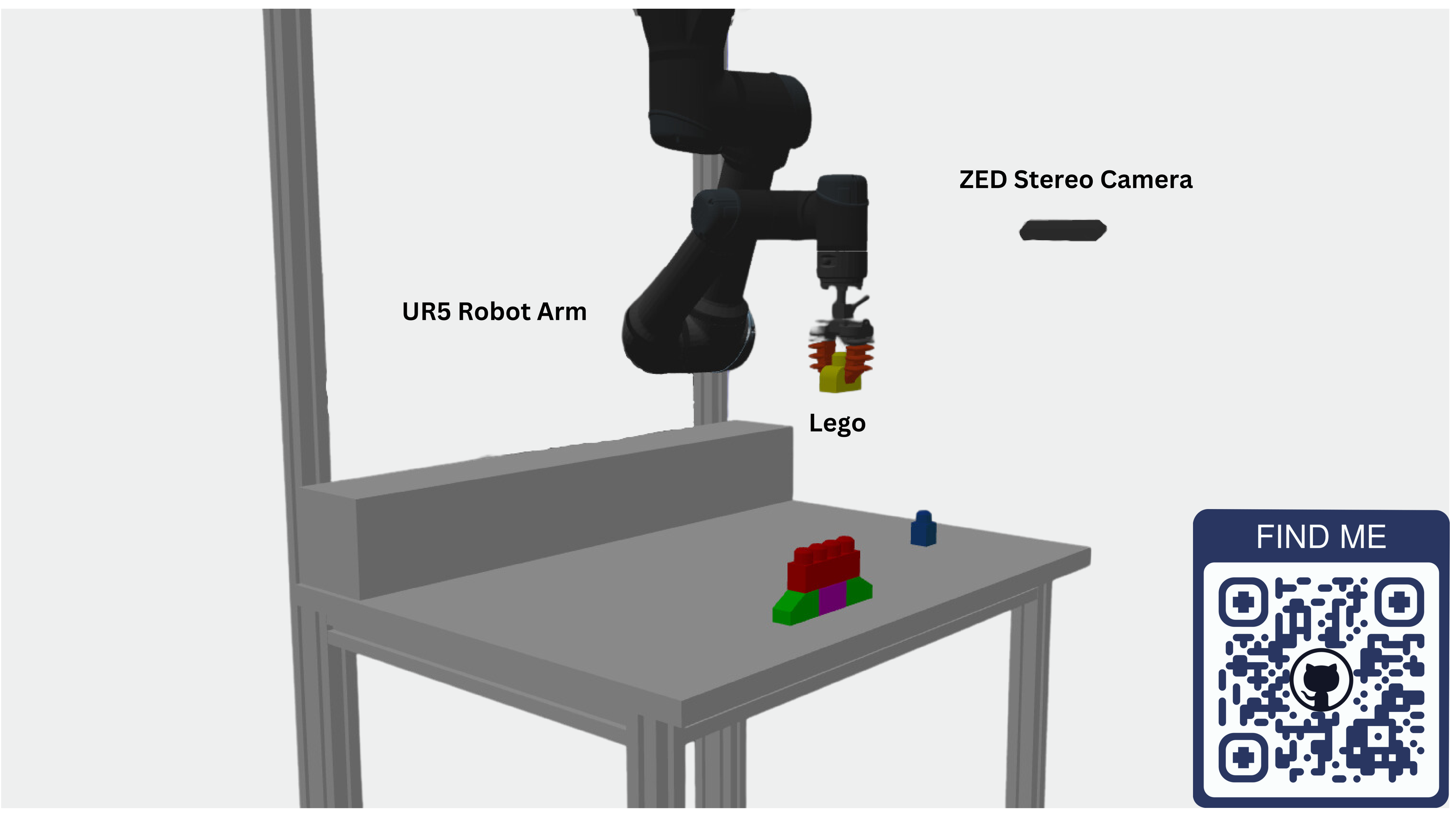

Fig 1. Main scene: UR5 robot performs pick-and-place Lego on the table

Methodology

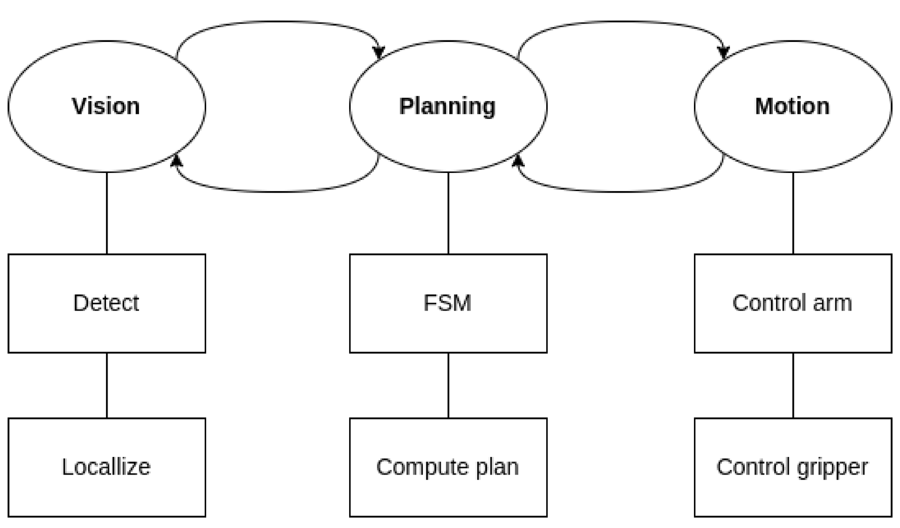

Fig 2. Main workflow

1. Vision component

- Dataset Collection: Compilation of 14,000 images (10% real, 90% synthetic) for training.
- YOLO Training: Utilization of YOLOv8L
- Point Cloud Representation: Transformation of 3D Lego models into point clouds.

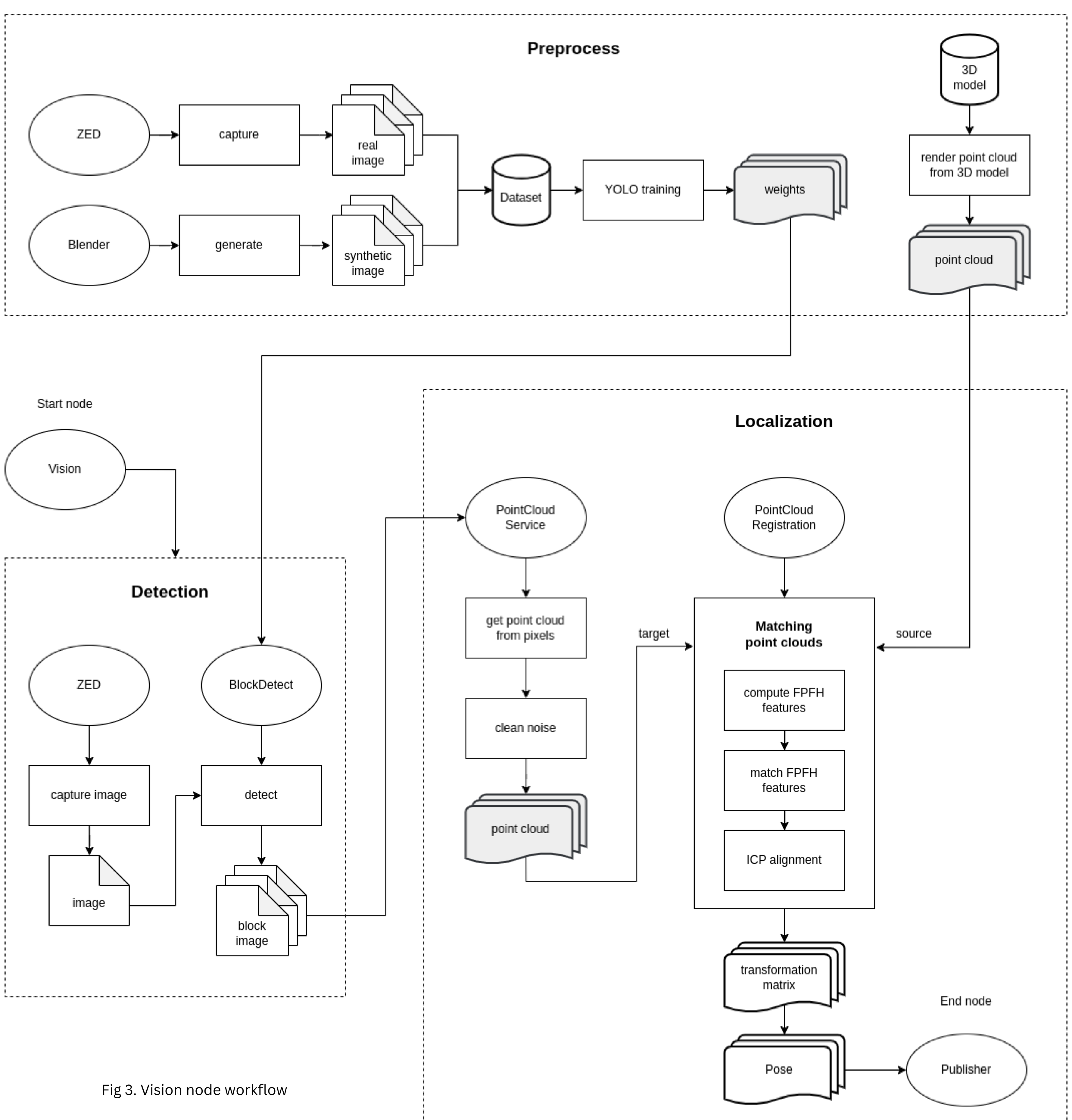

Fig 3. Vision node workflow

- Detection: The process involved capturing ZED camera images, using the trained YOLOv8L model for object detection, and passing bounding box information for localization.
- Localization: FPFH and ICP algorithms facilitated point cloud alignment, allowing accurate localization of Lego blocks.

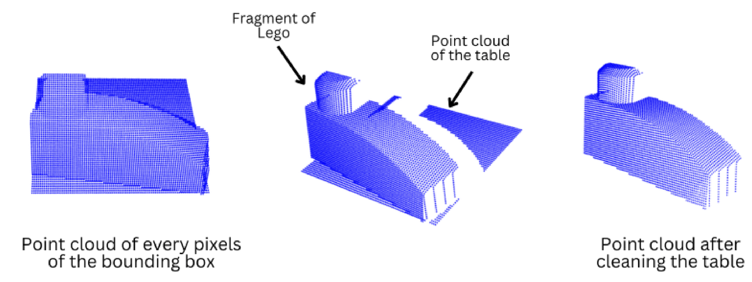

Fig 4. From image of detected Lego to point cloud (target)

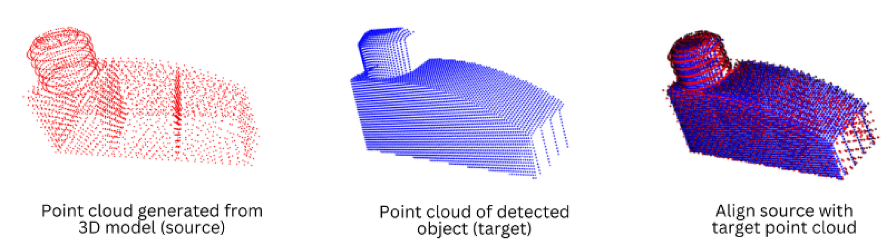

Fig 6. Align point cloud of 3D model to target

2. Motion component

- Command Extraction: Designed API commands for robot arm and gripper control.
- Trajectory Planning: Utilized ROS MoveIt for generating feasible arm trajectories based on Planner's commands.

3. Planning component

- (Working in progress)
- High-level task planning and error handling, integrating Vision and Motion nodes.
- Challenge in converting real image of target Lego building to planning configuration.

Result

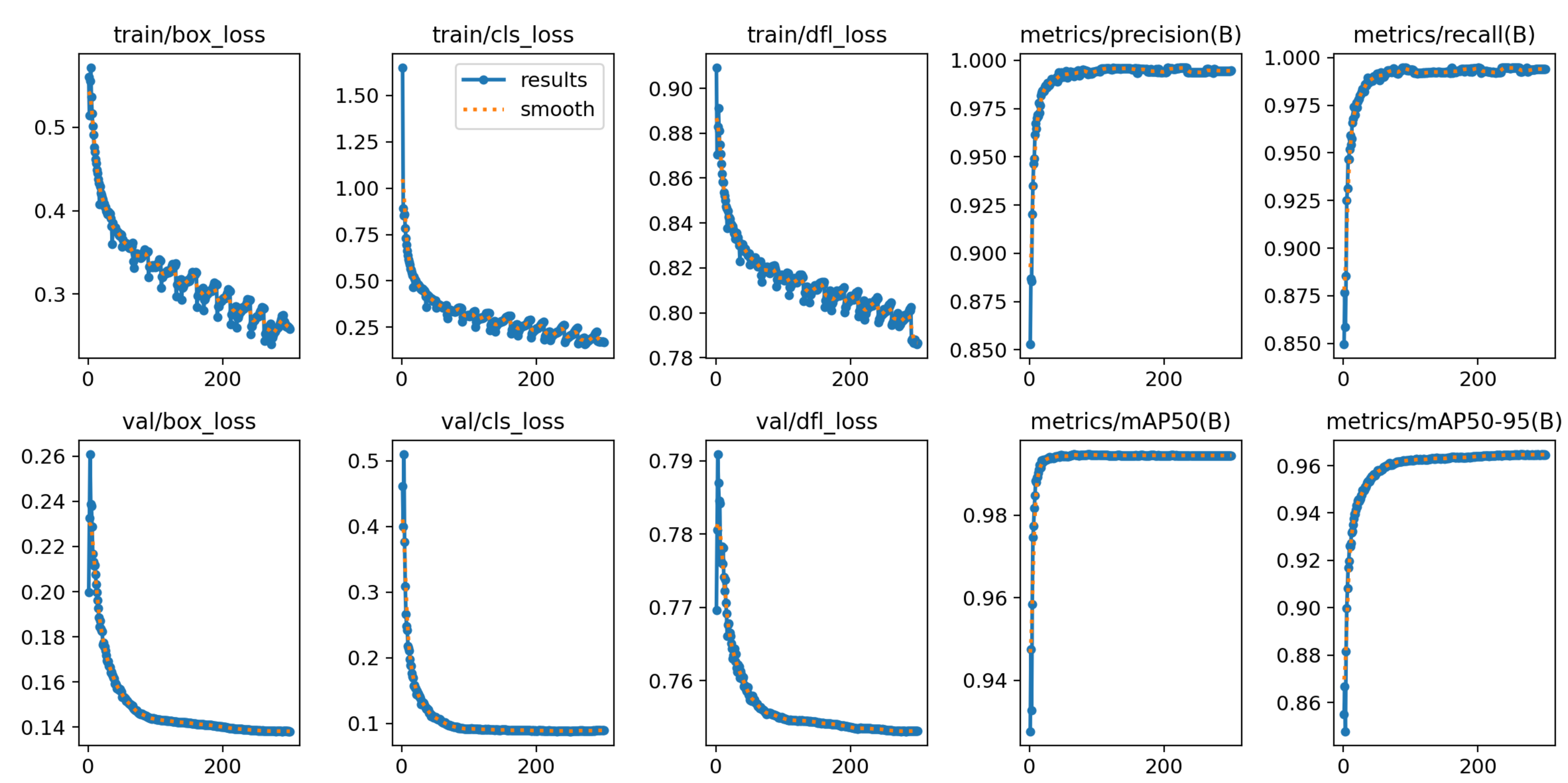

Fig 4. YOLOv8 training results

- YOLO Training: Utilizing YOLOv8L led to an object detection model achieving a remarkable accuracy of 98.5% after 300 epochs.
- Shape Representation: Successful transformation of 3D Lego models into precise point clouds, instrumental for object localization.
- Object Detection and Localization: Effective detection and precise localization of Lego blocks from ZED camera images, facilitated by FPFH and ICP algorithms.

The simulation in Gazebo provided promising outcomes, indicating successful manipulation and placement of Lego blocks. However, during experimentation with the real robot, minor challenges were encountered regarding inaccuracies in object localization. These challenges are currently being addressed to enhance the precision and reliability of the localization process. Overall, the project's success in simulation highlights its potential for achieving autonomous Lego assembly, with a focus on resolving minor challenges encountered in real robot operations.

Conclusion and Future Work

The project has made significant strides in implementing the Vision and Motion components for achieving autonomous Lego assembly using the UR5 robotic arm. However, challenges persist, particularly in the Planning component, which remains under development.

The successes in dataset compilation, YOLO training accuracy, and precise object localization set a strong foundation for the project. Despite successful Gazebo simulation, challenges emerged in the real robot's inaccuracy in object localization, necessitating future refinements.

Moving forward, the focus will be on:

- Improving Object Localization: Addressing accuracy issues in the real robot to ensure precise pick-and-place actions.
- Developing the Planning Component: Integrating Vision and Motion nodes, enhancing high-level coordination and error handling.
- Testing, Refinement, and Documentation: Iterative testing for system refinement and a comprehensive final report detailing the project's methodology and outcomes.

UNIVERSITÀ DI TRENTO
Department of Information Engineering and Computer Science
FONDAZIONE BRUNO KESSLER
FUTURE BUILT ON KNOWLEDGE