­Data Collection and Annotation using ALOHA v2 Stationary Kit

This document outlines the complete process for collecting and annotating data using the ALOHA v2 Stationary Kit. The collected and annotated dataset can later be used for training robotic imitation learning, intention estimation, action recognition, or manipulation policy learning models.

# 1. Data Collection Using ALOHA v2

To record task-specific episodes using ALOHA v2, follow these structured steps:

## 1.1. ROS Environment Setup

Before running any recording scripts, configure the required ROS 2 (Humble) environment and workspace for ALOHA v2.

* Step 1: Source the ROS system install
* source /opt/ros/humble/setup.bash
* Step 2: Source the Interbotix workspace
* source ~/interbotix\_ws/install/setup.bash

## 1.2. Navigating to ALOHA Scripts

* cd ~/interbotix\_ws/src/aloha/scripts/

## 1.3. Running the Episode Recording Script

* Command:
* python3 record\_episodes.py \  
   --task\_name <task\_name> \  
   --robot <robot\_configuration> \  
   [--episode\_idx <episode\_idx>] \  
   [-b, --enable\_base\_torque] \  
   [-g, --gravity\_compensation]

## 1.4. Example Command

* python3 record\_episodes.py \  
   --task\_name pick\_and\_place \  
   --robot dual-arm \  
   --episode\_idx 3 \  
   --gravity\_compensation

## 1.5. Data Storage Structure

~/interbotix\_ws/src/aloha/data/<task\_name>/<robot\_configuration>/episode\_<episode\_idx>/

# 2. Data Annotation

Annotation is critical to identify and label actions within each episode for learning algorithms.

## 2.1. Annotation Methods

### A. Semi-Automatic Annotation (Recommended) — Using CVAT

1. Steps for CVAT Annotation:
2. Install CVAT.
3. Upload dataset.
4. Label bounding boxes, segmentation masks, and action labels.
5. Export annotations.

### B. Manual Annotation (Frame-by-Frame Labeling)

1. Steps for Manual Annotation:
2. Use custom scripts to load frames sequentially.
3. Manually label actions and objects.
4. Save annotations in JSON, CSV, or YAML.

## 2.2. Suggested Annotation Labels

reach, grasp, lift, move, place, release

## 2.3. Organizing Annotated Data

~/interbotix\_ws/src/aloha/data/<task\_name>/<robot\_configuration>/  
├── episode\_0/  
│ ├── images/  
│ ├── joint\_states.json  
│ ├── actions.json  
│ └── other\_data/

# 3. Best Practices and Recommendations

* Calibrate cameras and arms before data collection.
* Use consistent task naming conventions.
* Record 10-20 episodes per task.
* Combine CVAT and manual annotation for accuracy.
* Backup datasets periodically.

# 4. Useful Tools and Links

* CVAT: https://github.com/openvinotoolkit/cvat
* LabelImg: https://github.com/tzutalin/labelImg
* Roboflow: https://roboflow.com
* OpenCV: https://opencv.org/

# 5. Conclusion

By following this data collection and annotation pipeline, you can create a high-quality, well-structured dataset for training and evaluating advanced models for robotic manipulation and learning from demonstration.