

Latent Action Discretization Granularity and Generalization in Multimodal Robot Policies

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the impact of latent action discretization granularity on the generalization capabilities of multimodal robot policies across unseen tasks in the CALVIN dataset. 9 claims were extracted from source literature; 5 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.9/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Open X-Embodiment: Robotic Learning Datasets and RT-X Models. Research question: What is the impact of latent action discretization granularity on the generalization capabilities of multimodal robot policies across unseen tasks in the CALVIN dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.9/10.

3 Results

11 papers retrieved. 9 claims extracted; 5 independently verified. Quality review score: 6.9/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The Open X-Embodiment dataset is assembled from data collected from 22 different robots.	×	0.13
The data collection for the Open X-Embodiment dataset involved a collaboration between 21 institutions.	×	0.12
The Open X-Embodiment dataset demonstrates 527 distinct skills.	×	0.06
The Open X-Embodiment dataset comprises 160,266 tasks.	×	0.04
The RT-X model is a high-capacity model trained on the Open X-Embodiment dataset.	✓	0.20
The RT-X model exhibits positive transfer when applied to multiple robots.	✓	0.15
The RT-X model improves the capabilities of multiple robots by leveraging experience from other platforms.	✓	0.22
Conventionally, robotic learning methods train a separate model for every application, robot, and environment.	✓	0.30
In domains from NLP to Computer Vision, large pretrained models have led to a consolidation of pretrained backbones.	✓	0.24

References

- <https://doi.org/10.48550/arxiv.2412.13877>
- <https://doi.org/10.48550/arxiv.2310.08864>
- <https://doi.org/10.1109/access.2025.3609980>