Sample Efficiency of Continuous Latent Action Models vs. Discrete Tokens in RT-1 Training

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 7 peer-reviewed papers addressing the following research question: What is the difference in sample efficiency between continuous latent action models and discrete token approaches when training on limited labeled data within the RT-1 framework. 14 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: CLAM: Continuous Latent Action Models for Robot Learning from Unlabeled Demonstrations. Research question: What is the difference in sample efficiency between continuous latent action models and discrete token approaches when training on limited labeled data within the RT-1 framework?.

2 Methodology

Systematic literature search across multiple databases yielded 7 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.0/10.

3 Results

7 papers retrieved. 14 claims extracted; 1 independently verified. Quality review score: 5.0/10.

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
CLAM outperforms all baselines and nearly matches the performance of BC with expert data in both state- and image-based	×	0.05
CLAM improves upon the best baseline VPT by more than 2$\times$ average normalized return on the DMControl (locomotion) tasks a	×	0.12
Transformer-CLAM achieves performance close to or even better than that of BC-Expert which uses the same amount of privi	×	0.08
All variants of CLAM outperform the best baseline VPT [11], highlighting the fact that latent action models scale with |	×	0.11
CLAM outperforms state-of-the-art methods in the problem setting where only play data is available as action-labeled dat	✓	0.16
The data regime enables scalable learning from easy-to-collect, cheap play data [21] avoiding the need for expensive tas	×	0.05
BC-AL using action-labeled data unsurprisingly does not perform well due to imitating suboptimal demonstrations.	×	0.09
For state-based inputs, the additional difficulty introduced by not training on ground-truth actions could regularize th	×	0.05
In the image domain, transfer from the pretrained IDM image encoder might cause performance improvements.	×	0.07
CLAM scales to learn capable robot policies in real-world scenarios.	×	0.11
The Transformer-CLAM model uses 6 encoder layers, 6 decoder layers, a feedforward dimension of 2048, 8 attention heads,	×	0.02
The CALVIN environment has a max episode steps of 200, state dimension of 39, action dimension of 7, image shape of [84,	×	0.02
The MetaWorld environment has a max episode steps of 100, state dimension of 39, action dimension of 4, image shape of [	×	0.03
The evaluation environments include locomotion tasks from the DMControl benchmark (Hopper and HalfCheetah) and manipulat	×	0.04

References

- http://arxiv.org/abs/2505.04999v1
- http://arxiv.org/abs/1904.03438v2
- http://arxiv.org/abs/2503.00653v1