Scalable Latent Action Labeling in CLAM vs. Discrete Tokens for Multi-Domain Robotics

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How does the scalability of CLAM's latent action labeling compare to discrete token approaches when evaluated on a multi-domain robotics benchmark (e.g., X-Embodiment) with limited labeled data. 10 claims were extracted from source literature; 4 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.6/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: FATE: A Prompt-Tuning-Based Semi-Supervised Learning Framework for Extremely Limited Labeled Data. Research question: How does the scalability of CLAM's latent action labeling compare to discrete token approaches when evaluated on a multi-domain robotics benchmark (e.g., X-Embodiment) with limited labeled data?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.6/10.

3 Results

16 papers retrieved. 10 claims extracted; 4 independently verified. Quality review score: 5.6/10.

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
FATE achieves state-of-the-art (SOTA) performance on all seven benchmarks compared to current SOTA SSL methods.	×	0.14
FATE improves test accuracy by 4.41% to 50.25% compared with SoftMatch.	×	0.03
FATE's accuracy improved by 3.16% to 34.35% compared to directly fine-tuning the vision model with labeled data.	×	0.09
Directly fine-tuning a vision model with only one labeled sample is insufficient to train a strong backbone from scratch	×	0.08
The proposed method achieved optimal performance on four benchmarks and suboptimal results on four others among ten eval	×	0.04
The proposed method surpasses all other Parameter-Efficient Fine-Tuning (PEFT) methods in overall performance.	×	0.10
General SSL approaches struggle to train effectively from scratch when labeled data is extremely scarce (e.g., a single	✓	0.30
Methods utilizing pre-trained models often fail to find an optimal balance between leveraging limited labeled data and a	✓	0.37
FATE employs a two-stage prompt tuning paradigm that first adapts a pre-trained model to the feature distribution of dow	✓	0.31
FATE applies an SSL method after adapting the pre-trained model to downstream data.	✓	0.21

References

- http://arxiv.org/abs/2505.04999v1

- http://arxiv.org/abs/2504.09828v1
- http://arxiv.org/abs/1910.03560v2