Computational Overhead of Follower-Aware Speaker Models vs Single-Turn Policy Gradients in Navigation

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: How does the computational overhead of the follower-aware speaker model (FOAM) compare to single-turn policy gradient methods in terms of inference time and memory usage during deployment on the. This paper develops LongNav-R1, an end-to-end multi-turn reinforcement learning (RL) framework designed to optimize Visual-Language-Action (VLA) models for long-horizon navigation. Unlike existing single-turn paradigm, LongNav-R1 reformulates the navigation decision process as a. 13 claims were extracted from source literature; 5 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: LongNav-R1: Horizon-Adaptive Multi-Turn RL for Long-Horizon VLA Navigation. Research question: How does the computational overhead of the follower-aware speaker model (FOAM) compare to single-turn policy gradient methods in terms of inference time and memory usage during deployment on the RxR-CE benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.5/10.

3 Results

14 papers retrieved. 13 claims extracted; 5 independently verified. Quality review score: 6.5/10.

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
LongNav-R1 outperforms state-of-the-art methods in long-horizon VLA navigation.	✓	0.22
LongNav-R1 demonstrates zero-shot performance in long-horizon real-world navigation settings.	✓	0.23
Current state-of-the-art methods for navigation adopt a single-turn imitation learning paradigm.	×	0.10
Single-turn imitation learning methods lack causal reasoning and lead to behavioral rigidity.	×	0.09
LongNav-R1 reformulates navigation as a multi-turn Reinforcement Learning (RL) process.	✓	0.24
LongNav-R1 provides comprehensive state and objective awareness, learning causal relationships between actions and dista	×	0.07
LongNav-R1 encourages exploration of diverse trajectories, improving robustness against environmental stochasticity.	×	0.07
Multi-turn RL for long-horizon VLA navigation faces challenges in temporal credit assignment.	✓	0.26
LongNav-R1 uses a horizon-adaptive multi-turn RL approach to address temporal credit assignment without auxiliary critic	✓	0.20
LongNav-R1 has been experimentally validated in real-world and diverse navigation benchmarks, outperforming existing met	×	0.14
Early semantic navigation methods focused on imitation learning or RL, suffering from poor generalization due to domain	×	0.05
Recent semantic navigation approaches leverage LLMs and VLMs for greater flexibility and adaptability in novel environme	×	0.02
LongNav-R1 trains VLA models end-to-end with navigation objectives, offering task-aware performance.	×	0.12

References

- http://arxiv.org/abs/2602.12351v1
- http://arxiv.org/abs/2007.04309v3

- `http://arxiv.org/abs/2507.02910v1`