Multi-Turn Reinforcement Learning in LongNav-R1 Outperforms Single-Turn Approaches on RxR-CE Benchmark

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 9 peer-reviewed papers addressing the following research question: How does the multi-turn RL framework in LongNav-R1 compare to single-turn approaches in terms of accuracy on the RxR-CE benchmark when evaluated with Success Weighted by Path Length (SPL) and goal. 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. 14 claims were extracted from source literature; 5 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.5/10. 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 multi-turn RL framework in LongNav-R1 compare to single-turn approaches in terms of accuracy on the RxR-CE benchmark when evaluated with Success Weighted by Path Length (SPL) and goal completion rate metrics?.

2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.5/10.

3 Results

9 papers retrieved. 14 claims extracted; 5 independently verified. Quality review score: 5.5/10.

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 tasks.	✓	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.	×	0.08
LongNav-R1 reformulates navigation as a multi-turn Reinforcement Learning (RL) process.	✓	0.22
Multi-turn RL in LongNav-R1 provides comprehensive state and objective awareness.	×	0.11
LongNav-R1 learns directly from online interactions, improving robustness against environmental stochasticity.	×	0.10
Multi-turn RL deployment is bottlenecked by the challenge of temporal credit assignment.	×	0.13
LongNav-R1 uses a horizon-adaptive mechanism to manage temporal credit assignment.	✓	0.17
LongNav-R1 allows large VLA models to improve multi-step decision-making without significant computational burden.	×	0.08
LongNav-R1 significantly outperforms existing methods in real-world and diverse navigation benchmarks.	✓	0.15
Early methods for semantic navigation focused on acquiring task-specific skills via imitation learning or RL.	×	0.04
Recent approaches leverage LLMs and VLMs to improve multi-task navigation but lack optimized task execution and navigati	×	0.04
LongNav-R1 trains VLA models end-to-end with navigation objectives, offering both task-awareness and navigation efficien	×	0.12

References

- http://arxiv.org/abs/2409.02392v2

- http://arxiv.org/abs/2601.06757v1
- http://arxiv.org/abs/2602.12351v1