Horizon-Adaptive Multi-Turn Reinforcement Learning for Robust VLA Navigation Under Visual Obscurations

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: What is the impact of horizon-adaptive multi-turn reinforcement learning on the robustness of VLA navigation policies against visual obscurations in simulated 3D environments. 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. 18 claims were extracted from source literature; 3 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.2/10. 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: What is the impact of horizon-adaptive multi-turn reinforcement learning on the robustness of VLA navigation policies against visual obscurations in simulated 3D environments?.

2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.2/10.

3 Results

10 papers retrieved. 18 claims extracted; 3 independently verified. Quality review score: 5.2/10.

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 navigation tasks.	×	0.12
LongNav-R1 demonstrates zero-shot performance in long-horizon real-world navigation settings.	✓	0.24
All source code for LongNav-R1 will be open-sourced upon publication.	×	0.04
Historically, navigation systems relied on modular pipelines involving separate perception, mapping, and planning compon	×	0.01
Recent progress in navigation has shifted toward end-to-end Vision-Language-Action (VLA) models.	×	0.13
Existing state-of-the-art navigation methods adopt a single-turn imitation learning paradigm.	×	0.10
Single-turn imitation learning treats navigation steps independently, overlooking sequential dependencies.	×	0.05
Single-turn imitation learning leads to behavioral rigidity by strictly imitating expert trajectories instead of optimiz	×	0.07
LongNav-R1 reformulates navigation as a multi-turn Reinforcement Learning (RL) process.	✓	0.23
LongNav-R1 treats the navigation task as a continuous conversation between the VLA policy and the physical environment.	×	0.13
Multi-turn RL deployment is bottlenecked by the challenge of temporal credit assignment.	×	0.15
Actor-critic methods like PPO manage temporal credit assignment via learned value functions but incur prohibitive comput	×	0.06
LongNav-R1 allows large VLA models to improve multi-step decision-making without the significant computational burden of	×	0.08
LongNav-R1 was experimentally validated in real-world and diverse navigation benchmarks.	✓	0.15
Early semantic navigation methods largely focused on acquiring task-specific skills via imitation learning or RL.	×	0.04
Early semantic navigation methods often suffer from poor generalization due to domain gaps.	×	0.02
Recent approaches leveraging LLMs and VLMs for navigation offer greater flexibility but often lack optimized task execut	×	0.03
LongNav-R1 trains the VLA model end-to-end with a navigation objective.	×	0.12

References

- http://arxiv.org/abs/1604.02485v1
- http://arxiv.org/abs/2603.13782v1
- http://arxiv.org/abs/2602.12351v1