LongNav-R1 Zero-Shot Performance in Long-Horizon Embodied Navigation Tasks

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How does the task completion accuracy of small-scale 3B multimodal policies scale relative to 7B and 13B models when faced with increasing instruction complexity in embodied navigation 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. 20 claims were extracted from source literature; 5 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.3/10. 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 task completion accuracy of small-scale 3B multimodal policies scale relative to 7B and 13B models when faced with increasing instruction complexity in embodied navigation environments?.

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 5.3/10.

3 Results

16 papers retrieved. 20 claims extracted; 5 independently verified. Quality review score: 5.3/10.

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 demonstrates zero-shot performance in long-horizon real-world navigation settings.	✓	0.23
All source code for LongNav-R1 will be open-sourced upon publication.	×	0.06
Historically, navigation systems relied on modular pipelines involving separate perception, mapping, and planning compon	×	0.04
Recent progress in navigation has shifted toward end-to-end Vision-Language-Action (VLA) models.	✓	0.16
Existing state-of-the-art navigation methods adopt a single-turn imitation learning paradigm.	×	0.10
The single-turn imitation learning paradigm reduces navigation to a sequence of isolated action predictions based on imm	×	0.06
LongNav-R1 is an end-to-end framework that reformulates navigation as a multi-turn Reinforcement Learning (RL) process.	✓	0.27
LongNav-R1 treats the navigation task as a continuous conversation between the VLA policy and the physical environment.	×	0.14
The deployment of multi-turn RL is bottlenecked by the challenge of temporal credit assignment.	✓	0.17
Actor-critic methods like PPO manage temporal credit assignment via learned value functions.	×	0.05
Actor-critic methods incur prohibitive computational overhead due to auxiliary critic networks.	×	0.03
LongNav-R1 allows large VLA models to improve multi-step decision-making without the significant computational burden of	×	0.08
LongNav-R1 significantly outperforms existing methods in real-world and diverse navigation benchmarks.	✓	0.15
Semantic navigation requires agents to navigate to a specific target in unseen environments based on human instructions.	×	0.06
Object Goal Navigation and Vision-and-Language Navigation are two representative tasks involving both visual information	×	0.09
Early semantic navigation methods largely focused on acquiring task-specific skills via imitation learning or RL.	×	0.06
Early navigation methods often suffer from poor generalization due to domain gaps.	×	0.03
Recent approaches leverage Large-Language Models (LLMs) and Vision-Language Models (VLMs) to improve multi-task navigati	×	0.10
Recent LLM- and VLM-based navigation ap	×	0.05

References

- http://arxiv.org/abs/2602.12351v1
- http://arxiv.org/abs/2402.04177v3
- http://arxiv.org/abs/2603.21389v1