

RxR and R2R Agent Performance on ALFRED Long-Horizon Navigation Tasks

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the performance of RxR-trained agents compare to those trained on Room-to-Room (R2R) when evaluated on the ALFRED benchmark for long-horizon language-grounded navigation tasks. Existing Vision-Language Navigation (VLN) methods primarily focus on single-stage navigation, limiting their effectiveness in multi-stage and long-horizon tasks within complex and dynamic environments. To address these limitations, we propose a novel VLN task, named Long-Horizon. 0 claims were extracted from source literature; 0 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: Towards Long-Horizon Vision-Language Navigation: Platform, Benchmark and Method. Research question: How does the performance of RxR-trained agents compare to those trained on Room-to-Room (R2R) when evaluated on the ALFRED benchmark for long-horizon language-grounded navigation tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 13 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

13 papers retrieved. 0 claims extracted; 0 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.

References

- <http://arxiv.org/abs/2412.09082v3>
- <http://arxiv.org/abs/2010.07954v1>
- <http://arxiv.org/abs/2604.17019v1>