Reinforcement Learning-Based VLN Models: Sample Efficiency and Convergence on RxR vs. R2R

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does the sample efficiency and convergence speed of reinforcement learning-based VLN models trained on RxR compare to those trained on R2R when scaling instruction complexity and language. This report presents the methods of the winning entry of the RxR-Habitat Competition in CVPR 2022. The competition addresses the problem of Vision-and-Language Navigation in Continuous Environments (VLN-CE), which requires an agent to follow step-by-step natural language. 13 claims were extracted from source literature; 3 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.1/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: 1st Place Solutions for RxR-Habitat Vision-and-Language Navigation Competition (CVPR 2022). Research question: How does the sample efficiency and convergence speed of reinforcement learning-based VLN models trained on RxR compare to those trained on R2R when scaling instruction complexity and language diversity?.

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 5.1/10.

3 Results

13 papers retrieved. 13 claims extracted; 3 independently verified. Quality review score: 5.1/10.

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
The RxR-Habitat Competition features a continuous action space where agents perform low-level controls such as moving fo	×	0.10
Grounding instructions directly to low-level actions leads to poor performance and loss of state tracking because a sub-	×	0.07
Single-vector state representation is insufficient for the long-trajectory instruction-following task in the RxR dataset	×	0.08
In the RxR-Habitat Competition, sliding along obstacles on collision is forbidden.	×	0.11
Transferring controlling methods from sliding-allowed to sliding-forbidden environments leads to a dramatic performance	×	0.02
Existing VLN controlling methods are unaware of collisions and unable to escape getting stuck, resulting in navigation f	×	0.07
The proposed solution uses a modular plan-and-control model consisting of a candidate waypoints predictor, a history enh	✓	0.23
The candidate waypoints predictor generates a local adjacent navigation graph to bridge differences between discrete and	×	0.11
The history enhanced planner uses a transformer-based architecture with an additional history encoder to track navigatio	✓	0.16
The tryout controller is a non-parametric controller that executes low-level actions to reach planned subgoals.	✓	0.17
The tryout controller detects collisions and helps the agent escape getting stuck using a heuristic trial-and-error mech	×	0.15
The system employs in-domain vision-and-language pretraining using approximately 1 million RxR-like synthetic instructio	×	0.07
The system adopts style-transfer-augmentation on RGB observations.	×	0.03

References

- http://arxiv.org/abs/2206.11610v2
- http://arxiv.org/abs/2104.04167v2
- http://arxiv.org/abs/1911.09615v1