Quality-Diversity Neuroevolution vs. Gradient-Based RL in MuJoCo Locomotion Benchmarks

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How do quality-diversity neuroevolution algorithms compare to gradient-based RL methods in terms of sample efficiency and final reward on MuJoCo locomotion benchmarks. Achieving fast and stable off-policy learning in deep reinforcement learning (RL) is challenging. Most existing methods rely on semi-gradient temporal-difference (TD) methods for their simplicity and efficiency, but are consequently susceptible to divergence. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Deep Reinforcement Learning with Gradient Eligibility Traces. Research question: How do quality-diversity neuroevolution algorithms compare to gradient-based RL methods in terms of sample efficiency and final reward on MuJoCo locomotion benchmarks?.

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 3.7/10.

3 Results

16 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.7/10.

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/2211.02193v1
- http://arxiv.org/abs/2211.13742v2
- http://arxiv.org/abs/2507.09087v2