

Distributionally Robust Contrastive Learning for Stable Cross-Modal Alignment

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: Does combining distributionally robust optimization with contrastive learning enhance the cross-modal alignment stability of multimodal models when evaluated on corrupted retrieval benchmarks. 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Group Distributionally Robust Optimization-Driven Reinforcement Learning for LLM Reasoning. Research question: Does combining distributionally robust optimization with contrastive learning enhance the cross-modal alignment stability of multimodal models when evaluated on corrupted retrieval benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.8/10.

3 Results

12 papers retrieved. 10 claims extracted; 0 independently verified. Quality review score: 2.8/10.

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 DAPO 14.1k English dataset is used for all training runs.	×	0.02
Metrics reported include mean@8 and pass@8 for benchmark accuracies.	×	0.02
Prompt-GDRO keeps the rollout budget fixed and changes training pressure through bin-wise reweighting of GRPO updates.	×	0.08
Rollout-GDRO keeps the mean rollout budget fixed and redistributes rollouts across bins.	×	0.11
The framework demonstrates robust gains across all settings, with Prompt-GDRO achieving a peak gain of +13.13% on the 4B	×	0.05
Rollout-GDRO achieves comparable or superior results, improving the 1.7B model by +10.64% and the 8B model by +9.20%.	×	0.05
Standard instruction tuning paradigms are characterized by static uniformity: a uniform distribution over prompts and a	×	0.12
The Multi-Adversary Framework decomposes the training process into dynamic distributional levers.	×	0.06
The framework includes an Adversarial Sampler (qprompt) and an Adversarial Budgeter (nrollout).	×	0.04
Dynamic Grouping via Online Pass@k defines groups solely through empirical interaction.	×	0.05

References

- <http://arxiv.org/abs/2405.13160v2>
- <http://arxiv.org/abs/2601.19280v1>

- <http://arxiv.org/abs/2207.05796v1>