Multimodal Language Model Alignment and Diminishing Returns in Visual Question Answering

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: Do larger multimodal language models exhibit diminishing returns in alignment quality compared to text-only performance gains when fine-tuned on standard VQA datasets. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Making Large Language Models Better Reasoners with Alignment. Research question: Do larger multimodal language models exhibit diminishing returns in alignment quality compared to text-only performance gains when fine-tuned on standard VQA datasets?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

11 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 3.8/10.

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
AFT significantly outperforms VFT on all three datasets, improving the average accuracy by 1.91% $\sim$2.57% for all models.	×	0.04
Our concurrent work RFT also expresses notable improvement compared with VFT.	×	0.01
Our proposed two constraint alignment strategies slightly outperform RFT with the binary feedback.	×	0.10
AFT can be easily extended to utilize the ranking feedback that RFT cannot well utilize.	×	0.05
AFT surpasses all other methods in the GSM8k-RANK, demonstrating its effectiveness with ranking feedback.	×	0.04
AFT exceeds the strongest baseline RFT by 0.88% in average accuracy in the GSM8k-RANK.	×	0.03
AFT helps LLMs recognize quality differences among any given pair in a ranking context.	×	0.05
The scoring behaviors of vanilla fine-tuned LLMs exhibit misalignment with the gold standard assessment.	×	0.08
AFT introduces an alignment objective LA on top of the VFT objective LVFT.	×	0.03
AFT requires generating multiple COTs for each question in the training set.	×	0.08
VFT-LLMs fail to give reasonable scores to COTs in GP and GN.	×	0.07
The alignment objective in AFT aims to ensure that the scores of all positive COTs in GP are larger than that of negativ	×	0.12

References

- http://arxiv.org/abs/2305.14882v2
- http://arxiv.org/abs/2309.02144v1
- http://arxiv.org/abs/2110.06500v2