

Adaptive Reasoning Strategies in Multimodal LLMs: Accuracy Gains on Dynamic Visual Benchmarks

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the impact of ARS on the reasoning accuracy of multimodal LLMs (e.g., Flamingo or GIT) when evaluated on dynamic visual reasoning benchmarks such as VQA-v2 or TextVQA, compared to static. Recent advances in Multimodal Large Language Models (MLLMs) have significantly improved performance on tasks such as visual grounding and visual question answering. However, the reasoning processes of these models remain largely opaque; they typically output only final. 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: Visual Reasoning Tracer: Object-Level Grounded Reasoning Benchmark. Research question: What is the impact of ARS on the reasoning accuracy of multimodal LLMs (e.g., Flamingo or GIT) when evaluated on dynamic visual reasoning benchmarks such as VQA-v2 or TextVQA, compared to static fine-tuned baselines?.

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 6.5/10.

3 Results

12 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/2603.03437v1>
- <http://arxiv.org/abs/2305.14882v2>
- <http://arxiv.org/abs/2512.05091v1>