

Scaling Model Size Enhances Actionable Counterfactual Learning in Complex Visual Scenes

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 1 peer-reviewed paper addressing the following research question: What is the effect of scaling model size on the ability to learn actionable counterfactual representations from the Causal Triplet benchmark's complex visual scenes. 9 claims were extracted from source literature; 9 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 8.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Explainable and Interpretable Multimodal Large Language Models: A Comprehensive Survey. Research question: What is the effect of scaling model size on the ability to learn actionable counterfactual representations from the Causal Triplet benchmark's complex visual scenes?.

2 Methodology

Systematic literature search across multiple databases yielded 1 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.8/10.

3 Results

1 papers retrieved. 9 claims extracted; 9 independently verified. Quality review score: 8.8/10.

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 rapid development of Artificial Intelligence (AI) has revolutionized numerous fields, with large language models (LL	✓	0.45
The convergence of these technologies has catalyzed the rise of multimodal AI, enabling richer, cross-modal understandin	✓	0.37
Multimodal large language models (MLLMs) have emerged as a powerful framework, demonstrating impressive capabilities in	✓	0.44
The complexity and scale of MLLMs introduce significant challenges in interpretability and explainability, essential for	✓	0.34
This paper provides a comprehensive survey on the interpretability and explainability of MLLMs, proposing a novel framew	✓	0.41
We systematically analyze interpretability from token-level to embedding-level representations.	✓	0.24
We assess approaches related to both architecture analysis and design.	✓	0.19
We explore training and inference strategies that enhance transparency.	✓	0.21
By comparing various methodologies, we identify their strengths and limitations and propose future research directions t	✓	0.25

References

- <https://doi.org/10.48550/arxiv.2412.02104>