Structural Graph Priors and Scaling Laws in Multimodal Vision-Language Models

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the integration of structural graph priors affect the scaling laws of multimodal models compared to pure attention architectures on vision-language benchmarks. Multimodal Transformers serve as the backbone for state-of-the-art vision-language models, yet their quadratic attention complexity remains a critical barrier to scalability. In this work, we investigate the viability of Linear Attention (LA) as a high-efficiency alternative. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: On The Application of Linear Attention in Multimodal Transformers. Research question: How does the integration of structural graph priors affect the scaling laws of multimodal models compared to pure attention architectures on vision-language benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.3/10.

3 Results

15 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 2.3/10.

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/2306.09265v1
- http://arxiv.org/abs/2604.10064v1
- http://arxiv.org/abs/2203.05985v1