

Multimodal vs. Unimodal Small Language Models on SLM-Bench: Accuracy and Efficiency Trade-offs

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How do multimodal SLMs under 10B parameters perform on SLM-Bench compared to unimodal SLMs in terms of accuracy and computational efficiency. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Toward More Reliable Artificial Intelligence: Reducing Hallucinations in Vision-Language Models. Research question: How do multimodal SLMs under 10B parameters perform on SLM-Bench compared to unimodal SLMs in terms of accuracy and computational efficiency?.

2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.7/10.

3 Results

8 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 6.7/10.

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/2602.00982v1>
- <http://arxiv.org/abs/2508.19294v2>
- <http://arxiv.org/abs/2512.07564v1>