

# Multimodal Model Performance on Low-Resource Diagram Understanding in HumanEval-V

Assignee Research

June 6, 2026

## Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How do multimodal models compare in performance on HumanEval-V tasks when evaluated for low-resource diagram understanding with limited training data. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Phi-3 Technical Report: A Highly Capable Language Model Locally on Your Phone. Research question: How do multimodal models compare in performance on HumanEval-V tasks when evaluated for low-resource diagram understanding with limited training data?.

## 2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.7/10.

## 3 Results

13 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.7/10.

## 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

- <https://doi.org/10.48550/arxiv.2307.06435>
- <https://doi.org/10.48550/arxiv.2403.05530>
- <https://doi.org/10.48550/arxiv.2404.14219>