Multimodal Input Integration Enhances DeepSeek R1 Vulnerability Repair Performance

Assignee Research

May 30, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does the integration of multimodal inputs like AST and control flow graphs affect the vulnerability repair capabilities of DeepSeek R1 compared to Codestral, when evaluated on the Big-Vul dataset. The advances of deep learning (DL) have paved the way for automatic software vulnerability repair approaches, which effectively learn the mapping from the vulnerable code to the fixed code. Nevertheless, existing DL-based vulnerability repair methods face notable limitations: 1). 10 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Multi-LLM Collaboration + Data-Centric Innovation = 2x Better Vulnerability Repair. Research question: How does the integration of multimodal inputs like AST and control flow graphs affect the vulnerability repair capabilities of DeepSeek R1 compared to Codestral, when evaluated on the Big-Vul dataset with respect to accuracy and throughput?.

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 4.5/10.

3 Results

12 papers retrieved. 10 claims extracted; 1 independently verified. Quality review score: 4.5/10.

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.

5 Extracted Claims

Claim	Verified	Confidence
Vulnerable-fixed code pairs associated with the input CWE type are deemed more relevant for repair than those linked to	×	0.07
The relevance prediction module performs a binary classification task to distinguish between 'most related' and 'less re	×	0.04
In the relevance prediction module, pairs belonging to the input CWE type are labeled as class 1 ('most related'), while	×	0.04
The relevance prediction module uses an MLP classifier to predict relevance scores based on contextual embeddings genera	×	0.02
The model updates its parameters by minimizing Cross-Entropy loss defined as the sum of -(g_k * log(p_k) + (1 - g_k) * l	×	0.02
VulMaster constructs a concatenated contextual embedding (C_encoder) comprising input vulnerable function segments, AST	×	0.05
VulMaster utilizes the Abstract Syntax Tree (AST) as part of its input to capture the structural aspects of vulnerable c	×	0.08
VulMaster follows the Fusion-in-Decoder (FiD) architecture to process and fuse diverse data inputs within the Transforme	×	0.05
Previous approaches VRepair and VulRepair incorporate CWE types into their input data but do not utilize CWE names or vu	×	0.09
The title claims that Multi-LLM Collaboration combined with Data-Centric Innovation results in a 2x improvement in vulne	✓	0.21

References

- http://arxiv.org/abs/2411.07586v1
- http://arxiv.org/abs/2401.15459v3
- http://arxiv.org/abs/2008.13369v1