

Multimodal Feature Integration Enhances Emotional Alignment in EchoMind Models

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: To what extent does the integration of multimodal features (e.g., audio-visual cues) improve emotional alignment scores in EchoMind-trained models compared to text-only fine-tuning on the IEMOCAP. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Orthogonal Disentanglement with Projected Feature Alignment for Multimodal Emotion Recognition in Conversation. Research question: To what extent does the integration of multimodal features (e.g., audio-visual cues) improve emotional alignment scores in EchoMind-trained models compared to text-only fine-tuning on the IEMOCAP dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.2/10.

3 Results

16 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 2.2/10.

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/1809.04931v1>
- <http://arxiv.org/abs/2511.22463v3>
- <http://arxiv.org/abs/2511.14969v1>