

Audio-Visual vs. Vision-Language Models in Corruption-Robust Deception Detection

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the corruption robustness of audio-visual deception detection models compare to vision-language models when evaluated on perturbed multimodal benchmarks. 11 claims were extracted from source literature; 11 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Deep learning modelling techniques: current progress, applications, advantages, and challenges. Research question: How does the corruption robustness of audio-visual deception detection models compare to vision-language models when evaluated on perturbed multimodal 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 7.4/10.

3 Results

15 papers retrieved. 11 claims extracted; 11 independently verified. Quality review score: 7.4/10.

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
Deep learning (DL) is revolutionizing evidence-based decision-making techniques that can be applied across various sectors.	✓	0.29
DL possesses the ability to utilize two or more levels of non-linear feature transformation of the given data via representations.	✓	0.35
DL is a multidisciplinary field that is still in its nascent phase.	✓	0.17
Articles that survey DL architectures encompassing the full scope of the field are rather limited.	✓	0.24
Many DL models exhibit a highly domain-specific efficiency and could be trained by two or more methods.	✓	0.24
Training DL models can be very time-consuming, expensive, and requires huge samples for better accuracy.	✓	0.27
DL is susceptible to deception and misclassification and tends to get stuck on local minima.	✓	0.24
Improved optimization of parameters is required to create more robust DL models.	✓	0.21
DL has already been leading to groundbreaking results in the healthcare, education, security, commercial, industrial, and other sectors.	✓	0.28
Some DL models, like the convolutional neural network (CNN), generative adversarial networks (GAN), recurrent neural networks (RNN), and others.	✓	0.39
The potential of other DL models remains widely unexplored.	✓	0.18

References

- <https://doi.org/10.1007/s10462-023-10466-8>
- <https://doi.org/10.1002/widm.1520>
- <https://doi.org/10.48550/arxiv.2402.10350>