

# Can the TAE method be extended to multimodal LLMs (e.g., Qwen-VL) without significant accuracy degradation, an

Assignee Research

May 29, 2026

## Abstract

Mohammad Mahdi Abootorabi, Amirhosein Zobeiri, Mahdi Dehghani, Mohammadali Mohammadkhani, Bardia Mohammadi, Omid Ghahroodi, Mahdiah Soleymani Baghshah, Ehsaneddin Asgari. Findings of the Association for Computational Linguistics: ACL 2025. 2025.

## 1 Introduction

This paper examines: Ask in Any Modality: A Comprehensive Survey on Multimodal Retrieval-Augmented Generation. Research question: Can the TAE method be extended to multimodal LLMs (e.g., Qwen-VL) without significant accuracy degradation, and how does it affect cross-modal retrieval speed on LAVIS benchmarks like COCO-Captioning?.

## 2 Methodology

Systematic literature search across multiple databases yielded 6 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.6/10.

## 3 Results

6 papers retrieved. 5 claims extracted; 3 independently verified. Quality review score: 6.6/10.

## 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
Multimodal Retrieval-Augmented Generation (RAG) is a comprehensive approach that combines retrieval and generation techn	✓	0.16
The survey covers various aspects of multimodal RAG, including architectures, datasets, evaluation metrics, and applicat	×	0.03
The authors discuss the challenges and future directions in the field of multimodal RAG.	×	0.06
The paper is published in the Findings of the Association for Computational Linguistics: ACL 2025.	✓	0.36
The authors of the paper are Mohammad Mahdi Abootorabi, Amirhosein Zobeiri, Mahdi Dehghani, Mohammadali Mohammadkhani, B	✓	0.68

## References

- <https://doi.org/10.18653/v1/2022.findings-emnlp.67>
- <https://doi.org/10.18653/v1/2025.findings-acl.861>
- <https://doi.org/10.48550/arxiv.2405.02358>