Oracle-RLAIF Training Effects on Cross-Domain Video Understanding Generalization

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does Oracle-RLAIF training affect cross-domain generalization performance on video understanding tasks beyond the MSVD benchmark, measured by zero-shot accuracy on MSR-VTT and DiDeMo datasets. Recently, zero-shot learning (ZSL) emerged as an exciting topic and attracted a lot of attention. ZSL aims to classify unseen classes by transferring the knowledge from seen classes to unseen classes based on the class description. 14 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.1/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Generalized Continual Zero-Shot Learning. Research question: How does Oracle-RLAIF training affect cross-domain generalization performance on video understanding tasks beyond the MSVD benchmark, measured by zero-shot accuracy on MSR-VTT and DiDeMo datasets?.

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 3.1/10.

3 Results

15 papers retrieved. 14 claims extracted; 0 independently verified. Quality review score: 3.1/10.

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
The proposed method outperforms the baselines and existing Continual Zero-Shot Learning (CZSL) methods.	×	0.12
Earlier Zero-Shot Learning (ZSL) approaches were primarily discriminative or non-generative (embedding-based).	×	0.12
Non-generative ZSL methods learn an embedding from visual space to semantic space or vice versa via a linear compatibili	×	0.04
Embedding-based ZSL approaches represent an image class as a point and are unable to capture intra-class variability.	×	0.06
In Generalized ZSL (GZSL), embedding-based approaches show a strong bias towards seen classes because models are trained	×	0.11
Generative models transform a ZSL problem into a typical supervised learning problem by using synthesized examples of un	×	0.06
Continual learning literature is divided into regularization-based, memory-based, and experience replay-based approaches	×	0.10
In the CZSL setting, training data contains unseen classes with their descriptions in textual form.	×	0.10
Chaudhry et al. developed an average gradient episodic memory (A-GEM) based CZSL method for a multi-head setting.	×	0.07
On the CUB dataset, the AGEM+CZSL method achieved a mean Seen Accuracy (mSA) of 24.66, mean Unseen Accuracy (mUA) of 8.5	×	0.02
On the AWA2 dataset, the CZSL-CA+res method achieved a mean Seen Accuracy (mSA) of 57.69, mean Unseen Accuracy (mUA) of	×	0.02
On the SUN dataset, the CZSL-CA+mof method achieved a mean Seen Accuracy (mSA) of 81.86, mean Unseen Accuracy (mUA) of 6	×	0.02
For the CUB dataset, the Offline (Upper bound) CZSL-CV method achieved a score of 34.50 and CZSL-CA achieved 52.40.	×	0.02
For the AWA1 dataset, the Sequential (Lower bound) CZSL-CV method achieved a score of 14.89 and CZSL-CA achieved 28.90.	×	0.02

References

- http://arxiv.org/abs/2507.02910v1
- http://arxiv.org/abs/2011.08508v3
- http://arxiv.org/abs/2205.00049v2