

Frozen Video Representations with Auxiliary Temporal Reconstruction for Long-Form Video Understanding

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the integration of auxiliary temporal reconstruction losses affect the accuracy of frozen video representations on long-form video understanding benchmarks compared to end-to-end trained. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: T*: Re-thinking Temporal Search for Long-Form Video Understanding. Research question: How does the integration of auxiliary temporal reconstruction losses affect the accuracy of frozen video representations on long-form video understanding benchmarks compared to end-to-end trained baselines?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.2/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 7.2/10.

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/2510.03955v1>
- <http://arxiv.org/abs/2504.02259v3>
- <http://arxiv.org/abs/2503.11495v1>