

Synthetic First-Person Video Features in KNN vs. Graph-Based Models for Zero-Shot Action Recognition

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 7 peer-reviewed papers addressing the following research question: How does the performance of k-nearest neighbors classification using features from synthetic first-person video data compare to that of graph-based models on the AVA benchmark for zero-shot. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: An Evaluation of Large Pre-Trained Models for Gesture Recognition using Synthetic Videos. Research question: How does the performance of k-nearest neighbors classification using features from synthetic first-person video data compare to that of graph-based models on the AVA benchmark for zero-shot multimodal action recognition?.

2 Methodology

Systematic literature search across multiple databases yielded 7 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

7 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.3/10.

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/1705.08421v4>
- <http://arxiv.org/abs/2410.02152v1>
- <http://arxiv.org/abs/2105.01563v5>