

# Auxiliary Temporal Prediction Tasks Improve Domain Shift Robustness in Video-JEPA Models

Assignee Research

June 8, 2026

## Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: To what extent do auxiliary temporal prediction tasks enhance robustness against domain shift in video-JEPA models evaluated via linear probing on Kinetics-400. 18 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Video Unsupervised Domain Adaptation with Deep Learning: A Comprehensive Survey. Research question: To what extent do auxiliary temporal prediction tasks enhance robustness against domain shift in video-JEPA models evaluated via linear probing on Kinetics-400?.

## 2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.0/10.

## 3 Results

12 papers retrieved. 18 claims extracted; 0 independently verified. Quality review score: 3.0/10.

## 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 UCF-Olympic dataset is built across the UCF50 and the Olympic Sports datasets.	×	0.01
The UCF-HMDBsmall dataset is built across the UCF50 and the HMDB51 datasets.	×	0.02
The UCF-HMDBfull dataset is built across the UCF50 and the HMDB51 datasets, but with more than doubled number of classes	×	0.01
The UCF-HMDBfull dataset has become one of the most commonly used benchmark datasets for VUDA research.	×	0.13
The average performance of methods on their respective benchmarked datasets is reported in Tables 5, 6, 7, 8, and 9.	×	0.04
The UCF-Olympic dataset has 10 action classes and 1,000 training/testing videos.	×	0.02
The UCF-HMDBsmall dataset has 10 action classes and 1,000 training/testing videos.	×	0.02
The UCF-HMDBfull dataset has 21 action classes and 2,000 training/testing videos.	×	0.02
The Kinetics-Gameplay dataset has 100 action classes and 10,000 training/testing videos.	×	0.02
The HMDB-ARID dataset has 51 action classes and 5,000 training/testing videos.	×	0.01
The Kinetics $\rightarrow$ NEC-Drone dataset has 100 action classes and 10,000 training/testing videos.	×	0.01
The Mixamo $\rightarrow$ Kinetics dataset has 100 action classes and 10,000 training/testing videos.	×	0.02
The ActorShift dataset has 100 action classes and 10,000 training/testing videos.	×	0.02
The Kinetics $\rightarrow$ BABEL dataset has 100 action classes and 10,000 training/testing videos.	×	0.02
The UCF-HMDBpartial dataset has 10 action classes and 1,000 training/testing videos.	×	0.02
The MiniKinetics-UCF dataset has 10 action classes and 1,000 training/testing videos.	×	0.02
The HMDB-ARIDpartial dataset has 51 action classes and 5,000 training/testing videos.	×	0.01
The EPIC Kitchens dataset has 100 action classes and 10,000 training/testing videos.	×	0.02

## References

- <http://arxiv.org/abs/2410.09380v1>
- <http://arxiv.org/abs/2606.05008v1>
- <http://arxiv.org/abs/2211.10412v3>