# Past-Token Prediction Enhances Robustness in Long-Context Robotic Policy Learning

Assignee Research

June 7, 2026

**Abstract**

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: Does adopting past-token prediction as a training objective improve robustness against spurious correlations in long-context policy learning for robotic agents compared to context truncation methods. 13 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: Learning Long-Context Diffusion Policies via Past-Token Prediction. Research question: Does adopting past-token prediction as a training objective improve robustness against spurious correlations in long-context policy learning for robotic agents compared to context truncation methods?.

## 2 Methodology

Systematic literature search across multiple databases yielded 10 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

10 papers retrieved. 13 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 (PTP) achieves an average improvement of over 30% compared to no-history diffusion policies across s | × | 0.06 |
| The proposed method (PTP) achieves an average improvement of over 60% compared to no-PTP diffusion policies across six s | × | 0.07 |
| Modern diffusion-based policies exhibit a clear drop in performance when conditioned on historical observations compared | × | 0.05 |
| The evaluation includes four tasks from existing benchmarks: square, tool hang, and transport from RoboMimic, and Push-T | × | 0.03 |
| Two new long-horizon simulation tasks were introduced: long-horizon square and long-horizon aloha. | × | 0.14 |
| In the long-horizon square task, the robot must place and remove a square onto the peg twice before finally dropping it | × | 0.06 |
| In the long-horizon aloha task, one arm must pick up a block, move it to the center of the field of view, and return it | × | 0.06 |
| By default, both diffusion and regression policies receive visual and proprioceptive observations from the past 16 time | × | 0.06 |
| The method uses Past-Token Prediction (PTP) as an auxiliary objective to jointly predict action tokens from past time st | × | 0.15 |
| The method employs a multi-stage training recipe with feature caching. | × | 0.03 |
| Policies are evaluated under a single-sample inference setting unless otherwise specified. | × | 0.01 |
| The method achieves a 10x training speedup compared to the naive history conditioned approach. | × | 0.03 |
| The method achieves 3x performance increase compared to the naive history conditioned approach. | × | 0.02 |

## References


- http://arxiv.org/abs/2605.11134v2
- http://arxiv.org/abs/2505.09561v2
- http://arxiv.org/abs/2403.03375v3