Diffusion Trajectory-Guided Policy and VLA-Adapter Success Rates Across RoboBench Time Horizons

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 9 peer-reviewed papers addressing the following research question: What is the impact of trajectory length on the success rate of Diffusion Trajectory-guided Policy versus VLA-Adapter in RoboBench, measured by the degradation rate of accuracy over increasing time. 13 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: VLM-TDP: VLM-guided Trajectory-conditioned Diffusion Policy for Robust Long-Horizon Manipulation. Research question: What is the impact of trajectory length on the success rate of Diffusion Trajectory-guided Policy versus VLA-Adapter in RoboBench, measured by the degradation rate of accuracy over increasing time horizons?.

2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.0/10.

3 Results

9 papers retrieved. 13 claims extracted; 0 independently verified. Quality review score: 4.0/10.

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
VLM-TDP improved performance by 3% on average across seven tasks compared to state-of-the-art 3D policies.	×	0.09
VLM-TDP outperformed the diffusion policy in tasks like Open Drawer, Open Wine Bottle, and Sweep to Dustpan.	×	0.08
VLM-TDP was less effective than 3D policies in tasks like Open Drawer, Open Wine Bottle, and Sweep to Dustpan.	×	0.04
Point cloud can handle distinguishable objects more effectively than RGB information.	×	0.02
3D policies struggled to accurately evaluate scenes with crowded or obstructed objects, such as in the Water Plants task	×	0.02
Diffusion policy and VLM-TDP surpassed 3D policies in tasks with crowded and cluttered situations.	×	0.11
VLM-TDP achieved higher success rates in tasks like Phone on Base and Put Item on Drawer by accurately targeting objects	×	0.04
VLM-TDP excelled in the longer-horizon task Put Item in Drawer, which consists of four subtasks.	×	0.10
The performance of VLM-TDP showed a smaller drop compared to other policies as the complexity of choices increased.	×	0.05
Diffusion policies approximate action distributions using denoising diffusion probabilistic models (DDPM).	×	0.07
Diffusion policies have shown significant promise in multimodal and high-dimensional action spaces.	×	0.08
Existing works are limited to short-horizon tasks and are vulnerable to noise and input variability.	×	0.11
Large Language Models (LLMs) and Vision-Language Models (VLMs) have gained attention in the field of robotics.	×	0.11

References

- http://arxiv.org/abs/2507.04524v1
- http://arxiv.org/abs/2502.10040v2
- http://arxiv.org/abs/2409.00588v3