Novelty Assessment Report

Paper: Real-Time Robot Execution with Masked Action Chunking
PDF URL: https://openreview.net/pdf?id=r0RGJ1j9on
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Real-time execution is essential for cyber-physical systems such as robots. These systems operate in dynamic real-world environments where even small delays can undermine responsiveness and compromise performance. Asynchronous inference has recently emerged as a system-level paradigm for real-time robot manipulation, enabling the next action chunk to be predicted while the current one is being executed. While this approach achieves real-time responsiveness, naive integration often results in execution failure. Previous methods attributed this failure to inter-chunk discontinuity and developed test-time algorithms to smooth chunk boundaries. In contrast, we identify another critical yet overlooked factor: intra-chunk inconsistency, where the robot's executed action chunk partially misaligns with its current perception. To address this, we propose REMAC, which learns corrective adjustments on the pretrained policy through masked action chunking, enabling the policy to remain resilient under mismatches between intended actions and actual execution during asynchronous inference. In addition, we introduce a prefix-preserved sampling procedure to reinforce inter-chunk continuity. Overall, our method delivers more reliable policies without incurring additional latency. Extensive experiments in both simulation and real-world settings demonstrate that our method enables faster task execution, maintains robustness across varying delays, and consistently achieves higher completion rates.

Core Task Landscape

This paper addresses: **Real-Time Robot Manipulation Under Asynchronous Inference**

A total of **33 papers** were analyzed and organized into a taxonomy with **20 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Asynchronous Inference Architectures for Vision-Language-Action Models**
- **Deep Reinforcement Learning Training for Real-Time Manipulation**
- **Real-Time Motion Planning and Trajectory Optimization**
- **Real-Time System Infrastructure and Middleware**
- **Perception and Inference Optimization for Real-Time Control**
- **Human-in-the-Loop Post-Training and Refinement**

Complete Taxonomy Tree

- Real-Time Robot Manipulation Under Asynchronous Inference Survey Taxonomy
- Asynchronous Inference Architectures for Vision-Language-Action Models
  - Future-State-Aware and Chunk Correction Approaches ★ (3 papers)
  - [0] Real-Time Robot Execution with Masked Action Chunking (Anon et al., 2026) View paper
  - [1] VLASH: Real-Time VLAs via Future-State-Aware Asynchronous Inference (Jiaming Tang, 2025) View paper
  - [7] Leave no observation behind: Real-time correction for vla action chunks (Alvarez, 2025) View paper
  - Asynchronous Action Execution Frameworks (2 papers)
  - [3] Real-Time Execution of Action Chunking Flow Policies (Black, 2025) View paper
  - [23] Asynchronous Fast-Slow Vision-Language-Action Policies for Whole-Body Robotic Manipulation (Teqiang Zou, 2025) View paper
  - Hierarchical Language-Based Control (1 papers)
  - [4] Rt-h: Action hierarchies using language (Suneel Belkhale, 2024) View paper
- Deep Reinforcement Learning Training for Real-Time Manipulation
  - Asynchronous Off-Policy and Real-World Training (3 papers)
  - [2] Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates (Gu Shixiang, 2017) View paper
  - [15] Non-blocking asynchronous training for reinforcement learning in real-world environments (Peter BÃ¶hm, 2022) View paper
  - [30] Asynchronous Reinforcement Learning for Real-Time Control of Physical Robots (Yufeng, 2022) View paper
  - Model-Based RL with Inference Delay Handling (2 papers)
  - [20] RT-HCP: Dealing with Inference Delays and Sample Efficiency to Learn Directly on Robotic Platforms (Zakaria El Asri, 2025) View paper
  - [33] A Real-Time Model-Based Reinforcement Learning Architecture for Robot Control (Hester, 2011) View paper
  - Neuromorphic Event-Based Control (2 papers)
  - [17] Event-based Vision for Early Prediction of Manipulation Actions (Fermuller, 2023) View paper
  - [24] Population-Coded Spiking Neural Networks for High-Dimensional Robotic Control (Jiang Xiao-yang, 2025) View paper
- Real-Time Motion Planning and Trajectory Optimization
  - Asynchronous MPC and Footstep Planning (3 papers)
  - [11] ASAP-MPC: an asynchronous update scheme for online motion planning with nonlinear model predictive control (Dries Dirckx, 2025) View paper

    - [25] Incremental Model Predictive Control Exploiting Time-Delay Estimation for a Robot Manipulator (Yongchao Wang, 2022) View paper
    - [26] Asynchronous Real-Time Optimization of Footstep Placement and Timing in Bipedal Walking Robots (Chappell, 2022) View paper
    - Predictive Collision Avoidance and Delayed Decision-Making (3 papers)
    - [6] Delayed-Decision Motion Planning in the Presence of Multiple Predictions (David Isele, 2025) View paper
    - [13] Spatio-Temporal Avoidance of Predicted Occupancy in Human-Robot Collaboration (Jared Flowers, 2023) View paper
    - [19] Occlusion-aware trajectory prediction for time-delayed teleoperation of space manipulators (Xiaoyi Qu, 2025) View paper
    - Kinodynamic Planning with Motion Primitives (1 papers)
    - [14] Delayed Expansion AGT: Kinodynamic Planning with Application to Tractor-Trailer Parking (Zheng Dongliang, 2025) View paper
    - Decentralized Multi-Robot Planning (1 papers)
    - [27] Asynchronous Real-time Decentralized Multi-Robot Trajectory Planning (Baskin ÅenbaÅlar, 2022) View paper
    - UAV Real-Time Path Planning (1 papers)
    - [9] Deep Reinforcement Learning based Real-Time Path Planning and Flight Validation of small UAS Application (Jinhyuk Park, 2024) View paper
- Real-Time System Infrastructure and Middleware
    - Real-Time Middleware and Hardware Abstraction (2 papers)
    - [12] The xbot2 real-time middleware for robotics (Arturo Laurenzi, 2023) View paper
    - [31] Probabilistic Articulated Real-Time Tracking for Robot Manipulation (Cristina Garcia Cifuentes, 2022) View paper
    - Real-Time Task Scheduling and Resource Management (1 papers)
    - [8] Red: A systematic real-time scheduling approach for robotic environmental dynamics (Zexin Li, 2023) View paper
    - Distributed Software Infrastructure for Coordination (1 papers)
    - [28] A real-time distributed software infrastructure for cooperating mobile autonomous robots (Santos, 2009) View paper
    - Formal Verification and Model-Checking (1 papers)
    - [10] Modelling and Model-Checking a ROS2 Multi-Robot System using Timed Rebeca (Hiep Hong Trinh, 2025) View paper
    - Supervisory Control and Analog Systems (2 papers)
    - [18] Built-in Analog Automatic Controls for Small Robots (N. L. Dembitsky, 2023) View paper
    - [21] A structural approach to the non-blocking supervisory control of discrete-event systems (Lei Feng, 2009) View paper
- Perception and Inference Optimization for Real-Time Control
    - Active Vision-Action Coordination (1 papers)
    - [5] Observe Then Act: Asynchronous Active Vision-Action Model for Robotic Manipulation (Guokang Wang, 2025) View paper
    - Object Detection Inference Optimization (1 papers)
    - [16] Performance Analysis of YOLOv5 for Real-Time Object Detection and Manipulation in Unity on CPU and GPU (Kankipati Venkata Meghana, 2025) View paper
    - Rule-Based Inference Engines (2 papers)
    - [29] Asynchronous production system for control of an autonomous mobile robot in real-time environment (S. Iyengar, 1992) View paper
    - [32] ADA Implementation of Concurrent Execution of Multiple Tasks in the Strategic and Tactical Levels of the Rational Behavior Model for the NPS Phoenix Autonomous Underwater Vehicle (AUV). (M. J. Holden, 1995) View paper
- Human-in-the-Loop Post-Training and Refinement (1 papers)
    - [22] Human-in-the-loop Online Rejection Sampling for Robotic Manipulation (Lu, 2025) View paper

Narrative

Core task: real-time robot manipulation under asynchronous inference. The field addresses the challenge of executing manipulation policies when perception and decision-making modules run at different rates or with variable latency. The taxonomy organizes work into several main branches: asynchronous inference architectures for vision-language-action models, which handle the mismatch between slow neural network inference and fast control loops; deep reinforcement learning training methods that account for asynchronous data collection; real-time motion planning and trajectory optimization that must react quickly despite delayed observations; system infrastructure and middleware designed to coordinate asynchronous components; perception and inference optimization techniques that reduce latency; and human-in-the-loop refinement approaches that improve policies post-deployment. Representative efforts include asynchronous off-policy RL methods (Asynchronous Off-Policy[2]) for training, middleware solutions (xbot2 Middleware[12], ROS2 Timed Rebeca[10]) for coordination, and model predictive control adaptations (ASAP-MPC[11], Incremental MPC Time-Delay[25]) for planning under delay.

A particularly active line of work focuses on future-state-aware and chunk correction approaches within the asynchronous inference architectures branch, where methods predict or compensate for the time lag between observation and action execution. Masked Action Chunking[0] sits squarely in this cluster, addressing how to generate and refine sequences of actions when inference cannot keep pace with control frequency. It shares thematic concerns with Real-Time Correction VLA[7], which also emphasizes correcting action sequences on-the-fly, and with approaches like Action Chunking Flow[3] that structure action generation to respect temporal dependencies. Nearby work such as Observe Then Act[5] and VLASH[1] similarly grapple with the trade-off between waiting for fresh observations versus acting on potentially stale information. The central tension across these methods is balancing reactivity—how quickly the system can respond to new sensory input—against the computational cost of frequent re-inference, with different solutions offering varying degrees of look-ahead prediction, action buffering, and online correction.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. VLASH: Real-Time VLAs via Future-State-Aware Asynchronous Inference

Authors: Jiaming Tang, Yufei Sun, Yilong Zhao, Shang Yang, Yujun Lin, et al. (10 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Vision-Language-Action models (VLAs) are becoming increasingly capable across diverse robotic tasks. However, their real-world deployment remains slow and inefficient: demonstration videos are often sped up by 5-10x to appear smooth, with noticeable action stalls and delayed reactions to environmental changes. Asynchronous inference offers a promising solution to achieve continuous and low-latency control by enabling robots to execute actions and perform inference simultaneously. However, becaus...

Relationship Analysis

Both papers belong to the Future-State-Aware and Chunk Correction Approaches category, addressing real-time robot manipulation under asynchronous inference by handling temporal misalignment between prediction and execution. While REMAC (the original paper) focuses on learning corrective adjustments through masked action chunking during training to handle intra-chunk inconsistency and inter-chunk discontinuity, VLASH takes a different approach by estimating future execution-time states through forward simulation of robot states with previously generated action chunks, requiring no training modifications or architectural changes to the pretrained VLA model.

2. Leave no observation behind: Real-time correction for vla action chunks

Authors: Alvarez, Maxime, Kohei Sendai, Matsushima, Tatsuya, et al. (13 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

To improve efficiency and temporal coherence, Vision-Language-Action (VLA) models often predict action chunks; however, this action chunking harms reactivity under inference delay and long horizons. We introduce Asynchronous Action Chunk Correction (A2C2), which is a lightweight real-time chunk correction head that runs every control step and adds a time-aware correction to any off-the-shelf VLA's action chunk. The module combines the latest observation, the predicted action from VLA (base actio...

Relationship Analysis

Both papers belong to the Future-State-Aware and Chunk Correction Approaches category, addressing real-time robot manipulation under asynchronous inference by correcting action chunks to handle inference delays. They overlap in targeting intra-chunk inconsistency and inter-chunk discontinuity through correction mechanisms applied to VLA action chunks. The key difference is that the original paper (REMAC) uses a training-time adaptation approach with masked action chunking and LoRA fine-tuning of the base policy, while the candidate paper (A2C2) introduces a lightweight real-time correction head that runs at every control step without retraining the base VLA model, making it a plug-and-play solution orthogonal to the base policy.

Contributions Analysis

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: REMAC: Real-time Execution with Masked Action Chunking

Description: The authors introduce REMAC, a training-time method that adapts pretrained vision-language-action policies for asynchronous inference by learning corrective adjustments through masked action chunking. This approach addresses intra-chunk inconsistency by masking arbitrary portions of action chunks during training, enabling the policy to handle misalignments between observations and executed actions without introducing additional inference latency.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. : a VLA That Learns From Experience

URL: View paper

Brief Assessment

VLA Learns Experience[34] focuses on learning from experience and correction episodes, not on masked action chunking for asynchronous inference with corrective adjustments.

2. Discrete Diffusion VLA: Bringing Discrete Diffusion to Action Decoding in Vision-Language-Action Policies

URL: View paper

Brief Assessment

Discrete Diffusion VLA[37] focuses on discrete diffusion for action decoding in vision-language-action models, not on corrective adjustments for asynchronous inference with masked action chunking.

3. Leave no observation behind: Real-time correction for vla action chunks

URL: View paper

Brief Assessment

Real-Time Correction VLA[7] proposes a lightweight correction head that runs at every control step to adjust action chunks, requiring no retraining of the base policy. In contrast, the original paper's REMAC is a training-time method that adapts pretrained policies through masked action chunking during the learning phase.

4. A Survey on Reinforcement Learning of Vision-Language-Action Models for Robotic Manipulation

URL: View paper

Brief Assessment

VLA Survey[36] appears to be a survey paper discussing general challenges in vision-language-action models. The provided context mentions action chunking and reward sparsity but does not present a specific method for corrective adjustments under asynchronous inference with masked action chunking.

5. AsyncVLA: Asynchronous Flow Matching for Vision-Language-Action Models

URL: View paper

Brief Assessment

AsyncVLA[35] focuses on asynchronous flow matching with temporal flexibility and self-correction for action token generation, not on learning corrective adjustments through masked action chunking for handling perception-action misalignments during asynchronous inference.

Contribution 2: Identification of intra-chunk inconsistency as a critical failure mode

Description: The authors identify and formalize intra-chunk inconsistency as a previously overlooked challenge in asynchronous inference with action chunking. This occurs when executed actions from a previous chunk are conditioned on outdated observations, creating a perception-action mismatch within a single chunk that degrades policy performance.

This contribution was assessed against **4 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. ImplicitRDP: An End-to-End Visual-Force Diffusion Policy with Structural Slow-Fast Learning

URL: View paper

Brief Assessment

ImplicitRDP[39] focuses on integrating visual and force modalities for contact-rich manipulation, not on perception-action misalignment in asynchronous action chunking execution.

---

2. Mobile robot programming using natural language

URL: View paper

Brief Assessment

Natural Language Programming[40] focuses on natural language interfaces for mobile robot programming, not on asynchronous inference or action chunking in robot policies. The candidate's mention of 'chunking' and 'mismatch' appears in a completely different context unrelated to the perception-action misalignment problem identified in the original paper.

---

3. VLASH: Real-Time VLAs via Future-State-Aware Asynchronous Inference

URL: View paper

Prior Art Analysis

VLASH[1] identifies the same fundamental problem of temporal misalignment between prediction and execution in asynchronous inference with action chunking. The candidate paper explicitly describes how 'a temporal misalignment arises between the prediction and execution intervals' during asynchronous inference, which directly corresponds to the original paper's concept of intra-chunk inconsistency. Both papers recognize that this misalignment occurs because 'the robot and environment continue to evolve during inference,' creating a perception-action mismatch. The original paper's formalization of 'intra-chunk inconsistency, where the robot's executed action chunk partially misaligns with its current perception' is substantively equivalent to VLASH[1]'s observation that actions become misaligned with the state at execution time due to the temporal gap.

Evidence

Evidence 1 - **Rationale**: Both papers identify the same core problem: during asynchronous inference, actions predicted at one time are executed at a later time when the robot state has changed, creating a misalignment. The original paper calls this 'intra-chunk inconsistency' while VLASH[1] describes it as 'temporal misalignment between the prediction and execution intervals.' - **Original**: intra-chunk inconsistency, where the robot's executed action chunk partially misaligns with its current perception. to address this, we propose remac, which learns corrective adjustments on the pretrained policy through masked action chunking, enabling the policy to remain resilient under mismatches... - **Candidate**: asynchronous inference offers a promising solution to achieve continuous and low-latency control by enabling robots to execute actions and perform inference simultaneously. however, because the robot and environment continue to evolve during inference, a temporal misalignment arises between the pred...

Evidence 2 - **Rationale**: The original paper's detailed formalization of how actions conditioned on past observations (ot-h) become suboptimal when executed at current state (ot) is conceptually identical to VLASH[1]'s recognition that the robot state evolves during inference, causing temporal misalignment and action instability. - **Original**: intra-chunk inconsistency. assume the policy vπ(at|ot) perfectly captures the underlying environment dynamics and therefore yields the optimal action sequence. given perception ot, the optimal executed chunk should fully correspond to at. however, under inference delay d with execution horizon h, th... - **Candidate**: however, because the robot and environment continue to evolve during inference, a temporal misalignment arises between the prediction and execution intervals. this leads to significant action instability, while existing methods either degrade accuracy or introduce runtime overhead to mitigate it.

---

4. SmolVLA: A Vision-Language-Action Model for Affordable and Efficient Robotics

URL: View paper

Brief Assessment

SmolVLA[38] focuses on efficient VLA model design and asynchronous inference implementation for deployment, not on identifying or formalizing intra-chunk inconsistency as a failure mode.

Contribution 3: Prefix-preserved sampling procedure for inter-chunk continuity

Description: The authors propose a prefix-preserved sampling procedure that initializes action generation using previously executed actions as priors and preserves the overlapping segment between consecutive chunks during sampling. This method enhances inter-chunk continuity by maintaining coherence across chunk boundaries during asynchronous execution.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Training-Time Action Conditioning for Efficient Real-Time Chunking

URL: View paper

Prior Art Analysis

Training-Time Action Conditioning[41] demonstrates that similar prior work exists for maintaining inter-chunk continuity through prefix-based conditioning. Both papers address the same fundamental problem of using previously executed actions as priors during action generation. The candidate paper explicitly describes conditioning on action prefixes to enhance continuity, which directly overlaps with the original paper's prefix-preserved sampling approach. The candidate achieves this through training-time simulation rather than test-time sampling, but the core concept of leveraging executed action prefixes for inter-chunk coherence is substantively similar.

Evidence

Evidence 1 - **Rationale**: Both papers explicitly describe using previously executed/committed actions as conditioning information to improve inter-chunk continuity in asynchronous action chunking systems. - **Original**: we introduce a prefix-preserved sampling procedure to reinforce inter-chunk continuity - **Candidate**: rtc enables vision-language-action models (vlas) to generate smooth, reactive robot trajectories by asynchronously predicting action chunks and conditioning on previously committed actions via inference-time inpainting

Evidence 2 - **Rationale**: The candidate paper's approach of conditioning on action prefixes directly addresses the same technical challenge as the original paper's use of executed actions as priors, demonstrating prior work on this concept. - **Original**: we enhance inter-chunk continuity by refining the sampling pipeline to incorporate previously executed actions as informative priors - **Candidate**: we propose a simple alternative: simulating inference delay at training time and conditioning on action prefixes directly

Evidence 3 - **Rationale**: While the original paper describes a specific sampling initialization procedure, the candidate demonstrates that prefix-based conditioning for inter-chunk continuity was already being explored as a practical solution, suggesting this is not a novel contribution. - **Original**: the initial action state a0 t is no longer drawn from the gaussian prior. instead, it is initialized using the executed actions during the delayed time steps, denoted as ap t ∈ rpxd, whose first p - h entries are filled with the last p - h actions from the previous predicted chunk - **Candidate**: our method requires no modifications to the model architecture or robot runtime, and can be implemented with only a few additional lines of code

---

2. LLM Inference Scheduling: A Survey of Techniques, Frameworks, and Trade-offs

URL: View paper

Brief Assessment

LLM Inference Scheduling[42] discusses prompt chunking and prefix-based schemes in the context of LLM inference optimization, not robotic action chunking or maintaining continuity across action sequences in physical control systems.

3. From Logs to Insights: Exploring User Behavior in RobotStudio

URL: View paper

Brief Assessment

RobotStudio User Behavior[43] focuses on analyzing user interaction logs in ABB RobotStudio software using n-gram analysis, Markov chains, and clustering. It does not address action chunking, robotic control policies, or prefix-preserved sampling procedures for maintaining inter-chunk continuity in robot execution.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Real-Time Robot Execution with Masked Action Chunking View paper
- [1] VLASH: Real-Time VLAs via Future-State-Aware Asynchronous Inference View paper
- [2] Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates View paper
- [3] Real-Time Execution of Action Chunking Flow Policies View paper
- [4] Rt-h: Action hierarchies using language View paper
- [5] Observe Then Act: Asynchronous Active Vision-Action Model for Robotic Manipulation View paper
- [6] Delayed-Decision Motion Planning in the Presence of Multiple Predictions View paper
- [7] Leave no observation behind: Real-time correction for vla action chunks View paper
- [8] Red: A systematic real-time scheduling approach for robotic environmental dynamics View paper
- [9] Deep Reinforcement Learning based Real-Time Path Planning and Flight Validation of small UAS Application View paper
- [10] Modelling and Model-Checking a ROS2 Multi-Robot System using Timed Rebeca View paper
- [11] ASAP-MPC: an asynchronous update scheme for online motion planning with nonlinear model predictive control View paper
- [12] The xbot2 real-time middleware for robotics View paper
- [13] Spatio-Temporal Avoidance of Predicted Occupancy in Human-Robot Collaboration View paper
- [14] Delayed Expansion AGT: Kinodynamic Planning with Application to Tractor-Trailer Parking View paper
- [15] Non-blocking asynchronous training for reinforcement learning in real-world environments View paper
- [16] Performance Analysis of YOLOv5 for Real-Time Object Detection and Manipulation in Unity on CPU and GPU View paper
- [17] Event-based Vision for Early Prediction of Manipulation Actions View paper
- [18] Built-in Analog Automatic Controls for Small Robots View paper
- [19] Occlusion-aware trajectory prediction for time-delayed teleoperation of space manipulators View paper
- [20] RT-HCP: Dealing with Inference Delays and Sample Efficiency to Learn Directly on Robotic Platforms View paper
- [21] A structural approach to the non-blocking supervisory control of discrete-event systems View paper
- [22] Human-in-the-loop Online Rejection Sampling for Robotic Manipulation View paper
- [23] Asynchronous Fast-Slow Vision-Language-Action Policies for Whole-Body Robotic Manipulation View paper
- [24] Population-Coded Spiking Neural Networks for High-Dimensional Robotic Control View paper
- [25] Incremental Model Predictive Control Exploiting Time-Delay Estimation for a Robot Manipulator View paper
- [26] Asynchronous Real-Time Optimization of Footstep Placement and Timing in Bipedal Walking Robots View paper
- [27] Asynchronous Real-time Decentralized Multi-Robot Trajectory Planning View paper
- [28] A real-time distributed software infrastructure for cooperating mobile autonomous robots View paper
- [29] Asynchronous production system for control of an autonomous mobile robot in real-time environment View paper
- [30] Asynchronous Reinforcement Learning for Real-Time Control of Physical Robots View paper
- [31] Probabilistic Articulated Real-Time Tracking for Robot Manipulation View paper
- [32] ADA Implementation of Concurrent Execution of Multiple Tasks in the Strategic and Tactical Levels of the Rational Behavior Model for the NPS Phoenix Autonomous Underwater Vehicle (AUV). View paper
- [33] A Real-Time Model-Based Reinforcement Learning Architecture for Robot Control View paper
- [34] : a VLA That Learns From Experience View paper
- [35] AsyncVLA: Asynchronous Flow Matching for Vision-Language-Action Models View paper
- [36] A Survey on Reinforcement Learning of Vision-Language-Action Models for Robotic Manipulation View paper
- [37] Discrete Diffusion VLA: Bringing Discrete Diffusion to Action Decoding in Vision-Language-Action Policies View paper
- [38] SmolVLA: A Vision-Language-Action Model for Affordable and Efficient Robotics View paper
- [39] ImplicitRDP: An End-to-End Visual-Force Diffusion Policy with Structural Slow-Fast Learning View paper
- [40] Mobile robot programming using natural language View paper
- [41] Training-Time Action Conditioning for Efficient Real-Time Chunking View paper
- [42] LLM Inference Scheduling: A Survey of Techniques, Frameworks, and Trade-offs View paper
- [43] From Logs to Insights: Exploring User Behavior in RobotStudio View paper