

Master–Micro Residual Correction with Adaptive Tactile Fusion and Force-Mixed Control for Contact-Rich Manipulation

Xingting Li^{1,*}, Yifan Xie^{1,*}, Han Liu^{1,*}, Wei Hou¹, Guangyu Chen¹, Shoujie Li^{2,†}, and Wenbo Ding^{1,†}

Abstract—Robotic contact-rich and fine-grained manipulation remains a significant challenge due to complex interaction dynamics and the competing requirements of multi-timescale control. While current visual imitation learning methods excel at long-horizon planning, they often fail to perceive critical interaction cues like friction variations or incipient slip, and struggle to balance global task coherence with local reactive feedback. To address these challenges, we present M^2 -ResiPolicy, a Master-Micro residual control architecture that synergizes high-level action guidance with low-level correction. The framework consists of a Master-Guidance Policy (MGP) operating at 10 Hz, which generates temporally consistent action chunks via a diffusion-based backbone and employs a binary contact-gated cross-attention mechanism to modulate visual–tactile fusion. Simultaneously, the Micro-Residual Corrector (MRC) uses TCP wrench feedback sampled at 60 Hz and a lightweight GRU to provide residual action compensation. This policy is further integrated with a fixed, non-learned force-mixed Position-Based Hybrid Impedance Control (PBIC) execution layer, which is used as an execution interface rather than a learned controller, providing a physically interpretable safety envelope for stable interaction. Experiments across several demanding tasks including fragile object grasping and precision insertion show that M^2 -ResiPolicy achieves higher average success rates than diffusion baselines, including a 93% damage-free success rate in chip transfer, with more stable force-regulation profiles.

Index Terms—Contact-Rich manipulation, Visuotactile fusion, Residual policy learning.

I. INTRODUCTION

ROBOTIC contact-rich and fine-grained manipulation [1], [2], [3], [4], such as peeling vegetables or handling fragile objects, remains a significant challenge due to complex and partially observable interaction dynamics. Humans achieve such dexterity by naturally integrating vision and touch [5]: vision provides global scene context and task understanding, while tactile feedback supplies localized, contact-specific information essential for fine-grained control. Despite substantial

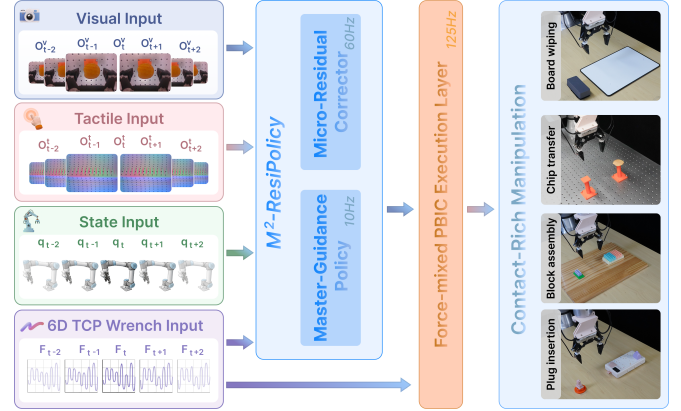


Fig. 1. M^2 -ResiPolicy overview. Multimodal inputs are processed by MGP/MRC before fixed PBIC execution in contact-rich tasks.

progress in long-horizon visual imitation learning (IL) [6], [7], vision-only policies often fail to capture frictional variations or incipient slip, limiting precision and stability in unstructured environments.

To overcome these limitations, prior work explores various visuo-tactile fusion strategies. Reactive Diffusion Policy [2] adopts a slow–fast hierarchical control paradigm that combines diffusion-based action chunking with high-frequency tactile feedback, while ViTacFormer [8] learns unified visuo–tactile representations via cross-attention and autoregressive tactile prediction for robust long-horizon manipulation. Nevertheless, existing hierarchical or temporally structured approaches still face a fundamental challenge: reconciling task-level temporal coherence with rapid, contact-sensitive force feedback. In particular, how perceptual focus should be dynamically modulated across contact phases and how physical safety can be systematically enforced at the execution layer remain largely unexplored.

To address these challenges, we propose M^2 -ResiPolicy, a Master–Micro residual control framework for contact-rich manipulation that couples multi-timescale policy learning with compliant execution. At the high level, a *Master-Guidance Policy* (MGP) uses contact-gated visuotactile fusion to generate temporally coherent action chunks from multimodal observations. At the low level, a high-frequency *Micro-Residual Corrector* (MRC) predicts residual compensation from TCP wrench feedback for rapid local refinement under transient contact. During execution, the policy output is further passed through a fixed, non-learned force-mixed *Position-Based Hybrid Impedance Control* (PBIC) layer, which converts motion references into compliant pose commands and provides a

Manuscript received: March 14, 2026; Revised: June 30, 2026; Accepted: September 17, 2026. This paper was recommended for publication by Editor Lorenzo Jamone upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the National Key R&D Program of China (Grant No. 2024YFB3816000), the Guangdong Innovative and Entrepreneurial Research Team Program (2021ZT09L197), the Tsinghua Shenzhen International Graduate School–Shenzhen Pengrui Young Faculty Program of Shenzhen Pengrui Foundation (No. SZPR2023005), and Meituan. (†Corresponding authors: Shoujie Li and Wenbo Ding.)

*Xingting Li, Yifan Xie, and Han Liu contributed equally to this work.

¹Xingting Li, Yifan Xie, Han Liu, Wei Hou, Guangyu Chen, and Wenbo Ding are with the Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China (e-mail: ding.wenbo@sz.tsinghua.edu.cn).

²Shoujie Li is with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639956, Singapore (e-mail: lsj20@mails.tsinghua.edu.cn).

Digital Object Identifier (DOI): see top of this page.

physically interpretable safety envelope for stable interaction. Our contributions are threefold:

- (1) A contact-gated visuotactile cross-attention fusion mechanism that adaptively modulates modality contributions across contact phases;
- (2) A high-frequency micro-residual correction module that uses TCP force/torque feedback for rapid local refinement under transient contact;
- (3) A fixed PBIC execution interface that converts learned MGP/MRC references into force-protected pose commands.

II. RELATED WORK

A. Learning Perception for Contact: Tactile and Force Representations

Modeling physical interaction requires inferring latent contact states, including contact location, friction, and incipient slip, that vision alone cannot directly observe. This motivates multimodal sensing: vision provides global context, force/torque captures interaction intensity, and tactile sensing provides localized contact structure. **Vision-based tactile sensing** supports adaptive multimodal fusion, representation learning, and dual-rate closed-loop control [4], [9], [2], [10], where slow and fast loops couple task-level progression with rapid local adjustment. These tactile streams help reveal contact onset and local slip beyond geometric observations.

In parallel, **proprioceptive force/torque (F/T) sensing** offers a low-dimensional yet highly responsive signal for contact transitions and interaction directions. Force has been incorporated into imitation learning as an explicit control target [11], [12], a force-aware training signal [13], [1], or a condition for generative policies [14], [15], [16]. Despite these advances, many frameworks still use single-modality inputs or static fusion, limiting their ability to adjust perceptual emphasis across contact phases. This leaves open how to shift from pre-contact visual guidance to contact-phase tactile/force feedback without destabilizing execution.

B. Multimodal Policy Learning for Contact-Rich Manipulation

Recent work increasingly adopts **end-to-end multimodal policy learning**, directly mapping vision, tactile, and force observations to control commands. Early approaches combine vision and touch in reinforcement learning [17], while later work uses large-scale visuotactile data or decomposes manipulation into reaching and contact-centric stages [18], [19], [20]. However, many methods remain pose-centric at the execution level, issuing position or trajectory commands without explicitly modeling interaction compliance. Recent work therefore integrates impedance- or compliance-aware controllers into low-level execution for safer contact-rich manipulation [16], [21].

Beyond deterministic regression, **conditional generative policies** model action distributions conditioned on perceptual context. Flow- and diffusion-based methods enable long-horizon, multi-modal action generation [22], [23], [24]. A key

challenge remains the mismatch between temporal abstraction and control bandwidth: long-horizon policies smooth rapid contact variations, while highly reactive controllers sacrifice long-term coherence, motivating two-time-scale or hierarchical designs. **M²-ResiPolicy** addresses this mismatch by coupling contact-gated visuotactile fusion, high-rate wrench-driven residual correction, and a fixed, non-learned PBIC execution interface.

III. METHOD

Contact-rich manipulation imposes competing multi-timescale requirements: the policy must maintain temporally consistent motion over long horizons, while remaining sufficiently reactive to rapid contact variations such as impacts and abrupt friction changes. To address this challenge, we propose **M²-ResiPolicy**, a Master-Micro framework illustrated in Fig. 2. At the high level, the *Master-Guidance Policy* (MGP) predicts an H -step, 7D end-effector action chunk from multimodal observations at 10 Hz, providing a smooth global motion prior. At the low level, the *Micro-Residual Corrector* (MRC) processes 6D TCP wrench feedback sampled at 60 Hz and predicts residual corrections aligned with the master action steps to refine execution under fast contact dynamics.

During deployment, the policy output is interpreted as a motion-reference pose x_{ref} . To improve robustness under modeling errors and out-of-distribution contacts, we execute this reference through a position-based hybrid impedance controller, which combines an impedance-derived force intention with the measured external wrench to generate a compliant pose command $x_{\text{cmd}} = x_{\text{ref}} + \Delta x_{\text{comp}}$. This execution layer preserves accurate free-space tracking while introducing compliance upon contact, thereby improving stability and repeatability in contact-rich interaction.

A. Data Collection and System Setup

As shown in Fig. 3, our data-collection platform consists of a UR7e manipulator teleoperated via a GELLO device [25], which maps the operator's end-effector pose intent to robot motion. Visual observations are provided by two RealSense D435 cameras, including a third-person view for global context and a wrist-mounted view for local details. Tactile observations are captured by two fingertip-mounted XENSE-G1 visuotactile sensors. Instead of using raw tactile images, we adopt frame-differenced tactile representations to improve signal consistency and stability. Each sensor provides six-axis force/torque estimates via the Xense SDK interface using manufacturer-provided calibration. These estimates gate MGP, while MRC and PBIC use separately acquired UR7e TCP wrench feedback. All modalities are timestamped and segmented into episodes.

To support cross-rate learning, we record multimodal observations and end-effector action labels at 10 Hz, while logging TCP wrench measurements at 60 Hz. During offline pre-processing, the high-rate wrench stream is timestamp-aligned with the action steps so that each step is associated with the corresponding force/torque information while preserving temporal consistency. Demonstrations are collected under

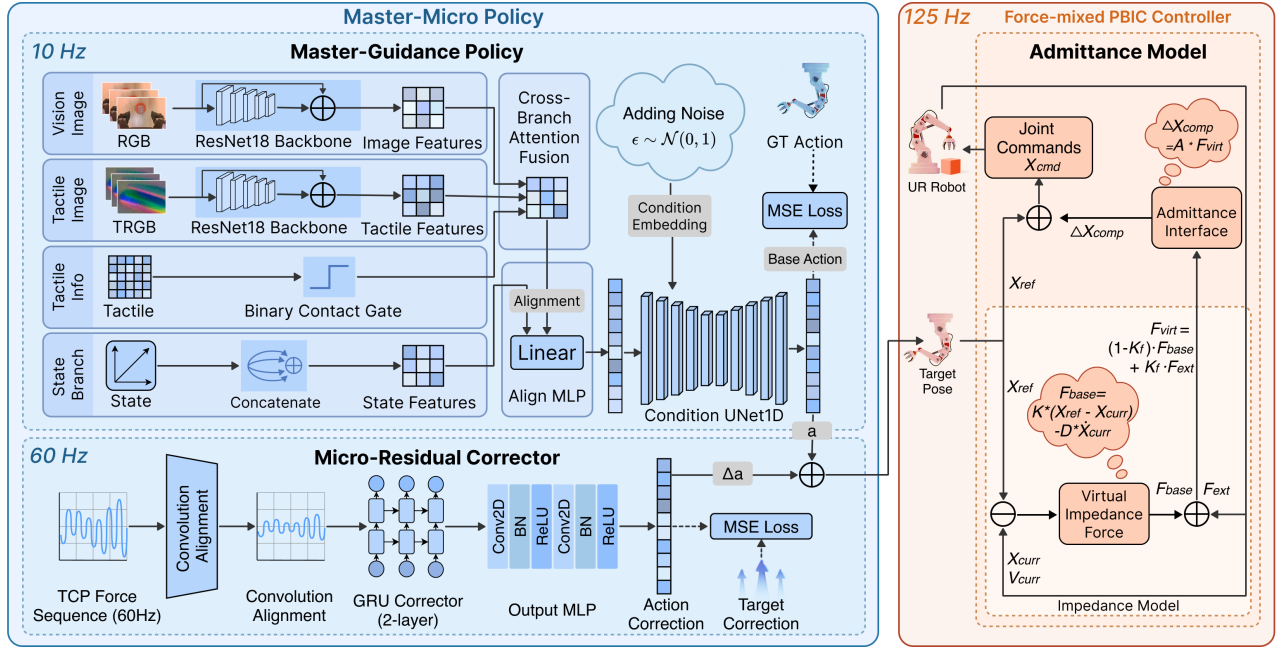


Fig. 2. **Overview of M²-ResiPolicy.** *Left:* MGP generates contact-gated action chunks and MRC predicts wrench-based residuals. *Right:* Fixed PBIC maps x_{ref} to x_{cmd} ; blue/orange blocks denote learned/fixed modules.

Position-Based Hybrid Impedance Control, with task-specific stiffness, damping, and force-mixing gains tuned according to contact intensity and desired compliance. As a result, the dataset captures not only geometric motion intent but also stable interaction regularities induced by compliant regulation, enabling the learned policy to reproduce safe and compliant behaviors when deployed through the same execution layer.

B. Contact-Gated Cross-Attention Fusion in the Master-Guidance Policy

Master-Guidance Policy (MGP) runs at 10 Hz and predicts temporally coherent end-effector action chunks from multi-modal observations, providing a smooth global motion prior for downstream execution. Let $\mathbf{o}_t$ denote the observation at time t , including third-person and wrist-view RGB images, two fingertip XENSE-G1 visuotactile observations, the robot proprioceptive state, and fingertip tactile readings for contact gating. The policy outputs an H -step 7D action chunk $\mathbf{a}_{t:t+H-1} \in \mathbb{R}^{H \times 7}$, modeled in a normalized action space during training.

As shown in Fig. 2 (top left), the RGB streams are encoded into visual tokens and the fingertip visuotactile streams into tactile tokens, while the proprioceptive state is retained as a low-dimensional auxiliary feature. Instead of applying generic self-attention over all modalities, MGP adopts a contact-gated cross-attention module, where visual tokens act as the primary geometric anchor and tactile cues are incorporated according to the current contact state.

Let $\mathbf{F}_{v,t} \in \mathbb{R}^{N_r \times D}$ and $\mathbf{F}_{tac,t} \in \mathbb{R}^{N_r \times D}$ denote the visual and tactile feature tokens at time t , respectively, and let $\mathbf{f}_t^{tac}$ denote the concatenated fingertip tactile readings. In our implementation, $\mathbf{f}_t^{tac} = [\mathbf{f}_t^{L,tac}, \mathbf{f}_t^{R,tac}] \in \mathbb{R}^{12}$, where the left and right XENSE-G1 fingertip six-axis readings are

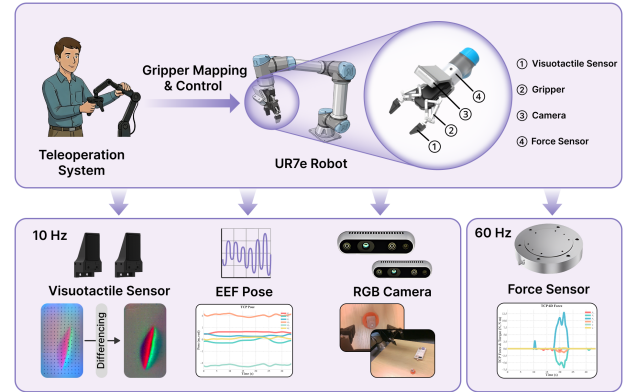


Fig. 3. **System setup.** UR7e teleoperation records RGB, visuotactile, pose, and TCP wrench streams; fingertip tactile readings gate MGP, while TCP wrench feeds MRC/PBIC.

concatenated. A binary contact gate is first derived from the tactile readings as

$$g_t = \mathbb{I}(\|\mathbf{f}_t^{tac}\|_2 > \tau_{tac}), \quad (1)$$

where τ_{tac} is a tactile threshold fixed across tasks and trials. We set $\tau_{tac} = 0.05$ for all tasks, methods, and trials. This threshold is an empirical threshold in the same preprocessed sensor scale used by the policy input and is used only to determine whether contact occurs; the 12D norm is not interpreted as a physical wrench magnitude.

The visual and tactile tokens are then concatenated into a joint key-value pool, while the visual tokens are used as queries [26]. Specifically, with $\mathbf{Q}_t = \mathbf{F}_{v,t} \mathbf{W}_Q$, $\mathbf{K}_t = [\mathbf{F}_{v,t}; \mathbf{F}_{tac,t}] \mathbf{W}_K$, and $\mathbf{V}_t = [\mathbf{F}_{v,t}; \mathbf{F}_{tac,t}] \mathbf{W}_V$, the pre-softmax score tensor is multiplicatively modulated as

$$\hat{\mathbf{S}}_t = \left(\frac{\mathbf{Q}_t \mathbf{K}_t^\top}{\sqrt{d_h}} \right) \odot \mathbf{M}_t. \quad (2)$$

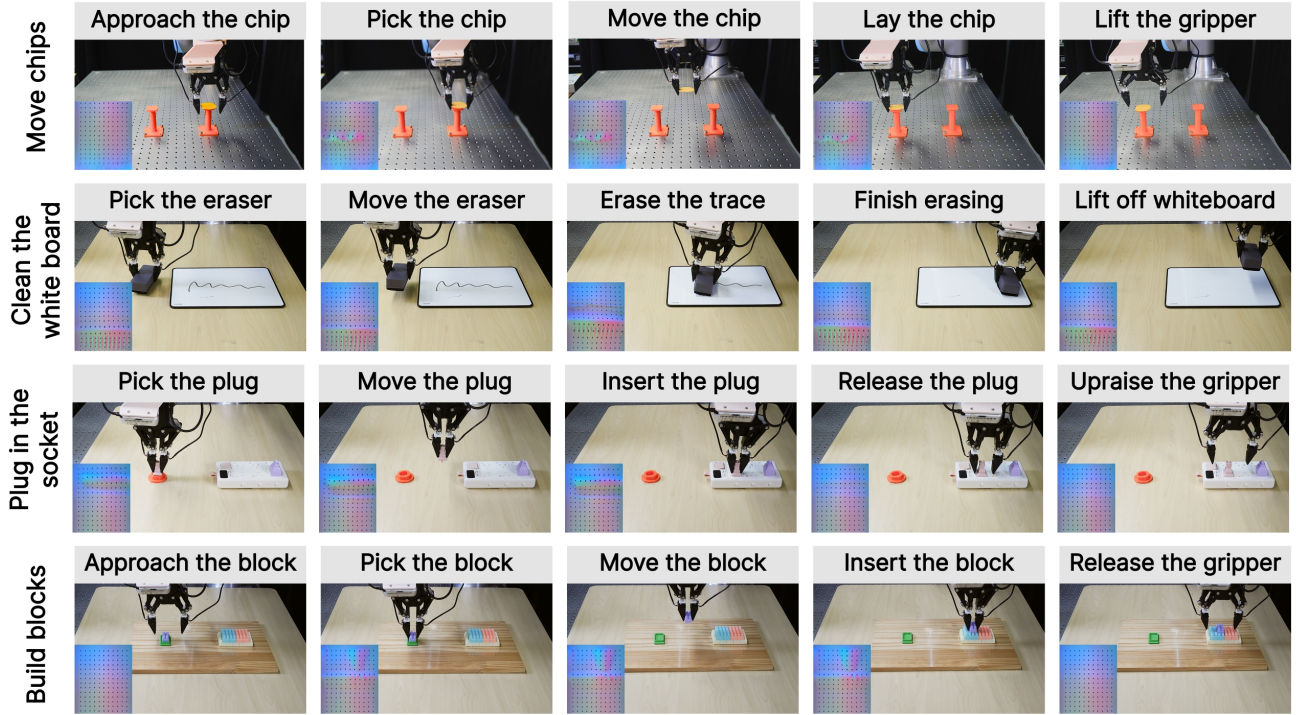


Fig. 4. **Evaluation tasks.** Four benchmarks from top to bottom: *Chip Transfer*, *Whiteboard Wiping*, *Plug Insertion* and *Block Assembly*. Insets show fingertip visuotactile observations.

Here, $\mathbf{M}_t$ is a logit-modulation mask constructed from the binary contact gate. It preserves the visual-key logits and multiplies the tactile-key logits by g_t before softmax. When $g_t = 0$, the attention scores corresponding to tactile keys are set to 0 before softmax, so the tactile branch no longer expresses contact-dependent score variations. However, these entries still participate in softmax normalization, and the operation modulates tactile logits rather than directly setting tactile attention weights to zero. When $g_t = 1$, tactile logits are unchanged. For $\mathbf{S}_t \in \mathbb{R}^{B \times H_a \times N_r \times (N_r + N_\tau)}$, $\mathbf{M}_t \in \mathbb{R}^{B \times 1 \times 1 \times (N_r + N_\tau)}$ has visual-key entries of 1 and tactile-key entries of g_t , and is broadcast across attention heads and visual queries. Thus, g_t denotes a binary contact gate rather than a probabilistic confidence score or an attention weight. The fused feature is then obtained by

$$\mathbf{F}_t^{\text{fused}} = \mathbf{F}_{v,t} + \text{softmax}(\hat{\mathbf{S}}_t) \mathbf{V}_t \mathbf{W}_O. \quad (3)$$

In implementation, MHA uses $H_a = 8$ attention heads with $D = 512$ -dimensional features, corresponding to $d_h = 64$ per head. In the current network, $N_r = 2$ denotes two encoded RGB feature tokens corresponding to the third-person and wrist cameras, and $N_\tau = 2$ denotes two encoded tactile feature tokens corresponding to the left and right fingertip tactile image streams. We use $T_o = 4$ observation steps and attention dropout 0.1. This design preserves vision as the global reference while multiplicatively modulating tactile attention logits according to contact state. The fused features are finally flattened, concatenated with the low-dimensional proprioceptive features, and mapped by a linear alignment layer to obtain the conditioning context $\mathbf{c}_t$.

The diffusion denoiser is trained with a conditional noise-prediction objective [27]. Let $\tilde{\mathbf{a}}$ denote the normalized ex-

pert action chunk. We sample a diffusion step k and Gaussian noise $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and construct the noisy action as $\tilde{\mathbf{a}}_k = \sqrt{\alpha_k} \tilde{\mathbf{a}} + \sqrt{1 - \alpha_k} \epsilon$. The training objective is $\mathcal{L}_{\text{MGP}} = \|\pi_{\text{MGP}}(\tilde{\mathbf{a}}_k, k, \mathbf{c}_t) - \epsilon\|_2^2$. At inference time, the same conditioning pathway guides iterative denoising to generate an H -step action chunk for subsequent compliant execution.

C. Reactive Micro-Residual Correction and Force-Mixed Compliant Execution

The *Micro-Residual Corrector* (MRC) is designed to capture high-frequency contact dynamics and suppress local execution deviations, such as those induced by impacts, abrupt friction transitions, and incipient slip. MRC operates in conjunction with the master action chunk predicted by the *Master-Guidance Policy* (MGP). The system continuously acquires 6D TCP force/torque (wrench) measurements at 60 Hz and compresses this high-rate data stream into a force-feature sequence aligned with the master action steps through strided temporal convolution. This ensures that each action step is paired with a compact summary of recent contact feedback. On top of these step-aligned features, MRC utilizes a lightweight GRU to model temporal dependencies, accumulating short-horizon contact history in its hidden state and predicting residual corrections aligned with the master action. This recurrent approach enables high-frequency contact information to be injected into residual predictions. To ensure consistency between training and inference, residuals are defined in the normalized action space, and the supervised target is constructed relative to a frozen MGP baseline: $\Delta \tilde{\mathbf{a}}^* = \tilde{\mathbf{a}}^{\text{exp}} - \tilde{\mathbf{a}}^{\text{MGP}}$. This formulation encourages MRC to learn local corrections around the master baseline, thereby improving the stability and transferability of contact-rich closed-loop execution.

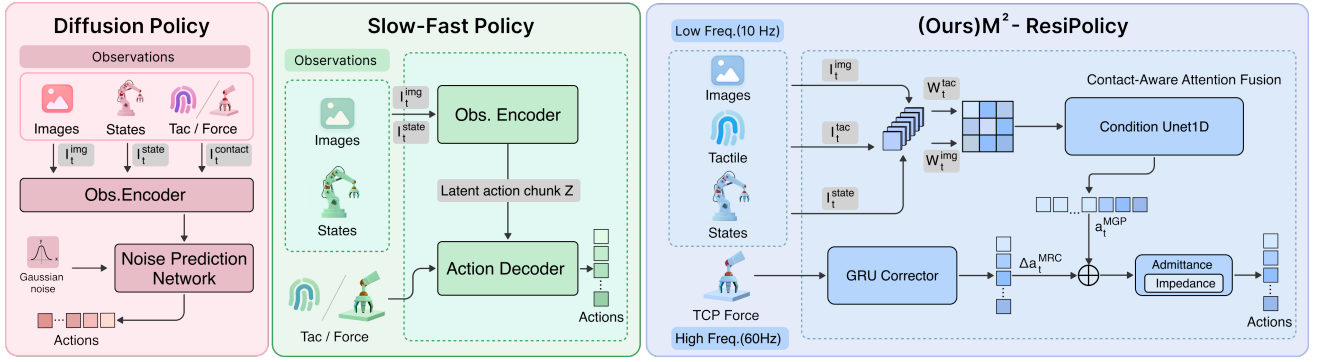


Fig. 5. **Baseline Architectures.** We compare our hierarchical MGP+MRC framework against open-loop Diffusion Policy (DP), vision-force diffusion policy (DP w. Force-CrossAttn), and end-to-end Reactive Diffusion Policy (RDP) baselines.

TABLE I
MAIN TASK SUCCESS RATES

Method	Chip Transfer			Surface Wiping			Plug Insertion			Block Assembly			Score
	Pick	Move	Place	Pick	Contact	Wipe	Grasp	Align	Insert	Grasp	Align	Stack	
DP (Baseline)	43%	40%	37%	80%	67%	63%	70%	63%	43%	57%	53%	43%	0.55
DP w. tactile	63%	57%	53%	83%	73%	70%	77%	70%	53%	67%	60%	50%	0.65
DP w. Force-CrossAttn	57%	57%	50%	87%	83%	80%	80%	73%	53%	73%	67%	50%	0.68
RDP (Tactile)	90%	87%	83%	97%	93%	90%	90%	87%	73%	93%	87%	77%	0.87
RDP (Force)	70%	63%	60%	97%	93%	93%	87%	87%	80%	87%	83%	80%	0.82
M²-ResiPolicy	100%	97%	93%	97%	97%	97%	93%	93%	80%	97%	93%	87%	0.94

During execution, a *force-mixed* position-based impedance controller (PBIC) is integrated as the fixed, non-learned execution layer to convert the policy's motion reference into a compliant end-effector pose command. Specifically, PBIC follows classical impedance/admittance control principles [28], [29]: it constructs a nominal virtual spring-damper wrench from the tracking error, incorporates measured TCP wrench feedback, and blends the two to compute a compliance offset, yielding $x_{\text{cmd}} = x_{\text{ref}} + \Delta x_{\text{comp}}$. The gains in this execution layer are empirically selected and kept fixed during training and evaluation. The resulting command is streamed to the position servo at 125 Hz. This execution layer provides a physically interpretable safety envelope and a compliant interface for learned MGP/MRC references: it tracks accurately in free space and injects force-closed-loop compliance upon contact.

D. Staged Training Strategy

We adopt a staged training scheme to decouple master action generation from micro-level residual correction. The master policy is first trained to produce reliable action chunks and define the action-normalization mapping. The master branch is then frozen, and the GRU-based residual module is trained to predict local corrections from the 6D end-effector force/torque signal around this fixed baseline. This training strategy reduces optimization interference and improves rapid contact adaptation while preserving long-horizon motion consistency.

IV. EXPERIMENTS

This section presents a systematic evaluation of the M²-ResiPolicy framework, designed to address the competing

multi-timescale requirements of contact-rich manipulation. Our experiments are structured to evaluate the three core contributions: the slow-fast hierarchical architecture, the tactile-force-adaptive fusion, and the impedance-controlled execution. We specifically address the following Research Questions (RQs):

RQ1 (Perception Effectiveness): Can the binary contact-gated fusion mechanism regulate the contribution of multi-modal information across different contact phases (e.g., free space versus contact)?

RQ2 (Residual Correction): Does the high-frequency *GRU-Based Micro-Residual Corrector (MRC)* achieve higher success rates than single-frequency baselines in compensating for contact dynamics and accumulated drift?

RQ3 (Interaction Safety): Does integrating *Position-Based Hybrid Impedance Control* improve safety and stability compared to rigid position control in constrained environments?

A. Experimental Setup

1) **System and Data Collection:** As detailed in Sec. III, our experimental platform comprises a UR7e manipulator teleoperated via a GELLO device. Perception is provided by two RealSense D435 cameras (global and wrist views) and two fingertip-mounted XENSE-G1 visuotactile sensors. To support our cross-frequency architecture, we record multimodal observations and master actions at 10 Hz, while synchronously logging high-fidelity TCP wrench data at 60 Hz for residual learning.

2) **Task Definitions:** To comprehensively evaluate policy performance under varying contact conditions, we select two categories of representative tasks, as visualized in Fig. 4:

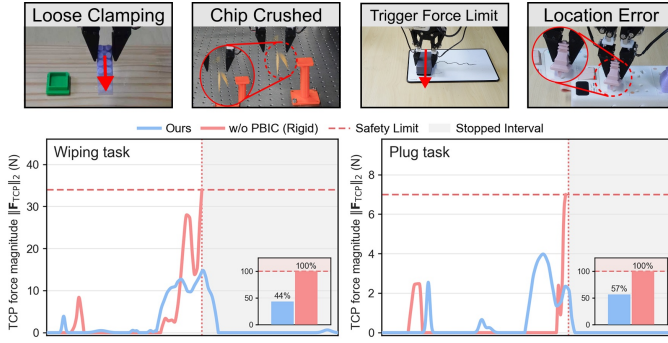


Fig. 6. **Failure analysis.** (Top) Representative baseline failures. (Bottom) Force profiles; red/blue denote baseline/ours. The dashed safety-limit line denotes the TCP force-threshold stop used for hardware protection. Insets show peak TCP force as a percentage of the task-specific safety threshold; 100% indicates the protective-stop threshold.

Category I: Force-Sensitive Manipulation. Includes *Fragile Chip Transfer* and *Surface Wiping*. The primary challenge here is precise force regulation to prevent object damage or excessive contact pressure.

Category II: Geometry-Constrained Assembly. Includes *Dual-Peg Insertion* and *Block Assembly*. These tasks involve millimeter-level tolerances and severe visual occlusion, requiring agile local correction to resolve geometric conflicts.

3) **Baselines:** We compare our method against Diffusion Policy (DP), DP w. tactile, DP w. Force-CrossAttn adapted from the open-source vision–force diffusion policy of Kang *et al.* [15], Reactive Diffusion Policy (RDP), and three ablations: Ours (w/o Adaptive Fusion), Ours (w/o MRC), and Ours (Rigid Control). Our physical-baseline selection follows a same-platform criterion, preserving the robot platform, action space, control interface, and evaluation protocol whenever possible. Although FoAR [1] and FILIC [21] are relevant, FoAR requires RGB-D point-cloud perception and a 3D encoder, while FILIC relies on torque-level impedance control; both would change our UR7e end-effector pose/position-servo experimental interface, and are therefore not used as direct physical-reproduction baselines.

4) **Evaluation Protocols:** Each task is evaluated over 30 independent trials per method. For methods with the impedance-based execution layer, the stiffness and damping parameters are tuned per task and kept fixed across all trials. Safety-stopped trials are counted as failures, and the safety-limit marker denotes a TCP force-threshold protective stop. The monitored scalar safety quantity is the translational TCP force norm $\|\mathbf{F}_{\text{TCP}}\|_2 = \sqrt{F_x^2 + F_y^2 + F_z^2}$. Task thresholds are fixed at 3.0/34.0/7.0/18.0 N for Chip/Wiping/Plug/Block, and crossing them triggers a protective stop. In addition to binary success rates, we analyze failure modes and force profiles to provide a mechanistic understanding of policy behavior.

B. Main Results

We summarize the quantitative success rates of complete deployed pipelines across all tasks in Table I and analyze the underlying failure mechanisms using the qualitative visualization in Fig. 6. Based on these results, we address the three core research questions.

TABLE II
ABLATION ANALYSIS: SUCCESS RATE DROPS

Variants	Chip Transfer	Surface Wiping	Plug Insertion	Block Assembly
Full Model (Ours)	93%	97%	80%	87%
w/o Adaptive Fusion	80% ($\downarrow_{13}$)	93% ($\downarrow_{04}$)	73% ($\downarrow_{07}$)	77% ($\downarrow_{10}$)
w/o MRC (Residual)	90% ($\downarrow_{03}$)	93% ($\downarrow_{04}$)	67% ($\downarrow_{13}$)	70% ($\downarrow_{17}$)
w/o Impedance (Rigid)	87% ($\downarrow_{06}$)	90% ($\downarrow_{07}$)	70% ($\downarrow_{10}$)	77% ($\downarrow_{10}$)

Addressing RQ1 (Perception Effectiveness): Experimental results suggest that the *binary contact-gated visuotactile fusion* mechanism contributes to regulating multimodal feature contributions across contact phases. As shown in Table I, the baseline methods exhibit lower success rates during the grasp-related stages of challenging tasks. As qualitatively illustrated in Fig. 6 (Bottom, Row 1), the open-loop baseline fails to adapt to varying contact conditions, resulting in crushing of fragile chips due to excessive closure (Left) and slippage of small objects due to insufficient contact regulation (Right). In contrast, M²-ResiPolicy maintains higher grasp success rates across these settings. By multiplicatively modulating tactile logits with the binary contact gate, our method balances the sensitivity required to avoid over-compression with the stability needed to prevent slip.

Addressing RQ2 (Residual Correction): Results on precision-critical tasks indicate the utility of the *Micro-Residual Corrector (MRC)*. As shown in Table I, single-frequency baselines exhibit lower success rates during the execution stages of *Insertion* and *Assembly*. Without high-frequency residual correction, these methods struggle to compensate for millisecond-scale contact dynamics, leading to the **geometric jamming** shown in Fig. 6 (Bottom, Row 2, Right). In contrast, with the GRU-based residual module using TCP wrench feedback sampled at 60 Hz, M²-ResiPolicy maintains higher execution success rates, indicating that the proposed slow–fast architecture helps mitigate accumulated drift and supports more robust contact-rich execution. We therefore interpret MRC as a learned policy-level residual corrector around the MGP reference, not as a replacement for the execution-level compliant controller.

Addressing RQ3 (Interaction Safety): Results on constrained contact tasks indicate that the proposed fixed PBIC execution layer provides a useful safety layer for physical interaction. As shown in Fig. 6, rigid control can lead to unsafe behaviors in both qualitative failures and force profiles. In the *Wiping* task, where sustained contact is required, high position stiffness can produce abnormally large normal-force peaks and trigger a TCP force-threshold safety stop (Bottom, Left). In the *Plug Insertion* task, the baseline exhibits a sharp force spike during insertion, indicating poor tolerance to contact misalignment and constrained interaction (Bottom, Right). In contrast, M²-ResiPolicy leverages compliant execution to keep contact forces within safe operating limits, thereby helping improve task continuity and interaction stability.

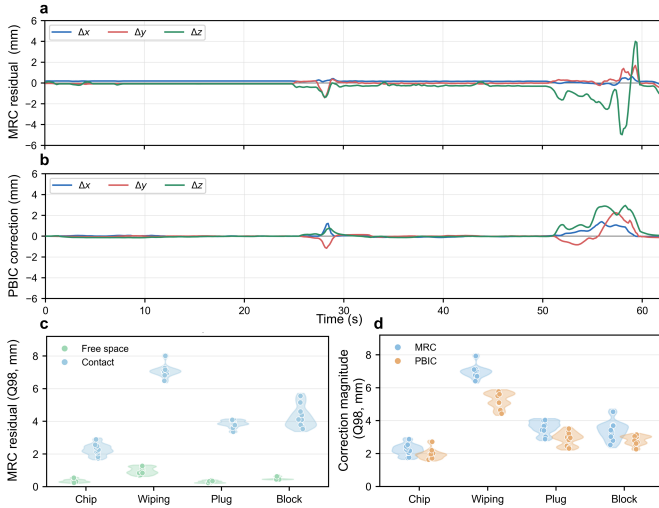


Fig. 7. **Residual-behavior analysis.** (a,b) Successful Block Assembly trajectory examples. (c) MRC residuals in free space and contact. (d) Command-level MRC/PBIC magnitudes. Statistics use trajectory-level Q_{98} ; each point is one successful trajectory.

C. Ablation Studies

To quantify the individual contributions of the core components in M^2 -ResiPolicy, we evaluate three ablation variants. Table II details the success rates in the most challenging phases of each task, illustrating the performance changes caused by removing specific modules.

1) **Impact of Tactile-Force-Adaptive Fusion:** As shown in Table II, removing the adaptive weighting mechanism (*w/o Adaptive Fusion*) is associated with lower success rates across all tasks, with larger drops observed in *Chip Transfer* and *Block Assembly*. In this variant, the dynamic contact-gate tensor is replaced by a constant all-one tensor, so tactile logits are preserved before softmax in all phases. This result supports the role of phase-adaptive modality weighting for balancing contact sensitivity and grasp stability.

2) **Impact of GRU Micro-Residual Correction:** Removing the high-frequency residual module (*w/o MRC*) is associated with lower success rates in *Plug Insertion* and *Block Assembly*, with success rates decreasing by 13% and 17%, respectively. This result suggests that the 10 Hz master policy alone is insufficient for geometry-constrained tasks with millimeter-level tolerances. Fig. 7 further shows that MRC residuals remain small in free space and increase after contact, consistent with policy-level residual correction around the MGP reference. Reducing MRC wrench feedback from 60 Hz to 10 Hz lowers average success from 89.2% to 84.2%, while increasing MRC capacity at fixed 60 Hz does not produce a monotonic gain.

3) **Impact of Impedance Control:** Replacing the compliant execution layer with rigid position servoing (*w/o Impedance*) leads to lower performance, particularly in *Plug Insertion*, *Block Assembly*, and *Surface Wiping*. This result indicates that rigid position control has limited tolerance to contact misalignment and force spikes in sustained or constrained interaction.

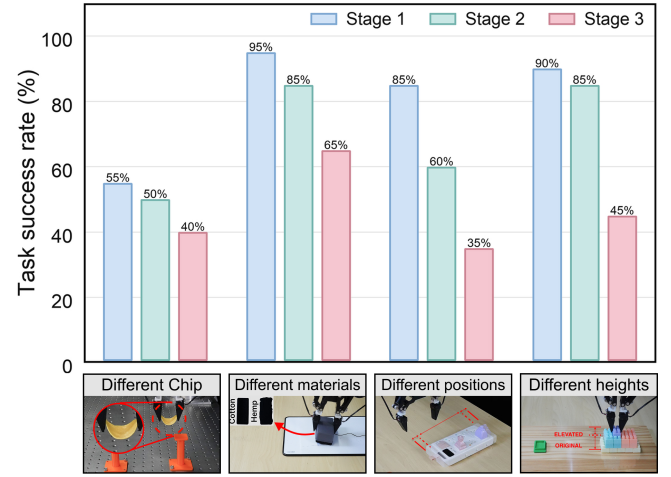


Fig. 8. **Zero-shot generalization under fixed perturbations.** Four unseen test conditions with stage-wise success rates from 20 trials each.

D. Generalization Analysis

To evaluate the behavior of M^2 -ResiPolicy under out-of-distribution conditions, we introduce physical and geometric perturbations that are absent from the training data. The zero-shot performance of the policy is assessed under four perturbation scenarios, as illustrated in Fig. 8. Each perturbation is evaluated over 20 fixed-condition real-robot trials without fine-tuning, and Fig. 8 reports the corresponding stage-wise success rates.

1) Adaptation to Unseen Object Variations:

Stiffness and Shape Variation: For chips with different stiffness and shapes, rigid baselines tend to crush the object due to altered contact-force gradients and load distribution. In contrast, M^2 -ResiPolicy benefits from contact-gated visuotactile fusion in the master policy, which adjusts perceptual emphasis after contact and generates more appropriate grasp actions, thereby reducing the risk of breakage.

Wiping Material Variation: Different wiping materials alter tangential resistance and contact conditions, often leaving visible residue under conventional control. M^2 -ResiPolicy leverages compliant execution to maintain more stable contact.

2) Robustness to Geometric Uncertainty:

Vertical Error (+0.8cm): While rigid policies often trigger protective stops under premature contact, the hybrid impedance layer regulates downward motion by converting position error into a safe normal force.

Lateral Offset (1cm): Under lateral misalignment, the insertion trajectory predicted by the master policy deviates from the true socket position. After initial contact, asymmetric contact cues are captured by the contact-gated visuotactile fusion, which triggers the residual policy to generate local corrections (Δa) and steer the peg back toward the correct insertion path. The lower success rate indicates that residual correction can support adaptation to unseen scenarios but does not fully eliminate sensitivity.

V. CONCLUSIONS

We presented M^2 -ResiPolicy, a Master-Micro residual control framework for contact-rich and fine-grained manipulation.

The proposed framework combines contact-gated visuotactile action generation in the master policy, high-rate residual correction from TCP wrench feedback, and a fixed, non-learned force-mixed PBIC execution layer as a compliant execution interface. In this way, it unifies long-horizon motion consistency with rapid contact adaptation and provides a physically grounded execution mechanism for safe interaction under contact. Experiments on multiple challenging tasks show that $M^2\text{-ResiPolicy}$ achieves higher average performance than diffusion baselines in both task success rate and force-regulation stability, supporting the benefit of integrating dynamic contact-gated perceptual adjustment, multi-timescale policy learning, and compliant execution.

Despite these encouraging results, several limitations remain. In particular, the low-level force-control parameters are still tuned empirically and do not adapt online to changing interaction conditions. The current evaluation is also tied to an end-effector pose-control interface rather than a torque-level control stack. Future work will explore adaptive low-level control mechanisms, tighter integration of higher-rate tactile and force feedback, and more unified optimization between policy generation and compliant execution to further improve robustness and generalization.

REFERENCES

- [1] Z. He, H. Fang, J. Chen, H.-S. Fang, and C. Lu, "Foar: Force-aware reactive policy for contact-rich robotic manipulation," *IEEE Robotics and Automation Letters*, 2025.
- [2] H. Xue, J. Ren, W. Chen, G. Zhang, Y. Fang, G. Gu, H. Xu, and C. Lu, "Reactive diffusion policy: Slow-fast visual-tactile policy learning for contact-rich manipulation," *arXiv preprint arXiv:2503.02881*, 2025.
- [3] S. Jiang, S. Zhao, Y. Fan, and P. Yin, "Gelfusion: Enhancing robotic manipulation under visual constraints via visuotactile fusion," *arXiv preprint arXiv:2505.07455*, 2025.
- [4] J. Li, T. Wu, J. Zhang, Z. Chen, H. Jin, M. Wu, Y. Shen, Y. Yang, and H. Dong, "Adaptive visuo-tactile fusion with predictive force attention for dexterous manipulation," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3232–3239, IEEE, 2025.
- [5] J. Jones, O. Mees, C. Sferrazza, K. Stachowicz, P. Abbeel, and S. Levine, "Beyond sight: Finetuning generalist robot policies with heterogeneous sensors via language grounding," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5961–5968, IEEE, 2025.
- [6] J. Yu, H. Liu, Q. Yu, J. Ren, C. Hao, H. Ding, G. Huang, G. Huang, Y. Song, P. Cai, *et al.*, "Forcevla: Enhancing vla models with a force-aware moe for contact-rich manipulation," *arXiv preprint arXiv:2505.22159*, 2025.
- [7] Q. Zhao, Y. Lu, M. J. Kim, Z. Fu, Z. Zhang, Y. Wu, Z. Li, Q. Ma, S. Han, C. Finn, *et al.*, "Cot-vla: Visual chain-of-thought reasoning for vision-language-action models," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 1702–1713, 2025.
- [8] L. Heng, H. Geng, K. Zhang, P. Abbeel, and J. Malik, "Vitacformer: Learning cross-modal representation for visuo-tactile dexterous manipulation," *arXiv preprint arXiv:2506.15953*, 2025.
- [9] F. Liu, C. Li, Y. Qin, J. Xu, P. Abbeel, and R. Chen, "Vitamin: Learning contact-rich tasks through robot-free visuo-tactile manipulation interface," *arXiv preprint arXiv:2504.06156*, 2025.
- [10] H. Zhao, Y. Xie, J. Gong, Y. Sun, K. Zhu, W. He, S. Li, H. Fu, and W. Ding, "Tacauchy: An extensible fem framework for vision-based tactile simulation," *arXiv preprint arXiv:2606.20426*, 2026.
- [11] W. Liu, J. Wang, Y. Wang, W. Wang, and C. Lu, "Forcecemic: Force-centric imitation learning with force-motion capture system for contact-rich manipulation," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1105–1112, IEEE, 2025.
- [12] C. Chen, Z. Yu, H. Choi, M. Cutkosky, and J. Bohg, "Dexforce: Extracting force-informed actions from kinesthetic demonstrations for dexterous manipulation," *IEEE Robotics and Automation Letters*, 2025.
- [13] J. J. Liu, Y. Li, K. Shaw, T. Tao, R. Salakhutdinov, and D. Pathak, "Factr: Force-attending curriculum training for contact-rich policy learning," *arXiv preprint arXiv:2502.17432*, 2025.
- [14] Y. Wu, Z. Chen, F. Wu, L. Chen, L. Zhang, Z. Bing, A. Swikir, S. Haddadin, and A. Knoll, "Tacdiffusion: Force-domain diffusion policy for precise tactile manipulation," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 11831–11837, IEEE, 2025.
- [15] J. H. Kang, S. Joshi, R. Huang, and S. K. Gupta, "Robotic compliant object prying using diffusion policy guided by vision and force observations," *IEEE Robotics and Automation Letters*, 2025.
- [16] Y. Hou, Z. Liu, C. Chi, E. Cousineau, N. Kuppaswamy, S. Feng, B. Burchfiel, and S. Song, "Adaptive compliance policy: Learning approximate compliance for diffusion guided control," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4829–4836, IEEE, 2025.
- [17] J. Hansen, F. Hogan, D. Rivkin, D. Meger, M. Jenkin, and G. Dudek, "Visuotactile-rl: Learning multimodal manipulation policies with deep reinforcement learning," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 8298–8304, IEEE, 2022.
- [18] F. Yang, C. Ma, J. Zhang, J. Zhu, W. Yuan, and A. Owens, "Touch and go: Learning from human-collected vision and touch," *arXiv preprint arXiv:2211.12498*, 2022.
- [19] Z. Zhao, S. Haldar, J. Cui, L. Pinto, and R. Bhirangi, "Touch begins where vision ends: Generalizable policies for contact-rich manipulation," *arXiv preprint arXiv:2506.13762*, 2025.
- [20] B. Huang, J. Xu, I. Akinola, W. Yang, B. Sundaralingam, R. O'Flaherty, D. Fox, X. Wang, A. Mousavian, Y.-W. Chao, *et al.*, "Vt-refine: Learning bimanual assembly with visuo-tactile feedback via simulation finetuning," *arXiv preprint arXiv:2510.14930*, 2025.
- [21] H. Ge, Y. Jia, Z. Li, Y. Li, Z. Chen, L. Shi, L. Han, R. Huang, and G. Zhou, "FILIC: Dual-loop force-guided imitation learning with impedance torque control for contact-rich manipulation tasks," *arXiv preprint arXiv:2509.17053*, 2025.
- [22] G. Yan, J. Zhu, Y. Deng, S. Yang, R.-Z. Qiu, X. Cheng, M. Memmel, R. Krishna, A. Goyal, X. Wang, *et al.*, "Maniflow: A general robot manipulation policy via consistency flow training," *arXiv preprint arXiv:2509.01819*, 2025.
- [23] E. Helmut, N. Funk, T. Schneider, C. de Farias, and J. Peters, "Tactile-conditioned diffusion policy for force-aware robotic manipulation," *arXiv preprint arXiv:2510.13324*, 2025.
- [24] Y. Xie, Y. Wang, G. Chen, J. Liu, Y. Sun, and W. Ding, "Learning human-intention priors from large-scale human demonstrations for robotic manipulation," *arXiv preprint arXiv:2604.24681*, 2026.
- [25] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel, "Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 12156–12163, IEEE, 2024.
- [26] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, E. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [27] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," *The International Journal of Robotics Research*, vol. 44, no. 10-11, pp. 1684–1704, 2025.
- [28] N. Hogan, "Impedance control: An approach to manipulation: Part i—theory," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 1–7, 1985.
- [29] R. J. Anderson and M. W. Spong, "Hybrid impedance control of robotic manipulators," *IEEE Journal on Robotics and Automation*, vol. 4, no. 5, pp. 549–556, 1988.