PanoTac: A Dense Three-Axis Tactile Sensor for Robust Touch-Centric Manipulation

Binglin (Andre) Wang[1,†], Xiaole (Alex) Jiang[2,†], Zicheng Teng[1], Xinqi Liu[1], Yunzhu Li[3], Lerrel Pinto[1], Zichen Jeff Cui[1,‡], and Binghao Huang[3,‡,*]

[1]New York University [2]City University of New York
[3]Columbia University

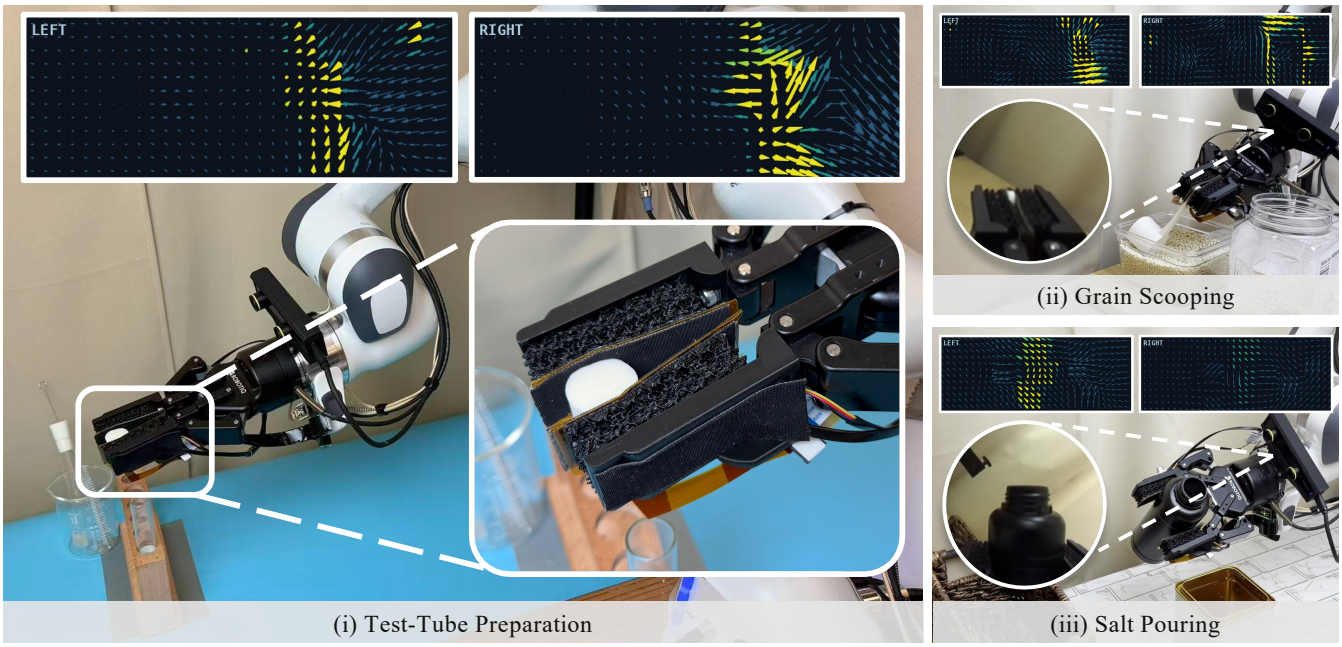


Fig. 1: We propose **PanoTac**, a low-cost, compliant tactile fingertip for dense three-axis tactile sensing. We showcase PanoTac with three tasks *(i)* test-tube preparation, where in-hand state helps infer tube orientation and shear feedback helps detect stopper release; *(ii)* grain scooping, where contact patterns help infer spoon orientation and guide robot–environment interaction; and *(iii)* salt pouring, where weight cues distinguish filled from empty bottles and inform pouring decisions.

***Abstract*— Contact-rich manipulation benefits from tactile feedback that captures both detailed contact geometry and interaction forces. When scooping grains, contact patterns help infer spoon orientation, while resistance transmitted through the grasp guides motion. To provide these complementary sensing capabilities, we present *PanoTac*, a low-cost, compliant tactile fingertip designed for dense three-axis tactile sensing. Each fingertip provides a $12 \times 32 \times 3$ tactile representation, with one normal (z) and two shear (x, y) response components at each taxel. The sensor operates at 60 Hz and costs less than $50 per fingertip. Sensor characterization quantifies response repeatability and consistency across sensor instances and contact locations. We integrate *PanoTac* into a robotic system with a visuo-tactile pipeline. Across test-tube preparation, grain scooping, and salt pouring, the resulting policies outperform vision-only baselines by an average of 46.7 percentage points. Ablations show that removing either spatial contact information or shear feedback reduces overall performance across the task suite. Our project page is available at https://panotac.github.io.**

† Equal contribution.
‡ Equal advising.
∗ Corresponding author: Binghao Huang.
Email: binghao.huang@columbia.edu.

I. Introduction

Humans rely on vision and touch to manipulate daily objects, using tactile feedback to understand contact states and determine how an interaction should proceed. Consider inserting a test tube into a rack and removing its stopper (Fig. 1(i)). Vision guides positioning, while contact patterns help infer the tube's in-hand orientation under partial occlusion. Stopper removal requires sensing whether the stopper remains engaged or has come free. Changes in resistance transmitted through the fingers help us decide when to stop wiggling and lift the stopper free. These stages highlight two complementary aspects of touch: spatial contact information helps estimate in-hand object orientation, while shear feedback helps detect changes in interaction state.

These requirements motivate dense three-axis tactile sensing, but realizing this capability in a practical robotic fingertip presents two challenges. *(i) Spatial resolution and force sensing.* Dense piezoresistive arrays provide detailed responses to normal contact but do not directly capture shear loading [1]–[3]. Magnetic tactile sensors provide responses to

normal and shear loading [4], [5], but the size and spacing of discrete sensing elements constrain spatial sampling density, making fine contact patterns difficult to resolve directly [6]. *(ii) Hardware integration and reproducibility.* Optical tactile sensors capture detailed contact geometry and support force estimation, but accommodating the camera and optical path imposes constraints on fingertip packaging and geometry [7]. Practical deployment also requires affordable, compliant hardware with straightforward fabrication [8], repeatable responses, and consistency across independent instances.

To address these challenges, we present **PanoTac**, a low-cost, compliant tactile fingertip designed for dense three-axis tactile sensing. Our contributions are threefold:

(i) A dense three-axis tactile fingertip. We combine spatial contact information from a dense piezoresistive sensing array [2] with magnetic responses to fingertip deformation [4] on a shared grid. This yields a $12 \times 32 \times 3$ tactile representation per fingertip, with one normal (z) and two shear (x, y) response components at each taxel. We design the magnet arrangement to reduce magnetic cross-talk between opposing fingertips. The fingertip is modular and straightforward to fabricate, with each sensor costing below \$50, excluding readout boards. Fabrication requires less than 3 hours of unattended printing and 30 minutes of hands-on assembly.

(ii) Sensor characterization and system integration. We evaluate response repeatability under repeated loading and characterize consistency across sensor instances and contact locations. We integrate *PanoTac* into a robotic system that conveys tactile information through auditory feedback, enabling operators to adjust actions during data collection.

(iii) Visuo-tactile policy learning and modality analysis. We learn visuo-tactile manipulation policies and evaluate them on test-tube preparation, grain scooping, and salt pouring. The resulting policies achieve an overall success rate of 91.7% across these three tasks, compared with 45.0% for vision-only policies. Through modality ablations, we evaluate the complementary contributions of dense spatial contact information and shear feedback, showing that removing either reduces overall task success.

We describe the sensor design, fabrication, tactile representation, and policy learning pipeline in Sec. III, report hardware characterization in Sec. IV, and present robot learning experiments and modality ablations in Sec. V.

II. Related Work

We first review tactile sensing mechanisms and the trade-offs that motivate dense three-axis sensing. We then discuss how prior work incorporates touch into visuo-tactile manipulation and learning.

A. Tactile Sensing Hardware

Optical sensing is a widely used approach to tactile sensing [7], [9]–[12] and is also incorporated into multimodal sensors [13], [14]. These sensors have a high-dimensional observation space (an image) and extract tactile information from a camera observing deformation of the elastomer contact surface [15], [16]. While this provides detailed contact geometry, the optical path between camera and elastomer sets a minimum sensor thickness, decreasing its manipulation clearance compared to other sensing mechanisms, and the elastomer surface abrades with use and is treated as a consumable [9]. Capacitive sensors offer a thinner profile and high sensitivity [17], [18], and recent designs [19], [20] deliver multi-axis force readings from a compact package. A practical challenge is calibration: the capacitance-to-force mapping is nonlinear and varies between units, requiring individual sensor calibration. Contact-microphone sensors [21], [22] detect tactile events by registering mechanical vibrations propagating through the contact surface. This yields a naturally event-driven signal, but a passive microphone carries no static component of the contact state and cannot directly measure sustained loads.

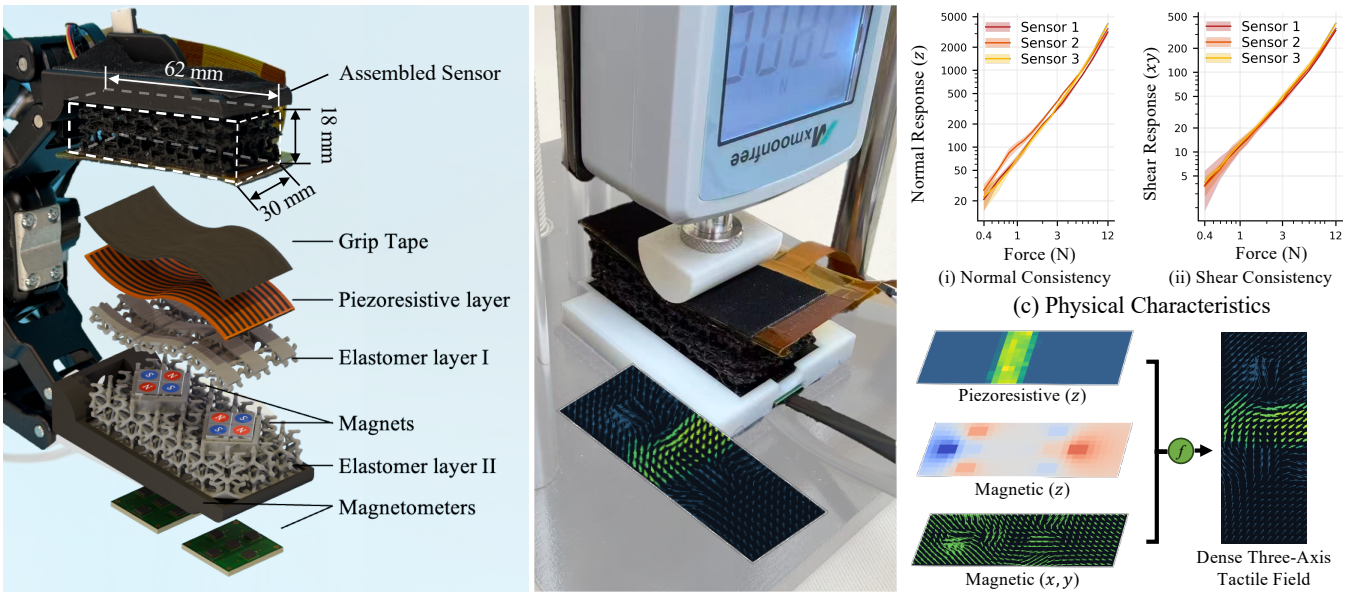
Piezoresistive arrays [1], [2], [23], [24] are flexible, easy to fabricate, and provide a high-resolution map of normal pressure, including contact location and shape, but cannot directly measure shear. Magnetic sensors [4], [5], [25], [26] infer aggregate normal and shear force from the deformation of a magnetized skin or microstructure, but their spatial resolution is bounded by the magnetometer count, so they provide a sparse local contact description. They also suffer from cross-talk between neighboring sensing units and from sensitivity to ferromagnetic objects and external fields.

Our sensor draws the best of both worlds. It keeps the low-cost fabrication of piezoresistive and microstructure-based magnetic skins, and it produces a three-axis tactile representation combining normal and shear responses with contact location and geometry. Sec. III describes the design.

B. Visuo-Tactile Manipulation Systems and Learning

Recent work has explored tactile representation learning and adapting generalist policies to additional sensory modalities [27]. One line learns a unified representation space across sensor classes [28]–[33], so that data collected on one sensor benefits a policy deployed on another; this is complementary to our work, and a sensor like ours is another sensor type these representations could incorporate. Another line reconstructs one tactile modality from another, to extract more tactile information from a limited sensor setup [34], [35]. Our proposition is the opposite: when affordable hardware can directly measure the tactile information needed for the task, one should favor direct measurement over reconstruction.

Work on visuo-tactile manipulation [1], [5], [13], [20], [36]–[40] usually builds the learning pipeline around a particular sensor, with a fusion module or pre-trained encoder designed for it. The task suites in this line of work also tend to be vision-centric, with tactile feedback primarily refining visual information. For example, task designs often focus on one axis of tactile information, such as force feedback helping with soft-touch and insertion, or contact pose helping with reorientation. We instead treat the sensor as one more observation stream that enters a standard end-to-end policy alongside the camera. We verify it on touch-centric tasks that a vision-only policy cannot solve, such as weighing a bottle



(a) Tactile Integrated Soft Gripper Design (b) Sensor Characterization Setup (d) Three-Axis Tactile Field Construction

Fig. 2: **PanoTac Tactile Sensing Platform.** *Part (a)* shows the assembled sensor and an exploded view of the multilayer design. *Part (b)* shows the indentation setup for sensor characterization. *Part (c)* shows the consistency of (i) the normal response z and (ii) the RMS shear (x, y) magnitude across three sensor under normal loads of 0.4–12 N. Shaded bands indicate ± 1 standard deviation. *Part (d)* illustrates how we construct the $12 \times 32 \times 3$ dense three-axis tactile representation.

or determining the orientation of an occluded tool before using it. Further details are provided in Sec. V.

III. PanoTac: Design, Fabrication, and Learning

A. Overview

PanoTac is a compliant, modular fingertip for dense three-axis tactile sensing. As shown in Fig. 2, its piezoresistive array captures detailed contact patterns, while magnetic measurements detect field changes caused by deformation of the supporting elastomer body. We construct a $12 \times 32 \times 3$ tactile representation per fingertip from these signals, with one normal (z) and two shear (x, y) response components at each taxel. This section describes the hardware design, tactile representation, and integration into a robotic system for demonstration collection and policy learning.

B. Hardware Design and Layout

As shown in Fig. 2(a), **PanoTac** comprises a piezoresistive sensing array, two compliant elastomer layers with magnets embedded between them, and magnetometers fixed at the base. This structure supports the contact surface while allowing deformation under normal and shear loading. The fingertip has a 62×30 mm contact surface and an 18 mm-thick contact stack. Each fingertip is carried by a rigid PLA housing that attaches to a parallel-jaw gripper finger. The housing holds the magnetometer boards at the rear and locates the sensing body through a keyed front slot, allowing tool-free replacement of the assembled sensing module.

(i) Normal and shear sensing principles. We use the dense piezoresistive array [2], providing 12×32 taxels at a pitch of 2 mm. Normal loading changes the local electrical resistance measured at each taxel. To capture shear loading, we support the array with a compliant elastomer body embedded with neodymium magnets. External loading deforms the elastomer and displaces the magnets relative to the fixed magnetometers, changing the measured magnetic fields [4], [5].

(ii) Compliant support structure. To support the contact array while accommodating fingertip deformation, we use coplanar thermoplastic polyurethane (TPU) tiles separated by relief gaps. These tiles also provide bonding surfaces for the array. We use TPU 95A above the magnet pockets to support the contact array and retain the magnets. Below the pockets, TPU 85A provides a more compliant layer that permits load-induced magnet displacement, including under shear. We apply grip tape (3M Gripping Material TB641) to the contact-facing surface to increase friction and use double-sided adhesive to bond the array to the elastomer support.

(iii) Magnet layout and interference reduction. We arrange two magnet groups along the fingertip, each centered within an elastomer pocket directly above its magnetometer board. This placement strengthens local signals and increases separation between neighboring groups. Each group contains four magnets in a 2×2 checkerboard pattern, with adjacent magnets magnetized in opposite directions along the surface normal to reduce magnetic interference. We rotate each group 45° about the surface normal relative to the magnetometer axes to distribute shear-induced magnetic responses across x and y channels and reduce the risk of channel saturation.

To explain the far-field behavior, we model the four magnets in each group as equal-magnitude point dipoles located at the corners of a square, with magnetic moments normal to the square plane. In the ideal symmetric checkerboard configuration, the dipole and quadrupole moments cancel, leaving an octupole as the leading multipole. For the same

magnet positions and moment magnitudes, an idealized multipole expansion gives the far-field magnitudes for aligned and checkerboard configurations:

$$|\mathbf{B}_{\text{aligned}}| = B_0\, g_{\text{aligned}}(\Omega) \left(\frac{a}{r}\right)^3 \left[1 + \mathcal{O}\left(\frac{a^2}{r^2}\right)\right], \quad (1)$$

$$|\mathbf{B}_{\text{checker}}| = B_0\, g_{\text{checker}}(\Omega) \left(\frac{a}{r}\right)^5 \left[1 + \mathcal{O}\left(\frac{a^2}{r^2}\right)\right]. \quad (2)$$

Here, $2a$ is the side length of a single 2×2 magnet arrangement, r is the distance from the center of that arrangement, and B_0 sets the magnetic field strength scale. The dimensionless functions $g_{\text{aligned}}(\Omega)$ and $g_{\text{checker}}(\Omega)$ describe the angular dependence on the observation direction Ω from that center. Under this idealized model, the leading checkerboard far-field term scales as r^{-5}, compared with r^{-3} for the aligned configuration. This faster decay motivates the alternating-polarity layout for reducing magnetic cross-talk between neighboring sensing units and opposing fingertips.

C. Three-Axis Tactile Representation

We construct a dense three-axis tactile representation by mapping contact and magnetic measurements onto a shared taxel grid, as shown in Fig. 2(d). We first low-pass filter both sensing streams and subtract their respective no-contact baselines. We denote the resulting magnetic response at magnetometer m by $\Delta\mathbf{B}_m$.

(i) Normal component. The normal component combines spatial contact patterns from the piezoresistive array with a global magnitude from the magnetic z responses. We suppress piezoresistive readings below a noise threshold and apply Gaussian smoothing with standard deviation 0.5 taxel to obtain the contact activation map A. We set $N = 0$ when no contact is detected. Otherwise, we normalize A by its maximum and scale it by the mean absolute magnetic z response across all magnetometers:

$$N_{ij} = \left(\frac{1}{M}\sum_{m=1}^{M} |\Delta B_{m,z}|\right) \frac{A_{ij}}{\max_{u,v} A_{uv}}, \quad (3)$$

where $M = 10$ is the number of magnetometers per fingertip and (i, j) indexes the taxel grid.

(ii) Shear components. We assign each magnetometer's $\Delta B_{m,x}$, $\Delta B_{m,y}$ readings to two taxels in the central column of its corresponding footprint on the contact grid. We fill the remaining taxels by interpolating each component from the magnetometer centers using inverse-distance weighting with power $p = 2$, obtaining shear-related responses S^x_{ij} and S^y_{ij}.

(iii) Three-axis tactile field. Combining normal and shear response maps yields the three-axis tactile representation:

$$\mathbf{T}_{ij} = \left(S^x_{ij}, S^y_{ij}, N_{ij}\right), \qquad \mathbf{T} \in \mathbb{R}^{12\times 32\times 3}. \quad (4)$$

We use this field for sensor characterization and real-time tactile visualization. Throughout the paper, we use z to denote the normal response component and (x, y) to denote the two shear response components. Our policy takes this three-axis representation as tactile input, as detailed in Sec. III-D.

D. System Integration and Policy Learning

We equip both fingers of a Robotiq 2F-85 gripper, mounted on a Franka Panda arm, with **PanoTac**. We acquire tactile measurements at 60 Hz.

(i) Teleoperation and data collection. We collect demonstrations through VR-based teleoperation. We display the contact map as a real-time heatmap and provide stereo tones whose pitch and volume increase with the deviation of each fingertip's magnetic readings from their moving average. This visual and auditory feedback helps the operator adjust robot motion during teleoperation. For each demonstration, we record wrist-camera RGB images, robot proprioception, tactile measurements, and action commands. We align the recorded streams at 30 Hz and process the tactile measurements as described in Sec. III-C. Details of the teleoperation interface are given in Appendix A.

(ii) Policy architecture. We train a U-Net diffusion policy [3], [41] jointly with its observation encoders to predict an action sequence from a single multimodal observation frame. An ImageNet-pretrained ResNet-18 encodes the visual observations and is fine-tuned jointly with the policy. A single three-layer CNN followed by fully connected layers jointly encodes the dense three-axis tactile representations from both fingertips, constructed as described in Sec. III-C. The visual and tactile features are concatenated with the robot proprioception to condition the U-Net through FiLM. For the tactile baselines (Sec. V-A), we use an MLP to encode the sparse three-axis magnetic readings and a CNN to encode measurements from the dense normal array. Encoder, training, and deployment details are given in Appendix B.

IV. Sensor Characterization

We evaluate the repeatability and cross-instance consistency of *PanoTac* under loading (Sec. IV-B) and its response variation across contact locations (Sec. IV-C).

A. Experimental Setup and Response Metrics

We characterize *PanoTac* using a force gauge with an experiment-specific indenter. For load-response experiments, the curved surface of a D-shaped indenter produces a rectangular contact area, as shown in Fig. 2(b). For spatial characterization, a spherical indenter produces localized contacts at selected positions on the fingertip. The force gauge provides the reference normal load for both experiments.

We record the piezoresistive and magnetic streams synchronously during each indentation and process them as described in Sec. III-C. From these we obtain a normal response and a shear response, the latter defined as the RMS of the 2-D (x, y) shear vector over all taxels. Sensor responses are derived from the readout boards' digital outputs, while applied loads are measured in newtons. Appendix C details the response computation and indentation protocols.

B. Response Consistency under Loading

We test three independently fabricated sensor instances under normal loads at 15 levels from 0.4 to 12 N, with three repeated sweeps per instance. Fig. 2(c) shows the normal z

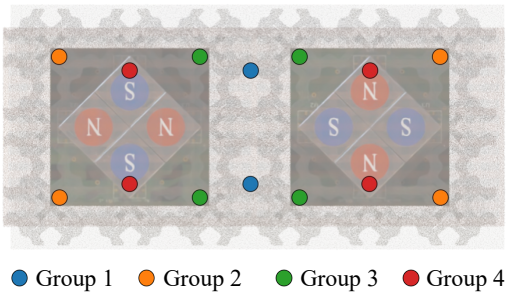


Fig. 3: **Probe locations for spatial response characterization.** Colored dots indicate probe locations, grouped according to their positions relative to the magnet layout. Each dot is probed 10 times. The corresponding response statistics are reported in Table II.

and RMS shear (x, y) responses. For each instance, lines and shaded bands indicate the mean and ± 1 standard deviation across the three sweeps.

We quantify response consistency using the coefficient of variation, $\mathrm{CV} = 100\,\sigma/\mu$. At each load, within-instance CV measures variation across the three sweeps of an instance, while cross-instance CV measures variation across the mean responses of the three instances. We average these CVs within three load bands: 0.4–1 N, 1.2–4 N, and 6–12 N. In Table I, repeatability improves with load, with mean within-instance CVs reaching 4.8% for the normal z response and 4.3% for the shear (x, y) response at 6–12 N. Cross-instance mean CVs for the shear (x, y) response remain below 9%, while those for the normal z response decrease from 22.9% to 12.5% and 10.8% across the three load bands.

C. Spatial Response Consistency

The piezoresistive layer samples on a highly symmetric 4 mm^2 taxel grid, while the magnetometers sense load only at sparse key points. Due to the nature of our magnetic sensor, the magnetometer reading enjoys fewer spatial symmetries than its piezoresistive counterpart. To probe for spatial consistency, we therefore define four probe location classes by their symmetric position on the sensor (Fig. 3): two points on the median line (G1), four points at the corners (G2), four points near the median line on the edge of the sensor (G3), and four interior points on the magnet groups (G4). Every site within a class is equivalent to the others under the symmetry of the magnet layout and is therefore amenable to statistical averaging. At every site we sample 10 presses. Both modalities are recorded from the same press. Table II reports the summary statistics. The responses exhibit moderate variation relative to their means, with response magnitudes varying across probe groups.

V. Learning Touch-centric Manipulation

Our experiments address two questions: whether dense three-axis tactile feedback improves manipulation performance, and how spatial contact information and shear feedback contribute to different manipulation behaviors. We compare **PanoTac** with a vision-only baseline and visuo-tactile baselines using different tactile inputs, analyzing full-task success, stage-level performance, and failure cases.

TABLE I: **Force-response consistency** ($\mathrm{CV} = 100\sigma/\mu$, sample SD). Repeatability: CV over the three sweeps of one instance; cross-instance: CV over the three instance means. Entries average per-load CVs within each band, repeatability also over the three instances.

		0.4–1 N	1.2–4 N	6–12 N
Repeatability	Normal	15.5 %	6.9 %	4.8 %
	Shear	21.1 %	8.3 %	4.3 %
Cross-instance	Normal	22.9 %	12.5 %	10.8 %
	Shear	4.2 %	5.3 %	8.8 %

TABLE II: **Spatial response** by point group (a.u., mean $\pm$ SD over presses).

	Group 1 (n=20)	Group 2 (n=40)	Group 3 (n=40)	Group 4 (n=40)
Normal z	646.1±68.7	992.7±79.5	588.8±82.0	683.6±79.0
Shear (x, y)	92.3±3.4	146.3±17.6	91.4±8.7	59.4±10.6

A. Experimental Setup

We evaluate all policies on three touch-centric manipulation tasks: *(i) test-tube preparation*, a long-horizon task in which dense contact patterns support tube reorientation and insertion, while shear feedback helps detect stopper release; *(ii) grain scooping*, in which contact patterns along the spoon handle help infer spoon orientation and guide the scooping direction; and *(iii) salt pouring*, in which shear feedback provides weight cues from an opaque bottle to guide the choice between pouring and discarding.

For each task, we train four policies on the same demonstrations: vision only, vision with dense piezoresistive contact maps, vision with sparse three-axis magnetic readings, and vision with *PanoTac*'s dense three-axis tactile representation. We refer to these as the vision-only, dense normal, sparse three-axis, and PanoTac policies, respectively. Each policy is evaluated over 20 rollouts per task, with at most 800 steps per rollout. Stage-level success rates are computed over all rollouts for the corresponding task or condition, including those that fail before reaching the stage. Overall success is the full-task success rate pooled across all three tasks (60 rollouts per policy). Task details follow.

B. Test-Tube Preparation

1) Task Setup: The robot retrieves a test tube from a storage box, reorients it upright against a rack, and inserts it into the rack. It then wiggles the stopper to loosen it, extracts it from the tube, and returns it to the box. Contact patterns provide cues about the tube's in-hand orientation, while shear feedback helps detect stopper release. We collect 50 demonstrations varying tube location and height, stopper color, and stopper fit tightness. Variations in fit tightness change the force required for removal. Fig. 4(a) shows a representative rollout. **Evaluation Metrics:** Reorientation

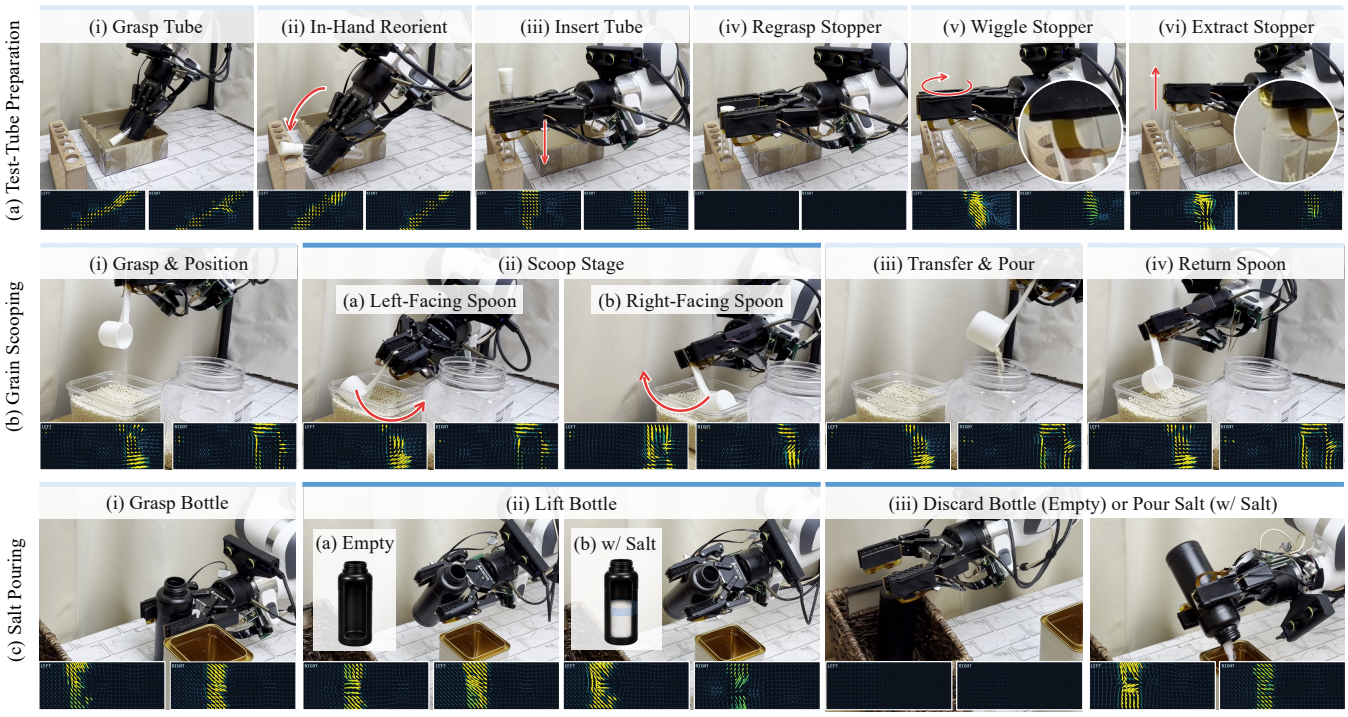


Fig. 4: **Qualitative Results.** Representative stages of (a) test-tube preparation, (b) grain scooping, and (c) salt pouring. Paired tactile visualizations show responses from both fingertips, and red arrows indicate motion directions. Blue bars mark stages shown for both task branches. Detailed task descriptions are provided in Sec. V.

TABLE III: **Quantitative Results.** Success rates (%) over 20 rollouts per policy per task, with 10 per spoon orientation or bottle condition. Overall reports full-task success averaged across the three tasks. Stage rates are not conditioned on reaching the stage. Bold indicates the best result in each column.

Policy	**Test-Tube Preparation**			**Grain Scooping**				**Salt Pouring**			**Overall**
	Reorient	Insert	**Unplug (END)**	Pre-Scoop Direction		Scoop	**Return (END)**	Sort Bottle Stage		**Return (END)**	
				Left	Right			With Salt	Empty		
Vision Only	**100**	90	35	20	80	50	45	50	60	55	45.0
Vision w/ Dense Normal Array	**100**	**100**	65	80	**100**	**85**	**85**	90	50	70	73.3
Vision w/ Sparse Three-Axis	**100**	**100**	85	90	70	70	60	90	**100**	90	78.3
Vision w/ PanoTac	**100**	**100**	**90**	**100**	**100**	**85**	**85**	**100**	**100**	**100**	**91.7**

succeeds when the tube is upright. Insertion succeeds when the tube is placed in the rack. Unplugging succeeds when the stopper is fully removed from the tube before the robot begins the return motion. A premature return attempt is counted as an unplugging failure. Every policy returns the stopper to the box in all rollouts that reach this stage, so we report unplugging as the final stage.

2) Analysis and Results: This task examines whether tactile feedback supports tube alignment and the transition from stopper loosening to extraction.

Tactile-refined alignment. All four policies achieve 100% reorientation success. Insertion additionally requires aligning the tube with the rack opening. All three tactile policies achieve 100% insertion success, compared with 90% for the vision-only policy (Table III), consistent with tactile feedback supporting alignment during insertion.

Force-sensitive wiggling. This stage requires applying sufficient force to loosen the stopper and detecting when to transition from wiggling to extraction. PanoTac achieves 90% unplugging success, compared with 35% for the vision-only policy, 65% for the dense normal policy, and 85% for the sparse three-axis policy. Representative vision-only failures involve lifting before stopper release, which can pull the tube out of the rack or cause it to slip from the grasp (Fig. 5(a)). The dense normal policy is the only policy observed to continue wiggling after stopper release (Fig. 5(a)), revealing difficulty detecting stopper release and coordinating the transition to extraction using dense normal feedback alone. Both policies with magnetic feedback achieve higher unplugging success than the dense normal policy, consistent with magnetic responses providing useful cues for stopper release. PanoTac achieves the highest observed unplugging success, with one more successful rollout than the sparse three-axis policy.

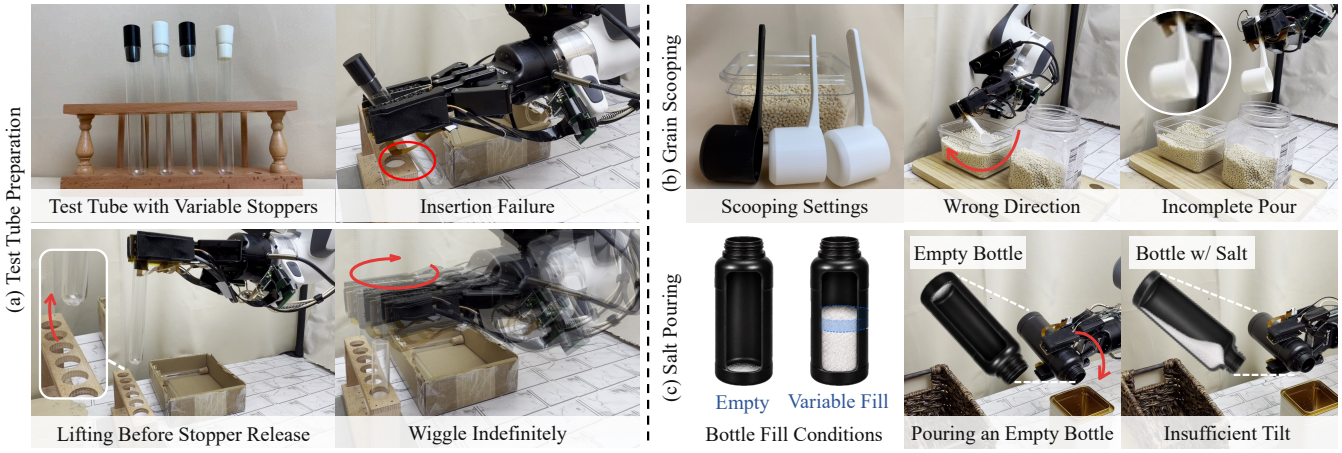


Fig. 5: **Representative Failure Cases of Baseline Policies.** We show variations in (a) test tubes, (b) spoons, and (c) bottle fill levels, together with representative manipulation failures. Detailed failure analysis is provided in Sec. V.

C. Grain Scooping

1) Task Setup: The robot must scoop grain from one container, pour it into another, and return the spoon. The spoon initially faces either left or right, requiring different scooping directions. After grasping, the spoon bowl is **occluded from the wrist camera**, while contact patterns along the handle provide orientation cues. We collect 45 demonstrations varying spoon color, handle length and width, spoon placement, grain surface level, and container height. The container is raised by 3 cm in approximately half of the demonstrations, and the two spoon orientations are balanced. These variations require the robot to adjust its motion to reach the grain. Fig. 4(b) shows a representative rollout. **Evaluation Metrics:** Direction selection succeeds when the robot initiates scooping in the direction appropriate for the spoon's orientation. Scooping succeeds when grain is lifted from the source container. Return succeeds when the spoon is returned to the source container.

2) Analysis and Results: This task examines how dense contact geometry guides the scooping direction and how tactile feedback supports execution under varying grain levels and container heights.

Tool-geometry-based action selection. The policy must infer the spoon's orientation from its in-hand contact pattern. Moving in the wrong direction is the dominant failure of the baseline policies (Fig. 5(b)). Averaged over both spoon orientations, direction selection succeeds in 100% of rollouts for PanoTac, 90% for the dense normal array policy, 80% for the sparse three-axis policy, and 50% for vision only (Table III). The dense normal array policy selects the wrong direction in only two of 20 trials, suggesting that dense contact patterns are informative for handle orientation. Magnetic responses may also provide orientation cues, since the sparse three-axis policy performs above chance. PanoTac combines both and selects the correct direction in all 20 rollouts. The two modalities appear to resolve ambiguities that remain when either is used alone. The vision-only policy performs at chance level, with a bias toward scooping to the right.

Contact-aware execution. Full-task success favors the policies with dense contact information: PanoTac and the dense normal array policy both achieve 85%, compared with 60% for the sparse three-axis policy and 45% for vision only. Choosing the right direction does not guarantee successful grain transfer. Even when the robot moves spoon in the intended direction, it may pass through the grain without collecting a full scoop, and pouring may leave residual grain in the spoon. The sparse three-axis policy also loses track of the task stage. In one rollout it pours twice, although every demonstration contains a single pour. In another, it begins returning the spoon before the grain is fully dispensed (Fig. 5(b)). Its sparse magnetic feedback may not always reveal which stage the task is in. Return failures occur in both this policy and the vision-only policy. For the vision-only policy, one possible explanation is the variation in container height across demonstrations. An action suited to a lower container leaves the spoon poorly positioned for return, and the camera sees too little below the gripper to correct it.

D. Salt Pouring

1) Task Setup: The opaque bottle's **contents are fully occluded from view**. The robot must lift the bottle and use its weight to choose between pouring and discarding. If the bottle contains salt, the robot pours it into an opaque central container and returns the bottle to its original position. Otherwise, it discards the bottle into a trash basket. Fig. 4(c) shows a representative rollout. We collect 40 demonstrations: 20 with near-empty bottles containing residual salt and 20 with variable amounts of salt, approximately half a bottle. Half a bottle of salt weighs around 150 grams. Each pouring episode is followed by an empty-bottle episode that retains any leftover salt. Evaluation follows a scrambled order, with 10 rollouts per bottle condition. **Evaluation Metrics:** We evaluate bottle sorting and final placement. For an empty bottle, sorting succeeds if the bottle is tilted no further than the onset of pouring. For a bottle with salt, sorting succeeds if the robot produces a steady stream and begins returning the bottle only after the stream stops. Full-task success additionally requires returning the bottle to its original position after pouring, or discarding it into the basket when empty.

2) Analysis and Results: This task examines whether tactile shear cues support the choice between pouring and discarding when the bottle's contents are visually hidden.

Weight-based action selection. The PanoTac policy achieves 100% sorting success, compared with 95% for the sparse three-axis policy, 70% for the dense normal array policy, and 55% for vision only (Table III). The strong performance of both policies with magnetic feedback supports its utility for distinguishing bottle conditions through their shear responses. The sparse three-axis policy fails once in sorting and once in placement, likely because its shear response varies with grasp location. PanoTac's dense contact map localizes the grasp, which helps disambiguate these readings. The dense normal array policy sorts 90% of bottles with salt but only 50% of empty bottles, attempting to pour in every empty-bottle failure (Fig. 5(c)). This bias may reflect the greater length of pouring trajectories. Its single failure with salt is an insufficient tilt that releases no salt (Fig. 5(c)), possibly because a weak weight signal blurs the boundary between pouring and discarding, leading the policy to average the two actions.

Gripper compliance may provide indirect weight cues. In informal trials with a bottle filled to capacity, we observe successful vision-only execution, suggesting that larger weight-induced deformation may provide a visible weight cue. Smaller deformations at the evaluated salt levels could likewise affect the visual observations and normal contact map, although these cues are difficult to discern by eye. This offers a possible explanation for some successful decisions without magnetic feedback. Appendix D extends this task to pouring liquids, where we test out-of-distribution capability of our policy by including liquids of unseen color and couscous, which have a different pouring dynamics.

VI. Conclusion

In this paper, we present **PanoTac**, a low-cost, compliant tactile fingertip for dense three-axis tactile sensing. Hardware characterization showed spatially consistent and temporally repeatable responses across loads and sensor instances. Integrated into a visuo-tactile learning pipeline, our policies outperformed vision-only baselines by an average of 46.7 percentage points on three manipulation tasks, and modality ablations indicated that spatial contact information and shear feedback each contribute to performance. Given its cost, fabrication simplicity, and customizable form factor, we believe **PanoTac** is a practical tactile interface for manipulation.

Limitations. First, the magnetometer readings are susceptible to ferromagnetic objects and external magnetic fields; training on interaction data that include such objects is the most direct remedy. Second, force resolution is bounded by the compliance of the elastomer and the number of magnetometers; scaling the magnetometer array and tuning the elastomer are natural next steps. Third, the current fingertip is shaped for a parallel-jaw gripper; adapting it to dexterous hands and evaluating in-hand manipulation would test the representation beyond two-finger grasping. Finally, the sensor's low cost and simple installation make it a candidate for handheld, UMI-style data collection, which we leave to future work.

Acknowledgment

We thank Venkatesh Pattabiraman, Zifan Zhao, Bowen Tan, Hengkai Pan, Kanad Patel, Yibo Yan, Nikhil Bhattasali, and Kevin Wu for helpful discussions.

We used ChatGPT (OpenAI) and Claude Fable (Anthropic) to assist with language polishing and Blender to create the 3D illustration in Fig. 2.

References

[1] B. Huang, Y. Wang, X. Yang, Y. Luo, and Y. Li, "3D-ViTac: Learning fine-grained manipulation with visuo-tactile sensing," *arXiv preprint arXiv:2410.24091*, 2024.

[2] B. Huang and Y. Li, "FlexiTac: A low-cost, open-source, scalable tactile sensing solution for robotic systems," *arXiv preprint arXiv:2604.28156*, 2026.

[3] X. Zhu, B. Huang, and Y. Li, "Touch in the wild: Learning fine-grained manipulation with a portable visuo-tactile gripper," *Advances in Neural Information Processing Systems*, vol. 38, pp. 153 783–153 812, 2026.

[4] V. Pattabiraman, Z. Huang, D. Panozzo, D. Zorin, L. Pinto, and R. Bhirangi, "eFlesh: Highly customizable magnetic touch sensing using cut-cell microstructures," *arXiv preprint arXiv:2506.09994*, 2025.

[5] R. Bhirangi, V. Pattabiraman, E. Erciyes, Y. Cao, T. Hellebrekers, and L. Pinto, "AnySkin: Plug-and-play skin sensing for robotic touch," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 16 563–16 570.

[6] P. Hou, D. Sun, M. Wang, Y. Huang, Z. Zhang, H. Liu, W. Li, and Z. Jiao, "Supermag: Vision-based tactile data guided high-resolution tactile shape reconstruction for magnetic tactile sensors," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 8561–8567.

[7] W. Yuan, S. Dong, and E. H. Adelson, "GelSight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[8] A. Sipos, W. v. d. Bogert, and N. Fazeli, "Gelslim 4.0: Focusing on touch and reproducibility," *arXiv preprint arXiv:2409.19770*, 2024.

[9] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer *et al.*, "DIGIT: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.

[10] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "GelSlim: A high-resolution, compact, robust, and calibrated tactile-sensing finger," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1927–1934.

[11] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, and N. F. Lepora, "The TacTip family: Soft optical tactile sensors with 3D-printed biomimetic morphologies," *Soft Robotics*, vol. 5, no. 2, pp. 216–227, 2018.

[12] A. Padmanabha, F. Ebert, S. Tian, R. Calandra, C. Finn, and S. Levine, "OmniTact: A multi-directional high-resolution touch sensor," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 618–624.

[13] J. Zhao, N. Kuppuswamy, S. Feng, B. Burchfiel, and E. Adelson, "PolyTouch: A robust multi-modal tactile sensor for contact-rich manipulation using tactile-diffusion policies," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 104–110.

[14] J. Shan, J. Zhao, J. Liu, X. Wang, Z. Xia, G. Xu, and B. Fang, "MagicGel: A novel visual-based tactile sensor design with MagneticGel," *arXiv preprint arXiv:2503.23345*, 2025.

[15] M. Lambeta, T. Wu, A. Sengul, V. R. Most, N. Black, K. Sawyer, R. Mercado, H. Qi, A. Sohn, B. Taylor *et al.*, "Digitizing touch with an artificial multimodal fingertip," *arXiv preprint arXiv:2411.02479*, 2024.

[16] Z. Xu, Y. Li, X. Wu, T. Qiu, and L. Shao, "FingerEye: Learning dexterous manipulation with continuous vision-tactile sensing," *arXiv preprint arXiv:2604.20689*, 2026.

[17] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, and G. Metta, "Methods and technologies for the implementation of large-scale robot tactile sensors," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 389–400, 2011.

[18] P. Maiolino, M. Maggiali, G. Cannata, G. Metta, and L. Natale, "A flexible and robust large scale capacitive tactile system for robots," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3910–3917, 2013.

[19] H. Choi, J. E. Low, T. M. Huh, S. Hong, G. A. Uribe, K. A. Hoffmann, J. Di, T. G. Chen, A. A. Stanley, and M. R. Cutkosky, "CoinFT: A coin-sized, capacitive 6-axis force torque sensor for robotic applications," *arXiv preprint arXiv:2503.19225*, 2025.

[20] H. Choi, Y. Hou, C. Pan, S. Hong, A. Patel, X. Xu, M. R. Cutkosky, and S. Song, "In-the-wild compliant manipulation with umi-ft," IEEE, 2026. [Online]. Available: https://arxiv.org/abs/2601.09988

[21] Z. Liu, C. Chi, E. Cousineau, N. Kuppuswamy, B. Burchfiel, and S. Song, "ManiWAV: Learning robot manipulation from in-the-wild audio-visual data," *arXiv preprint arXiv:2406.19464*, 2024.

[22] J. Liu and B. Chen, "SonicSense: Object perception from in-hand acoustic vibration," in *Proceedings of The 8th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, vol. 270. PMLR, 2025, pp. 4332–4353.

[23] S. Sundaram, P. Kellnhofer, Y. Li, J.-Y. Zhu, A. Torralba, and W. Matusik, "Learning the signatures of the human grasp using a scalable tactile glove," *Nature*, vol. 569, no. 7758, pp. 698–702, 2019.

[24] J. Yin, H. Qi, Y. Wi, S. Kundu, M. Lambeta, W. Yang, C. Wang, T. Wu, J. Malik, and T. Hellebrekers, "OSMO: Open-source tactile glove for human-to-robot skill transfer," *IEEE Robotics and Automation Letters*, 2026.

[25] T. P. Tomo, M. Regoli, A. Schmitz, L. Natale, H. Kristanto, S. Somlor, L. Jamone, G. Metta, and S. Sugano, "A new silicone structure for uSkin—a soft, distributed, digital 3-axis skin sensor and its integration on the humanoid robot iCub," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2584–2591, 2018.

[26] R. Bhirangi, T. Hellebrekers, C. Majidi, and A. Gupta, "ReSkin: versatile, replaceable, lasting tactile skins," *arXiv preprint arXiv:2111.00071*, 2021.

[27] 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)*, 2025.

[28] F. Yang, C. Feng, Z. Chen, H. Park, D. Wang, Y. Dou, Z. Zeng, X. Chen, R. Gangopadhyay, A. Owens *et al.*, "Binding touch to everything: Learning unified multimodal tactile representations," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2024, pp. 26 330–26 343.

[29] C. Higuera, A. Sharma, C. K. Bodduluri, T. Fan, P. Lancaster, M. Kalakrishnan, M. Kaess, B. Boots, M. Lambeta, T. Wu *et al.*, "Sparsh: Self-supervised touch representations for vision-based tactile sensing," *arXiv preprint arXiv:2410.24090*, 2024.

[30] C. Higuera, A. Sharma, T. Fan, C. K. Bodduluri, B. Boots, M. Kaess, M. Lambeta, T. Wu, Z. Liu, F. R. Hogan, and M. Mukadam, "Tactile beyond pixels: Multisensory touch representations for robot manipulation," 2025. [Online]. Available: https://arxiv.org/abs/2506.14754

[31] J. Zhao, Y. Ma, L. Wang, and E. H. Adelson, "Transferable tactile transformers for representation learning across diverse sensors and tasks," *arXiv preprint arXiv:2406.13640*, 2024.

[32] R. Feng, J. Hu, W. Xia, T. Gao, A. Shen, Y. Sun, B. Fang, and D. Hu, "AnyTouch: Learning unified static-dynamic representation across multiple visuo-tactile sensors," in *International Conference on Learning Representations*, vol. 2025, 2025, pp. 31 265–31 285.

[33] J. Park, S. Bhadang, C. Sferrazza, S. Yi, and X. Wang, "TactX: Learning shared tactile representations across diverse sensors," *arXiv preprint arXiv:2606.31236*, 2026.

[34] S. Rodriguez, Y. Dou, M. Oller, A. Owens, and N. Fazeli, "Touch2Touch: Cross-modal tactile generation for object manipulation," *arXiv preprint arXiv:2409.08269*, 2024.

[35] Z. Li, Y. Ling, Y. Gu, B. Huang, C. Liang, S. Islam, H. Bedri, J. Chirikjian, Y. Li, S. Nikolaidis, and D. Seita, "Felt: Generating tactile signals from vision for visuo-tactile manipulation," 2026. [Online]. Available: https://arxiv.org/abs/2607.20683

[36] V. Pattabiraman, Y. Cao, S. Haldar, L. Pinto, and R. Bhirangi, "Learning precise, contact-rich manipulation through uncalibrated tactile skins," 2024. [Online]. Available: https://arxiv.org/abs/2410.17246

[37] 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," in *Robotics: Science and Systems*, 2025.

[38] H. Chen, J. Xu, H. Chen, K. Hong, B. Huang, C. Liu, J. Mao, Y. Li, Y. Du, and K. Driggs-Campbell, "Multi-modal manipulation via multi-modal policy consensus," *arXiv preprint arXiv:2509.23468*, 2025.

[39] J. Huang, Y. Ye, Y. Gong, X. Zhu, Y. Gao, and K. Zhang, "Spatially anchored tactile awareness for robust dexterous manipulation," *arXiv preprint arXiv:2510.14647*, 2025.

[40] 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*, vol. 10, no. 6, pp. 5625–5632, 2025.

[41] 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.

[42] E. Perez, F. Strub, H. De Vries, V. Dumoulin, and A. Courville, "FiLM: Visual reasoning with a general conditioning layer," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.

Appendix

A. Teleoperation Interface

The demonstrations are collected through VR teleoperation with a Meta Quest 3 at 30 Hz.

Tactile feedback. To help the operator regulate contact during manipulation, we present the tactile measurements as visual and auditory feedback. A real-time heatmap of each fingertip's contact map shows the spatial distribution of contact. In parallel, the left and right fingertips drive tones in the corresponding stereo channels, allowing the operator to distinguish force changes at the two fingers. For each fingertip, we compute the Euclidean norm of the deviation of its 30 magnetometer values (10 three-axis readings) from their exponential moving average. Values below a deadzone produce silence; above it, the signal is logarithmically compressed and mapped to increasing volume and pitch over 300–1200 Hz. The audio therefore emphasizes changes in force rather than its absolute level.

Recording. For each demonstration, we record wrist-camera RGB images, robot proprioception, tactile measurements from both fingertips, and the commanded end-effector poses and gripper actions. Tactile measurements are acquired at 60 Hz, and the recorded streams are aligned at 30 Hz. We process the tactile measurements into the three-axis representation described in Sec. III-C; policy inputs and subsampling are detailed in Appendix B.

B. Policy Implementation

All four policies in Sec. V share the same vision encoder, diffusion backbone, and training recipe; they differ in their tactile inputs and encoders, which are absent in the vision-only policy.

B.1 Observations and Encoders

Table IV lists the per-frame inputs and their encoders. The policy conditions on a single observation frame.

TABLE IV: **Observation inputs and encoders.** Shapes are per frame; L and R denote the two fingertips.

Modality	Input	Encoder	Dim
RGB image	$3 \times 256 \times 256$	ResNet-18	512
Proprioception	9 (position, 6D rotation)	identity	9
Dense normal	$2 \times 12 \times 32$	CNN	128
Sparse 3-axis	$2 \times 10 \times 3$	MLP	256
PanoTac	$3 \times 24 \times 32$	CNN	512

Vision. The image encoder is an ImageNet-pretrained ResNet-18, fine-tuned jointly with the policy.

PanoTac representation. The fields $\mathbf{T}$ of the two fingertips (Sec. III-C) are stacked along the row axis into a single $3 \times 24 \times 32$ image.

Conditioning. The vision, tactile, and proprioceptive features are concatenated into a single vector, which conditions every residual block of the U-Net through FiLM [42].

B.2 Diffusion Policy and Training

We use the convolutional U-Net diffusion policy of [41] with the hyperparameters in Table V.

TABLE V: **Diffusion policy hyperparameters**, shared across all policies and tasks.

Parameter	Value
U-Net channels	$[512, 1024, 2048]$
Diffusion steps	100 (train), 16 DDIM (inference)
Noise schedule	squared cosine, ϵ-prediction
Observation horizon	1 frame
Prediction horizon	16 steps (≈ 2.7 s at 6 Hz)
Optimizer	AdamW, lr 10^{-4}, weight decay 0.01
LR schedule	10-epoch warmup, cosine decay
Batch size / epochs	128 / 150
EMA	on U-Net, max decay 0.9999

B.3 Actions and Deployment

The policy predicts a 10-dimensional absolute action: end-effector position (3), rotation in the 6D representation (6), and a binary gripper command (1), normalized to $[-1, 1]$. Recorded streams are aligned at 30 Hz and subsampled by a factor of 5, so the policy operates at 6 Hz. At deployment, the policy runs inference at every step, and overlapping action chunks are combined by exponential temporal ensembling with a half-life of 4 policy steps, weighting the most recent prediction highest. The ensembled action is interpolated to 30 Hz for the robot controller.

C. Sensor Characterization Details

This section details the procedures behind Sec. IV. In both experiments, the fingertip lies flat on the bench surface.

C.1 Response Computation

For each recording, we average both sensing streams over the recording window, subtract the no-contact baseline, and compute $\mathbf{T}$ as in Sec. III-C. The normal response is the Frobenius norm $\|N\|_F$ of the normal map in (3). The shear response is the RMS of $\|(S^x_{ij}, S^y_{ij})\|$ over all 12×32 taxels, where S^x and S^y are obtained by inverse-distance-weighted projection of the magnetometer readings, as in Sec. III-C.

C.2 Load Response

Each instance undergoes three sweeps over 15 loads in increasing order: 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 6.0, 8.0, 10.0, and 12.0 N. Before each sweep, with the indenter lifted, we record a one-second no-contact baseline. At each load, the indenter is driven past the target force and backed off until the gauge reads the target. We hold the load until the readings stop drifting, then record both layers for one second. The indenter stays fixed within each sweep.

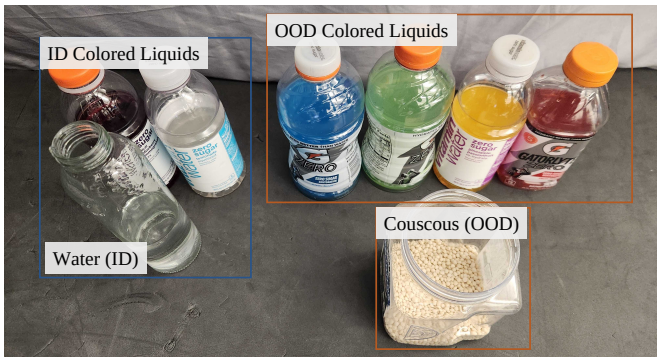


Fig. 6: **Tested Contents.** In-distribution (ID) liquids comprise water and colored liquids; out-of-distribution (OOD) contents comprise colored liquids and couscous. The retail bottles and jar display the contents; the robot uses the small transparent bottle shown in front.

C.3 Spatial Response

The spherical indenter is pressed by hand onto the fingertip. Before every press, with the indenter resting just above the site, we record a two-second no-contact baseline, which removes drift between presses. The site is then loaded to 5 N and held while both layers are recorded for two seconds. Each site is pressed 10 times, and the resulting presses are pooled within each group listed in Table II.

D. Bonus Task: Pour Anything

The salt-pouring task shows that PanoTac uses shear cues to distinguish filled from empty bottles and decide whether to pour, much as humans judge a bottle's contents by lifting it without seeing inside. We next ask whether this capability extends to different visual conditions and contents. As a bonus task, we consider liquid pouring, where spills have greater consequences. We train a policy using 36 demonstrations and evaluate it on unseen liquid colors and unseen couscous, testing whether the learned pour-or-discard behavior generalizes across appearance, density, and flow.

As in salt pouring, the robot must lift a bottle and choose between pouring and discarding. If the bottle contains material, the robot pours it into the receiving beaker and returns the bottle to its starting position; otherwise, it discards the bottle into the basket. Although the bottle is transparent, **the wrist camera cannot see the liquid until after the policy decides to pour**. The pour-or-discard decision therefore has the same visual occlusion as in the salt task. Fig. 6 shows the liquids and couscous used in these trials.

We report all 15 consecutive autonomous trials from a single continuous 9 min 44 s recording, which also captures manual preparation and resets between trials. The website presents these trials in a 2×2 grid, grouped as empty, ID, OOD color, and OOD material. Fig. 7 shows representative frames. We show only one couscous trial, at the end, because the grains cannot be cleaned out for repeated trials, as shown in the continuous recording.

Success requires discarding an empty bottle without pouring, or pouring its contents into the beaker and returning the bottle. Table VI reports successful/recorded trials, pooling OOD liquids and couscous. All 15 trials succeed.

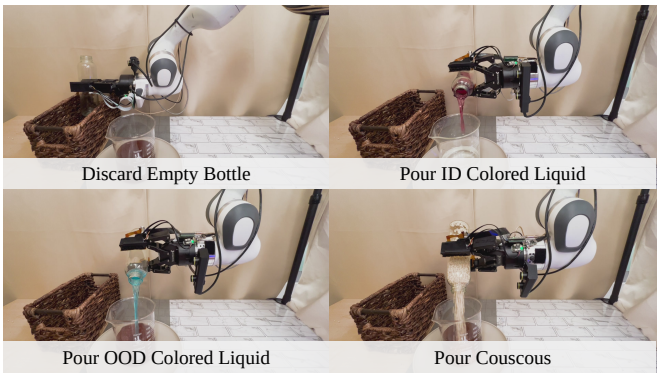


Fig. 7: **Bonus Task Rollouts.** Representative frames show empty-bottle discard, ID and OOD colored-liquid pouring, and couscous pouring.

TABLE VI: **Bonus Task Results.** Successful/recorded trials by condition and overall success rate in the continuous recording. ID comprises two water trials, one dark-red-liquid trial, and one faint-white-liquid trial. OOD pools four colored-liquid trials and one couscous trial.

Policy	**Empty**	**ID**	**OOD**	**Total**
PanoTac	6/6	4/4	5/5	100%