

DEX-Mouse: A Low-cost Portable and Universal Interface with Force Feedback for Data Collection of Dexterous Robotic Hands

Anonymous Authors



Fig. 1: An overview of DEX-Mouse. The proposed USD 150, fully open-sourced teleoperation interface enables scalable collection of high-quality demonstration data. Its calibration-free and highly portable design seamlessly accommodates diverse operators, various target robot hands, and diverse real-world environments. Furthermore, collected demonstrations can be directly deployed to diverse dexterous manipulation tasks.

Abstract—Data-driven dexterous hand manipulation requires large-scale, physically consistent demonstration data. Simulation and video-based methods suffer from sim-to-real gaps and retargeting problems, while MoCap glove-based teleoperation systems require per-operator calibration and lack portability, as the robot hand is typically fixed to a stationary arm. Portable alternatives improve mobility but lack cross-platform and cross-operator compatibility. We present DEX-Mouse, a portable, calibration-free hand-held teleoperation interface with integrated kinesthetic force feedback, built from commercial off-the-shelf components under USD 150. The operator-agnostic design requires no calibration or structural modification, enabling immediate deployment across diverse environments and platforms. The interface supports a configuration in which the target robot hand is mounted directly on the forearm of an operator, producing robot-aligned data. In a comparative user study across various dexterous manipulation tasks, operators using the proposed system achieved an 86.67% task completion rate under the attached configuration. Also, we found that the attached configuration reduced the perceived workload of the operators compared to spatially separated teleoperation setups across all compared interfaces. The complete hardware and software stack, including bill of materials, CAD models, and firmware, is open-sourced at <https://dex-mouse.github.io/> to facilitate replication and adoption.

I. INTRODUCTION

Data-driven approaches such as Imitation Learning (IL) and Vision-Language-Action (VLA) models have shown promise for dexterous hand manipulation. However, collecting high-quality demonstration data at scale remains a key challenge, and the difficulty is compounded by three practical requirements. First, the teleoperation interface must be easy enough to use that human operators can perform

demonstrations repeatedly without undue burden [1]. Second, demonstrations must be physically valid with how the target robot moves and interacts, which is difficult to capture from video or simulation alone [2]. Third, the dataset must have sufficient distributional diversity to support generalization across novel objects and environments [3].

Simulation-based approaches (e.g., [4]) suffer from sim-to-real gaps caused by inaccurate modeling of contact dynamics, friction, and material properties. Video-based approaches (e.g., [5], [6]) face retargeting problems stemming from morphological differences between the human hand and the target robot. Teleoperation via MoCap gloves (e.g., [7], [8]) collects demonstrations on physical robots, ensuring the data captures real contact dynamics and interactions. However, these systems require per-operator calibration. More broadly, conventional teleoperation setups are difficult to relocate. The robot hand, arm, and supporting structures form a bulky integrated system, which strictly limits data collection to fixed environments. To mitigate these mobility and scalability issues, portable hand-held motion-capture interfaces have emerged as a recent alternative ([9], [10]). While these interfaces improve portability, they require per-operator hardware customization due to varying hand sizes and lack compatibility across different robot platforms.

As illustrated in Fig. 1, we propose DEX-Mouse, a low-cost, calibration-free, and portable hand-held teleoperation interface with integrated force feedback. To improve physical alignment between the operator and the robot hand, our interface supports a configuration in which the target robot hand is mounted directly on the forearm of the operator, while a spatially separated teleoperation setup is also available. Also,

its current-based force feedback renders physical resistance for contact-rich manipulation, resulting in improved kinesthetic awareness and more precise force regulation by the operator. This configuration ensures that collected demonstrations directly capture the actual kinematics and dynamics of the target embodiment. Thus, DEX-Mouse produces robot-aligned data through simple proportional retargeting, thereby circumventing complex morphological retargeting. Furthermore, the operator-agnostic design of the interface enables immediate deployment across diverse operators.

DEX-Mouse offers three key advantages over existing interfaces:

- **Universal data-collection system:** An operator- and robot-agnostic teleoperation system accommodates diverse human anatomies and dexterous hands without per-operator calibration or structural modifications, enabling immediate deployment across various platforms.
- **Scalable real-robot data collection:** A self-contained, forearm-mounted design enables deployment in arbitrary environments without requiring a stationary robot arm. This portability reduces laborious work to relocate the system, accelerating the acquisition of large-scale demonstrations.
- **Low-cost open-sourced device:** The complete hardware and software stack, featuring a USD 150 bill of materials, is open-sourced. This cost-effectiveness lowers the barrier to entry for large-scale dexterous manipulation research, democratizing efficient high-quality data collection.

II. RELATED WORK

Simulation environments enable large-scale data collection but introduce sim-to-real gaps due to inaccurate modeling of contact dynamics [4]. Video-based approaches extract demonstrations directly from visual observations [5], [6], [11]–[13], leveraging real-world data. However, contact-rich manipulation frequently causes occlusion, and translating human motion to robotic systems introduces retargeting errors due to morphological differences.

Teleoperation enables direct transfer of human demonstrations to robotic systems. Existing interfaces include vision-based systems [14]–[16], VR/AR controllers [17]–[19], MoCap gloves [8], [20], and exoskeletons [7], [21], which typically retarget human hand motion to the robot via joint-space or vector-based mapping. Vision and VR/AR systems suffer from occlusion during contact-rich tasks, while MoCap gloves and exoskeletons require per-operator calibration and customized manufacturing. More broadly, these hardware-intensive systems involve heavy form factors and restricted mobility, limiting data collection to fixed environments.

Portable hand-held interfaces address the mobility constraints of traditional teleoperation. Operating the end-effector directly eliminates control latency and captures natural physical interactions [1]. Subsequent work extended this paradigm to dexterous manipulation [9], [10]. However, these interfaces require hardware customization for specific operators and robots, and rely on complex post-hoc visual

processing. DEX-Mouse addresses these limitations with an operator-agnostic, cross-embodiment compatible design that operates in real time without per-operator calibration or post-hoc processing.

Across these approaches, a consistent gap remains: no existing system simultaneously achieves portability, operator-agnosticism, cross-embodiment compatibility, and physical consistency of collected data without per-operator calibration or post-hoc processing. DEX-Mouse is designed to address this gap by combining a forearm-mounted, hand-held interface with kinesthetic force feedback in a low-cost, open-sourced package accessible to a broad range of operators and robot platforms.

III. DESIGN PHILOSOPHIES

To overcome the practical bottlenecks characteristic of existing teleoperation systems—namely, limited compatibility across diverse operators and robot platforms, mechanical complexity with high maintenance overhead, and prohibitive equipment costs—we establish three core design principles for DEX-Mouse: universal applicability, mechanical simplicity, and cost-effectiveness.

A. Universal Applicability

Inspired by the standard computer mouse, which functions identically across different computer architectures without hardware modification, we aim to design a device that is agnostic to both operator anatomy and target robot morphology. The goal is to eliminate dependency on anthropometric calibration, enabling true plug-and-play deployability across diverse operators and robot platforms without structural modifications. (This principle is reflected in the design described in Sec. IV-A, IV-B, IV-E and IV-F.)

B. Mechanical Simplicity

Highly articulated haptic mechanisms tend to introduce mechanical fragility and maintenance overhead. We adopt a minimalist design philosophy that reduces potential failure points, ensuring the durability required for long-term data collection. Minimal assembly complexity also allows the broader research community to reproduce, modify, and maintain the device without specialized manufacturing equipment. (This principle is reflected in the design described in Sec. IV-A, IV-B, and IV-C.)

C. Cost-Effectiveness

High-fidelity motion capture gloves and VR controllers often cost thousands of dollars, limiting the number of devices a research laboratory can deploy and restricting the scale and diversity of data collection efforts. We target a design built entirely from off-the-shelf components with a total bill of materials under USD 150, lowering the financial and technical barriers to entry for the broader research community. (This principle is reflected in the design described in Sec. IV-A, IV-B, and IV-C.)

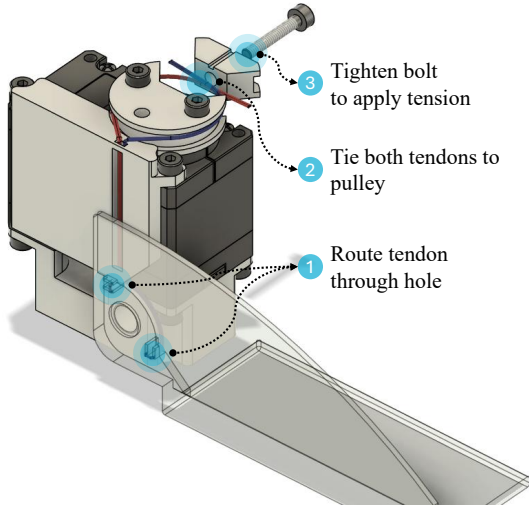


Fig. 2: Dual tendons facilitate bidirectional finger movement. A single thread is folded and assembled in three steps: (1) routing through guide holes, (2) tying to the pulley, (3) tightening via the tensioning bolt.

IV. PROPOSED SYSTEM: DEX-MOUSE

The physical dimensions of the system are derived from the ANSUR II anthropometric dataset [22] to ensure compatibility across diverse operators. The system employs a modular architecture using off-the-shelf components.

A. Tendon-Driven Finger Module

To teleoperate the fingers (excluding the thumb), we employ the Dynamixel XL330-M077-T smart actuator owing to its minimal footprint ($20 \times 34 \times 26$ mm) and high torque transparency.

1) *Anthropometric Compatibility*: The width of the finger module determines the physical compatibility across operators. According to the ANSUR II database, the 50th percentile hand breadth is approximately 90 mm for males and 78 mm for females. The inclusion of structural housing and manufacturing tolerances results in a module width of 23.5 mm, yielding a total system width of 92.8 mm. This total width exceeds the 50th percentile female hand breadth. We handle this geometric discrepancy by omitting rigid Abduction/Adduction (AA) joints. The absence of lateral kinematic constraints allows operators to naturally splay the fingers, mechanically compensating for the width difference and ensuring usability across a wide range of adult hand sizes.

2) *Actuation and Transmission*: To implement force feedback without expensive torque sensors, maximizing the back-drivability of the system is essential. The M077 variant is selected for its low gear ratio (77.5:1), which provides higher force transparency compared to higher-reduction models. The transmission employs an internal routing architecture that translates the rotational motion of the actuator into the linear displacement of the finger links. A spool is directly coupled to the output shaft of the actuator to drive the tendon through a 1:1 mapping.

Within the module housing, guide channels define the routing path to prevent derailment. These channels ensure that the high torque transparency of the actuator is maintained throughout the transmission path while keeping the overall footprint compact. For the tendon material, Polyethylene (PE) braided lines are used to minimize elongation under tension and prevent hysteresis, ensuring accurate transmission of the position of the finger to the robot even after prolonged operation. The complex tendon routing mechanism of the module is illustrated in Fig. 2.

B. Direct-Driven Thumb Module

Unlike the linear arrangement of the finger modules, the anatomical position of the thumb provides sufficient spatial clearance for a direct actuation mechanism. The Flexion/Extension (FE) joint of the thumb module is therefore directly driven by the actuator, ensuring rigid and immediate torque transmission and improving control bandwidth compared to the tendon-driven counterparts.

To capture the Abduction/Adduction (AA) joint motion—which is critical for dexterous manipulation primitives such as opposition and precision grasping [23], [24]—an AS5600 non-contact magnetic encoder is integrated at the base of the AA joint. This provides high-resolution measurement of the orientation of the thumb for mapping the opposition workspace of the human hand to the robot. Owing to spatial constraints within the palm, haptic feedback is provided for the FE joint only, and the AA joint functions as a passive sensing degree of freedom.

C. Embedded Control Architecture

The firmware runs on an STM32F410C8 at a control loop frequency of 100 Hz. Direct Memory Access is used for RS-485 communication with the host PC and UART communication with the actuators to minimize latency. Raw data from the AS5600 magnetic encoder at the AA joint of the thumb is processed via an Exponential Moving Average filter ($\alpha = 0.1$) to suppress sensor noise while maintaining responsiveness.

D. Force Feedback Implementation

The force feedback logic balances transparency during free motion and stiffness during contact using a current-based position control mode, where the actuator acts as a virtual spring.

1) *Dynamic Gain Scheduling*: The stiffness gain is modulated based on the velocity of the finger of the operator to minimize the feeling of mechanical resistance during free motion.

$$K_p(v) = \begin{cases} K_{\text{nominal}} & \text{if } |v| \leq v_{\text{th}} \text{ (Contact/Static)} \\ \gamma K_{\text{nominal}} & \text{if } |v| > v_{\text{th}} \text{ (Free Motion)} \end{cases} \quad (1)$$

where v is the finger velocity, v_{th} is the speed threshold (set to 20 tick / 10 ms, $\approx 175^\circ/\text{s}$), $K_{\text{nominal}} = 5.0$ is the nominal stiffness gain, and $\gamma \in [0, 1]$ is a reduction ratio (set to 0.1 in our firmware). The operator experiences low resistance during rapid movements but feels increased stiffness when grasping an object.

2) *Unidirectional Force Rendering*: In bilateral teleoperation, the robot can impose reverse forces on the hand of the operator when it retracts or opens, disrupting natural force perception. To prevent this, a unidirectional constraint is implemented where feedback force is generated only when the physical finger of the operator penetrates the virtual finger position of the robot (i.e., when the robot is blocked by an object).

$$\tau_{\text{cmd}} = \begin{cases} K_p(v) \cdot (q_{\text{robot}} - q_{\text{operator}}) & \text{if } (q_{\text{robot}} - q_{\text{operator}}) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $\epsilon = 100$ tick ($\approx 8.8^\circ$) is a dead-zone threshold to prevent oscillation. The goal position of the actuator is dynamically updated relative to the current position of the operator, effectively creating a virtual wall that resists further penetration but allows the finger of the operator to withdraw freely.

E. Simple Proportional Retargeting

DEX-Mouse employs simple proportional retargeting. The structural design of the interface naturally aligns the flexion and extension of the finger of the operator with the corresponding actuation modules. Consequently, the measured actuator positions are directly scaled to the joint limits of the target robot hand using simple linear interpolation. This mapping strategy eliminates the need for per-operator anthropometric calibration, reduces computational latency, and ensures real-time responsiveness during highly dynamic manipulation tasks. The system requires no per-operator calibration. For target hands whose finger FE is driven by a single active DoF, the mapping reduces to direct 1:1 scaling. For hands with multiple active FE DoF per finger, a proportional profile is specified once per target hand and reused across operators.

F. Portable System Integration

The portable system integration encompasses all components beyond the hand interface: the ergonomic mounting, pose tracker, visual perception unit, and onboard power supply, collectively enabling untethered data collection across diverse environments.

1) *Ergonomic Mounting Interface*: The data collection system integrates the DEX-Mouse interface and a target robot hand by an operator holding DEX-Mouse while attaching the hand to his/her forearm. The DEX-Mouse is secured to the palm of the operator via adjustable velcro straps. Concurrently, the target robot hand is mounted directly on the forearm of the operator. Placing the heavy robot hand proximally on the forearm, rather than distally near the wrist, reduces the moment of inertia and gravitational torque acting on the shoulder and elbow joints.

2) *Pose Tracking*: Unlike robot arms where end-effector pose can be computed through forward kinematics from joint encoder readings, the forearm-attached nature of DEX-Mouse precludes the use of a known kinematic chain. To obtain the pose of the hand, VIVE Ultimate Tracker [25]

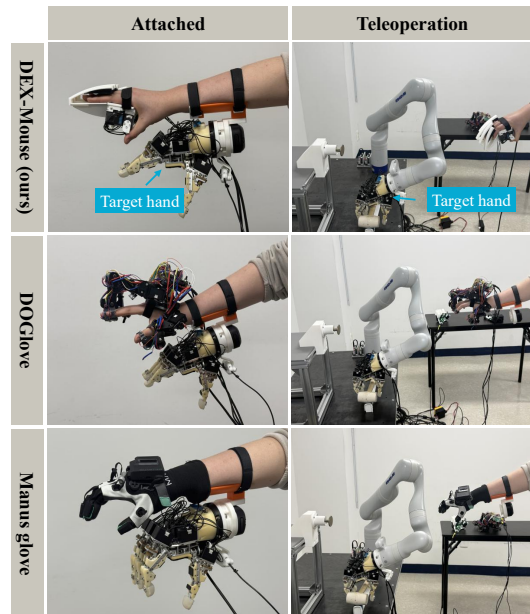


Fig. 3: Attached (left column): The teleoperation interface and the target robot hand are co-located on the arm of the operator. Teleoperation (right column): The operator wears the interface to control a target robot located in a spatially separated, fixed workspace. The rows correspond to the three different interface devices evaluated.

is integrated directly onto the device. The tracker employs inside-out tracking technology to estimate global pose independently, allowing the operator to move freely without any external infrastructure.

3) *Visual Information*: A Logitech HD Webcam C270 camera is mounted under the target robot hand. This camera captures aligned RGB streams at 30 Hz, providing visual context for training manipulation policies.

4) *Self-contained Design*: The entire system is untethered, powered by a portable USB-PD battery, with logging handled on a standard laptop; the full kit fits in a single backpack.

V. EXPERIMENTS

To quantitatively evaluate the proposed interface beyond qualitative demonstrations, we conduct a comparative user study against state-of-the-art devices. To address practical requirements for scalable data collection, we measure task success rate and completion time to verify the efficiency of data collection, alongside perceived workload to ensure minimal operator fatigue. Finally, we train a policy using the collected demonstrations to validate downstream learning utility. The study protocol was IRB-approved.

A. Experimental Setup

1) *Participants*: A total of eight participants over the age of 19 participated in the study, including four males ($M = 27.25$, $SD = 3.96$) and four females ($M = 23.75$, $SD = 1.92$). All participants received 12 US dollars as compensation.

2) *Experimental Design*: Our study employed a 3×2 within-subjects design to examine the effects of interface type and teleoperation configuration. The first factor

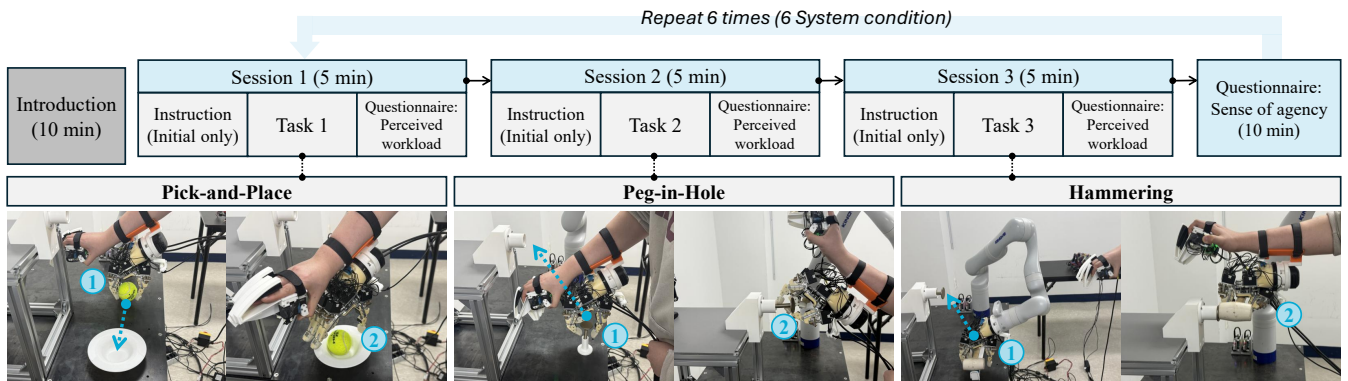


Fig. 4: An overview of the experimental procedure and evaluated manipulation tasks. The study follows a 3×2 within-subjects design across six randomized conditions. The procedural workflow (top) illustrates the sequence for each condition: participants complete one practice trial and five main evaluation trials per task, followed immediately by a perceived workload assessment. A SoA questionnaire concludes each condition. The evaluated tasks (bottom) include Pick-and-Place for basic reach-and-grasp performance, Peg-in-Hole for precise visuomotor coordination and fine force control, and Hammering for grasp stability under dynamic impacts.

was Interface Type, which included DEX-Mouse (proposed), DOGlove [7], and the Manus Quantum Metagloves (hereafter, Manus glove) [26]. The second factor was Collection Configuration, which included Attached, where the interface and the robot-hand tracking frame were physically co-located on the arm of the operator, and Teleoperation, where the interface was worn but the robot operated in a spatially separated fixed workspace (Fig. 3). Each participant therefore experienced six data-collection conditions in total.

Across all experimental conditions, we used a four-fingered Blue Robin dexterous hand [27], where each finger is driven by two active joints. The fingers also employ a spring-loaded multi four-bar linkage mechanism. The design of DEX-Mouse is not specific to this robot hand. The current implementation uses it solely due to hardware availability. For Teleoperation, a Kinova Gen3 7-degree-of-freedom (DoF) manipulator served as the follower arm where the robotic hand was attached. End-effector commands from VIVE Ultimate Tracker were resolved into joint targets via Pinocchio inverse kinematics solver [28] at 20Hz. In Attached, the target robot hand was directly mounted on the forearm of the operator.

3) *Experimental Procedure*: Participants were first briefed about the study and provided informed consent. They then experienced six system conditions in a randomized order, following the experimental procedure detailed in Fig. 4. For conditions utilizing the DOGlove and Manus glove, a per-operator calibration was performed prior to the trials. In the first system condition, participants watched an instructional video that explained how to perform the tasks. They then completed one practice trial and five main trials for each task. In the remaining system conditions, participants completed the same sequence of tasks with one practice trial followed by five main trials for each task, without additional instructions. We instructed all participants to complete each task as quickly and accurately as possible.

Immediately after completing the five main trials for a given task, participants assessed their perceived workload using NASA Raw Task Load Index (NASA-RTLX) [29].

Once all 15 main trials for a system condition were finished, participants completed a Sense of Agency (SoA) questionnaire [30] during a rest period before moving to the next condition. This questionnaire evaluates the subjective feeling of controlling the actions of the robot.

Participants completed three tasks in a fixed order across all system conditions: pick-and-place, peg-in-hole, and hammering (Fig. 4). These tasks were strategically selected to evaluate a comprehensive range of dexterous manipulation capabilities, encompassing distinct grasp taxonomies and dynamic force requirements.

- **Pick-and-Place**: Grasp a spherical ball and transfer it to a designated receptacle. This task evaluated basic reach-and-grasp performance and spatial positioning. A trial was marked successful if the ball was stably deposited inside the plate.
- **Peg-in-Hole**: Grasp a nail using a precision pinch grip and insert it into a fitted hole. This task assessed fine visuomotor coordination and multi-finger force regulation. A trial was considered successful when the nail was inserted into the hole.
- **Hammering**: Grasp a hammer using a stable power grasp and strike a pre-positioned nail. This task evaluated grasp stability and interface transparency under dynamic impacts. Success was defined as driving the nail flush with the surface while maintaining a stable, slip-free grasp on the hammer.

4) *Measures*: We evaluated system performance using objective and subjective measures. The objective measures were task success rate across the five main trials for each task and task completion time (recorded only for successful trials). The subjective measures were perceived workload, assessed using NASA-RTLX (six items, Cronbach's $\alpha = .83$), and SoA (11 items, Cronbach's $\alpha = .67$). All items were assessed using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

TABLE I: Detailed Task Performance ($N = 8$). Success Rate (SR) is measured in %, and Completion Time (Time) in seconds. All values are presented as Mean (SD). The bottom rows present the overall average of each configuration across all interfaces.

Interface Type	Collection Configuration	Pick & Place		Peg-in-Hole		Hammering		Overall Average	
		SR (SD)	Time (SD)	SR (SD)	Time (SD)	SR (SD)	Time (SD)	SR (SD)	Time (SD)
DEX-Mouse (Ours)	Attached	95.0 (13.23)	5.57 (2.46)	72.5 (28.17)	14.29 (6.41)	92.5 (13.92)	10.29 (3.86)	86.67 (22.11)	10.05 (5.78)
	Teleoperation	80.0 (20.00)	17.80 (4.98)	17.5 (27.27)	25.21 (3.87)	60.0 (31.62)	17.13 (7.37)	52.5 (37.33)	18.77 (6.56)
DOGlove	Attached	100.0 (0.00)	7.53 (2.23)	60.0 (30.00)	16.19 (5.79)	72.5 (33.07)	11.24 (3.31)	77.5 (30.72)	11.67 (5.47)
	Teleoperation	80.0 (24.49)	19.79 (6.24)	12.5 (13.92)	27.44 (6.97)	57.5 (32.31)	22.73 (6.91)	50.0 (37.42)	22.48 (7.24)
Manus glove	Attached	92.5 (13.92)	7.97 (3.59)	32.5 (26.34)	17.22 (4.84)	62.5 (29.05)	11.93 (4.34)	62.5 (34.31)	11.93 (5.59)
	Teleoperation	62.5 (33.82)	14.59 (3.25)	2.5 (6.61)	30.00 (0.00)	45.0 (23.98)	17.93 (4.77)	36.67 (34.96)	17.01 (5.39)
Configuration Average	Attached	95.83 (11.52)	7.02 (3.01)	55.00 (32.79)	15.90 (5.91)	75.83 (29.43)	11.15 (3.94)	75.56 (31.13)	11.22 (5.68)
	Teleoperation	74.17 (27.98)	17.39 (5.42)	10.83 (19.13)	27.55 (5.68)	54.17 (30.27)	19.26 (6.91)	46.39 (37.24)	19.42 (6.89)

B. Results and Findings

We evaluated six conditions defined by collection configuration and interface type across the three tasks. For each task, we measured task success rate and task completion time. We also assessed perceived workload and SoA.

1) *Impact of Collection Configuration*: The physical setup consistently affected teleoperation performance. As summarized in the bottom rows of Table I, Attached yielded a higher overall success rate and a faster completion time compared to Teleoperation.

This performance gap was observed across all manipulation tasks. In the pick-and-place task, Attached demonstrated higher spatial positioning efficiency. This performance difference became more pronounced in tasks requiring precise physical interactions. In the peg-in-hole task, Teleoperation resulted in lower success rates, likely due to the mental coordinate transformations required between the interface of the operator and the remote robotic workspace. In contrast, Attached mitigated this spatial discrepancy, improving force regulation. Similarly, in the hammering task, Attached outperformed the remote setup. These results indicate that co-locating the robotic workspace with the proprioception of the operator enables reflexive motor responses and reduces the workload of the operator.

2) *Impact of Interface Type*: As shown in Table I and Fig. 5 (focusing on Attached, which outperformed Teleoperation), DEX-Mouse achieved the highest overall success rate and fastest completion time among the three tested interfaces. This performance was achieved with a total component cost of under USD 150, compared to the commercial Manus glove (retail price \approx USD 7,000) and the research-grade DOGlove (manufacture price \approx USD 600).

In the pick-and-place task, all three interfaces achieved high success rates, indicating sufficient visual-motor mapping for basic reaching. However, the performance gap widened during contact-rich tasks. In the peg-in-hole task, the Manus glove resulted in lower success rates, likely due to the absence of kinesthetic feedback. While DOGlove provides haptic feedback, its full high DoF tracking occasionally induced unintended finger micro-motions. In contrast, the structurally constrained 6-DoF design of DEX-Mouse effectively suppressed unnecessary movements, resulting in a higher success rate. Per-participant inspection further revealed that most of the within-cell variance in peg-in-hole performance across all three interfaces originated from two participants who self-reported difficulty with precise thumb-

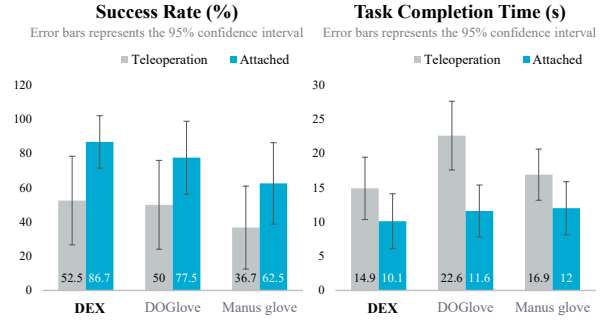


Fig. 5: Overall task success rates (%) and completion times (s) across different interfaces and collection configurations.

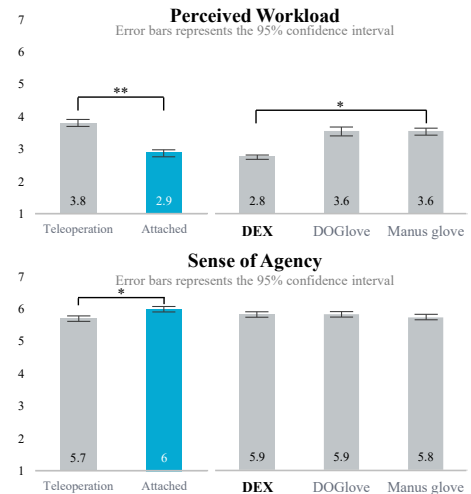


Fig. 6: Subjective operator experience results evaluating perceived workload and SoA. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

index opposition due to their short thumb length. In the hammering task, the simplified structure of DEX-Mouse enabled the operator to concentrate on thumb engagement for a secure power grasp, while the kinesthetic force feedback provided physical resistance to prevent the tool from slipping.

3) *Operator Experience*: Fig. 6 illustrates the subjective questionnaire results for perceived workload and SoA. For perceived workload, we found significant main effects of collection configuration [$F(1, 7) = 24.453, p = 0.002, \eta_p^2 = 0.777$] and interface type [$F(2, 14) = 7.720, p = 0.005, \eta_p^2 = 0.524$]. The interaction between collection configuration and interface type was not significant [$F(2, 14) = 1.952, p = 0.218, \eta_p^2 = 0.218$]. Bonferroni post hoc tests showed that perceived workload was significantly

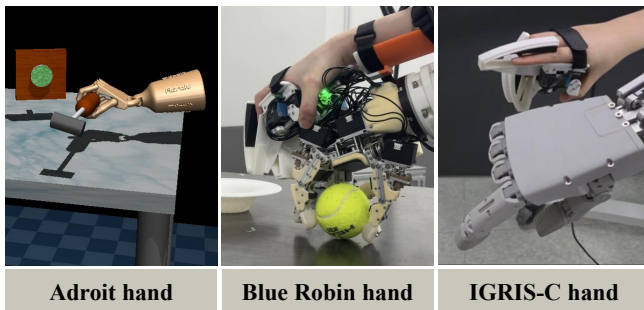


Fig. 7: Demonstration of cross-embodiment compatibility. By employing simple proportional retargeting, DEX-Mouse enables calibration-free teleoperation across distinct robot morphologies: (left) a simulated 5-finger active 30-DoF Adroit hand, (center) a 4-finger 8-DoF Blue Robin hand, and (right) a 5-finger 11-DoF (active 6-DoF) humanoid hand (IGRIS-C).

lower in the Attached condition than in the Teleoperation condition (95% CI [0.438, 1.242], $p < 0.01$). In addition, perceived workload was significantly lower with DEX-Mouse than with the Manus glove (95% CI [0.085, 0.360], $p < 0.05$). No significant differences were found for the other pairwise comparisons.

For SoA, we found a significant main effect of collection configuration [$F(1, 7) = 6.881, p = 0.034, \eta_p^2 = 0.496$]. The main effect of interface type was not significant [$F(2, 14) = 0.588, p = 0.569, \eta_p^2 = 0.077$], and the interaction between collection configuration and interface type was also not significant [$F(2, 14) = 0.145, p = 0.866, \eta_p^2 = 0.020$]. Bonferroni post hoc tests showed that SoA was significantly lower in the Teleoperation condition than in the Attached condition (95% CI [0.025, 0.490], $p < 0.05$). No other pairwise comparisons showed significant differences.

C. Cross-Embodiment Compatibility

To validate cross-embodiment compatibility, we deployed DEX-Mouse on a simulated 5-finger 30-DoF Adroit hand [31] and a physical 5-finger 11-DoF humanoid hand [32]. As illustrated in Fig. 7, using simple proportional retargeting without per-operator calibration or structural modifications, an operator successfully performed complex multi-finger manipulations on both platforms. These qualitative results demonstrate the universal applicability of our interface.

D. Downstream Policy Learning

To verify that demonstrations collected with DEX-Mouse are suitable for policy training, we trained a diffusion policy [9] across the three evaluated tasks: pick-and-place, peg-in-hole, and hammering. For each task, we collected 200 demonstrations under Attached configuration. The data collection process required approximately 1–1.5 hours per task, demonstrating that hundreds of demonstrations can be collected efficiently.

The policy architecture follows the DexUMI framework with two modifications. First, while DexUMI relies on relative positional representations (mainly due to the differences between DexUMI and the target hand), our policy uses absolute joint positions for the robot hand. This absolute mapping

is feasible because DEX-Mouse captures demonstrations directly on the target embodiment, inherently eliminating the morphological gap. Second, instead of using external tactile sensors, we incorporated the measured motor torque values as proprioceptive inputs to represent contact interactions.

To evaluate the robustness of the learned policy, we attached the hand to an FR5 [33] cobot to execute tasks. We measured the task success rate over 20 trials for each task. During these trials, the initial poses of the objects were randomized to assess spatial generalization. For the pick-and-place task, the positions of the objects were randomly varied within a 10×10 cm workspace around their nominal initial states, preserving the relative spatial arrangement between the ball and the receptacle. For the peg-in-hole task, the initial position of the nail was randomized within a 5×5 cm region, while the hole remained fixed. Similarly, for the hammering task, the target objects remained fixed, whereas the initial position of the robot hand was randomized within a $10 \times 10 \times 10$ cm volume. Under these randomized conditions, the learned policy achieved success rates of 90% for the pick-and-place task, 50% for the peg-in-hole task, and 95% for the hammering task. These results demonstrate that the proposed interface provides robust, robot-aligned data for downstream imitation learning. The 50% success rate in the peg-in-hole task reflects the compound difficulty of executing a precise pinch grasp followed by a millimeter-tolerance insertion. Under initial spatial randomization, successfully chaining these sequential precision tasks relying solely on a single camera and proprioceptive torque feedback remains inherently challenging.

VI. LIMITATIONS AND FUTURE WORK

While DEX-Mouse demonstrates high task success rates in dexterous teleoperation, future work will focus on three key areas to further enhance system capabilities and scalability.

First, the current calibration-free design omits rigid AA joints to ensure an operator-agnostic fit. To capture lateral finger motion during tasks requiring substantial finger splay without compromising this universality, we plan to incorporate passive flexure sensors. This addition aims to further improve the kinematic precision of the interface.

Second, while the forearm-mounted configuration mitigates wrist load, long-duration data collection requires further strain reduction. Future iterations will explore lightweight composite materials and active gravity compensation to improve wearability.

Finally, the variance observed in peg-in-hole performance was primarily driven by two participants who self-reported difficulty with precise thumb-index opposition due to short thumb length. Motivated by this observation, future iterations will incorporate an adjustable-length thumb module. The limited sample size ($N = 8$) also renders formal assumption checks underpowered, so reported effect sizes should be interpreted as upper bounds.

VII. CONCLUSION

This paper presented DEX-Mouse, a low-cost, portable, and universal interface designed for dexterous manipulation

data collection. By abstracting the kinematics of the human hand into a normalized tendon-driven mechanism, the system achieves a calibration-free architecture that seamlessly accommodates diverse human anatomies and robot morphologies. The integration of sensor-less kinesthetic force feedback and a forearm-mounted configuration significantly improved task success rates and reduced operator workload during contact-rich manipulations. Empirical evaluations showed that DEX-Mouse attains task performance competitive with significantly more expensive commercial MoCap gloves.

Ultimately, DEX-Mouse lowers the technical and financial barriers to entry for dexterous manipulation research. With an open-source design and a total component cost under USD 150, the proposed interface enables the scalable collection of high-fidelity, robot-aligned demonstration data. This accessibility will facilitate the rapid advancement of data-driven robotics across diverse, real-world environments.

REFERENCES

- [1] C. Chi, Z. Xu, C. Pan, E. Cousineau, B. Burchfiel, S. Feng, R. Tedrake, and S. Song, "Universal manipulation interface: In-the-wild robot teaching without in-the-wild robots," in *Proc. of Robot. Sci. Syst. (RSS)*, 2024.
- [2] T. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," in *Proc. of Robot. Sci. Syst. (RSS)*, 2023.
- [3] H. R. Walke, K. Black, T. Z. Zhao, Q. Vuong, C. Zheng, P. Hansen-Estruch, A. W. He, V. Myers, M. J. Kim, M. Du, A. Lee, K. Fang, C. Finn, and S. Levine, "Bridgedata v2: A dataset for robot learning at scale," in *Proc. Conf. Robot Learn. (CoRL)*, 2023, pp. 1723–1736.
- [4] A. Maddukuri, Z. Jiang, L. Y. Chen, S. Nasiriany, Y. Xie, Y. Fang, W. Huang, Z. Wang, Z. Xu, N. Chernyadev, S. Reed, K. Goldberg, A. Mandlekar, L. Fan, and Y. Zhu, "Sim-and-real co-training: A simple recipe for vision-based robotic manipulation," in *Proc. of Robot. Sci. Syst. (RSS)*, 2025.
- [5] Y. Hu, Y. Guo, P. Wang, X. Chen, Y.-J. Wang, J. Zhang, K. Sreenath, C. Lu, and J. Chen, "Video prediction policy: A generalist robot policy with predictive visual representations," in *Proc. of Int. Conf. Mach. Learn. (ICML)*, 2025, pp. 24328–24346.
- [6] W. Ye, F. Liu, Z. Ding, Y. Gao, O. Rybkin, and P. Abbeel, "Video2policy: Scaling up manipulation tasks in simulation through internet videos," *arXiv preprint arXiv:2502.09886*, 2025.
- [7] H. Zhang, S. Hu, Z. Yuan, and H. Xu, "Doglove: Dexterous manipulation with a low-cost open-source haptic force feedback glove," in *Proc. of Robot. Sci. Syst. (RSS)*, 2025.
- [8] K. Shaw, Y. Li, J. Yang, M. K. Srirama, R. Liu, H. Xiong, R. Mendonca, and D. Pathak, "Bimanual dexterity for complex tasks," in *Proc. Conf. Robot Learn. (CoRL)*, 2025, pp. 5166–5183.
- [9] M. Xu, H. Zhang, Y. Hou, Z. Xu, L. Fan, M. Veloso, and S. Song, "Dexumi: Using human hand as the universal manipulation interface for dexterous manipulation," in *Proc. Conf. Robot Learn. (CoRL)*, 2025, pp. 437–459.
- [10] H.-S. Fang, B. Romero, Y. Xie, A. Hu, B.-R. Huang, J. Alvarez, M. Kim, G. Margolis, K. Anbarasu, M. Tomizuka *et al.*, "Dexop: A device for robotic transfer of dexterous human manipulation," *arXiv preprint arXiv:2509.04441*, 2025.
- [11] S. Nair, A. Rajeswaran, V. Kumar, C. Finn, and A. Gupta, "R3m: A universal visual representation for robot manipulation," in *Proc. Conf. Robot Learn. (CoRL)*, 2023, pp. 892–909.
- [12] Y. J. Ma, S. Sodhani, D. Jayaraman, O. Bastani, V. Kumar, and A. Zhang, "VIP: Towards universal visual reward and representation via value-implicit pre-training," in *Proc. of Int. Conf. Learn. Represent. (ICLR)*, 2023.
- [13] I. Radosavovic, T. Xiao, S. James, P. Abbeel, J. Malik, and T. Darrell, "Real-world robot learning with masked visual pre-training," in *Proc. Conf. Robot Learn. (CoRL)*, 2023, pp. 416–426.
- [14] A. Handa, K. Van Wyk, W. Yang, J. Liang, Y.-W. Chao, Q. Wan, S. Birchfield, N. Ratliff, and D. Fox, "Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system," in *Proc. of IEEE Int. Conf. Robot. Automat. (ICRA)*, 2020, pp. 9164–9170.
- [15] Y. Qin, W. Yang, B. Huang, K. Van Wyk, H. Su, X. Wang, Y.-W. Chao, and D. Fox, "Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system," in *Proc. of Robot. Sci. Syst. (RSS)*, 2023.
- [16] S. Yang, M. Liu, Y. Qin, R. Ding, J. Li, X. Cheng, R. Yang, S. Yi, and X. Wang, "Ace: A cross-platform and visual-exoskeletons system for low-cost dexterous teleoperation," in *Proc. Conf. Robot Learn. (CoRL)*, 2025, pp. 4895–4911.
- [17] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang, "Open-television: Teleoperation with immersive active visual feedback," in *Proc. Conf. Robot Learn. (CoRL)*, 2025, pp. 2729–2749.
- [18] R. Ding, Y. Qin, J. Zhu, C. Jia, S. Yang, R. Yang, X. Qi, and X. Wang, "Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning," in *Proc. of IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2025, pp. 12248–12255.
- [19] A. Iyer, Z. Peng, Y. Dai, I. Guzey, S. Halder, S. Chintala, and L. Pinto, "Open teach: A versatile teleoperation system for robotic manipulation," in *Proc. Conf. Robot Learn. (CoRL)*, 2025, pp. 2372–2395.
- [20] C. Wang, H. Shi, W. Wang, R. Zhang, L. Fei-Fei, and C. K. Liu, "Dexcap: Scalable and portable mocap data collection system for dexterous manipulation," *arXiv preprint arXiv:2403.07788*, 2024.
- [21] E. Sung, S. You, S. Moon, J. Kim, and J. Park, "Snu-avatar haptic glove: Novel modularized haptic glove via trigonometric series elastic actuators," in *Proc. of IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2024, pp. 3573–3580.
- [22] C. C. Gordon, C. L. Blackwell, B. Bradtmiller, J. L. Parham, P. Barrientos, S. P. Paquette, B. D. Corner, J. M. Carson, J. C. Venezia, B. M. Rockwell *et al.*, "2012 anthropometric survey of U.S. army personnel: Methods and summary statistics," U.S. Army Natick Soldier Research, Development and Engineering Center, Natick, MA, Tech. Rep. NATICK/TR-15/007, 2014.
- [23] M. R. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Trans. on Robot. and Automat.*, vol. 5, no. 3, pp. 269–279, 1989.
- [24] T. Feix, J. Romero, H.-B. Schmiemayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy: Orthogonality status in human grasp types," *IEEE Trans. on Human-Machine Syst.*, vol. 46, no. 1, pp. 66–77, 2015.
- [25] HTC Corporation, "VIVE ultimate tracker," [Online]. Available: <https://www.vive.com/kr/accessory/vive-ultimate-tracker/>. Accessed: Apr. 4, 2026.
- [26] Manus Meta, "Quantum metagloves," [Online]. Available: <https://www.manus-meta.com/products/quantum-metagloves>. Accessed: Feb. 12, 2025.
- [27] Blue Robin, "Blue Robin dexterous robot hand," [Online]. Available: https://www.bluerobin.co.kr/EN/include/pages/products/robot_hand.php. Accessed: Feb. 12, 2025.
- [28] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiraux, O. Stasse, and N. Mansard, "The Pinocchio C++ library – a fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Paris, France, Jan. 2019.
- [29] S. G. Hart, "Nasa-task load index (NASA-TLX); 20 years later," in *Proc. of Hum. Factors Ergonom. Soc. Annu. Meet. (HFES)*, vol. 50, no. 9, 2006, pp. 904–908.
- [30] A. Tapal, E. Oren, R. Dar, and B. Eitam, "The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment," *Frontiers in Psychology*, vol. 8, p. 1552, 2017.
- [31] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," in *Proc. of Robot. Sci. Syst. (RSS)*, 2018.
- [32] ROBROS, "IGRIS-C: Compact artificial intelligence humanoid robot," [Online]. Available: <https://robros.co.kr/igrisc/>. Accessed: Apr. 4, 2026.
- [33] FAIR Innovation, "FR5 collaborative robot," [Online]. Available: <https://www.frtech.fr/FR/4.html>. Accessed: Apr. 14, 2026.