

Gloivity: Learning Dexterous Contact-Rich Manipulation via Spatial Wrench Feedback Teleoperation System

Yuyang Gao, Haofei Ma and Pai Zheng*

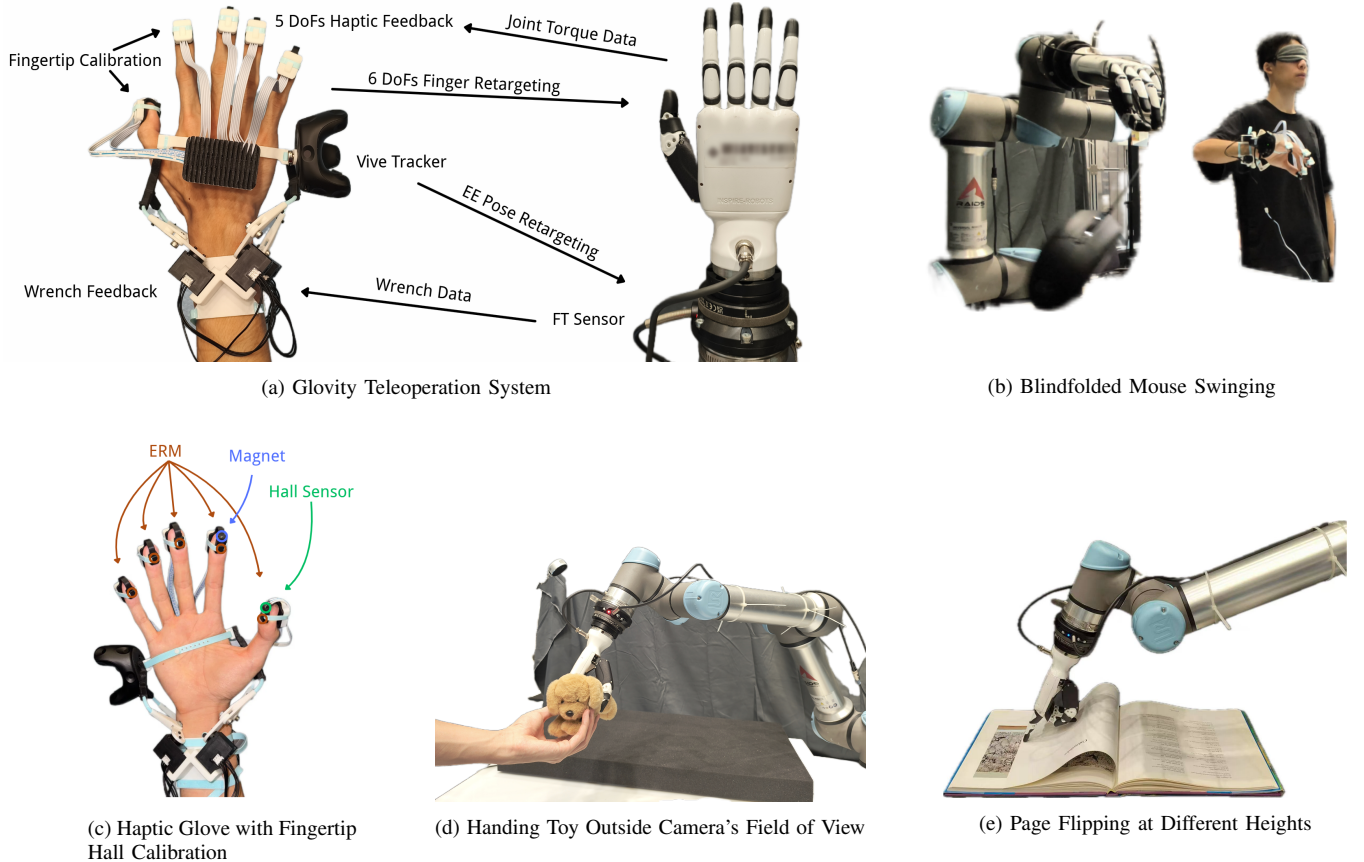


Fig. 1: Overview of Gloivity. (a) and (c) illustrates the overall system architecture, comprising a wrist-mounted wrench feedback system and a haptic glove, designed to facilitate intuitive force feedback and precise teleoperation. (b) Performance of the wrench feedback in high-dynamic scenarios, enabling the operator to accomplish mouse swinging tasks relying solely on the provided feedback. (d) and (e) showcase two imitation learning tasks, demonstrating the capability of the Gloivity system to collect high-quality data and its potential for imitation learning with force.

Abstract—We present Gloivity, a novel, low-cost wearable teleoperation system that integrates a spatial wrench (force-torque) feedback device with a haptic glove featuring fingertip Hall sensor calibration, enabling feedback-rich dexterous manipulation. Gloivity addresses key challenges in contact-rich tasks by providing intuitive wrench and tactile feedback, while overcoming embodiment gaps through precise retargeting. User studies demonstrate significant improvements: wrench feedback boosts success rates in book-flipping tasks from 48% to 78% and reduces completion time by 25%, while fingertip calibration enhances thin-object grasping success significantly compared to commercial glove. Furthermore, incorporating wrench signals into imitation learning (via DP-R3M) achieves high success rate in novel contact-rich scenarios, such as adaptive page flipping and force-aware handovers. All hardware designs, software will be open-sourced. Project website: gloivity.github.io

I. INTRODUCTION

Robotic teleoperation has emerged as a cornerstone for enabling robots to perform complex, contact-rich manipulation tasks, bridging the gap between human dexterity and robotic precision. Recent advancements in imitation learning have demonstrated the potential of human demonstrations to train robots for tasks requiring intricate interactions. However, existing teleoperation systems face significant challenges that limit their efficacy in such scenarios. First, the lack of multimodal feedback—beyond visual cues—restricts operators’ ability to perceive and adjust to real-time force and tactile interactions, which are critical for tasks involving delicate or dynamic contact [1], [2]. Second, structural disparities between human and robotic manipulators, coupled

with retargeting latency, introduce an embodiment gap that complicates precise control [3]. These limitations result in inefficient data collection and suboptimal performance in dexterous manipulation tasks.

To address these challenges, we introduce Glovity, a novel, low-cost, and open-source teleoperation system designed to enhance contact-rich manipulation through spatial wrench feedback and haptic glove integration. Glovity combines a wearable wrench feedback mechanism, which provides intuitive force and torque cues without restricting arm mobility, with a haptic glove equipped with fingertip Hall calibration for precise grasping. Costing less than \$300 and utilizing 3D-printed components, Glovity is accessible and replicable, requiring only a few hours for assembly. Unlike traditional systems that rely heavily on expensive haptic devices or bulky exoskeletons [4], [5], Glovity achieves effective feedback and efficient retargeting through innovative hardware and algorithmic designs, ensuring compatibility with diverse robotic platforms.

The main contribution of this work are as follows:

- **Wearable Spatial Wrench Feedback System:** We develop a compact, wearable system that delivers real-time, multi-dimensional force-torque feedback, enabling intuitive operator adjustments in contact-rich tasks without compromising arm mobility.
- **Haptic Glove with Fingertip Hall Calibration:** We propose a haptic glove integrated with linear Hall sensors for precise fingertip calibration, significantly improving the accuracy of grasping thin objects by mitigating structural misalignments.
- **Efficient Demonstration Data for Imitation Learning:** We demonstrate Glovity’s capability to collect high-quality demonstrations, incorporating wrench signals into diffusion-based imitation learning policies to achieve robust performance in novel, contact-rich tasks.

Through comprehensive user studies and imitation learning experiments, we demonstrate Glovity’s ability to provide intuitive and effective feedback for precise control in contact-rich and high-dynamic tasks, as well as its efficient grasping capabilities for thin objects. All hardware and software will be open-sourced, fostering community-driven innovation in robotic learning.

II. RELATED WORKS

A. Teleoperation with End-effector Force Feedback

Currently, the majority of teleoperation systems rely solely on visual feedback, including leader-follower arm setups [6], [7], VR-based interfaces [8], [9], and camera-based tracking [10], [11]. This exclusive dependence on visual cues renders these systems inadequate for precise contact-rich tasks. Some works have improved upon leader-follower arm systems by incorporating bilateral force feedback through methods like shared joint torque [4], [12], [13], [14]. However, these systems are typically tailored to specific robots and, due to structural constraints, are challenging to adapt

for dexterous manipulation. Other approaches utilize haptic devices to enable force feedback teleoperation [15], [5], [16], but such devices are typically expensive and have limited operation space, making them difficult to apply to a wide range of tasks. In contrast, Glovity’s wrench feedback system achieves low-cost implementation while being wearable on the human wrist and compatible with dexterous operations, providing operators with highly enriched feedback.

B. Dexterous Hand Teleoperation

Efficient teleoperation of dexterous hands has long been a challenge in robot learning, primarily due to structural differences between robotic and human hands, which necessitate effective mapping algorithms to bridge the gap. Several approaches have been developed to address this. Vision-based systems, such as AnyTeleop [10], achieve effective hand mapping through a flexible framework supporting diverse robotic arms, hands, and camera setups. Bunny-VisionPro [17] leverages VR devices and haptic feedback to provide real-time bimanual dexterous control, enhancing immersion for imitation learning. However, vision-based methods often suffer from occlusion issues and increased latency due to real-time pose estimation. Alternatively, some works integrated commercial gloves [18], [19], [20], and there are also Low-cost alternatives, such as DOGlove [21], employ an exoskeleton-encoder scheme with motion capture system and cable-driven torque feedback, achieving precise hand teleoperation and fingertip positioning. Yet, such exoskeleton-based designs can be bulky, compromising user comfort. In contrast, Glovity utilizes an IMU+Hall sensor approach, combining minimal components with innovative structural and algorithmic designs to achieve efficient hand mapping and precise fingertip calibration. This wearable, low-latency solution avoids occlusion issues and bulky hardware, offering enhanced performance for contact-rich dexterous tasks while maintaining low-cost and adaptability.

C. Imitation learning with Force/Tactile Data

Research on imitation learning has been around for many years, originating from foundational works in behavioral cloning and apprenticeship learning in the early 2000s. As the tasks assigned by humans to robots become progressively more complex and involve richer contact interactions and longer time horizons, integrating force/tactile signals into imitation learning algorithms has become an irreversible trend [22], [23], [24], [14], [25]. Among these, RDP [24] incorporates force/tactile as conditional inputs while employing a fast-slow architecture to fine-tune actions, delivering impressive results in contact-rich tasks. ForceMimic [23] uses a handheld gripper for data collection, predicts position-force pairs, and uses hybrid position-force control, excelling in the peeling task. The above methods show great potential for incorporating force/tactile sensing into imitation learning, but unfortunately they are not compatible with dexterous manipulation. In this project, we incorporate wrench data as

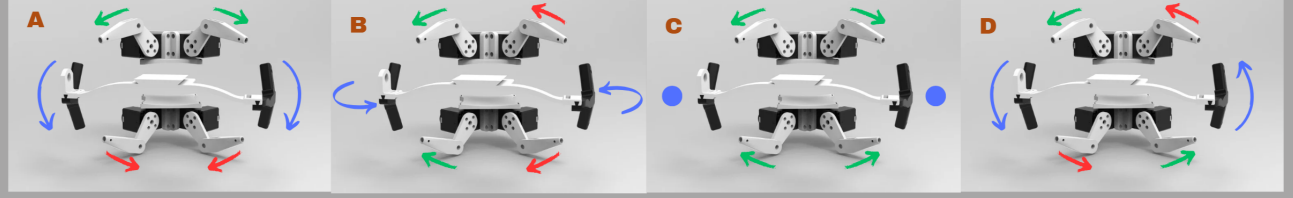


Fig. 2: Illustration of the structure and four base feedback modes of the wrench feedback system in response to base components. (A) Represents the force applied along the x-axis, inducing a Vertical motion pattern. (B) Depicts the force along the y-axis, driving a horizontal motion. (C) Indicates the force along the z-axis, eliciting a forward-backward motion. (D) Shows the torque around the z-axis, resulting in a rotational response.

observations into the DP-R3M model, exploring novel scenarios and further demonstrating the potential of multimodal imitation learning in dexterous manipulation.

III. TELEOPERATION SYSTEM DESIGN

A. Design Philosophy

Glovity is designed to provide rich, intuitive feedback along with excellent retargeting performance for dexterous manipulation. It significantly enhances the efficiency and quality of data collection while ensuring the entire system is low-cost, easy to replicate and assembly. Glovity's design adheres to the following principles:

- 1) **High Cost-Performance:** Glovity's core goal is to be affordable and highly effective. By using a minimal number of commercial high-quality servos and sensors, combined with 3D-printed parts, we keep the total cost around \$300 (\$230 for wrench feedback and \$70 for haptic glove). Through ingenious structural and algorithmic designs, it achieves rich feedback, efficient retargeting and calibration, delivering impressive performance in real-world use.
- 2) **Ease of Replication and Assembly:** All electronic and structural components of Glovity can be easily obtained online and via 3D printing. Additionally, it adopts a Lego-style component design, allowing the entire assembly process to be completed in just a few hours with a few screws.
- 3) **Versatility and modularity:** Glovity fits hands of various sizes and suitable for different robots and dexterous hands, offering comfortable wear and low latency (up to 100Hz for wrench feedback, retargeting, and tactile feedback) to handle diverse manipulation tasks. Glovity also adopts a modular design, allowing the wrench feedback and the tactile glove to be used separately, making it adaptable to various scenarios.

In summary, through these principles, Glovity offers an efficient and economical solution for dexterous operations, suitable for robotics researchers and developers.

B. Wrench Feedback System

Glovity's wrench feedback system is designed to provide intuitive wrench feedback in a compact, ergonomic design.

For practical considerations, the system is designed to provide feedback in **four primary dimensions**, with forces and torques along the x and y axes (where the x-axis passes through the palm, while the y-axis represents the axis of finger flexion) implemented using a fused feedback approach (detailed in later section). This design accounts for the fact that, in most manipulation tasks, forces and torques at the robot's end-effector along the x and y axes tend to occur simultaneously and in the same direction.

1) *Hardware and Mechanism Design:* The system employs four XL330 servomotors, selected for their high torque and rapid response, coupled with a dual-linkage mechanism to convert rotational motion into linear displacement and rotation of a palm-mounted fixator, this fixator is connected via a TPU-printed linkage, which minimizes restrictions on wrist movement while maintaining effective force feedback. The main body of the system is 3D-printed using PLA for structural rigidity, complemented by a TPU strap to securely fasten the device to the user's forearm.

The servomotors are arranged in pairs above and below the distal end of the forearm, with adjacent motor pairs oriented at a 90-degree angle, as illustrated in Figure 2. This configuration is driven by two key design considerations:

- **Compact Multi-Directional Feedback:** The orthogonal servomotor arrangement enables the encoding of forces and torques from multiple directions within a compact form factor, enhancing the realism of haptic feedback.
- **Ergonomic Optimization:** Positioning servomotors above and below aligns with the elliptical shape of the human forearm, reducing pressure points and improving wearability.

2) *Wrench Feedback Mapping:* The primary principle of the wrench feedback is to control the rotation of servomotors to corresponding positions based on real-time 6D force-torque data acquired from the FT sensor mounted at the robot's end-effector.

For each servomotor $i \in \{1, 2, 3, 4\}$, the mapping algorithm can be expressed as:

$$\theta_i = \theta_{\text{init}_i} + \sum_{j \in \{f_y, f_x, f_z, \tau_z\}} w_j \cdot \kappa_j \cdot s_{j_i} \cdot c_{j_i} \cdot j \quad (1)$$

where θ_{init_i} is the initial/no-load angle for servomotor i ; w_j is

the weight of component j , based on its proportion among all components; κ_j is the sensitivity of component j , manually set based on task requirements; s_{j_i} is the sign of component j for servomotor i , determined by the placement of the servo; j represents the force/torque component, where f_y , f_x , f_z represent forces (N) and τ_z represents torque (Nm); and c_{j_i} is the dynamic coefficient for component j on servomotor i , defined as:

$$c(r) = \max(c_{\min}, 1 + \sigma \cdot \kappa_r \cdot (r + \delta)) \quad (2)$$

with $r \in \{\tau_x/f_y, \tau_y/f_x\}$ being the torque-force ratio; κ_r the sensitivity of the ratio, manually set based on task requirements; σ the sign factor (± 1), determining the direction of adjustment based on the ratio type; δ the offset parameter, adjusting the pivot point of the ratio's effect, typically set based on system geometry (e.g., distance from hand center to wrist); and $c_{\min}$ the minimum coefficient value, ensuring stability by preventing excessive attenuation. Equation (1) defines a servo encoding method that maps the base wrench components f_x , f_y , f_z , and τ_z to each servomotor's rotation. As shown in Fig. 2, each component contributes to the servo angle through a specific combination of rotation direction, with their effects summed in proportion to produce the final angle increment. The dynamic coefficient in (2) further modulates the contributions of f_y and f_x based on the torque-to-force ratios τ_x/f_y and τ_y/f_x , respectively, as defined in the second equation. This adjustment primarily aims to simulate sensory differences caused by contact points farther from or closer to the wrist through asymmetric rotations. Additionally, the algorithm employs dynamic time-window filtering to enhance stability, ensuring sensitive responses to sudden changes and stable responses under sustained forces.

C. Haptic Glove

The IMU-based haptic glove is designed to deliver a cost-effective and user-friendly solution for dexterous teleoperation, enabling immersive experiences through low-latency finger motion capture, precise fingertip calibration, and intuitive haptic feedback. We elaborate on this system across three key aspects: hardware design, retargeting and calibration algorithms, and haptic feedback mechanisms.

1) *Hardware Design*: The Glovity haptic glove employs a compact, modular assembly akin to LEGO blocks, eliminating the need for screws or adhesives during assembly. Drawing on ergonomic principles, it integrates PLA-printed rigid structural components with flexible TPU straps, offering comfort comparable to commercial-grade gloves. The core hardware comprises six WitMotion JY61P IMU modules (five positioned at the fingertips and one on the back of the hand), a 49E Hall sensor (mounted on the thumb pad, paired with a magnet on the index finger), five 3V ERMs (8×2 mm) with L9110s drivers, an XIAO ESP32-S3 microcontroller, and a custom circuit board. All components are integrated within the glove, resulting in a sleek and cohesive design.

2) *Retargeting and Calibration Algorithms*: The retargeting process utilizes the differences in angular velocities and accelerations between the fingertip IMUs and the IMU on

the back of the hand to compute finger bend angles. For the index to little fingers, a hybrid acceleration-angular velocity fusion algorithm ensures robustness across diverse postures. The fused angle k is computed using a complementary filter that blends gyroscope integration with accelerometer-derived angles:

$$\theta_k = \alpha \cdot (\theta_k + \Delta\theta_{\text{gyro}}) + (1 - \alpha) \cdot \theta_{\text{acc}} \quad (3)$$

where: θ_k is the previous fused angle, $\Delta\theta_{\text{gyro}}$ is the gyroscope increment (with corrections for cross-axis influences and damping), θ_{acc} is the angle derived from accelerometer projections, α is an adaptive blending factor, to prioritize gyro when acceleration is not reliable.

This fusion ensures robustness against gyroscope drift and accelerometer noise, with additional smoothing via a moving average window of size 6. For the thumb, relative angular velocity is employed, supplemented by Hall sensor-assisted calibration. This leverages sensor readings to quantify the proximity distance between thumb and index fingertips: the sensor is mounted on the thumb pad and paired with a neodymium magnet on the index finger, causing the output value (h) to increase as magnetic field strength grows with approaching fingertips. Based on sensor sensitivity and practical needs, two key thresholds are predefined: a lower threshold (e.g., h_{low}) to detect proximity and trigger calibration when $h \geq h_{\text{low}}$, and a higher threshold (e.g., h_{high}) to indicate fingertip contact when $h \geq h_{\text{high}}$, setting specific angles per the robotic hand's structure for pinching the fingers together. Within the interval ($h_{\text{low}} \leq h < h_{\text{high}}$), transitional angles are computed via linear interpolation and smoothed with low-pass filtering, ensuring the dexterous hand's fingertip pose matches the operator's gesture. This threshold-based method uses empirical sensor calibration to adapt to varying user hand shapes and environmental noise, enabling efficient grasping of thin objects.

3) *Haptic Feedback Mechanisms*: The haptic feedback is derived from force sensors on the dexterous hand, with sensor readings linearly mapped to ERM motor speeds via PWM control, which similar to previous works like [17]. This feedback enhances the teleoperation's immersion and precision, facilitating intuitive control in complex tasks.

D. Hand pose tracking

Glovity is designed to be compatible with the Vive Tracker for precise hand gesture tracking, with dedicated mounting holes integrated into the haptic glove for seamless attachment. By leveraging the tracker's high-precision positional data, we employ inverse kinematics (IK) to compute joint angles, coupled with servo control, to achieve low-latency retargeting of hand poses to the robotic dexterous hand.

IV. EXPERIMENTS

By conducting experiments using Glovity, we hope to find answers to the following questions:

- Does the wrench feedback system provide accurate and intuitive feedback while improving the efficiency of teleoperation for contact-rich tasks?

- How does the haptic glove perform in object grasping tasks, especially thin objects?
- What is the potential of Glovity for imitation learning, especially for learning with force and torque?

To answer the above questions, several experiments were designed and all experiments were conducted using a UR5e robotic arm equipped with an FT300s force/torque sensor and an Inspire Hand RH56. For the user studies, we recruited 10 participants (5 males, 5 females) from the university, none of whom had prior professional teleoperation training.



Fig. 3: The force application regions for the perception experiment. The blue regions correspond to Level 2, comprising eight distinct areas, while the orange regions represent Level 3, consisting of two areas.

A. User Study on Wrench Feedback

To address the questions outlined above, we designed three sub-experiments: a perception to evaluate the accuracy and intuitiveness of the feedback, a case study to explore the potential of glovity in highly dynamic tasks, and a teleoperation experiment to assess efficiency improvements in real-world operations.

1) *Perception Experiment*: In the perception experiment setup, participants wore the wrench feedback system and completed a 3-minute familiarization period to become accustomed to the device. Subsequently, they performed three level perception tasks, ordered from least to most challenging. First, forces were applied along the x, y, or z axis of the FT sensor, and participants identified the force direction. Second, forces were applied to one of 8 distinct areas on the middle section of the Inspire Hand, as shown in Fig 3, and participants identified the specific area. Finally, forces were applied to regions near the fingertips or wrist on the palm, and participants distinguished between these two locations. Each task was repeated 10 times without visual input, with **success rate** as the evaluation metric, recognition errors or inability to identify as failure.

Results: As shown in Fig 4, participants in the first level of the perception experiment demonstrated high accuracy in distinguishing forces applied along the x, y, and z axes. In the second level, most participants successfully identified

force application areas among the eight designated regions on the middle section of the Inspire Hand, though some showed reduced sensitivity to specific regions, likely due to a mismatch between the device’s size and participants’ hand dimensions, causing slight looseness. In the third level, only a few participants accurately differentiated between force applications near the fingertips and wrist on the palm. This lower performance is attributed to the wrench feedback system’s fixed mechanism at the palm’s middle section, which, despite significant servo rotation differences, provided insufficiently intuitive feedback, hindering participants’ ability to precisely identify the force application region.

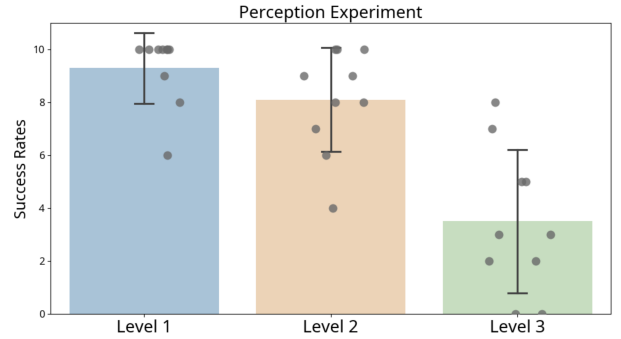


Fig. 4: Success rates across three experiment levels, where the average success rates of level 1, level 2, and level 3 is 93%, 81%, and 35%

2) *Teleoperation Experiment*: In this experiment, participants performed a page-flipping task five times using the Glovity system, under two conditions: with and without wrench feedback. Flipping thin deformable objects poses challenges in maintaining stable contact and adapting to surface deformation without precise control [26]. The task’s difficulty arose from the need to apply consistent downward force with the index finger, leveraging friction to lift and flip the page successfully. To minimize order effects, five participants initiated the task with the wrench feedback condition, while the remaining began without it. Haptic feedback remained active throughout the experiment; however, its utility was limited due to the near-vertical orientation of the index finger. Performance was evaluated based on two metrics: **task completion time and success rate**. A trial was deemed successful if the page was flipped in a single attempt without slipping, while failures were recorded if the page slipped mid-process or if excessive force caused the robotic system to stop.

Results: As shown in Table I, the experiment results indicate that wrench feedback significantly enhanced the success rate of the book-flipping task while also reducing completion time. The absence of feedback led to a more pronounced increase in instances of excessive force, a trend supported by user reports highlighting the difficulty in accurately assessing contact force magnitude through visual cues alone. Without feedback, operators struggled to perceive whether their fingertip was adequately pressing on the page, often resulting in overcompensation. Wrench feedback effectively addressed

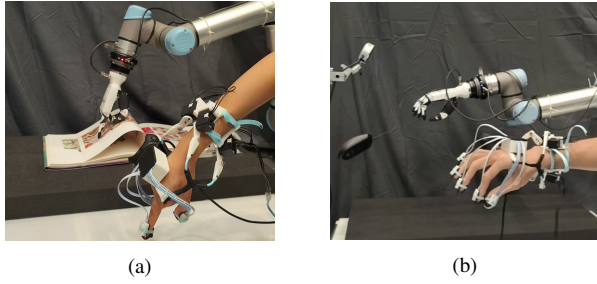


Fig. 5: (a) Teleoperation Experiment on Page Flipping. (b) Case Study on Dynamic Mouse Swinging

this by providing clear sensory cues about fingertip-book contact and pressure levels, enabling users to adjust their force application more precisely and maintain better control throughout the task.

Wrench Feedback	✓	✗
Success Rate	39/50	24/50
Mean Task Time (s)	11.8	17.7
Failures (Slipping)	7	15
Failures (Excessive Force)	4	11

TABLE I: Teleoperation Experiment Results

3) *Case Study of Dynamic Mouse Swinging*: To investigate the capability of the wrench feedback system in facilitating dynamic tasks, we conducted a case study involving a swinging motion task. A broken computer mouse was selected for its suitable weight and asymmetric mass distribution (heavier at the end). The mouse’s cable was securely attached to the index finger of the Inspire Hand. As shown in Fig 1b and 5b, Operators performed the task via teleoperation, swinging the mouse in a circular motion **without visual feedback**, relying solely on wrench feedback.

Results: Operators could successfully swing the mouse and modulate both the radius and speed of its motion. This highlights the wrench feedback system’s potential to support high-dynamic tasks, enabling effective control in scenarios where visual input is absent.

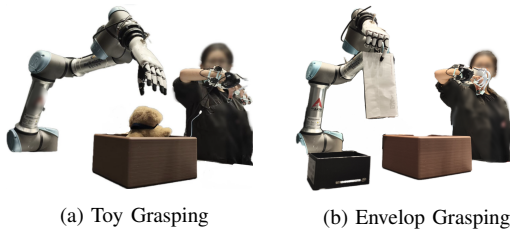


Fig. 6: User study on haptic glove

B. User Study on Haptic Glove

To evaluate the grasping capabilities of the haptic glove, participants performed two tasks: picking up a toy dog and an envelope, and placing each into a box, with each task repeated five times. The experiment used AnyTeleop [10]

and Vrtrix glove pro 11 [27] as benchmarks, with the robotic arm’s end-effector uniformly controlled via a Vive Tracker. The performance metrics includes the success rate and the average time to complete the grasp.

Results: The experimental results demonstrate that, in the toy dog grasping task, the three methods—Glovity, AnyTeleop, and Vrtrix Glove—exhibited comparable performance, characterized by high success rates and short completion times. However, significant differences emerged in the envelope grasping task. AnyTeleop achieved the highest success rate, yet its average completion time was approximately 4 seconds longer than our method. This delay arose from AnyTeleop’s latency and inconsistent detection of thumb-index finger pinching, necessitating users to adjust their hand gestures and retry the pinching action, exacerbated by the absence of haptic feedback. In contrast, the Vrtrix Glove showed a marked decline in success rate and increased completion time, primarily due to its inability to accurately locate fingertip pinching, requiring users to repeatedly adjust thumb and index finger positions to ensure that the Inspire Hand successfully grasped the envelope.

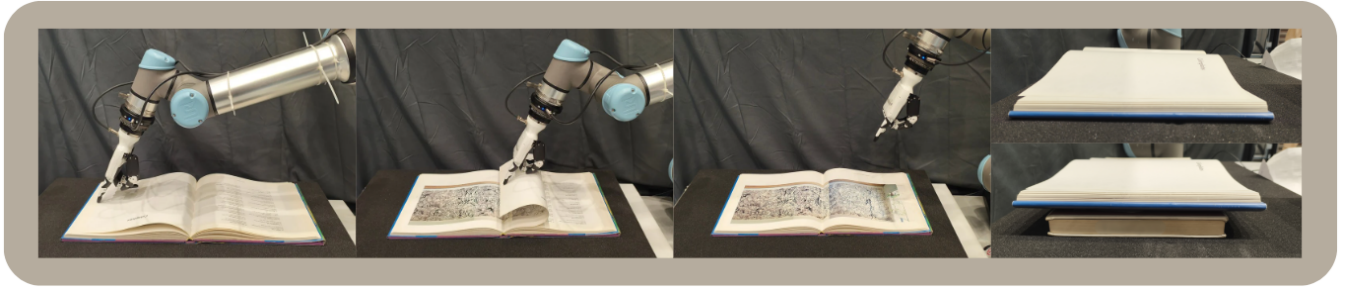
Method	Mean Completion Time (s)		Success Rate	
	Toy Dog	Envelope	Toy Dog	Envelope
Ours	5.2	5.9	46/50	41/50
AnyTeleop	5.8	9.3	48/50	44/50
Vrtrix Glove	5.3	18.4	45/50	17/50

TABLE II: Grasping Experiment Results

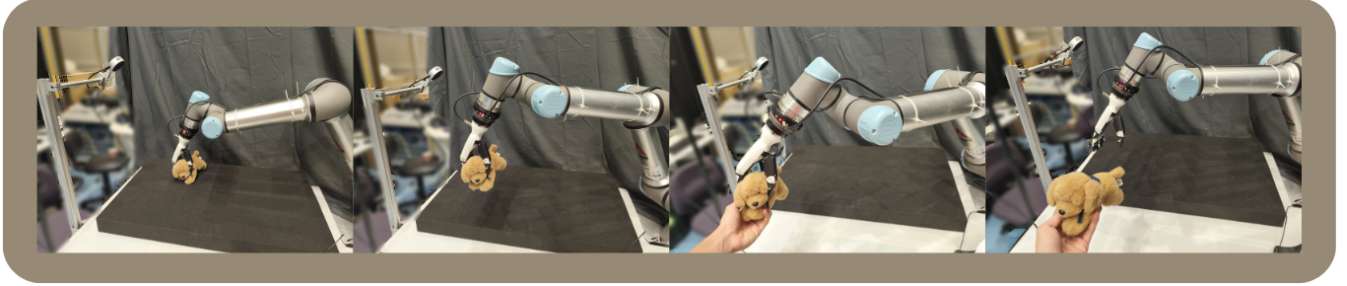
C. Imitation Learning

We conducted two experiments to evaluate the Glovity system for imitation learning in contact-rich and human-robot interaction tasks, using the IDP3 framework [28]. Fig 8 shows the data flow in this experiments. An Intel RealSense L515 camera provided RGB input at 640×480, resized to 224×224, the policy integrated RGB images processed via a pretrained R3M encoder and wrench data (6D force-torque vector) as low-dimensional observations, alongside 12D agent pose (6D for UR5e and 6D for Inspire Hand).

1) *Page-Flipping Task*: The first experiment assessed the Glovity system’s ability to capture high-fidelity wrench data for training policies in contact-rich tasks requiring both force regulation and force-based decision-making. In this setup, a picture album (300×300 mm) was placed on a flat table, with a 6 cm thick hard sponge pad placed between the table and the album to enhance the smoothness of force variations. The first side of the album was randomly elevated to a height between 0 and 5 cm using supports. The paper material and page length required the policy to press with sufficient force and slide backward 6 - 8 cm to initiate page lift, followed by leveraging the page’s elasticity to generate friction between the fingertip and page for completing the flipping action. We collected 30 demonstrations, each lasting 150 timesteps (approximately 13s), and trained two policies for 300 epochs: one with wrench data (6D force-torque) as



(a) Page-Flipping: The last figure shows two different height settings. The robot needs to learn to judge the timing of page flipping based on contact force.



(b) Hand-Over: The robot grasp a toy dog within a 20×20 cm workspace, moved it to a designated position outside the camera's field of view, and released it upon perceiving a force from receiver.

Fig. 7: Imitation Learning Experiment

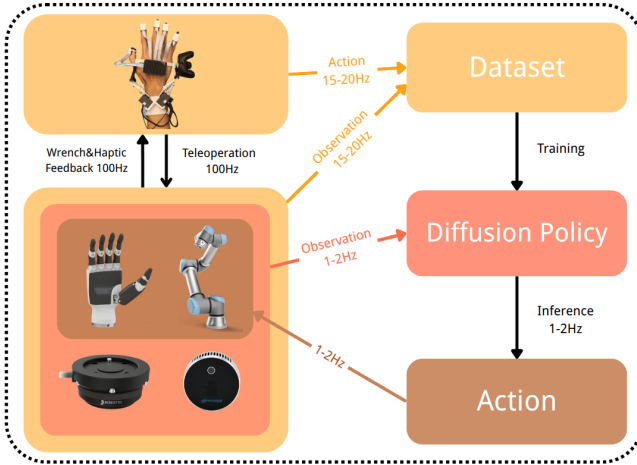


Fig. 8: Data Flow in Imitation Learning Experiments

part of the observation space and one without wrench data. Success rate, defined as a complete page flipping without slippage or excessive force causing the robotic arm to stop protectively, was the primary metric.

Results: For book elevations ranging from 0 to 5 cm, the DP-R3M policy augmented with wrench data achieved a success rate of 80% (16/20 trials), demonstrating robust adaptation to varying book elevations by using force feedback to gauge contact stability before initiating the flip. Ablation studies revealed that removing wrench data from the observations degraded performance to 45% (9/20 trials), as the policy failed to dynamically adjust the rubbing force and timing, often resulting in incomplete flips or excessive

force. This underscores the effectiveness of wrench data in policy learning in enabling adaptive, force-aware behaviors in unstructured environments and the successful utilization of wrench data collected via the Glovity system for effective policy training.

2) *Hand-Over:* The second experiment explored force-based imitation learning in human-robot interaction. We collected 50 demonstrations, each spanning 200 timesteps (approximately 17s), where the operator grasped a toy dog within a 20×20 cm workspace, moved it to a designated position outside the camera's field of view (50 cm away), and released it upon perceiving a distinct force from another human via wrench and haptic feedback. Two policies were trained for 300 epochs: one with wrench data and one without. Success was defined as the robot releasing the toy only after a pulling force was applied to the toy by a human.

Results: The wrench-augmented policy achieved a 75% success rate (15/20 trials), reliably holding the toy at the designated position until a human applied sufficient force, then releasing it and returning to the initial pose. Of the five failures, three were due to the robot failing to grasp the toy dog, and two were due to insufficient grip strength, causing the robot to not return to the initial pose after the human removed the toy. Without wrench data, the policy prematurely released the toy upon reaching the designated position, failing to wait for human intervention. However, qualitative analysis revealed that the wrench-augmented policy exhibited slightly increased jitter in the robot's motion, likely due to the diffusion model's iterative inference process exacerbating discontinuities in actions, which amplified noise in the force-torque sensor readings. This issue could

potentially be mitigated in future work through optimized data preprocessing or refined model architectures to enhance action continuity.

V. DISCUSSION&CONCLUSION

Glovisity exhibits several limitations that warrant further exploration. First, the precision of the wrench feedback system in force rendering remains sub-optimal while the effective feedback area is limited, potentially constraining its utility in tasks demanding precise force interactions, such as microsurgery or fragile object handling. Second, the haptic glove is currently restricted to six degrees-of-freedom tracking, making it challenging to perform complex hand manipulation tasks. Third, our integration of wrench signals into imitation learning frameworks, while effective for basic contact-rich tasks, employs a straightforward observational approach that may not generalize to highly complex, long-horizon manipulations involving variable dynamics.

In conclusion, Glovisity represents a significant advancement in cost-effective, open-source teleoperation for contact-rich dexterous manipulation. By integrating spatial wrench feedback with a haptic glove, it facilitates intuitive force-torque and tactile interactions, enhancing data collection and imitation learning. Looking forward, Glovisity's intuitive feedback can be leveraged to advance imitation learning algorithms, allowing models to learn from operators' force-feedback-guided action adjustments, thereby enabling robots to dynamically adapt their movements based on force cues and achieve more complex, long-horizon contact-rich tasks. Furthermore, this force and haptic feedback can facilitate human-in-the-loop learning, empowering humans to correct robotic actions in real-time based on intuitive sensory cues. Additionally, Glovisity's ultra-low latency positions it as a promising tool for high-dynamic scenarios, such as locomotion and rapid manipulation. By open-sourcing all hardware and software, we invite the research community to build upon this foundation, addressing limitations like force rendering intuitiveness and expanding degrees-of-freedom, to unlock new frontiers in adaptive, collaborative, and autonomous robotic systems.

REFERENCES

- [1] K. K. Babarahmati, C. Tiseo, Q. Rouxel, Z. Li, and M. Mistry, "Robust high-transparency haptic exploration for dexterous telemanipulation," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 10 146–10 152.
- [2] C. Pacchierotti and D. Prattichizzo, "Cutaneous/tactile haptic feedback in robotic teleoperation: Motivation, survey, and perspectives," *IEEE Transactions on Robotics*, vol. 40, pp. 978–998, 2023.
- [3] A. O'Neill, A. Rehman, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta, A. Mandlekar, A. Jain *et al.*, "Open x-embodiment: Robotic learning datasets and rt-x models: Open x-embodiment collaboration 0," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 6892–6903.
- [4] T. Buamane, M. Kobayashi, Y. Uranishi, and H. Takemura, "Bi-act: Bilateral control-based imitation learning via action chunking with transformer," in *2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2024, pp. 410–415.
- [5] P. Peuchpen, H. Liu, and J. Ma, "Real-time teleoperation in motion mapping for medical robot based on robot manipulator and haptic device," in *2024 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2024, pp. 2221–2226.
- [6] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel, "Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators," 2023.
- [7] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," 2023. [Online]. Available: <https://arxiv.org/abs/2304.13705>
- [8] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang, "Open-television: Teleoperation with immersive active visual feedback," 2024. [Online]. Available: <https://arxiv.org/abs/2407.01512>
- [9] S. P. Arunachalam, I. Güzey, S. Chintala, and L. Pinto, "Holo-dex: Teaching dexterity with immersive mixed reality," 2022. [Online]. Available: <https://arxiv.org/abs/2210.06463>
- [10] Y. Qin, W. Yang, B. Huang, K. V. Wyk, H. Su, X. Wang, Y.-W. Chao, and D. Fox, "Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system," 2024. [Online]. Available: <https://arxiv.org/abs/2307.04577>
- [11] S. Song, A. Zeng, J. Lee, and T. Funkhouser, "Grasping in the wild: learning 6dof closed-loop grasping from low-cost demonstrations," 2020. [Online]. Available: <https://arxiv.org/abs/1912.04344>
- [12] P. Wu, Y. Shentu, Q. Liao, D. Jin, M. Guo, K. Sreenath, X. Lin, and P. Abbeel, "Robocopilot: Human-in-the-loop interactive imitation learning for robot manipulation," 2025. [Online]. Available: <https://arxiv.org/abs/2503.07771>
- [13] A. Bazhenov, S. Satsevich, S. Egorov, F. Khabibullin, and D. Tsetserukou, "Echo: An open-source, low-cost teleoperation system with force feedback for dataset collection in robot learning," 2025. [Online]. Available: <https://arxiv.org/abs/2504.07939>
- [14] J. J. Liu, Y. Li, K. Shaw, T. Tao, R. Salakhutdinov, and D. Pathak, "Factr: Force-attending curriculum training for contact-rich policy learning," *arXiv preprint arXiv:2502.17432*, 2025.
- [15] Z. He, H. Fang, J. Chen, H.-S. Fang, and C. Lu, "Foar: Force-aware reactive policy for contact-rich robotic manipulation," *IEEE Robotics and Automation Letters*, 2025.
- [16] L. Meli, C. Pacchierotti, and D. Prattichizzo, "Experimental evaluation of magnified haptic feedback for robot-assisted needle insertion and palpation," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, no. 4, p. e1809, 2017.
- [17] 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," 2024. [Online]. Available: <https://arxiv.org/abs/2407.03162>
- [18] H. Xu, M. Chen, G. Li, L. Wei, S. Peng, H. Xu, and Q. Li, "An immersive virtual reality bimanual telerobotic system with haptic feedback," 2025. [Online]. Available: <https://arxiv.org/abs/2501.00822>
- [19] K. Shaw, Y. Li, J. Yang, M. K. Srirama, R. Liu, H. Xiong, R. Mendonca, and D. Pathak, "Bimanual dexterity for complex tasks," 2024. [Online]. Available: <https://arxiv.org/abs/2411.13677>
- [20] C. Lu, L. Jin, Y. Liu, J. Wang, and W. Li, "Teleoperated grasping using data gloves based on fuzzy logic controller," *Biomimetics*, vol. 9, no. 2, p. 116, 2024.
- [21] H. Zhang, S. Hu, Z. Yuan, and H. Xu, "Doglove: Dexterous manipulation with a low-cost open-source haptic force feedback glove," 2025. [Online]. Available: <https://arxiv.org/abs/2502.07730>
- [22] X. Zhu, B. Huang, and Y. Li, "Touch in the wild: Learning fine-grained manipulation with a portable visuo-tactile gripper," *arXiv preprint arXiv:2507.15062*, 2025.
- [23] W. Liu, J. Wang, Y. Wang, W. Wang, and C. Lu, "Forcemimic: Force-centric imitation learning with force-motion capture system for contact-rich manipulation," *arXiv preprint arXiv:2410.07554*, 2024.
- [24] 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 *Proceedings of Robotics: Science and Systems (RSS)*, 2025.
- [25] H. Kim, Y. Ohmura, and Y. Kuniyoshi, "Goal-conditioned dual-action imitation learning for dexterous dual-arm robot manipulation," *IEEE Transactions on Robotics*, vol. 40, pp. 2287–2305, 2024.
- [26] C. Zhao, C. Jiang, L. Luo, S. Yuan, Q. Chen, and H. Yu, "Learning thin deformable object manipulation with a multi-sensory integrated soft hand," *IEEE Transactions on Robotics*, 2025.
- [27] VRTRIX Inc., "VRTRIX® Data Glove," <http://www.vrtrix.com/>, 2025.
- [28] Y. Ze, Z. Chen, W. Wang, T. Chen, X. He, Y. Yuan, X. B. Peng, and J. Wu, "Generalizable humanoid manipulation with 3d diffusion policies," 2025. [Online]. Available: <https://arxiv.org/abs/2410.10803>