ACE-F: A Cross Embodiment Foldable System with Force Feedback for Dexterous Teleoperation

Rui Yan*   Jiajian Fu*   Shiqi Yang*   Lars Paulsen   Xuxin Cheng   Xiaolong Wang
UC San Diego
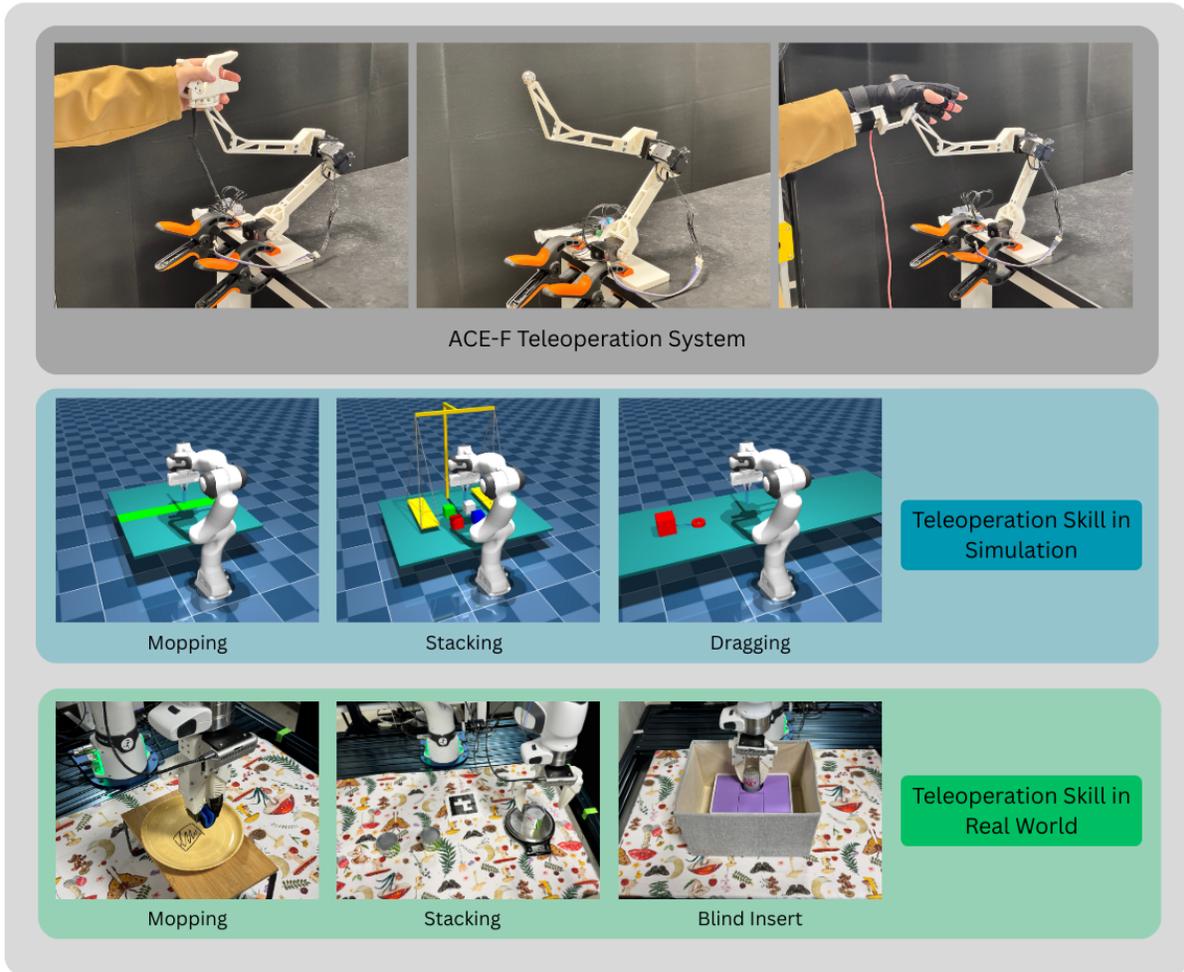
Fig. 1: ACE-F enables users to complete contact-rich tasks with force feedback in both simulator and real-world settings. It excels at mopping, stacking, and dragging in simulation, and at mopping, stacking, and blind insertion tasks in real-world scenarios.

Abstract—Teleoperation systems are essential for efficiently collecting diverse and high-quality robot demonstration data, especially for complex, contact-rich tasks. However, current teleoperation platforms typically lack integrated force feedback, cross-embodiment generalization, and portable, user-friendly designs, limiting their practical deployment. To address these limitations, we introduce ACE-F, a cross embodiment foldable teleoperation system with integrated force feedback. Our approach leverages inverse kinematics combined with a carefully designed human-robot interface (HRI), enabling users to capture precise and high-quality demonstrations effortlessly. We further propose a generalized soft-controller pipeline integrating PD control and inverse dynamics to ensure robot safety and precise motion control across diverse robotic embodiments. Critically, to achieve cross-embodiment generalization of force feedback without additional sensors, we innovatively interpret end-effector positional deviations as virtual force signals, which enhance data collection and enable applications in imitation learning. Extensive teleoperation experiments confirm that ACE-F significantly simplifies the control of various robot embodiments, making dexterous manipulation tasks as intuitive as operating a computer mouse.

I. Introduction

Teleoperation systems have shown great potential for collecting high-quality, diverse demonstration data for complex,

contact-rich robotic tasks. However, existing platforms suffer from three main limitations: (1) lack of integrated force feedback—either providing no haptic cues or relying on expensive, hard-to-integrate force/torque (FT) sensors [20, 34]; (2) poor cross-embodiment generalization—joint-copying schemes must be redesigned for each new robot morphology [33, 19]; and (3) bulky, non-portable hardware that hinders rapid deployment in real-world scenarios [50, 19].

To address these challenges, we propose **ACE-F**, a cross-embodiment foldable teleoperation system with integrated force feedback. First, ACE-F infers real-time 3-DoF external forces by monitoring end-effector (EE) trajectory deviations, no additional sensors required, and applies active gravity and friction compensation on the leader and follower arms to deliver smooth, intuitive haptic cues [22, 34]. Second, we combine inverse kinematics (IK)–based leader-arm control with glove-based hand tracking to build a universal retargeting algorithm that adapts to diverse robot platforms; a magnetic quick-swap interface further enables future integration of tactile gloves [46, 26, 49]. Finally, our soft-controller pipeline fuses proportional-derivative (PD) control with custom inverse dynamics (ID), ensuring stability, responsiveness, and safety across embodiments, and allowing rapid deployment via minor tuning of URDF parameters [46].

These innovations raise a critical question: *How can we obtain accurate hand poses and end-effector positions at low cost to enable dexterous manipulation across a wide range of robot platforms?*

The answer lies in combining cross-platform primary-arm control based on IK and ID with a lightweight, foldable human–robot interface. Our ACE-F system not only achieves high-precision demonstrations but also maintains affordability and portability.

In this paper, we describe the hardware design and software architecture of ACE-F, integrating real-time three-degree-of-freedom (3-DoF) force estimation via EE trajectory deviation, active gravity and friction compensation on the leader arm, IK-driven Cartesian control with retargeting algorithms, and a PD+ID soft-controller pipeline. Extensive experiments demonstrate two key advantages of ACE-F: (1) users can rapidly adapt and efficiently, accurately perform cross-platform teleoperation tasks under varying precision and workspace requirements; and (2) at a relatively low cost, ACE-F significantly outperforms systems without force feedback in complex contact-rich tasks.

II. RELATED WORK

Force Feedback Teleoperation. Force feedback has become a widely recognized enabler for contact-rich teleoperation, allowing operators to perceive interaction forces and improve manipulation performance [35, 13, 10]. Although many commercial robot arms incorporate built-in 6-DoF FT sensors, their high cost and integration complexity make them impractical for general deployment [37]. Furthermore, the majority of low-cost teleoperation systems forgo force

feedback entirely, relying solely on position or joint commands with no haptic cues [51, 45, 16, 46]. Traditional force-feedback teleoperation implementations therefore depend on external FT sensors mounted on the secondary device, imposing hardware and calibration burdens [12, 1, 5, 31]. Virtual force feedback schemes approximate contact forces via kinematic or impedance models, but cannot fully capture true interaction dynamics [2]. Some teleoperation systems have also attempted to convey force information through non-haptic channels—e.g., visual overlays on the video stream, audio alerts, or controller vibration cues—but these indirect modalities often lack intuitiveness and can increase operator cognitive load [13]. To overcome these limitations, ACE-F infers real-time 3-DoF end-effector forces from trajectory deviations and applies gravity and friction compensation on the primary arm—providing sufficiently accurate force cues for daily-life teleoperation tasks without any additional sensors.

Cross-Embodiment Teleoperation. Mapping human motions to robots with differing kinematic structures is necessary for proper teleoperation [33]. Direct joint-copying approaches build a small primary arm or mobile controller that mirrors the target robot's kinematics, providing intuitive, low-latency mapping, but require rebuilding the hardware for each new robot [51, 45, 16, 5, 31, 47]. In contrast, IK-based Cartesian control naturally generalizes across embodiments, allowing a single primary interface to drive robots of varied morphologies without any hardware changes [39, 50, 46, 24, 38, 18, 3, 23]. IK-driven teleoperation systems typically use four main interfaces to obtain wrist and hand pose: motion-capture devices [44, 14, 41, 30, 8], cameras [39], VR equipment[17, 4, 24, 11, 9, 27, 32], or exoskeleton hardware [50, 46, 6, 15, 42, 21, 25, 7]. The first three approaches can capture complete wrist and hand information to enable dexterous end-effector control [40, 48, 36, 18, 39, 29, 28], but their interfaces make integrating force feedback difficult. Exoskeleton-based systems offer a direct way to add force feedback, yet their bulky size and mechanical complexity significantly increase torque requirements—raising motor costs and reducing wearability. By contrast, ACE-F combines a compact, foldable 3-DoF primary arm with glove-based hand tracking to achieve precise, occlusion-free full-hand pose capture while minimizing device volume and torque demands—thereby lowering motor performance requirements and overall system cost, and enabling seamless integration of tactile gloves in the future.

III. SYSTEM DESIGN

Hardware Design. The ACE-F system, illustrated in Fig. 6, is a robotic manipulator designed for precise 3-DoF force feedback. It features three independent joints: a base joint for foundational rotation and two perpendicular elbow joints for compact and robust force rendering.

The manipulator employs DYNAMIXEL XM430-W350-T motors with U2D2 controllers, ensuring precise and responsive joint control. At its endpoint, a passive quick-release spherical joint supports interchangeable end-effectors via a 3D-printed ball-and-socket design embedded with neodymium magnets.
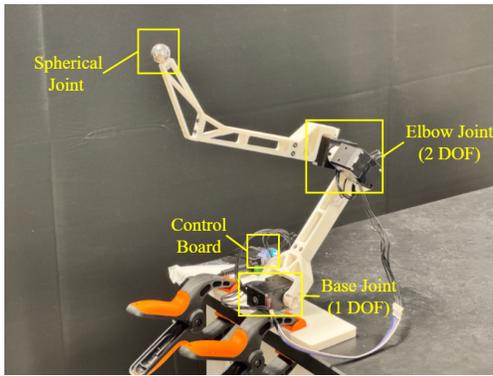
Fig. 2: Overview of the ACE-F system. **Left**: Annotated view of the ACE-F arm showing the base joint (1 DoF), perpendicular elbow joints (2 DoF), and the magnetic spherical joint for interchangeable end-effectors. **Right**: Three representative end-effector configurations are enabled by the spherical joint, using a bare setup, gripper attachment, or wearable glove.

An integrated elastic safety lock mechanism prevents unintended detachment during operation.

We validated three end-effector configurations:

- **Bare Configuration**: General-purpose setup for force feedback.
- **Gripper Configuration**: Designed for controlling single-arm robots with grippers.
- **Glove Configuration**: Tailored for humanoid platforms equipped with dexterous hands.

This modular design significantly enhances the system's flexibility and adaptability for diverse real-world applications.

End-Effector Control and Feedback. One of the core challenges in teleoperation is achieving full 6-DoF control (position and orientation) of the robot end-effector using compact and low-DoF input devices. A 3-DoF arm alone cannot simultaneously define both the position and orientation of the end-effector in space. To address this limitation, ACE-F decouples position and orientation control: the foldable 3-DoF primary arm is used to determine the end-effector's Cartesian position, while a glove-based tracking module captures the wrist and finger orientation in real time. By combining these two streams, ACE-F reconstructs a complete 6-DoF in-hand pose of the operator, which can then be retargeted to the robot.

This hybrid control design not only preserves portability and affordability but also enables platform-agnostic retargeting. The system uses inverse kinematics (IK) to convert the desired position from the primary arm and the orientation from the glove into robot-specific end-effector commands. As a result, users can define both position and rotation of the end-effector naturally within the workspace, enabling seamless teleoperation across robots with different kinematic structures.

Additionally, ACE-F incorporates virtual force feedback by interpreting trajectory deviations between the commanded and actual end-effector positions as 3-DoF force signals. These inferred forces are rendered on the primary arm via active torque control with gravity and friction compensation, providing intuitive haptic cues without requiring external sensors. This sensorless feedback loop improves the precision of manipulation and makes contact events such as collisions, object slippage, or resistance perceptible to the operator.

This design not only preserves the flexibility of the original ACE system, such as the ability to adapt to different workspace scales through simple geometric transformations—but also offers key advantages in force feedback implementation. By reducing the degrees of freedom on the primary arm to 3-DoF, ACE-F simplifies both mechanical design and real-time torque control. This reduction makes it significantly easier to implement reliable and low-latency force feedback, as the system only needs to estimate and render 3D translational forces rather than full 6D wrenches. At the same time, the combination of glove-based orientation tracking and IK-based position mapping ensures that the user can still define arbitrary 6-DoF in-hand poses within the workspace. Together, these properties enable intuitive and high-fidelity teleoperation across diverse robotic platforms while maintaining portability, low cost, and modular expandability.

Augmented Inverse Kinematics Solver. Unlike conventional IK solvers that purely minimize end-effector position and orientation errors, we propose an augmented IK approach tailored to the unique teleoperation challenges of the Franka arm. Our solver introduces two additional "tasks" to improve robustness and avoid singularities:

First, we compute the projection angle of the end-effector onto the base plane and match it to the first joint angle of the robot. This ensures a natural alignment between the operator's intended direction and the robot's base rotation.

Second, to prevent the Franka elbow from bending outward when the end effector approaches the base, an action that can lead to kinematic singularities, we introduce a soft constraint on the fourth joint. Specifically, we assign it a higher target value in the vertical (z) axis, encouraging a posture that avoids such configurations.

These task-level modifications enhance the solver's reliability and stability, allowing intuitive and continuous teleoperation even near the robot's kinematic limits.

Force Feedback Calculation. Instead of relying on the conventional Jacobian-based wrench-to-torque mapping for force feedback calculation, we adopt a simpler, yet more
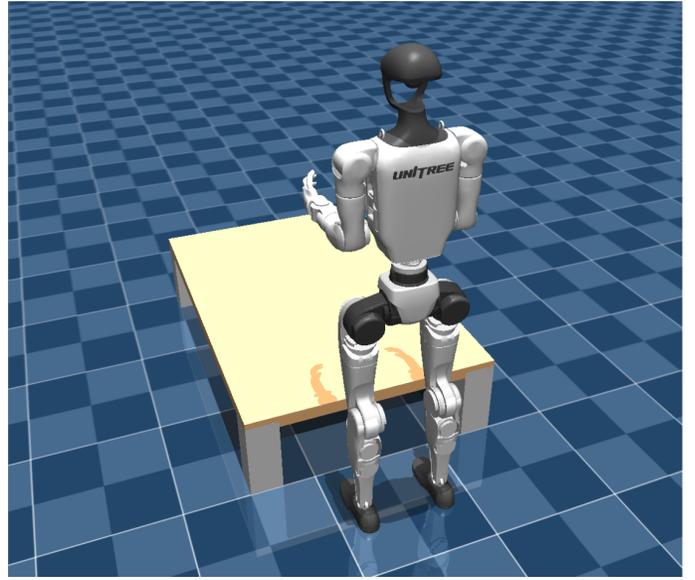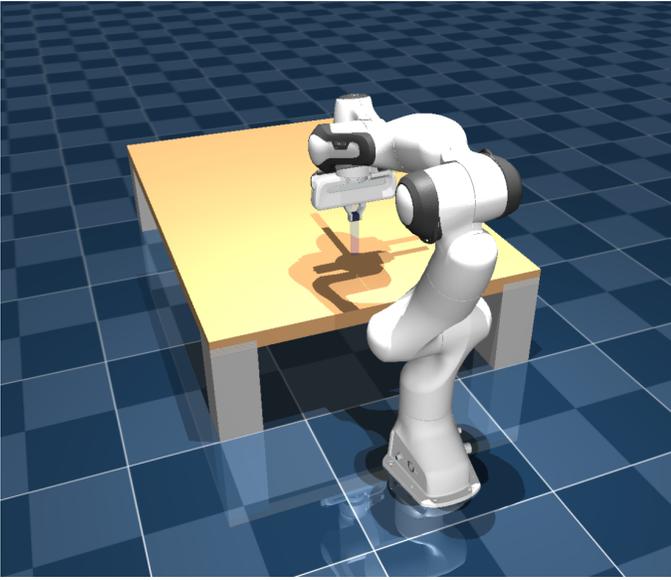
Fig. 3: Using a single ACE-F arm, we can control the Franka robot in simulated and real-world environments. When using two ACE-F arms, we can control a bimanual robot, such as the Unitree G1 in simulation.

robust approach to maintain system stability. Specifically, we compute the deviation between the secondary arm end-effector's target and actual positions:

$$\Delta \mathbf{ee} = \mathbf{target} - \mathbf{current}$$

This deviation, $\Delta \mathbf{ee}$, serves as the core indicator of the feedback force magnitude. Traditional wrench-to-torque mappings can be highly sensitive to transient forces, where even short-duration impulses result in substantial torque spikes that destabilize the primary arm. To circumvent this issue, we introduce a virtual target pose for the primary arm. By scaling $\Delta \mathbf{ee}$ and applying it to the primary arm's current end-effector pose, the primary arm naturally "tries" to align with the secondary's pose via force feedback. Since the operator's hand firmly grips the end-effector, this alignment manifests as tangible forces rather than significant positional displacements, thereby avoiding large oscillations.

To further prevent destabilizing effects at high speeds—where $\Delta \mathbf{ee}$ may inflate due to dynamic motion rather than contact—we modulate the feedback magnitude by the secondary's Cartesian velocity:

$$\text{Force Feedback Factor} = \sqrt{\frac{\alpha \cdot \|\Delta \mathbf{ee}\|^2}{1 + \|\mathbf{v}_{\text{cartesian}}\|^2}}$$

This ensures that the feedback force remains negligible during smooth or low-speed motions, only becoming significant during actual contact interactions. Additionally, we apply this feedback factor not only to generate the virtual target pose but also to adaptively modulate the primary arm's impedance gains (K_p and K_d), implementing stable and intuitive haptic feedback across tasks.

Algorithm 1 Force Feedback–Enhanced Teleoperation Loop

1: **repeat**
2:    **Primary Arm:** Compute target $\mathbf{ee}_{\text{target}}$ based on operator's 3-DoF arm pose and glove orientation
3:    **Primary Arm:** Solve inverse kinematics (IK) for $\mathbf{q}_{\text{target}}$ of the secondary arm
4:    **Primary Arm:** Send $\mathbf{q}_{\text{target}}$ to secondary arm
5:    **Secondary Arm:** Send current joint positions $\mathbf{q}_{\text{current}}$ to primary
6:    **Secondary Arm:** Receive $\mathbf{q}_{\text{target}}$ and start moving toward it
7:    **Primary Arm:** Compute current secondary $\mathbf{ee}_{\text{current}}$ using forward kinematics (FK) from $\mathbf{q}_{\text{current}}$
8:    **Primary Arm:** Compute deviation $\Delta \mathbf{ee} = \mathbf{ee}_{\text{target}} - \mathbf{ee}_{\text{current}}$
9:    **Primary Arm:** Compute force feedback factor:

$$\text{Factor} = \sqrt{\frac{\alpha \cdot \|\Delta \mathbf{ee}\|^2}{1 + \|\mathbf{v}_{\text{cartesian}}\|^2}}$$

10:   **Primary Arm:** Update virtual target pose and impedance gains K_p, K_d using the force feedback factor
11: **until** task complete

IV. EXPERIMENTS

A. Experiment Design.

ACE-F was evaluated according to its performance controlling a Franka Emika Panda robot arm in both virtual and real-world experiments, and a Unitree G1 robot was controlled in simulation to demonstrate ACE-F's cross-platform capabilities. We aimed to answer the following questions through these
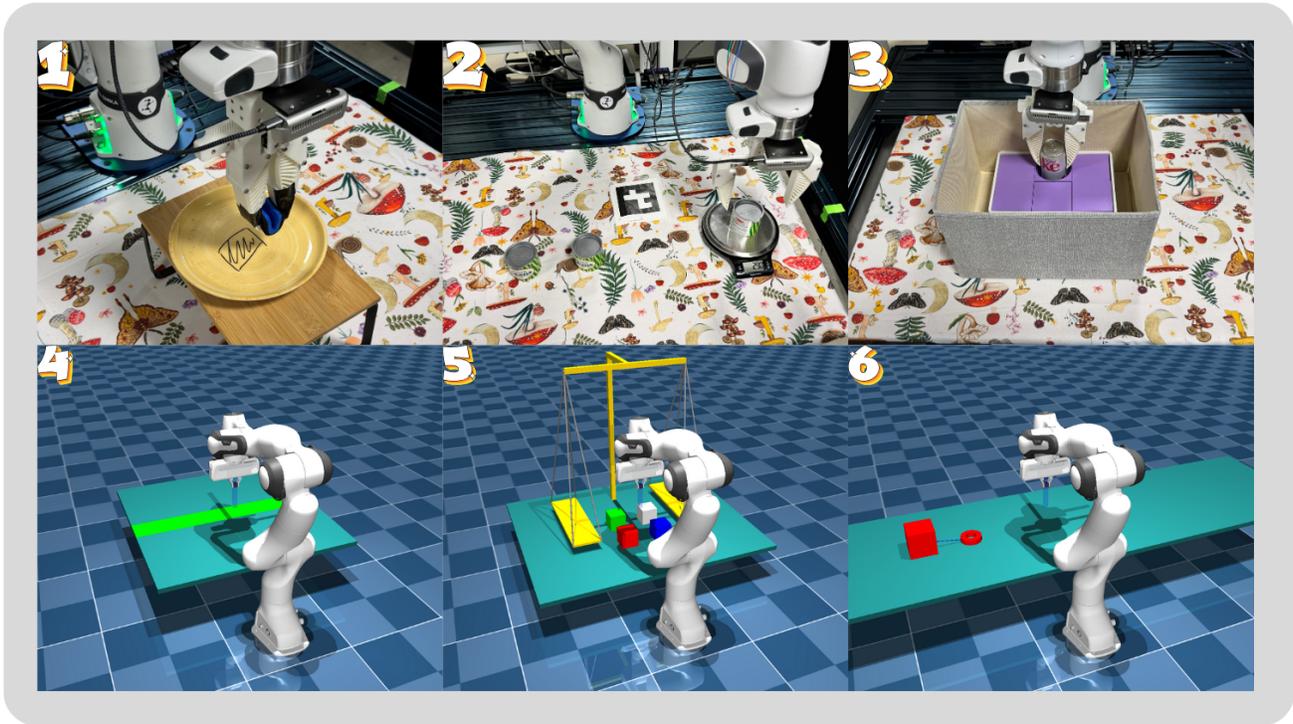
Fig. 4: Overview of the six tasks in the evaluation suite: real-world mopping, real-world can stacking, real-world blind can insertion, simulated table mopping (left-right and forward-backward), simulated box stacking, and simulated box dragging.

trials:

1. How well does ACE-F perform compared to joint-copy methods, like Gello, in both simulation and real-world environments?

2. How effective is the sensor-less force feedback in improving task performance and operator awareness?

3. How well does ACE-F perform under real-world physical constraints, when compared to its simulation?

1) Ablation Study.: Before performing the simulated and real world experiments, we conducted an ablation study to explore how different substitutions of the velocity term in the feedback compensation formula impact system stability and user experience. Specifically, we tested four variations: absolute value $|v|$, squared velocity v^2, exponential $\exp(v)$, and hyperbolic tangent $\tanh(v)$. Table I shows that $|v|$ produced the lowest high frequency energy ratio (0.123), lowest maximum local jerk (0.00064), and highest feedback correlation (0.758), indicating it provides the most stable and precise force feedback. In contrast, $\exp(v)$ introduced the highest high frequency energy ratio (0.211) and jerk anomalies (4.59%), suggesting a more aggressive but potentially destabilizing feedback response. $\tanh(v)$ and v^2 offered intermediate results, with $tanh(v)$ achieving a feedback correlation of 0.662 and relatively low jerk anomalies (0.81%). These findings demonstrate the trade-offs between stability, responsiveness, and user sensitivity across different velocity formulations in the feedback term.

2) Simulated Experiments: First, we compared ACE-F and Gello on three contact-rich simulations built in MuJoCo – Box Stacking, Box Dragging, and Table Mopping. Each task involved varying levels of physical interaction, visual reasoning, and force modulation, which we could use to measure the two platforms. Four test users with varying levels of experience on both systems were selected and placed into one of two groups, A and B. The groups began each task using a different system and switched platforms halfway through the task. Additionally, users were allowed up to five minutes of practice every time they switched platform. Practice sessions were conducted in a task neutral practice arena, which contained elements from each of the three actual experiments. Users were provided with a description of their goal before each task and loaded into the practice scene. Once they reached the end of their practice time or self-determined that they were ready to begin, they were given control of a Franka Emika Panda robotic arm within the scene. A description of all three of these simulated tasks is provided below:

a) Simulated Box Stacking.: This scene consisted of a fixed camera overlooking a stationary raised platform and a Franka Emika Panda robot. Four weighted blocks and a balance scale were placed on the table within the arm's reach, and users were asked to determine the relative weight of each block before stacking them from heaviest to lightest. The users were allowed to use any combination of force feedback, visual cues, and tools available in the simulation and were graded on

TABLE I: Teleoperation Stability Metrics for Different Velocity Transformations

Feedback Param	High Freq. Energy Ratio	Max Local Jerk	Primary-only Jerk Anomaly (%)	Feedback Correlation
$\|v\|$	**0.123**	**0.00064**	**0.00**	**0.758**
v^2	0.132	0.00219	0.39	0.521
$\exp(v)$	0.211	0.00273	4.59	0.686
$\tanh(v)$	0.131	0.00139	0.81	0.662

four metrics: speed, number of scale uses, number of times they knocked one or more blocks off their tower, and whether or not they were successful in stacking all four blocks in the correct order. Each user performed this experiment three times on each system, and the order of the block weights was randomly selected from a pre-generated set of configurations before each attempt. This test should clearly demonstrate one of ACE-F's major advantages over Gello – it gives the user force feedback proportional to the weight of the block in the gripper, possibly allowing users to forgo using the scale to compare block weights and thus accelerating the task.

b) Simulated Box Dragging.: This scene consisted of a fixed camera overlooking a stationary raised platform and a Franka Emika Panda robot. A singular weighted block attached to a handle was placed within the arm's reach, and users were asked to use the handle to drag the block from one end of the table to the other as quickly as possible. Somewhere along the way, they would encounter an invisible obstacle, and they should try to stop the block as soon as they found it, without crossing over it. They were allowed to use any combination of force feedback and visual cues available in the simulation and were graded on three metrics: speed, distance from the obstacle, and whether or not they were successful in stopping the box before it crossed the line. Each user performed this experiment three times on each system as above, and an additional three times on ACE-F with their vision disabled after they gripped onto the block's handle. The location of the hidden obstacle was changed between each experiment by randomly selecting the scene from a pre-generated set of configurations. ACE-F allows users to feel the friction of the table and the sudden force imposed on the robotic arm when the box collides with the hidden obstacle, which should give them the ability to detect the ledge more confidently with and without visual cues. This is not possible under Gello's framework.

c) Simulated Table Mopping.: This scene consisted of a fixed camera overlooking a stationary raised platform and a Franka Emika Panda robot. Users performed this experiment three times per robotic platform on two different table configurations. The table in configuration one was marked with a strip of green from left to right and the table in configuration two was marked from the closest edge to the farthest edge. Users were instructed to drag the robot's end-effector back and forth along the line from one edge of the table to the other twice per table configuration, while trying to maintain steady speed and pressure on the table. They were allowed to use any force feedback and visual cues available to them and were graded on their speed and force consistency, defined

as the variation of the normal forces experienced by the table in the simulation. Since this task does not benefit from prior knowledge, this experiment did not require a randomized scene structure. ACE-F should demonstrate a clear advantage over Gello because the user is able to "feel" the table to provide a more consistent force input to the system.

3) Real-World Experiments: Next, we compared ACE-F and Gello in the real-world on the following tasks: Can Stacking, Marker Erasing, and Hidden Insertion. Once again, these tasks involve physical interactions, visual reasoning, and force modulation. Due to the limited availability of testing equipment and concerns regarding the safety of inexperienced volunteers, only two testers were selected from the previous group of four. Users were once again allowed up to five minutes of practice every time they switched platform, and practice sessions were conducted in the same environment where the tests took place. Once users reached the end of their practice time or self-determined that they were ready to begin, they were recorded attempting the task three times. A description of all three real tasks is provided below:

a) Real Can Stacking.: Similar to the simulated stacking scene, this experiment consisted of three weighted cans and a digital balance scale. Users were instructed to determine the relative weight of each can before stacking them from heaviest to lightest. The users were allowed to use any combination of force feedback, visual cues, and tools available to them and were graded on four metrics: speed, number of scale uses, number of times they knocked one or more cans off their tower, and whether or not they were successful in stacking all three cans in the correct order. Each user performed this experiment three times on each system, and the order of the block weights was scrambled by a helper between each attempt. ACE-F should be able to stack the blocks more confidently and quickly, since the users can detect the cans' weight differences without using the scale.

b) Real Marker Erasing.: Similar to the simulated mopping scene, this experiment consisted of a raised platform holding a textured ceramic dish with a 2-inch by 2-inch square of dry-erase marker. Users were instructed to use a whiteboard eraser, which would start clamped between the robot's grippers, to remove the visible markings from the plate. Testers were allowed to use any combination of force feedback, visual cues, and tools available. They were graded on three metrics: speed, the number of times they triggered Franka's built-in safety features, and whether or not they were successful in removing all of the marker from the plate. Each user performed this experiment three times on each system. This task demonstrates ACE-F's ability to generate consistent

TABLE II: Aggregate simulation performance across three teleoperation tasks. ACE-F demonstrates advantages over Gello in every task.

Virtual Stacking Task			
Method	**Avg. Time (s)**	**Avg. Scale Uses**	**Success Rate (%)**
ACE	**102.1 ± 27.2**	**0.7 ± 0.5**	**90.0**
Gello	187.0 ± 121.2	2.3 ± 0.7	70.0

Virtual Box Dragging Task			
Method	**Avg. Time (s)**	**Light Status**	**Success Rate (%)**
ACE	16.7 ± 2.5	On	**100**
ACE	**16.4 ± 2.5**	Off	94.4
Gello	21.9 ± 4.8	On	75.0

Virtual Mopping Task: Left-Right		
Method	**Normal Force**	**Max/Avg. Force Ratio**
ACE	**196.8 ± 55.9**	**4.59**
Gello	231.9 ± 69.6	7.19

Virtual Mopping Task: Forward-Backward		
Method	**Normal Force**	**Max/Avg. Force Ratio**
ACE	**229.0 ± 146.8**	**7.19**
Gello	251.7 ± 183.1	9.73

TABLE III: Aggregate real-world performance across three teleoperation tasks. ACE-F consistently outperforms the joint-copy Gello method in success rate and stability, while reducing reliance on external tools like scales.

Real Stacking Task			
Method	**Avg. Time (s)**	**Avg. Scale Uses**	**Success Rate (%)**
ACE	90.81 ± 17.97	**0.0 ± 0.00**	**83.3**
Gello	**88.36 ± 23.46**	2.0 ± 0.63	66.7

Real Erasing Task			
Method	**Avg. Time (s)**	**# Safety Warnings**	**Success Rate (%)**
ACE	26.54 ± 7.89	**0**	100.0
Gello	**22.13 ± 4.30**	1	100.0

Real Blind Can Insertion Task		
Method	**Avg. Time (s)**	**Success Rate (%)**
ACE	43.26 ± 20.87	**100.0**
Gello	**34.42 ± 12.30**	50.0

forces for precision tasks.

c) Real Blind Can Insertion.: This task was performed exclusively in the real-world and made use of a rectangular box with interchangeable blocks that could be oriented so there was only one can-sized hole at the top of the box. It was placed behind a cardboard screen that blocked the robot operator from being able to see where the hole was located. Users were instructed to pick up a can from the visible portion of the table and move it behind the screen, where they needed to place the can into the open slot. They were allowed to use visual cues to grab and move the can to behind the screen, however due to their severely limited vision of the task itself, they mostly had to rely on force feedback when their device provided it. Each user performed this experiment three times on each system and was graded on their speed, the number of times they triggered Franka's built-in safety features, and whether or not they were successful in placing the can into the hole. ACE-F's force feedback gives users an additional sense, which is incredibly useful for blind tasks. It should allow users to find the correct location for the can more quickly and release it with a higher degree of confidence, since Gello can only operate by visuals.

B. Experimental Results.

1) Simulated Experiments:

a) Simulated Box Stacking.: When we evaluated both platforms on their speed, number of scale uses, number of tower topples (labeled as blunders in Table II), and success rate, ACE-F demonstrated a clear advantage over Gello in all categories. Not only were ACE-F testers able to complete the stacking task 54.62% faster than Gello users, they were also able to do so far more consistently (a quarter the standard deviation). Additionally, ACE-F users were 28.57% more successful, despite using the scale less than half as much as

Gello users. Since ACE-F allows users to feel the weight of the cube without having to rely on the scale, users spent less time testing each block. This also means ACE-F users could limit their motions to a smaller area of the task space, which contributed to the large time difference. The reduced motion and ACE-F's high-precision, gravity compensated control could also explain why ACE-F users made fewer blunders during stacking, since they had to do less work controlling the arm configuration and could use a softer touch.

b) Simulated Box Dragging.: ACE-F performed significantly better than Gello in the box dragging task, where it was consistently 23.72% faster than the same tests performed on Gello, even without vision for the latter half of the task. It was also far more successful, only failing the task in 5.6% of the blind tasks, compared to Gello's 25% failure rate with the lights on. This test in particular demonstrates the advantages of ACE-F in tasks where tactile feedback is more useful than visual cues alone.

c) Simulated Table Mopping.: ACE-F outperformed Gello in both configurations of the mopping task, as well. By providing the user with an even force when the arm collides with the table, the user is able to sense how much force they are applying and more easily regulate their downward pressure. This is clearly indicated by a 36.16% smaller maximum force to average normal force ratio in the left-right configuration and a 41.86% smaller ratio in the forward-backward test.

2) Real-World Experiments:

a) Real Can Stacking.: Gello performed the can stacking task faster on average, however it had a larger standard deviation in completion times and a lower success rate. This can be attributed to two things: the platform's unstable configuration and the increased movements associated with moving cans to the scale. ACE-F did not have to use the scale in any of its tests since the user could test each can's weight purely through force feedback, so it avoided making large movements which could shake the can loose from its gripper. Similarly, Gello's structure makes it difficult for the user to keep the gripper perfectly upright, which contributed to the cans falling from

its grasp more frequently. If the can landed in an awkward orientation, it could roll outside of the robot's work space, causing it to immediately fail. Sometimes, this happened early on in the task, which also explains Gello's faster task average completion time.

b) Real Marker Erasing.: Both ACE-F and Gello had a 100% success rate in this task, however Gello performed 16.62% faster than ACE-F on average and had a smaller standard deviation in its times. Compared to the previous task, where its easy rotations were a disadvantage, Gello users could rotate the whiteboard eraser so that it was better able to conform to the textured surface of the plate and remove the hard-to-reach marks more quickly. Gello users could also move the arm more quickly since there was zero resistance to their movements. This accelerated their performance but also triggered a warning in Franka's safety system during one of the tasks.

c) Real Blind Can Insertion.: ACE-F performed significantly better than Gello in this task because the force feedback allowed users to compensate for their poor vision by feeling around inside the box. Gello users generally completed this task faster, however they were only successful 50% of the time, while ACE-F users correctly identified the can's target location 100% of the time. This is likely because ACE-F users could modulate how much force they applied when inserting the can, which prevented them from pressing too hard and losing their grasp on the can.

V. DISCUSSION AND CONCLUSION

ACE-F displayed clear advantages over joint-copy methods, like Gello, when compared in virtual and real-world environments. It was capable of performing at least as well as Gello, if not better in every task because its inverse-kinematic controller removes the burden of monitoring the robot's configuration from the operator and the force-feedback from the inverse-dynamics controller gives the user an extra sense, which improves environmental awareness. Additionally, by reducing the complexity of the overall system, ACE-F remains very compact and portable compared to other teleoperation systems, which is beneficial when the user wants to operate a mobile robot, like the Unitree G1. The sensor-less force feedback also enables users to complete tasks in an entirely new domain – one where the user has limited vision of their workspace – as demonstrated by the virtual box dragging task and the real-world hidden can insertion task. This solves one of the major drawbacks of joint-copy methods and results in a surprise benefit by limiting user speed through resistance forces, which reduces the likelihood of the user triggering speed-based warnings. Lastly, ACE-F was able to perform at a similar level both in and out of simulation, which is beneficial for training new models through imitation learning. This reliability was demonstrated by the consistency of its friction and gravity compensation models in both environments. ACE-F comes with its own drawbacks, however. The XM430-W350 motors can only generate up to 46 rpm under no load, which is relatively slow for most robot systems. Fortunately, this is sufficient for most tasks because the motors rarely need to move on their own. Instead, we use the motors to apply impedance to the user's motion, which provides two main benefits: the motor does not need to generate high speeds and is usually operating in a stalled state, so it can apply up to its stalled torque (4.1 Nm). Nevertheless, better motors could supply stronger feedback to the user, as long as potential safety concerns are addressed. Also, this project prioritizes low-cost teleoperation over full 6-DoF force feedback, since cartesian forces are generally considered sufficient for most tasks [43]. That leaves a gap in the scope of our project, which can be improved in future releases. Additionally, initial implementations using the rotation glove were spotty due to poor calibration, requiring more attention. Future works should focus on improving the implementation of the gloves, replacing current hardware with more precise devices, and incorporating torque feedback for the user in addition to the cartesian forces currently provided as force feedback.

REFERENCES

[1] Sachith Abeywardena, Qi Yuan, Anna Tzemanaki, Electra Psomopoulou, Leonidas Droukas, Chris Melhuish, and Sanja Dogramadzi. Estimation of tool–tissue forces in robot-assisted minimally invasive surgery using neural networks. *Frontiers in Robotics and AI*, 6:56, 2019. doi: 10.3389/frobt.2019.00056.

[2] Ahmad Abiri, John Pensa, Anna Tao, Ji Ma, Yuan-Yu Juo, Syed J. Askari, James Bisley, Jacob Rosen, Erik Dutson, and Warren Grundfest. Multi-modal haptic feedback for grip force reduction in robotic surgery. *Scientific Reports*, 9(1):5016, 2019. doi: 10.1038/s41598-019-40821-1.

[3] Reuben M Aronson and Henny Admoni. Gaze complements control input for goal prediction during assisted teleoperation. In *Robotics science and systems*, 2022.

[4] Sridhar Pandian Arunachalam, Sneha Silwal, Ben Evans, and Lerrel Pinto. Dexterous imitation made easy: A learning-based framework for efficient dexterous manipulation. In *2023 ieee international conference on robotics and automation (icra)*, pages 5954–5961. IEEE, 2023.

[5] Artem Bazhenov, Sergei Satsevich, Sergei Egorov, Farit Khabibullin, and Dzmitry Tsetserukou. Echo: An open-source, low-cost teleoperation system with force feedback for dataset collection in robot learning. *arXiv preprint arXiv:2504.07939*, 2025.

[6] Qingwei Ben, Feiyu Jia, Jia Zeng, Junting Dong, Dahua Lin, and Jiangmiao Pang. Homie: Humanoid loco-manipulation with isomorphic exoskeleton cockpit. *arXiv preprint arXiv:2502.13013*, 2025.

[7] Domenico Buongiorno, Domenico Chiaradia, Simone Marcheschi, Massimiliano Solazzi, and Antonio Frisoli. Multi-dofs exoskeleton-based bilateral teleoperation with the time-domain passivity approach. *Robotica*, 37(9): 1641–1662, 2019.

[8] Manuel Caeiro-Rodríguez, Iván Otero-González, Fernando A Mikic-Fonte, and Martín Llamas-Nistal. A

systematic review of commercial smart gloves: Current status and applications. *Sensors*, 21(8):2667, 2021.

[9] Xuxin Cheng, Jialong Li, Shiqi Yang, Ge Yang, and Xiaolong Wang. Open-television: Teleoperation with immersive active visual feedback. *arXiv preprint arXiv:2407.01512*, 2024.

[10] Runyu Ding, Yuzhe Qin, Jiyue Zhu, Chengzhe Jia, Shiqi Yang, Ruihan Yang, Xiaojuan Qi, and Xiaolong Wang. Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning. *arXiv preprint arXiv:2407.03162*, 2024.

[11] Frederik Ebert, Yanlai Yang, Karl Schmeckpeper, Bernadette Bucher, Georgios Georgakis, Kostas Daniilidis, Chelsea Finn, and Sergey Levine. Bridge data: Boosting generalization of robotic skills with cross-domain datasets. *arXiv preprint arXiv:2109.13396*, 2021.

[12] Armin Ehrampoosh, Bijan Shirinzadeh, Joshua Pinskier, Julian Smith, Randall Moshinsky, and Yongmin Zhong. A force-feedback methodology for teleoperated suturing task in robotic-assisted minimally invasive surgery. *Sensors*, 22(20):7829, 2022. doi: 10.3390/s22207829.

[13] Issam El Rassi and Jean-Michel El Rassi. A review of haptic feedback in tele-operated robotic surgery. *Journal of Medical Engineering & Technology*, 44(5):247–254, 2020. doi: 10.1080/03091902.2020.1772391.

[14] Zicong Fan, Omid Taheri, Dimitrios Tzionas, Muhammed Kocabas, Manuel Kaufmann, Michael J Black, and Otmar Hilliges. Arctic: A dataset for dexterous bimanual hand-object manipulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12943–12954, 2023.

[15] Hongjie Fang, Hao-Shu Fang, Yiming Wang, Jieji Ren, Jingjing Chen, Ruo Zhang, Weiming Wang, and Cewu Lu. Airexo: Low-cost exoskeletons for learning whole-arm manipulation in the wild. *arXiv preprint arXiv:2309.14975*, 2023.

[16] Zipeng Fu, Tony Z. Zhao, and Chelsea Finn. Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation. *arXiv preprint arXiv:2401.02117*, 2024.

[17] Shangchen Han, Beibei Liu, Randi Cabezas, Christopher D Twigg, Peizhao Zhang, Jeff Petkau, Tsz-Ho Yu, Chun-Jung Tai, Muzaffer Akbay, Zheng Wang, et al. Megatrack: monochrome egocentric articulated hand-tracking for virtual reality. *ACM Transactions on Graphics (ToG)*, 39(4):87–1, 2020.

[18] Ankur Handa, Karl Van Wyk, Wei Yang, Jacky Liang, Yu-Wei Chao, Qian Wan, Stan Birchfield, Nathan Ratliff, and Dieter Fox. Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9164–9170. IEEE, 2020.

[19] Mahdi Hejrati, Pauli Mustalahti, and Jouni Mattila. Robust immersive bilateral teleoperation of dissimilar systems with enhanced transparency and sense of embodiment. *arXiv preprint arXiv:2505.14486*, 2025.

[20] Adam Imdieke and Karthik Desingh. Spark-remote: A cost-effective system for remote bimanual robot teleoperation. *arXiv preprint arXiv:2504.05488*, 2025.

[21] Yasuhiro Ishiguro, Tasuku Makabe, Yuya Nagamatsu, Yuta Kojio, Kunio Kojima, Fumihito Sugai, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Bilateral humanoid teleoperation system using whole-body exoskeleton cockpit tablis. *IEEE Robotics and Automation Letters*, 5(4):6419–6426, 2020.

[22] Alaa Khalifa, Mohamed Fanni, and Ahmed Khalifa. Sensorless contact force estimation and robust impedance control for a quadrotor manipulation system. *Scientific Reports*, 14(1):28772, 2024.

[23] Jonathan Kofman, Xianghai Wu, Timothy J Luu, and Siddharth Verma. Teleoperation of a robot manipulator using a vision-based human-robot interface. *IEEE transactions on industrial electronics*, 52(5):1206–1219, 2005.

[24] Vikash Kumar and Emanuel Todorov. Mujoco haptix: A virtual reality system for hand manipulation. In *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pages 657–663. IEEE, 2015.

[25] Hooman Lee, Joongbae Kim, and Taewoo Kim. A robot teaching framework for a redundant dual arm manipulator with teleoperation from exoskeleton motion data. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 1057–1062. IEEE, 2014.

[26] Danilo Leonardis, Giovanni Furferi, and Luciano Governi. A novel exoskeleton system integrating an exoskeleton glove for teleoperation. *Robotics*, 13(8):119, 2024.

[27] Jialong Li, Xuxin Cheng, Tianshu Huang, Shiqi Yang, Ri-Zhao Qiu, and Xiaolong Wang. Amo: Adaptive motion optimization for hyper-dexterous humanoid whole-body control. In *Proceedings of Robotics: Science and Systems (RSS)*, 2025.

[28] Shuang Li, Xiaojian Ma, Hongzhuo Liang, Michael Görner, Philipp Ruppel, Bin Fang, Fuchun Sun, and Jianwei Zhang. Vision-based teleoperation of shadow dexterous hand using end-to-end deep neural network. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 416–422. IEEE, 2019.

[29] Shuang Li, Jiaxi Jiang, Philipp Ruppel, Hongzhuo Liang, Xiaojian Ma, Norman Hendrich, Fuchun Sun, and Jianwei Zhang. A mobile robot hand-arm teleoperation system by vision and imu. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10900–10906. IEEE, 2020.

[30] Hangxin Liu, Zhenliang Zhang, Xu Xie, Yixin Zhu, Yue Liu, Yongtian Wang, and Song-Chun Zhu. High-fidelity grasping in virtual reality using a glove-based system. In *2019 international conference on robotics and automation (icra)*, pages 5180–5186. IEEE, 2019.

[31] Jason Jingzhou Liu, Yulong Li, Kenneth Shaw, Tony Tao, Ruslan Salakhutdinov, and Deepak Pathak. Factr: Force-attending curriculum training for contact-rich policy learning. *arXiv preprint arXiv:2502.17432*, 2025.

[32] Chenhao Lu, Xuxin Cheng, Jialong Li, Shiqi Yang, Mazeyu Ji, Chengjing Yuan, Ge Yang, Sha Yi, and Xiaolong Wang. Mobile-television: Predictive motion priors for humanoid whole-body control. *arXiv preprint arXiv:2412.07773*, 2024.

[33] Honghao Lv, Depeng Kong, Gaoyang Pang, Baicun Wang, Zhangwei Yu, Zhibo Pang, and Geng Yang. Gulim: A hybrid motion mapping technique for teleoperation of medical assistive robot in combating the covid-19 pandemic. *IEEE Transactions on Medical Robotics and Bionics*, 4(1):106–117, 2022.

[34] Xin Ma, Jianshu Zhou, and Masayoshi Tomizuka. Few-shot sim2real based on high fidelity rendering with force feedback teleoperation. *arXiv preprint arXiv:2503.01301*, 2025.

[35] Johanna Miller, Manuel Braun, Johannes Bilz, Sebastian Matich, Carsten Neupert, Wolfgang Kunert, and Andreas Kirschniak. Impact of haptic feedback on applied intracorporeal forces using a novel surgical robotic system—a randomized cross-over study with novices in an experimental setup. *Surgical Endoscopy*, 35(8):3554–3563, 2021. doi: 10.1007/s00464-020-07818-8.

[36] Takehiko Ohkawa, Kun He, Fadime Sener, Tomas Hodan, Luan Tran, and Cem Keskin. Assemblyhands: Towards egocentric activity understanding via 3d hand pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12999–13008, 2023.

[37] Rodrigo Pérez Úbeda, Sebastián C. Gutiérrez-Rubert, Radovan Zotović-Stanišić, and Álvaro Perles-Ivars. Design and manufacturing of an ultra-low-cost custom torque sensor for robotics. *Sensors*, 18(6):1786, 2018.

[38] Yuzhe Qin, Hao Su, and Xiaolong Wang. From one hand to multiple hands: Imitation learning for dexterous manipulation from single-camera teleoperation. *IEEE Robotics and Automation Letters*, 7(4):10873–10881, 2022.

[39] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dieter Fox. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. In *Robotics: Science and Systems (RSS)*, Daegu, Republic of Korea, July 2023.

[40] Yu Rong, Takaaki Shiratori, and Hanbyul Joo. Frankmocap: Fast monocular 3d hand and body motion capture by regression and integration. *arXiv preprint arXiv:2008.08324*, 2020.

[41] Omid Taheri, Nima Ghorbani, Michael J Black, and Dimitrios Tzionas. Grab: A dataset of whole-body human grasping of objects. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16*, pages 581–600. Springer, 2020.

[42] Alexander Toedtheide, Xiao Chen, Hamid Sadeghian, Abdeldjallil Naceri, and Sami Haddadin. A force-sensitive exoskeleton for teleoperation: An application in elderly care robotics. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12624–12630. IEEE, 2023.

[43] Lawton N. Verner and Allison M. Okamura. Force torque feedback vs force only feedback. In *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 406–410, 2009. doi: 10.1109/WHC.2009.4810880.

[44] Chen Wang, Haochen Shi, Weizhuo Wang, Ruohan Zhang, Fei-Fei Li, and C. Karen Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. *arXiv preprint arXiv:2403.07788*, 2024.

[45] Philipp Wu, Yide Shentu, Zhongke Yi, Xingyu Lin, and Pieter Abbeel. Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators. *arXiv preprint arXiv:2309.13037*, 2023.

[46] Shiqi Yang, Minghuan Liu, Yuzhe Qin, Runyu Ding, Jialong Li, Xuxin Cheng, Ruihan Yang, Sha Yi, and Xiaolong Wang. Ace: A cross-platform visual-exoskeletons system for low-cost dexterous teleoperation. *arXiv preprint arXiv:2408.11805*, 2024.

[47] Yanjie Ze, Zixuan Chen, João Pedro Araújo, Zi-ang Cao, Xue Bin Peng, Jiajun Wu, and C. Karen Liu. Twist: Teleoperated whole-body imitation system. *arXiv preprint arXiv:2505.02833*, 2025.

[48] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*, 2020.

[49] Han Zhang, Songbo Hu, Zhecheng Yuan, and Huazhe Xu. Doglove: Dexterous manipulation with a low-cost open-source haptic force feedback glove. *arXiv preprint arXiv:2502.07730*, 2025.

[50] Liang Zhao, Tingting Yang, Yang Yang, and Peng Yu. A wearable upper limb exoskeleton for intuitive teleoperation of anthropomorphic manipulators. *Machines*, 11(4):441, 2023.

[51] Tony Z. Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. *arXiv preprint arXiv:2304.13705*, 2023.

VI. SUPPLEMENTARY
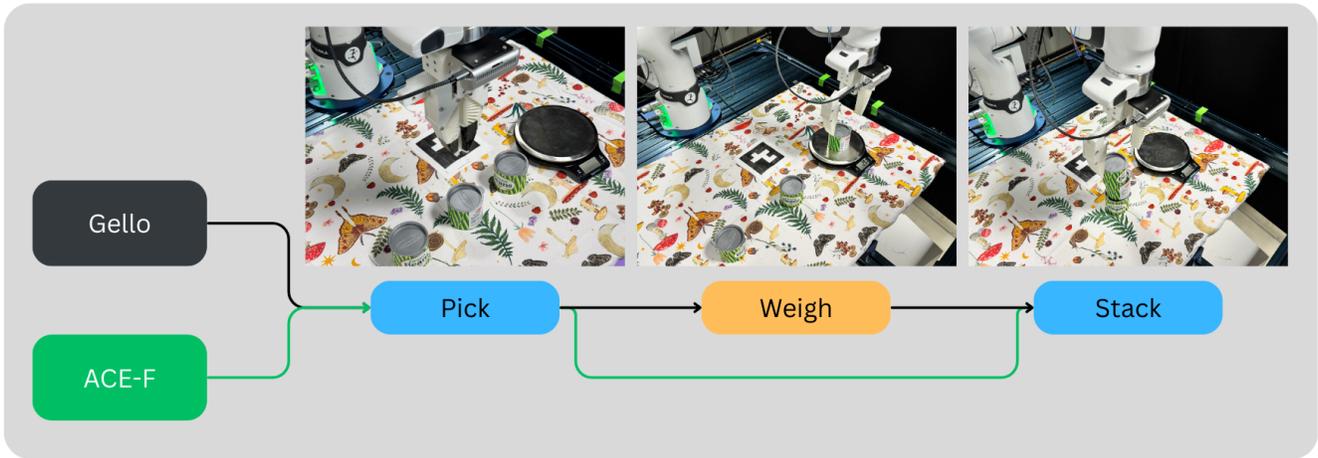
Fig. 5: Real Can Stacking task setup: Users stack three weighted cans in descending order, leveraging available force feedback, visual cues, and tools. Performance is measured by stacking speed, number of scale uses, stability, and stacking accuracy. ACE-F can detect the weight differences – it does not need extra tools so it saves more time.
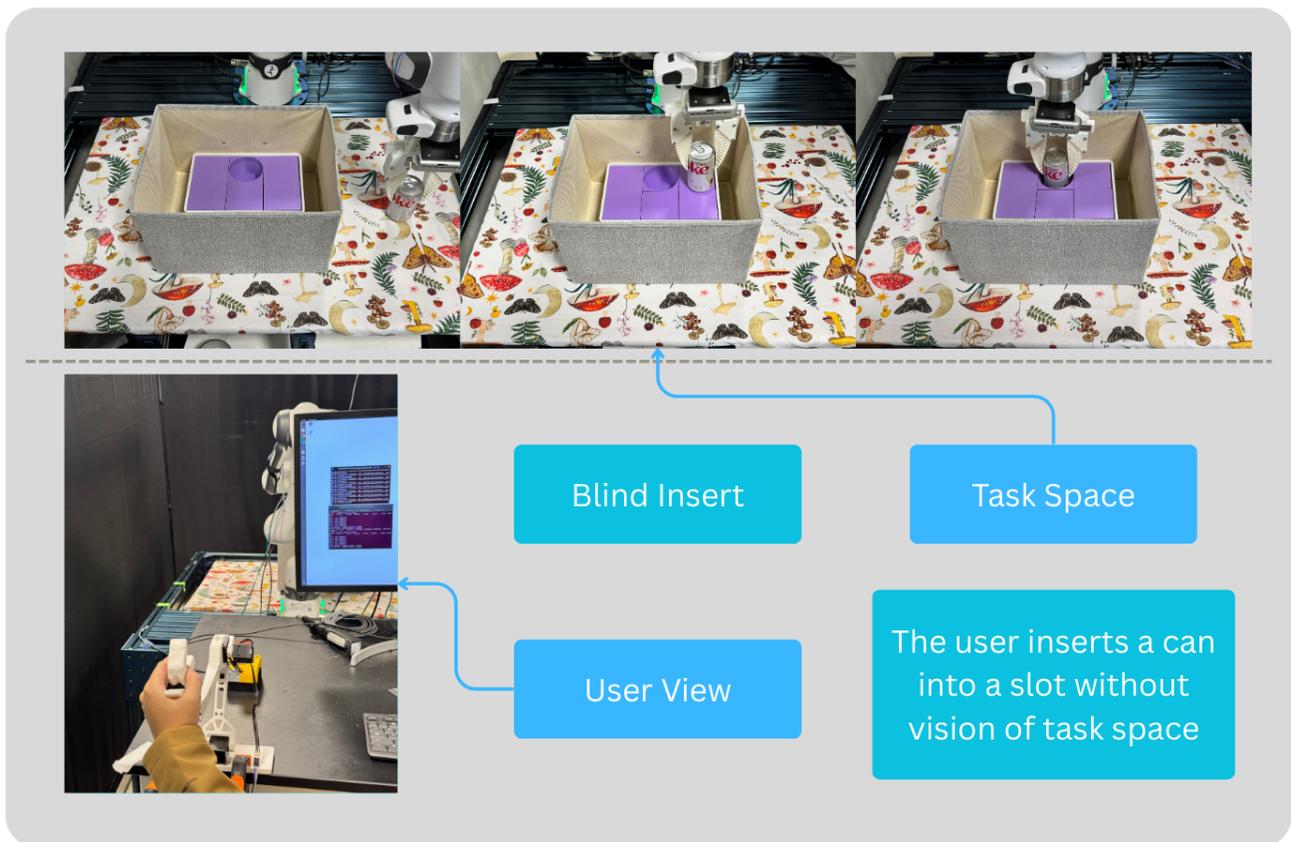
Fig. 6: Real Blind Can Insertion task: Users must insert a can into a concealed hole behind a screen, primarily relying on force feedback for accurate alignment and insertion.