

GELLO: A General, Low-Cost, and Intuitive Teleoperation Framework for Robot Manipulators

Philipp Wu Yide Shentu Zhongke Yi Xingyu Lin Pieter Abbeel
University of California, Berkeley

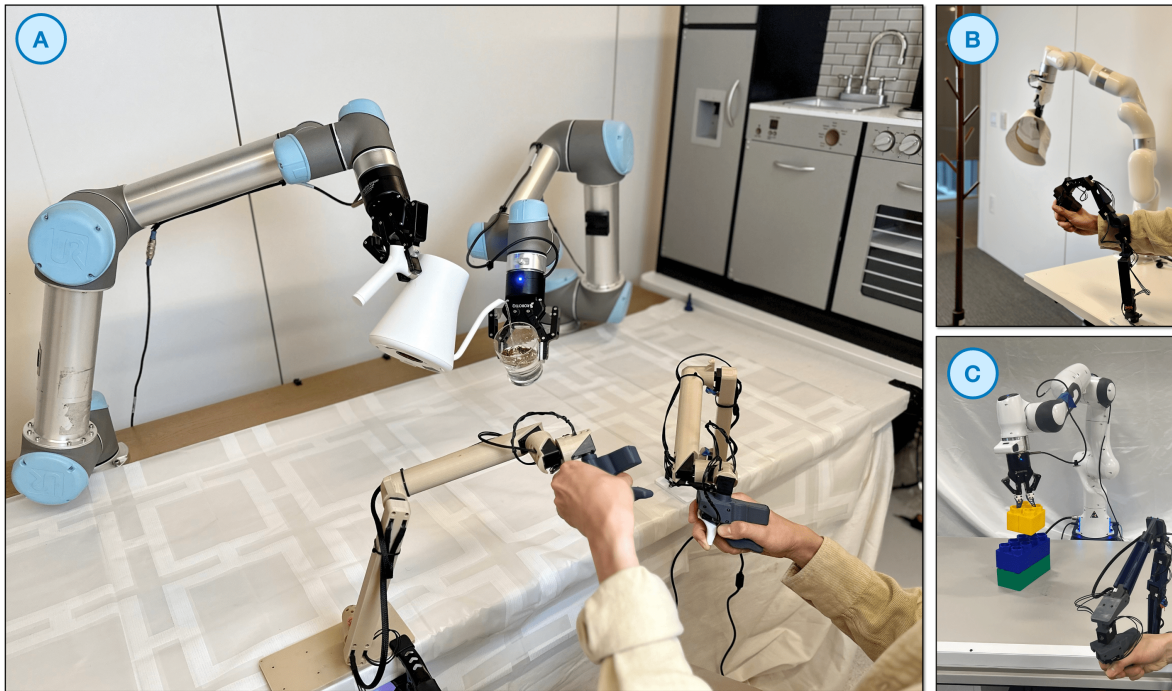


Fig. 1: We show GELLO teleoperation systems built for three different types of robots: (A) two UR5s, (B) an xArm, and (C) a Franka. The user can teleoperate the robot arms by controlling the GELLO devices. The bill of materials for each GELLO device is less than \$300.

Abstract—Humans can teleoperate robots to accomplish complex manipulation tasks. Imitation learning has emerged as a powerful framework that leverages human teleoperated demonstrations to teach robots new skills. However, the performance of the learned policies is bottlenecked by the quality, scale, and variety of the demonstration data. In this paper, we aim to lower the barrier to collecting large and high-quality human demonstration data by proposing a GEneral framework for building LOw-cost and intuitive teleoperation systems for robotic manipulation (GELLO). Given a target robot arm, we build a GELLO controller device that has the same kinematic structure as the target arm, leveraging 3D-printed parts and economical off-the-shelf motors. GELLO is easy to build and intuitive to use. Through an extensive user study, we show that GELLO enables more reliable and efficient demonstration collection compared to other cost efficient teleoperation devices commonly used in the imitation learning literature such as virtual reality controllers and 3D spacemouses. We further demonstrate the capabilities of GELLO for performing complex bi-manual and contact-rich manipulation tasks. To make GELLO accessible to everyone, we have designed and built GELLO systems for 3 commonly used robotic arms: Franka, UR5, and xArm. All software and hardware are open-sourced and can be found on our website: <https://wuphilipp.github.io/gello/>.

I. INTRODUCTION

In recent years, robotics has gone through a remarkable transformation driven by the increasing integration of data-

driven methods in every component, ranging from perception to control. The ability to learn from diverse data helps robots generalize to new scenarios which would be difficult to achieve by a manually designed system. For manipulation, there have been great successes in leveraging imitation learning [1], [2], [3], [4] where system performance improves with larger datasets [5], [6], [7], [8]. However, existing systems are still bottlenecked by the dataset size as well as the complexity and diversity of the tasks within the dataset.

Human teleoperation has consistently proven to be one of the best ways to control robot manipulators to collect demonstrations for imitation learning. Teleoperation has a long history and enabled impressive robot capabilities in a diverse range of challenging situations such as fine grain manipulation of everyday tasks [9], robot surgery [10], [11], [12], underwater exploration [13], [14], and disaster relief [15], [16], all of which would be useful skills for a robot to learn. A common feature of the teleoperation devices used in these systems is their bilateral capability, which enables the operator to not only direct commands at the robot but also receive force feedback from the robot’s environment. While performant, these teleoperation systems typically require high-fidelity sensors due to their strict system requirements, leading to higher costs that ultimately limit accessibility.

As a result, a more commonly used approach is to build a teleoperation system that captures control signals from lower-cost commodity electronic devices like 3D-mouse [17], VR controllers [18], [19] or cameras, which is then converted to a robot action. However, these systems abstract away the kinematic constraints of the robot and can be unintuitive to new users. “A Low-cost Open-source Hardware System for Bimanual Teleoperation” (ALOHA) presents a teleoperation system leveraging an off-the-shelf servo-based arm to control a manipulator of similar size and kinematics, showing impressive teleoperation capabilities for fine-grained manipulation tasks despite being unilateral [9]. Nevertheless, the ALOHA system is tailored to a specific robot arm and has a higher cost due to having additional robot arms as controllers for the user.

In this paper, we introduce GELLO, a **GE**neral, **LO**w-cost, and intuitive teleoperation framework for robot manipulators, targeted towards enabling a scalable way to collect demos by exploring the limits of affordability and simplicity. GELLO is designed to be low-cost, easy to build, and intuitive for humans to use. The key principle is to build miniature, kinematically equivalent controllers with 3D-printed parts and off-the-shelf motors as joint encoders. We clarify that the ideas behind GELLO are not new, rather our contributions can be summarized in the three points below:

- 1) We present practical implementations of GELLO as a teleoperation system for three commonly used robot arms with simple and low-cost designs.
- 2) We perform a comprehensive user study demonstrating the system’s effectiveness compared to other prevalent low-cost teleoperation systems in the literature.
- 3) We fully open-source the hardware and software needed to replicate and operate GELLO to ensure accessibility for the research community to build upon, complete with a bill of materials and assembly instructions¹.

II. RELATED WORK

A. Teleoperation Systems for Manipulation

Low-cost Controllers. Teleoperation systems have a long-standing history and various low-cost sensors have been used to provide the interface for human-robot interaction. Commonly used teleoperation systems include joysticks and spacemouses [20], [21], [22], commercial VR controllers [4], [23], [24], [25], RGB cameras [26], [27], [28], [29] or IMU sensors [30], [31], [32]. However due to the morphological differences between these control devices and the robots, the user often can only perform teleoperation in the more abstracted end-effector space. The kinematic constraints of the robot arms therefore are not perceived by the human operators. This prevents the operator from precisely controlling the arm near the areas of kinematic singularities and self-collisions, which reduces the demonstration throughput and increases failures. Moreover, both VR and camera-based solutions can suffer from occlusion and additional latency.

¹Website <https://wuphilipp.github.io/gello/>

TABLE I: An approximate cost comparison of the price of commonly used teleoperation systems for robot learning research. In our user study, we show that GELLO compares favorably to other low-cost options (e.g. spacemouse and VR) while being orders of magnitudes cheaper than other options.

Teleop Device	Approximate Cost
3D Mouse (SpaceNavigator[17])	\$150
GELLO (Ours)	\$300
VR (Meta Quest 2) [18]	\$300
Robot-to-robot Teleop (e.g. UR5)	\$30,000
Haptic Device (Omega7 [45])	\$40,000

Notably, the recent ALOHA system showcases impressive fine-grained bi-manual manipulation with Dynamixel-based servo arms, where an additional two arms of similar form factor function as controllers [9]. Similar to other more conventional but costly teleoperation systems [33], [34], [35], the teleoperation device is another fully-fledged robot arm, with size and capability comparable to the manipulator arm, which can increase the cost. In comparison, we use low-cost components to design a scaled replica of the target arm, resulting in an economical solution that still maintains the advantages of using a kinematically isomorphic arm as the controller.

Bilateral Teleoperation Systems. In contrast to unilateral teleoperation approaches, bilateral teleoperation enables the user to feel force feedback from the target arm [36], [37], an active area of research with a wide range of methodologies. One such approach uses additional robot arms that are isomorphic to the target robots serving as controllers [33], [34], [35], [38], [39], [40]. This approach enables environment feedback from the manipulator to be directly relayed back to the joints of the control device. Also, the user can easily sense the kinematic constraints of the robot arm, as the controller arm has the same kinematic constraints. Additionally, there have been efforts to design 1-to-1 exoskeletons for teleoperation purposes [41], [42], [43]. These systems, though effective, are typically bespoke for specific robots, leading to a varied design approach across different robot types. Another avenue explored in the literature is the use of special input devices with haptic feedback [44], [45]. While these devices offer a tangible sense of the robot’s kinematic constraints, they often have a very tight operation space and additionally require translating the robot’s kinematic constraints into tangible force feedback increasing system complexity [46], [13]. In contrast to these approaches, our contribution presents a generalized framework for designing affordable and easily accessible exoskeleton-like unilateral controllers. We provide instances of our approach for three widely-used robot arms. Our system, GELLO, stands out for its affordability, portability, and replicability, reducing the challenges associated with collecting quality teleoperated human demonstrations.

B. Learning from Human Demonstrations

Learning from demonstrations has been a popular framework for enabling robots to perform a wide range of tasks [5], [6], [7], [4], [29], [47], [48]. Prior works have observed

that the performance of the learning system scales with the size of the dataset. As such, there are substantial ongoing efforts at collecting larger and larger datasets [49], [50]. However, collecting human demonstrations can be expensive and time-consuming. For example, the data collection process as done by Brohan *et al.*[5] spanned over 17 months with a team of researchers. On the other hand, significant efforts have been made towards better human-robot interaction to address the bottleneck. Some examples include sharing control between the human and the robot [20], or enabling a human to operate multiple robots simultaneously [51]. These approaches are complementary to our objective, which is to build teleoperation systems that are more accessible and intuitive to use. Finally, a promising direction is learning directly from human videos [52], [53], [54], [55], [56], [57]. While collecting videos of humans performing the tasks directly is relatively inexpensive, overcoming the morphology gap between robots and humans remains challenging.

III. TELEOPERATION DEVICE DESIGN

The focus for the design of GELLO is to create an interface that is both economically accessible and easy-to-use for users aiming to render high-quality demonstrations in robot learning. The primary design principles are summarized as follows:

- **Low-cost:** We aim to show that a capable system can be constructed at an affordable price, thus minimizing the entry barrier. This is achieved through the use of economical backdrivable servo motors, 3D printed components, and a minimalist design, making it possible to construct a teleoperation solution for under \$300. A cost comparison is shown in Table I. We will show that GELLO outperforms other low-cost options in our user study while being much cheaper than the other systems.
- **Capable:** GELLO is designed to be easy to use for human operators. We demonstrate GELLO’s capabilities on a range of complex bi-manual manipulation and contact-rich tasks.
- **Portable:** The diversity of demonstration data collected in different tasks and environments is critical to the final performance of the learning system. As such, we design GELLO to be compact, and self-contained, facilitating easy transportation. We show GELLO performing tasks across both lab and in-the-wild environments.
- **Simple to replicate:** The sourcing for parts is minimal beyond off-the-shelf motors and 3D printing components. The assembly process is also straightforward, requiring minimal technical expertise.

We follow these design principles to make GELLO for three types of robot arms, as shown in Figure 1. The instantiation of our approach centers around critical components like motor selection, kinematically equivalent structure, joint regularization, and 3D printed parts.

a) Servo selection: The critical component to enable GELLO’s construction is the availability of low-cost, fully-featured servos. Specifically, we used the DYNAMIXEL XL330 series [58]. Despite their affordability, these servos are equipped with high-resolution 12-bit encoders, enabling

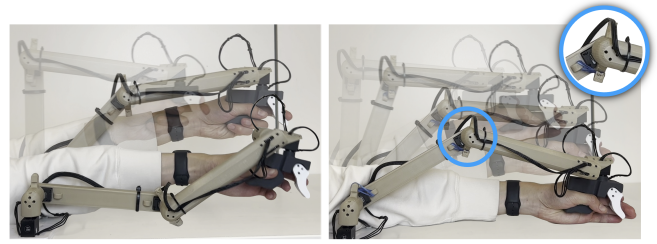


Fig. 2: This figure illustrates the trajectory of GELLO both with and without joint regularization. Left: Without joint regularization the elbow joint drops. This results in an unfavorable joint configuration for subsequent tasks or even collision with the table. Right: With joint regularization in place, the elbow exhibits minimal movement, leading to a more advantageous joint configuration. We find that simple rubber bands or springs are effective.

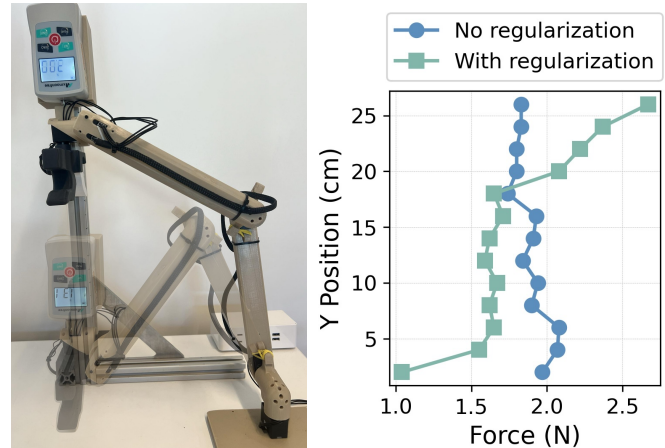


Fig. 3: A quantitative analysis of the effects of implementing simple joint regularization on the force necessary to maintain the teleoperation device in a stationary position at various heights. Without the joint regularization, the measured force required to resist gravity is approximately constant around 1.9N. This can make it more difficult for the operator to feel the kinematic limits of the robot. With regularization, we see that near the table, the restorative force helps reduce the required force to compensate for gravity while when far away, a higher restorative force is provided, encouraging the user to avoid singularities.

joint measurements within 0.088 of a mechanical degree. These encoders provide measurements of the servo’s position, allowing accurate mapping of the controller’s configuration to the target arm. In principle, a servo is not even necessary for the construction of GELLO, as we only need to read joint positions. However, in practice, a servo package provides an easy off-the-shelf, self-contained solution that has an encoder and communication protocol, simplifying construction, usage, and maintenance, furthering the goal of easy replication. In addition, the servo actuator provides physical resistance as the user backdrives it, which acts as natural damping and improves stability for the user. For this reason, we use the XL-330-288T, which provides the highest gear ratio offering the most resistance.

b) A scaled kinematically equivalent structure: We build GELLO as a small-scale version of the target arm which possesses a kinematically equivalent structure. This means that the joints and links of GELLO correspond directly to those of the target arm, allowing the user to control the GELLO manipulator as if they were directly controlling the target arm, as in kinesthetic teaching [59]. The kinematically equivalent structure is generated by taking the DH parameters

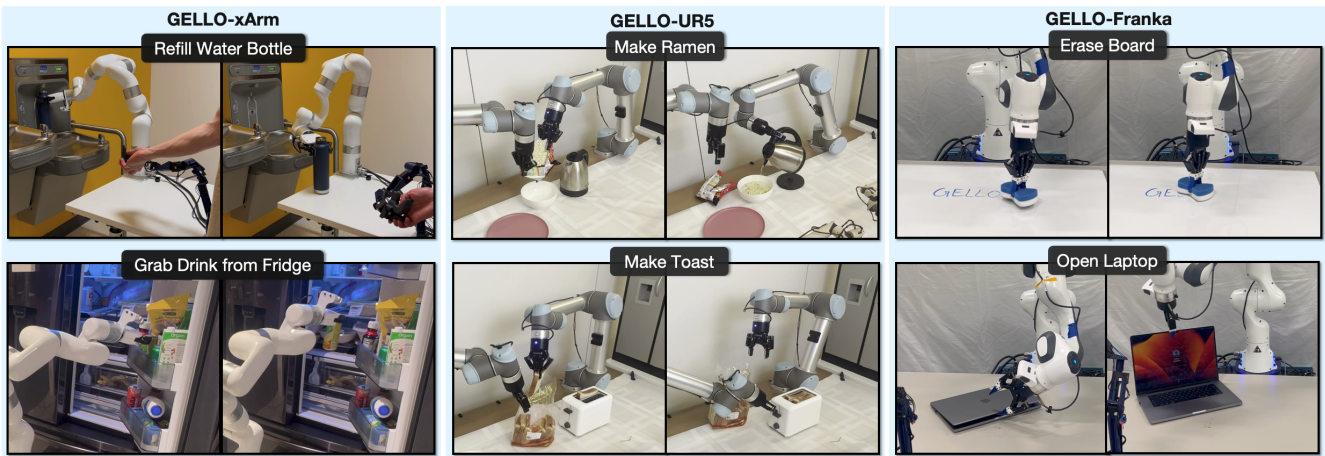


Fig. 4: We design and instantiate GELLO for 3 different robot arms. We experiment with GELLO across these three arms on a range of tasks to qualitatively observe the teleportation behavior. GELLO’s portable and self-sufficient design enables us to gather data across a wide range of environments.

of the target arm, and then scaling the lengths by a factor of α . Although the optimal α for comfort will be user and robot dependent, we used $\alpha = 0.5$ in our implementation and found it to be effective. Joint positions are read from the GELLO device and directly sent as a joint command to the target arm for operation, avoiding the need to compute inverse kinematics. The user can feel resistance from the controller when the joints are close to kinematic singularities or joint limits and is thus more aware of these failures, leading to more reliable teleoperation. At the same time, the miniature design makes the controller more portable while still allowing the user to operate full-scale robot arms.

c) Joint regularization: With only passive servo motors, the arm is constantly dragged by gravity to undesirable configurations during operation. We find that by adding simple joint regularizers, we can counteract the force of gravity on the manipulator, making it easier for the user to control. We employ rudimentary but effective passive joint regularization that involves the use of mechanical components such as springs or rubber bands to ensure that the device maintains a “natural” posture. This prevents the arm from adopting other kinematically viable yet unconventional positions, as illustrated in Figure 2, which may result in collisions. We only add joint regularization elements to the joints that exhibit the most significant resistance against gravity in the arm’s default resting position, which for the UR design, is the second and third joint. We quantitatively study this in Figure 3. We also find that joint regularization provides passive force feedback to the user which differs near the extremities of the joint range. This can help inform the user about the current configuration of the arm.

d) 3D printed parts: The use of 3D printed parts in GELLO allows a high degree of customization, enabling users to design and print parts that match the specific robot hardware. 3D printing allows us to easily design GELLO systems for 3 kinematically different robots. 3D printing is also a cost-effective method of producing parts, further contributing to the low-cost nature of GELLO.

Following these simple design principles, we instantiate

and test GELLO for three commonly used robot arms, the Universal Robot UR5, uFactory xArm7, and Franka Panda. Example tasks that we can perform with GELLO on the different robots are illustrated in Figure 4.

The control setting of direct joint control results in a very simple software stack. Joint angles are read directly from the GELLO device using the DYNAMIXEL provided python API and are commanded to the follower robot. We use the various python APIs for each robot type to send commands to the follower robot. We use ZMQ [60] for message passing between processes, and provide a simple protocol for extending to new robot types.

IV. EXPERIMENTS

We conduct experiments to evaluate GELLO as a teleoperation system. Quantitatively, we compare the performance of GELLO with other common low-cost teleoperation systems used in the literature across 5 tasks exploring various aspects of manipulation. Qualitatively, we study GELLO through tests on robots from 3 different manufacturers across a variety of manipulation tasks in diverse settings.

A. User Study Procedure

We conducted a user study involving 12 participants, focusing on bi-manual robot teleoperation using two UR robots to assess the comparative effectiveness of GELLO, 3D mice, and VR controllers under controlled conditions. All participants were volunteers from our university and none of the participants had professional training in robot teleoperation prior to the study. An overview of our experimental study is shown in Figure 5, where details of the 5 tasks we study are provided. These tasks cover a wide range of different manipulation challenges such as control over large work spaces, deformable objects, and fine-grain manipulation. Control with a 3D mouse or VR controller requires additional tuning of gain parameters that affect how sensitive the controller is to human input. A larger gain results in a larger robot motion for a fixed human control input. We tune each by testing the device across the 5 tasks, ensuring control is sensitive enough to achieve the fine-grain

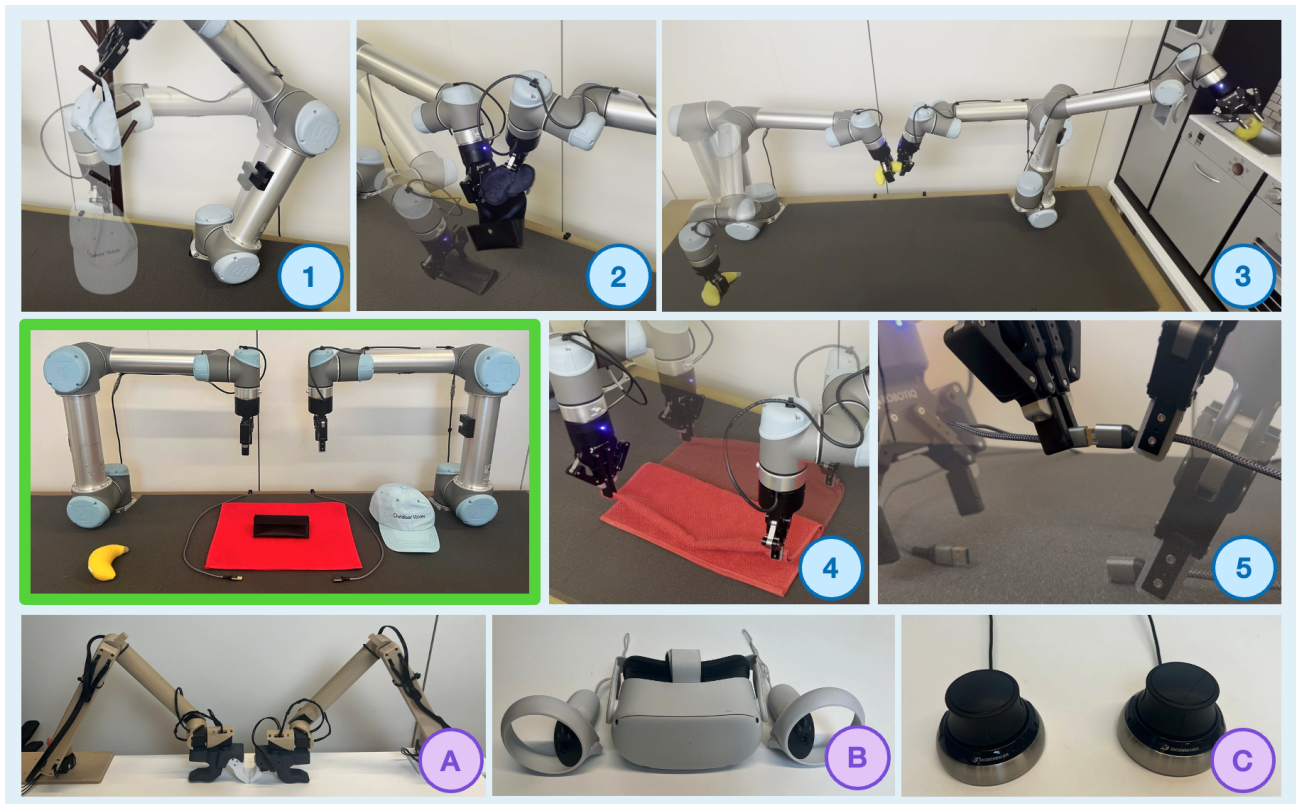


Fig. 5: An illustration of the experimental setup for our user study. The primary workspace of our experiments, indicated in green, showcases the scene of a bimanual robot station comprising of two UR5 robots with the complete task suit. The figure details the five tasks which are represented by the numerically labelled frames in blue: (1) Place a hat on a rack, (2) Open a case and fetch the sleeping mask inside, (3) Hand over a banana to the kitchen area, (4) Fold a towel, and (5) Plug in a USB cable. These tasks are designed to explore different teleoperation challenges such as articulated object interaction, large workspaces, deformable objects, and precise insertion. The three different bimanual teleoperation devices are shown by the letter-labelled images: (A) GELLO, (B) VR (Meta Quest2) controllers, and (C) 3D Mice (SpaceMouse). Each user attempts to accomplish the 5 tasks with all 3 teleoperation devices.

USB insertion tasks while still being responsive enough for quickly traversing across the workspace in the banana handoff task.

Before experimenting with the teleoperation tools, each user was granted a brief 6-minute orientation session, introducing them to the basics of the robot and teleoperation itself, as well as the task requirements. To reduce potential biases, no device-specific instructions were given in this orientation, and video demonstrations of the task do not show any particular device. This is then followed by the sequential introduction of all 3 different teleoperation devices. Upon introducing a new device, users were given a 5-minute practice phase, allowing them to gain familiarity with the device and its usage. During this time, users are allowed to experiment with the device and practice the tasks however they please. Next, the participants begin the task execution phase, with a time limit of 45 seconds for the hat task, and 90 seconds for the remaining tasks. This is repeated for all 3 teleoperation devices. *The order in which each participant learns to use the three devices is randomized.* This eliminates any potential bias from a fixed ordering. There are 6 possible orderings that the user can operate the devices. Each order is seen twice. Users were instructed to solve the task as best they could while avoiding self-collision (collision between the follower robot with itself or the other follower robot) or collision with

TABLE II: The task success rate for different teleoperation systems. GELLO achieves the top success rate across the board.

Device	Hat	Mask	Banana	Towel	USB	Avg
Gello	0.92	0.92	1.0	0.92	0.83	0.92
3D Mice	0.75	0.58	0.67	0.58	0.58	0.63
VR	0.92	0.83	0.75	0.58	0.5	0.72

the environment. The robot will be stopped when failure happens and the current task is terminated immediately. We record the task success and failure mode or task completion time as applicable. In the end, we have the user accomplish the same tasks without a robot using their hands. For privacy, no other user-specific data was collected.

B. User Study Results

In Table II, we show the success rate across distinct tasks for each teleoperation device. Using GELLO consistently results in the top success rate. For the simplest task, which only requires controlling a single arm, placing the hat on the rack, GELLO and VR perform comparably with respect to mean completion time. In more complex tasks, such as the banana handoff which requires both arms to maneuver across the large workspace, using spacemouse or VR leads to more failures like self-collisions or hitting singularities, suggesting that GELLO offers easier control in more broad use cases. For the task of towel folding, GELLO also shows a large

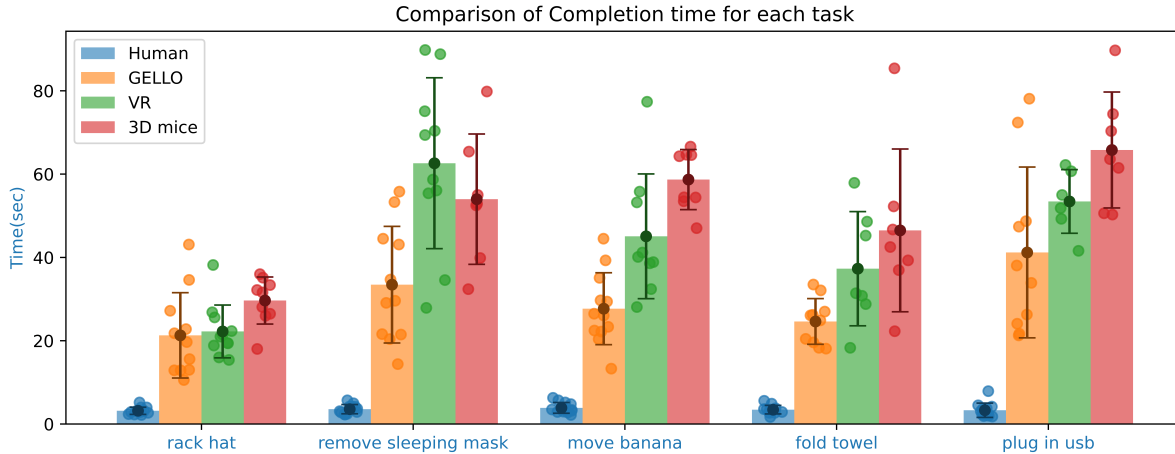


Fig. 6: Comparison with other teleoperation systems on the required duration for each task. For every combination of task and system, the average completion time (smaller is better) is plotted only for successful trials. Colored dots indicates the completion time for each user under each task-system pairing. Human, plotted in blue, gives a lower bound on teleoperation completion time, where the user directly accomplished each task with their hands.

TABLE III: The failure count for each failure mode aggregated across all 5 tasks. Each \times indicates a single failure of that type from our trials. The “Other” category captures all other irrecoverable task failure modes such as dropping the working item outside the reach of the robots.

Failure Mode	GELLO	3D Mice	VR
Timeout	\times	$\times\times\times\times\times\times\times\times\times$	$\times\times\times\times\times$
Self Collision	\times	$\times\times\times\times\times$	$\times\times\times\times\times$
Env Collision	$\times\times\times$	$\times\times\times\times\times\times$	$\times\times\times\times\times\times$
Other		$\times\times$	\times

advantage over the other two devices. We hypothesize that this is because GELLO offers more intuitive control and thus better coordination between the two arms. Interestingly, for the task of plugging in the USB cable, using GELLO also resulted in a much higher success rate. This is despite VR or spacemouse having the advantage of simplifying the problem by keeping one controller static while focusing their attention on the other, a common strategy users employed.

In Table III, we present a breakdown on the different failure modes for each teleoperation system. Our observations suggest that, unlike other devices that operate in Cartesian space, GELLO requires minimal user expertise. This is further corroborated by the lowest timeout count for GELLO in comparison to other teleoperation systems. Furthermore, GELLO’s isomorphic joint structure design ensures that teleoperation with GELLO has the least collision risk. Qualitatively we observed that self collisions were a common problem for the spacemouse and VR controllers due to the teleoperator paying attention only to the end effector and its relative motion. These devices, unlike GELLO, are unobservant of kinematic and robot-to-robot constraints which could explain the high amount of self-collisions.

Figure 6 provides a summary of the completion times taken by each device across all five tasks, given successful task execution. Utilizing GELLO results in consistently faster completion times. This not only signifies that GELLO is easier to use with a higher success rate but also indicates its efficiency; faster completion times would enable users to achieve more successful operations in a given time frame.

C. Teleoperation System Capabilities

We further demonstrate the capabilities of GELLO on more challenging manipulation tasks, including in real-world environments across the three different robot platforms. We show some of them in Figure 4 and put more videos on our project website. These tasks include contact-rich tasks, long-horizon tasks, and challenging bi-manual coordination tasks. Tasks like filling water bottles require a certain payload on the robot arm and would be difficult for smaller arms to perform, such as the ViperX arm used in the ALOHA system [9]. GELLO is also effective for 7 DOF arms, such as the Panda and xArms, despite the extra degree of freedom. We find that the kinematically equivalent structure enables a user to directly manage the arm’s null space if required, which can be advantageous when operating in cluttered spaces.

V. DISCUSSION

Due to limited output torque of the motors used, GELLO does not provide force feedback to the users, which limits GELLO’s capabilities when teleoperating for more contact-rich tasks. We made this compromise to keep GELLO low-cost, more accessible, and more applicable to all robot arms, as bilateral devices also require force sensing capability for the target robot. However, we hope to incorporate this as an optional capability in the future for more advanced usage.

Our user study is limited to inexperienced users who are only briefly taught about teleoperation and who only practiced for a limited time. Additional training can significantly improve the user’s proficiency in using teleoperation devices and we leave such study to future work.

In this paper, we introduced GELLO, a general low-cost teleoperation platform for manipulation. Our results demonstrate its effectiveness through a user study on teleoperation with a bi-manual robot system using two UR5s. To demonstrate versatility and make GELLO more accessible, we design GELLO for 3 robots. We hope GELLO will lower the barrier to collecting large and high-quality demonstration datasets, and thus accelerate progress in robot learning.

ACKNOWLEDGMENTS

Philipp Wu was supported in part by the NSF Graduate Research Fellowship Program. We thank Rachel Thomasson for providing valuable insights on teleoperation and feedback on early drafts of the manuscript. We thank Youngwoon Lee, Kevin Zakka, and Laura Smith for valuable discussions regarding GELLO. We thank Hannah Chong for feedback on the presentation of the manuscript. We thank Tingfan Wu and Charles Xu, for help with testing on the Franka robot.

REFERENCES

- [1] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in Cognitive Sciences*, vol. 3, no. 6, pp. 233–242, 1999.
- [2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, no. 5, pp. 469–483, 2009.
- [3] B. Siciliano and O. Khatib, Eds., *Springer Handbook of Robotics*. Berlin, Heidelberg: Springer, 2008.
- [4] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 5628–5635.
- [5] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, J. Dabis, C. Finn, K. Gopalakrishnan, K. Hausman, A. Herzog, J. Hsu, J. Ibarz, B. Ichter, A. Irpan, T. Jackson, S. Jesmonth, N. J. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, I. Leal, K.-H. Lee, S. Levine, Y. Lu, U. Malla, D. Manjunath, I. Mordatch, O. Nachum, C. Parada, J. Peralta, E. Perez, K. Pertsch, J. Quiambao, K. Rao, M. Ryoo, G. Salazar, P. Sanketi, K. Sayed, J. Singh, S. Sontakke, A. Stone, C. Tan, H. Tran, V. Vanhoucke, S. Vega, Q. Vuong, F. Xia, T. Xiao, P. Xu, S. Xu, T. Yu, and B. Zitkovich, "Rt-1: Robotics transformer for real-world control at scale," 2022.
- [6] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, X. Chen, K. Choromanski, T. Ding, D. Driess, A. Dubey, C. Finn, P. Florence, C. Fu, M. G. Arenas, K. Gopalakrishnan, K. Han, K. Hausman, A. Herzog, J. Hsu, B. Ichter, A. Irpan, N. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, I. Leal, L. Lee, T.-W. E. Lee, S. Levine, Y. Lu, H. Michalewski, I. Mordatch, K. Pertsch, K. Rao, K. Reymann, M. Ryoo, G. Salazar, P. Sanketi, P. Sermanet, J. Singh, A. Singh, R. Soricutt, H. Tran, V. Vanhoucke, Q. Vuong, A. Wahid, S. Welker, P. Wohlhart, J. Wu, F. Xia, T. Xiao, P. Xu, S. Xu, T. Yu, and B. Zitkovich, "Rt-2: Vision-language-action models transfer web knowledge to robotic control," 2023.
- [7] E. Jang, A. Irpan, M. Khansari, D. Kappler, F. Ebert, C. Lynch, S. Levine, and C. Finn, "BC-z: Zero-shot task generalization with robotic imitation learning," in *5th Annual Conference on Robot Learning*, 2021.
- [8] S. Reed, K. Zolna, E. Parisotto, S. G. Colmenarejo, A. Novikov, G. Barth-Maron, M. Gimenez, Y. Sulsky, J. Kay, J. T. Springenberg, T. Eccles, J. Bruce, A. Razavi, A. Edwards, N. Heess, Y. Chen, R. Hadsell, O. Vinyals, M. Bordbar, and N. de Freitas, "A generalist agent," 2022.
- [9] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning fine-grained bimanual manipulation with low-cost hardware," 2023.
- [10] J. Guo, C. Liu, and P. Poignet, "A scaled bilateral teleoperation system for robotic-assisted surgery with time delay," *Journal of Intelligent Robotic Systems*, vol. 95, 07 2019.
- [11] F. Pugin, P. Bucher, and P. Morel, "History of robotic surgery : From aesop® and zeus® to da vinci®," *Journal of Visceral Surgery*, vol. 148, no. 5, Supplement, pp. e3–e8, 2011, robotic surgery.
- [12] J. Liang, J. Mahler, M. Laskey, P. Li, and K. Goldberg, "Using dvrc teleoperation to facilitate deep learning of automation tasks for an industrial robot," in *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*, 2017, pp. 1–8.
- [13] G. Brantner and O. Khatib, "Controlling ocean one: Human–robot collaboration for deep-sea manipulation," *Journal of Field Robotics*, vol. 38, no. 1, pp. 28–51, 2021.
- [14] L. Barbieri, F. Bruno, A. Gallo, M. Muzzupappa, and M. L. Russo, "Design, prototyping and testing of a modular small-sized underwater robotic arm controlled through a master-slave approach," *Ocean Engineering*, vol. 158, pp. 253–262, 2018.
- [15] A. Wang, J. Ramos, J. Mayo, W. Ubellacker, J. Cheung, and S. Kim, "The hermes humanoid system: A platform for full-body teleoperation with balance feedback," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, 2015, pp. 730–737.
- [16] K. D. Katyal, C. Y. Brown, S. A. Hechtman, M. P. Para, T. G. McGee, K. C. Wolfe, R. J. Murphy, M. D. Kutzer, E. W. Tunstel, M. P. McLoughlin, *et al.*, "Approaches to robotic teleoperation in a disaster scenario: From supervised autonomy to direct control," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 1874–1881.
- [17] D. Connexion. (2023) Spacemouse. [Online]. Available: <https://3dconnexion.com/us/spacemouse/>
- [18] Meta. (2023) Meta quest. [Online]. Available: <https://www.meta.com/quest>
- [19] Vive. (2023) Vive vr products. [Online]. Available: <https://www.vive.com/us/product/>
- [20] H. Liu, S. Nasiriany, L. Zhang, Z. Bao, and Y. Zhu, "Robot learning on the job: Human-in-the-loop autonomy and learning during deployment," *RSS*, 2023.
- [21] X. Lin, J. So, S. Mahalingam, F. Liu, and P. Abbeel, "Spawnet: Learning generalizable visuomotor skills from pre-trained networks," 2023.
- [22] Y. Zhu, A. Joshi, P. Stone, and Y. Zhu, "Viola: Imitation learning for vision-based manipulation with object proposal priors," *CoRL*, 2022.
- [23] D. Rakita, B. Mutlu, and M. Gleicher, "A motion retargeting method for effective mimicry-based teleoperation of robot arms," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 361–370.
- [24] M. Shridhar, L. Manuelli, and D. Fox, "Perceiver-actor: A multi-task transformer for robotic manipulation," in *Conference on Robot Learning*. PMLR, 2023, pp. 785–799.
- [25] S. P. Arunachalam, I. Güzey, S. Chintala, and L. Pinto, "Holo-dex: Teaching dexterity with immersive mixed reality," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 5962–5969.
- [26] 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 *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 9164–9170.
- [27] A. Sivakumar, K. Shaw, and D. Pathak, "Robotic telekinesis: Learning a robotic hand imitator by watching humans on youtube," *arXiv preprint arXiv:2202.10448*, 2022.
- [28] 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," *RSS*, 2023.
- [29] S. Song, A. Zeng, J. Lee, and T. Funkhouser, "Grasping in the wild: Learning 6dof closed-loop grasping from low-cost demonstrations," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4978–4985, 2020.
- [30] M. Laghi, M. Maimeri, M. Marchand, C. Leparoux, M. Catalano, A. Ajoudani, and A. Bicchi, "Shared-autonomy control for intuitive bimanual tele-manipulation," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–9.
- [31] S.-K. Kim, S. Hong, and D. Kim, "A walking motion imitation framework of a humanoid robot by human walking recognition from imu motion data," 01 2010, pp. 343 – 348.
- [32] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3593–3600, 2019.
- [33] T. Hulin, K. Hertkorn, P. Kremer, S. Schätzle, J. Artigas, M. Sagardia, F. Zacharias, and C. Preusche, "The dlr bimanual haptic device with optimized workspace," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 3441–3442.
- [34] B. Katz, "A low cost modular actuator for dynamic robots," Ph.D. dissertation, 01 2018.
- [35] M. Schwarz, C. Lenz, A. Rochow, M. Schreiber, and S. Behnke, "Nimbro avatar: Interactive immersive telepresence with force-feedback telemanipulation," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 5312–5319.
- [36] D. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 624–637, 1993.

- [37] B. Hannaford, "A design framework for teleoperators with kinesthetic feedback," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 4, pp. 426–434, 1989.
- [38] J. Elsner, G. Reinert, L. Figueredo, A. Naceri, U. Walter, and S. Haddadin, "Parti-a haptic virtual reality control station for model-mediated robotic applications," *Frontiers in Virtual Reality*, vol. 3, p. 925794, 2022.
- [39] C. Lenz and S. Behnke, "Bimanual telemanipulation with force and haptic feedback and predictive limit avoidance," *CoRR*, vol. abs/2109.13382, 2021.
- [40] E. Ackerman. (2014) Beautiful fluid actuators from disney research make soft, safe robot arms.
- [41] I. Jo, Y. Park, and J. Bae, "A teleoperation system with an exoskeleton interface," in *2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. IEEE, 2013, pp. 1649–1654.
- [42] A. Toedtheide, X. Chen, H. Sadeghian, A. Naceri, and S. Haddadin, "A force-sensitive exoskeleton for teleoperation: An application in elderly care robotics," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 12 624–12 630.
- [43] Y. Ishiguro, T. Makabe, Y. Nagamatsu, Y. Kojio, K. Kojima, F. Sugai, Y. Kakiuchi, K. Okada, and M. Inaba, "Bilateral humanoid teleoperation system using whole-body exoskeleton cockpit tablis," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6419–6426, 2020.
- [44] D. H. Lab. (2023) "phantom omni" - 6 dof master device. [Online]. Available: <https://delfthapticslab.nl/device/phantom-omni/>
- [45] F. Dimension. (2023) Haptic devices. [Online]. Available: <https://www.forcedimension.com/products>
- [46] G. Zhu, X. Xiao, C. Li, J. Ma, G. Ponraj, A. V. Prituja, and H. Ren, "A bimanual robotic teleoperation architecture with anthropomorphic hybrid grippers for unstructured manipulation tasks," *Applied Sciences*, vol. 10, no. 6, 2020.
- [47] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 763–768.
- [48] K. Bousmalis, G. Vezzani, D. Rao, C. Devin, A. X. Lee, M. Bauza, T. Davchev, Y. Zhou, A. Gupta, A. Raju, A. Laurens, C. Fantacci, V. Dalibard, M. Zambelli, M. Martins, R. Pevceviciute, M. Blokzijl, M. Denil, N. Batchelor, T. Lampe, E. Parisotto, K. Žořna, S. Reed, S. G. Colmenarejo, J. Scholz, A. Abdolmaleki, O. Groth, J.-B. Regli, O. Sushkov, T. Rothörl, J. E. Chen, Y. Aytar, D. Barker, J. Ortiz, M. Riedmiller, J. T. Springenberg, R. Hadsell, F. Nori, and N. Heess, "Robocat: A self-improving foundation agent for robotic manipulation," 2023.
- [49] H. Walke, K. Black, A. Lee, M. J. Kim, M. Du, C. Zheng, T. Zhao, P. Hansen-Estruch, Q. Vuong, A. He, *et al.*, "Bridgedata v2: A dataset for robot learning at scale," *arXiv preprint arXiv:2308.12952*, 2023.
- [50] H.-S. Fang, H. Fang, Z. Tang, J. Liu, J. Wang, H. Zhu, and C. Lu, "Rh20t: A robotic dataset for learning diverse skills in one-shot," 2023.
- [51] S. Dass, K. Pertsch, H. Zhang, Y. Lee, J. J. Lim, and S. Nikolaidis, "Pato: Policy assisted teleoperation for scalable robot data collection," *RSS*, 2023.
- [52] C. Finn, T. Yu, T. Zhang, P. Abbeel, and S. Levine, "One-shot visual imitation learning via meta-learning," in *CoRL*, 2017.
- [53] L. Smith, N. Dhawan, M. Zhang, P. Abbeel, and S. Levine, "Avid: Learning multi-stage tasks via pixel-level translation of human videos," *RSS*, 2019.
- [54] S. Bahl, A. Gupta, and D. Pathak, "Human-to-robot imitation in the wild," *RSS*, 2022.
- [55] Y. Qin, Y.-H. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, and X. Wang, "Dexmv: Imitation learning for dexterous manipulation from human videos," in *European Conference on Computer Vision*. Springer, 2022, pp. 570–587.
- [56] C. Wang, L. Fan, J. Sun, R. Zhang, L. Fei-Fei, D. Xu, Y. Zhu, and A. Anandkumar, "Mimicplay: Long-horizon imitation learning by watching human play," *arXiv preprint arXiv:2302.12422*, 2023.
- [57] C. Wen, X. Lin, J. So, K. Chen, Q. Dou, Y. Gao, and P. Abbeel, "Any-point trajectory modeling for policy learning," *arXiv preprint arXiv:2401.00025*, 2023.
- [58] I. ROBOTIS. (2023) Dynamixel xl. [Online]. Available: <https://www.robotis.us/xl/>
- [59] A. G. Billard, S. Calinon, and F. Guenter, "Discriminative and adaptive imitation in uni-manual and bi-manual tasks," *Robotics and Autonomous Systems*, vol. 54, no. 5, pp. 370–384, 2006, the Social Mechanisms of Robot Programming from Demonstration.
- [60] ZeroMQ, "Zeromq," <https://zeromq.org/>, 2024.