MEVION: Low-Cost Open-Source Data Collection System for Powerful and High-Speed Dual-Arm Manipulation

Kento Kawaharazuka[1], Yoshiki Obinata[1], Hirokazu Ishida[1], Jihoon Oh[1], Temma Suzuki[1], Shintaro Inoue[1], Keita Yoneda[1], Ayumu Iwata[1], and Kei Okada[1]

***Abstract*— The global competition for developing robotic foundation models is intensifying. Among the data collection systems used for dual-arm robots, ALOHA is representative of being low-cost and open-source, and is widely adopted by researchers as a de facto standard. However, due to its limited ability to generate high forces and speeds, it is difficult to handle heavy objects or perform fast manipulations. To address this, we developed MEVION, a low-cost and open-source dual-arm robot data collection system capable of generating greater force and speed. All parts of this robot can be sourced through e-commerce, and by extensively utilizing sheet metal welding, its large body structure is constructed with a small number of components at low cost, while also simplifying assembly. MEVION is equipped with four 6-DoF arms with parallel grippers. Each arm weighs 7.0 kg and has a maximum torque of 60 Nm, and the entire system can be constructed for about USD 14,000. The elbow joint adopts a closed-link mechanism similar to those used in quadruped robots, which reduces the distal mass and enables higher force and speed output at the end-effector. We demonstrate that MEVION enables data collection for object manipulation tasks not previously possible and supports imitation learning-based motion generation. All hardware and software of this work are included in the Supplementary Materials or github.com/haraduka/mevion.**

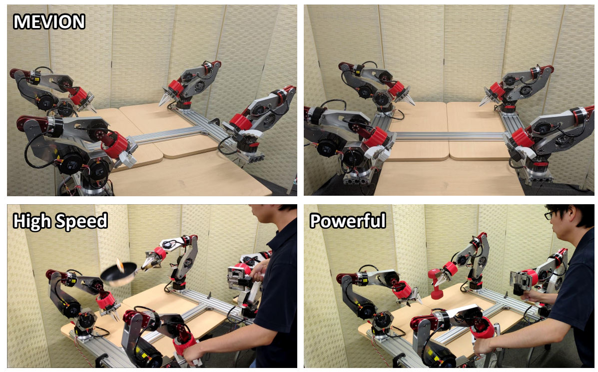

Fig. 1. MEVION: Low-cost open-source data collection system for powerful and high-speed dual-arm manipulation, developed in this study.

I. INTRODUCTION

Competition in the development of robotic foundation models is intensifying, with research institutes and companies worldwide vying for progress. In particular, the advancement of imitation learning and its extension to Vision-Language-Action (VLA) models [1]–[3] has spurred active research on manipulation using dual-arm robots. Various types of dual-arm robots and corresponding data collection systems have been proposed for this purpose.

Broadly speaking, two main teaching methods exist. The first involves operating robots via VR devices or motion capture systems [4]–[6], while the second adopts a leader-follower scheme in which humans directly move the robot's arms [7]–[13]. Both approaches have their advantages and drawbacks, but the latter is generally considered more intuitive and has therefore been widely adopted by researchers. In terms of robot morphology, three categories can be identified: tabletop-type [7]–[9], [11], [12], mobile-type [10], [14], and standing-type [5], [6], [13]. The tabletop-types are placed on a table; mobile-types have a wheeled base; and standing-types resemble a humanoid upper body. While tabletop-types and mobile-types typically have arms extending upward from below, standing-types have arms extending downward from the shoulders. Another important distinction is whether a system is open-source or closed-source. Commercially released systems are generally closed-source, making it difficult to modify link lengths or change the setup. In contrast, open-source systems offer the advantage of customizability. Considering the increasing complexity of future manipulation tasks, the ability to freely reconfigure systems through open-source design is a crucial factor.

Within this landscape, this study focuses on tabletop-type dual-arm robot data collection systems employing the leader-follower method. Among these, ALOHA [7] is representative of being both low-cost and open-source, and has become a de facto standard widely used by researchers. However, since it relies on small servo motors, it cannot generate high forces or speeds, making it difficult to handle heavy objects or perform rapid manipulations. As a result, studies using ALOHA and similar systems have primarily focused on tasks involving slow manipulation of lightweight objects, posing challenges for collecting more diverse manipulation data.

To address these limitations, we developed MEVION, a low-cost and open-source dual-arm robot data collection system capable of generating greater force and speed. All components of this robot can be sourced through e-commerce, and by leveraging sheet metal welding, a large body structure can be constructed with a small number of parts at low cost. MEVION is equipped with four 6-DoF arms with parallel grippers. Each arm weighs 7.0 kg and provides a maximum torque of 60 Nm, and the entire system can be built for approximately USD 14,000. The elbow joint adopts a closed-link mechanism similar to those used in quadruped robots, reducing the distal mass and enabling higher speed and

[1] The authors are with the Department of Mechano-Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan. [kawaharazuka, obinata, h-ishida, oh, t-suzuki, s-inoue, yoneda, a-iwata, k-okada]@jsk.imi.i.u-tokyo.ac.jp

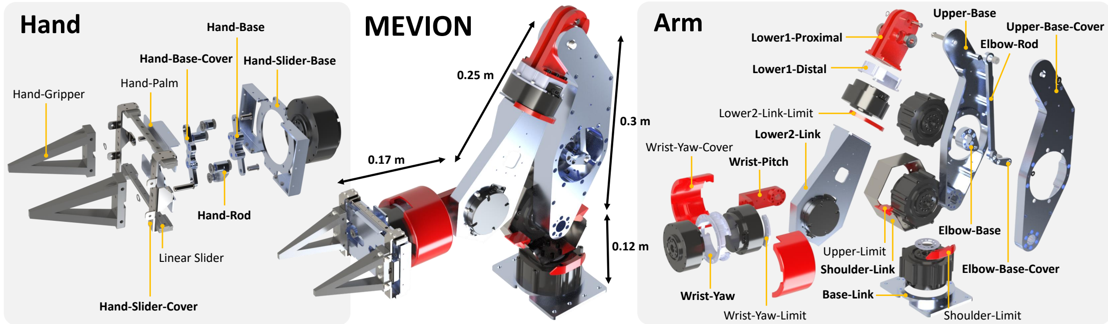

Fig. 2. Design overview of MEVION: The design is divided into the arm and hand, comprising 17 essential metal components in total — 11 in the arm and 6 in the hand, three of which are fabricated through sheet metal welding to form single integrated parts with complex and large geometries.

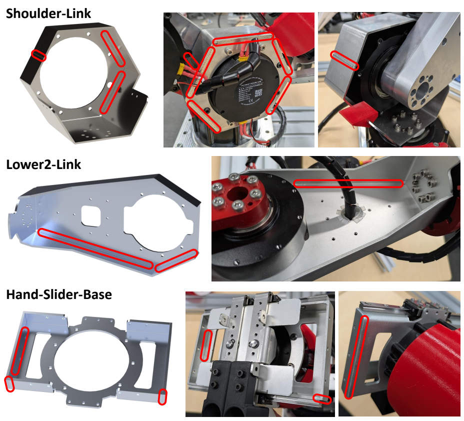

Fig. 3. Details of sheet metal welding for the Shoulder-Link, Lower2-Link, and Hand-Slider-Base. The welded sections are highlighted in red.

force output at the end-effector. With MEVION, it becomes possible to manipulate heavy objects and perform high-speed operations, enabling the collection of more diverse manipulation data. Furthermore, we demonstrated a complete imitation learning pipeline with MEVION, confirming its capability for autonomous motion generation. All hardware and software developed in this study are included in the Supplementary Materials or github.com/haraduka/mevion.

II. Hardware and Software of MEVION

A. Design Overview of MEVION

The design overview of MEVION is shown in Fig. 2. The design is primarily divided into the arm and hand components, with the essential metallic structural parts comprising 11 components for the arm and 6 for the hand, totaling 17 components. When including joint angle limiters and cover parts, the total number of metallic components amounts to 21. For materials, all machined metal components are made of A7075 aluminum alloy. Among the sheet metal components, only the Shoulder-Link and Hand-Slider-Cover are made of SUS304 stainless steel, while the rest use A5052 aluminum alloy. In addition, some joint angle limiters and cable covers are fabricated in flexible TPU, while the fingertip structures of the leader and follower hands are made of PLA, both manufactured using 3D printing.

The arm's joint configuration is identical to that of ALOHA, with six degrees of freedom consisting of Shoulder-Yaw, Shoulder-Pitch, Elbow-Pitch, Elbow-Yaw, Wrist-Pitch, and Wrist-Yaw. In general, each joint is driven by a single motor; however, for the elbow joint, a parallel-link mechanism, similar to those used in quadruped robots [15], [16], is employed to drive the joint from a proximal motor, thereby reducing the distal mass. The hand adopts a slider-crank parallel gripper mechanism, where a single motor simultaneously drives two fingers. The motors used are: RobStride03 for Shoulder-Yaw, Shoulder-Pitch, and Elbow-Pitch; RobStride02 for Elbow-Yaw, Wrist-Pitch, and Wrist-Yaw; and RobStride01 for the hand. Their respective maximum torques are 60 Nm, 17 Nm, and 17 Nm.

A notable feature of MEVION is its small number of parts, achieved through extensive use of sheet metal welding, as shown in Fig. 3. The Shoulder-Link and Lower2-Link of the arm, as well as the Hand-Slider-Base of the hand, integrate large and complex shapes into single units via sheet metal welding, while the remaining parts are manufactured through machining. By leveraging sheet metal welding that provides sufficient structural strength even for large-scale shapes, a large body structure is realized with a small number of components, achieving both low cost and ease of assembly. All of these metallic components are designed to be fully compatible with meviy [17], MISUMI's online machining and fabrication service, which automatically generates quotations and orders from 3D CAD files. As a result, the entire ordering process can be completed through e-commerce, enabling MEVION to be easily reproduced.

B. Comparison with Existing Robot Systems

The comparison between MEVION and existing open-source leader-follower dual-arm robot data collection systems is shown in Table I. Eight aspects were considered: morphology, number of DOFs, weight, length, materials, maximum torque, maximum speed, and price, comparing representative systems SO-101 [12], ALOHA [7], OpenArm [13], and MEVION. Among them, SO-101, ALOHA, and MEVION are tabletop-type systems, whereas OpenArm is standing-type.

Comparing the three tabletop-type systems, weight, length, and maximum torque increase stepwise in the order SO-101 < ALOHA < MEVION. In other words, these three robots clearly differ in terms of overall size and output capability. Regarding materials, SO-101's structural components are all plastic, ALOHA uses metal only in the arms, and MEVION

TABLE I

COMPARISON BETWEEN EXISTING OPEN-SOURCE LEADER-FOLLOWER BIMANUAL DATA COLLECTION SYSTEMS AND MEVION

Name	Morphology	DOFs	Weight	Length*1	Materials		Maximum Torque*2	Maximum Speed*2	Price*3
					Arm	Hand			
SO-101 [12]	Tabletop	4+1	0.7 kg	0.45 m	Plastic	Plastic	1.9 Nm	5.5 rad/s	$120
ALOHA [7]	Tabletop	6+1	4.3 kg	0.75 m	**Metal**	Plastic	21.2 Nm	3.1 rad/s	$4,900
OpenArm [13]	Standing	7+1	5.5 kg	0.63 m	**Metal**	**Metal**	40 Nm	16.8 rad/s	$3,200
MEVION (this study)	Tabletop	6+1	7.0 kg	0.83 m	**Metal**	**Metal**	60 Nm	20.4 rad/s	$3,500

*1The distance from the center of the first joint to the end-effector; *2The peak torque and no-load joint speed of the highest-output motor; *3The approximate cost of a single arm.

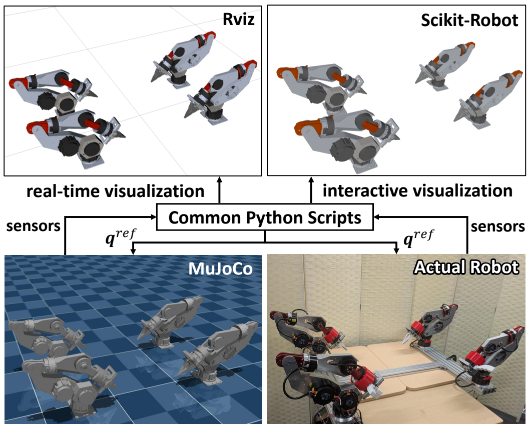

Fig. 4. Software of MEVION: MEVION is controlled by a unified Python script, allowing both simulation on MuJoCo and real-world control to be executed with the same code. It supports real-time visualization of the robot state in RViz via ROS, as well as interactive visualization prior to sending control commands using Scikit-Robot [18].

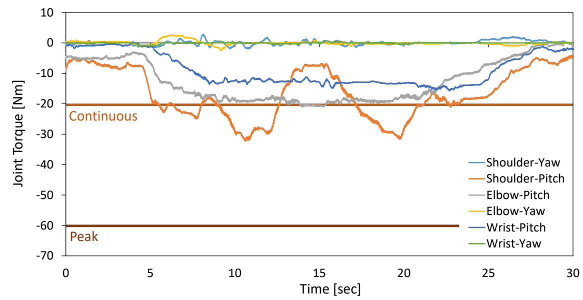

Fig. 5. The transition of joint torques during the dumbbell manipulation.

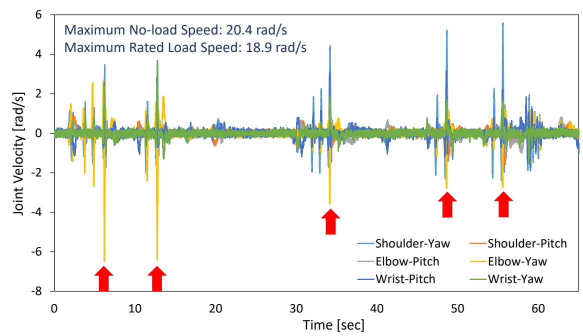

Fig. 6. The transition of joint velocities during the Daruma Otoshi.

uses all-metal components, reflecting a progressive increase in structural robustness. Notably, compared with ALOHA, MEVION is about 1.6 times heavier, yet achieves approximately three times the maximum torque and seven times the maximum speed, enabling significantly greater force and speed. Unlike SO-101 and ALOHA, MEVION places the elbow motor proximally and employs a parallel link mechanism, further raising the upper limits of torque and speed. Furthermore, MEVION remains low-cost, at around USD 3,500 per arm, while offering high performance. All of these systems are open-source and thus freely reconfigurable, and both SO-101 and ALOHA also have wheeled mobile-base variants [10].

Next, we compare MEVION with OpenArm, which has relatively similar specifications. While MEVION is tabletop-type and OpenArm is standing-type, leading to substantial differences in joint sequence and count that preclude direct comparison, their weight, length, maximum torque, maximum speed, and price are all very close. The major differences lie in end-effector weight reduction and the number of parts. As described earlier, MEVION reduces end-effector weight and required gravity-compensation torque by arranging the elbow motor proximally. The peak payload at maximum extension is 7.6 kg for MEVION and 4.5 kg for OpenArm, indicating that MEVION can handle heavier loads. In addition, OpenArm consists of 31 basic metallic components, whereas MEVION uses only 17 (21 including joint limiters and covers), representing a substantial reduction in the number of parts. This is enabled by sheet metal welding, which allows integration of large and complex shapes while maintaining strength.

C. Software of MEVION

The software architecture of MEVION is shown in Fig. 4. MEVION is primarily controlled through a unified Python script, enabling both MuJoCo-based simulation and real-world control to be executed with the same code. It supports real-time visualization of the robot's state in RViz via ROS, as well as interactive rendering before sending control commands using Scikit-Robot [18]. Unless otherwise specified, Intel RealSense D405 cameras are employed as both fixed and wrist-mounted cameras.

MEVION operates on either 24V or 48V power and adopts CAN communication. Four CAN-USB interfaces are connected to the individual arms, which are then linked to a PC for control. The basic control scheme is a PD controller that computes current commands based on joint angle commands, with gravity compensation torque added.

III. EXPERIMENTS

A. Teleoperation Experiments

The teleoperation experiment results are shown in Fig. 7. Five tasks were performed: (a) bottle cap opening, (b) object packing, (c) frying pan operation, (d) 3.6 kg dumbbell manipulation, and (e) the traditional Japanese game Daruma Otoshi. These tasks range from (a) bimanual cooperative

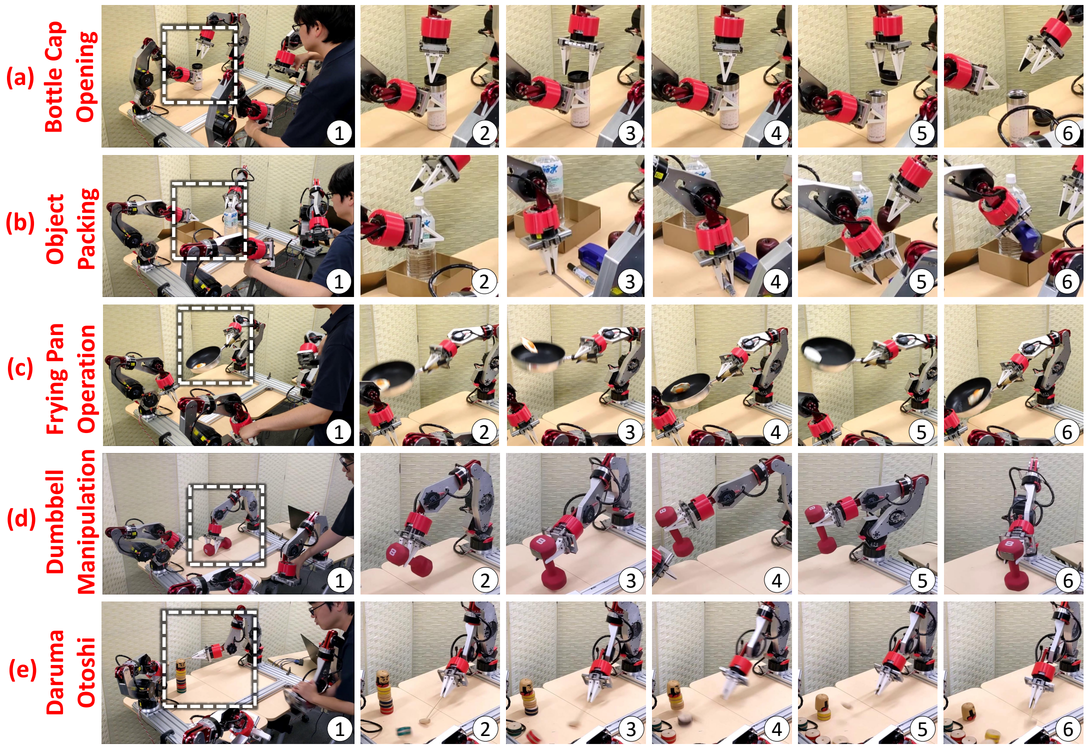

Fig. 7. Teleoperation experiments using MEVION: (a) bottle cap opening, (b) object packing, (c) frying pan operation, (d) 3.6 kg dumbbell manipulation, and (e) Daruma Otoshi.

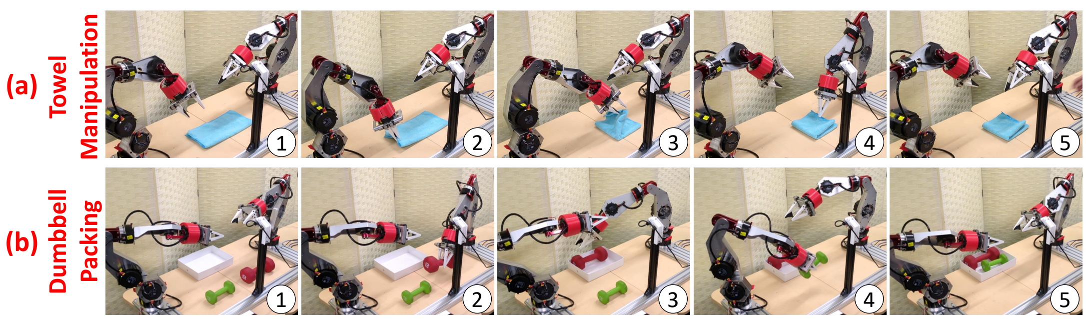

Fig. 8. The results of imitation learning using MEVION: (a) towel manipulation and (b) dumbbell packing.

actions, to (b) the transportation of slender and heavy objects including wrenches and 1.5 kg plastic bottles, and to (c) high-speed manipulations using a frying pan, confirming that MEVION can execute a wide variety of tasks. The joint torques during (d) dumbbell manipulation are presented in Fig. 5, and the joint velocities during (e) Daruma Otoshi are presented in Fig. 6. In the dumbbell task, all joints produced torque values with sufficient margin from their maximum limits. In the Daruma Otoshi task, the maximum joint velocity reached 6.5 rad/s, demonstrating that MEVION enables high-speed manipulation that would be difficult with systems such as ALOHA or SO-101.

B. Imitation Learning Experiments

The inference results of imitation learning are shown in Fig. 8. Two tasks were performed: (a) folding a towel and placing it in the center, and (b) packing dumbbells weighing 3.6 kg and 1.4 kg into a box. First, a human operator collected 15 demonstrations for each task via teleoperation. Using the collected data, imitation learning was conducted with the Action Chunking Transformer [7]. For both tasks, we confirmed that the robot successfully executed autonomous actions more than five times consecutively, even when the positions of the towel or dumbbells were slightly shifted.

IV. CONCLUSION

In this work, we developed MEVION, a low-cost and open-source tabletop dual-arm data collection platform that achieves greater force and speed than existing systems. All components of MEVION can be sourced via e-commerce, particularly through MISUMI and meviy, and its large body structure is realized with a small number of parts at low cost by leveraging sheet metal welding. With MEVION, it becomes possible to manipulate heavy objects and perform high-speed operations, enabling the collection of more diverse manipulation data. We expect that this open-source hardware and software will be widely adopted by researchers and contribute to advancing dual-arm manipulation as well as education in the field.

REFERENCES

[1] Y. Ma, Z. Song, Y. Zhuang, J. Hao, and I. King, "A Survey on Vision-Language-Action Models for Embodied AI," *arXiv preprint arXiv:2405.14093*, 2024.

[2] Y. Zhong, F. Bai, S. Cai, X. Huang, Z. Chen, X. Zhang, Y. Wang, S. Guo, T. Guan, K. N. Lui, *et al.*, "A Survey on Vision-Language-Action Models: An Action Tokenization Perspective," *arXiv preprint arXiv:2507.01925*, 2025.

[3] K. Kawaharazuka, J. Oh, J. Yamada, I. Posner, and Y. Zhu, "Vision-Language-Action Models for Robotics: A Review Towards Real-World Applications," *IEEE Access*, vol. 13, pp. 162 467–162 504, 2025.

[4] Y. Qin, W. Yang, B. Huang, K. V. Wyk, H. Su, X. Wang, Y. Chao, and D. Fox, "Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system," *arXiv preprint arXiv:2307.04577*, 2023.

[5] X. Cheng, J. Li, S. Yang, G. Yang, and X. Wang, "Open-television: Teleoperation with immersive active visual feedback," *arXiv preprint arXiv:2407.01512*, 2024.

[6] R. Ding, Y. Qin, J. Zhu, C. Jia, S. Yang, R. Yang, X. Qi, and X. Wang, "Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning," *arXiv preprint arXiv:2407.03162*, 2024.

[7] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware," in *2023 Robotics: Science and Systems*, 2023.

[8] P. Wu, Y. Shentu, Z. Yi, X. Lin, and P. Abbeel, "GELLO: A General, Low-cost, and Intuitive Teleoperation Framework for Robot Manipulators," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2024, pp. 12 156–12 163.

[9] J. Aldaco, T. Armstrong, R. Baruch, J. Bingham, S. Chan, K. Draper, D. Dwibedi, C. Finn, P. Florence, S. Goodrich, *et al.*, "Aloha 2: An enhanced low-cost hardware for bimanual teleoperation," *arXiv preprint arXiv:2405.02292*, 2024.

[10] Z. Fu, T. Z. Zhao, and C. Finn, "Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," *arXiv preprint arXiv:2401.02117*, 2024.

[11] M. Kobayashi, T. Buamanee, and T. Kobayashi, "Alpha-α and bi-act are all you need: Importance of position and force information/control for imitation learning of unimanual and bimanual robotic manipulation with low-cost system," *IEEE Access*, 2025.

[12] H. Face, "SO-101," https://huggingface.co/docs/lerobot/so101, accessed: 2026-07-05.

[13] Enactic, "OpenArm," https://github.com/enactic/openarm, accessed: 2026-07-05.

[14] H. Cui, Y. Yuan, Y. Zheng, and J. Hao, "AhaRobot: A Low-Cost Open-Source Bimanual Mobile Manipulator for Embodied AI," *arXiv preprint arXiv:2503.10070*, 2025.

[15] K. Kawaharazuka, S. Inoue, T. Suzuki, S. Yuzai, S. Sawaguchi, K. Okada, and M. Inaba, "MEVIUS: A Quadruped Robot Easily Constructed through E-Commerce with Sheet Metal Welding and Machining," in *2024 IEEE-RAS International Conference on Humanoid Robots*, 2024, pp. 631–636.

[16] K. Kawaharazuka, K. Yoneda, S. Inoue, T. Suzuki, J. Oda, and K. Okada, "MEVIUS2: Practical Open-Source Quadruped Robot with Sheet Metal Welding and Multimodal Perception," *IEEE Robotics and Automation Practice*, vol. 1, pp. 50–54, 2026.

[17] "meviy (MISUMI)," https://meviy.misumi-ec.com/worldwide/en/.

[18] "Scikit-Robot," https://github.com/iory/scikit-robot.

[19] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiraux, O. Stasse, and N. Mansard, "The Pinocchio C++ library – A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives," in *IEEE International Symposium on System Integrations*, 2019.

Description of Supplementary Materials

A. Overview

The Supplementary Materials for this study include the following files:

- **PartsList.xlsx**: A list of all components used in MEVION.
- **MEVION-STEP.tar.gz**: STEP files for each individual MEVION component.
- **MEVION-Assembly.STEP**: The full assembly file of MEVION.
- **Software.tar.gz**: The software package for MEVION.

B. Hardware

This section provides information related to the hardware, with a focus on key points regarding the parts list. Currently, the metallic parts of MEVION cost USD 1,465 for the arm and USD 421 for the hand, and including all motors and sensors, the total comes to USD 3,532. Since four sets are used, the entire system can be built for approximately USD 14,000. The parts list includes details such as part names, materials, quantities, and prices, but the most critical information is the part numbers used for ordering via meviy. All metal components used in this study can be ordered simply by specifying their part numbers. Of course, it is also possible to place orders by uploading the corresponding STEP files. All mechanical parts can be ordered through MISUMI and meviy, and for electronic components, individual URLs are provided in the list. Note that the parts list contains information for two types of hands: **hand** refers to a gripper driven by CyberGear, while **hand2** refers to a gripper driven by RobStride01. In addition, all components laid out are shown in Fig. 9. Assembly files for these components are provided, allowing users to assemble the system by referring to them. Owing to the small number of parts, the assembly process is straightforward, and once familiar with the procedure, it can be completed in one hour.

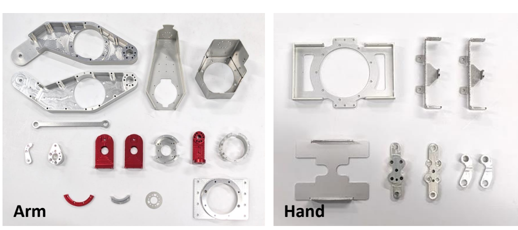

Fig. 9. All metal components used in MEVION.

C. Software

Here, we provide information on the software. All programs are written in Python and support both MuJoCo-based simulation and real-world control. In addition, visualization in RViz via ROS and interactive visualization using Scikit-Robot are also supported. Accordingly, the package includes XML model files for MuJoCo, URDF model files for RViz, and related resources.

MEVION, unlike ALOHA, does not employ hardware-based gravity compensation; instead, gravity compensation torques are computed and applied on the software side. The control input for each joint i is given by

$$\tau_i^{cmd} = K_{p,i}(q_i^{ref} - q_i) + K_{d,i}(\dot{q}_i^{ref} - \dot{q}_i) + \tau_i^{ff} \quad (1)$$

where the feedforward term τ^{ff} is computed using Pinocchio [19]. This enables smooth operation without dedicated gravity compensation mechanisms, as used in ALOHA. To ensure safe operation, limits are imposed on all of q_i^{ref}, $\dot{q}_i^{ref}$, and τ_i^{ff}, corresponding to joint angle, joint velocity, and joint torque, respectively. The control loop runs at 200 Hz, with basic unilateral control implemented.

While this approach eliminates the need for additional frames or mechanisms as used in ALOHA, it raises greater safety concerns. Therefore, safety mechanisms are incorporated both in hardware and software: a wireless emergency stop switch is implemented on the hardware side, while constraints such as maximum torque and velocity limits are enforced in software. In addition, sample programs for collision avoidance are provided.

D. Limitations and Future Challenges

Sheet metal welding. While sheet metal welding is highly effective for reducing the number of components and overall cost, it also presents several challenges. First, unlike metal machining, it is difficult to perform modifications or repairs independently. Although re-ordering parts is often the simplest solution due to the relatively low cost, it may be necessary to consider the trade-offs with metal machining in future work. Second, compared to metal machining, the quality of welding can vary depending on the fabrication vendor. In leader-follower systems, such variability often does not pose a significant issue, as the robot is directly operated by a human. However, when scaling to large-scale data collection, it may be necessary to account for the impact of such variability.

Large-scale data collection. This work ultimately aims to enable large-scale data collection. By performing gravity compensation in software, the physical footprint of the robot can be reduced, allowing a larger number of robots to be operated simultaneously. In practice, we are planning to construct multiple MEVION systems in collaboration with several researchers and utilize them for data collection (Fig. 10). On the other hand, scalability introduces challenges, particularly in terms of mass production. It may be necessary to revisit the design with a focus on simplifying the structure and improving assembly efficiency to better support large-scale deployment.

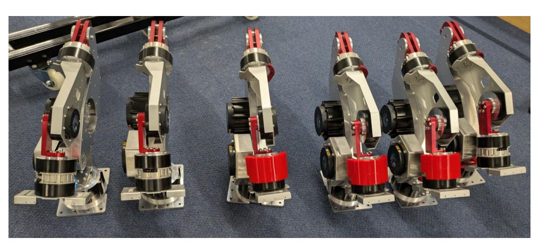

Fig. 10. Researchers are currently building multiple MEVIONs for large-scale data collection.