

Instrumentation for Better Demonstrations: A Case Study

Remko Proesmans¹, Thomas Lips¹ and Francis wyffels¹

Abstract—Learning from demonstrations is a powerful paradigm for robot manipulation, but its effectiveness hinges on both the quantity and quality of the collected data. In this work, we present a case study of how instrumentation, i.e. integration of sensors, can improve the quality of demonstrations and automate data collection. We instrument a squeeze bottle with a pressure sensor to learn a liquid dispensing task, enabling automated data collection via a PI controller. Transformer-based policies trained on automated demonstrations outperform those trained on human data in 78% of cases. Our findings indicate that instrumentation not only facilitates scalable data collection but also leads to better-performing policies, highlighting its potential in the pursuit of generalist robotic agents.

I. INTRODUCTION

Although robots can surpass humans for specific tasks in carefully constructed environments, human-level performance is still beyond the reach of general-purpose robots in unstructured environments [1]. One way to approach human-level manipulation performance with robots is to make them imitate human behaviour through demonstrations. This has proven to be effective in teaching complex manipulation skills to robots [2]–[5]. However, learning even a single task in a controlled environment can take 100 or more demonstrations [2], [3]. For a robot to handle multiple complex tasks in unstructured environments, the data requirements grow tremendously. Large datasets for robot pre-training exist [6]–[8], but current systems still require task-specific fine-tuning for complex tasks or diverse environments [4].

In addition to larger datasets and algorithmic advances, we can attempt to improve the quality of the demonstrations to increase imitation learning capabilities. Both [9] and [10] determined evaluation metrics for the quality of demonstration trajectories, and showed that better demonstrations can result in a better policy for the same amount of data. Xu et al. [11] have shown that reinforcement learning (RL) agents generate high-quality trajectories through reward maximisation, making them better suited for fine-tuning generalist policies compared to human demonstrations.

A different approach was followed in [12] for the task of raspberry harvesting: instead of teleoperating the robot and recording trajectories as demonstrations, human demonstrators manipulated an instrumented, i.e. sensorised, strawberry phantom. The robot behaviour was then tuned to match human behaviours as experienced by the strawberry. Instrumentation was also used in [13] for learning to solve a Rubik’s

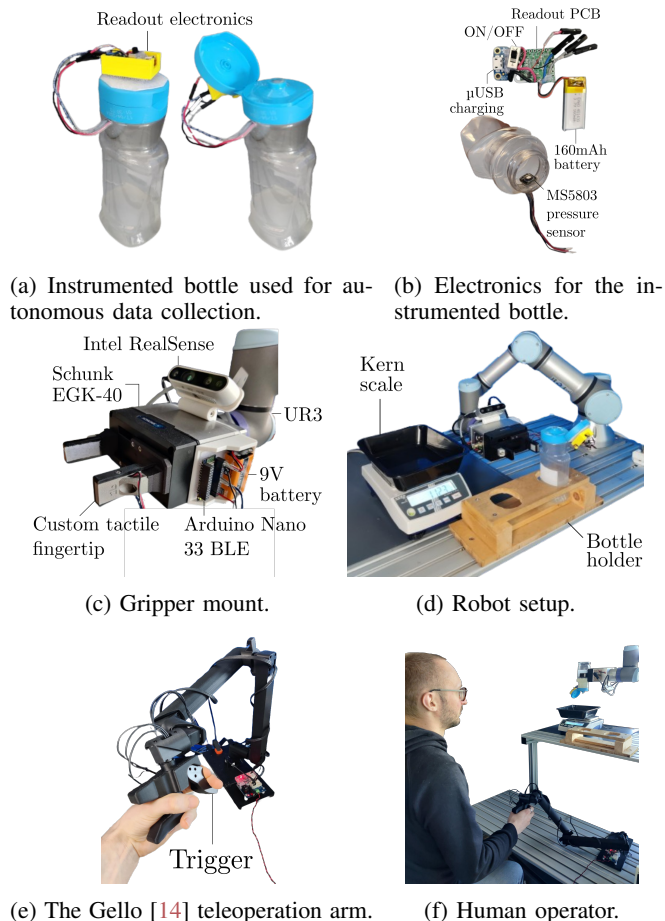


Fig. 1: Experimental setup for learning to squeeze a constant flow from a bottle based on demonstrations, collected either by a human or by an autonomous controller relying on instrumentation data.

cube with a single anthropomorphic robot hand. Specifically, the sensor data provided accurate state information of the Rubik’s cube to shape the reward function.

The idea of using additional state information during the learning phase, and later deploying without this privileged information, using only sensory inputs that are available in the real-world setup, is also found in works that learn robot behaviours in simulation [15], [16]. In [15], sim-to-real quadrupedal locomotion is achieved by first training a teacher policy that has access to privileged state information of the robot, allowing it to quickly achieve high performance. The teacher is then used to guide the learning of a student controller that only uses sensors available on the real robot. Similarly in [16], the student-teacher paradigm is used to

*This work was supported by the Research Foundation Flanders (FWO) under grant agreement no. 1S15925N and 1S56024N, and by the euROBIN Project (EU grant number 101070596).

¹Remko Proesmans, Thomas Lips and Francis wyffels are with the AI and Robotics Lab (IDLab-AIRO), Ghent University—imec, Ghent, Belgium remko.proesmans@ugent.be

learn in-hand reorientation. The teacher used the object orientation to quickly learn appropriate actions, after which the student was trained to match the teacher’s outputs while relying on occluded visual observations of the objects.

Instrumentation like in [12], [13] can be a way to obtain similar privileged information directly in the real world, potentially allowing for faster learning and/or more performant policies and avoiding the need for sim-to-real transfer, which already is a significant challenge on its own. We have previously applied it to learned cloth manipulation [17] and state estimation [18]. In this work, we present a case study on the use of instrumentation for demonstration learning, focusing on the task of squeezing a bottle to maintain a constant liquid flow. We will show that (1) instrumentation can be used for automated collection of demonstrations by a “teacher” agent, (2) that these demonstrations are more qualitative than those collected by human teleoperators, and (3) that this quality difference improves the performance of a policy trained on those demonstrations, compared to training on human demonstrations.

II. MATERIALS & METHODS

A. Task Description

With the aim of examining how instrumentation can be applied to demonstration learning, we chose the task of squeezing a constant flow of liquid from a bottle. The flow rate is directly related to the pressure inside the bottle near the opening, so a constant flow can be achieved by integration of a pressure sensor and controlling its output.

Plastic bottles are popular test objects in robotic manipulation [19]–[23] because they are ubiquitous in human environments and deformable, which presents unique challenges for e.g. state observation. In [21], [23] specifically, accurate dosing of liquids was tackled with tactile sensing. Potential applications of this skill include culinary arts [23] and the oiling or glueing of mechanical parts.

B. Hardware setup

1) *Instrumented Bottle*: A commercial condiment squeeze bottle is instrumented by integration of a MS5803 waterproof pressure sensor, see Fig. 1a. The supply and communication wires of the sensor are pulled through a small hole in the side of the bottle. To make the sensor PCB as well as the cable hole waterproof, silicone (Silicone Addition Colourless 50 by Silicones and more) is applied all around it with a syringe. The silicone is allowed to cure for 30 minutes before application, so that it is viscous enough to not seep away immediately. For readout, we repurposed a PCB from our open-source, modular, wireless Smart Textile [18], featuring an nRF52840 microcontroller. The PCB, along with a 160mAh LiPo battery, a micro USB charging interface, and an on/off button, is placed in a 3D printed holder and attached to the flip-top cap of the bottle with hook-and-loop tape. The bottle is filled with a mixture of water and 2 wt% xanthan gum for thickening. The bottle opening is a small hole without valve.

2) *Robot Setup*: A Schunk EGK-40-MB-M-B gripper is mounted to a UR3 robot, as seen in Fig. 1c. One of the fingers is a standard Robotiq 2F-85 fingertip, the other is one of our custom tactile fingertips based on a magnetic transduction principle [24]. The fingertip has a 2-by-2 grid of taxels with a 6 mm pitch, each taxel outputs local 3D contact force estimation. A custom 3D-printed adapter allows for attaching Robotiq-compatible fingertips to the Schunk. In addition, an Intel RealSense 435 camera and an Arduino Nano 33 BLE (for readout of the tactile fingertip) are attached to the Schunk gripper via custom mounts. The instrumented bottle is placed in a wooden holder before the robot grasps it. When the bottle is grasped, it is turned over above a Kern PCB2000-1 scale. The scale has a resolution of 0.1 g.

3) *Teleoperation Setup*: The trigger of a Gello [14] arm (Fig. 1e) is configured to control the opening of the Schunk gripper. A human operator (Fig. 1f) squeezes the trigger while visually maintaining a constant flow.

C. Data Collection Protocol

Five different datasets are collected. For the first, different people manually squeeze the instrumented bottle in whatever way they deem fit, while attempting to maintain a constant flow by looking at the stream. This lasts 10 seconds, during which time the pressure in the bottle is recorded. Five people were each given a full bottle, and performed 10 s trials until the bottle was empty. This resulted in 34 trials.

The other four datasets involve the setup in Fig. 1: a training dataset collected by a PI (PID with zero derivative gain) controller, a training dataset collected a human teleoperator, an evaluation dataset of rollouts from a policy Π_{PI} trained on the PI data, and an evaluation dataset of rollouts from a policy Π_{Teleop} trained on the teleoperation data.

For these four datasets, the data collection procedure is illustrated in Fig. 2. First, the robot grasps the bottle from the holder and moves it above the Kern scale (see Fig. 1d). Then, an initial pressure P_{init} is determined at random. P_{init} is to be reached before the agent attempts to maintain a constant flow, such that the training dataset contains trials with a variety of initial conditions (zero flow, drizzle, high flow). For the PI training set and the evaluation sets of Π_{PI} and Π_{Teleop} , P_{init} is sampled from a uniform distribution with a 5 kPa window around the pressure reading when the bottle is not squeezed, P_{rest} . With P_{init} determined, the gripper starts squeezing at 5 mm/s until P_{init} is reached, at which point the agent attempts to maintain a constant flow. The human teleoperator, not having access to instrumentation data, was tasked to start half of the trials from zero flow, half with a flow larger than what they would eventually settle on. All teleoperation trials are collected by the same person.

After initialisation, a 15 Hz control loop starts. Every iteration, an observation is recorded, composed of an RGB frame of the wrist camera, the internal pressure in the bottle, the current gripper opening, the tactile forces, and the Gello trigger position (teleoperation only). In addition, the weight reading of the Kern scale is recorded at 3–4 Hz. The scale reading is only used as an extra evaluation of the liquid flow

stability, its readout frequency is too low for use as the PI controller's process variable. Based on this observation, the agent will compute the required gripper action. The agent update method differs among the four agents (PI controller, human teleoperator, Π_{PI} , Π_{Teleop}):

- **PI:** The relative gripper movement $u[k]$ in mm at timestep k is computed as:

$$\begin{cases} u[k] = K_p e[k] + I[k] \\ I[k] = \alpha I[k-1] + K_i T_s e[k] \end{cases} \quad (1)$$

K_p and K_i are the proportional and integral gain respectively. $e[k]$ is the error on the process variable, which is the pressure in the instrumented bottle. The setpoint is 1.2 kPa above P_{rest} , which resulted in the lowest flow while still maintaining a constant stream. Specifying the setpoint with respect to P_{rest} ensures that changes in air pressure due to weather conditions have no effect. The sample time T_s is the inverse of the control frequency (15 Hz), and α is an exponential back-off factor to reduce the impact of previous errors. The values used for these parameters are included in Fig. 2.

- **Teleop:** The human operator looks at the bottle and attempts to keep the stream constant, squeezing the Gello trigger accordingly. The encoder in the Gello trigger outputs an angle that is linearly mapped to the Schunk opening width, such that a released trigger corresponds to a loose grasp of the bottle, and a fully squeezed trigger corresponds to the minimum gripper opening W_{min} of 2 cm. The operator chose the smallest flow that still maintained a constant stream, mimicking the chosen setpoint for the PI controller.
- **Π_{PI} & Π_{Teleop} :** These are neural networks (NN), see section II-D. An agent update constitutes a forward pass of the NN.

A squeezing trial lasts until the maximum scale weight S_{max} of 25 g or the minimum gripper opening W_{min} has been reached. The bottle is turned right side up to equalise the air pressure inside. If the bottle is empty, it is placed down to be manually refilled. If it is not yet empty, the robot turns the bottle upside down again, and a new trial starts immediately. Detecting whether the bottle is empty can be done by comparing the pressure readings before and after turning the bottle over: remaining liquid will add to the pressure reading when the bottle is upside down. If both readings are approximately equal, the bottle is empty.

For each of the training datasets collected with the PI and Teleop agents, 42 trials were collected, which corresponds to five full bottles squeezed until empty. For the evaluation datasets, collected with the Π_{PI} and Π_{Teleop} agents, four full bottles were squeezed until empty, resulting in 32 trials each.

D. Learning Architecture

Π_{PI} and Π_{Teleop} are Action Chunking Transformer (ACT) models [2] with a ResNet18 vision backbone. They both predict a relative gripper move, given an observation that consists of a 340x480 wrist camera image and a 13-dimensional state vector. The state vector is composed of

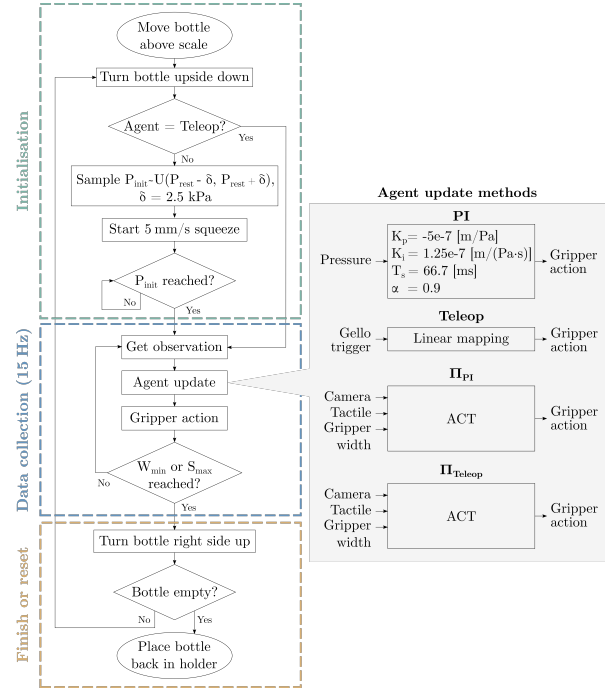


Fig. 2: Control flow diagram for data collection.

the current gripper opening and the 3D force readings of the four taxels in the tactile fingertip. We set the chunk size to 10, such that the model predicts a sequence of 10 actions, which corresponds to 0.6 seconds of wall-clock time at 15 Hz. Training a model took about 1.5 h on an NVIDIA RTX4080 GPU. Inference takes 12 ms on an NVIDIA 3090 GPU. Unlike the original ACT paper, we do not use temporal ensembling nor do we execute multiple actions for each chunk: we found that this highly dynamic task benefited from frequent predictions, and only execute the first action of every predicted chunk.

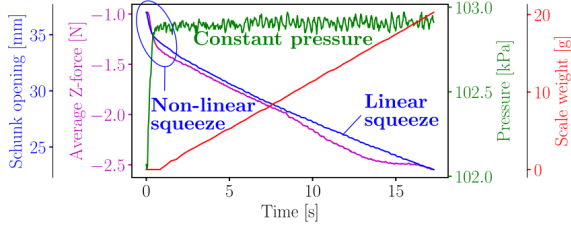
E. Evaluation

The measure of quality for a squeeze trial is the standard deviation σ of the time series pressure reading as a proxy for the flow rate. The first two seconds of a trial are discarded, to exclude transient effects. We refer to σ as the “score”.

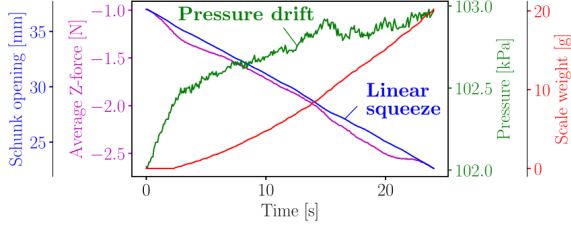
III. RESULTS & DISCUSSION

A. Task Characterisation

Fig. 3a shows a typical trial of the PI controller. This trial started from a P_{init} below the setpoint, so the gripper quickly closes at the start, increasing the internal pressure. The PI controller then maintains an approximately constant pressure, with a σ of 26 Pa. The scale clearly shows that the liquid flow is constant: the weight curve rises at 1.22 g/s with a root mean square error (RMSE) of 0.07 g. It appears that the PI controller settles on a linear squeeze of 0.6 mm/s. This raises the question whether it is sufficient to squeeze at a constant speed, rather than attempting to learn reactive behaviour. Fig. 3b shows that reactive behaviour is necessary: if the gripper squeezes at a constant linear speed of 0.6 mm/s



(a) The PI controller settles on a gripper speed of 0.6 mm/s.



(b) Forcing a gripper speed of 0.6 mm/s from the start results in an internal pressure drift.

Fig. 3: A constant flow requires non-linear squeezing.

from the start, the internal pressure drifts. The RMSE of a linear fit on the scale weight curve increases more than tenfold, to 0.85 g. A similar result was obtained for speeds of 1.1 and 1.25 mm/s. The gripper opening in Fig. 3a resembles the manually tuned finger motion used in [23].

B. Agent Comparison

For each collected trial, the standard deviation σ is computed. The results are compiled in histograms with a bin width of 25 Pa, see Fig. 4. The manual trials show a very wide spread, some even reaching a σ of over 1000. This is because some people naturally performed better than others, but also because people tried different squeezing strategies, some of which were ineffective. For clarity of the figure, the final histogram bin aggregates all scores over 400. The lowest σ achieved was 105, which serves as a reference point to compare the performance of other agents. Notably, 40 % of the teleoperation demonstrations score better than the best-case manual performance. This is attributed to the Gello trigger taking less force to squeeze than the bottle itself, making it physically less demanding for the teleoperator. On average, the teleoperated trials had a σ of 126 ± 58 Pa. We argue that our teleoperated demonstrations are “good”, in the sense that they are comparable to best-case manual performance. The PI training set scores by far the best, with an average σ of 30 ± 7 Pa.

The evaluation set of the Π_{PI} policy has a mean score of 93 ± 30 Pa, compared to 144 ± 56 Pa for the Π_{Teleop} policy. Assuming Gaussian priors, we can calculate the probability that Π_{PI} performs better than Π_{Teleop} . Let X_{PI} be the σ of a random Π_{PI} trial and X_{Teleop} of a Π_{Teleop} trial:

$$\begin{cases} X_{PI} \sim N(93, 30) \\ X_{Teleop} \sim N(144, 56) \end{cases} \Rightarrow X_{Teleop} - X_{PI} \sim N(51, 64) \quad (2)$$

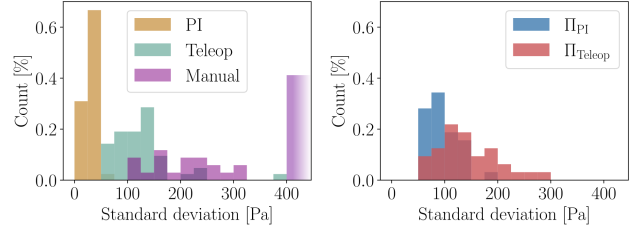


Fig. 4: Comparing agent scores.

$$\begin{aligned} P[\Pi_{PI} \text{ better than } \Pi_{Teleop}] &= P[X_{PI} < X_{Teleop}] \\ &= P[X_{Teleop} - X_{PI} > 0] \quad (3) \\ &\stackrel{(2)}{=} 78 \% \end{aligned}$$

We conclude that, given the choice between Π_{PI} and Π_{Teleop} , choosing Π_{PI} will lead to a better result in 78 % of cases. This is an added benefit on top of the partial automation of demonstrations for Π_{PI} , because of which data collection required significantly less human effort. However, instrumentation requires hardware design effort before data collection can start. This raises important trade-offs: Is the performance gain achieved by exploiting instrumentation worth the initial design effort? Is data collection prohibitively time consuming at scale, so that an initial investment in instrumentation can save time overall? The answer to these questions may change on a case-by-case basis, but we believe that, in the search for generalist agents, instrumentation has a lot of unexplored potential.

IV. CONCLUSION AND FUTURE WORK

We developed an instrumented squeeze bottle to aid in learning the task of squeezing a constant flow from the bottle. Five datasets were collected. It was found that manual human performance varied wildly, and that human teleoperation can be easier than manual operation if the required grasping force to actuate the teleoperation mechanism is less than for the bottle itself. Still, a PI controller exploiting the instrumentation data significantly outperformed the teleoperator. In addition, like [9] and [10], we explicitly showed that better demonstrations lead to a better policy: an ACT policy trained on the automated PI demonstrations performs better than an ACT policy trained on the teleoperated demonstrations in 78 % of cases. Hence, instrumentation not only reduces human effort during data collection, but can also improve the performance of learned policies.

In future work, the generalisation capabilities of the policies should be evaluated. Our results are promising, but for practical applications the true quality of our squeezing policy should be measured on out-of-distribution bottles and environments. In addition, we have shown just one possible approach for incorporating instrumentation in imitation learning. Our aim is to develop similar case studies to build an overview of the possible benefits of instrumentation for robot learning.

REFERENCES

- [1] Y. Hu, Q. Xie *et al.*, “Toward general-purpose robots via foundation models: A survey and meta-analysis,” 2024. [Online]. Available: <https://arxiv.org/abs/2312.08782>
- [2] T. Z. Zhao, V. Kumar *et al.*, “Learning fine-grained bimanual manipulation with low-cost hardware,” 2023. [Online]. Available: <https://arxiv.org/abs/2304.13705>
- [3] C. Chi, Z. Xu *et al.*, “Diffusion policy: Visuomotor policy learning via action diffusion,” *The International Journal of Robotics Research*, 2024.
- [4] K. Black, N. Brown *et al.*, “ π_0 : A vision-language-action flow model for general robot control,” 2024. [Online]. Available: <https://arxiv.org/abs/2410.24164>
- [5] M. J. Kim, K. Pertsch *et al.*, “Openvla: An open-source vision-language-action model,” 2024. [Online]. Available: <https://arxiv.org/abs/2406.09246>
- [6] E. Collaboration, A. O’Neill *et al.*, “Open x-embodiment: Robotic learning datasets and rt-x models,” 2024. [Online]. Available: <https://arxiv.org/abs/2310.08864>
- [7] A. Khazatsky, K. Pertsch *et al.*, “Droid: A large-scale in-the-wild robot manipulation dataset,” 2024. [Online]. Available: <https://arxiv.org/abs/2403.12945>
- [8] H. Walke, K. Black *et al.*, “Bridgedata v2: A dataset for robot learning at scale,” 2024. [Online]. Available: <https://arxiv.org/abs/2308.12952>
- [9] M. Sakr, Z. J. Li *et al.*, “Quantifying demonstration quality for robot learning and generalization,” *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9659–9666, 2022.
- [10] M. Bilal, N. Lipovetzky *et al.*, “Beyond success: Quantifying demonstration quality in learning from demonstration,” in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, pp. 5120–5127.
- [11] C. Xu, Q. Li *et al.*, “Rldg: Robotic generalist policy distillation via reinforcement learning,” 2024.
- [12] K. Junge, C. Pires, and J. Hughes, “Lab2field transfer of a robotic raspberry harvester enabled by a soft sensorized physical twin,” *Communications Engineering*, vol. 2, no. 1, p. 40, Jun 2023.
- [13] OpenAI, I. Akkaya *et al.*, “Solving rubik’s cube with a robot hand,” 2019.
- [14] P. Wu, Y. Shentu *et al.*, “Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators,” 2024.
- [15] J. Lee, J. Hwangbo *et al.*, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, Oct. 2020. [Online]. Available: <http://dx.doi.org/10.1126/scirobotics.abc5986>
- [16] T. Chen, M. Tippur *et al.*, “Visual dexterity: In-hand reorientation of novel and complex object shapes,” *Science Robotics*, vol. 8, no. 84, p. eadc9244, 2023. [Online]. Available: <https://www.science.org/doi/abs/10.1126/scirobotics.adc9244>
- [17] A. Verleysen, T. Holvoet *et al.*, “Simpler learning of robotic manipulation of clothing by utilizing diy smart textile technology,” *Applied Sciences*, vol. 10, no. 12, 2020. [Online]. Available: <https://www.mdpi.com/2076-3417/10/12/4088>
- [18] R. Proesmans, A. Verleysen *et al.*, “Modular piezoresistive smart textile for state estimation of cloths,” *Sensors*, vol. 22, no. 1, 2022.
- [19] T. Li, X. Sun *et al.*, “Robot grasping system and grasp stability prediction based on flexible tactile sensor array,” *Machines*, vol. 9, no. 6, 2021.
- [20] N. Kuppuswamy, A. Alspach *et al.*, “Soft-bubble grippers for robust and perceptive manipulation,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 9917–9924.
- [21] T. Hellebrekers, K. Zhang *et al.*, “Localization and force-feedback with soft magnetic stickers for precise robot manipulation,” in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 8867–8874.
- [22] L. Pozzi, M. Gandolla *et al.*, “Grasping learning, optimization, and knowledge transfer in the robotics field,” *Scientific Reports*, vol. 12, no. 1, p. 4481, Mar 2022.
- [23] H.-J. Huang, J. Xiang, and W. Yuan, “Kitchen artist: Precise control of liquid dispensing for gourmet plating,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 13 933–13 939.
- [24] L. V. den Stockt, R. Proesmans, and F. wyffels, “Automatic calibration for an open-source magnetic tactile sensor,” 2024.