Update of the SCARAB robot to sort valuable items in containers of residual waste

Daniel Reischl[1], Johannes Wenninger, Simon Zwirtmayr and Johannes Schröck

***Abstract*— In this paper the features of the autonomous mobile robot SCARAB are extended. SCARAB is now not only exchanging full waste containers with empty ones but also sorting out the valuable objects of the waste. For this task, a gripper was added to the robot's end-of-arm tool. The fingers of the gripper have a Fin Ray design to robustly grasp the objects. Adaptions of the waste container allow to empty the waste onto a sorting table without additional actuators. Object detection is done with a YOLOv8 model which was initially trained with an open data set and improved with additional training data. In order to label this training data a standalone tool based on the Segment Anything Model (SAM) was developed. The paper shows the mechanical design of the gripper fingers, the adaption of the waste container as well as the design of a suitable sorting table. It is demonstrated that the waste sorting task is carried out robustly without the need of any additional expensive equipment.**

***Index Terms*— object detection, segmentation, waste sorting**

I. INTRODUCTION

Automated image-based recognition and sorting of waste using robots is already being used commercially worldwide. Companies such as ZenRobotics, WasteRobotics, AMP Robotics, Recycleye, Machinex, Bollegraaf, Green Machine and many others offer solutions for efficient sorting on a conveyor belt. However, efficient object recognition is also still a topic of research [7].

This paper, however, is not about a highly efficient implementation of a waste sorting system with expensive cameras and fast delta robots. Our focus is on the subsequent and cost-effective retrofitting of an existing robot, which is used already to autonomously exchange full waste containers for empty ones.

The development platform SCARAB [10] was able to collect full waste containers on demand autonomously and bring them back to a garage. With this setup however it was not possible to sort the waste and all the waste was treated as "residual waste". In January 2025 a deposit on non-returnable containers was put into force in Austria [1], which changed the requirements for the SCARAB platform. As minimum requirement, at least the containers (bottles, cans, etc.) which are subject of the deposit have to be identified and separated of the waste automatically. In this paper the challenges of adapting an existing mobile robot to this new task are described as well as the technical solutions applied for a successful implementation.

[1]All authors are with Linz Center of Mechatronics GmbH, Altenberger Straße 69, 4040 Linz, Austria `daniel.reischl@lcm.at`

Fig. 1. SCARAB during operation while changing the container.

II. SCARAB DEVELOPMENT PLATFORM

The mobile robot SCARAB shown in Fig. 1 was designed to drive autonomously in a semi-public area and exchange the full waste containers. As presented in [10], a sensor in the waste container reports the filling height and a mission to exchange the container is initiated, if the boundary conditions (e.g. weather) are fulfilled. The entire process is not time-critical and the main focus is on personal safety. The new task of sorting waste is therefore carried out in a locked garage to which no passers-by have access. The garage door is controlled automatically via the higher-level mission control system.

After returning back to the garage, SCARAB is now driving to a sorting table. The full waste container is emptied onto the sorting table with the robot arm. No additional actuators or sensors are necessary for the robotic arm or the waste container as shown in section III in more detail. The pile of waste on the sorting table is slightly distributed by a statically programmed movement of the robot arm to facilitate object recognition. A picture of the waste is taken with the wrist camera of the robotic arm. Based on this picture, the valuable items in the waste are detected, as shown in section IV. The recognized objects are sorted out of the waste one by one and separated in the appropriate containers. The waste remaining after the sorting process can then be fed into an appropriate residual waste container by tilting the sorting table. Once the sorting process is complete, SCARAB picks up the empty waste bin and moves to the charging

https://doi.org/10.34749/3061-0710.2025.1

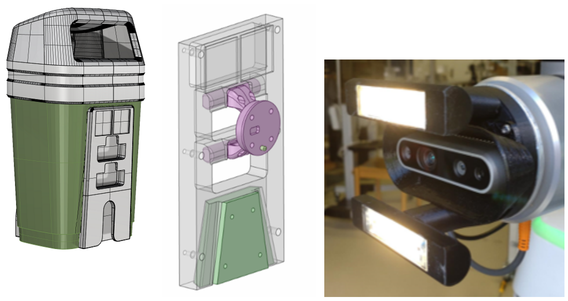

Fig. 2. Already existing passive end of arm tool with Realsense camera and LED lights to manipulate the containers.

station to wait for its next mission.

III. MECHANICAL ADAPTIONS

The design of SCARAB should not be changed, but additional features are necessary to perform the sorting process. In order to solve this challenge, mechanical adaptions of the waste containers were necessary as well as to add a sorting table and a gripper.

A. Container

The waste bin has a lid with an integrated fill level sensor. This configuration with lid and the robot end effector, which picks up the waste bin via a form-fit connection shown in Fig. 2, do not allow the bin to be emptied by turning it over.

A mechanism has therefore been developed that allows the base of the container to be opened. This mechanism opens the base when the container is pressed against the rear wall of the sorting table with the robot arm, shown in Fig. 3 and Fig. 4. After the contents of the container have fallen out, the bottom of the container is closed again with a suitable trajectory. Both processes, opening and closing, are carried out without additional actuators but solely by pressing the container against the sorting table. The empty waste container is put on a fixture and the robotic arm with the gripper is now free for the sorting task.

B. Sorting table

A suitable sorting table was set up, which allows SCARAB to attach the waste container to the table and then move partially under the table itself. In this way, it is possible to optimize the working space of the robot arm. The sorting table has 2 storage bins, to the left and right of the sorting surface, into which the cans and bottles are deposited. Once the sorting process is complete, the sorting area can be tilted with the robot arm and the remaining waste falls into a residual waste container, as shown in Fig. 5. The sorting table is not equipped with any actuators or electronics.

C. Gripper

A passive gripper system was originally developed for manipulating the waste containers to ensure the most robust

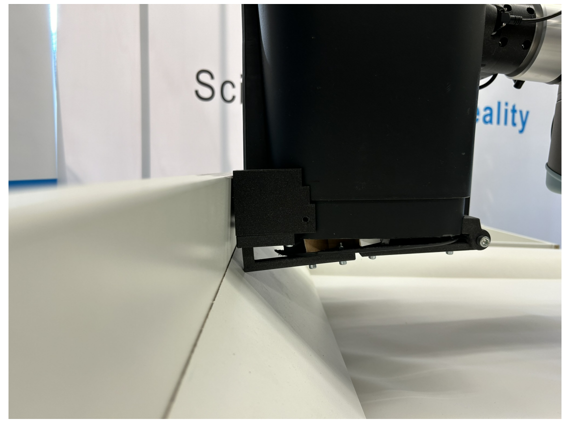

Fig. 3. Mechanism to open the container at the bottom without actuators by pressing the container to the rear wall of the sorting table.

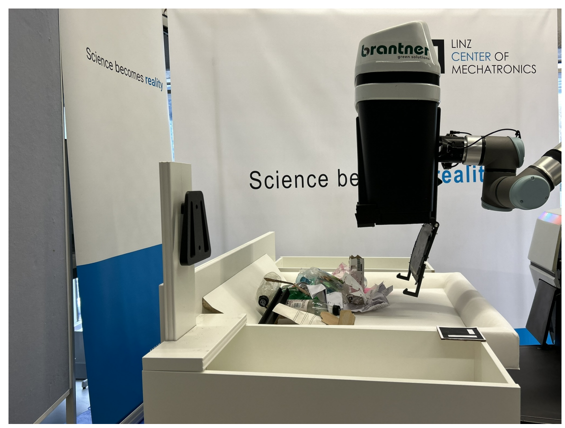

Fig. 4. Container with opened bottom. The container will be placed on the fixture after closing the bottom.

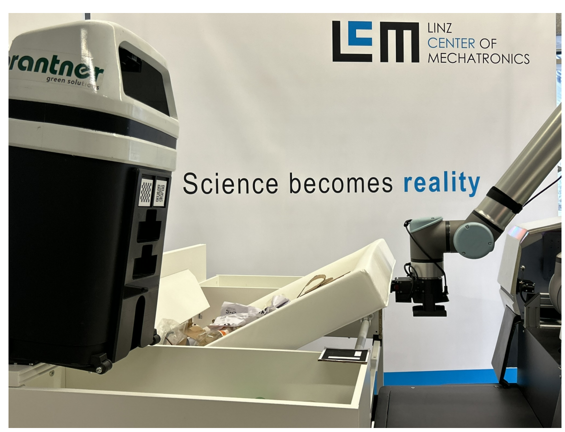

Fig. 5. The sorting table is also operated by the robotic arm without additional actuators.

https://doi.org/10.34749/3061-0710.2025.1

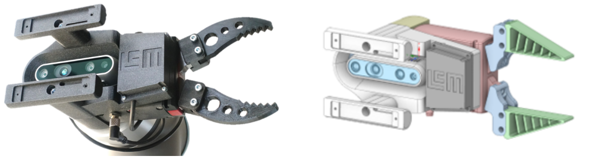

Fig. 6. Different designs of the gripper: Left with rigid fingers, right with fin ray design. In the right figure, the Gimatic gripper is shown in red and blue and the capacitor box is under the lid with the LCM logo.

Fig. 7. Elastic fingers with fin ray design in opened and closed configuration. (The LED lights still need to be installed next to the Realsense camera.)

and safe handling possible. However, a gripper is now required for sorting the waste.

A self-centring electric angle gripper from Gimatic was used as the gripper for sorting the waste. This fulfils the special requirements in terms of available installation space, closing force and compatibility with the robot arm. The exact type of gripper is 'MPBM3240'. The gripper requires additional external control electronics (Capacitor Box CAPBOX3200-03), which must be used to provide the power for the Gimatic gripper. Without these electronics, the power requirement at the pins on the wrist of the UR10e robot arm could not be covered. Furthermore, different fingers can be used thanks to the modular design. The gripper also allows rapid adaptation to other problems, as the fingers can be created and customised using rapid prototyping.

As the available installation space is very limited in the folded state, the angular gripper was integrated into the existing robot end effector to save as much space as possible. An adapter plate was designed for this purpose, which must be fitted to the robot's wrist in the first assembly step. The Gimatic angular gripper can then be screwed onto this adapter and fixed in place. The electronics of the Capacitor Box are located directly in front of the gripper on the robot end effector. The original robot end effector has been adapted accordingly so that it can be mounted on the adapter with the Intel Realsense camera fitted.

The first version of the fingers was 3D-printed from TPU (thermoplastic material) and is shown in Fig. 6. As the narrow design of the fingers led to twisting when gripping and different objects were not always gripped correctly, a new finger design was tested.

The new design of the fingers was based on the so-called Fin Ray design, which has already been successfully used in the literature to grasp variable shapes, [8], [3], [11]. This design is originally biologically inspired by the tail fins of fish and patented by the company Evologics GmbH. The company Festo offers commercial products of gripper fingers based on this concept. The soft gripper used in our studies is lightweight (entirely 3D-printed from TPU), has a simple structure, high compliance and adaptability, and is capable of grasping objects of any geometry. Fig. 7 shows how the principle of the Fin Ray design works: In the unloaded state, the design retains its original shape. If any object is gripped (for example an already deformed aluminium can), the gripper automatically adapts to the shape of the object. This enables various objects to be gripped safely. As the inner gripper surfaces are aligned parallel to each other in the open state, the object is automatically pressed towards the gripper when gripping.

It is also important to mention that the material of the functional model (except the fingers) is PLA (polylactide). PLA is not resistant to ultraviolet radiation (UV) and should be replaced with a UV-resistant material if necessary. If the first tests are successful, a change to the commercial product of FESTO will be considered.

IV. MANIPULATING THE OBJECTS

In order to sort out the valuable objects, it is necessary to identify them within the residual waste, grab them robustly and place them in separate containers.

A. Segmentation with YOLOv8 model

An instance segmentation model was selected in order to not only obtain a bounding box of the objects, but also to detect the exact contour of the waste object. This property is important in the later calculation of the gripping point in order to be able to analyse the shape of the object. Therefore, a YoloV8 model [5] was used to detect the valuable parts of the waste.

The model was first trained with the TACO dataset [9], an open image dataset of waste in the wild. This dataset has 63 classes of objects but only "clear plastic bottle", "drink can" and "food can" are used in our work.

To create additional training images, the waste container was filled and opened several times from a defined height in the center of the table. In total 200 photos were taken of different waste distributions on the table. 160 photos were used for training and 40 for validation. A semi-automatic labeling tool was developed as all contours of the objects must first be labeled for each photo in order to be able to train the network later. This would be very time-consuming with manual labeling. The Segment Anything Model (SAM) [6] implemented by Meta was used for this purpose, which saves a great deal of time when labeling the waste objects. The online version of SAM can not be used for generating the training data as no labels are available. Meta provides the code as open source and it was possible to use this code for developing a standalone offline tool for labeling the images taken in our lab. The workflow is the following:

1) Click on a single object in the image and SAM will highlight automatically (at least a part) of the object.

Fig. 8. Validation of the segmentation: The training data labeled with the offline tool based on SAM.

Fig. 9. Validation of the segmentation: Result of the YOLOv8 model with the same picture as in the training.

2) Add or substract parts of the object by continue clicking with the mouse.
3) When the entire object is highlighted, enter the appropriate label for the object.
4) Continue with the next object in the image.
5) When all objects in the image are labeled, continue to the next image.

Comparing Fig. 8 and Fig. 9 shows the good results of the segmentation algorithm, including concealed and deformed objects. For the manual labeling only the classes ”food can” and ”clear plastic bottle” have been used, which will be called ”bottle” and ”can” in the following.

B. Gripping pose

The Realsense camera is used to find the gripping positions of the objects. The first step is to take a 2D photo and a depth image with the camera mounted on the robotic arm from a well defined position right above the sorting table.

With the YoloV8 model, the objects are segmented in the 2D photo and processed one after the other. As output of the YoloV8 model the contour of each object is provided in 2D together with a label and a numerical value for the confidence, as shown in Fig. 10 for a bottle which is obstructed by a sheet of paper. With the function minAreaRect of OpenCV library [2] the center point, orientation and main axis of the object contour are computed. The distance between the camera and the gripping point is determined with an ArUco marker [4]. The 3D position of this gripping point can then be calculated using the usual camera calibration algorithms. The following assumptions are made in order to calculate the 6D pose of the gripping point from the position: The gripper is parallel to the image plane and rotated around the global vertical axis corresponding to the rotation of the 2D object contour, as shown in Fig. 11. A safe gripping of the objects was observed, even in the cases when only small parts of the object are visible, as shown in Fig. 12.

V. TEST RESULTS

The robustness of the waste sorting process described above was tested extensively. The objects to be sorted out of the residual waste were not part of the training data and can be seen in Fig. 13. The test data consists of 4 bottles and 5 cans. In addition to these desired objects, the test waste contains 15 disturbing objects, which were also not part of the training data: Plastic packaging films, cardboard and paper.

The tests were done in the following way:

1) fill the waste (desired and disturbing objects) into the bin and mix thoroughly
2) empty the waste on the sorting table
3) distribute the waste with the robotic arm
4) take a picture of the waste
5) grasp a desired object and put it into the bins next to the sorting table
6) repeat steps (4) and (5) until no more desired objects are detected

https://doi.org/10.34749/3061-0710.2025.1

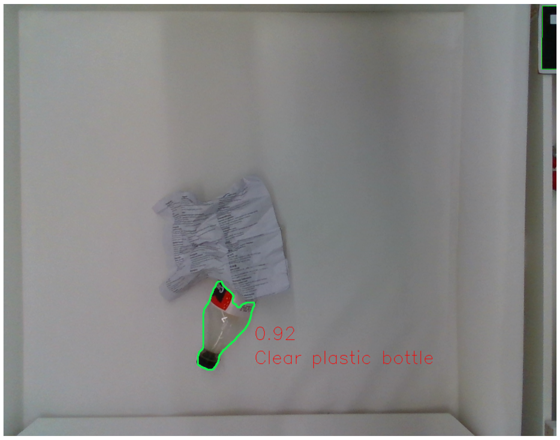

Fig. 10. Output of the YoloV8 model: visible contour (shown in green) and label (with confidence) of the object.

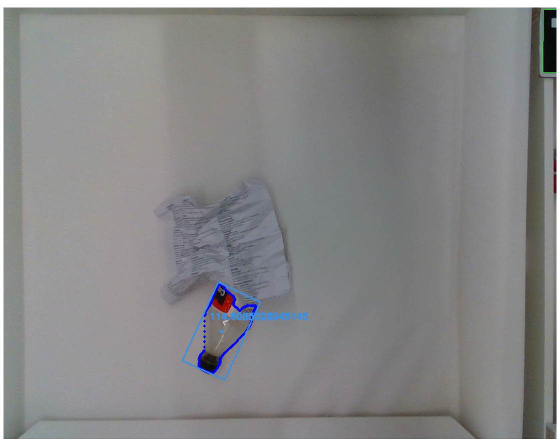

Fig. 11. Computation of the gripping point: bounding box (light blue rectangular) with center point and its rotation in degree.

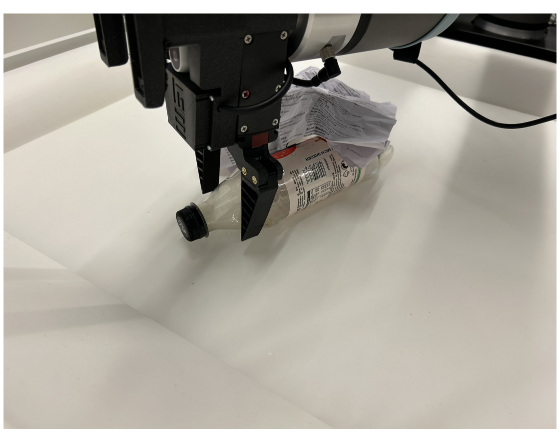

Fig. 12. Gripping in the center of the bounding box of the (visible) contour of the bottle.

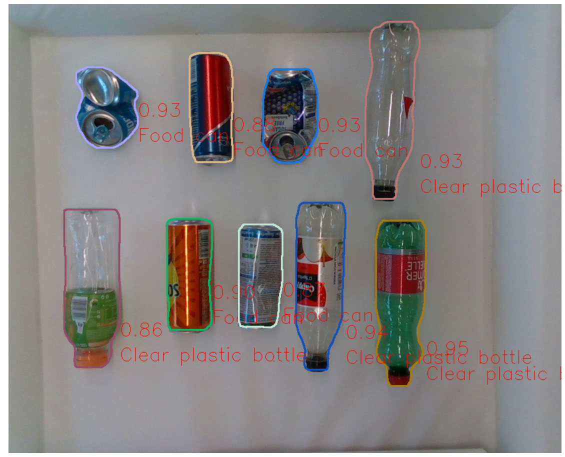

Fig. 13. Test objects with their classes according to the YoloV8 model.

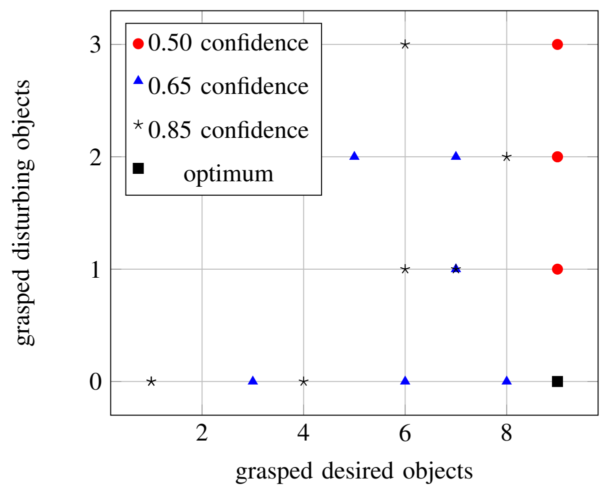

Fig. 14. Results of the tests with different settings of the confidence parameter

These tests were repeated multiple times with different settings. The most significant parameter regarding the performance of the sorting process was the confidence of the segmentation step. A low confidence value leads to a high number of successfully picked objects. However, you have to accept that a few unwanted objects will also be picked up. In Fig. 14 the results of the tests are shown. The optimal result would be to pick 9 out of 9 desired objects and 0 out of 15 disturbing objects. If the same result was observed multiple times with the same setting, the result is still just shown as a single point in the graph. The graph shows, that it was not possible to reach the optimal result with any setting and that it was not possible to strictly avoid grasping disturbing objects. However with setting the confidence to 50% it was possible to reach a robust result of picking all desired objects while accepting to pick 1 to 3 of the 15 disturbing objects.

A more detailed evaluation of the detection process was done to study the influence of the confidence value. For each of the 150 photos taken during the tests, it was analyzed

https://doi.org/10.34749/3061-0710.2025.1

TABLE I

INFLUENCE OF THE CONFIDENCE SETTINGS ON THE DETECTION RATE. AVERAGE VALUES BASED ON 150 IMAGES.

confidence setting	correct detected objects	wrong objects per image
0.50	83%	0.57
0.65	40%	0.19
0.85	34%	0.17

TABLE II

RECORDED TIME FOR THE OBJECT DETECTION IN SECONDS. AVERAGE VALUES BASED ON THE RECORDING OF 16 OBJECTS.

take picture	0.046
preprocess picture	0.056
segment object (incl. saving the picture)	2.182
compute bounding box	0.158
compute gripping pose (incl. 2 coordinate transformations)	0.010

how many of the desired objects depicted were correctly recognized and how many of the undesired objects were erroneously marked. The average values for the 3 different settings can be seen in Table I. The value for the ”wrong objects per image” is an absolute value and is between 0.17 and 0.57 objects per image. The correctly detected objects are given as percentage of the desired objects in the image and differs between 34% for a high confidence value and 83% for a low confidence value.

The time required for object recognition depends heavily on the hardware used. In the tests shown here, the photo was taken using a Realsense camera, the data was read out via the RTDE interface of the Universal Robot and then analyzed with a Python script. The evaluation was carried out on a NUC (Next Unit Computing). All computations are performed locally with hardware located in the SCARAB platform. The duration of the individual steps is shown in Table II.

VI. SUMMARY AND OUTLOOK

It was demonstrated how the functionality of an existing mobile robot was extended with low cost hardware to add the feature of waste sorting. The low amount of training data in the lab still limits the quality of the overall performance but was sufficient to find the most significant parameter. A good choice of the limit for the confidence in the segmentation step has large impact on the results. During operation SCARAB will collect much more (and more realistic) training data on a daily basis which will lead to a more robust performance.

ACKNOWLEDGMENT

This work has been supported by the COMET-K2 Center of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

REFERENCES

[1] P. Beigl and A. Allesch, “Einwegpfand in österreich: Gestaltung und herausforderungen im internationalen vergleich,” *Österreichische Wasser-und Abfallwirtschaft*, pp. 1–11, 2024.

[2] G. Bradski, A. Kaehler, *et al.*, “Opencv,” *Dr. Dobb’s journal of software tools*, vol. 3, no. 2, 2000.

[3] J. M. Gandarias, J. M. Gómez-de Gabriel, and A. J. García-Cerezo, “Enhancing perception with tactile object recognition in adaptive grippers for human–robot interaction,” *Sensors*, vol. 18, no. 3, p. 692, 2018.

[4] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, “Automatic generation and detection of highly reliable fiducial markers under occlusion,” *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[5] G. Jocher, J. Qiu, and A. Chaurasia, “Ultralytics yolo (version 8.0.0) [computer software],” 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

[6] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, “Segment anything,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.

[7] W. Lu and J. Chen, “Computer vision for solid waste sorting: A critical review of academic research,” *Waste Management*, vol. 142, pp. 29–43, 2022.

[8] O. Pfaff, S. Simeonov, I. Cirovic, P. Stano, *et al.*, “Application of fin ray effect approach for production process automation,” *Annals of DAAAM & Proceedings*, vol. 22, no. 1, pp. 1247–1249, 2011.

[9] P. F. Proença and P. Simões, “Taco: Trash annotations in context for litter detection,” *arXiv preprint arXiv:2003.06975*, 2020.

[10] D. Reischl, J. Wenninger, J. Schröck, A. Reininger, M. Trinkl, J. Parz, and M. Weywoda, “Scarab - autonomous demand-oriented waste disposal with mobile robot,” in *Proceedings of The First Austrian Symposium on AI, Robotics, and Vision (AIROV24)*, 2024.

[11] Y. Yang, K. Jin, H. Zhu, G. Song, H. Lu, and L. Kang, “A 3d-printed fin ray effect inspired soft robotic gripper with force feedback,” *Micromachines*, vol. 12, no. 10, p. 1141, 2021.

https://doi.org/10.34749/3061-0710.2025.1