Analysis of an Intelligent Robotic Grasping System for Waste Sorting

Ziang Xiao[1], Meixin Wang[2], Zhiyi Yuan[3,*]

{18990908765@163.com[1], mayyyxwang@gmail.com[2], zhiyi.yuan.1@student.unimelb.edu.au[3,*]}

Sino-German College of Applied Sciences, Tongji University, Shanghai, 200092, China[1], School of Information Science and Technology, Southwest Jiaotong University, Chengdu Sicuan, 611756, China[2], Melbourne School of Design, The University of Melbourne, Melbourne, Victoria, 3010, Australia[3]

Abstract: The increasing global municipal solid waste generation and the limitations of manual sorting underscore the urgent need for intelligent robotic solutions. This paper systematically analyzes intelligent robot grasping systems for garbage sorting, focusing on the synergistic integration of key technologies: multimodal perception fusion (vision and tactile), adaptive shared control, and soft gripper design. The review highlights that combining visual perception for global localization with tactile feedback for fine force adjustment significantly enhances grasping success in cluttered scenes. Furthermore, human-machine collaborative control balances operational efficiency with safety, while compliant soft grippers improve adaptability to irregular objects. Experimental validations demonstrate that such an integrated system architecture effectively improves robustness and operational performance in unstructured environments. This analysis provides a comprehensive technological reference and a viable pathway for deploying reliable intelligent grasping systems in practical waste sorting applications, contributing to environmental protection and the circular economy.

Keywords: Intelligent robot grasping system, Garbage sorting, Multimodal perception fusion, Soft gripper, Human-machine collaborative control

1. Introduction

In recent years, the annual generation of municipal solid waste has been up to two billion metric tons globally. However, only a small fraction is recycled effectively, with a more concerning situation in developing countries, particularly [1]. Due to numerous challenges in the sorting stage, a significant volume of potentially recyclable materials ultimately ends up being disposed of by landfill or incineration. Moreover, manual sorting is inefficient and costly [2], which leads to the increasing occupational health risks to workers due to prolonged and direct exposure to toxic and harmful wastes [3,4]. Mechanisation and intelligence substitution have become an inevitable choice to mitigate these occupational harms. Robotic grasping technology, characterised by its continuous operation capability and full traceability of data, is regarded as a key solution to overcome current challenges increasingly.

This article aims to systematically outline the key technologies of robotic grasping in waste sorting, analyse the progress and limitations of existing research, and prospect future development directions. The subsequent sections will first review the critical technologies of robotic grasping, then explore the design principles of integrated system architectures.

ICMEEA 2026, April 24-26, Singapore, Singapore

DOI 10.4108/eai.24-4-2026.2364936

Validation of the effectiveness of various methods through experimental comparisons can point out the limitations and improvement space of the current system. Finally, the application potential of intelligent grasping technology in waste sorting and its role in promoting environmental protection and the circular economy will be discussed. The structure of this paper is as follows: Chapter 2 will elaborate on.

2. Key Technologies of Robot Grasping Systems

In the complex application scenario of garbage sorting, the realization of an intelligent robot grasping system requires the synergistic support of multiple key technologies. Currently, visual perception and localization, tactile perception and soft sensors, shared control and human-machine collaboration, as well as soft gripper design, constitute the four technical pillars in this field, collectively enabling solutions to grasping challenges in unstructured environments.

2.1 Visual Perception and Localization Technology

The vision system is a critical step in robot environmental perception, undertaking the task of target recognition and pose localization in garbage sorting [5]. In practical applications, most systems choose RGB-D cameras as the primary sensor, providing a complete data foundation for subsequent processing by simultaneously acquiring color images and depth information.

At the algorithmic level, traditional recognition and localization methods often face challenges such as model library limitations and complex scenes. With the rise of deep learning, researchers have begun exploring approaches that do not rely on pre-existing knowledge bases and generate solutions directly from sensor data. For example, in 6-degree-of-freedom (6-DoF) grasp pose generation, Gui et al. proposed a fast detection algorithm [6]. They designed a point cloud preprocessing method to reduce the input size for grasp pose detection. Xie et al. proposed a universal framework[7], FlexLoG, optimizing the local grasping model and improving efficiency. Sundermeyer et al. proposed Contact-GraspNet, transforming the complex 6-DoF pose estimation problem into classification and rotation prediction of potential contact points in the observed point cloud, greatly simplifying the learning process [8]. It achieved an over 90% grasping success rate for unknown objects in cluttered scenes, demonstrating the potential of end-to-end generation methods in handling unstructured objects(Fig.1).

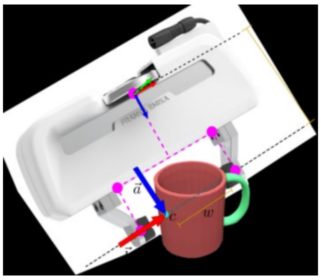

Fig. 1. The method of representing the grab posture used in Contact-GraspNet c depicts an observed contact point. A and b constitute the 3-DoF rotation, w is the predicted grasp width, and d is the distance from baseline to base frame[8]

Meanwhile, in six-degree-of-freedom object pose estimation, researchers are also working to make the process simpler [9]. A single forward-pass end-to-end network has been proposed. It combines object detection with a differentiable PnP (BP-n-P) solving layer. This network directly outputs an object's 6-DoF pose from a single RGB image. It removes complex multi-stage post-processing. It offers a new method for fast pose estimation.

These technological advances show that visual perception is moving in a more direct, flexible, and efficient direction. The goal is to rely less on precise models and complex reconstruction. This helps adapt to the high uncertainty and real-time needs in garbage sorting scenarios.

2.2 Tactile Perception and Soft Sensor Technology

In the densely populated situations, where the eye fails to create an image, there is a task that tactile sense cannot perform. There have been studies based on the concept that grasping success rates can be significantly enhanced when tactile sensors are added when a robot is in a soft grasp position. This is particularly so when dealing with light and free floating targets. The tactile feedback can successfully prevent the rebellious nature of targets who will be shoved away using undue contact forces [10]. Robot recognition and grasping have been on the increase over the recent years. Soft sensors have been the focus of researchers in an attempt to make them work better. They have attained great sensitivity, large detection range, quick dynamic response among other properties [11, 12]. One design supports elastic sensory components such as optical fibers in soft fingers. This design retains the flexible nature of the gripper. It is also capable of providing low-cost scalable solutions to tactile sensing.

Such tactile systems have the ability in real-time to measure the distribution of contact forces and slip conditions in real-time during grasping in terms of function. This itself forms a direct control source of grasp force. Slip can be detected by high-sensitivity sensors. This enables the grasp to be adjusted in order to avoid the dropping of items. It develops steady grasping without receiving any harm due to excessive forces [13]. What is more, recording the touch patterns, the device is also able to detect the surface features of objects. These are the hardness, texture, shape, roughness, and other physical property [14, 15, 16]. This enhances recognitions. It facilitates making different types of garbage adhere to classification [17].

Experience reveals that the combination of visual and tactile information increases the knowledge of the system concerning the grasped objects. It plays a major role in enhancing the accuracy and precision of recognizing the object and distinguishing between different types of garbage [9]. It is a multimodal sense model that prevents blind grasping by the robots. Rather, they have the ability to make changes in the strategies towards dynamic adjustment to real-time sensory feedback. This forms a significant enhancement of the flexibility of the system in complicated settings.

2.3 Shared Control and Human-Machine Collaborative Architecture

Since the environment of garbage sorting is full of uncertainty, full autonomy grasping is likely to struggle with stability. In a bid to manage this, there has emerged shared control architectures. They combine the ability of a human operator to make a judgment with the execution ability of a robot. This creates complimentary strengths.

A standard shared controller provides operators with indication by force feedback control or virtual fixtures. The robot will intervene when the operator tries to execute an action that it perceives to be at odds with the optimal strategy to grasp. It employs appropriate visual feedback or force. This would guarantee that the grasping process is guided by the intention of the operator but also in respect of mechanical logic. This layout is particularly suitable with delicate products or complicated packing situations. In this case, the operator deals with macro-level decisions. The micro-level stability is maintained through the robot.

It is essential to mention that even more developed systems of shared control may also vary their level of assistance in question dynamically. This is done depending on the expertise of the operator and the complexity of the task. They give freedom to more skilled operators. They provide tighter guidance to the beginners [18]. This adaptive measure will enable this system to balance operational efficiency and safety. It presents positive flexibility in practice.

In recent years, control strategies using deep reinforcement learning, for example the Soft Actor-Critic algorithm, have shown strong potential in unstructured grasping tasks. They allow adaptive grasping strategies through simulation training. This happens without human help [8,10].

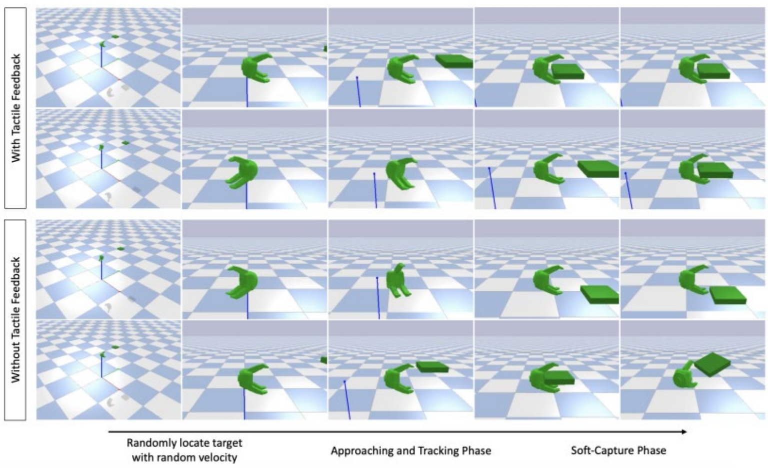

Fig. 2. Snapshot comparison of the grasping process for the two types of agents in simulation[10]

As shown in the Fig.2[10], the agent with tactile feedback demonstrates higher cumulative reward and success rate in the soft capture phase, validating the key role of tactile information in dynamic grasping.

2.4 Soft Gripper Design and Optimization

The nature of soft grippers is that of natural flexibility. This provides them with privileges in garbage sorting cases. According to Chen et al., there is a difference between old rigid grippers and the difference explained [19]. Soft grippers contact the objects by deforming their entire body or wrapping over the surface. This facilitates understanding more clearly. It allows them to manipulate objects on different shapes and sizes. It also renders control to be a lot less complicated.

According to research, there is dynamic grasp advantage. Fitting soft capture techniques reduce the impact force of a contact. They achieve this by dedicating more time in moving with the target. This prevents the target being chased off by some abrupt contact. This fact is much significant in picking up delicate or light items [20, 21].

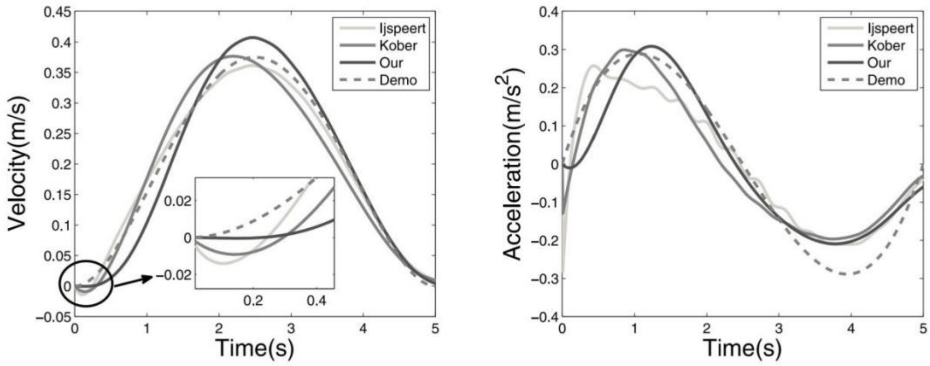

Fig.3. Comparison of velocities and accelerations generated by different DMP formulations[20]

Fig.3 shows this [20]. The soft capture method greatly lowers contact impact forces. It achieves this by extending the time of synchronized tracking.

Looking at structural design, common multi-finger soft grippers today use two main methods. One method uses air pressure. The other uses tendons. The air pressure method works by changing the air pressure inside to bend the fingers. This meets needs for flexible grasping [22,23]. The tendon method creates bending by using motors to pull flexible cables. It provides quick reaction and precise control. Both approaches have their own strong points. The choice depends on the exact situation where they will be used.

When talking about making them better, researchers keep working on soft gripper design. They use finite element analysis and real-world tests. How often a grab succeeds, how well it handles different objects, and how long it lasts are now main measures of gripper quality. Tests show that improved soft grippers work well with many waste items. They can grab stiff cans and also floppy plastic bags. They keep a steady hold while causing very little harm to the objects.

To sum up, vision, touch, control, and the gripper itself are four key technical areas. Each has its own job, but they all connect closely. Making their roles and links clear involves a comparison. The main points are looked at from four sides: seeing, feeling, how control is set up, and the physical hardware that does the work(Table 1).

Table 1: Comparison of Core Technologies for Robot Grasping in Garbage Sorting

Technical Pillar	**Representative Methods/Types**	**Core Contribution**	**Current Main Limitations**	**Role in Integrated System**
Visual Perception	Deep Learning Grasp Detection, 6D Pose Estimation	Provides target localization capability in unstructured scenes	Sensitive to occlusion and transparent objects; lacks physical interaction information	Global perception vanguard, providing an action basis for the system
Tactile Perception	Flexible Optical, Capacitive Sensors	Provides real-time feedback on physical attributes such as contact force, slip, and texture	Susceptible to interference; complex algorithms for information interpretation and fusion	Core of local interaction, enabling fine force control and stable grasping
Control Architecture	Adaptive Shared Control	Dynamically fuses human and machine intelligence, balancing operational efficiency and safety.	May limit operator autonomy; system latency requires strict control	Decision-making hub, coordinating perception and execution to handle uncertainty
Execution Hardware	Adaptive Soft Grippers	Forms stable enveloping contact with objects through compliant	Difficulty grasping specific objects (e.g., sharp, sticky);	Physical interaction terminal, translating control commands into

Technical Pillar	Representative Methods/Types	Core Contribution	Current Main Limitations	Role in Integrated System
		structures	durability needs further verification	robust actions

Overall, robot grasping technology is developing towards multimodal perception fusion, human-machine intelligence collaboration, and flexible design. The vision system provides global environmental awareness, the tactile system enables fine local feedback, the shared control architecture achieves complementary human-machine advantages, and the soft gripper ensures adaptability in physical interaction. The continuous evolution and system integration of these key technologies will promote garbage sorting robots to better cope with various challenges in real-world environments, achieving the leap from laboratory research to practical application.

3. Integrated Architecture of Intelligent Grasping Systems

In messy places like waste sorting, one technology alone usually cannot get a stable and efficient grasp. Because of this, recent work has slowly moved towards system designs that combine several technologies. This chapter will look at the architecture design of smart grasping systems for waste sorting. The focus is on three parts: mixing different senses, control methods, and hardware-software teamwork. It will also look at how these technologies are combined and the real-world use possibilities of such systems.

3.1 Multimodal Perception Fusion Framework

Single types of sensing have limits. To deal with this, common system designs usually build a sensing framework that works with both vision and touch. The vision system uses RGB-D cameras. Its job is early target recognition and finding where things are. It keeps high recognition rates even when objects are piled up in complex ways. The touch system uses optical fiber sensors placed inside soft fingers. It watches mechanical changes during grasping in real time.

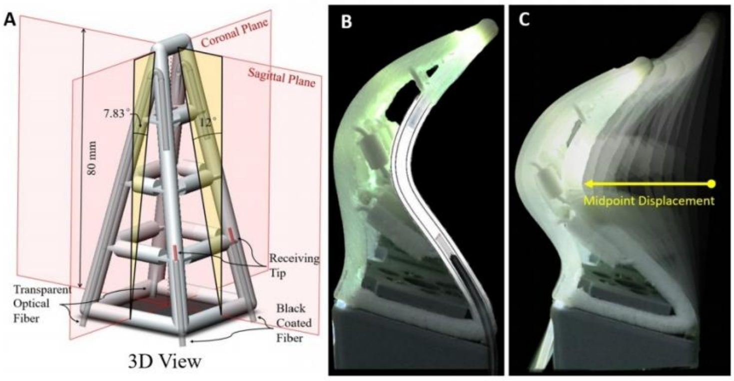

Fig.4 Soft finger and sensor design. The Sagittal plane and the coronal plane are the main bending planes (A). The soft optical fiber passes through the inner tube to transmit the light, which will be received by a black-coated optical fiber (A). The tube wall will hinder light transmission while bending (B), which relates to the midpoint displacement (C)[10]

The teamwork between the two shows up in data helping each other on different levels(Fig.4). Vision gives big-picture location information. Touch takes care of small, local fine-tuning. When vision cannot work because something is in the way, the system changes on its own to a control mode led by touch. This makes a smooth changeover. Tests show this mixing method makes grasping success rates go up by 18.3%. It works especially well for grabbing see-through objects and shiny surfaces [10].

3.2 Adaptive Shared Control Strategy

In shared control system design, changing things dynamically is now a common plan. This change is based on how skilled the operator is and how hard the task is. The system looks at past operation data. It builds a model of what the user can do. Then it changes how much help it gives in real time. For new users, the system gives strong guidance using virtual fixtures. For experienced users, it steps in less to keep their freedom to operate. At the task level, the system picks control modes by itself. This choice is based on object features like shape and material [18]. Normal objects get handled with self-sufficient grasping. Fragile or complicated objects turn on shared control. This layered plan makes mistakes in operation drop by 27.6%. It keeps speed high while also making the operator's mental work much less.

3.3 Software-Hardware Co-optimization

At the hardware level, a modular design is used. Soft grippers allow for fast replacement. Touch sensors are built right inside the fingers. This makes sensing and doing one combined action. This design does two things. It makes the system more reliable. It also cuts maintenance costs a lot . The software setup has three layers. A system diagram shows this. The bottom layer handles real-time control. The middle layer deals with mixing data from many sources. The application layer gives a visual interface. Each layer talks through standard interfaces. This keeps the system easy to expand. Real-world tests show this setup keeps system response time under 150ms. This meets the needs for real-time operation .

4. Challenges and Future Directions

Current research has moved forward a lot. Even so, smart robot grasping systems for garbage sorting have many problems left. Future work will need to look into several paths.

One issue is at the sensing level. Vision systems still have problems like heavy blocking and trouble seeing clear objects. Touch sensors have their own issues. Calibration mistakes and noise from the environment are technical problems that need fixing. Another issue is at the control level. Shared control systems make work faster. But force feedback can get in the operator's way. Mixing different types of information needs a lot of computer power. This can sometimes cause delays. A third issue involves the grippers themselves. Soft grippers adapt well. But they still have a hard time grabbing things that are very soft, sticky, or sharp. How long they last also needs more checking.

Considering the deficiency in the present work, they believe that future research may take a few directions. One of them is to achieve the understanding of strategies that adjust with the reinforcement learning. This allows robots to learn how to grasp the most optimally as they interact with the environment. The other method is by examining systems with more than two robots to sort. The entire sorting of machines might work as a team. The third option is the creation of a hybrid framework of real-time sensing and control. This would make use of AI and edge computing to reduce system delays. The fourth option is to develop grippers with long lives in highly challenging locations. Wet or corrosive environment is an example. This would ensure the systems are more reliable. Lastly, they might seek alternatives of people and machines interactions that are more natural. Examples are mixed reality displays and hand gesture interfaces. This would facilitate the training process and minimize the operator fatigue.

5. Conclusion

The system design which was examined in this chapter constructs a good and reliable grasping system model. It achieves this by integrated new concepts in a variety of sensing, adaptive control and software-hardware synergy. This design does two things. It addresses the key technical issues of real-life applications. Its modular construction also leaves room to improve on its designs. This provides a good assistance in applying smart sorting technology within the industry.

Present paper thoroughly studies the smart robot grasping systems to sort garbage. It demonstrates the full advantages of collaborating. These include conglomeration of sensory abilities, shared control of adaptation, and soft grippers. To combine senses, it is better to mix sight and touch to make the system stronger. It assists the system to adapt more suitably in disorganized environments. When it comes to control design, the shared control design is best as both machines and people receive the best thereof. It makes work safe and at the same time ensures that jobs are faster accomplished. Soft grippers naturally are flexible insofar as software-hardware teamwork is concerned. They deal with strange-shaped objects at ease. This further helps in making the system useful even in unexpected locations.

This mixed system is proved to be much easier to understand success and work speed. There is one point that is emphasized in this paper. The combination of the systems and the new thinking which is joint is highly important in actual application. The point is this. An effective approach is presented to implement smart grasping systems with effectiveness in real conditions. This occurs in profound collaboration among sensing, decision making and action stages.

Looking into the future, there is a number of technologies which continue to increase. These involve reinforcement, multi-robot learning and edge learning. The emerging smart grasping systems can probably be widely used in numerous other domains. These include environmental care, health services and transportation of goods. They will provide significant technological assistance to the development that can withstand.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

Reference

[1] Kaza, S.; Yao, L.; Bhada-Tata, P.; Van Woerden, F.; Ionkova, K.: What a waste 2.0: A global snapshot of solid waste management to 2050. World Bank Group, Washington DC (2018)

[2] Abdollahi Saadatlu, E.; Barzinpour, F.; Yaghoubi, S.: A sustainable model for municipal solid waste system considering global warming potential impact: A case study. Computers & Industrial Engineering. p. 108127 (2022)

[3] Kumar, B.N.; Patel, M.; Kumar, K.S.; Kushwaha, P.: Advancing sustainable electronic waste management: An overview of mechatronics solutions for health, environment, and recycling. Chemosphere. p. 144501 (2025)

[4] Kurniawan, T.A.; Meidiana, C.; Othman, D.; Goh, H.H.; Chew, K.W.: Strengthening waste recycling industry in Malang (Indonesia): Lessons from waste management in the era of Industry 4.0. Journal of Cleaner Production. p. 135296 (2023)

[5] Tong, L.; Qian, K.; Zhou, B.; Fang, F.: A lightweight prior encoding–decoding cascade framework for robust depth completion in robotic grasping of transparent objects using RGB-D sensors. Engineering Applications of Artificial Intelligence. p. 112934 (2026)

[6] Gui, H.; Pang, S.; He, X.; Wang, L.; Zhai, X.; Yu, S.; Zhang, K.: GraspFast: Multistage lightweight 6-DoF grasp pose fast detection with RGB-D image. Pattern Recognition. p. 111318 (2025)

[7] Xie, P.; Chen, S.; Tang, W.; Yang, K.; Wang, G.: Rethinking 6-DoF grasp detection: A flexible framework for high-quality grasping. Pattern Recognition. p. 112088 (2026)

[8] Sundermeyer, M.; Mousavian, A.; Triebel, R.; Fox, D.: Contact-GraspNet: Efficient 6-DoF grasp generation in cluttered scenes. arXiv preprint (2021)

[9] Liu, F.; Shao, R.; Zhao, Y.; Li, C.; Wang, Y.: Multimodal tactile sensing fused with vision for object recognition. Chemical Engineering Journal. p. 169249 (2025)

[10] Beigomi, B.; Zhu, Z.H.: Improving SoftCapture phase success in space debris removal missions: Leveraging deep reinforcement learning and tactile feedback. Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE). pp. 1554–1561 (2024)

[11] Ali, A.R.; Selim, M.Y.: Design and integration of a multiaxial tactile sensor for dexterous manipulation by humanoid robots for industrial applications. Results in Engineering. p. 108494 (2026)

[12] Liang, T.; Liu, Z.; Zhang, H.; Zhou, X.; Liang, Y.: The structure, material and performance of multifunctional tactile sensor and its application in robot field: A review. Materials Today. pp. 452–481 (2025)
[13] Yu, L.; Xiao, W.; Wang, Q.; Liu, D.: Soft microtubular sensors as artificial fingerprints for incipient slip detection. Measurement. p. 117729 (2025)
[14] Wang, S.; Wang, X.; Gao, Z.; Zhang, Y.; Wang, L.; Liu, Y.; Liu, J.; Rong, L.; Peng, Z.; Ouyang, Q.: A smart finger for robotic tactile sensing of surface/subsurface patterns based on a high-density piezoresistive sensor array. Biosensors and Bioelectronics. p. 117858 (2025)
[15] Marzani, M.; Khatibi, S.; Masinjila, R.; Prado, A.; Alves, J.: A dataset for tactile textures on uneven surfaces collected using a BioInTacto sensing module. Data in Brief. p. 111312 (2025)
[16] Du, X.T.; Yang, Y.L.; Yang, Q.C.; Ma, K.X.; Zhang, Q.Z.: Multiparameter tactile sensing and recognition for robotic finger based on BiFBG. Optics & Laser Technology. p. 113634 (2025)
[17] Zhou, Y.; Zhao, X.; Xu, J.; Chen, G.; Tat, T.; Li, J.; Chen, J.: A multimodal magnetoelastic artificial skin for underwater haptic sensing. Science Advances. p. eadj8567 (2024)
[18] Lie, W.; Aing, L.: Efficient 6-DoF pose estimation for multi-instance objects from a single image. Image and Vision Computing. p. 105882 (2026)
[19] Chen, H.; Li, Y.; Xu, P.; Li, J.; Noor, A.; Zhou, X.; He, W.; Wang, T.; Mou, Z.; Song, L.; Xu, M.: Octopus-inspired soft gripper with embedded triboelectric tactile sensor for underwater target recognition and grasp. Nano Energy. p. 111007 (2025)
[20] Yu, M.; Luo, J.; Wang, M.; Li, J.; Liu, C.; Sun, J.: Directly skill-transferred DMP-based trajectory planning for robot soft capture of space tumbling targets. Acta Astronautica. pp. 611–629 (2023)
[21] Fu, H.; Xu, D.; Wu, J.: Robotic arm intelligent grasping system for garbage recycling. Proceedings of the China Automation Congress (CAC). pp. 6821–6826 (2021)
[22] Sun, T.; Chen, Y.; Han, T.; Jiao, C.; Lian, B.; Song, Y.: A soft gripper with variable stiffness inspired by pangolin scales, toothed pneumatic actuator and autonomous controller. Robotics and Computer-Integrated Manufacturing. p. 101848 (2020)
[23] Liu, X.; Song, Z.; Xu, L.: Bioinspired soft gripper with SMA-based variable stiffness. Mechatronics. p. 103417 (2025)