

**The 11th Asian-Australasian Conference on Precision Agriculture (ACPA 11)
October 14-16, 2025, Chiayi, Taiwan**

Six-Axis Robotic Arm and Object 3D Detection Technique for Supporting Mobility-Limited People on Grasping Objects

Din-Yuen Chan¹, Chao-Min Hsu¹, Joseph E. Chou^{1*}

¹Department of Computer Science and Information Engineering, National Chiayi University,
Chiayi City, Taiwan

*Corresponding Author: s1110311@mail.ncyu.edu.tw

ABSTRACT

This paper presents the integration of a six-axis semi-industrial robotic arm with a real-time 3D object detection system to enable intuitive, contactless human-robot interaction, with a particular focus on healthcare applications. The robotic arm, powered by high-precision stepper motors, delivers enhanced accuracy and reliability compared to traditional servo-based systems, making it ideal for tasks that demand precision and consistency. The system is designed for extensibility, supporting secondary development through serial communication in languages such as C++, Java, and Python. This allows for flexible customization beyond the standard human-machine interface (HMI), enabling broader application adaptability. To facilitate precise control and interactive performance, the system incorporates a real-time 3D object detection approach based on a lightweight monocular depth estimation network (MonoNet-Lite) and a MobileNet-YOLO (M-YOLO) network, utilizing an RGB-D camera. These models work in tandem to determine the 3D location of target objects. As the robotic gripper approaches a target, MonoNet-Lite helps navigate around surface protrusions to ensure a stable, top-down grasp using a bird's-eye view (BEV) perspective. This contactless control method is particularly advantageous in clinical settings, such as sterile environments, rehabilitation centers, and assistive care facilities, where minimizing physical contact is essential. The integration of artificial intelligence, computer vision, and robotics in this system demonstrates strong potential for intelligent automation in healthcare, supporting applications such as remote patient care, interactive therapy, and assistive robotic support.

Keywords: MobileNet, YOLO, Monocular depth estimation, Object 3D detection, Six-axis robotic arm

INTRODUCTION

The rapid advancement of AI and robotics (Lee and An, 2022) is transforming healthcare, requiring precise, safe, and adaptable systems. This study presents a robotic arm integrated with an RGB camera and MobileNet-YOLO for real-time object recognition and 3D localization. A bird's-eye view (BEV) grasping strategy enables stable top-down manipulation, demonstrating the potential of combining AI, computer vision, and robotics for healthcare automation in rehabilitation, assistive care, and remote patient support.

MATERIALS AND METHODS

Figure 1 present the overall system architecture and workflow of the proposed robotic arm with

a three-finger gripper. The system incorporates an NPU-embedded notebook and an NVIDIA Jetson Orin Nano as the primary computing platforms. Visual input is acquired through a webcam and an Intel RealSense D-435 depth camera, which simultaneously captures RGB images and depth maps. The RGB images are processed using Reduced MobileNet-SSD v2 (Chiu et al., 2020) or YOLOv7 (Wang et al., 2023) to perform real-time object recognition, generating bounding box coordinates and class information. By aligning these bounding boxes with the corresponding depth maps, the system estimates the spatial position and depth of the detected objects. The resulting object categories, positions, and depth information are then transmitted to the robotic arm, enabling top-down grasping and precise manipulation. This architecture ensures efficient and accurate object handling, supporting contactless assistance and healthcare applications.

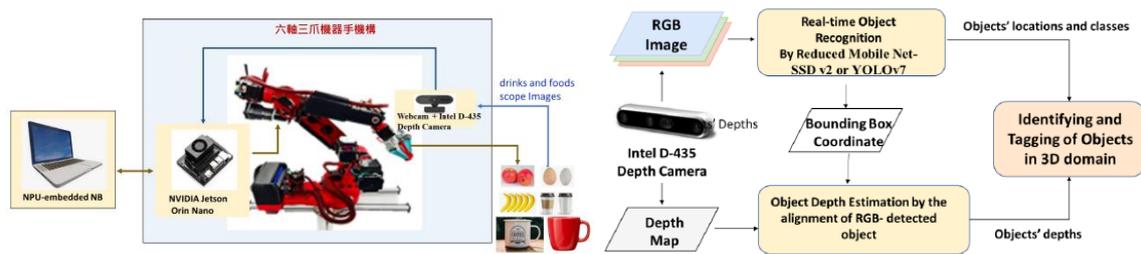


Fig.1 (a) System configuration (b) Process of Target Object Identification and 3D Positioning.

RESULTS & DISCUSSION

The robotic arm system was evaluated in healthcare-focused scenarios (Chan et al., 2025) for real-time object detection, contactless interaction, and precise grasping. Results show that RGB-based vision with AI detection provides an efficient and cost-effective solution, though performance may decline in low-light or reflective conditions. Future improvements include adaptive sensing and multi-object coordination to enhance robustness and clinical applicability.

CONCLUSIONS

This work integrates a six-axis robotic arm with real-time object detection for precise, contactless human-robot interaction in healthcare. Experiments demonstrate stable manipulation and highlight the potential of robotics and vision for intelligent, patient-centered automation.

REFERENCES

- Chan, D. Y., J. F. Wang, C. I. Chang, and J. W. Du. 2025. A low-cost autonomous outdoor robot with stabilized controller and deep-learning integrated GPS navigator under end-to-end implementation. *Journal of the Chinese Institute of Engineers*. 1-15.
- Chiu, Y. C., C. Y. Tsai, M. D. Ruan, G. Y. Shen, and T. T. Lee. 2020. Mobilenet-SSDv2: An improved object detection model for embedded systems. In *2020 International conference on system science and engineering (ICSSE)*. 1-5.
- Lee, C., and D. An. 2022. AI-based posture control algorithm for a 7-DOF robot manipulator. *Machines*. 10(8): 651.
- Wang, C. Y., A. Bochkovskiy, and H. Y. M. Liao. 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 7464-7475.