International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 12, Issue 9 [September. 2023] PP: 95-100

Application of tomato picking robots based on deep learning detection methods in plant factories

Wu Zhenwei, Zhang Mengting, Wang xinfa
School of Information Engineering, Henan Institute of Science and Technology, Henan, China
Corresponding Author: Wang xinfa.

ABSTRACT: *Agricultural factories represent an advanced form of controlled-environment farming, well-suited for mechanization due to their fully enclosed and controllable environments. In the context of agricultural factories, the application of robots is expected to be a growing trend. Therefore, a tomato harvesting robot designed for use in agricultural factories has been developed. This robot is capable of patrolling between plant racks within the agricultural factory, autonomously identifying ripe tomatoes, and completing the harvesting and collection process. The design features an ARM processor as the central controller, a tracked multi-directional mobile platform serving as the robot's base, and an RGB camera controlled by the ARM processor for tomato recognition. The camera captures images, which are then processed by a deep learning-based detection network to obtain the tomato's position. Mounted on the base is a 6-degree-of-freedom robotic arm equipped with a flexible gripper. The mechanical arm is scalable, allowing for adjustments in grasping distance and height. The ARM processor controls the camera for tomato identification and capture. The coordinates of ripe tomatoes are transmitted to the ARM controller via an API port. The ARM controller uses inverse kinematic analysis to drive the robotic arm and gripper in coordinated motion, including movements of the arm's end joint, thereby achieving tomato harvesting.*
Keywords: *Artificial Intelligence; Deep Learning; Tomato Harvesting Robot; Robot Design; Agricultural Machinery; Detection Algorithm.*

Date of Submission: 06-09-2023 Date of acceptance: 19-09-2023

I. INTRODUCTION

Plant factories represent an advanced form of facility agriculture and are characterized by intensive labor requirements. Even in highly automated facility agriculture operations in developed countries, crop harvesting still heavily relies on manual labor. Manual harvesting is time-consuming and labor-intensive, making the pursuit of fully automated harvesting a primary requirement in the current development of facility agriculture, especially within the context of plant factory environments [1].

As electronic information technology, artificial intelligence, image recognition technology, and robot manufacturing control technology have matured, countries worldwide have been conducting research on harvesting robots since the 1970s and 1980s. These robots are designed for the harvesting of various crops, including apples, strawberries, cabbages, cucumbers, tomatoes, and more [2].

However, research on agricultural harvesting robots in China started relatively late and is still in its early stages. Wang Shunwei et al. [3] designed a harvesting mechanical arm that uses vibrations to detach fruit stems and collects fruits using an umbrella-like mechanism. However, this mechanical arm is prone to missing fruit due to the flexibility of fruit stems, and it cannot avoid collisions between fruits during the collection process, which significantly impacts harvesting coverage and fruit quality.

Shao Kun [4] extended the mechanical arm of a harvesting robot to five degrees of freedom with a deviation of only 6.71 mm. However, the tracked chassis used in the robot leads to slow travel speeds and requires a significant amount of turning space, making it unsuitable for operation inside sunlight greenhouses. Furthermore, Tang Yadong [5] constructed a prototype tomato harvesting robot using binocular vision and robotic arm motion simulation. It completed tomato harvesting tests in complex environments but achieved only a 76.3% success rate due to the rigid gripper used.

Wang Xiaonan et al. implemented damage-free tomato harvesting by utilizing vacuum suction devices, flexible bladder devices, and twisting motors. However, interference from branches and leaves can disrupt the recognition system, resulting in reduced harvesting accuracy. Liu Fang et al. [6] employed an improved multi-scale YOLO algorithm for tomato training and recognition under various environmental conditions. Zhu Mingxiu [7] used K-means clustering, along with convolutional neural networks and binocular vision technology, to detect and locate fruits for harvesting robots. Hu Huiming [8] used binocular vision technology to obtain three-dimensional coordinates for fruits and vegetables, laying the foundation for fruit and vegetable robot harvesting in greenhouse environments.

In this study, we focus on tomatoes grown in artificial light plant factories and design a mobile robot capable of inspecting and identifying ripe tomato fruits, suitable for deployment in artificial light plant factories.

II. Harvesting Robot System Functional Design

The harvesting robot system consists of four main components: the chassis four-wheel drive system, harvesting system, and image recognition system [9].

The chassis drive system includes four driving wheels with their respective motors, chassis suspension structure, and bottom line-sensing camera. The harvesting system comprises flexible grippers and a sliding-rail-type mechanical arm. The liftable upper platform system includes drive motors, a mechanical arm with an image recognition system, mounting brackets, and more.

Currently, robots in artificial light plant factories operate in various modes, including ground-based free-moving, suspended rail-based movement, and ground rail-based movement [10]. Given the need for direct sunlight utilization in artificial light plant factories, installing suspended rails would occupy a significant portion of the upper space of the planting racks, significantly reducing light penetration and energy efficiency. On the other hand, installing ground rails would occupy a considerable area inside the greenhouse, affecting the layout of plant racks. Therefore, this study employs a four-wheel track-based ground rail mobile robot operation method. The four-wheel track platform is more flexible than traditional steering mechanisms, as it can move in any direction by controlling the speed and direction of the two-wheel sets. It can complete harvesting tasks without changing its own state in narrow greenhouses.

Tomatoes in artificial light plant factories are primarily grown in shelf and bed cultivation modes, with columnar distribution and vertical growth. Tomato fruits grow vertically on the plants. Therefore, the design of the chassis takes into account the factors mentioned above and uses a pendulum-type suspension structure. This structure enhances the robot's terrain adaptability, ensures the stability of the chassis during operation, and improves driving stability and precision [11]. The tomato cultivation mode in artificial light plant factories is shown in Fig 1.

Fig. 1 Tomato Cultivation Mode in Plant Factories

When the harvesting robot is in operation, the mechanical claw's role is to grip the fruit for stem separation. Considering the mechanical properties of ripe tomato fruit skins, the robot's harvesting mechanism employs flexible grippers made of rubber material using injection molding [12]. These flexible grippers use a two-fingered gripping approach driven by motors. The gripping action part of the gripper is 40 mm long, and it can grasp diameters ranging from 10 to 90 mm. Following an analysis and discussion of research on the separation force between tomato fruit and stem conducted by Huang Guowei and others, a harvesting scheme involving twisting and picking was determined. An end joint capable of continuous 180° rotation was designed. With this design, during the harvesting action, the mechanical hand can penetrate into the tomato plant at 0°. After stable gripping, the end joint rapidly rotates 180°, causing the tomato stem to move away from the stem in a direction 90° away, using minimal force to separate the fruit from the stem. The harvesting claw schematic is shown in Fig 2.

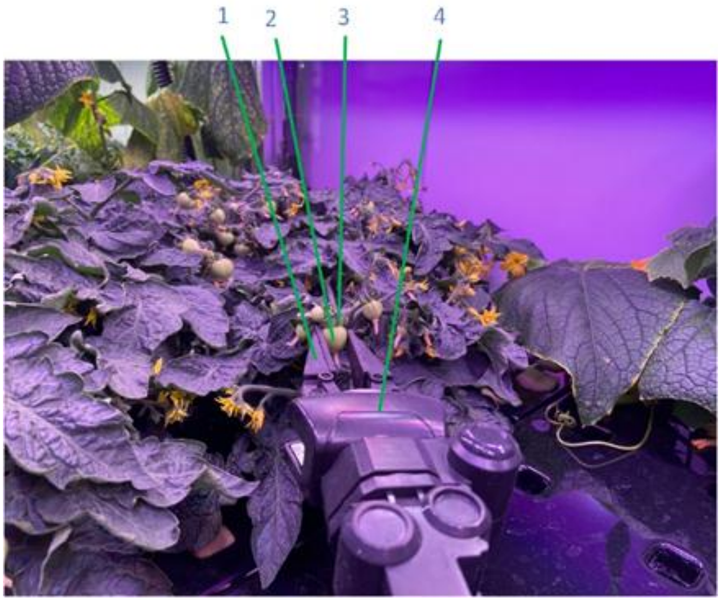

Fig. 2 Harvesting Claw Schematic

To account for the fact that tomatoes are not at a fixed height and may have different depths of growth relative to the robot, a lifting structure is used. The lifting structure is controlled by an electric motor, allowing the robot to harvest tomatoes at different heights. Additionally, considering that tomato plants may have different depths of growth (the horizontal distance between the tomato fruit and the centerline of the ridge), a retractable arm on the robot's upper platform is employed [13]. The controller, using sensors such as an RGB camera, obtains tomato data and calculates the depth of tomatoes relative to the screw platform's center position. The arm movement is adjusted by feedback from the screw-drive motor and ranging sensor to position the gripper accurately for fruit picking.

III. Harvesting Robot Software Design

The hand-eye coordination is a critical aspect that embodies the autonomous harvesting operation of fruit and vegetable harvesting robots. Based on the camera's installation location, the robot's vision system can be classified into Eye to Hand and Eye in Hand types, where Eye to Hand involves mounting the camera outside the robot body and fixing it in place, while Eye in Hand involves fixing the camera on the robot's arm, moving it together with the robot arm [14]. Height limitations in greenhouse environments, the vertical growth direction of tomato plants, and the relatively small size of tomato fruits have made Eye to Hand operation less suitable for greenhouse operations. Considering Eye in Hand, where the robot arm is in motion, real-time changes in the camera's position can be obtained, resulting in more accurate system calculations. Additionally, microcontrollers like ARM can meet the real-time control requirements of the system to some extent. Therefore, it was decided to place the camera on the robot's arm and keep it fixed in place.

An Intel RealSense D455 depth camera is used in conjunction with an ARM controller to capture images. Tomato recognition primarily consists of object detection based on RGB images and image data processing [15]. RGB images are processed using the YOLOv5 object detection network, which calculates the rectangular bounding boxes for tomatoes. Depth information is then fused with the two-dimensional information generated by object detection. Finally, the 3D center position of the tomatoes is computed. The overall algorithm flowchart is shown in Fig 3.

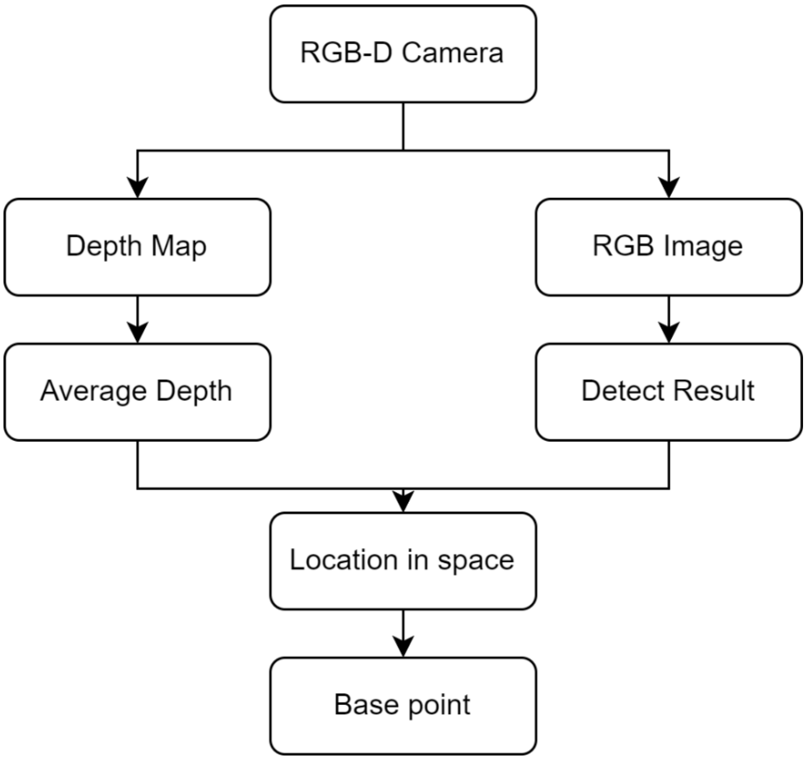

Fig. 3 Image Algorithm Flowchart

YOLOv5 is a high-precision, high-speed single-stage object detection algorithm [16]. It is the latest achievement in the YOLO series, significantly improving both speed and accuracy. The YOLOv5 network model consists of four main parts: the input, the backbone network, the neck network, and the prediction [17]. The input part processes the input image data and annotation data to adapt the image to the network size, using normalization, mosaic data augmentation, adaptive anchor calculation, and adaptive image scaling. The backbone network specializes in feature extraction, and its quality directly affects the accuracy of subsequent work. YOLOv5 introduces the Focus structure for the first time, with a key operation being the slice operation. It also uses the CSPDarknet53 feature extraction network. The neck network includes the SPP module and FPN+PAN module. The SPP module uses multi-scale max pooling to extract features at different scales for multi-scale fusion [18-20]. The FPN+PAN module enhances both expressive and positional features by fusing different detection layers from different backbone layers.

The Prediction module is used for final target classification, detection, and positioning output. It employs the GIOU_Loss function and DIOU_nms function [21]. The GIOU loss function solves the problem of distinguishing between non-intersecting and intersecting prediction boxes and different intersecting situations when the intersection ratio is the same. The DIOU_nms function further considers the bounding box and center position information when performing non-maximum suppression, optimizing the prediction output.

The process begins with labeling previously captured tomato images and dividing them into training and testing sets in a 4:1 ratio. These images are then used to train the YOLO object detection network, which is later converted into the NCNN (a lightweight neural network framework for mobile devices) framework. The network is deployed on the ARM controller. Images are obtained using a stereo infrared RGB depth camera, creating a point cloud map based on the camera's coordinate system. The three-dimensional point cloud information is filtered through algorithms like multi-plane segmentation and clustering [22]. The three-dimensional point cloud information is then projected onto two-dimensional coordinates and matched with the two-dimensional information obtained from the detected tomato keyframes. This process results in the three-dimensional bounding box for the target tomato, allowing the calculation of the tomato's center point. The tomato recognition results are shown in Fig 4.

Fig. 4 Tomato Recognition Results

Traditional harvesting robots with multi-degree-of-freedom mechanical arm motion use methods like differential interpolation and coordinate inverse kinematic analysis. These methods involve sensors, such as cameras, that provide information about the target's coordinates and orientation angles. The controller calculates the joint angles based on the collected data and generates control signals using pulse width modulation technology to drive the mechanical arm's rotation. This positions the end effector at the specified location to complete the harvesting task [23]. However, due to the relatively fixed growth patterns of crops like tomatoes, which grow vertically due to gravity, the end effector connected to joint DE always remains parallel to the ground, with an angle α of 0°. Therefore, in the inverse kinematic analysis process of the robot arm, α can be directly used as a known condition, significantly reducing the controller's computational load. Additionally, considering that the arm may be quite long, and joint B must withstand certain forces, the system fixed the angle $\beta 3$ between arm BC and the horizontal plane at 45°, without imposing restrictions on other joints. This meets the requirements after theoretical analysis.

IV. Experiments and Result Analysis

The experiments were conducted in May 2022 in the Artificial Light Plant Factory Laboratory at Henan University of Science and Technology. The robot prototype was placed in the gaps between planting trays, and the system was initiated. The robot correctly detected tomatoes and completed the harvesting. The experimental site is shown in Fig 5.

Fig. 5 Experimental Site

The tomato harvesting robot successfully completed the experimental operation. The control chip effectively controlled the robot arm to pick tomatoes and place them in the designated collection container. However, the designed tomato harvesting robot still has some errors and areas that need improvement. It should be noted that the full replication of the design was limited by the experimental conditions and equipment support.

REFRENCES

[1]. Li Zonggeng, Yang Qichang, Sha Dejian, Ma Wei. Research Progress on Full-Process Mechanization of Hydroponic Leafy Vegetable Production in Plant Factories. Chinese Journal of Agricultural University, 2022, 27(05): 12-21.

[2]. Zhang Wenxiang, Zhang Bingyuan, Gong Yu, Ren Ni. Research Status and Prospects of Robotic Arms for Fruit and Vegetable Harvesting. Transactions of the Chinese Society for Agricultural Machinery, 2022, 43(09): 232-237+244. DOI: 10.13733/j.jcam.issn.2095-5553.2022.09.031.

[3]. Wang Shunwei, Liu Hualei, Xu Tianyi, Jin Yifeng, Li Qianyin, Zhang Kai. Design of a Vibratory Fruit Picking Robot Arm. Science and Technology Innovation, 2020(16): 17-18.

[4]. Shao Kun. Tomato Intelligent Picking Robot with Five Degrees of Freedom. Jiangsu Agricultural Sciences, 2018, 46(21): 250-253. DOI: 10.15889/j.issn.1002-1302.2018.21.063.

[5]. Tang Yadong. Optimization Design and Experimental Research of Tomato Intelligent Picking Mechanism. Ph.D. Thesis, Henan Agricultural University, 2018.

[6]. Liu Fang, Liu Yukun, Lin Sen, Guo Wenzhong, Xu Fan, Zhang Bai. A Rapid Tomato Fruit Recognition Method in Complex Environments Based on Improved YOLO. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(06): 229-237.

[7]. Zhu Mingxiu. Research on Fruit Detection and Localization of Picking Robots Based on Image Processing and Convolutional Neural Networks. Agricultural Mechanization Research, 2022, 44(04): 49-53. DOI: 10.13427/j.cnki.njyi.2022.04.009.

[8]. Hu Huiming. Key Technology Research on Greenhouse Tomato Picking Based on Binocular Vision. Ph.D. Thesis, Hubei University of Technology, 2018.

[9]. Zhang Wenxiang, Zhang Bingyuan, Gong Yu, Ren Ni. Research Status and Prospects of Robotic Arms for Fruit and Vegetable Harvesting. Transactions of the Chinese Society for Agricultural Machinery, 2022, 43(09): 232-237+244. DOI: 10.13733/j.jcam.issn.2095-5553.2022.09.031.

[10]. Liu Jizhan, Li Pingping, Li Zhiguo. Hardware Design of Tomato Picking Robot End Effector. Journal of Agricultural Machinery, 2008(03): 109-112.

[11]. Wang Shanshan. Design of Wheeled Mobile Robot Control System. Master's Thesis, Nanjing University of Science and Technology, 2013.

[12]. Li Xianhua, Guo Yongcun, Song Tao. Kinematic Analysis and Simulation of Six-Degree-of-Freedom Industrial Robot Arm Forward Movement. Journal of Anhui University of Science and Technology (Natural Science Edition), 2013, 33(02): 34-38.

[13]. Zhai Jingmei, Dong Pengfei, Zhang Tie. Design of Vision-Guided Industrial Robot Positioning and Grasping System. Journal of Mechanical Design and Research, 2014, 30(05): 45-49. DOI: 10.13952/j.cnki.jofmdr.2014.0135.

[14]. Wang Junxiu, Kong Bin. Application of Computer Vision in Robot Target Localization. Microcomputer Development, 2003(12): 7-10.

[15]. [15] Du Xuedan, Cai Yinghao, Lu Tao, Wang Shuo, Yan Zhe. A Mechanical Arm Grasping Method Based on Deep Learning. Robot, 2017, 39(06): 820-828+837. DOI: 10.13973/j.cnki.robot.2017.0820.

[16]. Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 779-788. DOI: 10.48550/arXiv.1506.02640.

[17]. Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020. DOI: 10.48550/arXiv.2004.10934.

[18]. Yang J, Fu X, Hu Y, et al. PanNet: A deep network architecture for pan-sharpening. Proceedings of the IEEE international conference on computer vision, 2017: 5449-5457. DOI: 10.1109/ICCV.2017.193.

[19]. Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017: 2117-2125. DOI: 10.48550/arXiv.1612.03144.

[20]. Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2020: 390-391. DOI: 10.48550/arXiv.1911.11929.

[21]. Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the AAAI conference on artificial intelligence, 2020, 34(07): 12993-13000. DOI: 10.48550/arXiv.1911.08287.

[22]. Luo Xianbo, Zhong Yuexian, Li Renju. Data Registration Technology in Three-Dimensional Scanning System. Journal of Tsinghua University (Science and Technology), 2004(08): 1104-1106. DOI: 10.16511/j.cnki.qhdxxb.2004.08.026.

[23]. Jiang Hongchao, Liu Shirong, Zhang Botao. Inverse Kinematics Analysis of Six-Degree-of-Freedom Modular Robot Arm. Journal of Zhejiang University (Engineering Science Edition), 2010, 44(07): 1348-1354.