Received: 3 April 2023 | Revised: 21 August 2023 | Accepted: 28 October 2023

DOI: 10.1002/rob.22268

RESEARCH ARTICLE

WILEY

An automated apple harvesting robot—From system design to field evaluation

Kaixiang Zhang[1] | Kyle Lammers[2] | Pengyu Chu[2] | Zhaojian Li[1] | Renfu Lu[3]

[1]Department of Mechanical Engineering, Michigan State University, East Lansing, Michigan, USA

[2]Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, USA

[3]United States Department of Agriculture, Agricultural Research Service, East Lansing, Michigan, USA

Correspondence
Zhaojian Li, Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824, USA.
Email: lizhaoj1@egr.msu.edu

Funding information
U.S. Department of Agriculture-Agricultural Research Service; National Science Foundation, Grant/Award Number: ECCS-2024649

Abstract

Decreased availability and rising cost in labor poses a serious threat to the long-term profitability and sustainability of the apple industry in the United States and many other countries. Harvest automation is thus urgently needed. In this paper, we present the unified system design and field evaluation of a new apple harvesting robot. The robot is mainly composed of a specially designed perception component, a four-degree-of-freedom manipulator, an improved vacuum-based soft end-effector, and a dropping/catching component to receive and transport picked fruits. Software algorithms are developed to enable synergistic coordination of the hardware components for efficient, automated harvesting of apples in challenging orchard environments. Specifically, by integrating modified triangulation and image processing and analysis algorithms, a novel perception strategy is developed to achieve robust apple detection and precise localization. Improved planning and control algorithms are developed to guide the robot to the target positions. The performance of the robotic system was evaluated through field tests in two apple orchards with different tree architectures and foliage conditions. In the orchard where trees were young and well-pruned, the robot achieved 82.4% successful harvesting rate. In a second, older orchard with dense, clustered branches and foliage, the robot had 65.2% successful rate. The average cycle time to harvest a fruit was approximately 6 s, which included software algorithm processing and hardware execution. Moreover, through an in-depth analysis of the obtained results, limitations and planned future works are discussed.

KEYWORDS
agricultural robotics, apple, end-effector, harvesting, perception

1 | INTRODUCTION

With rising global demands, apple production has been steadily increasing and it has surpassed 80 million tons worldwide for the 2021/2022 harvest season (USDA, 2022), which needs a huge seasonal labor force to harvest. Efficient and timely harvesting of apples is critical for producing high-quality fruits to maximize profit for growers. It is estimated that in the United States alone, more than 10 million worker hours per year are needed for apple harvesting, which is mainly done by the seasonal, migrant labor force and accounts for about 15%

Kaixiang Zhang and Kyle Lammers contributed equally to this paper.

of the total production costs (Gallardo & Galinato, 2012). Like most horticultural sectors, the apple industry is facing severe labor shortage and rising labor costs, which raises significant concerns about its long-term sustainability and profitability. Meanwhile, the apple production systems have undergone profound changes in the two past decades; conventional unstructured orchards have been replaced by high-density orchards with smaller and structured trees (i.e., V-trellis, vertical fruiting wall, etc.). These modern tree structures would greatly facilitate orchard automation, and hence there is growing interest in exploring robotic harvesting as a possible alternative to reduce the industry's dependence on manual labor and harvesting costs.

Numerous robotic systems have been developed over the past decade to autonomously harvest a variety of horticultural crops, including kiwifruit (Williams et al., 2019, 2020), strawberry (Xiong et al., 2020), plum (Brown & Sukkarieh, 2021), citrus (Mehta et al., 2016), sweet pepper (Arad et al., 2020; Lehnert et al., 2017, 2020), and iceberg lettuce (Birrell et al., 2020). Comprehensive reviews of existing harvesting systems can be found in Bac et al. (2014) and Zhou et al. (2022). Hereafter, we only review relevant works on automated apple harvesting systems. The automation systems for apple harvesting can be grouped into two major categories. The first category is bulk harvesting (De Kleine & Karkee, 2015; Zhang et al., 2020), which is based on the principle of applying vibrations to the tree trunk and/or branches to detach apples. Although the bulk harvesting techniques are effective in separating fruits from trees, they generally produce excessive tree damage and apple bruising due to machine vibrations and extensive collisions (e.g., apple-to-apple, apple-to-tree, and apple-to-container collisions). Bruised apples are not acceptable for the fresh market, and thus there is an active research filed on reducing the damage caused by bulk harvesting (Fu et al., 2020). The other category is fruit-by-fruit selective harvesting, which could substantially alleviate or eliminate fruit damage by using appropriate manipulators to pick fruits in a controlled manner. Because of this advantage, along with the latest advances in robotics, sensors, and artificial intelligence technologies, the fruit-by-fruit robotic harvesting approach has attracted extensive attention from both academics and industry in recent years. A typical fruit-by-fruit selective harvesting system includes a perception module, a manipulator, and an end-effector. The perception module aims at identifying and localizing the apples by using sensors mounted on the robot (e.g., cameras and lidars). With the information for the target fruit positions generated by the perception module, the manipulator is regulated to approach the target apple, which is subsequently detached from the tree using a specialized end-effector (e.g., vacuum tube or gripper). The picked apple is then released to a receiver and transported to the final destination or a container. It is clear that the fruit-by-fruit harvesting system requires multidisciplinary advancements to enable a variety of synergistic functionalities, including perception, planning, and manipulation. Therefore, it is a great challenge to design such systems to achieve satisfactory picking efficiency and real-world viability.

So far, many different highly integrated robotic systems have been developed to harvest apples fruit-by-fruit (Baeten et al., 2008; Bu et al., 2022; Bulanon & Kataoka, 2010; Hohimer et al., 2019; Silwal et al., 2017; Zhang et al., 2022, 2021; Zhao et al., 2011). For example, Baeten et al. (2008) reported an apple harvesting robot, which consisted of an internal camera, a seven-degree-of-freedom (7-DOF) industrial manipulator, and a vacuum activated, funnel shaped gripper. The system scanned the tree from 40 look-out sectors, and for each sector the apples were detected and picked one-by-one with the harvesting cycle time of 8–10 s/fruit. In Silwal et al. (2017), an apple harvester was designed with a global camera, a 7-DOF manipulator, and a finger-based gripper. Field tests conducted on a V-trellis orchard showed that the system successfully picked 84% apples attempted, and the total cycle time to pick a single fruit was about 7.6 s. However, the orchard had been prethinned with clusters of foliage being removed to facilitate fruit harvesting by the robot, and a black curtain was used behind the tree canopy to facilitate apple detection. Hohimer et al. (2019) proposed a harvesting robot based on a pneumatic three-dimensional (3D)-printed soft-robotic end-effector, and the system achieved a successful detachment rate of 67% with an average time of 7.3 s/fruit that the system took from apple detachment to storage bin. Moreover, in Bu et al. (2022), two picking patterns were designed to evaluate the performance of a robotic apple harvester, and both of them achieved a successful rate of more than 80%. However, the harvester took at least 12 s to pick a single apple, and field tests were only conducted in an ideal condition where apples were fully exposed with few or no leave occlusions and flexible and long branches.

As identified in the aforementioned systems, fruit perception is a critical step in robotic apple harvesting. The perception procedure mainly includes two parts: detection and localization. The former is to segment apples from the background areas, and the latter aims at computing the spatial locations of the detected apples. RGB-D cameras have been a popular sensor choice for fruit detection and localization. Specifically, RGB image information is first consumed for apple detection by using conventional computer vision methods (Bulanon & Kataoka, 2010; Silwal et al., 2017; Zhao et al., 2005) or deep learning-based methods (Chu et al., 2021; Kang & Chen, 2020; Tian et al., 2019). Then, the depth information provided by RGB-D cameras is fused with the obtained detection results to calculate the apple location. Considerable research efforts have been devoted to improving apple detection in challenging orchard environments, while much less emphasis has been placed on fruit localization. Although consumer RGB-D cameras are compact and can provide dense image and depth information, several recent works (Bu et al., 2022; Fu et al., 2020; Zhang et al., 2021) reported that the depth measurements by these RGB-D cameras are not robust under varying lighting conditions or when apples are partially occluded by foliage, which would result in inaccurate fruit localization and eventually degrades the harvesting performance of the robotic system. As such, it is imperative to develop new hardware and software designs to enhance fruit localization accuracy and robustness.

Despite the significant progress, the existing robotic apple harvesting systems are still far from being commercially viable mainly because they are unreliable in performance and inefficient or too slow in picking fruit in the real orchard environment and are too complicated or expensive to be economically sound (Gongal et al., 2015). In this paper, we introduce the system design and field evaluations of an efficient and

high-performance robot platform for automated fruit-by-fruit apple harvesting. Specifically, the main components of the system include a novel perception component, a low-cost 4-DOF manipulator, a vacuum-based soft end-effector, and a dropping/catching component. The perception component is comprised of an RGB-D camera and a laser-camera unit. A deep learning-based perception algorithm is developed to robustly detect apples in clustered environments, and an active laser-camera scanning scheme is exploited to obtain precise target apple positions. The 4-DOF manipulator, the vacuum-based end-effector, and the dropping component are designed to approach, detach, and collect the apple, respectively. In contrast to high DOF industrial manipulators used in most existing works, the developed 4-DOF manipulator has a simpler design, making it easier to assemble and more effective at approaching apples. In addition, unlike finger-based or blade-based grippers that require bulky or complex actuating mechanisms, the soft end-effector is affixed to the inlet of a vacuum tube (i.e., the front end of the manipulator) for fruit gripping, which eliminates the need for an intricate mechanism, allows for conformity to different fruit contours and orientations, and avoids fruit bruising during the harvesting process. The major contributions of this paper are summarized as follows.

- We present the hardware and software development of the robotic apple harvesting system, which is not only compact in structure design but also effective in utilizing multidisciplinary advances to enable synergistic harvesting functionalities.
- We propose a novel perception strategy that integrates an RGB-D camera and a laser-camera unit to achieve robust apple detection and localization. To the best of our knowledge, this is the first work that fuses a line laser with a camera to achieve millimeter-level localization performance.
- We present performance results from the field testing and evaluation of the harvesting robot in two apple orchards with different tree architectures and discuss further improvements for the robot system.

The remainder of this paper is organized as follows. Section 2 presents the hardware components of the apple harvesting robot, whereas Section 3 details the software design. Field evaluation results are provided in Section 4. Finally, conclusions are drawn in Section 5.

2 | HARDWARE SYSTEM DESIGN

The principal objective of the hardware system design is to build a fully self-contained, modular harvesting platform that can support various payloads, endure rough terrains, and be easily modifiable and extensible. To facilitate the movement in the orchard environment, the whole system is built on a trailer base which can be hauled by a farm vehicle or robotic moving platform. The developed robotic apple harvesting system is shown in Figure 1, which consists of three major modules: the support hardware module, the computer and operation module, and the robotic harvester.

The support hardware module includes a 5.5 kW Honda gas-powered electric generator and a Delfin industrial vacuum. The generator provides a 240 V power source and can continuously support the whole system running for approximately 5.5 h if fully refilled. The Delfin industrial vacuum has two powerful bypass

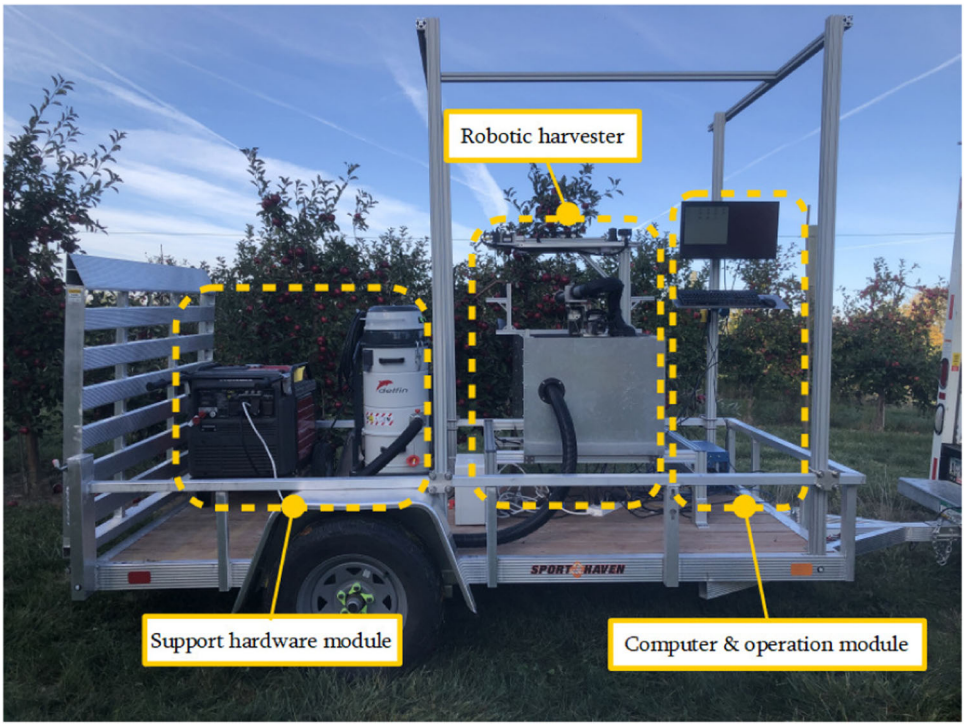

FIGURE 1 Hardware modules of the developed robotic apple harvesting system.

motors with independent cooling and can generate a peak horsepower of 5.5 HP. During fruit harvesting, this vacuum machine runs continuously to provide vacuum flow, generating suction forces to enable the soft effector to detach fruits (see Section 2.3 for descriptions on the end-effector).

The computer and operation module is comprised of a high-performance industrial computer and a workstation where users can monitor and control the robotic system. The industrial computer has an Intel®Xeon E2176G processor, 64 GB of RAM, and a NVIDIA GeForce RTX 2080 Ti graphic processing unit. This computer hosts all software algorithms and the communication connections to all components.

The main components of the robotic apple harvesting system are shown in Figure 2, including a perception component, a 4-DOF manipulator, a vacuum-based soft end-effector, and a dropping component. More detailed descriptions on these four components are given in the following sections.

2.1 | Perception component

For robotic apple harvesting, the first and foremost task is to detect and localize the fruits. As shown in Figure 3, a RealSense D435i RGB-D camera (Intel Corp.) and a custom-built laser-camera unit are integrated as the sensor set to achieve apple detection and localization. The RGB-D camera is mounted on a horizontal frame that is above the manipulator. Different from the other robotic harvesting systems (e.g., Baeten et al., 2008) that attach the camera to the manipulator or the end-effector, our installation scheme ensures that the RGB-D camera can provide a global view of the scene, which facilitates the use of multiple manipulators planned in our future versions. Since the depth measurement of consumer RGB-D camera is not stable under leaf/branch occlusions and/or challenging lighting conditions (see Fu et al., 2020; Neupane et al., 2021, for an in-depth review on localization performance of commercial RGB-D sensors), we design a new laser-camera unit to address this issue. The specially designed laser-camera unit is comprised of a red line laser with a 635 nm wavelength (Laserglow Technologies), a Flir RGB camera (Teledyne Flir), and a linear motion slide. The line laser is mounted on top of the linear motion slide which enables the laser to move back and forth horizontally with a full stroke of 20 cm. Meanwhile, the Flir RGB camera is installed at the rear end of the linear motion slide with a relative angle to the laser scanner. The RGB-D camera and the laser-camera unit are fused synergistically to achieve high perception accuracy and robustness, where the fusion scheme will be detailed in Section 3.

2.2 | Manipulator

A 4-DOF manipulator is designed and assembled with compact mechanical structure for efficient manipulation in the workspace. As illustrated in Figure 4, the manipulator consists of three revolute

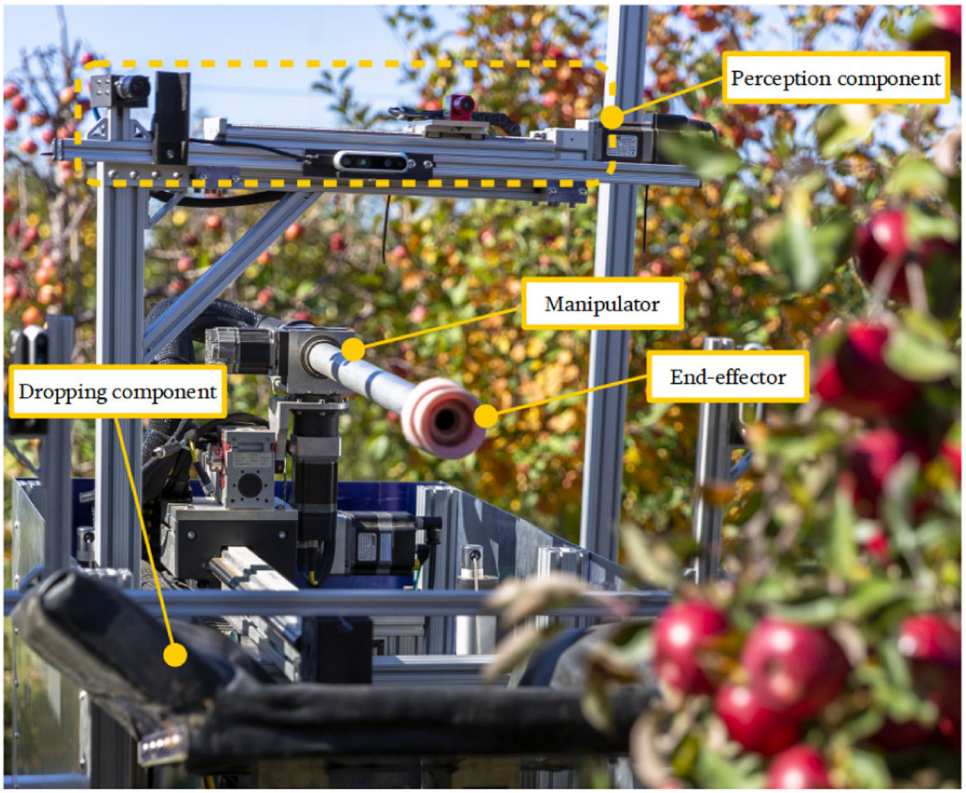

FIGURE 2 Main components of the robotic harvester for apple picking.

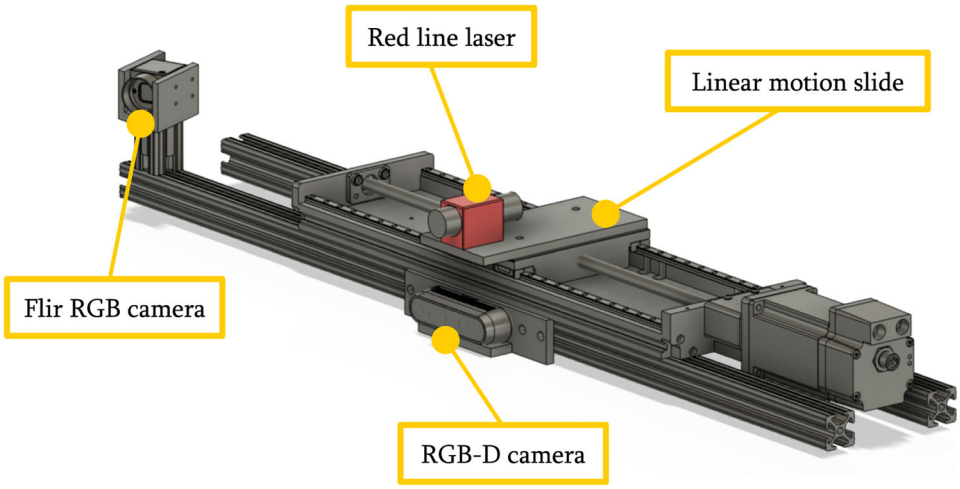

FIGURE 3 CAD model of the perception component.

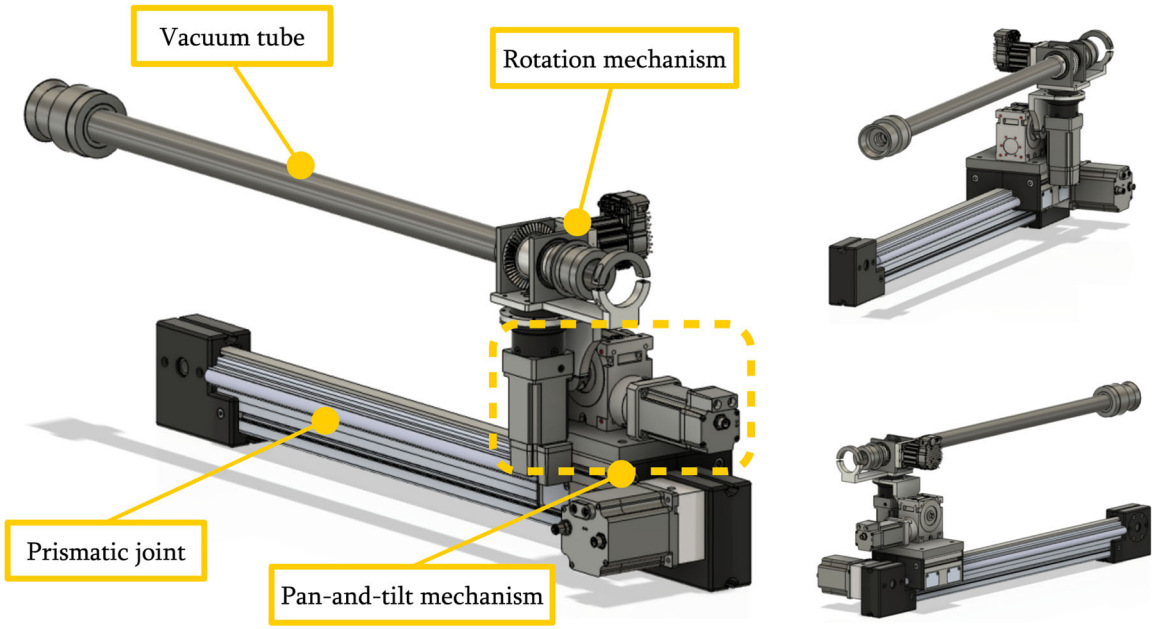

FIGURE 4 CAD model of the 4-DOF manipulator.

joints and one prismatic joint. The first and second revolute joints are connected by an *L*-shaped aluminum plate to form a pan-and-tilt mechanism. The prismatic joint is used as the base of the pan-and-tilt mechanism to extend the manipulator's workspace. A hollow aluminum link is installed on the pan-and-tilt mechanism to enable the end-effector to reach the apples and serve as a vacuum tube for grasping the fruits during the harvesting process. The third revolute joint is assembled at the rear end of the aluminum tube to create a rotation mechanism. After the end-effector has grasped the fruit, this rotation mechanism is triggered to rotate the aluminum tube to detach the fruit from the tree. In addition, the revolute joints and the prismatic joint are driven by SMD23E2 and SMD34E2 integrated motors (Advanced Micro Controls Inc.), respectively. These motors are equipped with onboard controllers and encoders. The onboard controllers can execute velocity commands transmitted via EtherNet, and the embedded encoders can provide real-time position feedback, enabling precise measurement of the manipulator's status during operation. The current manipulator design has transitioned from a hybrid pneumatic/motor actuation mechanism used in our previous iteration (Zhang et al., 2021) to a fully motor-driven one. This deliberate change simplifies the actuation process and facilitates the development of an integrated control scheme.

Different from most existing studies (Baeten et al., 2008; Bu et al., 2022; Hohimer et al., 2019; Silwal et al., 2017) that rely on high DOF industrial manipulators, the developed 4-DOF manipulator is simple and compact in structure and highly efficient in picking fruit.

This hardware design allows convenient development of adaptable control algorithms such that agile manipulation can be achieved. It also facilitates future extensions of multiple robotic arms for coordinated multiarm apple harvesting.

2.3 | End-effector

In our robotic system, a vacuum-based soft end-effector is used to grasp and detach fruits. The end-effector is a vacuum cup made of silicone rubber and is attached to the front end of the aluminum tube (i.e., the front end of the manipulator). Through laboratory experiments, a silicone material with a hardness of 40 Shore A and a cup shape geometric design shown in Figure 5 are selected for the end-effector. Compared with our previous designs (Lu et al., 2022), the current cup shape geometric design has a smaller inner diameter and a larger outer-lip. The end-effector is securely connected to the aluminum tube through an adapter, and the rear end of the aluminum tube is connected to the Delfin industrial vacuum via a flexible and expandable tube. Indoor testing reveals that when the vacuum machine operates in continuous mode, the soft end-effector can generate a suction force of about 47 N, which is sufficient for holding and detaching the fruit.

The motivation behind designing this vacuum-based soft end-effector stems from several aspects. First, the unpredictable nature of apple contours and hanging orientations poses a challenge for end-effector performance. Traditional finger-based or blade-based grippers often require meticulous orientation adjustments to facilitate fruit detachment. In contrast, our developed end-effector's soft nature and improved cup geometry allow it to flexibly conform to the diverse contours and orientations of apples. This adaptability ensures efficient gripping while reducing the risk of fruit bruising. Second, unlike existing vacuum-based harvesting systems, such as the one developed by Abundant Robotics (Simonit, 2017) that sucks apples into a large vacuum tube, our system utilizes a compact end-effector coupled with a small-sized vacuum tube (44.5 mm inner diameter) for fruit gripping. This design choice minimizes the chances of sucking leaves or other components into the vacuum chamber. Furthermore, a major advantage of employing the vacuum-based end-effector is that it can tolerate some approaching inaccuracy. Operating under the current vacuum flow, it can attract fruits within a distance of about 1.5 cm. This feature allows the manipulator to grasp the fruit even if it does not approach the target accurately.

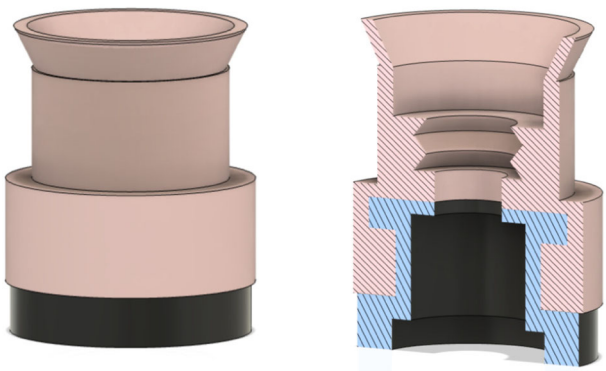

FIGURE 5 CAD model of the vacuum-based soft end-effector.

2.4 | Dropping/catching component

A dropping/catching component is assembled and attached to the robot platform to enable faster release or dropping of the picked fruit. The base of the dropping component is made up of a rectangular aluminum plate covered with a soft foam cushion of 50 mm in thickness. Laboratory tests showed that the foam cushion allows apples to drop from the highest position of the end-effector (approximately 80 cm) without causing fruit bruising, while keeping fruit bouncing to minimum. The manipulator can drop the picked apples at any spots from above the dropping component without fully returning to its home position, thus reducing the overall fruit picking cycle time. After an apple has fallen onto the sloped surface of the dropping component, it rolls down to a screw-driver conveyor, which transports the apple to the destination or a bin (Zhang et al., 2021).

3 | SOFTWARE/ALGORITHM DESIGN

The software suite is designed within the robot operating system (ROS) framework. Different software components are primarily communicated via custom messages sent through ROS actions and services. Figure 6 shows the main logic/algorithm flow of the software system during apple harvesting. It is apparent that the software design of our robotic system requires multi-disciplinary advances to enable various synergistic functionalities and coordination for achieving reliable automated apple harvesting. The logic flow of the apple harvesting cycle is detailed in the following.

At the beginning of each harvesting cycle, the RGB-D camera was triggered to acquire images at 30 fps. Based on the obtained image information, the deep learning-based technique was exploited to detect and localize the fruits within the manipulator's workspace. All identified apples were prioritized to generate a picking list of 3D apple locations. The criterion to determine the picking list was "closest apple first pick," that is, the apple closer to the robot has higher priority for harvesting. Other criteria, for example, travel cost based optimization (Zhang et al., 2022), can also be utilized. After obtaining the picking list, the one on top of the list was selected as the target fruit. Since location results provided by the RGB-D camera might not be sufficiently accurate, the developed laser-camera unit and corresponding perception scheme were triggered to scan the target fruit and calculate its 3D position. Given the ameliorative target apple location, the planning algorithm was used to generate a reference trajectory, and the control module will actuate the manipulator to follow this reference trajectory to reach the fruit.

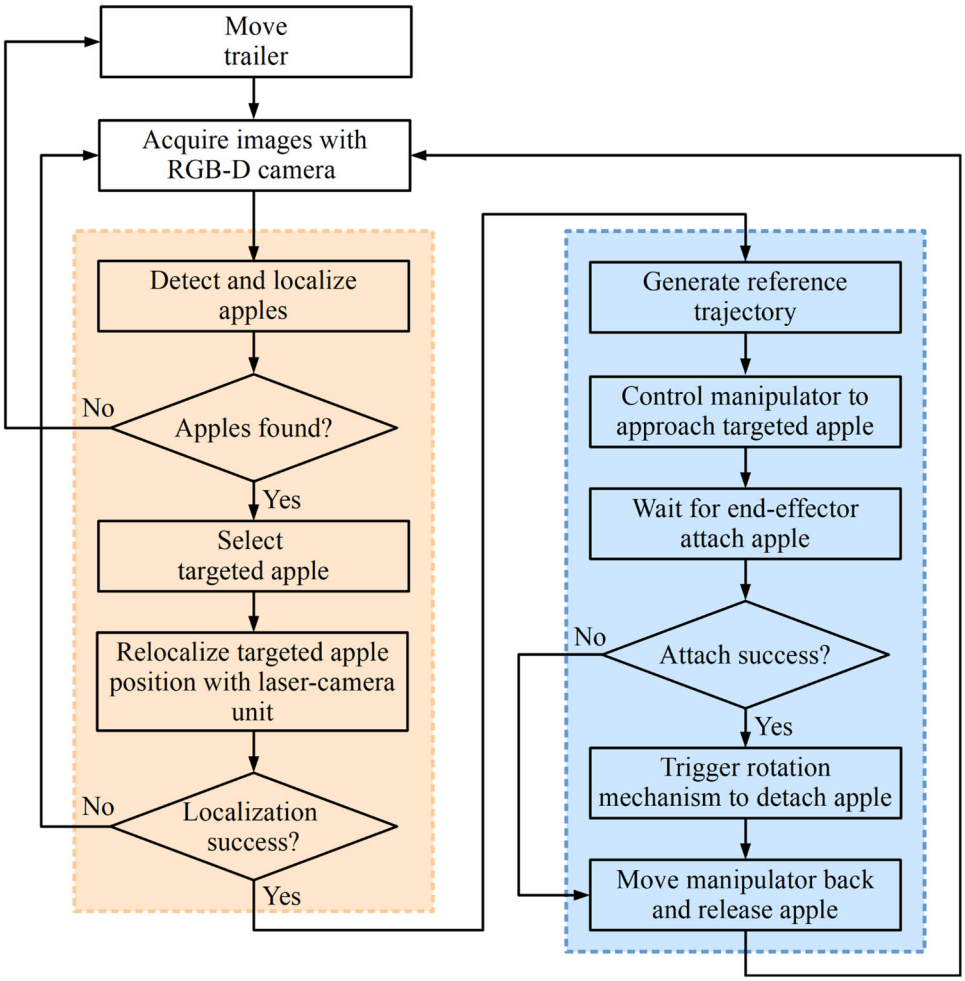

FIGURE 6 Logic flowchart in apple harvesting.

When the fruit was successfully attached to the end-effector, a sharp pressure drop will occur within the vacuum chamber which was monitored by a pressure sensor mounted inside the vacuum tube. Once the pressure dropped below a predetermined threshold, the rotation mechanism was triggered to rotate the whole aluminum tube by a certain angle, and the manipulator then retracted to pull and detach the apple (if the rotation action had not resulted in complete detachment of the fruit). After the manipulator reached a predetermined dropping spot (detected by embedded encoders of motors), a vacuum control valve installed between the outlet of the vacuum machine and the inlet of the flexible vacuum hose (not shown in Figure 4) was actuated, which caused rapid loss of vacuum pressure in the tube, thus enabling the fruit to fall off the end-effector by gravity to the dropping component. The fruit finally rolled down the slope of the dropping component to the screw conveyor, from which the apple was transported to the destination. Note that when the manipulator failed to approach or attach the fruit (e.g., in the case of a collision with obstacles), the pressure sensor will detect the absence of a significant pressure drop. As a result, the system will trigger the manipulator to return back and start the next harvesting cycle, ensuring continuous harvesting operation.

3.1 | Apple detection and localization with RGB-D camera

Apple detection and localization is one of the key functionalities in robotic harvesting. The former aims at segmenting apples from the complex background areas, while the latter is to subsequently compute the 3D locations of the detected fruits that will be used to guide the robot to reach the target. In our preliminary work (Chu et al., 2021), a deep learning-based method was developed that combines a Mask R-CNN backbone and a suppression end for apple detection, which achieved state-of-the-art performance on our collected data set. In the current version, we further refined the model by introducing an occlusion-aware structure to enhance the detection performance. This is motivated to address the major challenge in orchard perception as identified in our previous field tests—apple occlusion and overlap.

Specifically, in the cluttered orchard environment, apples tend to occlude each other, making it difficult for conventional deep learning-based approaches to detect both the occluder and occludee fruits. Therefore, a new network—occluder-occludee relational network (O2RNet)—is designed to enhance detection performance in occlusion scenarios. Figure 7 illustrates the network structure of the developed

method, which is composed of three parts. The first part is a backbone whose functionality is to extract features from the entire image. We choose residual network 101 (He et al., 2020) as the backbone due to its ability to keep strong semantic features at different resolution scales. The second part is a region proposal network (RPN) (Ren et al., 2017), which aims at identifying promising regions of interest (ROIs) that are likely to contain an object. The RPN utilizes the feature maps provided by the backbone to generate promising ROIs via a series of convolutional and fully connected layers. The RPN output is then fed into the third part, that is, an occlusion-aware modeling head, to decouple overlapping relations and segment the apples. The occlusion-aware modeling head consists of an occluder branch and an occludee branch. The occluder branch follows the design of the object detection head in Faster R-CNN (Ren et al., 2017) to prepare the cropped ROI features for the occludee branch. By combining the feature expansion structure and the cropped ROI features, the occludee branch can refine object boundaries and generate bounding boxes for the partially occluded apples. The O2RNet algorithm is tested on a comprehensive orchard image data set and achieves an identification accuracy of more than 90%. Figure 8 shows an example of the detection results where bounding boxes represent the identified apples. A detailed description on the development and implementation of the O2RNet algorithm for apple detection can be found in Chu et al. (2023).

Once the fruit is detected, apple localization is performed with the depth measurements provided by the RGB-D camera. Specifically, the image pixels of the detected apple are extracted for each bounding box to create a range matrix based on the disparity map. The mean value of the range matrix is then calculated, and it is used as the depth range of the apple. Back-projection (Hartley & Zisserman, 2003) is utilized to determine the apple's 3D position by fusing the depth range with the center of the bounding box pixels. This procedure is performed for every bounding box to obtain the positions of all detected apples in the image.

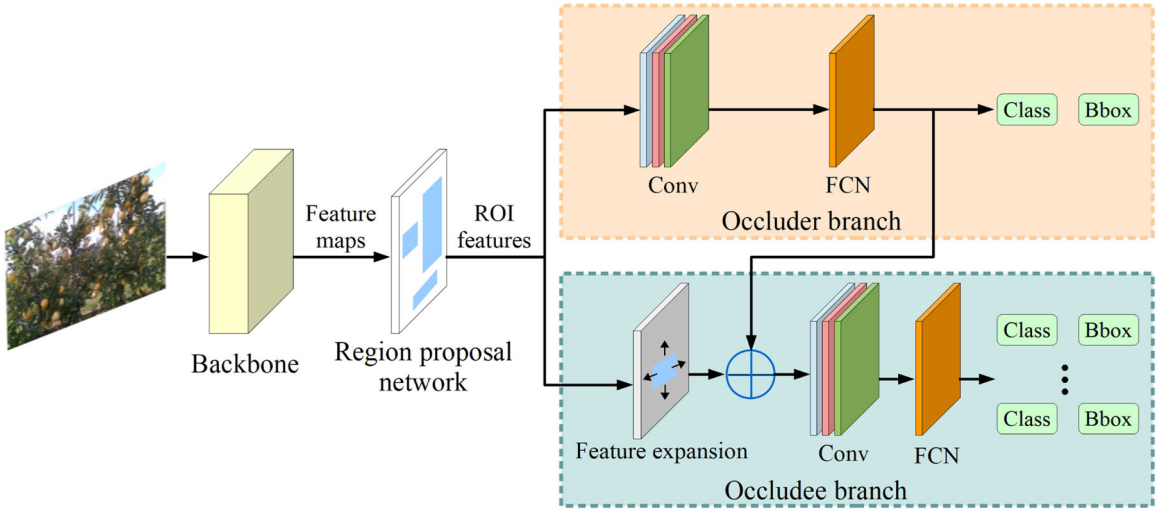

FIGURE 7 Network structure of the deep learning-based method for apple detection.

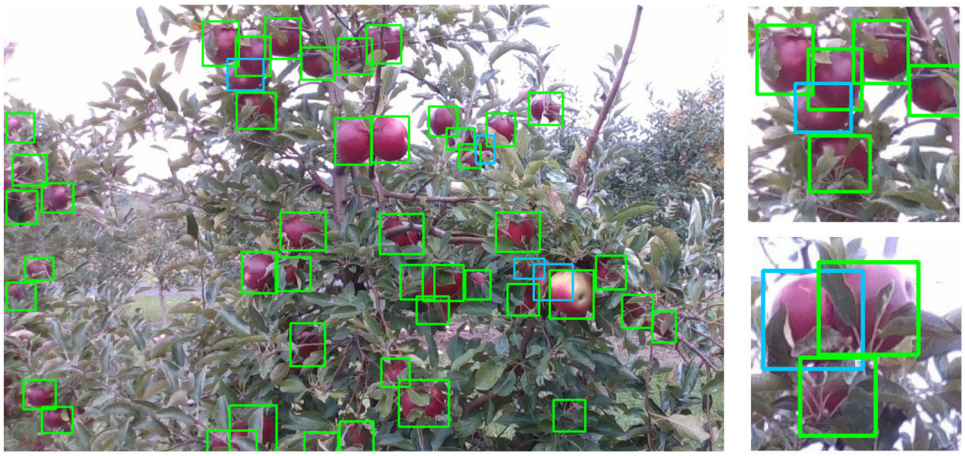

FIGURE 8 Example of apple detection through O2RNet algorithm.

3.2 | Refined apple localization with active laser-camera scanning

Although the RealSense RGB-D camera is compact and can provide dense environmental information, its depth measurements are sensitive to varying lighting conditions and lack stability when apples are partially occluded by branches or foliage. Such limitations would result in inaccurate apple localization, which has been identified as one of the primary causes for harvesting failure in our previous field tests. Considering this issue, we develop an active laser-camera scanning scheme to enhance the apple localization accuracy (Figure 9). More specifically, the rough apple location provided by the RGB-D camera will be used by the laser-camera unit to facilitate the scanning of the target apple and the recalculation of the 3D position. The pseudo code of the active laser-camera scanning scheme is shown in Algorithm 1, and more details are introduced below:

Algorithm 1 Active-laser camera scanning for apple localization

Input: rough target apple location provided by the RGB-D camera
Output: refined target apple location
1. Initialization
Based on the rough target apple location, compute an initial laser position
Adjust the laser to the initial position with laser line being projected on the left side of the target apple
Trigger Flir camera to capture image $\mathcal{I}_1$
2. Interval scanning
for k = 1, 2, 3, 4 **do**
Move the laser to the right by one centimeter
Trigger Flir camera to capture image $\mathcal{I}_{k+1}$
end for
3. Refinement of 3D position
for k = 1, 2, 3, 4, 5 **do**
Extract laser line from image $\mathcal{I}_k$
Compute 3D location candidate q_k based on laser triangulation techniques
end for
Select refined target apple location from $\{q_1, q_2, q_3, q_4, q_5\}$ based on a holistic evaluation function

1. Initialization: After the target apple has been selected, the linear motion slide is controlled to adjust the laser towards an initial position so that the red laser line would be projected on the left half region of the target. To calculate the initial laser position, the rough target apple location provided by the RGB-D camera is transformed to the laser-camera unit's coordinate frame, which is used to regulate the linear slider to direct the laser line onto the left side of the target. Here the relative pose/transformation between the RGB-D camera and the laser-camera unit is carefully calibrated offline. Once the laser reaches the initial position, the Flir camera is triggered to capture an image.
2. Interval scanning: The linear motion slide subsequently moves to the right by a total of 4 cm with 1 cm increments, and it pauses at each increment for the Flir camera to capture a corresponding image. Through this scanning procedure, a total of five images will be collected, and in each image the laser line will be projected on different positions. Figure 9 shows a set of example images of the target apple that are collected during the interval scanning. The motivation for using this scanning scheme is to alleviate the impact of occlusions caused by leaves, branches, and/or other apples to enhance the localization accuracy of the target fruit. Specifically, when the target apple is partially occluded by obstacles, moving the laser to various positions can reduce the possibility of laser lines being completely blocked by the obstructions, leading to a considerable improvement in the localization accuracy for the target fruit.
3. Refinement of 3D position: For each image captured by the Flir camera, the laser line projected on the target apple surface is extracted and then used to generate a 3D location candidate. Computer vision approaches and laser triangulation-based techniques are exploited to accomplish laser line extraction and position candidate computation, respectively. Five position candidates will be generated as a result, and a holistic evaluation function is designed to select one of the candidates as the final target apple location.

FIGURE 9 Example images of target apple during the interval laser scanning.

Laser line extraction and position candidate computations are two main challenges in Step 3. We design the corresponding algorithms to address the challenges by fully considering the hardware structure of the laser-camera unit. The algorithms are detailed in the following.

As illustrated in Figure 10, laser line extraction mainly includes five steps. In the first step, the red channel information of the raw RGB image is chosen as the substrate for the laser line extraction. The RGB image has three color channels, and through preliminary tests, it is found that under the red channel, the laser line has higher contrast with the background of apples or foliage. Therefore, we select the red component of the raw image to facilitate the following processing. In the second step, we use a threshold filter to segment the laser line from the background. The design of this filter is based on the observation that the pixel values of the laser line almost always approach saturation in the red channel. As there exists laser diffusion, we extract the centroid points of the segmented outline in the third step. The polynomial function is used in the fourth step to fit a continuous curve based on the discrete centroid points. Note that if the target apple is occluded by foliage, there may exist multiple segments of the laser line. Under this circumstance, only the centroid points extracted from the most prominent segment of the target apple will be selected for curve fitting. Finally, the curve is used to represent the laser line position, and the image coordinates $[u_i, v_i]^\top \in \mathbb{R}^2$ $(i = 1, 2, ..., n)$ of the curve portion belonging to the target apple will be outputted for the calculation of 3D position candidates.

The computation of position candidates is conducted with the laser triangulation-based technique. Laser triangulation-based technique is one of the most effective approaches to measure depth information from camera images and has been widely used in various applications (Kjaer & Ottosen, 2015; Schlarp et al., 2020). The basic idea of this technique is to capture depth measurements by pairing a laser illumination source with a camera. Both the laser beam and the camera are aimed at the target object, and based on the extrinsic parameters between the laser source and the camera sensor, the depth information can be collected with trigonometry. In our system, three extrinsic parameters are considered to describe the trigonometry model of the laser-camera unit. Figure 11 shows these parameters, where $\alpha \in \mathbb{R}$ is the rotating angle of the Flir camera with respect to the laser, $L \in \mathbb{R}$ is the horizontal distance between the camera and the laser, and $\beta \in \mathbb{R}$ is the angular offset of the emitted laser line. α and β are constants, while L is variable (but measured) as the linear motion slide will move to different positions. α, β, and the initial value of L are obtained via offline calibration, and L will be updated online based on its initial value and the movement distance of the linear motion slide. We now introduce the method used to calculate position candidates. As mentioned above, the laser line extraction will output a sequence of image coordinates $[u_i, v_i]^\top$

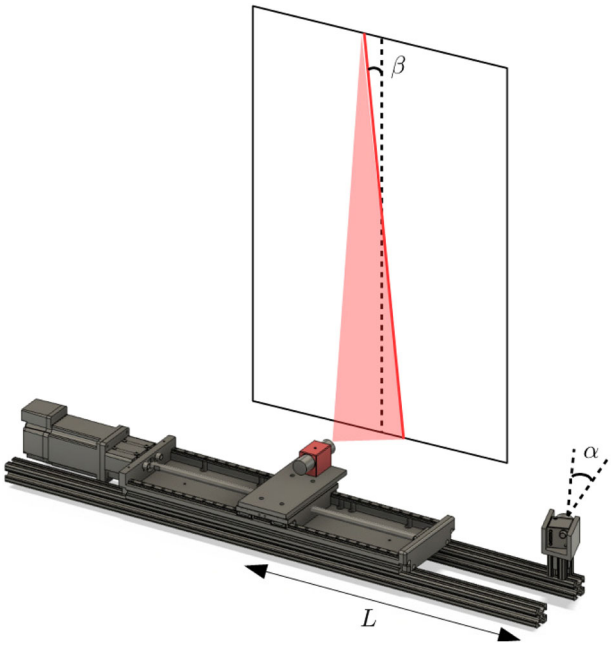

FIGURE 11 Extrinsic parameters between the laser source and camera sensor.

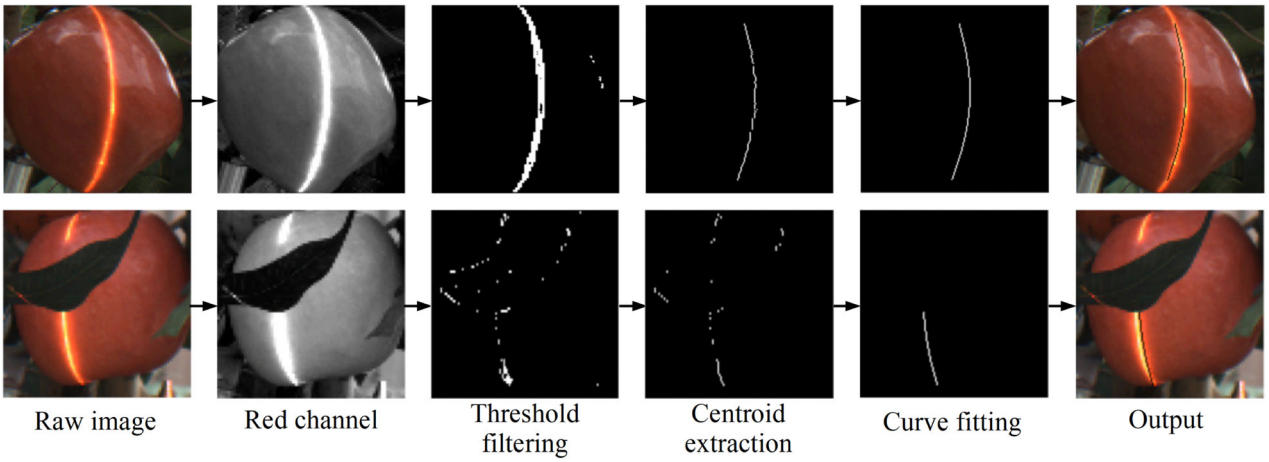

FIGURE 10 Diagram illustrating the steps performed to extract laser line.

($i = 1, 2, \ldots, n$) to represent the laser light projected on the target apple. By using intrinsic camera parameters, the image coordinates $[u_i, v_i]^\top$ are first transformed to the normalized image coordinates, denoted by $[\bar{u}_i, \bar{v}_i]^\top \in \mathbb{R}^2$ (Hartley & Zisserman, 2003). Then, based on the trigonometry model of the laser-camera unit (Zhang et al., 2023), the corresponding depth of each normalized image coordinate $[\bar{u}_i, \bar{v}_i]^\top$ can be computed by

$$z_i^c = \frac{L}{\sin(\alpha) - \bar{u}_i \cos(\alpha) - \bar{v}_i \tan(\beta)}, \; i = 1, 2, \ldots, n. \tag{1}$$

The mean value of z_i^c is calculated as the depth range $z^c \in \mathbb{R}$ of the target apple, that is,

$$z^c = \frac{1}{n}\sum_{i=1}^{n} z_i^c. \tag{2}$$

Finally, by combining the depth range with the center of the bounding box (i.e., the target apple center), back-projection (Hartley & Zisserman, 2003) is used to calculate the 3D position candidate. Laboratory experiments show that the laser-camera unit can achieve better apple localization accuracy than the RGB-D camera, with an error of less than 7 mm.

In summary, the new perception component has been tailored for the current robotic system to calculate the 3D apple position based on the laser triangulation-based method. To the best of our knowledge, this is the first work that extends the laser triangulation-based technique to enhance fruit localization accuracy for autonomous robotic harvesting.

3.3 | Planning and control

The apples within the workspace are localized by the perception component and then assigned in turn as the target positions $p_d = [x_d, y_d, z_d]^\top \in \mathbb{R}^3$, where the manipulator needs to reach. To facilitate the manipulation control, given a target position p_d, we use the quintic function (Siciliano et al., 2010) to generate a corresponding reference trajectory $p_r(t) = [x_r(t), y_r(t), z_r(t)]^\top$. This reference trajectory is a function of time with its terminus being the target position p_d. The introduction of the quintic function-based reference trajectory p_r brings the following advantages: First, the reference trajectory is continuously differentiable and its terminal velocity and acceleration are zero, which is conducive to ensuring that the end-effector approaches the desired position along a smooth path. Second, by adjusting function parameters, the velocity profile of the reference trajectory can be modified, and thus the end-effector can reach the desired position within a specific time interval.

Given a target apple position and the generated reference trajectory using the planning algorithm discussed above, we next introduce the control algorithms that drive the manipulator to follow the reference trajectory. To achieve this goal, a motion control strategy is developed by fully exploiting the mechanical structure of the developed 4-DOF manipulator. The kinematic model and parameters of the manipulator are presented in Figure 12 and Table 1, respectively. Denote $p = [x, y, z]^\top \in \mathbb{R}^3$as the position of the end-effector. Based on the Denavit-Hartenberg convention (Siciliano et al., 2010) and the kinematic diagram presented in Figure 12, the forward kinematics function of the manipulator can be derived, as follows:

$$\begin{aligned} x &= d_6 \cos(\theta)\cos(\varphi) + d_5 \sin(\theta) + d_2 + D, \\ y &= d_6 \sin(\varphi) + d_4 + d_1, \\ z &= -d_6 \sin(\theta)\cos(\varphi) + d_5 \cos(\theta) + d_3, \end{aligned} \tag{3}$$

where $d_1, d_2, d_3, d_4, d_5, d_6 \in \mathbb{R}$ are the link lengths, and $[\varphi, \theta, D]^\top \in \mathbb{R}^3$ are the joint variables. Note that the revolute joint ρ is assembled to rotate the vacuum tube for fruit detachment and has no bearing on the end-effector position.

The planning algorithm will provide the reference trajectory $p_r(t) = [x_r(t), y_r(t), z_r(t)]^\top$for the target position p_d. The objective of the manipulation control is to regulate the end-effector to follow the reference trajectory p_r and finally approach the target position p_d. The

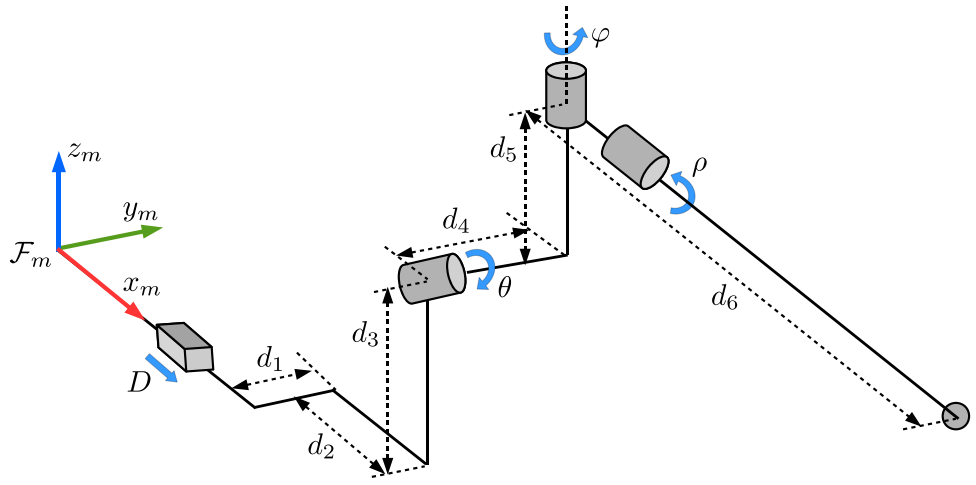

FIGURE 12 The kinematic model of the 4-DOF manipulator.

TABLE 1 Model parameters of the 4-DOF manipulator.

Parameter	Value
Link d_1	0.017 m
Link d_2	0.031 m
Link d_3	0.062 m
Link d_4	0.093 m
Link d_5	0.138 m
Link d_6	0.844 m
Revolute joint ρ	(−180°, 180°)
Revolute joint φ	(−25°, 25°)
Revolute joint θ	(−25°, 25°)
Prismatic joint D	(0 m, 0.55 m)

revolute joint parameters φ, θ, and prismatic joint parameter D are all driven by electrical motors, and a velocity-based control scheme is employed to generate explicit speed commands to smoothly adjust the joints based on real-time position feedback. Specifically, based on (3), the time derivative of $[x, y, z]^\top$can be calculated as

$$\begin{aligned}\dot{x} &= -d_6(\sin(\theta)\cos(\varphi)\omega_\theta + \cos(\theta)\sin(\varphi)\omega_\varphi) + d_5\cos(\theta)\omega_\theta + v_D,\\ \dot{y} &= d_6\cos(\varphi)\omega_\varphi,\\ \dot{z} &= -d_6(\cos(\theta)\cos(\varphi)\omega_\theta - \sin(\theta)\sin(\varphi)\omega_\varphi) - d_5\sin(\theta)\omega_\theta,\end{aligned} \tag{4}$$

where $\omega_\varphi, \omega_\theta \in \mathbb{R}$ are the angular velocity of the revolute joints φ and θ, respectively, and $v_D \in \mathbb{R}$ is the linear velocity of the prismatic joint D. Furthermore, the error signals $[e_x, e_y, e_z]^\top \in \mathbb{R}^3$ are constructed as

$$\begin{aligned}e_x &= x - x_r,\\ e_y &= y - y_r,\\ e_z &= z - z_r.\end{aligned} \tag{5}$$

Based on (4), (5) and by virtue of Lyapunov-based control techniques (Khalil, 2002) and robust control technique (Patil et al., 2022; Xian et al., 2004), the velocity controller is designed as

$$\begin{aligned}\omega_\varphi &= \frac{-k_y e_y + \dot{y}_r + \eta_y}{d_6\cos(\varphi)},\\ \omega_\theta &= \frac{k_z e_z + d_6\sin(\theta)\sin(\varphi)\omega_\varphi - \dot{z}_r + \eta_z}{d_6\cos(\theta)\cos(\varphi) + d_5\sin(\theta)},\\ v_D &= -k_x e_x + d_6(\sin(\theta)\cos(\varphi)\omega_\theta + \cos(\theta)\sin(\varphi)\omega_\varphi)\\ &\quad -d_5\cos(\theta)\omega_\theta + \dot{x}_r + \eta_x,\end{aligned} \tag{6}$$

where $k_x, k_y, k_z \in \mathbb{R}^+$ are positive constant gains, and $\eta_x, \eta_y, \eta_z \in \mathbb{R}$ are given by

$$\begin{aligned}\eta_x &= \iota_x\int_0^t (e_x(\tau) + \operatorname{sgn}(e_x(\tau)))d\tau,\\ \eta_y &= \iota_y\int_0^t (e_y(\tau) + \operatorname{sgn}(e_y(\tau)))d\tau,\\ \eta_z &= \iota_z\int_0^t (e_z(\tau) + \operatorname{sgn}(e_z(\tau)))d\tau.\end{aligned} \tag{7}$$

In (7), $\iota_x, \iota_y, \iota_z \in \mathbb{R}^+$ are positive constant gains and $\operatorname{sgn}(\cdot)$ is the standard signum function. The control scheme designed in (6) has several advantages compared to the ones used in our previous robotic prototypes (Zhang et al., 2022, 2021). Different from Zhang et al. (2021) that uses a hybrid velocity/position control method to adjust the manipulator's pneumatic/motor actuation joints, the controller (6) is unified and can coordinate with the current manipulator to achieve faster and more accurate movement. Furthermore, in contrast to the control method proposed in Zhang et al. (2022), the controller (6) has advantages in eliminating the influence of bounded disturbances while ensuring asymptotic error convergence. Rigorous stability analysis can be conducted based on the Lyapunov-based methods and by leveraging the techniques developed in Patil et al. (2022) and Xian et al. (2004).

4 | FIELD EVALUATION

4.1 | Experimental setup

To fully evaluate the performance of the developed apple harvesting robot, field tests were conducted in two Michigan State University's research orchards in East Lansing and Holt, Michigan, USA, respectively, during the 2022 harvest season (Figure 13). The first orchard had been planted with "Gala" apple trees of 2 years old, which is a popular bicolored variety with a red color on the foreground and a yellow background. There were fewer fruits grown on these young trees with less occlusions by branches and foliage, but many of the fruits were grown in clusters (Figure 13b). In the second orchard were "Ida Red" apple trees of 7 years old (Figure 13c). Since the trees had not been pruned and thinned during the winter and early spring seasons, there were dense and unstructured foliage and branches. A high percentage of apples were grown in clusters and occluded by leaves and branches, which presented a significantly more challenging environment for our robot. It should be mentioned that most reported studies tested their robotic apple harvesting systems in high-density tree orchards with well-trained tree architectures with few clustered apples and less dense foliage (either naturally or being removed).

Field evaluation of the robot was conducted using the following procedure. The robot platform was first positioned in front of a crop row, with the perception component distanced from the canopy ranging from 0.6 to 1 m. Since the manipulator could reach a wide range of depth, it was not necessary to position it in a specific location at a fixed distance from the canopy or in line with the tree trunk. The robot was then actuated to perform apple picking autonomously and continuously with fully integrated functionalities. In each harvesting cycle, the perception component detected and localized the target apple. The manipulator was controlled to reach the apple, and the vacuum-based end-effector sucked and held the fruit, followed with tube rotation by the rotation mechanism to detach the fruit from the tree. The manipulator finally returned to a dropping spot and released the fruit to the dropping module by actuating the inline vacuum control valve. Once the apples within the workspace were all attempted and picked,

the harvest platform was moved to the next harvesting position. The field test was conducted under both sunny and cloudy weather conditions with moderate to minimal wind and temperature between 5°C and 20°C. In addition, a tarp was placed on top of the robotic harvester to block partial sunlight so that the perception module could perform stable apple detection and localization.

4.2 | Harvesting performance evaluation

In the first orchard, a total of 187 apples were present and identified by the detection algorithm where 154 of them were picked successfully with the harvesting rate being 82.4%. In the second orchard, the robotic system picked 101 of the 155 fruit attempted with the harvesting rate being 65.2%. As mentioned above, the tree structures in the second orchard were more dense and complex, thus resulting in a lower harvesting rate by the robotic system. For comparison, our earlier version of the robot prototype only achieved a picking rate of 52.11% in a similar orchard environment (Zhang et al., 2022). Clearly, the current system has demonstrated significant improvements. The improved performance of the current system was mainly attributed to the adoption of the new perception component which provided more accurate localization information for target fruits under occlusions.

Table 2 shows statistics of the failed picking cases that are grouped into four categories.

1. Obstacle: The most frequent cause for the failed picking cases was due to the obstruction of trunk, branches, or foliage. In contrast to V-trellis or vertical fruiting wall orchards where the majority of fruits would be well exposed, the orchards (especially the second one) where the experiment was carried out had a higher percentage of apples that were obscured by trunk, branches, and/or foliage. The current robot demonstrated satisfactory performance in harvesting fruits that were fully exposed or only mildly occluded. However, it failed to pick apples heavily obstructed by obstacles due to the absence of obstacle (i.e., trunk and branches) detection and avoidance functions.
2. Position error: Localization errors occurred when the algorithm failed to recognize the laser line precisely. Inaccurate laser line

FIGURE 13 Field trails of the robotic apple harvesting system. (a) Image of the platform operating in the orchard environment; (b) Example images of young and well-pruned trees in the first orchard; and (c) Example images of older trees with dense foliage in the second orchard.

TABLE 2 Failed case analysis for autonomous harvesting experiment.

Orchard	Obstacle	Position error	Flexible shoot	Attachment failure
1	18/33 (54.6%)	9/33 (27.3%)	4/33 (12.1%)	2/33 (6.1%)
2	31/54 (57.4%)	19/54 (35.2%)	3/54 (5.6%)	1/54 (1.9%)

extractions occurred when the target apples were heavily occluded by leaves or when the bright sunlight was directly shed onto the apple's surface, which would have caused problems in extracting laser lines or segments from the target fruit.

3. Flexible shoot: The missed apples in this category were due to the presence of long and flexible shoots. Under such circumstances, the apples that were grasped by the end-effector could not be detached, due to the flexibility and large swinging of the branches.
4. Attachment failure: There were only three occurrences in the two orchards during which the vacuum-based end-effector failed to detach apples due to irregular apple contours/orientations or improper contact angle between the apple and end-effector. The results demonstrated that the vacuum-based end-effector is highly effective in detaching target fruits, once they are attached to it, which represents significant improvements over the previous version of the end-effector designs (Lu et al., 2022).

Harvesting cycle times were estimated from the recorded videos taken during the field testing. For each harvesting cycle, on average, 1.97 s was required to accomplish apple detection and localization. The proposed apple detection and localization scheme included two phases. During phase one, which involved using the RGB-D camera for initial apple detection and localization, it took approximately 0.3 s to accomplish this task. Phase two, where the active laser-camera unit was utilized to refine the target apple position, took about 1.67 s on average. The duration for the manipulator to approach an individual apple and return to the home position was approximately 2.3 s. The fruit attachment usually took about 0.5 s. Detaching and releasing the apple took 0.7 and 0.5 s, respectively. The average cycle time required to successfully harvest a fruit was approximately 5.97 s, which was the total time for perception, manipulation to approach the apple, fruit attachment, fruit detaching, manipulation to return to the home position, and fruit releasing.

As presented in Figure 14, the harvested fruits were classified into three fruit removal categories: stem-intact, spur-pulled, and stem-pulled. Stem-intact apples would be stored better in a controlled environment than apples without stems (De Kleine & Karkee, 2015). During the field test, a total of 255 apples were picked from the two orchards. 174 or 68.2% of the picked fruit were in the stem-intact category. The number of spur-pulled and stem-pulled apples was 28 and 53, respectively, which accounted for 11.0% and 20.8% of the total.

After the field trials, fruit bruising damage was assessed based on the USDA standard (Peterson et al., 2010). The bruising test was intended to evaluate if the robotic harvesting system would have caused mechanical damage to apples during the attachment, detachment, releasing, and collection. A total of 50 harvested apples were randomly selected and evaluated for bruise damage, and each apple was visually inspected for pre-existing bruises. The pre-existing bruises were marked with a permanent black marker so that they would not be considered as fresh ones caused by the harvesting system. After all 50 apples had been kept at room temperature for 24 h, they were visually inspected for new bruises in terms of numbers of bruises and total area for each fruit. According to the USDA standard (Peterson et al., 2010), fresh apples are classified into three quality grades: Extra Fancy, Fancy, and Downgraded, with the first two being considered fresh market quality and the last being only acceptable for processing. There are four classes within the Extra Fancy grade based on the severity of bruising, and two classes in the Downgraded grade. The classification of quality grades and the evaluation results are presented in Table 3. The results showed that 100% of the tested apples were graded Extra Fancy and the percentage of Extra Fancy Class 1 (no bruising at all) accounted for 78%. The harvesting system has exceeded the industrial bruise requirement that at least 95% of the harvested apples be graded as Extra Fancy. The results demonstrated that the use of soft and resilient materials for the end-effector and dropping component (see details in Sections 2.3 and 2.4) proved particularly advantageous in reducing or eliminating bruising to apples.

4.3 | Discussion

The field evaluation has shown that the synergistic hardware and software designs of the robotic system have achieved superior performance, as measured by multiple metrics. However, further improvements to the current system are needed to ultimately meet the commercial application requirements.

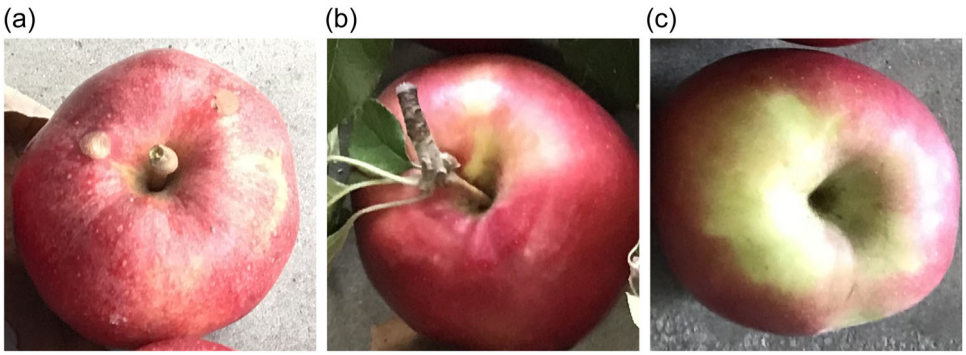

FIGURE 14 Fruit removal classifications: (a) stem-intact, (b) spur-pulled and (c) stem-pulled.

TABLE 3 Classification of apple bruise damage and evaluation results.

Class	USDA fresh market standard	Bruise specification	Results
1	Extra fancy	No bruising	39/50 (78%)
2	Extra fancy	Bruise diameter ≤3.2 mm	3/50 (6%)
3	Extra fancy	Bruise diameter 3.2 − 6.4 mm	5/50 (10%)
4	Extra fancy	Bruise diameter 6.4 − 12.7 mm or area of several bruises ≤127 mm^2	3/50 (6%)
5	Fancy	Bruise diameter 12.7 − 19 mm or total area of multiple bruises ≤283 mm^2	0/50 (0%)
6	Downgraded	Bruise larger than tolerances for "Fancy"	0/50 (0%)
7	Downgraded	Cuts or punctures of any size	0/50 (0%)

In the field tests, fruit obstruction caused by the trunk, branches, and foliage was the main cause for more than half of the failed cases. Several improvements are needed to address this issue. First, the perception system should be extended to identify obstacles around the apple, including the trunk, branches, and leaves, to provide a more comprehensive sensing of the environment. To accomplish this, an object detection algorithm based on panoptic segmentation technique (Kirillov et al., 2019) is under development. Second, based on the obstacle information, a comprehensive planning module needs to be designed to determine an optimal apple picking sequence and generate collision-free trajectories. The motion controller will then work in tandem with the collision-free trajectories to direct the manipulator toward target apples without colliding obstacles in its path to the target fruit. Various techniques, such as the traveling salesman algorithm and the potential field algorithm, are being investigated to plan the most efficient apple picking sequence and effectively harvest the apples with minimal movement and collision risks. Finally, we may need to increase the DOF for the manipulator to enable it to avoid obstacles and reach fruits with increased flexibility under a dense canopy structure.

The use of the laser-camera unit to localize fruits has shown promising results. However, further advancements are necessary to enhance its applicability and efficiency. A red laser source is used in the current design to facilitate the fruit localization since the preliminary test showed that the red laser line had higher contrasts with the background being red apples. It is unclear whether the red laser works on green and yellow apples. Further work will evaluate and compare more laser sources to ensure that the design can be applied to different apple varieties. In addition, the perception component is sensitive to intense sunlight, and thus a tarp was placed on top of the harvester during the field experiment to alleviate lighting impact. However, a more permanent, more sophisticated, or automated light shading module should be considered so as to create a more desirable lighting condition for the perception component for fruit detection and localization.

The cycle time to harvest a fruit was approximately 6 s, and reducing the cycle time would be of high priority for future work. Fruit detection and localization is one of the most time-consuming procedures, as the laser-camera unit utilizes an interval scan scheme (which requires multiple intermediate pauses) to refine the apple localization. To address this, it is essential to explore more effective scan patterns, such as implementing a continuous scan scheme, leading to a smoother and faster localization process. Additionally, optimizing the perception algorithm's flow can further expedite this procedure. Another crucial aspect to consider is the advanced planning and control techniques employed by the harvesting robot. In-depth research is needed to integrate state-of-the-art algorithms, aiming to streamline the robot's movement, minimizing unnecessary delays, and maximizing harvesting efficiency. Furthermore, a potential solution to accelerate the harvesting speed involves enhancing the system with multiple manipulators that can cooperate in fruit harvesting tasks. To achieve this goal, the hardware should be adapted to accommodate multiple manipulators while maintaining the current system's compact structure, and the software should be designed to enable seamless coordination and communication between the manipulators during the harvesting process.

5 | CONCLUSION

This paper has reported the system design and field evaluation of a novel robotic apple harvesting system. A specially designed perception component, a 4-DOF manipulator, a vacuum-based soft end-effector, and a dropping component were fully integrated as the main components of the hardware system. Effective software algorithms were designed to synergistically work with the hardware system to enable autonomous apple harvesting. Particularly, a new active laser-camera scanning scheme was developed by integrating deep learning with laser triangulation-based techniques to robustly detect and localize fruits in the challenging orchard environment. Improved designs for the end-effector and the robot's planning and control schemes further enhanced the overall performance of the robotic system. Field tests demonstrated that the robotic system was able to pick 82.4% of the target fruits in an orchard with less foliage and branch occlusions, while it successfully harvested 65.2% of the target fruits in a more challenging orchard with dense branches and foliage.

The robotic system had an average picking cycle time of around 6 s/fruit. The novel hardware and software designs have led to considerable improvements in harvesting success rates and cycle time compared with the state-of-the-art. We will continue to refine the current prototype to achieve further meaningful progress toward making a system that is practically viable and economically feasible and has an impact on the apple industry. The approaches presented in this paper could also be extended and applied to a variety of other high-value horticultural crops, and they lay a foundation for the development of fully autonomous and reliable crop harvesting systems that will lower labor costs, increase produce quality, and ultimately improve the profitability and sustainability of the horticultural industry.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. Department of Agriculture-Agricultural Research Service (No. 5050-43640-003-000D) and the National Science Foundation (No. 2024649). The findings and conclusions in this paper are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. Mention of commercial products in the paper does not imply endorsement by USDA over those not mentioned.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Arad, B., Balendonck, J., Barth, R., Ben-Shahar, O., Edan, Y., Hellström, T., et al. (2020) Development of a sweet pepper harvesting robot. *Journal of Field Robotics*, 37(6), 1027–1039. https://doi.org/10.1002/rob.21937

Bac, C.W., Van Henten, E.J., Hemming, J. & Edan, Y. (2014) Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *Journal of Field Robotics*, 31(6), 888–911. https://doi.org/10.1002/rob.21525

Baeten, J., Donné, K., Boedrij, S., Beckers, W. & Claesen, E. (2008) Autonomous fruit picking machine: a robotic apple harvester. In *Field and service robotics*, pp. 531–539.

Birrell, S., Hughes, J., Cai, J.Y. & Iida, F. (2020) A field-tested robotic harvesting system for iceberg lettuce. *Journal of Field Robotics*, 37(2), 225–245. https://doi.org/10.1002/rob.21888

Brown, J. & Sukkarieh, S. (2021) Design and evaluation of a modular robotic plum harvesting system utilizing soft components. *Journal of Field Robotics*, 38(2), 289–306. https://doi.org/10.1002/rob.21987

Bu, L., Chen, C., Hu, G., Sugirbay, A., Sun, H. & Chen, J. (2022) Design and evaluation of a robotic apple harvester using optimized picking patterns. *Computers and Electronics in Agriculture*, 198, 107092. https://doi.org/10.1016/j.compag.2022.107092

Bulanon, D.M. & Kataoka, T. (2010) Fruit detection system and an end effector for robotic harvesting of fuji apples. *Agricultural Engineering International: CIGR Journal*, 12(1), 203–210.

Chu, P., Li, Z., Lammers, K., Lu, R. & Liu, X. (2021) Deep learning-based apple detection using a suppression mask R-CNN. *Pattern Recognition Letters*, 1476, 206–211. https://doi.org/10.1016/j.patrec.2021.04.022

Chu, P., Li, Z., Zhang, K., Chen, D., Lammers, K. & Lu, R. (2023) O2RNet: occluder-occludee relational network for robust apple detection in clustered orchard environments. *Smart Agricultural Technology*, 5, 100284. https://doi.org/10.1016/j.atech.2023.100284

De Kleine, M.E. & Karkee, M. (2015) A semi-automated harvesting prototype for shaking fruit tree limbs. *Transactions of the ASABE*, 58(6), 1461–1470. https://doi.org/10.13031/trans.58.11011

Fu, H., Karkee, M., He, L., Duan, J., Li, J. & Zhang, Q. (2020) Bruise patterns of fresh market apples caused by fruit-to-fruit impact. *Agronomy*, 10(1), 59. https://doi.org/10.3390/agronomy10010059

Fu, L., Gao, F., Wu, J., Li, R., Karkee, M. & Zhang, Q. (2020) Application of consumer RGB-D cameras for fruit detection and localization in field: a critical review. *Computers and Electronics in Agriculture*, 177, 105687. https://doi.org/10.1016/j.compag.2020.105687

Gallardo, R.K. & Galinato, S.P. (2012) *2012 cost estimates of establishing, producing, and packing red delicious apples in Washington*. Washington State University Extension.

Gongal, A., Amatya, S., Karkee, M., Zhang, Q. & Lewis, K. (2015) Sensors and systems for fruit detection and localization: a review. *Computers and Electronics in Agriculture*, 116, 8–19. https://doi.org/10.1016/j.compag.2015.05.021

Hartley, R. & Zisserman, A. (2003) *Multiple view geometry in computer vision*. Cambridge University Press.

He, K., Gkioxari, G., Dollár, P. & Girshick, R. (2020) Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

Hohimer, C.J., Wang, H., Bhusal, S., Miller, J., Mo, C. & Karkee, M. (2019) Design and field evaluation of a robotic apple harvesting system with a 3D-printed soft-robotic end-effector. *Transactions of the ASABE*, 62(2), 405–414. https://doi.org/10.13031/trans.12986

Kang, H. & Chen, C. (2020) Fast implementation of real-time fruit detection in apple orchards using deep learning. *Computers and Electronics in Agriculture*, 168, 105108. https://doi.org/10.1016/j.compag.2019.105108

Khalil, H.K. (2002) *Nonlinear systems*. Upper Saddle River, NJ: Prentice-Hall.

Kirillov, A., He, K., Girshick, R., Rother, C. & Dollár, P. (2019) Panoptic segmentation. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9396–9405. https://doi.org/10.1109/CVPR41558.2019

Kjaer, K.H. & Ottosen, C.-O. (2015) 3D laser triangulation for plant phenotyping in challenging environments. *Sensors*, 15(6), 13533–13547. https://doi.org/10.3390/s150613533

Lehnert, C., English, A., McCool, C., Tow, A.W. & Perez, T. (2017) Autonomous sweet pepper harvesting for protected cropping systems. *IEEE Robotics and Automation Letters*, 2(2), 872–879. https://doi.org/10.1109/LRA.2017.2655622

Lehnert, C., McCool, C., Sa, I. & Perez, T. (2020) Performance improvements of a sweet pepper harvesting robot in protected cropping environments. *Journal of Field Robotics*, 37(7), 1197–1223. https://doi.org/10.1002/rob.21973

Lu, R., Dickinson, N., Lammers, K., Zhang, K., Chu, P. & Li, Z. (2022) Design and evaluation of end effectors for a vacuum-based robotic apple harvester. *Journal of the ASABE*, 65(5), 963–974. https://doi.org/10.13031/ja.14970

Mehta, S.S., MacKunis, W. & Burks, T.F. (2016) Robust visual servo control in the presence of fruit motion for robotic citrus harvesting. *Computers and Electronics in Agriculture*, 123, 362–375. https://doi.org/10.1016/j.compag.2016.03.007

Neupane, C., Koirala, A., Wang, Z. & Walsh, K.B. (2021) Evaluation of depth cameras for use in fruit localization and sizing: finding a successor to kinect v2. *Agronomy*, 11(9), 1780. https://doi.org/10.3390/agronomy11091780

Patil, O.S., Isaly, A., Xian, B. & Dixon, W.E. (2022) Exponential stability with RISE controllers. *IEEE Control System Letters*, 6, 1592–1597. https://doi.org/10.1109/LCSYS.2021.3127134

Peterson, D.L., Tabb, A.L., Baugher, T.A., Lewis, K. & Glenn, D.M. (2010) Dry bin filler for apples. *Applied Engineering in Agriculture*, 26(4), 541–549. https://doi.org/10.13031/2013.32057

Ren, S., He, K., Girshick, R. & Sun, J. (2017) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Schlarp, J., Csencsics, E. & Schitter, G. (2020) Optical scanning of a laser triangulation sensor for 3-D imaging. *IEEE Transactions on Instrumentation and Measurement*, 69(6), 3606–3613. https://doi.org/10.1109/TIM.2019.2933343

Siciliano, B., Sciavicco, L., Villani, L. & Oriolo, G. (2010) *Robotics: modelling, planning and control*. London: Springer-Verlag.

Silwal, A., Davidson, J.R., Karkee, M., Mo, C., Zhang, Q. & Lewis, K. (2017) Design, integration, and field evaluation of a robotic apple harvester. *Journal of Field Robotics*, 34(6), 1140–1159. https://doi.org/10.1002/rob.21715

Simonit, T. (2017) *Apple-picking robot prepares to compete for farm jobs*. MIT Technology Review.

Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E. & Liang, Z. (2019) Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and Electronics in Agriculture*, 157, 417–426. https://doi.org/10.1016/j.compag.2019.01.012

USDA. (2022) *Fresh apples, grapes, and pears: world markets and trade*. Foreign Agricultural Service. Available from: https://apps.fas.usda.gov/psdonline/circulars/fruit.pdf [accessed on 4 December 2022].

Williams, H., Jones, M.H., Nejati, M., Seabright, M.J., Bell, J., Penhall, N., et al. (2019) Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. *Biosystems Engineering*, 181, 140–156. https://doi.org/10.1016/j.biosystemseng.2019.03.007

Williams, H., Ting, C., Nejati, M., Jones, M.H., Penhall, N., Lim, J., et al. (2020) Improvements to and large-scale evaluation of a robotic kiwifruit harvester. *Journal of Field Robotics*, 37(2), 187–201. https://doi.org/10.1002/rob.21890

Xian, B., Dawson, D.M., de Queiroz, M.S. & Chen, J. (2004) A continuous asymptotic tracking control strategy for uncertain nonlinear systems. *IEEE Transactions on Automatic Control*, 49(7), 1206–1211. https://doi.org/10.1109/TAC.2004.831148

Xiong, Y., Ge, Y., Grimstad, L. & From, P.J. (2020) An autonomous strawberry-harvesting robot: design, development, integration, and field evaluation. *Journal of Field Robotics*, 37(2), 202–224. https://doi.org/10.1002/rob.21889

Zhang, K., Chu, P., Lammers, K., Li, Z. & Lu, R. (2023) Active laser-camera scanning for high-precision fruit localization in robotic harvesting: system design and calibration. *arXiv Preprint arXiv: 2311.02500*. https://arxiv.org/pdf/2311.02500.pdf

Zhang, K., Lammers, K., Chu, P., Dickinson, N., Li, Z. & Lu, R. (2022) Algorithm design and integration for a robotic apple harvesting system. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 9217–9224. https://doi.org/10.1109/IROS47612.2022.9981417

Zhang, K., Lammers, K., Chu, P., Li, Z. & Lu, R. (2021) System design and control of an apple harvesting robot. *Mechatronics*, 79(6), 102644. https://doi.org/10.1016/j.mechatronics.2021.102644

Zhang, X., He, L., Karkee, M., Whiting, M.D. & Zhang, Q. (2020) Field evaluation of targeted shake-and-catch harvesting technologies for fresh market apple. *Transactions of the ASABE*, 63(6), 1759–1771. https://doi.org/10.13031/trans.13779

Zhang, Z., Lu, Y. & Lu, R. (2021) Development and evaluation of an apple infield grading and sorting system. *Postharvest Biology and Technology*, 180(4), 111588. https://doi.org/10.1016/j.postharvbio.2021.111588

Zhao, D.-A., Lv, J., Ji, W., Zhang, Y. & Chen, Y. (2011) Design and control of an apple harvesting robot. *Biosystems Engineering*, 110(2), 112–122. https://doi.org/10.1016/j.biosystemseng.2011.07.005

Zhao, J., Tow, J. & Katupitiya, J. (2005) On-tree fruit recognition using texture properties and color data. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 263–268. https://doi.org/10.1109/IROS.2005.1545592

Zhou, H., Wang, X., Au, W., Kang, H. & Chen, C. (2022) Intelligent robots for fruit harvesting: recent developments and future challenges. *Precision Agriculture*, 23, 1–52. https://doi.org/10.1007/s11119-022-09913-3

How to cite this article: Zhang, K., Lammers, K., Chu, P., Li, Z. & Lu, R. (2024) An automated apple harvesting robot—From system design to field evaluation. *Journal of Field Robotics*, 41, 2384–2400. https://doi.org/10.1002/rob.22268