

Review

Advance of Target Visual Information Acquisition Technology for Fresh Fruit Robotic Harvesting: A Review

Yajun Li ^{1,2}, Qingchun Feng ^{1,2,*} , Tao Li ^{1,2} , Feng Xie ^{1,2}, Cheng Liu ^{1,2} and Zicong Xiong ^{1,2}

¹ Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; lyj20210043@stu.hunau.edu.cn (Y.L.); lit@nercita.org.cn (T.L.); 2111916009@stmail.ujes.edu.cn (F.X.); 2020317017@stu.sicau.edu.cn (C.L.); xzc@st.gxu.edu.cn (Z.X.)

² National Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China

* Correspondence: fengqc@nercita.org.cn; Tel.: +86-10-51503192

Abstract: In view of the continuous increase in labor costs for complex picking tasks, there is an urgent demand for intelligent harvesting robots in the global fresh fruit cultivation industry. Fruit visual information is essential to guide robotic harvesting. However, obtaining accurate visual information about the target is critical in complex agricultural environments. The main challenges include the image color distortion under changeable natural light, occlusions from the interlaced plant organs (stems, leaves, and fruits), and the picking point location on fruits with variable shapes and poses. On top of summarizing the current status of typical fresh fruit harvesting robots, this paper outlined the state-of-the-art advance of visual information acquisition technology, including image acquisition in the natural environment, fruit recognition from the complex backgrounds, target stereo locating and measurement, and fruit search among the plants. It then analyzed existing problems and raised future potential research trends from two aspects, multiple images fusion and self-improving algorithm model.

Keywords: fresh fruit; harvesting robot; visual information; stereo location



Citation: Li, Y.; Feng, Q.; Li, T.; Xie, F.; Liu, C.; Xiong, Z. Advance of Target Visual Information Acquisition Technology for Fresh Fruit Robotic Harvesting: A Review. *Agronomy* **2022**, *12*, 1336. <https://doi.org/10.3390/agronomy12061336>

Academic Editor: Simon Pearson

Received: 25 April 2022

Accepted: 28 May 2022

Published: 31 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Urgent Need of Fresh Fruit Robotic Harvesting

Fresh fruit is an excellent food source for human nutritional needs, essential for food safety, and is widely planted and produced all over the country. Fresh fruit plants can be grouped into two categories: herbaceous plants represented by tomato, strawberry, and sweet pepper, and woody plants represented by apple, citrus, and kiwi. Fresh fruit planting scale and economic benefits play an essential role in agricultural products internationally. They are an essential source for farmers to increase their income. For example, the annual global production of typical fresh fruit of tomato, citrus, apple, and strawberry is 182 million tons [1], 89 million tons [2], 86 million tons [3], and 9 million tons [4], respectively. Their planting scale and yield in China rank first globally, accounting for 34% [1], 15% [2], 46% [3], and 40% [4] of total production, respectively.

Since fresh fruit products require good eating and appearance quality, selective harvesting methods are required to ensure that mature fruits are harvested quickly and without damage. Fresh fruit harvesting involves a series of processes such as fruit maturity discrimination, fruit separation from plants, fruit collection, and transportation. It is a complex, labor-intensive, and minimally mechanized operation. With the aging population and the increasing loss of the agricultural labor force, the phenomenon of “difficult and expensive labor” in fresh fruit harvesting has become increasingly prominent, and the labor cost of harvesting has reached 30~50% of the total production cost [5–7].

Given this, the use of intelligent harvesting robots to replace or assist the manual harvesting of fresh fruits is significant for reducing the production costs and improving

economic returns [8]. As a typical representative of the agricultural robot, the fruit harvesting robot is considered to have good prospects in the future of smart agriculture and has attracted extensive attention around the world.

1.2. Target Visual Information Acquisition of Harvesting Robots

Fruit characteristics, such as color, shape, position, and posture, are prerequisites for autonomous picking, and the visual information acquired by the robot's camera is the comprehensive reflection of the fruit's reflection characteristics, imaging sensors, and lighting conditions. The fruit's image color and size are usually necessary for estimating its maturity. For example, mature fruits such as strawberries, apples, and tomatoes are relatively easy to identify from the background, where there is a significant difference in color from the background objects such as plant stems and leaves. Fruits such as cucumber, pepper, and watermelon, have similar colors to plant stems and leaves, making them difficult to recognize. Besides, based on the identification of mature fruit in the image, it is necessary to locate and measure the spatial posture of the fruit to guide the robot's end-effector achieving an expected picking operation. For fruit with long stems, such as strawberry [5], tomato [9], and sweet pepper [10], the stem is usually considered as the holding area for picking and separating the fruit from the plant to avoid damage. For fruit with short stems that are challenging to detect, such as apple [7], kiwi [11], and citrus [12], the fruit body is usually defined as the retention area to locate.

However, for the biomorphic plants in the natural environment, their fruits, leaves, and stems show the characteristics of cluster crisscross, random distribution, and mutual shielding, and their image color varies dynamically with the fluctuation of sunlight. Due to the special working conditions in an unstructured agricultural environment, the acquisition of target visual information has become one of the main bottlenecks restricting the application of harvesting robots in production [13,14].

Given the significance of visual information acquisition for harvesting robots, this paper reviews the advances in four aspects: image information acquisition in a natural environment, fruit recognition from complex background, fruit stereo location, and fruit search among the plant, based on the state-of-the-art robots for typical fruits. Additionally, the existing problems are summarized and analyzed, then the future development trend of harvesting robot's vision technology is proposed.

2. Current Status of Fresh Fruit Harvesting Robots

2.1. Typical Harvesting Robots

With the breakthrough development of basic technology theories such as artificial intelligence, deep learning, and intelligent control, the fresh fruit harvesting robot has entered a critical period from laboratorial research to industrial application. As shown in Figure 1, taking typical fresh fruits such as apple [15], strawberry [16], tomato [17], and kiwi [18] as objects, a series of commercial fresh fruit harvesting robots have been developed and applied in standardized greenhouses and orchards.

The fresh fruit harvesting robots in the greenhouse environment mainly focused on strawberry, tomato, sweet pepper, and cucumber. The widely concerned prototype robots include: The sweet pepper harvesting robot SWEEPER by Arad et al. [10] that could work in the greenhouse day and night with an average picking success rate of 61% and a picking efficiency of 24 s per fruit. The strawberry harvesting robot developed by Agrobot company [16] that realized the full-automatic picking of strawberries, with the picking efficiency of 3~5 s per fruit. The tomato harvesting robot developed by MetoMotion company [17] that was supposed to reduce labor costs by 50% with the multiple picking arms. The cucumber picking robot developed by Li et al. [19] that could identify cucumbers of similar color with plant leaves and stems, and its picking success rate is 85%, and the efficiency 8.6 s per fruit.

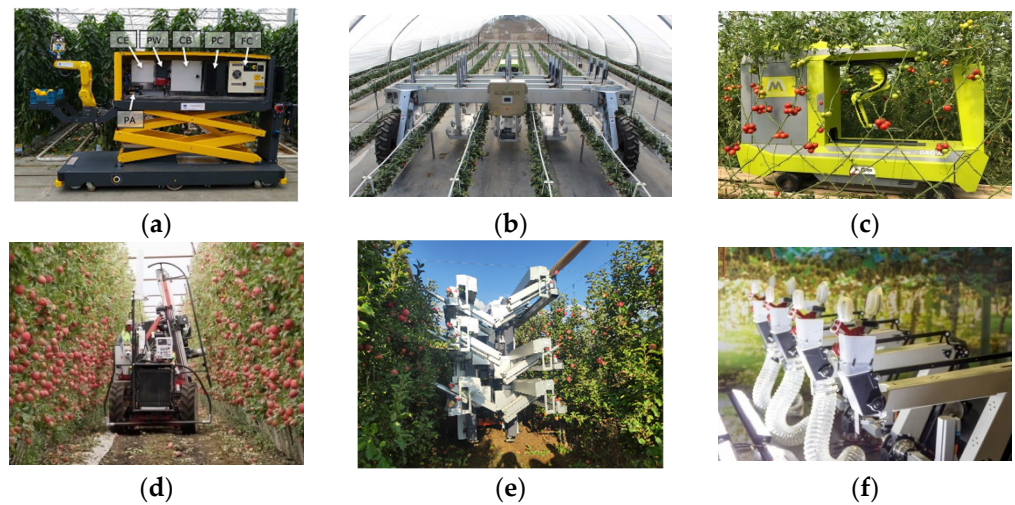


Figure 1. Typical harvesting robots. (a) Sweet pepper harvesting robot (Photo: Reprinted with permission from Ref. [10]. 2020, Arad B.); (b) strawberry harvesting robot (Photo: Reprinted with permission from Ref. [16]. 2019, Zitter L.); (c) tomato harvesting robot (Photo: Reprinted with permission from Ref. [17]. 2019, Leichman A.K.); (d) apple harvesting robot (Photo: Reprinted with permission from Ref. [15]. 2019, Thorne J.); (e) apple harvesting robot (Photo: Reprinted with permission from Ref. [20]. 2017, Dinny S.); (f) kiwi harvesting robot (Photo: Reprinted with permission from Ref. [18]. 2021, Saunders S.).

The fresh fruit harvesting robot in the orchard environment mainly took apple, kiwi, and citrus fruits as the objects. Representative robot products include: The apple harvesting robot by Abundant Robotics company [15], that had an average picking efficiency of one second per fruit for the standardized-shape trees, saving more than 60% of picking labor. The apple harvesting robot by Israel FFrobotics [20] that adopted six picking arms to achieve an efficiency of 8000 fruits per hour and a success rate of 80%. A kiwi harvesting robot by Robotics Plus [18], that had a picking efficiency of 5.5 s per fruit, and a success rate of 51%.

2.2. Characteristics of the Robot's Visual Unit

The vision system of the harvesting robots varies for different picking targets; its characteristics mainly include the imaging sensor, the specific content of the fruit visual information, the detection range, the view height, and so on. RGB-D, binocular camera, and ranging sensor are widely used for harvesting robots to obtain the color, size, position, and pose information of fruits in a specific operational area. As shown in Figure 2, RGB-D could output color images and depth point cloud data in the field of view; the binocular camera outputs two field images corresponding to the left and right cameras, respectively, and the target point's 3D coordinate in the public area of the two views calculated through image pixel matching; ranging sensors are usually used to obtain depth information of individual points based on the principle of reflection ranging.

The main characteristic parameters of some representative typical robot vision units are listed in the Table 1. Compared with the harvesting robots in a greenhouse whose target fruit exist in a small field of view, the orchard robot needs to obtain the image information from the large-scale canopy. To meet the locating needs for soft fruit stalks, except the cameras, auxiliary ranging sensors also could be used for strawberries and tomatoes.

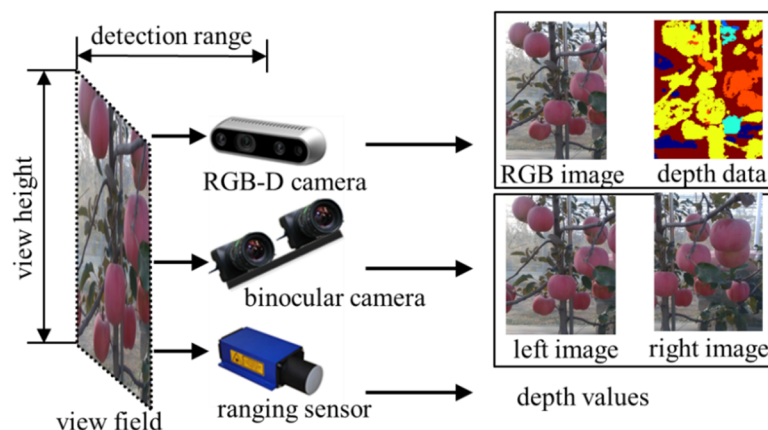


Figure 2. Typical visual units of harvesting robots.

Table 1. Characteristic list of some harvesting robots’ visual units.

Object Fruit	Sensor	Visual Information	View Field
Pepper	RGB-D [10,21], Binocular camera [22]	Fruit color [21], 3D point cloud [21], spatial coordinates [21], fruit stalk posture [21], plant main stem morphology [10]	Detection range 200~600 mm [22], Height range 1000 mm [23]
Strawberry	Laser ranging sensor [24], 3 CCD cameras [25], binocular camera [26], RGB-D [5], infrared sensor [5]	Fruit color [25], position [26], and stem posture [5]	Detection range 200~700 mm [26,27], Width range 350~670 mm [26,27], Height range 200~300 mm [28]
Tomato	Photoelectric sensor [29], binocular camera [30], laser sensor [9], RGB-D [31]	Fruit color [9], size [9] and position [30], fruit stalk posture [31]	Detection range 400~1000 mm [32], Height range 600 mm [33]
Apple	Binocular camera [34], RGB-D [35]	Fruit color [34,35], size [34,35] and position [34,35]	Detection range 1000~2000 mm [36], Height range 1000~1500 mm [7]
Citrus	Binocular camera [37], RGB-D [38]	Fruit color [37,38], size [37,38] and position [37,38]	Detection range 500~1000 mm [38], Height range 1850 mm [39]
Kiwi	Monocular camera + infrared position switch [40], binocular camera [41], RGB-D [42]	Fruit color [40], size [40] and position [40], trunk shape [43] and position [11]	Detection range 500~1000 mm [42], Visible area 3170 × 968 mm [43], 1250 mm × 1800 mm [11]

3. Image Acquisition under Agricultural Environment

3.1. Image Color Correction for Various Sunlight Conditions

Image color distortion under sunlight with temporal and spatial variation is an objective challenge for visual information acquisition in the agricultural natural environment. With the high dynamic range (HDR) [44] of the foreground target and background radiation, and the limited range of target radiation sensed by the camera under a certain exposure intensity, the exposure distortion (overexposure/underexposure) of the image needs to be corrected. The image color correction methods under the natural environment mainly involve two aspects: imaging hardware unit optimization and image data correction processing.

In terms of imaging hardware unit optimization, the imaging color is stabilized mainly through the artificial light source and camera imaging adjustment. Yuan [45] proposed a various sunlight compensation method based on the color constancy principle, which ensured the stable color of cucumber flower by dynamically adjusting the camera’s exposure gain and white balance parameters. Fu [46] set the foreground LED light to highlight the contour boundary of overlapping targets and reduce background interference for acquiring kiwi fruit images at night, and the detection accuracy was 88.3%. Zhang [47] addressed the problem of shadow and high brightness on the fruit surface through fusing the multi-view images, and the apple recognition accuracy was improved from 90.5% to 93.2%. To overcome the color distortion of the sweet pepper plant image under the background with intensive radiation, Arab [48] integrated the image under the condition of natural light and

artificial light by the Flash-No-Flash (FNF) controlled illumination unit (Figure 3), and the fruit recognition accuracy was improved by 4%.

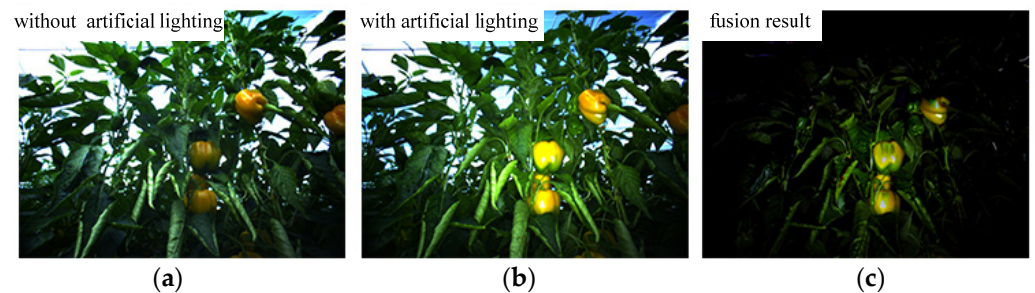


Figure 3. Color correction from images with/without artificial lighting (Photo: Reprinted with permission from Ref. [48]. 2019, Arad B.). (a) Image without artificial lighting; (b) Image with artificial lighting; and (c) Image after color correction.

In terms of the image data correction processing, the color distortion is mainly resolved according to the actual image data. Xiong et al. [49] proposed the Retinex enhancement algorithm to overcome the uneven brightness of lychee images under natural lighting conditions. Kurtulmus [50] established a neural network model for green peach recognition based on the color difference between the backlight and front-light images to weaken the influence of various light conditions. Yu [35] and Vitzrabin [51] transformed RGB images into HSV and NDI color spaces to separate image brightness channels, and then processed them based on chroma channels to improve the red apple's identification. Lv [52] used an adaptive gamma correction algorithm to correct the image color. Silwal [7] fused the optimal imaging area in different exposure images of the same view field to overcome the exposure distortion in specific image areas under natural light conditions. Feng [53] estimated the illumination radiation intensity according to four images of different exposure intensity, and then restored the image color of the global view (Figure 4).

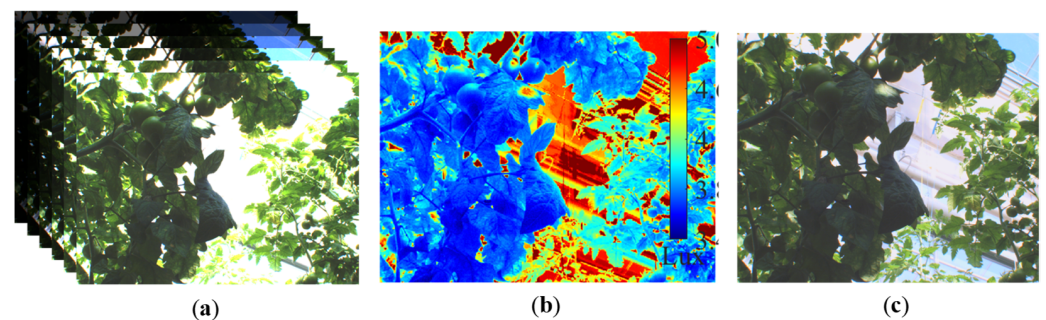


Figure 4. Color correction from multi-exposure images (Photo: Reprinted with permission from Ref. [53]. 2020, Feng Q.). (a) Multi-exposure images; (b) radiation intensity estimation; and (c) image color recovery.

3.2. Similar-Colored Target Image Acquisition

When the picking objects are cucumber, pepper, green orange, and so on, the fruit has a similar color to the plant stems and leaves. It is not easy to accurately recognize fruit targets based on the broad visible light image information. Fortunately, due to the variability in intrinsic micro-structure components such as carbohydrates, water, and fiber, the reflection characteristics of plant stems, leaves, and fruits in the specific waveband are significantly different [54].

An image is a visual representation of the spectral reflection characteristics of an object. Based on the spectral characteristics of the plant's similar-colored organs, it is sensible to obtain the optimal spectral image reflecting the similar-colored organs' microscopic

differences, so as to enrich the detection basis of the near color targets. Particularly, it is effective to select the strong reflection band as the imaging band. The target area would be highlighted in gray and background interference would be weakened. Gan et al. [55] distinguished the fruit and leaves of green orange according to the characteristics of the orange color image and thermal image (Figure 5). Bac et al. [56] established a binary tree classification model based on the images of six wavebands in the 447~900 nm range to identify the sweet pepper stems, leaves, and fruits. Li et al. [57] and Yuan et al. [58] selected 800 nm as the optimal wavelength to distinguish the cucumber fruit from the similar-colored leaves. Sa [59] fused the color image and near-infrared image as the input of the Fast R-CNN network, so that the detection accuracy of green pepper, green apple, and melon was improved from 0.816 to 0.838. Fernandez [60] constructed the visual system obtaining 635 nm, 880 nm, and depth images (Figure 6), to improve the recognition result of apple and grapefruit. Feng [61] selected 450, 600, and 900 nm bands as the optimal imaging wavelength. After the multi-band image fusion, the image grayscale difference between the fruit and background was 7.89 times that of a single-band image. Both Liu [62] and Choi [63] improved the image detection result of kiwi fruit and citrus by fusing RGB and NIR images.

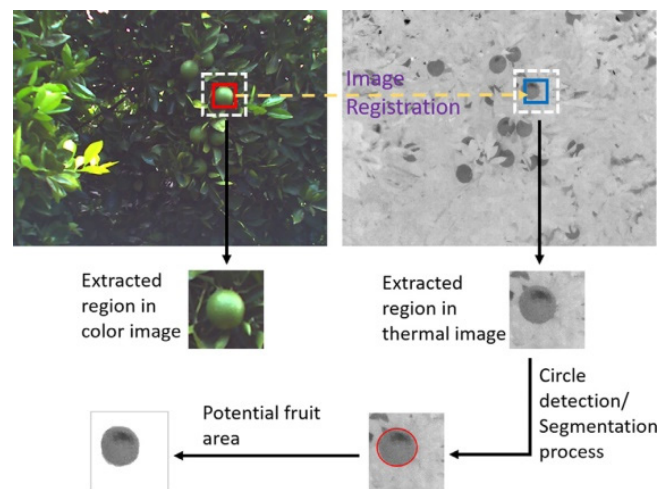


Figure 5. Green orange recognition based on a thermal infrared image (Photo: Reprinted with permission from Ref. [55]. 2018, Gan H.).

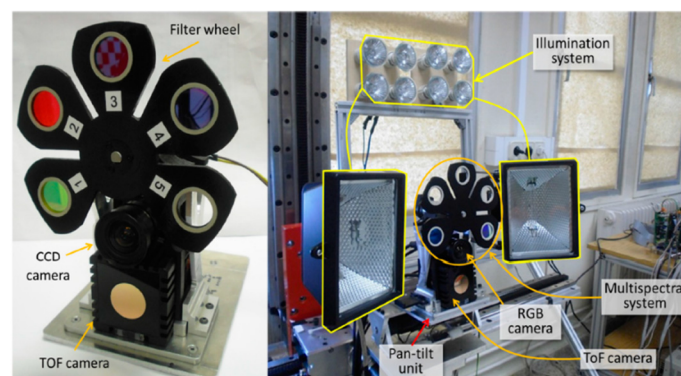


Figure 6. Multisensory system for fruit harvesting (Photo: Reprinted with permission from Ref. [60]. 2014, Fernández R.).

Some representative green fruit spectral image acquisition units are listed in Table 2.

Table 2. Visual unit for fruits in the similar-colored background.

Fruit Object	Sensor	Optimal Imaging Wavelength
Green Citrus	Color camera + thermal camera [55]	750~1400 nm [55], 827~850 nm [63]
Green Pepper	CCD camera + 6 wavelength filters [56] Multispectral camera [59]	447 nm [56], 562 nm [56], 624 nm [56], 692 nm [56], 716 nm [56], 900~1000 nm [56], 750~1400 nm [59]
Apple	Color camera + 2 band-pass interference filters [60] Thermal imager [64]	635 nm [60], 880 nm [60], 750~1400 nm [64]
Tomato	Near infrared camera + filter wheel with 3 filters [61]	450 nm [61], 600 nm [61], 900 nm [61]
Kiwi	Near infrared image from Kinect camera [63]	827~850 nm [63]

4. Fruit Target Identification from Complex Background

4.1. Visual Feature Extraction and Fusion

Because of the unstructured characteristics of agricultural objects, such as random spatial distribution, overlapping occlusion, and various shapes, it is necessary to establish an adaptive recognition model for improving the detection effect of fruit targets [65], which could integrate multiple features such as color, shape, texture, and posture. Mathematical definition of the visual features is the premise for the model. Gisellsson [66] classified the seedling leaf species of cornflower and eggplant according to the shortest Euclidean distance between the pixels on the leaf edge contour. Additionally, the highest discrimination accuracy of 97.5% was achieved with the Legendre polynomial feature set consisting of 10 numerical values. Pastrana [67] segmented the adhesive tobacco leaf through elliptical fitting. With no overlapping, the method was able to detect plants with 2~4 leaves with almost 100% accuracy. Senthilnath [68] took the green fruit in the large-scale tomato plant image by using the expectation–maximization algorithm (EM), and the segmentation accuracy reached 73.5%. Vitzrabin [69] proposed the adhesive sweet pepper fruit detection method according to the depth gradient on fruit, which obtained the true-positive rate of 0.909 under the natural illumination. Combining the normal vector with the depth data, Barnea [70] proposed the segmentation method of green sweet pepper fruit under the near color background. With the best combination of the symmetry detection and highlight-based pruning, the mean average precision reached 0.55. Rakun [71] used the Wigner Ville distribution algorithm to classify green apple and leaf features according to the texture features, obtaining at least 53% of all fruit pixels. Kurtulmusti [72] proposed a green orange recognition method based on Gabor texture analysis, and the detection accuracy reached 75.3%.

4.2. Classic Machine Learning Algorithms Application

The classic algorithm research mainly focused on application and improvement of machine learning algorithms with better computation efficiency and intelligibility. Song [73] established the support vector machine classifier (SVM) for pepper fruit recognition, combined with the most stable chroma (MSCR) and texture information, and the recognition rate was 74.2%. Ostovar [74] adopted the reinforcement learning Epsilon-Greedy algorithm to obtain the adaptive segmentation threshold, so as to improve the segmentation of yellow sweet pepper fruit under different lighting conditions. The performance with the Decaying Epsilon-Greedy algorithm reached 91.5% of the performance achieved by exhaustive search, with 73% fewer iterations than the benchmark. According to color features and shape features, Zhao [75] developed a support vector machine (SVM) algorithm with radial basis function kernel to recognize apple fruit, with the average recognition accuracy of 93.3%. Vitzrabin [51] proposed an image adaptive segmentation algorithm based on

multi-objective optimization, and the red sweet pepper recognition rate under various illumination was 87%. Lee [23] took the color parameters of hue, saturation, brightness Y, Cb, Cr as inputs of the neural network to recognize the red sweet pepper fruit with 82.16% accuracy.

4.3. Deep Learning Model Application

With the improvement in the performance of computing chips in recent years, the deep learning model [76] with a multi-layer convolution feature extraction network as the core concept has been widely used in target recognition for harvesting robots. Due to its “end-to-end” model structure and good portability to avoid the complex construction process, the model has higher recognition accuracy, and has unique advantages in the perceptual fusion of complex visual information with agricultural work objects [77,78]. According to the recognition results of the target area, the deep learning model can be categorized into the classification detection model and semantic segmentation model. The classification detection model takes the target category and its bounding box area as the output results. In addition, the semantic segmentation model can also output the target pixel area for accurate labeling.

The single-stage network model represented by the YOLO deep convolution neural network was widely used in fruit target detection. Zhao [79] and Yan [80] proposed an apple detection method based on YOLOv3 and YOLOv5, with fruit accuracy of 87.71% and 83.83%; Kounalakis et al. [81] used YOLOv3 to identify tomatoes with 98% accuracy and fruit stalks with 90% accuracy. Birrell [82] used YOLOv3 to detect cabbage in four growth stages with the overall detection accuracy of 0.91 and the classification accuracy of 0.82. Yu [83] adopted the improved rotated YOLO (R-YOLO) model, which can output the bounding box with rotating attitude parameters. Compared with the traditional YOLOv3, the harvest success rate of the harvesting robot was increased by 12%. Regarding the two-stage network model represented by fast R-CNN, Williams [11] obtained a recognition accuracy of 90.70% for kiwi calyx in a small area, but its real-time operation was 5 fps, which was not as good as the two-stage model. Combined with the characteristics of the two-type model, Kirk [84] adopted a new one-stage model RetinaNet, the recognition accuracy for mature strawberry fruit was 89.2%, and its performance was better than fast R-CNN for the same sample set. In order to further improve the real-time performance of the algorithm, Cui [85] built a lightweight model including a LeanNet backbone, feature enhancement module (P-Enhance), self-attention module, and four-scale perdition network. The recognition accuracy of green peach in the far and near view scale was 97.3%, which is better than YOLOv4 and Fast R-CNN models in terms of detection accuracy and real-time performance.

Compared with the traditional machine learning algorithm, the deep learning network model significantly improves the target detection accuracy. However, because the rectangular bounding box output by the detection network has difficulty fitting the edge of the fruit target accurately, the fruit position and posture information could not be obtained. If the central point of the bounding box is directly used as the central point of the fruit, there will be a significant positioning error. Therefore, the target segmentation network is a better choice to obtain accurate fruit target pose information.

Williams [43] proposed a semantic segmentation method for the kiwi calyx region based on the FCN-8S complete convolution network (Figure 7), and the detection accuracy of dense fruits reached 79.0%; to maintain the clarity of the segmentation results, Zhang [86] used the Deeplab v3 network to segment multiple targets such as branches, trunks, leaves, and apples, and the per-class accuracy was 97%. To meet the segmentation and recognition needs of individual targets of overlapping fruits, Yu [87] and Jia [88] used the instance segmentation model Mask R-CNN to identify the overlapping strawberry and apple fruits, which could determine not only categories but also individuals. To improve the time efficiency of multiple category target segmentation, Kang [34] proposed a multi-task deep neural network DaSNet, which could realize the semantic segmentation on the branch and

instance segmentation on apple fruit, as shown in Figure 8. Compared with mask R-CNN, the accuracy of fruit recognition was equal, but the detection time was reduced by 50%.

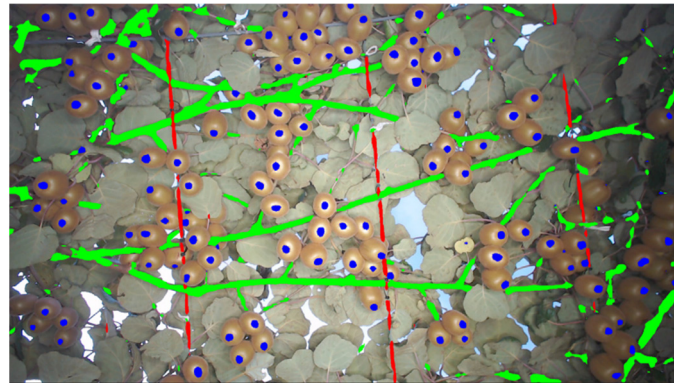


Figure 7. Semantic segmentation for kiwifruit's calyx (Photo: Reprinted with permission from Ref. [43]. 2019, Williams H.A.).



Figure 8. Instance segmentation for apple fruit (Photo: Reprinted with permission from Ref. [34]. 2019, Kang H.). (a) The original image; (b) the instance segmentation result.

5. Fruit's Stereo Location and Measurement

Based on the 2D image information recognition and segmentation of the fruit target area in the field of view, obtaining the 3D pose information of the target fruit is essential for guiding the robot operation. The positioning of the fruit picking operation part by the robot is mainly divided into three steps: the measurement of the 3D coordinates of the target area; the dynamic servo alignment control in the process of target approaching; and the accurate measurement of the pose parameters of the picking target area in the close-up field of view. The specific implementation methods of each step require selecting the appropriate hardware and algorithm model of the stereo measurement unit in combination with the requirements of the picking operation and the target characteristics.

5.1. Hardware Unit of Stereo Vision

Stereo vision technology is the most widely used stereo measurement method for fruit harvesting robots. Three products are relevant according to the measurement principle, including the RGB-D depth camera, dual/multi-camera, and structured light vision systems (monocular camera + optical auxiliary components). Wang [32] and Yu [35] built a fixed-mounted binocular vision system to measure the spatial coordinates of the fruit center according to the matching relationship between the left and right camera images. Kaczmarek [89] developed a five-eye stereo vision system (Figure 9a) and proposed a parallax map synthesis method based on error matching point exclusion (EEMM) to improve the measurement accuracy of the vision system. On top of obtaining the point coordinates, in order to further measure the spatial pose information of the fruit area, Ling [30] and

Xiang [90] reconstructed the discrete point coordinates obtained by the binocular camera-based on the parallax constraints and the grayscale feature matching of the target area to form the point cloud data of the target area in the camera field of view. In view of the high precision and fast response of the laser radar, Si [91] and Eizentals [92] matched and fused the laser detection depth image with the color image to measure the 3D shape and spatial attitude of the fruit (Figure 9b). Gongal [93] fused 2D image and PMD CamCube3D camera depth image data to locate apple targets spatially in a large field of view (Figure 9c). Feng [94] built a vision system (Figure 9d) composed of a single camera and line structured light to locate the overlapping fruit according to the image morphology of the fruit surface stripes to reduce the redundant information of the laser radar field of view. With the performance improvement and cost reduction of depth camera products in recent years, RGB-D cameras represented by Microsoft Kinect, Intel RealSense, and LIPS LIPSedge have become the best choice for picking robots. Lehnert [95] can realize the comprehensive measurement of fruit size, posture, and surface contour curvature by fitting the 3D geometric shape of fruit and solving the surface normal vector according to the RGB and depth point cloud information automatically aligned and matched by the RGB-D camera.

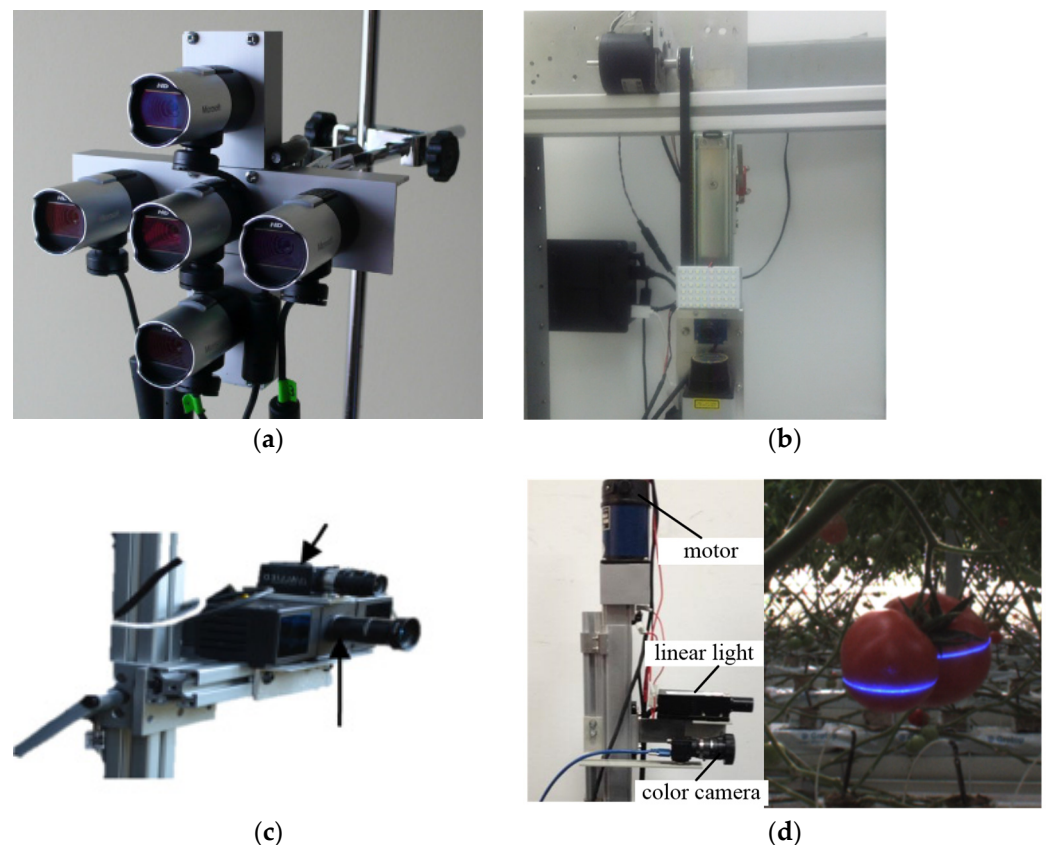


Figure 9. Stereo vision system for fruit harvesting. (a) Multiple cameras set for obtaining depth maps of plants (Photo: Reprinted with permission from Ref. [89]. 2017, Kaczmarek A.L.); (b) camera and laser sensor combined application for pepper 3D pose estimation (Photo: Reprinted with permission from Ref. [92]. 2016, Eizentals P.); (c) 2D and 3D cameras combined application for apple fruit yield estimation (Photo: Reprinted with permission from Ref. [93]. 2016, Gongal A.); (d) structured light stripe vision unit for overlapped tomato fruits (Photo: Reprinted with permission from Ref. [94]. 2014, Feng Q.).

In addition, according to the size of the image field of view obtained by the robot vision system, the vision unit can be divided into three types: distant view, close view, and distant–close view combination. The close-range vision system usually installs the camera at the end of the manipulator in the form of the eye-in-hand to collect the close range and minor field images of plants. This method has high accuracy for fruit pose positioning, but the effective detection range is small. Lehnert [21] can measure features such as the surface curvature of the sweet pepper and the posture of the fruit stem through the RGB-D camera installed at the end of the manipulator. The distant-range vision system usually fixes the camera on the robot in the form of eye-to-hand to collect the plant distant-range image with a large field of view, which has a sizeable effective detection range and low accuracy for fruit pose measurement. For example, Yu [35] used the distant-range binocular vision system to locate the center point of the apple which approximates a sphere. On this basis, Feng [96] developed a vision system for strawberry harvesting robots with a combination of distant and close views, reducing the long-range positioning error by 19.53% on the premise of obtaining a large operation field of view.

5.2. Measurement of Position and Posture

In view of the differences between various fruit picking methods, the 3D pose information that the robot needs to obtain is also different. For apple [20], kiwi [18], citrus [12], and other fruits with short-stalked approximately spherical fruits, it is usually only necessary to obtain the spatial coordinates of the central area of the fruit to guide the manipulator to operate. It is challenging to locate the center point of fruit by fitting the contour of sticky and occluded fruit based on the visible region. Nguyen [97] studied the apple pixel clustering segmentation method based on pixel Euclidean distance for apple tree images collected by an RGB-D camera. Xiang [90] generated a tomato plant depth map based on parallax constraints and grayscale feature matching of the target area. Through OTSU threshold segmentation and least squares contour fitting of the depth map, the accuracy of single fruit recognition in clustered lychee fruit clusters was 87.9%. Kang [98] proposed a recognition and grasping estimation method based on the PointNet model according to the RGB-D point cloud data in the visible area of apple fruit, and the success rate of harvesting reached 0.8.

However, for soft-skinned fruits such as tomato, strawberry, and sweet pepper, it is usually necessary to obtain the spatial pose information of the fruit and the stems to guide the manipulator to separate the fruit from the stems. Xiong [5] used an IR sensor to scan the fruit (Figure 10a), and Yu [83] established the R-YOLO fruit detection model (Figure 10b) to determine the inclined posture of the central axis of strawberry fruit, respectively, to improve the positioning accuracy of the cutting point of the fruit stem on the central axis of the fruit. Eizentals [92] matched and fused the depth image obtained by laser radar with the color image to locate the growth posture of pepper fruit. Lehnert [21], according to the RGB-D point cloud data (Figure 10c) of the sweet pepper fruit and stem area, as well as the normal vector gradient characteristics of the fruit surface, determined the optimal posture of the picking sucker for fruit adsorption and the cutting point of the stem. Kounalakis [81] positioned the tomato fruit using the distant view image, and further localized the fruit stalk through a close-up camera mounted on the picking claw, which guided the picking claw to precisely hold the fruit stalk. To improve the obstacle avoidance performance of robot picking operation, Bac [22] further obtained the spatial shape of the main plant stem based on the positioning of sweet pepper fruit and the stem and obtained the optimal operation posture of the fruit stalk from multiple perspectives (Figure 10d).

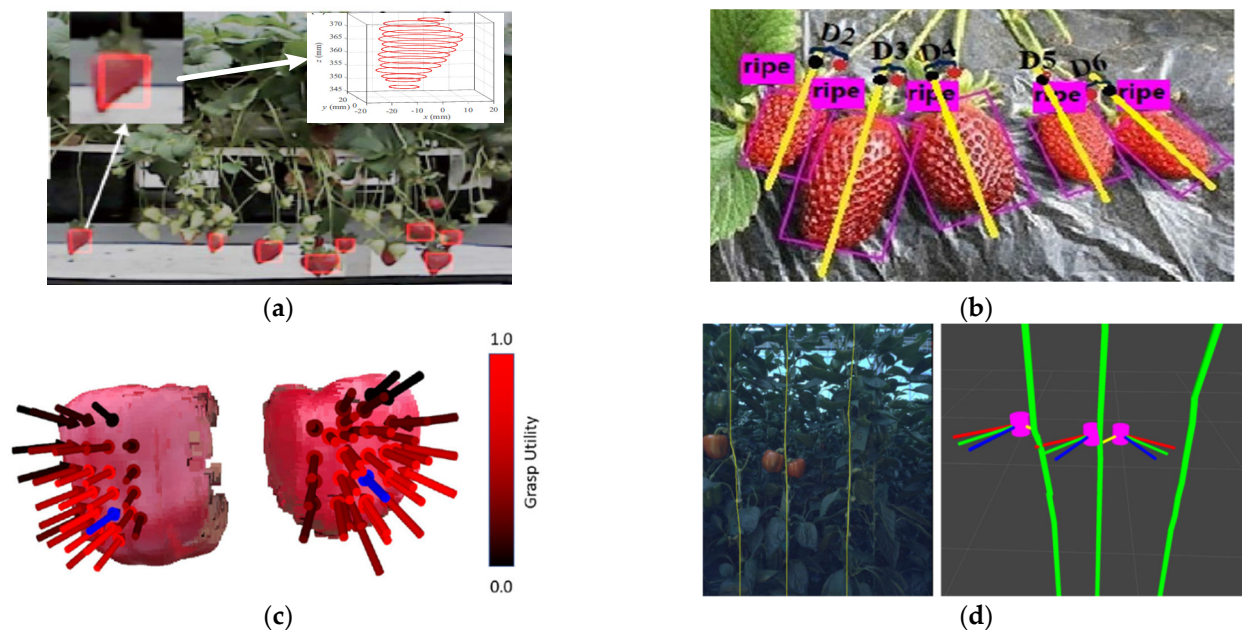


Figure 10. Method for obtaining spatial pose information of the fruit. (a) Shape reconstruction of strawberry fruit scanned by IR sensor (Photo: Reprinted with permission from Ref. [5]. 2017, Lehnert C.); (b) picking point localization in strawberry's inclined stem (Photo: Reprinted with permission from Ref. [83]. 2020, Yu Y.); (c) grasp pose estimation according to pepper's surface gradient (Photo: Reprinted with permission from Ref. [21]. 2020, Lehnert C.); (d) grasp pose estimation according to the position relationship between the main-stem and pepper fruit (Photo: Reprinted with permission from Ref. [22]. 2017, Bac C.W.).

6. Disordered Fruits Search from Plants

Plant stems, leaves, and fruits grow alternately and randomly distributed. Efficient search and detection of randomly distributed fruit targets in tall plants is of great significance for improving the effective working space of robots and improving work efficiency. The fruit target search methods are mainly categorized into passive detection and active detection.

6.1. Passive Detection with Fixed View Field

Passive detection means that the camera obtains the image of the picking area with a fixed attitude to distinguish and locate the fruit within the characteristic field of view. It is usually used in robot operation scenarios with large object distance space or high fruit density, such as apple, kiwi, and strawberry. For example, Zhang [6], Silwal [7], and Yu [35] used fixed cameras to collect the global image information of the robot operation area; Williams [43] collected images of dense fruit of kiwifruit cultivated on trellises with four sets of RGB-D depth cameras upward to realize the identification and positioning of the calyx; Xiong [5] and Feng [96] obtained the image information of strawberry fruits cultivated overhead with a fixed perspective camera, and determined the picking order of the fruits accordingly. The passive detection vision system is simple in configuration and easy to use, but the detection range is limited. It is difficult to cover the fruit distribution area accurately, and the images randomly collected from a fixed perspective will contain much redundant information.

6.2. Active Detection with Multiple View Field

Active detection refers to the identification and positioning of fruits in different field of view areas by a single camera with a moving posture or multiple cameras with different postures. It is usually used in operation scenes where it is difficult to obtain a large field of view image due to small object distance, or the distribution of fruits is sparse, such as

tomato and sweet pepper. The active detection vision system needs to integrate the image information from the multiple fields of view. The system configuration and processing operation are relatively complex, but it can realize the active search and detection of the target, which is conducive to expanding the operation range of the robot. At the same time, the directional search based on the fruit distribution law can reduce the robot's acquisition of redundant interference information.

Firstly, the active multi-view detection for a single fruit target can improve the accuracy of the picking operation, as shown in Figure 10. Yamamoto [28], based on the strawberry fruit recognition in the field of view of the distant-range camera, further determined the posture of the fruit stem by the close-range camera at the end of the manipulator to improve the operation accuracy of the fruit stem. Lehnert [95] installed the camera at the end of the manipulator and realized the search for fruits distributed in different areas by controlling the scanning movement according to a predetermined trajectory (Figure 11a). Lehnert [99] further controlled the movement of the manipulator dynamically according to the dynamic relationship between the occlusion of fruits in the dynamic field of view of the camera, and the posture of the manipulator to maximize the visible area of fruits in the field of view of the robot (Figure 11b). Mehta [12,100] adopted the camera in hand (CIH) configuration vision system to dynamically control the movement trajectory of the manipulator according to the coordinate changes in the citrus fruit in the dynamic field of view of the camera, so as to realize the directional harvesting movement of the fruit. To avoid the interference of the picking claws on the main stem of the plant, Bac [80], Arad [10], Barth [101], and Hemming [102] et al. collected the images of sweet pepper fruits from multiple perspectives through the camera installed at the end of the manipulator, to determine the optimal fruit clamping strategy. The success rate of the relative passive detection method was increased from 14% to 52% [101].

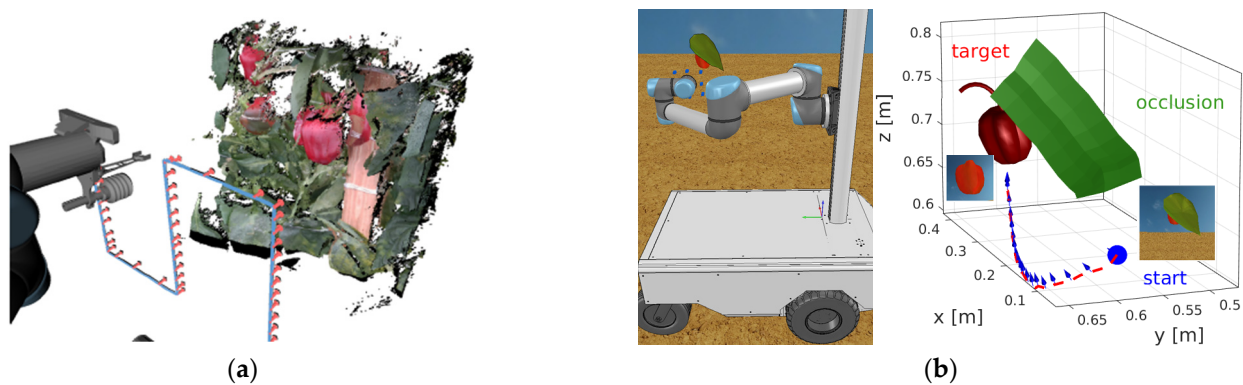


Figure 11. Active multi-view detection for a single fruit target. (a) Scanning trajectory for determining the pepper fruit from plants (Photo: Reprinted with permission from Ref. [95]. 2017, Lehnert C.); (b) multi-perspective visual observation for occluded fruit (Photo: Reprinted with permission from Ref. [99]. 2018, Lehnert C.).

Secondly, the multi-view active search guided by the growth form of the plant stem can improve the operation range and efficiency of the picking robot, as shown in Figure 11. Bac [103] took the cultivated nylon suspension wire with high brightness reflection characteristics as the visual mark to identify and locate the sweet pepper stem wrapped around it. Amatya [104] used the method of combining straight lines and exponential curves to fit the overall shape of the cherry stem (Figure 12a,b). Li [105] fitted the shape of the main lychee stem with dense branches and leaves through discrete target spatial coordinate clustering based on target segmentation and recognition of trunk and fruit. Feng [106] proposed a multi-view image tracking acquisition method constrained by the main stem of a tomato, which realizes the directional search of discrete targets such as fruits, side branches, and inflorescence (Figure 12c).

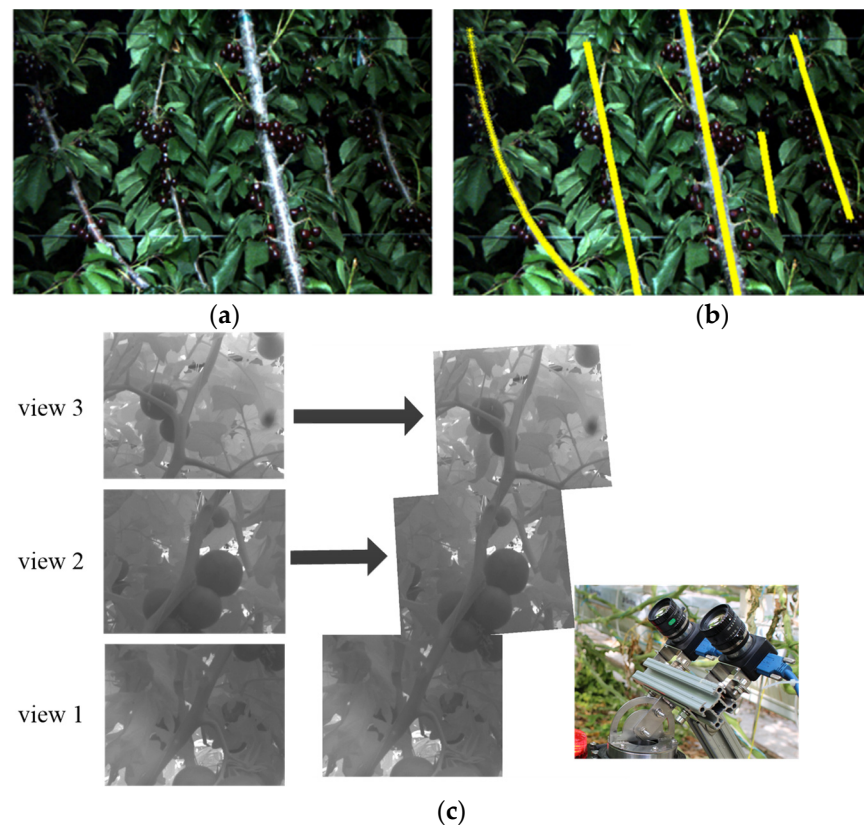


Figure 12. The multi-view active search guided by the growth form of plant stem. (a,b) Detection of cherry tree branches for automated sweet-cherry harvesting (Photo: Reprinted with permission from Ref. [105]. 2020, Li J.); (c) tomato plant main-stem tracking and measuring based on binocular pan-tilt vision unit (Photo: Reprinted with permission from Ref. [106]. 2020, Feng Q.).

7. Challenges and Trends

7.1. Challenge Summaries

The radiation intensity of foreground plants and background sky within the robot's view field presents a high dynamic range in the open agricultural environment, but the camera has a limited range of radiometric imaging at a particular exposure setting, which would result in exposure distortion (overexposure/underexposure). The existing research was mainly aimed at the unsaturated image color correction under specific exposure conditions, which lacked applicability for the exposure saturation distortion under natural light. In terms of similar-colored target identification, the current research mainly used the strong reflection band as the imaging band, which usually not only emphasized the image brightness of the target area but also the background. As the weak reflection image fusion is lacking, the background brightness cannot be effectively suppressed, which makes it not entirely valid to highlight the target and weaken the background interference.

The deep learning model with the multi-layer convolution feature extraction network has unique advantages for the perceptual fusion of multiple visual information of agricultural objects [77,78]. However, the current fruit recognition model was mainly obtained by off-line training, and the constant weight model was applied to target detection and segmentation. Therefore, the robustness of the model usually depended on the training data set. For biomorphic plants in the agricultural environment, the variation in environment, growth stage, tree shape, and imaging sensor all lead to the diversity of image information. Limited data sets were usually challenging to ensure the off-line trained model has broad applicability.

The fruits supported by branches/stem are distributed in the tall plants. The detection range and efficiency of the robot directly determine the performance. It is an effective

way to improve the working space and efficiency of the harvesting robot by detecting the discrete fruit individually from multi-view and multi-scale fields of view. However, the current research on cluster target detection mainly focused on the passive detection under fixed perspectives and scales. There was less research on active detection according to the distribution of target groups, limiting the robot's effective workspace. In addition, during the picking process, the branches/stems inevitably swung by external forces, which would shift the spatial posture of the fruits. For example, the picking operation of one manipulator usually caused the shaking of the picking target of other manipulators, especially for the multi-arm picking robot. In the current research, the harvesting robot usually assumed that the pose of the target was stationary when obtaining the close-up image information of the fruit. According to this, using the picking end-effector achieved to hold the fruit, there may be a significant positioning error or even picking failures.

7.2. Potential Trends

Since the visual features obtained under complex agricultural conditions are usually incomplete, it is an effective approach to acquire and fuse multiple visual information to improve the decision-making performance of the harvesting robot. For example, restoring the image color according to the multi-exposure image, measuring the fruit posture according to the distant–close combined view images, selecting the appropriate multi-band spectral image to highlight the difference between the similar-colored targets, combining the camera and structured light (laser, infrared, and visible light) to improve the target positioning accuracy, and determining the picking end-effector's posture according to the fruit and the main stem multi-target image, and so on.

Working with the biomorphic organisms in the natural environment, the robot's active learning and self-renewal for picking object feature recognition is necessary to ensure the practicability of intelligent picking operations. Reinforcement learning technology is actively learning according to task changes and reward feedback for providing generalization performance, which will help to improve the pick robots' visual information perception ability in the natural environment. With the popular application of 5G technology, the real-time updated model for fruit detection based on a cloud computing platform [107] will provide an essential guarantee for the online reinforcement learning, to improve the visual information perception model of the harvesting robot.

8. Conclusions

As the key component, the perception range and accuracy of the visual unit directly determined the picking robot's working space and successful harvesting rate. Due to the unique agricultural conditions, the research advances on visual information acquisition technology mainly focused on stable imaging, feature recognition, and pose measurement. As the recent in-depth application of AI algorithms and chips, the deep convolution network model had significant advantages for fruit target recognition. The performance of RGB-D and laser sensor products were continuously improved, which effectively reduced the cost and structural complexity of the robot's vision system.

However, there are still many challenges in the visual information acquisition for robotic harvesting, which is one of the common technical problems restricting the commercial application of harvesting robots. Facing the biomorphic plant groups, a visual information perception model with self-learning and self-renewal ability is necessary to ensure that the robots adapt well to different working objects. In addition, the fruit target active search along the plant growth morphology is essential for expanding the robot's workspace, obtaining multi-view images of fruits, and planning the robot's obstacle avoidance path.

Author Contributions: Conceptualization, Y.L. and Q.F.; methodology, Y.L. and Q.F.; formal analysis, Q.F., T.L. and C.L.; investigation, Y.L., Q.F. and Z.X.; resources, Y.L. and Q.F.; data curation, Y.L., Q.F. and F.X.; writing—original draft preparation, Y.L. and Q.F.; writing—review and editing, Y.L., T.L., F.X., C.L. and Z.X.; project administration, Q.F.; funding acquisition, Q.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Beijing Science and Technology Plan Project (grant number Z201100008020009), National Key Research and Development Plan Project (grant number 2019YFE0125200), BAAFS Innovation Capacity Building Project (KJCX20210414) and China Agriculture Research System of MOF and MARA (CARS—20).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tomato Production 2018. Available online: <https://ourworldindata.org/grapher/tomato-production> (accessed on 20 April 2022).
2. Crops and Livestock Products. Available online: <https://www.fao.org/faostat/en/#data/QCL> (accessed on 20 April 2022).
3. Apple Production 2018. Available online: <https://ourworldindata.org/grapher/apple-production> (accessed on 20 April 2022).
4. Strawberry. Available online: <https://en.wikipedia.org/wiki/Strawberry> (accessed on 20 April 2022).
5. Xiong, Y.; Ge, Y.; Grimstad, L.; From, P.J. An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation. *J. Field Robot.* **2020**, *37*, 202–224. [CrossRef]
6. Zhang, K.; Lammers, K.; Chu, P.; Li, Z.; Lu, R. System design and control of an apple harvesting robot. *Mechatronics* **2021**, *79*, 102644. [CrossRef]
7. Silwal, A.; Davidson, J.R.; Karkee, M.; Mo, C.; Zhang, Q.; Lewis, K. Design, integration, and field evaluation of a robotic apple harvester. *J. Field Robot.* **2017**, *34*, 1140–1159. [CrossRef]
8. King, A. Technology: The Future of Agriculture. *Nature* **2017**, *544*, S21–S23. [CrossRef]
9. Feng, Q.; Zou, W.; Fan, P.; Zhang, C.; Wang, X. Design and test of robotic harvesting system for cherry tomato. *Int. J. Agric. Biol. Eng.* **2018**, *11*, 96–100. [CrossRef]
10. Arad, B.; Balendonck, J.; Barth, R.; Ben-Shahar, O.; Edan, Y.; Hellström, T.; Hemming, J.; Kurtser, P.; Ringdahl, O.; Tielen, T.; et al. Development of a sweet pepper harvesting robot. *J. Field Robot.* **2020**, *37*, 1027–1039. [CrossRef]
11. Williams, H.; Ting, C.; Nejati, M.; Jones, M.H.; Penhall, N.; Lim, J.; Seabright, M.; Bell, J.; Ahn, H.S.; Scarfe, A.; et al. Improvements to and large-scale evaluation of a robotic kiwifruit harvester. *J. Field Robot.* **2020**, *37*, 187–201. [CrossRef]
12. Mehta, S.S.; Burks, T.F. Vision-based control of robotic manipulator for citrus harvesting. *Comput. Electron. Agric.* **2014**, *102*, 146–158. [CrossRef]
13. Wang, Z.H.; Xun, Y.; Wang, Y.K.; Yang, Q.H. Review of smart robots for fruit and vegetable picking in agriculture. *Int. J. Agric. Biol. Eng.* **2022**, *15*, 33–54. [CrossRef]
14. Tang, Y.; Chen, M.; Wang, C.; Luo, L.; Li, J.; Lian, G.; Zou, X. Recognition and Localization Methods for Vision-Based Fruit Picking Robots: A Review. *Front. Plant Sci.* **2020**, *11*, 510. [CrossRef]
15. Thorne, J. Apple-Picking Robots Gear up for U.S. Debut in Washington State. Available online: <https://www.geekwire.com/2019/apple-picking-robots-gear-u-s-debut-washington-state/> (accessed on 20 April 2022).
16. Zitter, L. Berry Picking at Its Best with AGROBOT Technology. Available online: <https://www.foodandfarmingtechnology.com/news/harvesting-technology/berry-picking-at-its-best-with-agrobot-technology.html> (accessed on 20 April 2022).
17. Leichman, A.K. World's First Tomato-Picking Robot Set to Be Rolled Out. Available online: <https://www.israel21c.org/israeli-startup-develops-first-tomato-picking-robot> (accessed on 20 April 2022).
18. Saunders, S. The Robots That Can Pick Kiwi-Fruit. Available online: <https://www.bbc.com/future/ bespoke/ follow-the-food/ the-robots-that-can-pick-kiwifruit.html> (accessed on 20 April 2022).
19. Ji, C.; Feng, Q.C.; Yuan, T.; Tan, Y.Z.; Li, W. Development and performance analysis on cucumber harvesting robot system in greenhouse. *Robot* **2011**, *33*, 726–730. [CrossRef]
20. The Latest on FF Robotics' Machine Harvester. Available online: <https://basinbusinessjournal.com/news/2021/apr/12/machine-picked-apples/> (accessed on 20 April 2022).
21. Lehnert, C.; McCool, C.; Sa, I.; Perez, T. Performance improvements of a sweet pepper harvesting robot in protected cropping environments. *J. Field Robot.* **2020**, *37*, 1197–1223. [CrossRef]
22. Bac, C.W.; Hemming, J.; Van Tuijl, B.A.J.; Barth, R.; Wais, E.; van Henten, E.J. Performance evaluation of a harvesting robot for sweet pepper. *J. Field Robot.* **2017**, *34*, 1123–1139. [CrossRef]
23. Lee, B.; Kam, D.; Min, B.; Hwa, J.; Oh, S. A Vision Servo System for Automated Harvest of Sweet Pepper in Korean Greenhouse Environment. *Appl. Sci.* **2019**, *9*, 2395. [CrossRef]

24. Han, K.S.; Kim, S.C.; Lee, Y.B.; Kim, S.C.; Im, D.H.; Choi, H.K.; Hwang, H. Strawberry harvesting robot for bench-type cultivation. *J. Biosyst. Eng.* **2012**, *37*, 65–74. [[CrossRef](#)]
25. De Preter, A.; Anthonis, J.; De Baerdemaeker, J. Development of a robot for harvesting strawberries. *IFAC-PapersOnLine* **2018**, *51*, 14–19. [[CrossRef](#)]
26. Feng, Q.; Wang, X.; Zheng, W.; Qiu, Q.; Jiang, K. New strawberry harvesting robot for elevated-trough culture. *Int. J. Agric. Biol. Eng.* **2012**, *5*, 1–8. [[CrossRef](#)]
27. Xiong, Y.; Peng, C.; Grimstad, L.; From, P.J.; Isler, V. Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. *Comput. Electron. Agric.* **2019**, *157*, 392–402. [[CrossRef](#)]
28. Yamamoto, S.; Hayashi, S.; Saito, S.; Ochiai, Y.; Yamashita, T.; Sugano, S. Development of robotic strawberry harvester to approach target fruit from hanging bench side. *IFAC Proc. Vol.* **2010**, *43*, 95–100. [[CrossRef](#)]
29. Kondo, N.; Yata, K.; Iida, M.; Shiigi, T.; Monta, M.; Kurita, M.; Omori, H. Development of an end-effector for a tomato cluster harvesting robot. *Eng. Agric. Environ. Food* **2010**, *3*, 20–24. [[CrossRef](#)]
30. Ling, X.; Zhao, Y.; Gong, L.; Liu, C.; Wang, T. Dual-arm cooperation and implementing for robotic harvesting tomato using binocular vision. *Robot. Auton. Syst.* **2019**, *114*, 134–143. [[CrossRef](#)]
31. Fujinaga, T.; Yasukawa, S.; Ishii, K. Development and Evaluation of a Tomato Fruit Suction Cutting Device. In Proceedings of the 2021 IEEE/SICE International Symposium on System Integration (SII), Fukushima, Japan, 11–14 January 2021; pp. 628–633. [[CrossRef](#)]
32. Wang, L.L.; Zhao, B.; Fan, J.; Hu, X.; Wei, S.; Li, Y.; Zhou, Q.; Wei, C. Development of a tomato harvesting robot used in greenhouse. *Int. J. Agric. Biol. Eng.* **2017**, *10*, 140–149. [[CrossRef](#)]
33. Feng, Q.; Wang, X.; Wang, G.; Li, Z. Design and test of tomatoes harvesting robot. In Proceedings of the 2015 IEEE International Conference on Information and Automation, Lijiang, China, 8–10 August 2015; pp. 949–952. [[CrossRef](#)]
34. Kang, H.; Zhou, H.; Chen, C. Visual perception and modeling for autonomous apple harvesting. *IEEE Access* **2020**, *8*, 62151–62163. [[CrossRef](#)]
35. Yu, X.; Fan, Z.; Wang, X.; Wan, H.; Wang, P.; Zeng, X.; Jia, F. A lab-customized autonomous humanoid apple harvesting robot. *Comput. Electr. Eng.* **2021**, *96*, 107459. [[CrossRef](#)]
36. Chu, P.; Li, Z.; Lammers, K.; Lu, R.; Liu, X. Deep learning-based apple detection using a suppression mask R-CNN. *Pattern Recognit. Lett.* **2021**, *147*, 206–211. [[CrossRef](#)]
37. Wang, Y.; Yang, Y.; Yang, C.; Zhao, H.; Chen, G.; Zhang, Z.; Fu, S.; Zhang, M.; Xu, H. End-effector with a bite mode for harvesting citrus fruit in random stalk orientation environment. *Comput. Electron. Agric.* **2019**, *157*, 454–470. [[CrossRef](#)]
38. Hu, X.; Yu, H.; Lv, S.; Wu, J. Design and experiment of a new citrus harvesting robot. In Proceedings of the 2021 International Conference on Control Science and Electric Power Systems (CSEPS), Shanghai, China, 28–30 May 2021; pp. 179–183. [[CrossRef](#)]
39. Yang, C.; Liu, Y.; Wang, Y.; Xiong, L.; Xu, H.; Zhao, W. Research and Experiment on Recognition and Location System for Citrus Picking Robot in Natural Environment. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 14–22. [[CrossRef](#)]
40. Zhang, F.; Li, Z.; Wang, B.; Su, S.; Fu, L.; Cui, Y. Study on recognition and non-destructive picking end-effector of kiwifruit. In Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 2174–2179. [[CrossRef](#)]
41. Barnett, J.; Duke, M.; Au, C.K.; Lim, S.H. Work distribution of multiple Cartesian robot arms for kiwifruit harvesting. *Comput. Electron. Agric.* **2020**, *169*, 105202. [[CrossRef](#)]
42. Mu, L.; Cui, G.; Liu, Y.; Cui, Y.; Fu, L.; Gejima, Y. Design and simulation of an integrated end-effector for picking kiwifruit by robot. *Inf. Process. Agric.* **2020**, *7*, 58–71. [[CrossRef](#)]
43. Williams, H.A.; Jones, M.H.; Nejati, M.; Seabright, M.J.; Bell, J.; Penhall, N.D.; Barnett, J.J.; Duke, M.D.; Scarfe, A.J.; Ahn, H.S.; et al. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. *Biosyst. Eng.* **2019**, *181*, 140–156. [[CrossRef](#)]
44. Debevec, P.E.; Malik, J. Recovering high dynamic range radiance maps from photographs. In Proceedings of the SIGGRAPH '08: Special Interest Group on Computer Graphics and Interactive Techniques, Los Angeles, CA, USA, 11–15 August 2008; pp. 1–10. [[CrossRef](#)]
45. Yuan, T.; Kondo, N.; Li, W. Sunlight fluctuation compensation for tomato flower detection using web camera. *Procedia Eng.* **2012**, *29*, 4343–4347. [[CrossRef](#)]
46. Fu, L.S.; Wang, B.; Cui, Y.J.; Su, S.; Gejima, Y.; Kobayashi, T. Kiwifruit recognition at nighttime using artificial lighting based on machine vision. *Int. J. Agric. Biol. Eng.* **2015**, *8*, 52–59. [[CrossRef](#)]
47. Zhang, K.; Lammers, K.; Chu, P.; Dickinson, N.; Li, Z.; Lu, R. Algorithm Design and Integration for a Robotic Apple Harvesting System. *arXiv* **2022**, arXiv:2203.00582. [[CrossRef](#)]
48. Arad, B.; Kurtser, P.; Barnea, E.; Harel, B.; Edan, Y.; Ben-Shahar, O. Controlled Lighting and Illumination-Independent Target Detection for Real-Time Cost-Efficient Applications. The Case Study of Sweet Pepper Robotic Harvesting. *Sensors* **2019**, *19*, 1390. [[CrossRef](#)]
49. Xiong, J.; Zou, X.; Wang, H.; Peng, H.; Zhu, M.; Lin, G. Recognition of ripe litchi in different illumination conditions based on Retinex image enhancement. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 170–178. [[CrossRef](#)]
50. Kurtulmus, F.; Lee, W.S.; Vardar, A. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precis. Agric.* **2014**, *15*, 57–79. [[CrossRef](#)]

51. Vitzrabin, E.; Edan, Y. Changing task objectives for improved sweet pepper detection for robotic harvesting. *IEEE Robot. Autom. Lett.* **2016**, *1*, 578–584. [[CrossRef](#)]
52. Lv, J.; Wang, Y.; Xu, L.; Gu, Y.; Zou, L.; Yang, B.; Ma, Z. A method to obtain the near-large fruit from apple image in orchard for single-arm apple harvesting robot. *Sci. Hortic.* **2019**, *257*, 108758. [[CrossRef](#)]
53. Feng, Q.; Wang, X.; Li, J.; Cheng, W.; Chen, J. Image Color Correction Method for Greenhouse Tomato Plant Based on HDR Imaging. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 235–242. [[CrossRef](#)]
54. Kondo, N.; Namba, K.; Nishiwaki, K.; Ling, P.P.; Monta, M. An illumination system for machine vision inspection of agricultural products. In Proceedings of the 2006 ASABE Annual International Meeting, Portland, OR, USA, 9–12 July 2006; p. 063078. [[CrossRef](#)]
55. Gan, H.; Lee, W.; Alchanatis, V.; Ehsani, R.; Schueller, J.K. Immature green citrus fruit detection using color and thermal images. *Comput. Electron. Agric.* **2018**, *152*, 117–125. [[CrossRef](#)]
56. Bac, C.; Hemming, J.; Henten, E. Robust pixel-based classification of obstacle for robotic harvesting of sweet-pepper. *Comput. Electron. Agric.* **2013**, *96*, 148–162. [[CrossRef](#)]
57. Li, W.; Feng, Q.; Yuan, T. Spectral imaging for greenhouse cucumber fruit detection based on binocular stereovision. In Proceedings of the 2010 ASABE Annual International Meeting, Pittsburgh, PA, USA, 20–23 June 2010; p. 1009345. [[CrossRef](#)]
58. Yuan, T.; Ji, C.; Chen, Y.; Li, W.; Zhang, J. Greenhouse Cucumber Recognition Based on Spectral Imaging Technology. *Trans. Chin. Soc. Agric. Mach.* **2011**, *42*, 172–176.
59. Sa, I.; Ge, Z.; Dayoub, F.; Upcroft, B.; Perez, T.; McCool, C. DeepFruits: A fruit detection system using deep neural networks. *Sensors* **2016**, *16*, 1222. [[CrossRef](#)] [[PubMed](#)]
60. Fernández, R.; Salinas, C.; Montes, H.; Sarria, J. Multisensory system for fruit harvesting robots. experimental testing in natural scenarios and with different kinds of crops. *Sensors* **2014**, *14*, 23885–23904. [[CrossRef](#)] [[PubMed](#)]
61. Feng, Q.; Chen, J.; Cheng, W.; Wang, X. Multi-band image fusion method for visually identifying tomato plant's organs with similar color. *Smart Agric.* **2020**, *2*, 126–134. [[CrossRef](#)]
62. Liu, Z.; Wu, J.; Fu, L.; Majeed, Y.; Feng, Y.; Li, R.; Cui, Y. Improved kiwifruit detection using pre-trained VGG16 with RGB and NIR information fusion. *IEEE Access* **2019**, *8*, 2327–2336. [[CrossRef](#)]
63. Choi, D.; Lee, W.S.; Schueller, J.K.; Ehsani, R.; Roka, F.; Diamond, J. A performance comparison of RGB, NIR, and depth images in immature citrus detection using deep learning algorithms for yield prediction. In Proceedings of the 2017 ASABE Annual International Meeting, Spokane, WA, USA, 16–19 July 2017; p. 1700076. [[CrossRef](#)]
64. Feng, J.; Zeng, L.; He, L. Apple Fruit Recognition Algorithm Based on Multi-Spectral Dynamic Image Analysis. *Sensors* **2019**, *19*, 949. [[CrossRef](#)]
65. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agric.* **2016**, *125*, 184–199. [[CrossRef](#)]
66. Gisellsson, T.M.; Midtby, H.S. Seedling discrimination with shape features derived from a distance transform. *Sensors* **2013**, *13*, 5585–5602. [[CrossRef](#)]
67. Pastrana, J.C.; Rath, T. Novel image processing approach for solving the overlapping problem in agriculture. *Biosyst. Eng.* **2013**, *115*, 106–115. [[CrossRef](#)]
68. Senthilnath, J.; Dokania, A.; Kandukuri, M.; Ramesh, K.N.; Anand, G.; Omkar, S.N. Detection of tomatoes using spectral-spatial methods in remotely sensed RGB images captured by UAV. *Biosyst. Eng.* **2016**, *146*, 16–32. [[CrossRef](#)]
69. Vitzrabin, E.; Edan, Y. Adaptive thresholding with fusion using a RGB-D sensor for red sweet-pepper detection. *Biosyst. Eng.* **2016**, *146*, 45–56. [[CrossRef](#)]
70. Barnea, E.; Mairon, R.; Ben-Shahar, O. Colour-agnostic shape-based 3D fruit detection for crop harvesting robots. *Biosyst. Eng.* **2016**, *146*, 57–70. [[CrossRef](#)]
71. Rakun, J.; Stajniko, D.; Zazula, D. Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry. *Comput. Electron. Agric.* **2011**, *76*, 80–88. [[CrossRef](#)]
72. Kurtulmus, F.; Lee, W.S.; Vardar, A. Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Comput. Electron. Agric.* **2011**, *78*, 140–149. [[CrossRef](#)]
73. Song, Y.; Glasbey, C.A.; Horgan, G.W.; Polder, G.; Dieleman, J.A.; Van der Heijden, G.W.A.M. Automatic fruit recognition and counting from multiple images. *Biosyst. Eng.* **2014**, *118*, 203–215. [[CrossRef](#)]
74. Ostovar, A.; Ringdahl, O.; Hellström, T. Adaptive image thresholding of yellow peppers for a harvesting robot. *Robotics* **2018**, *7*, 11. [[CrossRef](#)]
75. Zhao, D.A.; Lv, J.D.; Ji, W.; Zhang, Y.; Chen, Y. Design and control of an apple harvesting robot. *Biosyst. Eng.* **2011**, *110*, 112–122. [[CrossRef](#)]
76. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How Transferable Are Features in Deep Neural Networks? In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 3320–3328.
77. Sun, H.; Li, S.; Li, M.; Liu, H.; Qing, L.; Zhang, Y. Research Progress of Image Sensing and Deep Learning in Agriculture. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 1–17. [[CrossRef](#)]
78. Wan, S.; Goudos, S. Faster R-CNN for multi-class fruit detection using a robotic vision system. *Comput. Netw.* **2020**, *168*, 107036. [[CrossRef](#)]

79. Zhao, D.; Wu, R.; Liu, X.; Zhao, Y. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background. *Trans. Chin. Soc. Eng.* **2019**, *35*, 164–173. [[CrossRef](#)]
80. Yan, B.; Fan, P.; Lei, X.; Liu, Z.; Yang, F. A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. *Remote Sens.* **2021**, *13*, 1619. [[CrossRef](#)]
81. Kounalakis, N.; Kalykakis, E.; Pettas, M.; Makris, A.; Kavoussanos, M.M.; Sfakiotakis, M.; Fasoulas, J. Development of a Tomato Harvesting Robot: Peduncle Recognition and Approaching. In Proceedings of the 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 11–13 June 2021; pp. 1–6. [[CrossRef](#)]
82. Birrell, S.; Hughes, J.; Cai, J.Y.; Iida, F. A field-tested robotic harvesting system for iceberg lettuce. *J. Field Robot.* **2020**, *37*, 225–245. [[CrossRef](#)] [[PubMed](#)]
83. Yu, Y.; Zhang, K.; Liu, H.; Yang, L.; Zhang, D. Real-time visual localization of the picking points for a ridge-planting strawberry harvesting robot. *IEEE Access* **2020**, *8*, 116556–116568. [[CrossRef](#)]
84. Kirk, R.; Cielniak, G.; Mangan, M. L*a*b*Fruits: A Rapid and Robust Outdoor Fruit Detection System Combining Bio-Inspired Features with One-Stage Deep Learning Networks. *Sensors* **2020**, *20*, 275. [[CrossRef](#)]
85. Cui, Z.; Sun, H.M.; Yu, J.T.; Yin, R.N.; Jia, R.S. Fast detection method of green peach for application of picking robot. *Appl. Intell.* **2022**, *52*, 1718–1739. [[CrossRef](#)]
86. Zhang, X.; Karkee, M.; Zhang, Q.; Whiting, M.D. Computer vision-based tree trunk and branch identification and shaking points detection in Dense-Foliage canopy for automated harvesting of apples. *J. Field Robot.* **2021**, *38*, 476–493. [[CrossRef](#)]
87. Yu, Y.; Zhang, K.; Yang, L.; Zhang, D. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Comput. Electron. Agric.* **2019**, *163*, 104846. [[CrossRef](#)]
88. Jia, W.; Tian, Y.; Luo, R.; Zhang, Z.; Lian, J.; Zheng, Y. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. *Comput. Electron. Agric.* **2020**, *172*, 105380. [[CrossRef](#)]
89. Kaczmarek, A.L. Stereo vision with Equal Baseline Multiple Camera Set (EBMCS) for obtaining depth maps of plants. *Comput. Electron. Agric.* **2017**, *135*, 23–37. [[CrossRef](#)]
90. Xiang, R.; Jiang, H.; Ying, Y. Recognition of clustered tomatoes based on binocular stereo vision. *Comput. Electron. Agric.* **2014**, *106*, 75–90. [[CrossRef](#)]
91. Si, Y.; Liu, G.; Feng, J. Location of apples in trees using stereoscopic vision. *Comput. Electron. Agric.* **2015**, *112*, 68–74. [[CrossRef](#)]
92. Eizentals, P.; Oka, K. 3D pose estimation of green pepper fruit for automated harvesting. *Comput. Electron. Agric.* **2016**, *128*, 127–140. [[CrossRef](#)]
93. Gongal, A.; Silwal, A.; Amatya, S.; Karkee, M.; Zhang, Q.; Lewis, K. Apple crop-load estimation with over-the-row machine vision system. *Comput. Electron. Agric.* **2016**, *120*, 26–35. [[CrossRef](#)]
94. Feng, Q.; Cheng, W.; Zhou, J.; Wang, X. Design of structured-light vision system for tomato harvesting robot. *Int. J. Agric. Biol. Eng.* **2014**, *7*, 19–26. [[CrossRef](#)]
95. Lehnert, C.; English, A.; McCool, C.; Tow, A.W.; Perez, T. Autonomous sweet pepper harvesting for protected cropping systems. *IEEE Robot. Autom. Lett.* **2017**, *2*, 872–879. [[CrossRef](#)]
96. Feng, Q.; Wang, X.; Zhang, M.; Zhang, Z.; Chen, J. Visual system with distant and close combined views for agricultural robot. *Int. Agric. Eng. J.* **2019**, *28*, 324–329.
97. Nguyen, T.T.; Vandevoorde, K.; Wouters, N. Detection of red and bicoloured apples on tree with an RGB-D camera. *Biosyst. Eng.* **2016**, *146*, 33–44. [[CrossRef](#)]
98. Kang, H.; Zhou, H.; Wang, X.; Chen, C. Real-Time Fruit Recognition and Grasping Estimation for Robotic Apple Harvesting. *Sensors* **2020**, *20*, 5670. [[CrossRef](#)]
99. Lehnert, C.; Tsai, D.; Eriksson, A.; McCool, C. 3d move to see: Multi-perspective visual servoing for improving object views with semantic segmentation. *arXiv* **2018**, arXiv:1809.07896. [[CrossRef](#)]
100. Mehta, S.S.; MacKunis, W.; Burks, T.F. Robust visual servo control in the presence of fruit motion for robotic citrus harvesting. *Comput. Electron. Agric.* **2016**, *123*, 362–375. [[CrossRef](#)]
101. Barth, R.; Hemming, J.; Van Henten, E.J. Angle estimation between plant parts for grasp optimisation in harvest robots. *Biosyst. Eng.* **2019**, *183*, 26–46. [[CrossRef](#)]
102. Hemming, J.; Ruizendaal, J.; Hofstee, J.W.; Van Henten, E.J. Fruit detectability analysis for different camera positions in sweet-pepper. *Sensors* **2014**, *14*, 6032–6044. [[CrossRef](#)] [[PubMed](#)]
103. Bac, C.W.; Hemming, J.; Van Henten, E.J. Stem localization of sweet-pepper using the support wire as a visual cue. *Comput. Electron. Agric.* **2014**, *105*, 111–120. [[CrossRef](#)]
104. Amatya, S.; Karkee, M.; Gongal, A.; Zhang, Q.; Whiting, M.D. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting. *Biosyst. Eng.* **2016**, *146*, 3–15. [[CrossRef](#)]
105. Li, J.; Tang, Y.; Zou, X.; Lin, G.; Wang, H. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots. *IEEE Access* **2020**, *8*, 117746–117758. [[CrossRef](#)]
106. Feng, Q.; Wang, X.; Liu, J.; Cheng, W.; Chen, J. Tracking and Measuring Method of Tomato Main-stem Based on Visual Servo. *Trans. Chin. Soc. Agric. Mach.* **2020**, *51*, 221–228. [[CrossRef](#)]
107. Liu, J.; Zhou, F.Y.; Yin, L.; Wang, Y. A novel cloud platform for service robots. *IEEE Access* **2019**, *7*, 182951–182961. [[CrossRef](#)]