agriculture

MDPI

Article

A Novel Telescopic Cartesian Manipulator for Kiwifruit Harvesting with Hybrid Model–Vision Error Compensation

Bo Jia [1], Shuolin Kong [1], Juncai Huang [1], Xiaoyu Ma [1], Jiwei Zhang [1], Rui Li [1], Chen Li [1], Majeed Yaqoob [2], Shen Hin Lim [3] and Longsheng Fu [1,4,5,*]

[1] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China; jiabo27@nwafu.edu.cn (B.J.); kslberich@nwafu.edu.cn (S.K.); huangjc@nwafu.edu.cn (J.H.); maxiaoyu@nwafu.edu.cn (X.M.); jarvis_zhang@nwafu.edu.cn (J.Z.); ruili1216@nwafu.edu.cn (R.L.); c.li@nwafu.edu.cn (C.L.)
[2] Department of Electrical Engineering and Computer Science, University of Wyoming, Laramie, WY 82071, USA; ymajeed@uwyo.edu
[3] School of Engineering, University of Waikato, Hamilton 3240, New Zealand; hin.lim@waikato.ac.nz
[4] Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling 712100, China
[5] Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling 712100, China
* Correspondence: fulsh@nwafu.edu.cn

Abstract

Crops cultivated on trellis systems, such as kiwifruit, grapes, et al., create a partially structured workspace; this environment is highly suitable for Cartesian robotic harvesting. However, limited extension range of current Cartesian manipulators necessitates a large vertical space, which directly conflicts with the height constraints of trellis canopies. This paper presented a hollow telescopic Cartesian manipulator with a belt-driven cascaded differential transmission for single-degree-of-freedom kiwifruit operations. The two-stage nested carbon fiber structure offers an extension ratio of 1.83:1 and a 549 mm retracted length, requiring 45.36% less vertical space than its single-stage architecture. The hybrid model–vision error compensation (HMVEC) strategy is proposed to address nonlinear positioning errors inherent in the cantilever telescopic configuration. It combines a polynomial–Fourier kinematic model for feedforward correction with YOLO11n AprilTag detection. The results showed the HMVEC strategy reduced the positioning root mean square error (RMSE) by 42.28% (from 7.90 mm to 4.56 mm). The designed manipulator achieved a 95% kiwifruit-transfer success rate at a mean cycle time of 4.8 s per fruit. These results demonstrate the feasibility of the proposed structural design and HMVEC strategy for precise manipulator positioning and post-detachment fruit transfer.

Keywords: trellis systems; telescopic manipulator; hybrid model–vision error compensation; stereo vision; kinematic modeling; object detection

Check for updates

Academic Editor: Valentin Vlăduț

Received: 25 June 2026
Revised: 27 July 2026
Accepted: 29 July 2026
Published: 3 August 2026

Copyright: © 2026 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.

1. Introduction

The global industry for trellis-cultivated crops faces growing labor shortages and rising production costs, driving demand for automated harvesting. Manual harvesting accounts for a substantial proportion of fruit production costs [1,2]. Crops such as grapes, kiwifruit, and passion fruit are typically cultivated on standardized trellis systems, creating a partially structured workspace that is available for robotic harvesting [3–5]. However, the transition to robotic harvesting imposes stringent design constraints. Spatially, the manipulator must

operate within a vertically limited workspace, requiring a compact retracted structure for navigation and a high extension ratio to reach fruits at varying depths. Operationally, high-speed motion is essential to minimize cycle times, while precise end-effector alignment is critical to prevent collateral damage [6,7]. A manipulator capable of satisfying such demanding requirements is difficult to realize.

While serial manipulators offer high flexibility, their complex joint motions increase the risk of collisions with branches and leaves within the canopy space. Consequently, complex trajectory planning and real-time obstacle-avoidance strategies are necessitated, which may increase planning complexity and operational time [8–11]. Conversely, Cartesian manipulators are inherently better suited to the rectangular workspace of trellis canopies, yet traditional configurations require substantial vertical installation space [12]. Retractable chain mechanisms and telescopic links offer an effective solution to this spatial constraint, driving their increasing adoption in harvesting robots [13]. However, these conventional telescopic designs typically follow a discrete 'pick-and-place' paradigm, necessitating the complete retraction of the manipulator to deposit the fruit after each pick [14]. The resulting increase in operational cycle time highlights a critical need for integrated continuous fruit-transport mechanisms.

In traditional robotic harvesting scenarios, the manipulator is typically required to retract to a designated collection bin after each individual grasp. This repetitive stop-and-go motion pattern leads to a significant bottleneck that restricts operational throughput. This non-productive retrieval phase alone consumes approximately 2.2 s [15]. The better strategy to save time on redundant movements is to transfer fruit through the manipulator itself. A corrugated hose integrated into the end-effector provides a practical solution for continuous fruit collection [16]. Furthermore, such pneumatic or gravity-assisted transport systems streamline the workflow, allowing the robot to focus exclusively on detachment. These hollow structures can compress the harvesting cycle to a rapid range of 4.8 to 5.5 s [17,18]. Despite the improvement in time efficiency, the integration of a corrugated tube or the use of a gravity-assisted transport system inevitably compromises the positioning accuracy of the telescopic structure itself, necessitating specific compensation strategies to address this issue.

Some researchers have applied visual servoing to mitigate the positioning errors resulting from the cantilever and hollow structures of telescopic manipulators. These errors stem from the coupling of gravitational deflection and the cumulative structural clearances between nested segments, which jointly exhibit strong nonlinear kinematic characteristics [19]. Integrating deep learning or real-time visual feedback can optimize approach trajectories and reduce static errors [20–24]. Nevertheless, these vision-centric approaches often assume the manipulator as a rigid system, falling short in actively compensating for the high-frequency dynamic disturbances induced by mechanical backlash, especially when visual lines of sight are intermittent [25]. Previous studies have separately investigated compact telescopic mechanisms, continuous fruit-transfer devices, and model- or vision-based positioning control. However, the integration of these functions within a lightweight telescopic axis remains insufficiently investigated, particularly when the manipulator must simultaneously satisfy limited installation space, continuous internal fruit transport, and compensation for load-dependent nonlinear errors. Consequently, positioning is not an isolated mechanical or control issue, but a complex coupling of structural nonlinearities and control responsiveness. These limitations motivated the development of a hybrid model–vision error compensation (HMVEC) strategy to address structural nonlinearities and residual positioning errors.

This study proposes a novel telescopic Cartesian manipulator with a single-degree-of-freedom, specifically designed for terminal positioning-error compensation and post-

detachment fruit transfer in trellis-grown kiwifruit harvesting. The main novelty lies in the structural-control integration of a two-stage hollow telescopic mechanism and the HMVEC strategy. The hollow mechanism provides compact extension and an internal fruit-transport path, while the HMVEC strategy combines a polynomial–Fourier feedforward model with AprilTag-based visual feedback to compensate for systematic structural deformation and reduce residual positioning errors. AprilTag markers installed at both ends of the manipulator also provide operating-state verification. The integrated system achieved a 95% kiwifruit-transfer success rate at an operating speed of 0.5 m/s and reduced the overall positioning root mean square error (RMSE) by 42.28% relative to uncompensated control, while maintaining a lightweight construction of 1438 g. These results indicate the potential of the proposed structural-control architecture for kiwifruit-harvesting applications in standardized trellis orchards.

2. Materials and Methods

2.1. System Overview

Kiwifruit is typically cultivated on standardized trellis systems, where rattans climb along support wires to form a dense canopy approximately 1.8 m above the ground. It creates a distinct rectangular workspace between the canopy and the ground. Fruit clusters hang naturally within a vertical range of 0.2–0.3 m beneath the canopy, necessitating a bottom-up approach for automated harvesting. Low-hanging foliage constrains the platform height and limits the available vertical space for manipulator extension.

The telescopic manipulator system comprises a mechanical assembly and a hierarchical control architecture, integrated onto a dual-manipulator Cartesian mobile platform with a vertical mounting clearance of 695 mm. The mechanical assembly features a lightweight (1438 g) hollow nested carbon fiber structure with a retracted length of 549 mm and an inner channel diameter of 110 mm, which is larger than the maximum transverse diameter of the tested kiwifruit cultivars, the corrugated hose is made of flexible material, and harvested kiwifruit generally has relatively firm texture at harvest maturity. Therefore, the mild buffering provided by the flexible corrugated tube in the curved region may help reduce impact damage during fruit transport. The manipulator is actuated by a DJI M3508 brushless DC gear motor (DJI, Shenzhen, China) via a DJI C620 ESC. An overview diagram of the system is shown in Figure 1. The control architecture encompasses three core elements: first, a kinematic regression model derived from the Mars 1.3H motion capture system for feedforward error compensation; second, a visual perception module utilizing an ALG AVS-2812 stereo camera (3840 $\times$ 2160 resolution, Aili-Light Technology Co., Ltd., Shenzhen, China) and AprilTag fiducials (Tag36h11) for real-time pose feedback; and finally, a central controller executing a hybrid compensation strategy. The proposed strategy employs YOLO11n to detect tags and a PID loop to correct errors, ensuring highly accurate closed-loop positioning.

2.2. Manipulator Design

2.2.1. Nested Telescopic Structure

The proposed manipulator allows for a maximum reach with minimal inertia. The actuation unit is powered by a single M3508 brushless DC motor mounted to the base. The motor uses a cascaded differential transmission to drive the belt, which extends the first-stage manipulator. Simultaneously, a constrained telescopic belt, which is anchored to the base and connected to the pull plate of the second stage synchronously drives the second-stage manipulator in the same direction, as illustrated in Figure 2a. This differential coupling allows the end-effector to achieve rapid response with a high extension ratio. Meanwhile, an end-effector mounting bracket is attached to the distal end to accommodate

diverse operational tools. The manipulator adopts a two-stage nested structure made of hollow carbon fiber tubes, as illustrated in Figure 2b. This hollow structure creates a continuous internal pipe which, in conjunction with a fruit-gathering wire harness, facilitates immediate fruit transport while reducing potential fruit damage, thereby eliminating the retraction time associated with solid-link manipulators.

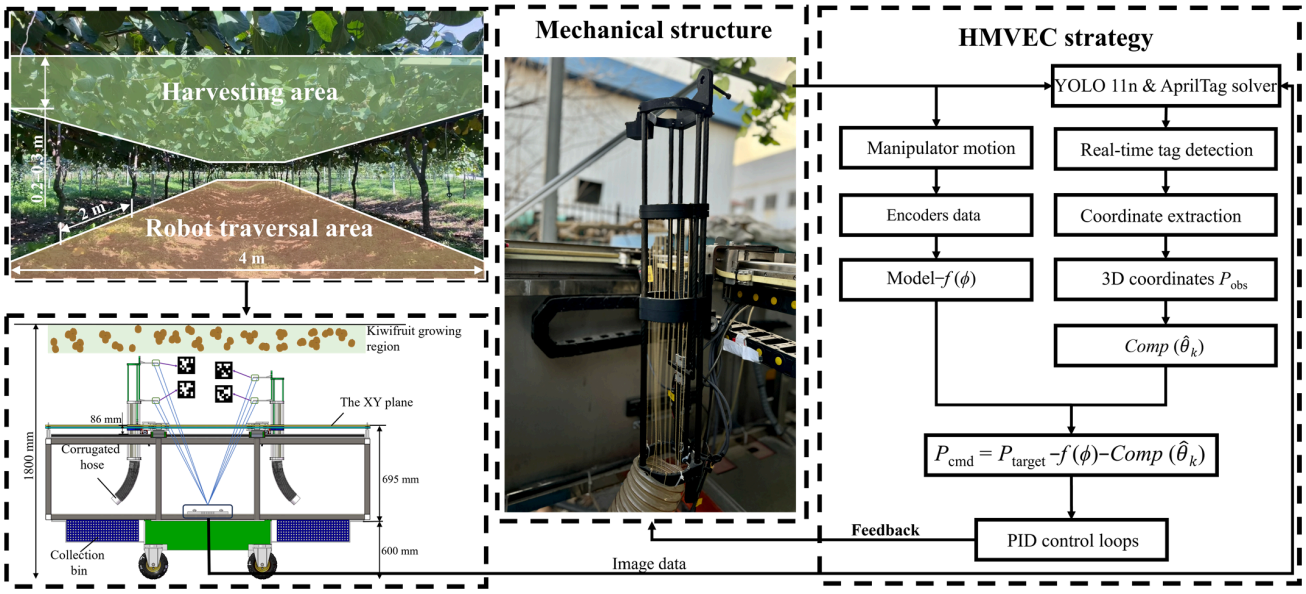


Figure 1. Overall architecture of the telescopic manipulator system for terminal positioning-error compensation and post-detachment fruit transfer in trellis-grown kiwifruit harvesting. The top-left panel illustrates the spatial constraints of the standardized kiwifruit trellis environment. The bottom-left schematic details the hardware integration, including the mobile platform and the binocular stereo vision system. The central photograph displays the physical prototype of the hollow telescopic manipulator. The right panel presents the control flowchart of the hybrid model–vision error compensation (HMVEC) strategy.

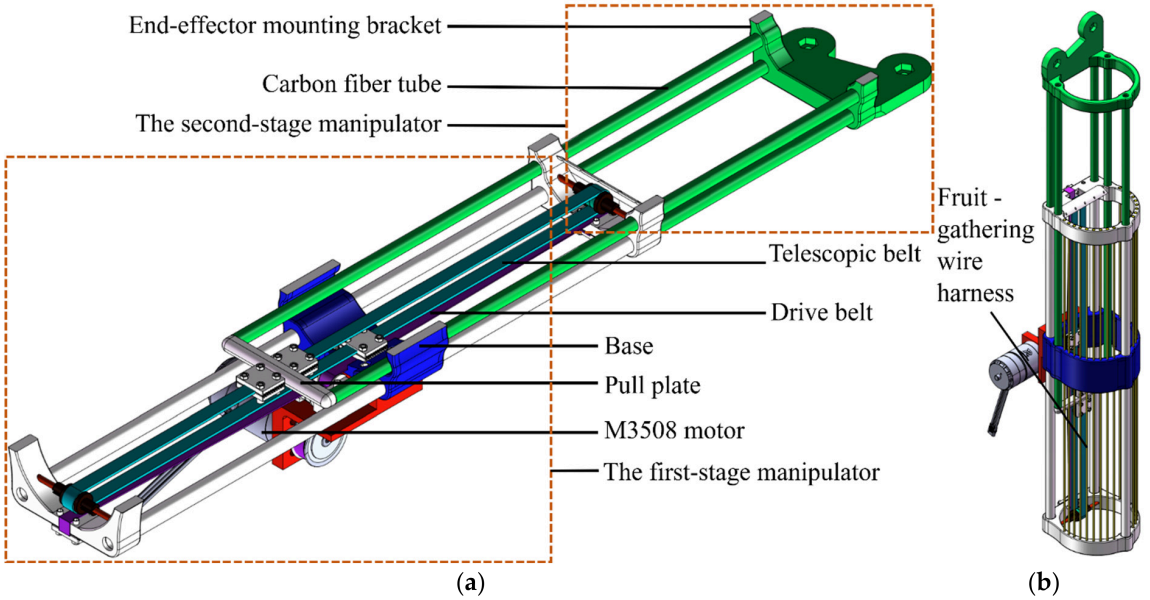


Figure 2. Hollow telescopic manipulator design. (**a**) Belt-driven cascaded mechanism with labeled components. The drive belt extends the first-stage manipulator while the constrained telescopic belt simultaneously extends the second-stage manipulator, achieving 1.83:1 kinematic ratio. (**b**) Overall structure of the telescopic manipulator.

2.2.2. Constraint-Based Dimensional Design

The manipulator dimensions (length L and inner diameter d) are determined by two independent geometric constraints. Length L is bounded by the platform's vertical space h_1 and the requirement to reach the canopy height H with extension ratio λ. The inner diameter d must accommodate kiwifruit dimensions with a safety margin to prevent blockage. The final design selects feasible dimensions while satisfying all agronomic and platform constraints. The geometric parameters of kiwifruit and platform parameters also include the maximum fruit transverse diameter d_1, maximum longitudinal dimension h_3, mounting loss t_1, nesting loss t_2, and platform height h_2, as illustrated in Figure 3. The kiwifruit cultivars investigated in this study include Hayward, Xuxiang, Huayou, and Cuixiang.

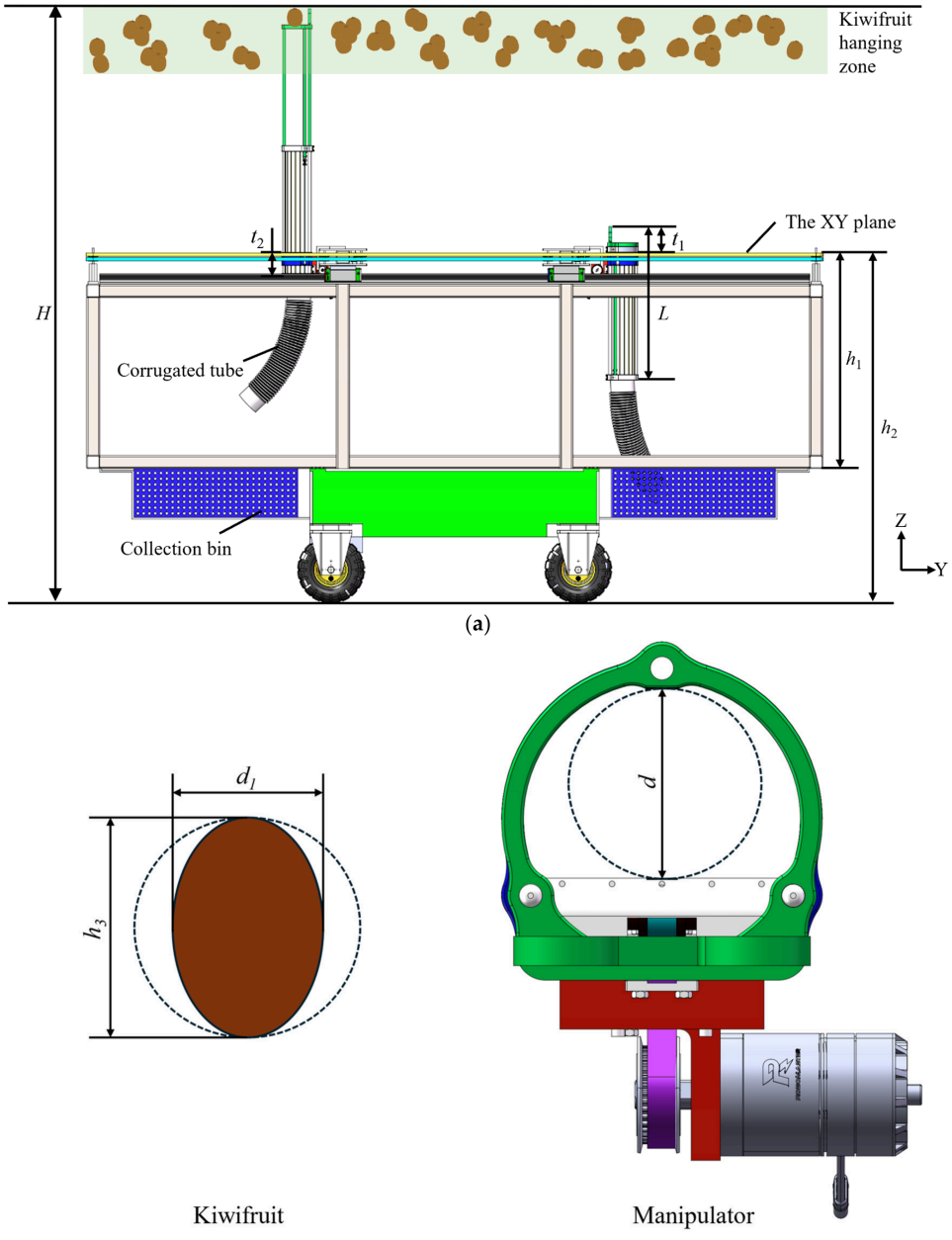


Figure 3. Dimensional design and spatial constraint analysis of the telescopic manipulator. (**a**) Vertical spatial constraints of the Cartesian mechanism operating under the kiwifruit hanging zone. (**b**) Dimensional matching between kiwifruit geometric parameters and the hollow manipulator.

Since the design variables L and d are governed by independent physical requirements, their feasible ranges can be evaluated separately. The retracted length is constrained by the available vertical installation space and required extension range:

$$\frac{H - h_2 + t_2}{\lambda} < L < h_1 + t_1 \quad (1)$$

The inner diameter is constrained by the maximum transverse and longitudinal dimensions of the kiwifruit:

$$d \geq \max\{d_1, h_3\} \quad (2)$$

Through dimensional measurements of the designed manipulator model in both fully extended and retracted states, an extension ratio of λ = 1.83 was obtained. Concurrently, field measurements of the general-purpose dual-manipulator kiwifruit production platform and relevant fruit characteristics yielded the parameters and dimensions listed in Table 1.

Table 1. Technical specifications of the dual-manipulator kiwifruit production platform and geometric dimensions of kiwifruit.

	Dual-Manipulator Kiwifruit Production Platform						**Kiwifruit**	
Parameter	λ	H	h_1	h_2	t_1	t_2	h_3	d_1
Dimension	1.83	1800 mm	695 mm	1295 mm	80 mm	86 mm	100 mm	80 mm

Substituting these empirical values into Equations (1) and (2) yields the feasible ranges for the fully retracted length L and inner diameter d:

$$323 \text{ mm} < L < 775 \text{ mm}$$

$$d \geq 100 \text{ mm}$$

Operational cycle analysis indicates that a manipulator velocity of 0.5 m/s satisfies the continuous task requirements. Given a synchronous pulley pitch diameter D of approximately 29 mm, the required rotational speed to achieve this target velocity is calculated via the kinematic mapping Equation (3) as:

$$n_{req} = \frac{60v}{\pi D} \approx 329.29 \text{ rpm} \quad (3)$$

where n_{req} denotes the required motor rotational speed. The M3508 motor features a rated load output speed of 469 rpm, with a safe continuous operating range of 330 to 400 rpm. Operating within this specific interval maximizes the intrinsic motor efficiency and minimizes thermal rise, while concurrently reducing gear wear and acoustic noise in the gearbox. These characteristics establish it as the optimal choice for prolonged continuous operation, thereby justifying its selection as the drive unit in this study. Furthermore, calculations based on Equation (3) demonstrate that operating the M3508 motor at its rated load output speed of 469 rpm enables the manipulator to achieve a linear velocity of 0.71 m/s.

2.3. Error Characterization and Kinematic Modeling

2.3.1. Error Characterization

The end-effector positioning error of this telescopic manipulator is primarily caused by the coupling of static deformation and dynamic disturbances. Static positioning errors arise primarily from gravitational loading, structural clearances between telescopic segments and eccentric payloads, which jointly induce lateral deflection of the telescopic manipulator.

Since the manipulator is mounted in a cantilever-like configuration, deflection increases with extension length and payload mass, rendering rigid-body kinematic assumptions inadequate. In addition to static effects, dynamic errors occur during rapid start-stop motions. Inertial forces excite structural flexibility, while transmission backlash introduces hysteresis during direction reversals. These effects result in transient oscillations superimposed on the static deflection, further degrading positioning accuracy. As shown in Figure 4, these complex coupling effects manifest primarily as a deflection angle θ of the end-effector relative to the ideal axis.

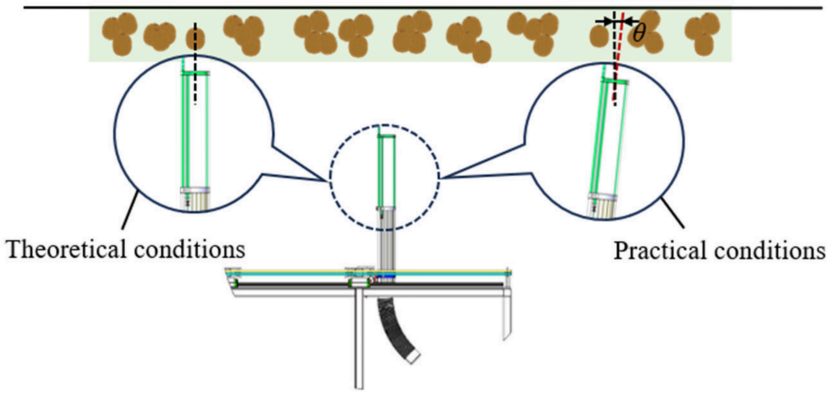


Figure 4. Structural deflection of the telescopic manipulator during vertical extension. The left side displays the theoretical conditions with perfect vertical alignment. The right side illustrates the practical conditions. The symbol θ denotes the actual deflection angle caused by structural bending.

Dynamic tests were conducted to visually evaluate the impact of end-effector weight on the ascending trajectory of the manipulator by sequentially mounting three end-effectors with incrementally increasing masses: bud thinning (310 g), pruning (820 g), and harvesting (1680 g). During the ascending motion, an external motion capture system (Mars 1.3H, 240 Hz, ±0.2 mm accuracy, Beijing NOKOV Science & Technology Co., Ltd., Beijing, China) tracked reflective markers attached to the manipulator in real time, thereby accurately recording and visualizing its six-degree-of-freedom (6-DOF) spatial pose variations, as illustrated in Figure 5. For each payload condition, considering that the designed manipulator can meet most operational requirements at a motion speed of 0.5 m/s, multiple complete motion trajectories from the fully retracted position to the fully extended position were recorded at this speed. To reduce the influence of random measurement fluctuations and obtain a representative motion trend, 20 complete trajectories were selected under each payload condition and aligned according to the motor encoder counts. At each identical encoder-count position, the mean displacements along the X-, Y-, and Z-axis directions were calculated to construct a fused mean trajectory. In addition, under the harvesting end-effector condition, five complete trajectories that were not used for model fitting were reserved as validation trajectories, providing a data basis for subsequent tilt-angle analysis and kinematic regression-model validation.

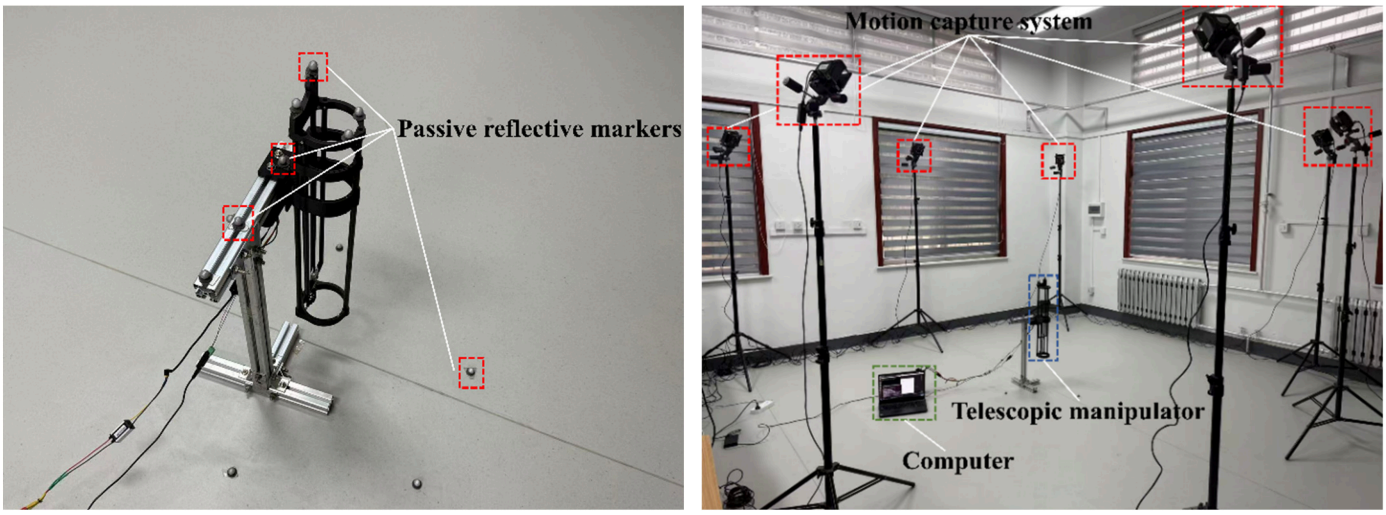
2.3.2. Kinematic Modeling

From a physical mechanism perspective, structural clearances amplify gravity-induced deflection, while the synchronous belt transmission introduces position-dependent fluctuations, thereby modulating the overall compliance of the system. Individual contributions of each error source cannot be isolated through direct measurement, as the manipulator exhibits a unified deformation pattern at the end-effector. Theoretical prediction via physics-based analytical modeling would require precise knowledge of structural parameters such as joint stiffness, clearance distributions, and friction coefficients, which are difficult to ob-

tain for assembled systems and subject to manufacturing variability and wear-induced drift. Therefore, a hybrid polynomial–Fourier regression model was formulated to map motor encoder values to predict positioning errors in Cartesian space, as shown in Equation (4):

$$y = \underline{ax^2 + bx} + \underline{c\sin(\omega x + \theta)} + k \quad (4)$$

where the $ax^2 + bx$ component corresponds to a truncated Taylor series, primarily employed to fit the monotonic quadratic trend induced by cantilever gravitational deflection and structural clearances as the extension length increases. The $c\sin(\omega x + \theta)$ component represents a Fourier series term, designed to independently characterize the periodic errors resulting from the eccentricity of rotating components or transmission meshing. The coefficients of the regression model (a, b, c, ω, θ, k) were determined using a nonlinear least squares optimization method. The dataset was constructed by synchronizing the motor encoder values x with the ground-truth positional deviations y obtained from the Mars 1.3H motion capture system. The Trust Region Reflective method was employed to iteratively minimize the Residual Sum of Squares between the model predictions and the experimental trajectories. The fitted model serves as a feed forward correction term in the control system. For a given target position and known payload mass, the model predicts the expected systematic error.



(**a**) (**b**)

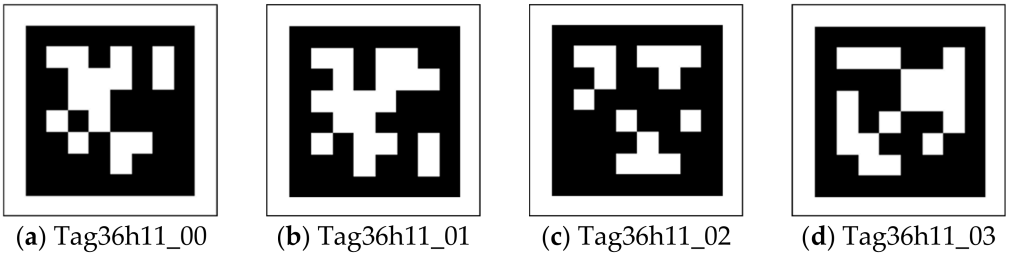
Figure 5. Experimental setup for kinematic trajectory acquisition. (**a**) The telescopic manipulator instrumented with passive reflective markers for pose tracking. (**b**) Global layout of the Mars 1.3H optical motion capture system environment, the system consists of a host computer, a telescopic manipulator, and seven optical cameras.

2.4. Hybrid Model–Vision Error Compensation Strategy

The HMVEC strategy comprises three core modules: (1) an offline-calibrated polynomial–Fourier kinematic model to compensate for systematic static positioning errors induced by the physical characteristics of the mechanical structure; (2) a stereo vision module for real-time pose tracking and safety verification; and (3) a PID feedback controller with load-dependent preset gains, which corrects residual errors according to the validity of visual feedback. This architecture ensures sub-centimeter positioning accuracy under both nominal conditions and transient occlusion scenarios.

2.4.1. Visual Perception Module

The visual perception system employs an ALG stereo camera (Model: AVS-2812-B20-M17G-G88-ABA-IP67) with a resolution of 3840 × 2160 pixels and a baseline of 200 mm. The camera was rigidly mounted in the lower-middle region of the mobile platform at a height of 600 mm, providing a field of view (FoV) of 120° that encompasses the entire workspace of both manipulators. The Tag36h11 family from the AprilTag library was selected as the visual fiducial system (Figure 6).



(**a**) Tag36h11_00 (**b**) Tag36h11_01 (**c**) Tag36h11_02 (**d**) Tag36h11_03

Figure 6. The four types of AprilTag markers used in this experimental.

Specifically, AprilTag markers (ID 00 and ID 01) were rigidly affixed to the distal and proximal ends of the left manipulator, respectively, as shown in the global image in Figure 7, with a symmetrical configuration applied to the right manipulator (ID 02 and ID 03). During operation, the system performs a real-time consistency check by computing the coordinate deviation along the X-axis (ΔX) for each tag pair. Given the structural rigidity of the telescopic mechanism, this deviation is theoretically constant. Consequently, if ΔX remains within the measurement precision threshold of the vision system, the operational state is deemed nominal. Conversely, should the deviation significantly exceed this threshold, the system identifies a structural anomaly or tracking failure and triggers an immediate emergency stop to ensure safety.

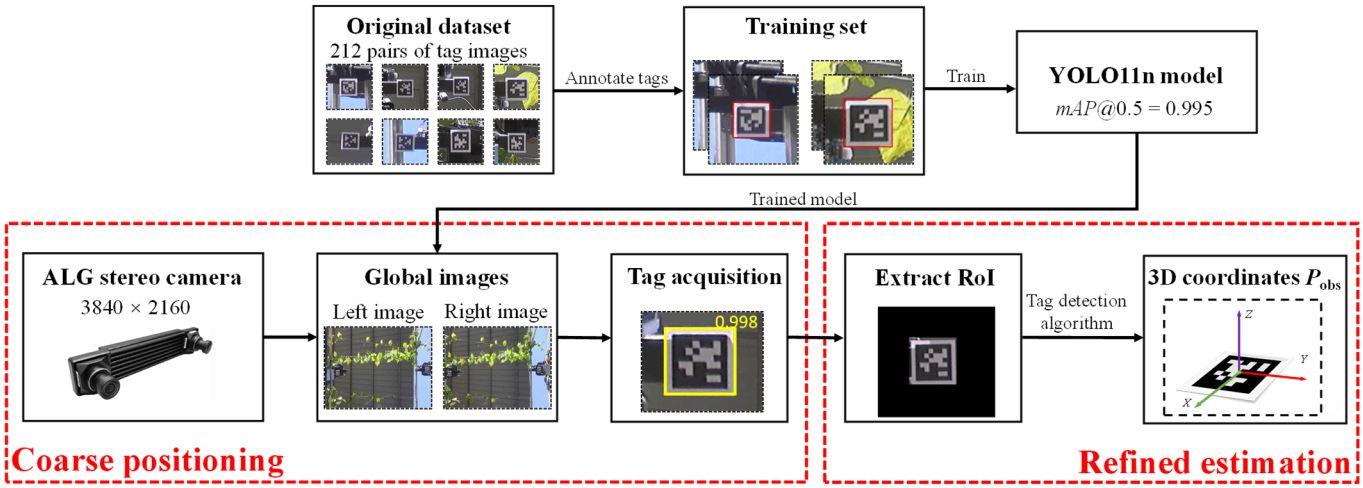


Figure 7. Two-stage detection module workflow. YOLO11n detection was used to acquire the RoI while the AprilTag detection processed the extracted tag pair to estimate the refined position.

Traditional detection methods rely heavily on the integrity of the tag's Quiet Zone (the white border) for quadrilateral extraction. When tags are subjected to strong specular highlights (causing overexposure) or partial foliage occlusion (which disrupts border continuity), traditional methods frequently fail to extract valid quadrilateral contours. This susceptibility inevitably results in detection failures or erroneous target identification [26], thereby compromising the overall reliability of the visual servoing system. In contrast,

this paper proposed a 'coarse-to-fine' two-stage detection strategy. It should be noted that the YOLO11n model in this study was used for AprilTag region of interest (RoI) detection rather than kiwifruit fruit detection. YOLO11n model was trained using 212 pairs of real-world images, while another 53 pairs of images were used for validation. These images were collected using the same stereo camera and robotic platform as those used in the experiments, and were acquired under outdoor simulated orchard conditions with natural illumination, vegetal backgrounds, partial occlusion, and reflective interference. The proposed method utilizes YOLO11n as a semantic feature extractor, with the aim of improving the robustness of AprilTag marker localization under complex backgrounds. Therefore, the performance of online kiwifruit detection across different cultivars and orchard conditions was not evaluated in this study. The specifications of the platform used for training and deployment are listed in Table 2.

Table 2. Specifications of the platform for YOLO11n model training and deployment.

Item	Details
CPU	Intel Core i5-13600KF
GPU	Nvidia GeForce RTX 4060 (8 GB)
Memory Size	32 GB
OS	Microsoft Windows 10 (64-bit)
Python Version	3.10.19
PyTorch Version	2.3.1
CUDA Version	11.8

The module swiftly filters out complex background clutter to output a RoI bounding box encapsulated the target tag. Based on these predicted boundaries, the original image is cropped and preprocessed to extract high signal-to-noise-ratio (SNR) local imagery. These preprocessed images are fed into the standard AprilTag detection for precise sub-pixel pose estimation, as shown in Figure 7. By combining YOLO11n assisted RoI extraction with AprilTag-based geometric pose estimation, the proposed coarse-to-fine visual strategy improved tag localization robustness compared with direct AprilTag detection under occlusion and shadow conditions. This more reliable visual feedback supports the accuracy of end-effector positioning error compensation.

2.4.2. Closed-Loop Control

The system presents a closed-loop adaptive control architecture that systematically integrates kinematic modeling with deep learning-based visual perception. A polynomial–Fourier kinematic model generates a feedforward term to actively compensate for inherent static nonlinearities, as shown in Figure 8. Specifically, gravitational deflection and structural clearances. Building upon this kinematic baseline, the visual perception module functions as a robust feedback mechanism. It utilizes a YOLO11n detection model to isolate the RoI from the global frame, after which a tag solver decodes the tag's 2D pixel coordinates. These coordinates are then reconstructed into the 3D observed position of the terminal center point of the telescopic manipulator via stereo disparity fusion and coordinate transformation. A valid tag ID, complete corner detection, and tag-pair consistency were considered as the validity criteria for AprilTag detection. The PID parameters were set as K_p = 2.01, K_i = 49.25, and K_d = 0.0019.

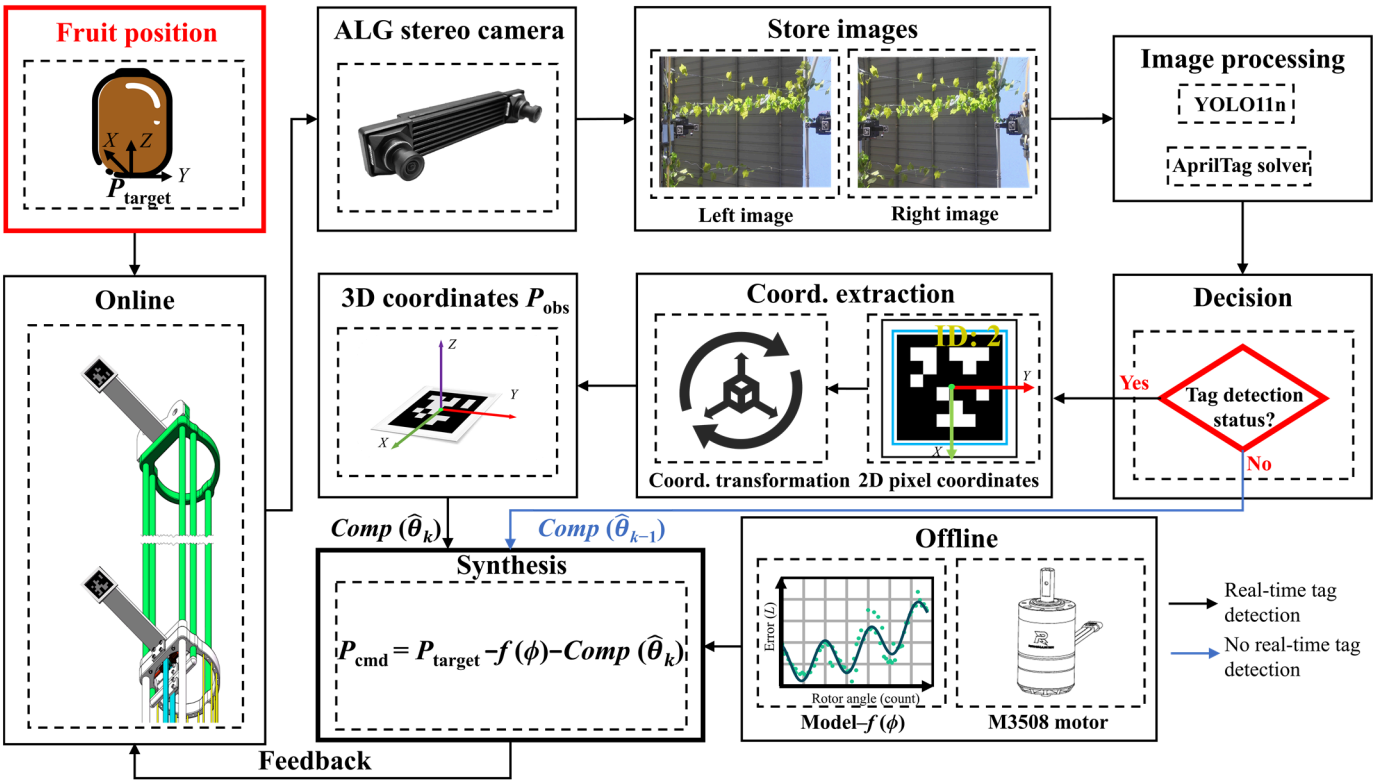


Figure 8. The proposed hybrid adaptive control architecture. It integrates an offline polynomial–Fourier feedforward model for static error compensation and a real-time visual feedback loop driven by YOLO11n and AprilTag methods to handle dynamic disturbances.

Subsequently, the positioning residual derived from this visual feedback drives the Adaptive Estimator, which modulates control logic based on signal integrity. When the AprilTag-based observation was valid, the residual between the reconstructed 3D observed position of the terminal center point of the telescopic manipulator and the position calculated from the polynomial–Fourier model was used to update the visual compensation term. When the AprilTag observation was invalid or interrupted, the Hold Strategy was activated to lock the parameter at its previous state ($\hat{\theta}_{k-1}$), as indicated by the blue arrow in Figure 8, thereby preventing control divergence. Consequently, the final control command is synthesized by incorporating the 3D fruit coordinates derived from binocular stereo matching, as shown in Equation (5).

$$P_{\text{cmd}} = P_{\text{target}} - f(\phi) - Comp(\hat{\theta}) \tag{5}$$

where P_{cmd} and P_{target} are the final control command and the 3D fruit coordinates, respectively; f (ϕ) is the kinematic feedforward term; and $Comp(\hat{\theta})$ represents the adaptive compensation parameters.

This final control command ensures high-precision, real-time compensation for both static gravitational deflection and dynamic transmission errors.

2.4.3. Outdoor Validation Under Simulated Orchard Conditions

All experiments were conducted on the selected robotic manipulation platform in an open outdoor site outside the laboratory under simulated kiwifruit-orchard operating conditions. A vision sensor was mounted on the platform to provide visual information for target detection and localization, while a laser rangefinder was used to obtain the ground-truth Cartesian position of the end-effector. During each trial, 53 predefined target points were distributed within the reachable workspace, and the actual end-effector positions

measured by the laser rangefinder were recorded for quantitative analysis. A comparative experiment evaluated the effectiveness of the proposed strategy by implementing three control strategies under identical conditions: open-loop control without compensation, model-based compensation using the predicted error model alone, and the proposed HMVEC strategy combining predictive compensation with vision-based feedback. The RMSE was adopted as the evaluation metric due to its sensitivity in measuring end-effector dispersion and system stability, particularly for capturing significant deviations [27].

Subsequently, a continuous-operation experiment was conducted using a single manipulator and artificial kiwifruit targets. Because the fruit-detection system was not deployed on the experimental platform, 60 target coordinates were predefined. These coordinates were evenly divided into 12 sets, with five target coordinates in each set. A corresponding artificial kiwifruit target was placed at each coordinate, and one operation cycle was performed for each target. For each target, the operation cycle included reading the target coordinates, performing feedforward compensation based on the polynomial–Fourier model, moving the telescopic arm toward the target position, acquiring AprilTag-based visual feedback, applying HMVEC correction to the terminal-center positioning error, aligning the existing bionic finger-type end-effector with the target fruit, achieving fruit–peduncle separation by gripping the target fruit, tilting the fruit, and pulling it away from the peduncle, and transporting the detached fruit through the hollow manipulator and corrugated tube into the collection bin. Fruit detection and 3D localization were not included. In addition, the end-effector was spatially matched with the hollow structure, and no motion interference occurred between them; therefore, the hollow structure did not restrict or affect the normal motion of the end-effector.

Within each set of the continuous-operation experiment, the first cycle was timed from the moment the manipulator started moving toward the first target. Each subsequent cycle began when the preceding target fruit entered the collection bin and ended when the current target fruit entered the collection bin. The mean cycle time per fruit was calculated from the 60 measured cycle times. During continuous operation, the manipulator moved directly from the current operating position to the next target; therefore, the cycle did not include a separate return-to-initial-position stage. During the same experiment, the number of artificial kiwifruit targets that, following successful separation by the end-effector, passed through the hollow manipulator and corrugated tube and ultimately reached the collection bin was recorded. The kiwifruit-transfer success rate, S_t, was calculated as follows:

$$S_t = \frac{N_c}{N_d} \times 100\% \tag{6}$$

where N_d denotes the number of artificial kiwifruit targets successfully detached by the harvesting end-effector, and N_c denotes the number of detached artificial kiwifruit targets that successfully reached the collection box.

3. Results and Discussion

3.1. Structural and Dynamic Performance Analysis of the Manipulator

A retracted length of 549 mm was selected as the midpoint of the feasible interval derived from Equation (1). This intermediate configuration provides a balance between structural rigidity and reaching capability. With a retracted length L of 549 mm, the proposed telescopic mechanism reduced the vertical installation space by 45.36% compared to a single-stage manipulator of equivalent reach, while weighing only 1438 g. In contrast to the manipulator developed by Salazar et al. [28], which is primarily fabricated from 3D-printed PLA and restricted to a maximum extension of 518 mm at a mass of 2.0 kg, the system proposed in this study adopts a hollow nested carbon fiber structure. This

design not only achieves an approximate 28% reduction in total moving mass relative to the Salazar prototype but also extends the maximum reach to 1004.67 mm. Furthermore, the superior strength-to-weight ratio of carbon fiber effectively overcomes the inherent rigidity limitations associated with PLA materials in large-span structures.

3.2. Deflection and Kinematic Error Characterization of the Manipulator

Motion trajectories for the telescopic Cartesian manipulator across various end-effectors were recorded via a motion capture system and subsequently visualized. The results revealed strong lateral stability but highlighted noticeable gravity-induced deflections under varying loads. Analysis of the projection on the X-Z plane showed that trajectory fluctuations along the X-axis remained within 3 mm for all configurations (as shown in Figure 9). This suggests that the manipulator possesses sufficient lateral rigidity perpendicular to the cantilever plane, with limited lateral interference caused by eccentric torque. In contrast, the Y-axis exhibited significant nonlinear deviation characteristics. As the weight of the end-effector increased, the tendency of the manipulator to tilt away from the fixed base became increasingly pronounced. This deflection reached a maximum deviation of 9.76 mm. This phenomenon aligns with the physical laws governing the deflection of cantilever beams under gravitational load.

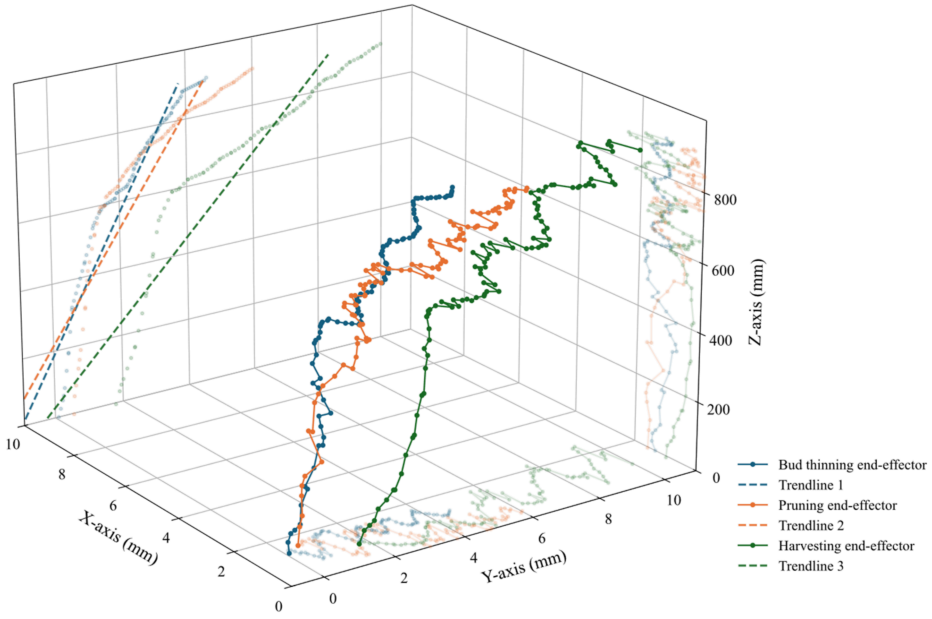


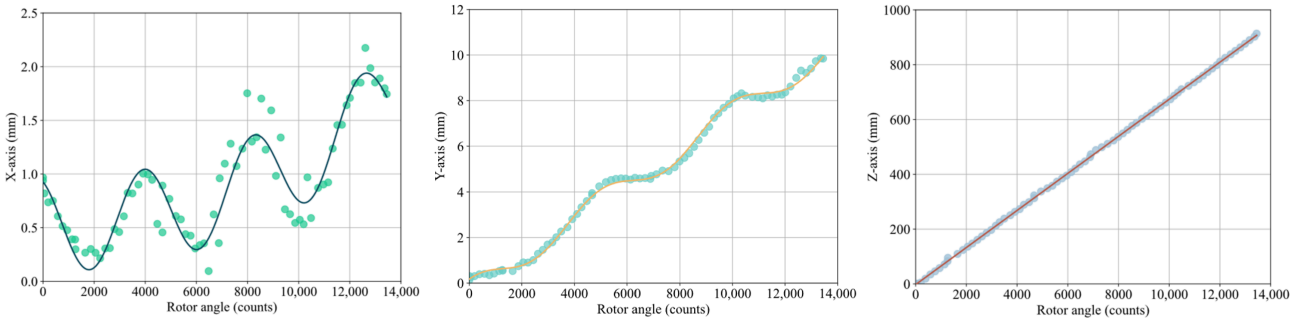
Figure 9. Three-dimensional motion trajectories of the telescopic manipulator during the extension process under varying end-effector loads. For each payload condition, the displayed trajectory was obtained by aligning 20 complete motion trajectories according to the motor encoder counts and averaging the X, Y, and Z axes displacements at identical encoder-count positions. Deviations along the X-axis indicate lateral instability, while deviations along the Y-axis reflect gravity-induced bending under different payload conditions.

Because the proposed manipulator was primarily designed for the harvesting end-effector, the polynomial–Fourier regression model was fitted using the fused mean trajectory obtained under the harvesting end-effector condition. The fitted equations for the X-, Y-, and Z-axis directions are reported in Table 3. The coefficient of determination (R^2) indicates the goodness of fit of the regression model to the experimental data. An R^2 value closer to 1 implies a stronger explanatory power of the model regarding the manipulator's motion trajectory. For the Y-axis, which was most significantly affected by the gravitational moment, the polynomial–Fourier model achieved an R^2 of 0.998, accurately reproducing

the nonlinear sagging trend of the manipulator terminal center point. Similarly, the model demonstrated robust adaptability in the X-axis direction, with an R^2 of 0.896, effectively capturing minor lateral oscillations, while the Z-axis, which served as the primary driving direction, exhibited excellent linear characteristics with an R^2 of 0.999. The prediction accuracy of the model was evaluated using five additional complete trajectories that were not involved in model fitting. The validation trajectories were aligned according to the motor encoder counts, and their mean validation points were compared with the corresponding coordinates calculated from the fitted polynomial–Fourier curves. The regression and validation results are shown in Figure 10.

Table 3. Polynomial–Fourier regression models and coefficients of determination fitted from the fused mean trajectory of 20 calibration trials under the harvesting end-effector condition.

Axis	Regression Models	R^2
X	$y = 6.892 \times 10^{-9}x^2 - 9.792 \times 10^{-6}x + 0.435\sin(1.473 \times 10^{-3}x + 2.040) + 0.539$	0.896
Y	$y = -1.303 \times 10^{-9}x^2 + 7.930 \times 10^{-4}x + 0.545\sin(1.270 \times 10^{-3}x + 1.533) - 0.370$	0.998
Z	$y = 0.068x - 2.422$	0.999



(**a**) (**b**) (**c**)

Figure 10. Regression analysis of the relationship between the 3D pose of the harvesting manipulator's end-effector and the motor rotor angle. (**a**) X-axis displacement, showing the lateral oscillation trend and the fitted polynomial–Fourier curve; (**b**) Y-axis displacement, showing the nonlinear deflection trend mainly caused by gravitational loading and the fitted polynomial–Fourier curve; and (**c**) Z-axis displacement, showing the approximately linear extension trend and the fitted linear curve.

Based on the fitted models provided in Table 3, the approximation errors were obtained by comparing the model-predicted values with the measured values. The mean, median, and maximum errors of the regression models within the encoder-count range of 0–13,500 are summarized in Table 4. It can be observed that the maximum error in the Z-axis direction reached 7.157 mm, which may be attributed to a local deviation at an individual measurement point. Nevertheless, the mean absolute error of the Z-axis model was 2.699 mm, accounting for only approximately 0.30% of the Z-axis displacement range. Therefore, the model still accurately captured the dominant linear relationship between the encoder counts and the Z-axis displacement.

Table 4. Axis-wise approximation errors of the regression models within the operating range.

Axis	Mean Error/mm	Median Error/mm	Maximum Error/mm
X	0.123	0.079	0.441
Y	0.109	0.013	0.423
Z	2.699	2.599	7.157

3.3. Two-Stage Tag Detection Using YOLO11n for Manipulator Spatial Positioning

The adaptability of the visual perception module was evaluated under simulated orchard conditions in an open outdoor site. Comparative experiments were conducted to assess the tag detection success rate. Specifically, the experiments targeted critical failure cases for traditional direct tag detection methods, such as structural occlusion by vines and strong light reflection. The proposed coarse-to-fine detection method is employed to detect markers in panoramic images captured by the stereo camera, and the results are compared with those of traditional marker detection methods, as shown in Figure 11.

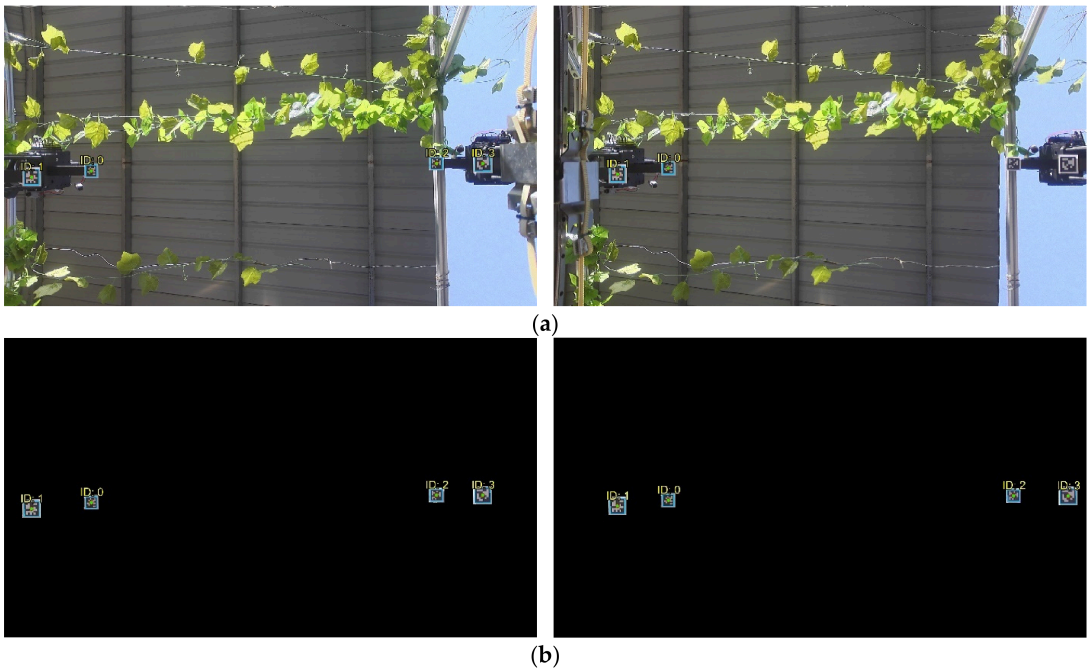


Figure 11. Comparative evaluation of AprilTag detection performance. (**a**) Traditional direct tag detection methods. (**b**) The proposed coarse-to-fine detection method. This strategy integrates the YOLO11n model for initial RoI extraction (coarse), followed by precise tag localization within the cropped bounding box (fine).

The YOLO11n demonstrates superior robustness because it was trained on a dataset explicitly containing samples with varying degrees of occlusion and shadow. As summarized in Table 5, the YOLO11n AprilTag marker detection achieved an *mAP*0.5 of 0.995 and an *mAP*0.5:0.95 of 0.944 on the validation set. Even under identical adverse conditions where the traditional method fails, the model network successfully localizes the tag and generates a precise RoI. This coarse-to-fine detection method isolates the tag from environmental noise, ensuring that the subsequent sub-pixel corner detection proceeds within a clean, interference-free window. The tag detection success rate was significantly improved, as shown in Table 6.

Table 5. Performance metrics of YOLO11n AprilTag marker detection.

Metric	Value
Training image pairs	212
Validation image pairs	53
Precision	1
Recall	0.9953
F1-score	0.9976
*mAP*0.5	0.9950
*mAP*0.5:0.95	0.9439
Inference time	3.8 ms/image
Tag size	40 × 40 mm

Table 6. Quantitative comparison of tag detection performance between the traditional method and the proposed YOLO11n integrated method.

	Left Camera	Right Camera	Combined Total	Detection Rate
Traditional method	190	153	343	80.90%
YOLO11n	212	210	422	99.53%

3.4. Feasibility Verification of the HMVEC Strategy

For each manipulator, 53 predefined target positions were selected within the reachable workspace. At each target position, three control strategies were evaluated sequentially: uncompensated open-loop control, model-based compensation, and the proposed HMVEC strategy. Therefore, 53 target-position trials were conducted for each compensation strategy on each manipulator, resulting in 159 positioning measurements per manipulator and 318 measurements for the dual-manipulator system. The RMSE was calculated from the positioning errors of the 53 target positions under each strategy. The corresponding spatial positions are shown in Figure 12, green dots represent the ideal target points in space. Meanwhile, triangles, squares, and stars show the actual spatial positions reached by the end of the manipulator under different strategies. These results demonstrate the effectiveness of the proposed compensation strategy.

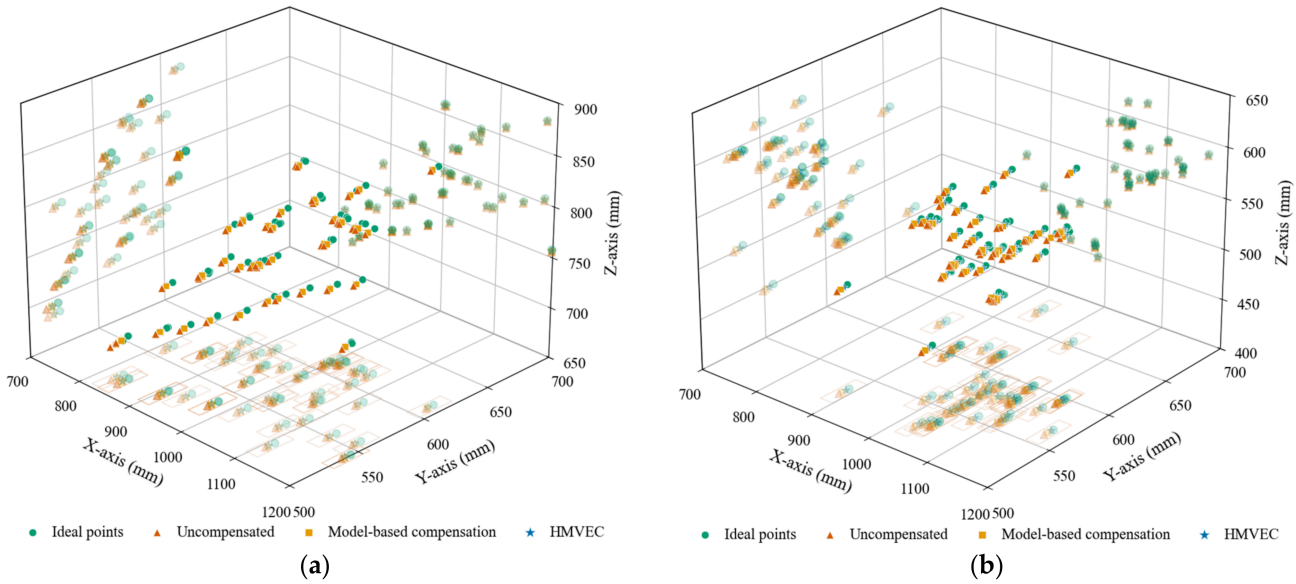


Figure 12. Comparison of 3D trajectories of the manipulator end point under different compensation strategies for the (**a**) left and (**b**) right manipulators.

The proposed HMVEC strategy minimized the spatial positioning errors of the dual-manipulator system. Figure 12 clearly illustrates that the ends of both the left and right manipulators progressively converged toward the ideal target point under different compensation strategies. Overall, the implementation of compensation strategies results in a stepwise reduction in the positioning error of the manipulator, as illustrated in Figure 13. Regarding the *Y*-axis, the direction most susceptible to errors, the hybrid compensation strategy achieved a reduction of 42.70% (from 7.40 mm to 4.24 mm), thereby substantially enhancing the manipulator's accuracy during dynamic operations.

Also, the study utilizes wireframe spheres to visualize the RMSE values as a volume boundary for the average error. This visualization intuitively demonstrates the effectiveness of the proposed HMVEC strategy in reducing the positioning error of the manipulator. The radius of the error spheres for both manipulators decreases significantly when the

system transitions from the uncompensated state (orange) to the hybrid compensated state (blue), as illustrated in Figure 14. This directly shows that the HMVEC strategy reduces the overall positioning error. With this strategy, the positioning error of the dual-manipulator system decreased by approximately 42.28%. The projected points represent the orthogonal projections of the three-dimensional positioning errors onto the X-Y, X-Z, and Y-Z coordinate planes, allowing the directional distribution and dispersion of the errors before and after compensation to be visualized more clearly. The highly consistent error convergence observed across both manipulators not only validates the model's efficacy but also underscores its robustness and universality, confirming its capacity to accommodate individual kinematic variations and ensure reliable precision for cooperative dual-manipulator tasks.

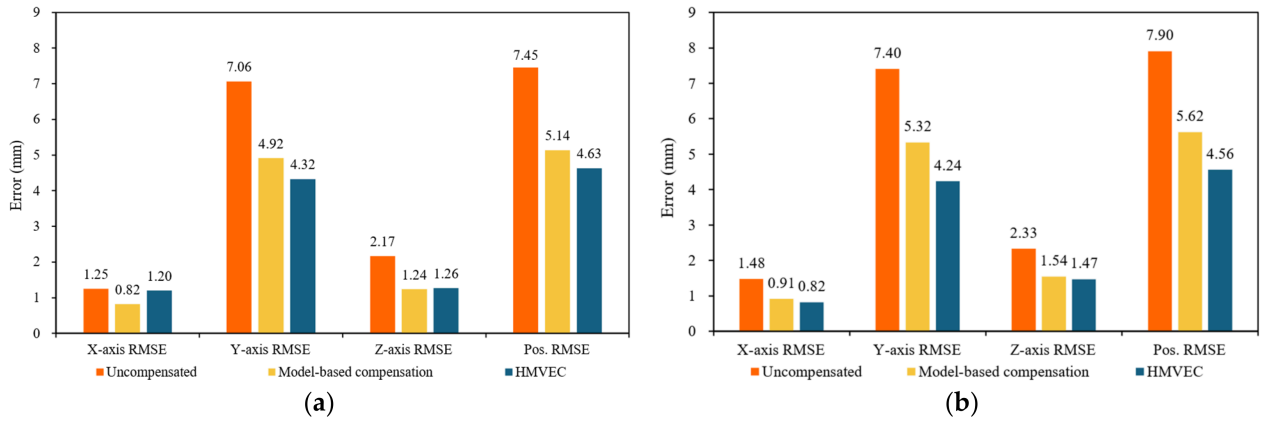


Figure 13. RMSE-based comparative analysis of positioning accuracy under different compensation strategies for the (**a**) left and (**b**) right manipulators, illustrating both individual-axis errors and overall positioning error.

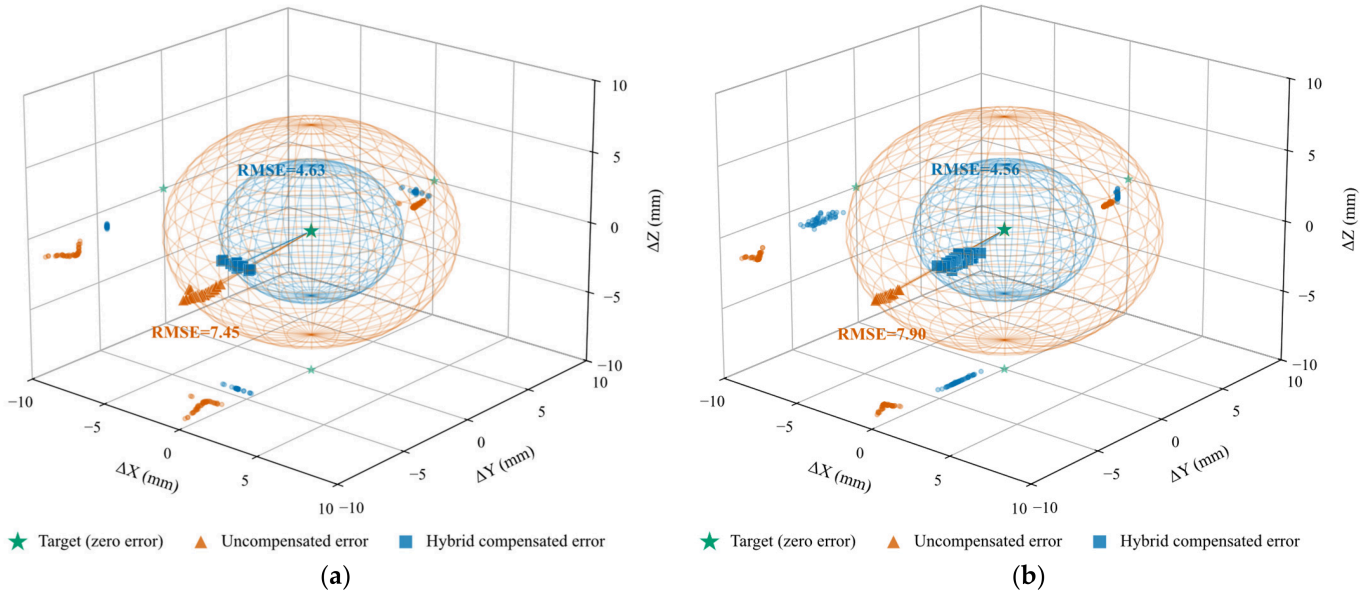


Figure 14. Three-dimensional distributions of positioning errors of the (**a**) left and (**b**) right manipulators under uncompensated and hybrid predictive compensated conditions, where RMSE spheres represent the overall error magnitude and visualize the reduction of positioning errors after compensation.

While the deep learning and MoveIt-based architecture proposed by Singh et al. leverages advanced 3D visual perception [29], its reliance on ideal rigid-body kinematics

renders it incapable of proactively suppressing the nonlinear gravity sag and structural backlash inherent to physical cantilevers. Conversely, empirical closed-loop methods, such as the two-stage fuzzy logic control (2s-FLC) employed by Chang et al., attempt to iteratively compensate for dynamic deviations using heuristic rules tied to visual bounding box areas [30]. Without an explicit mathematical decoupling of the underlying physical deformations, these "black-box" systems suffer from steady-state oscillations and severe tracking failures during partial foliage occlusion. By embedding a polynomial–Fourier kinematic model as a predictive feedforward baseline, the proposed HMVEC strategy analytically cancels out monotonic deflection and periodic transmission errors well before visual intervention occurs. Additionally, the adaptive Hold Strategy utilizes this physical model to seamlessly preserve trajectory integrity during transient visual occlusions. Table 7 further summarizes the main features, strengths, and limitations of the proposed strategy and representative visual feedback compensation methods. This comparison indicates that the proposed HMVEC strategy provides a structured model–vision compensation framework for the telescopic manipulator, although direct quantitative comparison among these methods remains limited by differences in robotic platforms, experimental conditions, and evaluation metrics.

Table 7. Comparison of the proposed HMVEC strategy with representative visual feedback compensation methods.

Method	Main Features	Strengths	Weaknesses
Singh et al. [29]	Deep-learning-based perception and MoveIt-based motion planning	Strong 3D perception and planning capability	Mainly relies on rigid-body kinematics; does not explicitly model structural deformation of telescopic manipulators
Chang et al. [30]	Two-stage fuzzy logic control based on visual feedback	Does not require an explicit physical model; can compensate for visual tracking deviations	Heuristic control rules; may be sensitive to occlusion and lacks explicit compensation for gravity-induced deformation
Proposed HMVEC strategy	Polynomial–Fourier feedforward model + YOLO11n-assisted AprilTag tracking + Hold Strategy	Compensates systematic structural deformation and reduces residual positioning errors; maintains compensation during temporary visual interruption	Currently depends on artificial AprilTag markers and was validated under simulated orchard conditions

3.5. Operational Efficiency and Kiwifruit-Transfer Performance

A simulated operation experiment comprising 60 individual runs was conducted using a single manipulator and artificial kiwifruit targets. The 60 predefined target coordinates were divided into 12 sets, with five coordinates in each set. A corresponding artificial kiwifruit target was placed at each coordinate. Based on the cycle definition described in Section 2.4.3, the data obtained from the 60 runs yielded a mean cycle time of 4.8 s per fruit.

During this experiment, the existing harvesting end-effector successfully separated all 60 artificial kiwifruit targets from their peduncles. Of these, 57 targets successfully entered the hollow section of the manipulator after separation and subsequently passed through the corrugated tube into the collection bin. The remaining three targets failed to enter the hollow section after fruit–peduncle separation and were therefore recorded as fruit-transfer failures. Accordingly, based on Equation (6), the kiwifruit-transfer success rate, S_t, was calculated as:

$$S_t = \frac{57}{60} \times 100\% = 95\%$$

Compared with the articulated robotic kiwifruit harvesting system developed by Williams et al. [17], which reported an average cycle time of approximately 5.5 s per fruit under orchard conditions, the proposed compact telescopic Cartesian structure combined

with the hollow continuous transport mechanism achieved a comparable cycle time of 4.8 s per fruit under simulated operation conditions. It should be noted that the present cycle time and fruit-transfer success rate did not include fruit detection or 3D localization, because the target coordinates were predefined in this experiment. Nevertheless, the results suggest that the proposed structure has the potential to reduce non-productive recovery motion and support continuous fruit transfer during trellis-grown kiwifruit harvesting.

4. Conclusions

This study proposed a novel hollow telescopic Cartesian manipulator, specifically designed for trellis-grown kiwifruit to address the spatial constraints and complex nonlinear positioning errors associated with traditional harvesting robots. The manipulator integrated a nested carbon fiber structure with the HMVEC strategy, which fused kinematic modeling and stereo vision, achieving terminal positioning-error compensation and post-detachment fruit transfer, as well as robust error suppression in structured environments. Its structure featured a two-stage hollow manipulator driven by motor, the system achieves an operational speed of up to 0.5 m/s and saves 45.36% of vertical installation space compared to conventional single-stage solutions. Experimental results obtained in the simulated harvesting environment showed that the HMVEC strategy reduced the overall positioning RMSE from 7.90 mm under uncompensated open-loop control to 4.56 mm, corresponding to a reduction of 42.28%. The HMVEC strategy was also applied during detachment trials using artificial kiwifruit targets to compensate for manipulator positioning errors. Following successful fruit–peduncle separation by the existing harvesting end-effector, the hollow transport structure achieved a kiwifruit-transfer success rate of 95%. The mean system cycle time was 4.8 s per fruit. Overall, this study integrated kinematic-model-based predictive compensation and visual feedback into a hollow telescopic manipulator, thereby offering a potential technical approach for precise operation and post-detachment fruit transfer in robotic kiwifruit harvesting.

While the proposed hollow telescopic manipulator demonstrated promising performance in the simulated operation tests, the current validation still has several limitations. First, the experiments were conducted under simulated orchard conditions using predefined target coordinates, because the online fruit-detection and 3D localization modules were not deployed on the experimental platform. Therefore, the reported cycle time and fruit-transfer success rate should be interpreted as the experimental results obtained when the telescopic manipulator was equipped with the existing harvesting end-effector to perform target alignment, artificial-fruit separation, and post-detachment fruit transfer under simulated conditions. Second, the manipulator currently moves only along a straight line, which restricts its flexibility when dealing with dense fruit clusters or irregular fruit distributions. Third, the current visual feedback system relies on artificial AprilTag markers, which may be susceptible to occlusion by leaves, branches, and dirt in practical orchard environments. Furthermore, prolonged high-speed operation can lead to mechanical wear, thereby degrading the accuracy of the offline compensation model over time.

To enhance the robustness and applicability of the system, future work will focus on several improvements. Mechanically, a compact multi-DOF joint will be integrated at the distal end to facilitate obstacle avoidance and precise pose adjustment. An adjustable tilting base will also be designed to accommodate diverse canopy architectures. In addition, wear-prone components, including the belt transmission, bearings, and mechanical joints, will be periodically inspected and maintained to reduce accumulated clearance and transmission backlash. At the software and control-logic level, future work will implement a residual-monitoring and event-triggered recalibration mechanism. The residual between the offline polynomial–Fourier compensation model prediction and the visual feedback

measurement will be continuously recorded. If the residual exhibits persistent drift or exceeds a preset threshold, maintenance and recalibration will be triggered to compensate for cumulative mechanical wear, clearance variation, and transmission looseness over time. In addition, although the experiments were conducted in an open outdoor site under natural illumination, the validation did not fully reproduce real orchard conditions, such as canopy occlusion, dappled illumination, irregular fruit distribution, branch interference, wind disturbance, and long-duration continuous field operation. Future work will include real-orchard trials to further evaluate the robustness and practical applicability of the proposed system.

Author Contributions: Conceptualization, B.J., C.L. and L.F.; Methodology, B.J., J.Z., R.L., M.Y. and S.H.L.; Software, J.Z.; Validation, S.K., J.H. and X.M.; Formal analysis, B.J.; Investigation, B.J. and S.K.; Data curation, B.J., S.K., J.H. and X.M.; Writing—original draft preparation, B.J.; Writing—review and editing, J.Z., C.L. and L.F.; Supervision, C.L. and L.F.; Funding acquisition, R.L. and L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the National Natural Science Foundation of China (32371999); the Chinese Universities Scientific Fund (2452026021); the Key Research and Development Program of Shaanxi (2024PT-ZCK-27, 2024NC-YBXM-195); the Science and Technology Program of Xi'an City, China (23NYGG0071); and the National Foreign Expert Project, Ministry of Human Resources and Social Security, China (H20240238, Y20240046).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the editors and reviewers for their constructive comments and suggestions. During the preparation of this manuscript, the authors used ChatGPT (gpt-5.1), Gemini (3.1 Pro) and DeepSeek (V3.2) for grammar checking to improve the readability of article. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

References

1. Li, L.; Li, K.; He, Z.; Li, H.; Cui, Y.J. Kiwifruit segmentation and identification of picking point on its stem in orchards. *Comput. Electron. Agric.* **2025**, *229*, 109748. [CrossRef]
2. Fountas, S.; Mylonas, N.; Malounas, I.; Rodias, E.; Hellmann Santos, C.; Pekkeriet, E. Agricultural robotics for field operations. *Sensors* **2020**, *20*, 2672. [CrossRef] [PubMed]
3. 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]
4. Yang, Y.X.; Zhang, M.; Ma, W.; Hu, Y.S. Intelligent batch harvesting of trellis-grown fruits with application to kiwifruit picking robots. *Agronomy* **2025**, *15*, 2499. [CrossRef]
5. He, Z.; Ma, L.; Wang, Y.C.; Wei, Y.Z.; Ding, X.T.; Li, K.; Cui, Y.J. Double-arm cooperation and implementing for harvesting kiwifruit. *Agriculture* **2022**, *12*, 1763. [CrossRef]
6. Ma, Y.H.; Feng, Q.C.; Sun, Y.H.; Guo, X.; Zhang, W.H.; Wang, B.W.; Chen, L.P. Optimized design of robotic arm for tomato branch pruning in greenhouses. *Agriculture* **2024**, *14*, 359. [CrossRef]
7. Mu, L.T.; Cui, G.P.; Liu, Y.D.; Cui, Y.J.; Fu, L.S.; Gejima, Y. Design and simulation of an integrated end-effector for picking kiwifruit by robot. *Inf. Process. Agric.* **2020**, *7*, 58–71. [CrossRef]
8. Lei, X.J.; Liu, J.Z.; Jiang, H.K.; Xu, B.C.; Jin, Y.C.; Gao, J.N. Design and testing of a four-arm multi-joint apple harvesting robot based on singularity analysis. *Agronomy* **2025**, *15*, 1446. [CrossRef]
9. Feng, Q.; Chen, J.; Zhang, M.; Wang, X. Design and test of harvesting robot for table-top cultivated strawberry. In *Proceedings of the 2019 WRC Symposium on Advanced Robotics and Automation*; IEEE: Piscataway, NJ, USA, 2019; pp. 80–85. [CrossRef]
10. Zhang, B.; Xu, S.H.; Xiong, Z.M.; Qin, H.; Ai, X.Y.; Yuan, T.; Li, W. Research on robot control technology of tomato plant lowering in greenhouses. *Agronomy* **2024**, *14*, 1966. [CrossRef]
11. Kang, M.L.; Chen, Q.H.; Fan, Z.M.; Yu, C.; Wang, Y.X.; Yu, X.J. A RRT based path planning scheme for Multi-DOF robots in unstructured environments. *Comput. Electron. Agric.* **2024**, *218*, 108707. [CrossRef]

12. Arikapudi, R.; Vougioukas, S.G. Robotic tree-fruit harvesting with arrays of Cartesian arms: A study of fruit pick cycle times. *Comput. Electron. Agric.* **2023**, *211*, 108023. [CrossRef]
13. Zhang, L.; Ye, T.X.; Zhang, F.R.; Li, J.F. Design and dynamic modelling of chain telescopic robot. *Mech. Mach. Theory* **2023**, *181*, 105204. [CrossRef]
14. Zahid, A.; Mahmud, M.S.; He, L.; Choi, D.; Heinemann, P.; Schupp, J. Development of an integrated 3R end-effector with a Cartesian manipulator for pruning apple trees. *Comput. Electron. Agric.* **2020**, *179*, 105837. [CrossRef]
15. 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]
16. Tian, Z.W.; Guo, X.Y.; Ma, W.; Xue, X.Y. Research on kiwifruit harvesting robot worldwide: A solution for sustainable development of kiwifruit industry. *Smart Agric. Technol.* **2025**, *10*, 100792. [CrossRef]
17. Williams, H.A.M.; 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]
18. Zhang, X.; Xun, Y.; Yang, Q.H.; Wang, Z.H. Design and experiment of the negative pressure adsorption Cartesian robot system for apple harvesting. *Int. J. Agric. Biol. Eng.* **2025**, *18*, 145–153. [CrossRef]
19. Shang, D.Y.; Li, X.P.; Yin, M.; Li, F.J.; Wen, B.C. Rotation angle control strategy for telescopic flexible manipulator based on a combination of fuzzy adjustment and RBF neural network. *Chin. J. Mech. Eng.* **2022**, *35*, 53. [CrossRef]
20. Li, T.; Xie, F.; Zhao, Z.Q.; Zhao, H.; Guo, X.; Feng, Q.C. A multi-arm robot system for efficient apple harvesting: Perception, task plan and control. *Comput. Electron. Agric.* **2023**, *211*, 107979. [CrossRef]
21. Yeshmukhametov, A.; Koganezawa, K.; Yamamoto, Y.; Buribayev, Z.; Mukhtar, Z.; Amirgaliyev, Y. Development of continuum robot arm and gripper for harvesting cherry tomatoes. *Appl. Sci.* **2022**, *12*, 6922. [CrossRef]
22. Li, H.; He, Z.; Wang, Y.C.; Ding, X.T.; Cui, Y.J. Research on the mechanized harvesting strategy for clustered kiwi fruits based on deep reinforcement learning. *Comput. Electron. Agric.* **2025**, *237*, 110686. [CrossRef]
23. Ren, J.; Wang, W.D.; Tian, Y.; He, J.R. YOLOv10-kiwi: A YOLOv10-based lightweight kiwifruit detection model in trellised orchards. *Front. Plant Sci.* **2025**, *16*, 1616165. [CrossRef] [PubMed]
24. Mu, L.T.; Gao, Z.B.; Cui, Y.J.; Li, K.; Liu, H.Z.; Fu, L.S. Kiwifruit detection of far-view and occluded fruit based on improved AlexNet. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 24–34. [CrossRef]
25. Zhou, H.Y.; Wang, X.; Au, W.; Kang, H.; Chen, C. Intelligent robots for fruit harvesting: Recent developments and future challenges. *Precis. Agric.* **2022**, *23*, 1856–1907. [CrossRef]
26. Olson, E. AprilTag: A robust and flexible visual fiducial system. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*; IEEE: Piscataway, NJ, USA, 2011; pp. 3400–3407. [CrossRef]
27. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
28. Salazar, M.; Portero, P.; Pérez, S. Design of a low-cost robotic prototype for fruit harvest with flexible gripping using 3D printing. *Acta Technol. Agric.* **2025**, *29*, 7–17. [CrossRef]
29. Singh, R.; Seneviratne, L.; Hussain, I. A deep learning-based approach to strawberry grasping using a telescopic-link differential drive mobile robot in ros-gazebo for greenhouse digital twin environments. *IEEE Access* **2024**, *13*, 361–381. [CrossRef]
30. Chang, C.L.; Huang, C.C. Design and implementation of an AI-based robotic arm for strawberry harvesting. *Agriculture* **2024**, *14*, 2057. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.