PRELIMINARY PREPRINT VERSION – DO NOT CITE
This is not the version of record. The final version will appear in the Proceedings of IJCAI-ECAI 2026.

Learning to Harvest: VR-Guided Expert Behaviour Capture for Decision Modelling in Agricultural Robots

Yining Lang[1,2], **Zhaoxin Li**[1,2], **Xiujuan Chai**[1*]
[1]Agricultural Information Institute, Chinese Academy of Agricultural Sciences
[2]Key Laboratory of Agricultural Big Data, Ministry of Agriculture and Rural Affairs
chaixiujuan@caas.cn

Abstract

While deep learning has significantly enhanced robotic perception in agriculture, autonomous decision-making in dense and occluded environments remains a persistent challenge. This paper proposes a VR-based expert motion capture framework to bridge this gap by integrating high-fidelity virtual environments with human expertise. By reconstructing a realistic tomato greenhouse in Unity and utilizing 27-point motion capture suits, we captured multi-dimensional expert demonstration data encompassing trajectories, postures, and velocities. In the VR environment, "hard cases" featuring frequent collisions can be selectively generated at a higher frequency, allowing experts to provide concentrated demonstrations of specialized postures that markedly accelerate algorithmic learning efficiency. Crucially, the virtual framework enables multiple experts to repetitively harvest identical targets to match optimal trajectories—a process precluded in reality by the destructive nature of harvesting. Moreover, it bypasses biological growth and seasonal constraints, facilitating continuous, high-frequency experimentation independent of crop maturation. These expert data were subsequently integrated into a multi-task reinforcement learning model as experience replay and reward templates. Experimental results demonstrate that our expert-driven algorithm significantly outperforms baseline methods in harvesting efficiency, collision avoidance, and stability, establishing VR-guided motion capture as a robust and scalable paradigm for advancing agricultural robotics.

Figure 1: VR-guided harvesting: From 3D greenhouse reconstruction (left) to expert motion capture (middle) and robotic decision-making (right), offering new paradigm for data-driven algorithms.

1 Introduction

The accelerated evolution of computer-related technologies has driven digital transformation across many domains, and agriculture has recently experienced explosive growth in intelligent applications [Singh *et al.*, 2022]. From large-scale farmland remote sensing [Weiss *et al.*, 2020] and greenhouse monitoring [Soheli *et al.*, 2022], to localized tasks such as crop phenotyping [Domingues *et al.*, 2022] and harvesting decision-making [Li *et al.*, 2024b], and even to microscopic levels such as genetic breeding [Farooq *et al.*, 2024], artificial intelligence (AI) has become one of the core driving forces for the future of agriculture. Virtual reality (VR) technologies have also played a significant role in smart agriculture [De Oliveira and Corrêa, 2020]. Owing to their unique capability for immersive and interactive simulation, VR-based systems have been successfully applied in crop cultivation education [Kim *et al.*, 2018] and agricultural machinery operation training [Cutini *et al.*, 2023].

In the expanding field of smart agriculture, harvesting robots have emerged as a primary application due to their mature infrastructure. While deep learning has significantly advanced crop perception—with YOLO-based models [Junos *et al.*, 2021; Aljaafreh *et al.*, 2023] and keypoint estimation [Li *et al.*, 2024a] identifying optimal picking positions—autonomous decision-making in occluded, high-density environments remains a critical bottleneck. In particular, algorithm-driven robotic manipulators remain markedly inferior to human experts in terms of posture control and decision-making. This raises a critical question: how can harvesting robots make decisions that more closely resemble those of experienced human pickers? Inspired by this challenge, we explore the integration of motion capture technologies within VR environments to bridge this gap. By constructing high-fidelity virtual spaces, our framework enables the extensive simulation of "hard cases" at scale, providing experts with ample opportunities to perform concentrated demonstrations that markedly accelerate algorithmic learning efficiency. Crucially, the virtual environment allows multiple experts to repetitively harvest identical targets to match optimal trajectories—a process precluded in reality by the destructive nature of harvesting. Moreover, this approach bypasses biological growth and seasonal constraints, permitting continuous experimentation independent of crop maturation. These multi-dimensional data are ultimately leveraged to drive robust robotic decision-making algorithms.

In this work, we focus on tomato harvesting, one of the most widely adopted tasks in agricultural robotics. We first scanned a real tomato greenhouse using a robot-mounted, automatically controlled sliding RGB-D camera. The captured data were reconstructed in Unity [Unity Technologies, 2025] through a series of visual perception and 3D modeling algorithms to approximate the spatial distribution and ripeness of plants and fruits. Next, multiple human experts, equipped with 27-point motion capture suits and VR headsets, performed extensive harvesting operations in the virtual greenhouse. These demonstrations covered a wide range of harvesting difficulties and encoded essential information such as picking order, posture, and speed. Finally, we employed this expert dataset to train a multi-task reinforcement learning model, enabling robots to acquire composite decision-making capabilities for tomato harvesting. We validated our approach in both simulation and real robotic environments. Experimental results show that the expert-data-driven decision algorithm consistently outperforms existing methods in terms of harvesting efficiency and collision rates, while also achieving superior success rates and stability.

In summary, the contributions of this paper are threefold:

- We introduce a novel task of VR-based expert motion capture to drive robotic harvesting decision-making.
- We design a system-level integration of algorithms and hardware through a multi-task reinforcement learning framework, enabling robots to accomplish harvesting decision.
- We demonstrate that expert-data-driven decision algorithms outperform existing approaches in terms of harvesting efficiency, stability, and collision reduction.

2 Related Work

2.1 Tomato Detection and Harvesting Algorithm

Early tomato harvesting robots primarily relied on traditional computer vision and graphics-based techniques, which performed coarse-grained recognition of tomatoes based on features such as color and simple geometric descriptors [Yin *et al.*, 2009; Kondo *et al.*, 1996]. Although these methods enabled basic fruit detection, their limited precision highlighted the constraints of relying solely on shallow visual cues. Later, machine learning algorithms were introduced for tasks such as ripeness estimation [Qiang *et al.*, 2014], offering incremental improvements over these coarse-grained approaches. With the advent of deep learning, perception in harvesting has made a qualitative leap [Liu *et al.*, 2020; Zheng *et al.*, 2022]. For instance, Jun et al. [2021] integrated 3D perception and reconstruction methods to alleviate occlusion issues during fruit detection.

Despite such advances, most current harvesting robots rely on standardized agronomy, such as defoliation, to maintain operability. In realistic, dense planting scenarios, leaf occlusion complicates the determination of viable harvesting postures. A naive approach of aligning the manipulator directly with the fruit along a straight trajectory (i.e., the normal direction of the plant wall) risks plant damage and leads to repeated failures, severely reducing efficiency. Consequently, while fruit perception is foundational, strategic manipulator posture decision is crucial for enhancing performance.

2.2 Robot Control and Collaboration Algorithm

Early research on harvesting robots primarily relied on analytical and optimization-based approaches. Li [Li *et al.*, 2008] studied inverse kinematics for computing joint configurations from end-effector positions. Liang et al. [2010] proposed an optimization-based motion planner for a seven-DOF manipulator, achieving smooth trajectories and precise end-effector positioning. Boryga et al. [2015] introduced the PR-APT method for efficient trajectory planning in only a few computational steps. However, these analytic approaches [Yin JianJun *et al.*, 2012] often lacked robustness in no-standard condition.

With the rise of reinforcement learning, it has become increasingly applied to planning and control in harvesting robots. Li et al. [2024c] applied RL algorithms to optimize parking positions and task allocation, achieving promising results in both areas, while posture control received comparatively less attention. More recently, Li et al. [2024b] proposed a deep RL-based peduncle grasping pose model, training an improved HER-SAC algorithm with digital tomato plant models reconstructed from greenhouse images and geometric measurements. This enabled dynamic, collision-free grasping motions in simulation or standard conditions. However, in more complex in-the-wild scenarios, identifying leaves and plants becomes significantly harder, making collision-free harvesting challenging or reducing completion rates due to skipped difficult targets. Additionally, relying solely on simulated data without large-scale expert demonstrations increases training difficulty and may limit performance compared to human-level harvesting. Furthermore, critical factors like speed modulation in human postures are not effectively captured. Motion capture, recognized as one of the most effective ways to accumulate expert demonstration data, has already been widely adopted in diverse application domains [Menolotto *et al.*, 2020; Salisu *et al.*, 2023]. Building on this foundation, our work aims to bring the integration of human–robot interaction and virtual reality simulation into digital agriculture, leveraging expert demonstration data to enhance decision-making in tomato harvesting tasks.

3 Method

3.1 Problem Formulation

This work develops two independent reinforcement learning (RL) models for a tomato harvesting robot: one for path planning and another for posture decision-making. The input includes the base pose $\mathbf{b}$, spatial positions $\mathbf{x}_i \in \mathbb{R}^3$, maturity

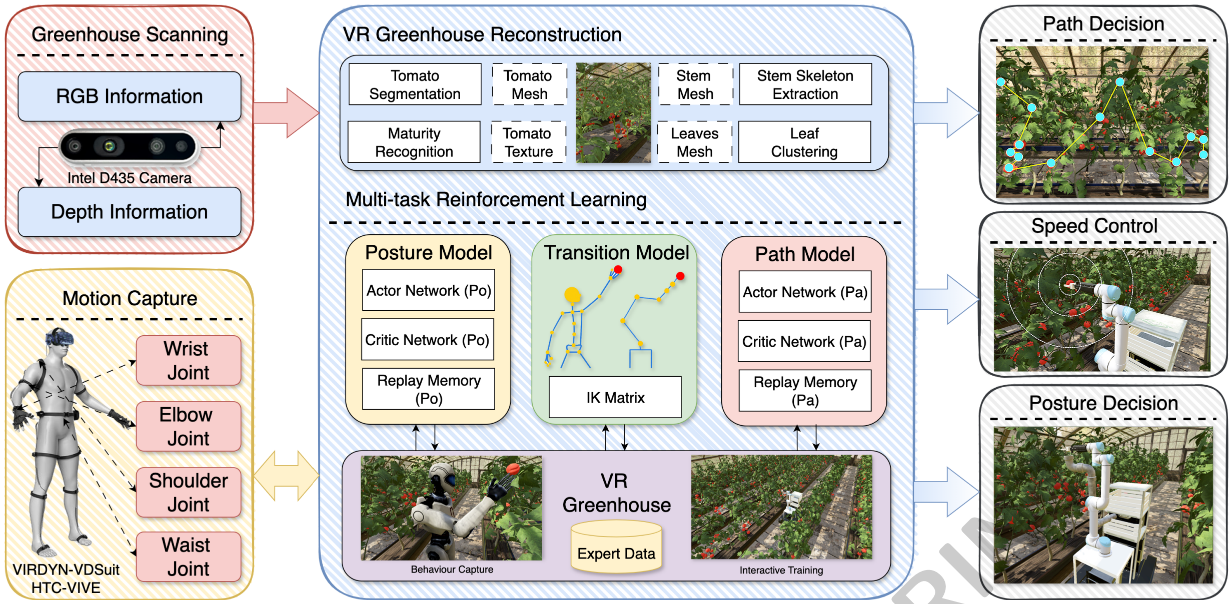

Figure 2: The overall framework of the proposed VR-guided harvesting system. The pipeline integrates three core stages: (1) Real-world perception, where a robot-mounted RGB-D camera scans the greenhouse for 3D plant reconstruction and maturity assessment; (2) Expert demonstration, where motion capture and VR headsets are used to record human harvesting trajectories, postures, and velocities in the reconstructed environment; and (3) Learning and control, where expert data are transformed via inverse kinematics and integrated into a multi-task reinforcement learning framework to drive autonomous robotic decision-making.

levels $m_i \in \{0, 1\}$, and occlusion estimates o_i for all tomatoes $i \in \mathcal{T}$. The output consists of a harvesting sequence $\pi = (g_1, g_2, \ldots, g_K)$, posture actions $\mathbf{u}_t$, and adaptive speed control parameters.

The problem is formulated as two independent reinforcement learning tasks, optimized sequentially. The overall objective is to first optimize path planning and then optimize posture decision-making based on the output of path planning:

$$J(\theta) = \mathbb{E}\left[\sum_{t=0}^{T-1} \gamma^t R_t^{\text{path}} + \gamma^T R^{\text{term}}\right] \quad \text{then} \quad \mathbb{E}_{u\sim\pi_\theta}\left[R_t^{\text{post}}\right], \tag{1}$$

The first term corresponds to the Path RL task, where the agent optimizes the harvesting sequence π based on the robot's state $\mathbf{s}_{\text{path}}$ and action $\mathbf{a}_{\text{path}}$, aiming to minimize travel distance and avoid occlusion. The second term corresponds to the Posture RL task, where the agent optimizes posture actions $\mathbf{u}_t$ based on the state $\mathbf{s}_{\text{posture}}$, ensuring efficient and precise harvesting.

In addition to interaction data from a virtual greenhouse environment, the system is augmented with expert demonstration data collected from professional human pickers wearing motion capture suits in reconstructed greenhouse scenarios. These expert trajectories are stored in the replay memory to accelerate training, while adaptive speed control, based on expert data, is directly encoded into the manipulator's control program to ensure safe and efficient harvesting in real-world operations.

3.2 VR Greenhouse Scanning and Reconstruction

Through a sliding-rail RGB-D camera system (Intel RealSense D435) mounted on the tomato harvesting robot, we scanned a real greenhouse with dimensions of 30 meters in length, 30 meters in width, and 2 meters of inter-row spacing. This real-scene scanning serves two purposes: it provides a realistic reconstructed environment for reinforcement learning training, while also enabling expert demonstration through VR-based motion capture to supply valuable replay experiences for the learning process.

The data collection begins with RGB and depth data captured in a real greenhouse. The RGB data plays a crucial role in identifying key characteristics, such as the ripeness of tomatoes, the density of leaves, and the locations of tomatoes, leaf clusters, and stems in a flat view. On the other hand, depth images offer vital insights into determining the 3D positioning and depth of these elements. The segmentation and identification of tomato fruits are performed through the Mask R-CNN [He *et al.*, 2017] algorithm applied to the RGB images. By integrating the segmentation outcomes with the associated depth information, we can precisely calculate the 3D spatial positions of the tomatoes in the environment. Once the 3D models of the tomatoes and plants are reconstructed, they are integrated into the Unity engine to create a virtual reality greenhouse.

3.3 Expert Demonstration via Motion Capture

Hardware Setup. The VR environment was rendered on a workstation equipped with an NVIDIA RTX 4090 GPU, an Intel i9 processor, and 64 GB of DDR5 memory. For immersive interaction, we used an HTC VIVE PRO 2 headset with

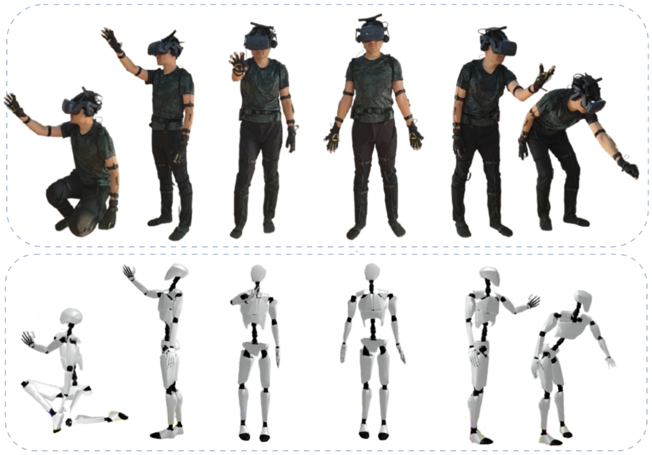

Figure 3: This figure illustrates expert motion-capture sequences that provide harvesting action demonstrations for decision-making.

a wireless receiver. Expert motions were captured using a 27-point inertial motion capture system produced by VIRDYN, which also includes data gloves capable of tracking 10 finger joints. The robotic platform was based on the ELITE EC60, a six-DOF manipulator, outfitted with a custom-designed end-effector specifically tailored for tomato harvesting.

Expert Data Collection. Five expert pickers (height 175–180 cm) were recruited to ensure uniform anthropometric reach and motion consistency. The experiments were conducted in a 10×10 m obstacle-free indoor environment, with participants equipped with an inertial motion capture suit and a VR headset. Following a 3-minute unrecorded warm-up phase for environmental adaptation, each participant performed harvesting tasks across 50 virtual scenes categorized into five difficulty levels based on planting density and fruit distribution. Each scene contained 5–20 standard tomatoes. Participants were granted a 10-second observation window to formulate a strategic harvesting trajectory before executing sequential picking. To align with the mechanical configuration of mainstream harvesting robots—which utilize collection tubes for transport—participants were instructed to release each fruit immediately upon detachment. This protocol ensures that the captured demonstrations reflect the operational logic and kinematic constraints of autonomous robotic systems.

Motion Capture Data Processing. The motion-capture (MoCap) output comprises three parts: (i) tomato-picking trajectories (for target sequencing), (ii) postures of the picking arm segment, and (iii) per-posture velocity profiles. Trajectories align naturally with the RL transition tuples and are directly placed into the replay buffer.

Postures require a kinematic conversion because a human arm and a 6-DOF manipulator differ structurally. We define the expert's effective posture segment as the forearm–hand chain (from elbow to fingertip). This segment is mapped to the robot's tool chain (from joint 6 along the cylindrical actuator to the rotary gripper), which matches in structure and effective length by design. Consequently, for a given tomato position, we set the human elbow position as the robot joint–6 origin, and solve a 6-DOF inverse kinematics (achieved by MATLAB Robotics Toolbox [2025]) to recover a full joint configuration consistent with the expert's forearm.

By this way, expert demonstrations captured via VR-based motion capture, as shown in Figure 3, are converted into joint-space trajectories using inverse kinematics and then injected into the replay buffers to accelerate convergence. These demonstrations provide initial guidance for the reinforcement learning process, allowing the policy to start with a better understanding of the task and environment, thus reducing training time and improving efficiency.

Adaptive Speed Control. Analysis of expert MoCap data reveals a distinct two-phase velocity structure: a target-transfer phase ($>$ 30 cm from target) characterized by high-speed movement ($\sim$ 125% baseline) for efficiency, and a target-approach phase ($\leq$ 30 cm) where speed is halved ($\sim$ 50% baseline) to ensure collision avoidance. This 30 cm threshold serves as a heuristic boundary separating coarse transfer from fine manipulation. To prioritize operational safety and stability, this velocity modulation is directly hard-coded into the manipulator's control firmware rather than learned via reinforcement learning.

3.4 Reinforcement Learning Framework

We formulate the harvesting decision-making problem as a two-level reinforcement learning (RL) framework, consisting of (i) Path RL for global sequence planning and (ii) Posture RL for local grasping. The Path module follows the standard DDPG [Lillicrap *et al.*, 2015] Actor–Critic framework, where the policy predicts continuous harvesting priorities for candidate tomatoes. The Posture module extends the same backbone into a single-step continuous posture optimization formulation tailored for local posture selection.

Path RL: Harvesting Sequence Planning. At each parking location, the robot perceives all tomatoes $\mathcal{T}$ and selects a feasible harvesting sequence that minimizes travel while avoiding severe occlusion. The harvestable set is

$$\mathcal{T}^{+} = \{\, i \in \mathcal{T} \mid m_i = 1 \,\}, \quad \mathcal{T}^{\star} = \{\, i \in \mathcal{T}^{+} \mid \tilde{o}_i \geq o_{\text{thr}} \,\}, \tag{2}$$

where $m_i \in \{0, 1\}$ indicates maturity (1 = harvestable, 0 = unripe), $\tilde{o}_i \in [0, 1]$ is the normalized occlusion score (derived from the confidence score of the tomato detection model, with lower confidence reflecting heavier occlusion), and o_{thr} is the threshold for severe occlusion. The state at time t is:

$$s_t = \Big[b,\ x_t^{ee},\ R_t^{ee},\ \{(x_i,\ \tilde{o}_i,\ \tilde{d}(x_t^{ee}, x_i),\ \text{picked}_i)\}_{i \in \mathcal{T}}\Big], \tag{3}$$

where b is the robot base pose, (x_t^{ee}, R_t^{ee}) are the end-effector position and orientation, x_i is the 3D position of tomato i, $\text{picked}_i \in \{0, 1\}$ indicates whether tomato i has been harvested, and

$$\tilde{d}(x_t^{ee}, x_i) = \frac{\|x_t^{ee} - x_i\|_2}{D_{\max}} \in [0, 1] \tag{4}$$

is the normalized Euclidean distance to the target, with $D_{\max}$ denoting the maximum reachable distance of the workspace.

At each decision step, the actor assigns continuous harvesting priorities to feasible candidate tomatoes, where higher

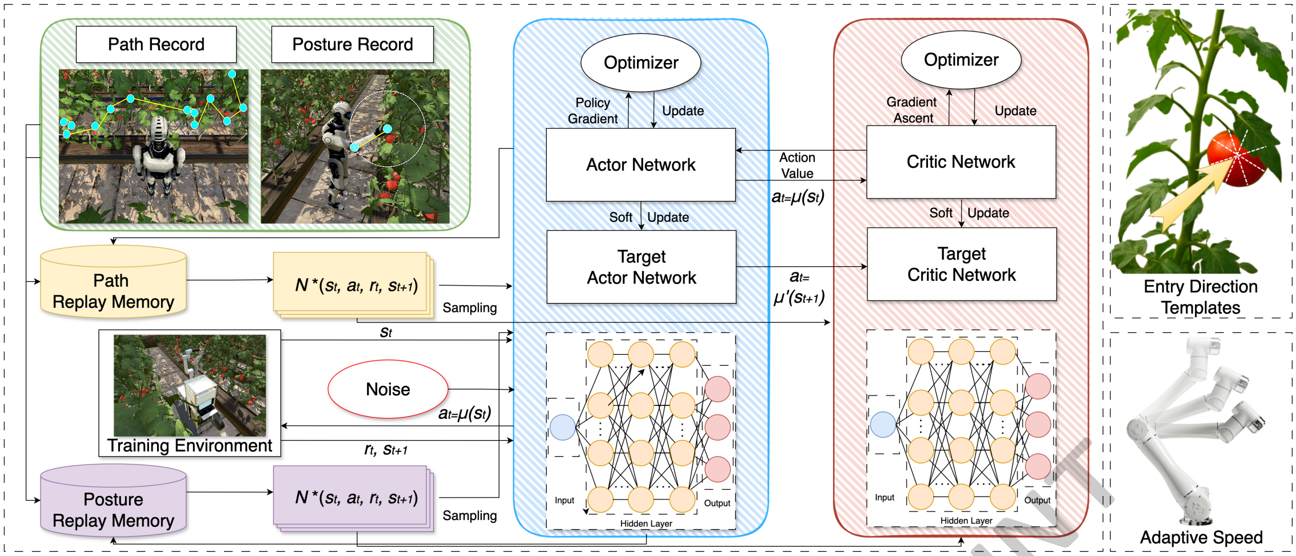

Figure 4: Overview of the RL Framework for Robot Manipulation: Trajectory and posture motion-capture data are organized into two separate replay buffers for Path RL and Posture RL training, where the harvesting robot interacts and trains in a VR environment. The top right corner showcases one of the 8 expert entry direction templates, while the bottom right corner displays adaptive speed control integrated with the reinforcement learning model, enabling tri-dimensional decision-making for harvesting path, posture, and speed.

values correspond to stronger harvesting preference. The episode terminates when $\mathcal{T}^\star \setminus \text{picked} = \varnothing$ (i.e., no harvestable and feasible targets remain, including those excluded due to severe occlusion). The per-step reward is

$$R_t^{\text{path}} = -\alpha_1 \left(1 - \tilde{o}_{g_t}\right) - \alpha_2\, \tilde{d}(x_t^{ee}, x_{g_t}), \tag{5}$$

where g_t is the selected tomato at step t, $\alpha_1, \alpha_2 > 0$ are weighting factors for occlusion and distance penalties (set to 1 by default). A positive terminal bonus R^{term} is granted only if all feasible targets are successfully harvested. The overall objective is

$$J(\theta^{\text{path}}) = \mathbb{E}\left[\sum_{t=0}^{T-1} \gamma^t R_t^{\text{path}} + \gamma^T R^{\text{term}}\right], \tag{6}$$

where $\gamma \in (0, 1]$ is the discount factor. The policy parameters θ^{path} are optimized using an Actor–Critic framework.

Posture RL: Harvesting Posture Planning. Given a target tomato at position $x_g \in \mathbb{R}^3$ and a fixed tool length $L_e > 0$, the contact-side of the end-effector is assumed to coincide with x_g. The policy selects an approach direction $u \in \mathbb{S}_+^2$ on the reachable hemisphere (the plant backside being infeasible). At time t, the state is defined as

$$s_t = \left[x_t^{ee},\ R_t^{ee},\ q_t,\ x_g,\ \omega_g\right], \tag{7}$$

where (x_t^{ee}, R_t^{ee}) is the current end-effector pose, $q_t \in \mathbb{R}^6$ are joint angles, x_g is the target position, and ω_g is an occlusion descriptor extracted from the segmentation results of the target tomato. The candidate endpoint is determined by

$$x^{ee}(u) = x_g - L_e u, \qquad u \in \mathbb{S}_+^2. \tag{8}$$

The end-effector orientation is aligned with the chosen u.

From VR–MoCap data, we construct a template bank of $K = 8$ expert groups, each corresponding to a distinct occlusion direction and containing the ten fastest entry trajectories demonstrated under that pattern. For a given tomato, the segmentation mask is analyzed: if multiple occlusion regions exist, the region with the highest proportion of green pixels (indicating the most severe leaf coverage) is chosen. The template group on the opposite side of this dominant occlusion is then selected, and its ten expert directions are averaged to form the expert-informed target direction n_{tmpl}.

The one-shot posture reward is

$$R_t^{\text{post}} = \beta_{\text{align}} \langle u, n_{\text{tmpl}} \rangle - \beta_{\text{joint}} \|\Delta q(u)\|_1 - \beta_{\text{col}}\, \mathbb{I}[\text{collision}], \tag{9}$$

where $\Delta q(u)$ is the IK displacement from the current configuration. This encourages alignment with expert templates, penalizes unnecessary joint displacement, and discourages collisions. The weighting factors are set by default to $\beta_{\text{align}} = 2$, $\beta_{\text{joint}} = 1$, and $\beta_{\text{col}} = 1$. Since posture decisions are one-shot per target, the optimization goal is simply

$$J(\theta^{\text{post}}) = \mathbb{E}_{u \sim \pi_\theta}\left[R_t^{\text{post}}\right], \tag{10}$$

where π_θ is the posture policy. The policy thus maximizes expected efficiency and safety at each harvesting step.

Training Procedure. Both Path RL and Posture RL are trained independently in the reconstructed VR greenhouse, with four hidden layers whose neuron counts start at 512 and progressively decrease. The training employs the Adam optimizer with a learning rate of 10^{-3} and a batch size of 32. During training, the Actor networks generate continuous actions (harvesting priorities or entry directions), while the Critic networks estimate their state–action values. Target networks are softly updated to stabilize training, and mini-batch sampling from the replay buffers ensures stability. After convergence, the two modules are integrated into a unified framework that produces both the harvesting sequence and the collision-free posture decision for each parking location.

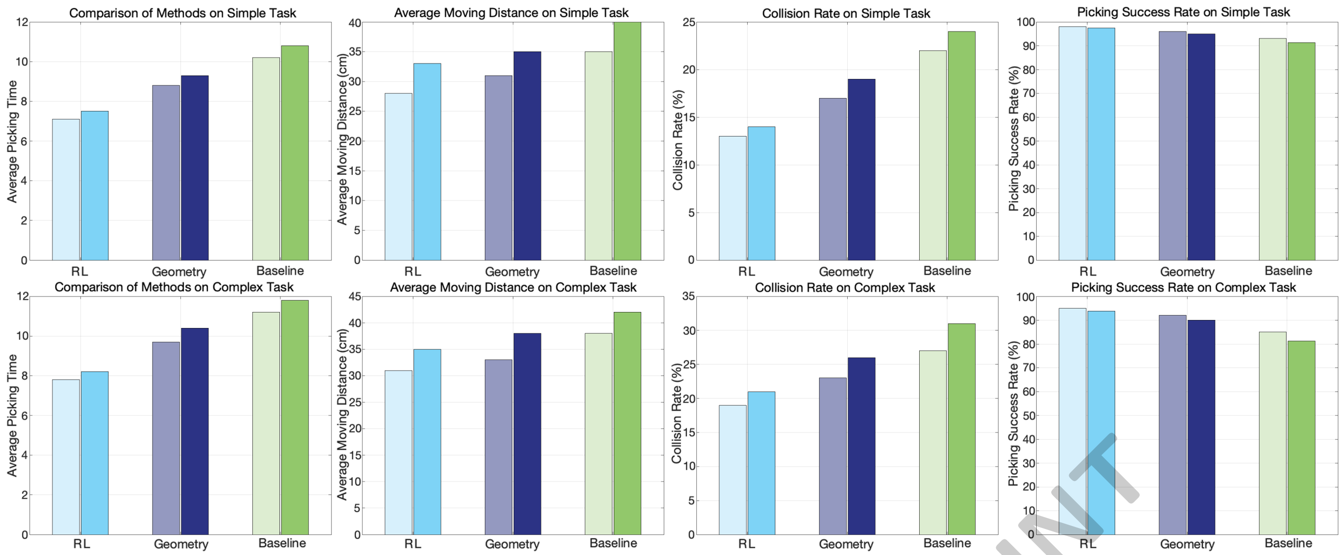

Figure 5: Comparative analysis of RL, geometry-based, and baseline methods. Bars within each group represent fully ripe (first) and mixed ripe (second) fruits. Results are categorized by scenario complexity (Simple vs. Complex) and assessed via four efficiency and safety metrics.

4 Experiments

4.1 Simulation Experiments

Experiment Setup. Four performance metrics were evaluated: average harvesting time per fruit, average end-effector travel distance, collision occurrence rate, and harvesting success rate. Three representative methods were benchmarked: (i) a baseline sequential harvesting strategy, which processes tomatoes from left to right and applies a conventional posture controller [Kondo *et al.*, 1996; Qiang *et al.*, 2014] by first positioning the end-effector directly in front of the target (i.e., along the normal direction of the plant wall) and then executing a straight approach; (ii) a geometry-based method, which combines the commonly adopted A* path planner [Hart *et al.*, 1968] for sequencing with the geometry-based posture algorithm widely used in prior tomato harvesting studies [Zeeshan and Aized, 2023; Boryga *et al.*, 2015]; and (iii) our proposed expert-driven reinforcement learning (RL) framework.

Two levels of task complexity were designed: simple and complex, which differ primarily in the degree of occlusion, with complex scenarios involving more severe leaf and branch interference. In addition, both fully ripe and mixed ripeness conditions were tested. For every scenario, the total number of tomatoes was fixed at 100 (excluding unripe ones from the success rate calculation). To account for safety and robustness, each collision between the manipulator and the plant was recorded. If three consecutive collisions occurred when attempting to harvest the same tomato, the attempt was marked as a failure, and the manipulator was reset to its initial pose before proceeding to the next target.

Results and Analysis. Figure 5 summarizes the performance across four evaluation metrics, where our expert-driven RL framework consistently outperformed both baseline and geometry-based methods, with its advantages becoming more pronounced as task difficulty increased. In terms of efficiency, RL achieved the lowest average harvesting time, ranging from 7.1,s (simple–fully ripe) to 8.2,s (complex–mixed); despite increased occlusion, it maintained a $\sim$20% improvement over the geometry-based method and a $>$ 30% gain over the baseline. This robustness stems from the joint optimization of harvesting sequences and postures, which effectively eliminates redundant motions and failed attempts. Kinematically, RL reduced end-effector travel distance to 28–35,cm, whereas geometry-based and baseline methods required 31–38,cm and 35–42,cm, respectively, as the learned policy exploited direct entry directions to bypass sub-optimal detours in complex scenes. The benefit of expert-informed control was most evident in safety metrics: RL maintained a low collision rate (13–21%) even in dense environments, compared to 17–26% for the geometry-based method and up to 31% for the baseline. Unlike traditional heuristics that often force collisions with occluded targets, our planner strategically skipped low-reward fruits to preserve plant integrity. Consequently, RL delivered a stable success rate of 93.8–98%, outperforming the geometry-based and baseline methods by nearly 4 and over 12 percentage points in the complex–mixed scenario.

4.2 Experiments on a Real Robotic System

Experiment Setup. In this section, we evaluate the expert-driven RL decision-making framework when deployed on a real robotic manipulator. The objective is to validate whether the harvesting trajectories and posture decisions learned in simulation, guided by expert demonstrations, can effectively transfer to a physical platform. The trained policy was integrated into the ELITE EC60 six-DOF manipulator equipped with our custom-designed pipeline-style end-effector.

Unlike the simulation environment, comparing decision-making algorithms on real hardware is inherently more challenging, as robot embodiments, end-effector structures, and experimental environments differ significantly across studies. To ensure fairness and comparability, we adopt two widely

Method	Time (s)	Success (%)
Ours	7.2	91.0
Novel-EE [Park *et al.*, 2023]	15.5	80.6
3MSP2 [Dai *et al.*, 2024]	8.6	70.9
Pose-Rec [Rong *et al.*, 2022]	14.6	72.1
PID [Lili *et al.*, 2017]	15.0	86.0
HIS [Feng *et al.*, 2015]	24.0	83.9
Light-FS [Shi *et al.*, 2023]	17.4	92.9
GA [Wang *et al.*, 2023]	20.0	88.0

Table 1: Harvest Time and Success Rate for Different Methods.

Method	T(s)	D(cm)	C(%)	S(%)
RL (expert)	8.2	35	21	93.8
Path only (expert)	9.8	36	28	87.5
Posture only (expert)	10.2	40	25	88.8
RL (no expert)	9.5	37	24	91.3
RL (no speed-control)	8.7	35	23	93.8
Baseline (sequential)	11.8	42	31	81.3

Table 2: Ablation study results under complex–mixed ripe scenario (Time, Distance, Collision rate, and Success rate).

recognized evaluation metrics in the agricultural robotics literature: (i) average harvesting time per fruit, which directly reflects operational efficiency, and (ii) harvesting success rate, which captures robustness and practical viability. These two metrics are also the most relevant for large-scale deployment in smart agriculture applications. In our experiments, each harvesting view contained 5–20 tomatoes, with a total of 100 mature fruits tested across trials. For the baseline comparisons, the reported results of alternative methods are directly taken from their papers, since their hardware configurations and experimental conditions could not be fully replicated.

Results and Analysis. Table 1 benchmarks our expert-driven RL framework against representative baselines. Our approach achieved the shortest average harvesting time per fruit (7.2,s) with a high success rate (91.0%). Compared to non-expert-driven methods, our framework demonstrated substantial temporal gains: reducing harvesting time by 53.5% versus Novel-EE (15.5,s), 50.7% versus Pose-Rec (14.6,s), 52.0% versus PID (15.0,s), and 70.0% versus HIS (24.0,s). While Light-FS and GA yielded relatively high success rates (92.9% and 88.0%, respectively), their cycle times—17.4,s and 20.0,s—were over twice as slow as our method. Conversely, the second-fastest method, 3MSP2 (8.6,s), exhibited a significantly lower success rate (70.9%), highlighting a speed–reliability trade-off that our method effectively circumvents. To evaluate overall productivity in large-scale scenarios, we analyzed the throughput metric—successful picks per minute. Our framework achieved 7.58 fruits/min, significantly outperforming PID (3.44), 3MSP2 (4.95), GA (2.64), Novel-EE (3.12), Pose-Rec (2.96), HIS (2.10), and Light-FS (3.20). Despite inherent hardware and protocol variations across studies, our approach consistently demonstrated superior throughput. This performance is attributed to a dual-synergy: (i) the unified expert-informed policy that optimizes path planning, posture selection, and variable-speed motion, and (ii) an integrated hardware design featuring a scalable soft-tube collection system. By extending the tube beneath the rotating claw immediately upon detachment, the system ensures the fruit rolls smoothly into a cushioned storage box.

4.3 Ablation Study

To quantify the individual contributions of the framework's components, we conducted ablation experiments under the most challenging complex–mixed ripe scenario, comparing the full expert-driven RL framework against five variants: path-only (expert), posture-only (expert), full RL without expert data, full RL without adaptive speed control, and a sequential baseline. As summarized in Table 2, the full framework achieved the optimal balance across all metrics, yielding an average harvesting time of 8.2,s—a significant improvement over the path-only (9.8,s), posture-only (10.2,s), and no-expert (9.5,s) variants. While the no-speed-control version maintained a comparable travel distance (35,cm) and success rate (93.8%), its harvesting time increased to 8.7,s, underscoring the efficiency gains from adaptive velocity modulation. Furthermore, the integration of expert demonstrations proved critical for operational safety; the full model maintained the lowest collision rate (21%) and the highest success rate (93.8%), notably outperforming the no-expert model (24% collision, 91.3% success) and the baseline (31% collision, 81.3% success). By incorporating rich experience data and fitting the fastest expert postures to provide high-quality priors, the agent effectively emulated professional strategies to resolve complex grasping tasks with fewer attempts.

5 Conclusion

Despite significant advancements in robotic perception, autonomous decision-making remains a bottleneck for harvesting in dense, occluded environments. This paper addresses the limitations of conventional heuristics and exploration-limited reinforcement learning by proposing a VR-based expert motion capture framework. By reconstructing real-world tomato greenhouses in Unity and mapping professional harvesting demonstrations to manipulator kinematics via inverse kinematics, we successfully integrated human expertise into a multi-task RL model for joint path planning and posture control. Comprehensive evaluations demonstrate that this expert-driven approach effectively bridges the gap between machine logic and human skill, achieving stable, efficient, and low-damage harvesting across diverse scenarios. While currently validated on tomatoes, the proposed paradigm offers a scalable solution for diverse crops.

Acknowledgments

This work was supported by the Basic Research Center, Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-BRC-SAE-2025-01), the Jiangsu Agriculture Science and Technology Innovation Fund (CX(24)1021), and the Central Public-interest Scientific Institution Basal Research Fund(JBYW-AII-2026-21).

References

[Aljaafreh *et al.*, 2023] Ahmad Aljaafreh, Ezzaldeen Y ELZAGZOUG, Jafar Abukhait, Abdel-Hamid Soliman, Saqer S ALJA'AFREH, Aparajithan Sivanathan, and James Hughes. A real-time olive fruit detection for harvesting robot based on yolo algorithms. *Acta Technologica Agriculturae*, 26(3):121–132, 2023.

[Boryga *et al.*, 2015] Marek Boryga, Andrzej Graboś, Paweł Kołodziej, Krzysztof Gołacki, and Zbigniew Stropek. Trajectory planning with obstacles on the example of tomato harvest. *Agriculture and Agricultural Science Procedia*, 7:27–34, 2015.

[Cutini *et al.*, 2023] Maurizio Cutini, Carlo Bisaglia, Massimo Brambilla, Andrea Bragaglio, Federico Pallottino, Alberto Assirelli, Elio Romano, Alessandro Montaghi, Elisabetta Leo, Marco Pezzola, et al. A co-simulation virtual reality machinery simulator for advanced precision agriculture applications. *agriculture*, 13(8):1603, 2023.

[Dai *et al.*, 2024] Nianzu Dai, Jiaming Fang, Jin Yuan, and Xuemei Liu. 3msp2: Sequential picking planning for multi-fruit congregated tomato harvesting in multi-clusters environment based on multi-views. *Computers and Electronics in Agriculture*, 225:109303, 2024.

[De Oliveira and Corrêa, 2020] Monique Emídio De Oliveira and Cléber Gimenez Corrêa. Virtual reality and augmented reality applications in agriculture: a literature review. In *2020 22nd symposium on virtual and augmented reality (svr)*, pages 1–9. IEEE, 2020.

[Domingues *et al.*, 2022] Tiago Domingues, Tomás Brandão, and João C Ferreira. Machine learning for detection and prediction of crop diseases and pests: A comprehensive survey. *Agriculture*, 12(9):1350, 2022.

[Farooq *et al.*, 2024] Muhammad Amjad Farooq, Shang Gao, Muhammad Adeel Hassan, Zhangping Huang, Awais Rasheed, Sarah Hearne, Boddupalli Prasanna, Xinhai Li, and Huihui Li. Artificial intelligence in plant breeding. *Trends in Genetics*, 40(10):891–908, 2024.

[Feng *et al.*, 2015] Qingchun Feng, Xiaonan Wang, Guohua Wang, and Zhen Li. Design and test of tomatoes harvesting robot. In *2015 IEEE international conference on information and automation*, pages 949–952. IEEE, 2015.

[Hart *et al.*, 1968] Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[He *et al.*, 2017] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.

[Jun *et al.*, 2021] Jongpyo Jun, Jeongin Kim, Jaehwi Seol, Jeongeun Kim, and Hyoung Il Son. Towards an efficient tomato harvesting robot: 3d perception, manipulation, and end-effector. *IEEE access*, 9:17631–17640, 2021.

[Junos *et al.*, 2021] Mohamad Haniff Junos, Anis Salwa Mohd Khairuddin, Subbiah Thannirmalai, and Mahidzal Dahari. An optimized yolo-based object detection model for crop harvesting system. *IET Image Processing*, 15(9):2112–2125, 2021.

[Kim *et al.*, 2018] Jun-Gyu Kim, In-bok Lee, Kwang-Sik Yoon, Tae-hwan Ha, Rack-woo Kim, Uk-hyeon Yeo, and Sang-yeon Lee. A study on the trends of virtual reality application technology for agricultural education. *Journal of Bio-Environment Control*, 27(2):147–157, 2018.

[Kondo *et al.*, 1996] Naoshi Kondo, Yoshiaki Nishitsuji, Peter P Ling, and K Chong Ting. Visual feedback guided robotic cherry tomato harvesting. *Transactions of the ASAE*, 39(6):2331–2338, 1996.

[Li *et al.*, 2008] Zhiguo Li, Jizhan Liu, Pingping Li, and Wei Li. Analysis of workspace and kinematics for a tomato harvesting robot. In *2008 International Conference on Intelligent Computation Technology and Automation (ICICTA)*, volume 1, pages 823–827. IEEE, 2008.

[Li *et al.*, 2024a] Xingxu Li, Nan Ma, Yiheng Han, Shun Yang, and Siyi Zheng. Ahppebot: Autonomous robot for tomato harvesting based on phenotyping and pose estimation. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 18150–18156. IEEE, 2024.

[Li *et al.*, 2024b] Yajun Li, Qingchun Feng, Yifan Zhang, Chuanlang Peng, Yuhang Ma, Cheng Liu, Mengfei Ru, Jiahui Sun, and Chunjiang Zhao. Peduncle collision-free grasping based on deep reinforcement learning for tomato harvesting robot. *Computers and Electronics in Agriculture*, 216:108488, 2024.

[Li *et al.*, 2024c] Yajun Li, Qingchun Feng, Yifan Zhang, Chuanlang Peng, and Chunjiang Zhao. Intermittent stop-move motion planning for dual-arm tomato harvesting robot in greenhouse based on deep reinforcement learning. *Biomimetics*, 9(2):105, 2024.

[Liang and Wang, 2010] Xi Feng Liang and Yong Wei Wang. Motion planning of tomato harvesting manipulators with 7 dof. *Advanced Materials Research*, 97:2840–2844, 2010.

[Lili *et al.*, 2017] Wang Lili, Zhao Bo, Fan Jinwei, Hu Xiaoan, Wei Shu, Li Yashuo, Qiangbing Zhou, and Wei Chongfeng. Development of a tomato harvesting robot used in greenhouse. *International Journal of Agricultural and Biological Engineering*, 10(4):140–149, 2017.

[Lillicrap *et al.*, 2015] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[Liu *et al.*, 2020] Guoxu Liu, Joseph Christian Nouaze, Philippe Lyonel Touko Mbouembe, and Jae Ho Kim. Yolo-tomato: A robust algorithm for tomato detection based on yolov3. *Sensors*, 20(7):2145, 2020.

[Matlab, 2025] Matlab. *Inverse Kinematics*. MathWorks, 2025. https://ww2.mathworks.cn/help/robotics/inverse-kinematics.html?s_tid=CRUX_lftnav.

[Menolotto *et al.*, 2020] Matteo Menolotto, Dimitrios-Sokratis Komaris, Salvatore Tedesco, Brendan O'Flynn, and Michael Walsh. Motion capture technology in industrial applications: A systematic review. *Sensors*, 20(19):5687, 2020.

[Park *et al.*, 2023] Yonghyun Park, Jaehwi Seol, Jeonghyeon Pak, Yuseung Jo, Jongpyo Jun, and Hyoung Il Son. A novel end-effector for a fruit and vegetable harvesting robot: mechanism and field experiment. *Precision Agriculture*, 24(3):948–970, 2023.

[Qiang *et al.*, 2014] Lü Qiang, Cai Jianrong, Liu Bin, Deng Lie, and Zhang Yajing. Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. *International Journal of Agricultural and Biological Engineering*, 7(2):115–121, 2014.

[Rong *et al.*, 2022] Jiacheng Rong, Pengbo Wang, Tianjian Wang, Ling Hu, and Ting Yuan. Fruit pose recognition and directional orderly grasping strategies for tomato harvesting robots. *Computers and Electronics in Agriculture*, 202:107430, 2022.

[Salisu *et al.*, 2023] Sani Salisu, Nur Intan Raihana Ruhaiyem, Taiseer Abdalla Elfadil Eisa, Maged Nasser, Faisal Saeed, and Hussain A Younis. Motion capture technologies for ergonomics: A systematic literature review. *Diagnostics*, 13(15):2593, 2023.

[Shi *et al.*, 2023] Yinggang Shi, Shiteng Jin, Yiming Zhao, Yujia Huo, Li Liu, and Yongjie Cui. Lightweight force-sensing tomato picking robotic arm with a "global-local" visual servo. *Computers and Electronics in Agriculture*, 204:107549, 2023.

[Singh *et al.*, 2022] Gurjeet Singh, Naresh Kalra, Neetu Yadav, Ashwani Sharma, and Manoj Saini. Smart agriculture: a review. *Siberian Journal of Life Sciences and Agriculture*, 14(6):423–454, 2022.

[Soheli *et al.*, 2022] Sultana Jahan Soheli, Nusrat Jahan, Md Bipul Hossain, Apurba Adhikary, Ashikur Rahman Khan, and M Wahiduzzaman. Smart greenhouse monitoring system using internet of things and artificial intelligence. *Wireless Personal Communications*, 124(4):3603–3634, 2022.

[Unity Technologies, 2025] Unity Technologies. Unity real-time development platform. https://unity.com/, 2025.

[Wang *et al.*, 2023] Dong Wang, Yongxiang Dong, Jie Lian, and Dongbing Gu. Adaptive end-effector pose control for tomato harvesting robots. *Journal of Field Robotics*, 40(3):535–551, 2023.

[Weiss *et al.*, 2020] Marie Weiss, Frédéric Jacob, and Grgory Duveiller. Remote sensing for agricultural applications: A meta-review. *Remote sensing of environment*, 236:111402, 2020.

[Yin *et al.*, 2009] Hongpeng Yin, Yi Chai, Simon X Yang, and Gauri S Mittal. Ripe tomato recognition and localization for a tomato harvesting robotic system. In *2009 International Conference of Soft Computing and Pattern Recognition*, pages 557–562. IEEE, 2009.

[Yin JianJun *et al.*, 2012] Yin JianJun Yin JianJun, Wu ChuanYu Wu ChuanYu, Simon X Yang, GS Mittal, and Mao HanPing Mao HanPing. Obstacle-avoidance path planning of robot arm for tomato-picking robot. *Transactions of the Chinese Society of Agricultural Machinery*, 2012.

[Zeeshan and Aized, 2023] Sadaf Zeeshan and Tauseef Aized. Performance analysis of path planning algorithms for fruit harvesting robot. *Journal of Biosystems Engineering*, 48(2):178–197, 2023.

[Zheng *et al.*, 2022] Taixiong Zheng, Mingzhe Jiang, Yongfu Li, and Mingchi Feng. Research on tomato detection in natural environment based on rc-yolov4. *Computers and Electronics in Agriculture*, 198:107029, 2022.