Online ISSN: 3104-9338
Print ISSN: 3104-932X

Design of a Dynamic Obstacle Avoidance System for Greenhouse Robots Based on the Fusion of YOLOv5 and Ultrasonic Sensing

Yixi Li*, Hengjia Xiang, Xu Wang

Chengdu College of University of Electronic Science and Technology of China

Received: January 25, 2026

Revised: January 26, 2026

Accepted: February 5, 2026

Published online: February 11, 2026

To appear in: *International Journal of Advanced AI Applications*, Vol. 2, No. 3 (March 2026)

*Corresponding Author: Yixi Li (2799480283@qq.com)

Abstract. To address the limitations of single-sensor obstacle avoidance in greenhouse environments characterized by complex lighting and diverse obstacle morphologies, this study proposes and implements a dynamic obstacle avoidance system based on the fusion of visual and ultrasonic multi-sensor information. The system employs an STM32F407 microcontroller as the real-time control core and a Canaan K230 edge AI processor as the dedicated visual processing unit, constructing a hardware platform that integrates a lightweight YOLOv5s model, multiple ultrasonic sensors, and a dual-track drive module. The model is trained and converted using "VSCode + PyTorch," while embedded software programming is accomplished with "Keil MDK + CanMV IDE," enabling a complete implementation from algorithm to hardware. This paper achieves a complementary advantage of long-range observation and short-range avoidance by fusing semantic recognition results with rangig data in real-time on the STM32. Experimental validation in a simulated greenhouse environment demonstrate that the proposed fusion system attains an obstacle avoidance success rate of 94%, with an average decision response time of 305 milliseconds, representing a significant performance enhancement over single-sensor solutions. This research provides a low-cost, robust, and practical solution for autonomous navigation of mobile robots in greenhouses.

Keywords: *Greenhouse Robot; Dynamic Obstacle Avoidance; YOLOv5; K230*

1. Introduction

The intellectualization of agriculture is a dominant trend in modern farming development and a crucial pathway for optimizing crop growth environments and improving yield quality. Consequently, research on greenhouse robots holds substantial economic and engineering

significance. In recent years, extensive research in smart agriculture has been conducted by universities and enterprises in China, yielding fruitful outcomes. For instance, in 2022, Wang et al. developed a mobile, parallel, multifunctional agricultural robot capable of efficient irrigation and stable operation on complex terrain [1]. In 2023, Li et al. designed a rail-based greenhouse inspection robot for rapid and precise acquisition of real-time crop growth data [2]. In 2024, Liang et al. enhanced the path planning for greenhouse mobile robots by introducing an adaptive adjustment function into the traditional ant colony algorithm [3]. Subsequently, in 2025, Meng et al. further reduced the number of robots turns by performing secondary optimization on paths generated by an improved ant colony algorithm [4]. While these studies have made considerable progress in terrain navigation and path planning, integrating machine vision can substantially improve the stability and efficiency of dynamic robot movement. Therefore, building upon these foundations, this paper designs and implements a dynamic obstacle avoidance system based on the fusion of YOLOv5 and ultrasonic sensing, tailored to the practical demands of greenhouse environments.

2. System Overall Design

Table 1. Functional and Modular Correspondence.

Function	Module
Visual Obstacle Avoidance	Visual Module
Bluetooth Remote Control	Bluetooth Module
Ultrasonic Obstacle Avoidance	Ultrasonic Module
Drive Robot Movement	Motor Drive Module
Display Relevant Data	Display Module

This study focuses on strawberry-cultivating greenhouses as the research object, designing an agricultural robot for internal inspection [5]. According to greenhouse requirements, the robot must incorporate visual, Bluetooth, ultrasonic, motor drive, and display modules. The corresponding implementation modules are detailed in Table 1.

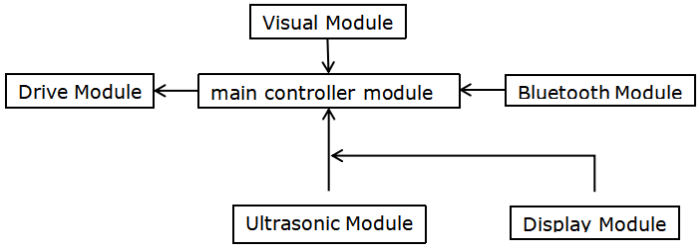

Figure 1. System Block Diagram.

Based on the system's functional and modular correspondence outlined in Table 1, the overall system layout is depicted in Figure 1

3. System Hardware Design

3.1. Construction of Robot 3D Model

The robot comprises the following modules: vision, Bluetooth, ultrasonic, motor drive, and display. The integrated 3D-printed assembly is presented in Figure 2.

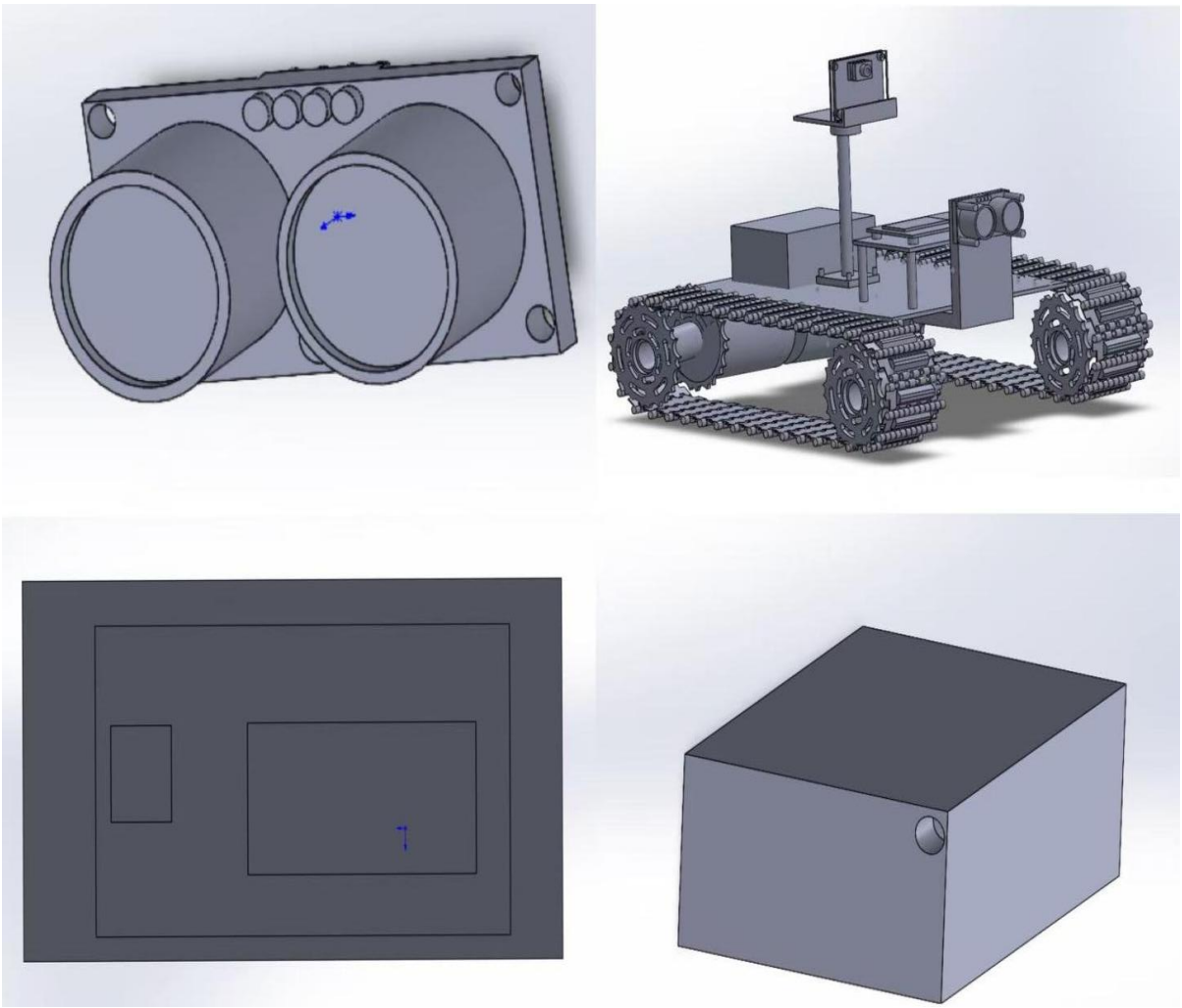

Figure 2. All 3D Printed Modules.

3.2. System Control Circuit Design

3.2.1. Main Control Unit

To achieve high-precision PWM control over multiple motors, efficient multi-channel serial communication, outstanding real-time performance in complex task processing, and flexibility for future external device expansions, the STM32F407ZGT6 was selected as the main control unit after comprehensive consideration (Figure 3).

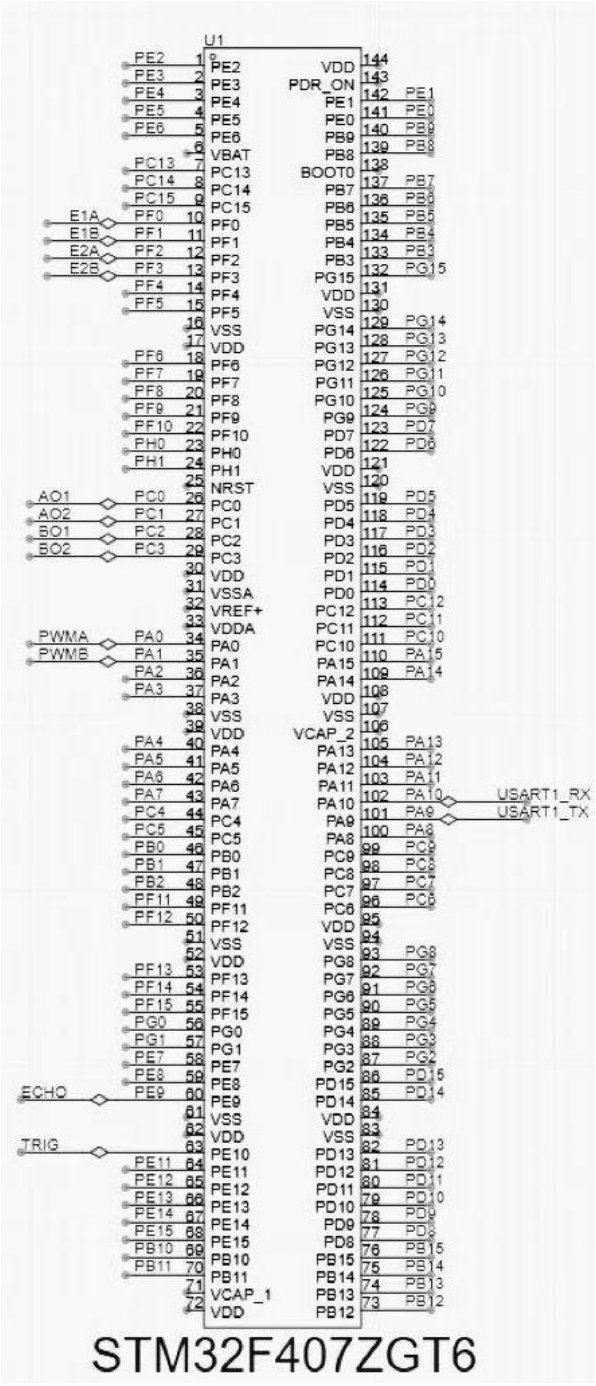

Figure 3. Minimum System Diagram.

Based on the high-performance ARM Cortex-M4 core with a main frequency of 168MHz, this main control unit rapidly processes complex tasks, providing a solid foundation for system real-time performance [6]. Its rich peripheral interfaces, including UART, SPI, I2C, and CAN, easily meet the demands of multi-motor PWM control and multi-channel serial communication, facilitating system expansions for various external devices.

Additionally, its mature ecosystem, supported by STM32CubeMX and HAL libraries, significantly simplifies the development process and enhances efficiency.

3.2.2. Drive and Motion Unit Control

The drive system adopts a dual-track differential structure, enabling flexible adaptation to various complex terrains and providing a stable moving foundation for the device.

Equipped with a 12V DC reduction motor, the system offers robust power, while a Hall encoder (390 pulses per revolution) accurately perceives motor rotation information. The STM32 collects pulse signals using the advanced timer's encoder interface, ensuring reliable high-precision speed feedback. For motion control, a dual-loop PID architecture is employed, where the inner current loop is hardware-implemented by the motor driver for precise and stable current control, and the outer speed loop is software-implemented by the STM32. The speed loop PID output is cleverly converted into a PWM duty cycle, which is then precisely controlled by the TB6612 driver to regulate motor speed (Figure 4). This design enables independent, rapid, and stable closed-loop regulation of the left and right wheels, ensuring precise execution of instructions for both straight-line travel and obstacle avoidance turns, effectively coping with various complex scenarios and laying a solid foundation for stable device operation and accurate manipulation [7].

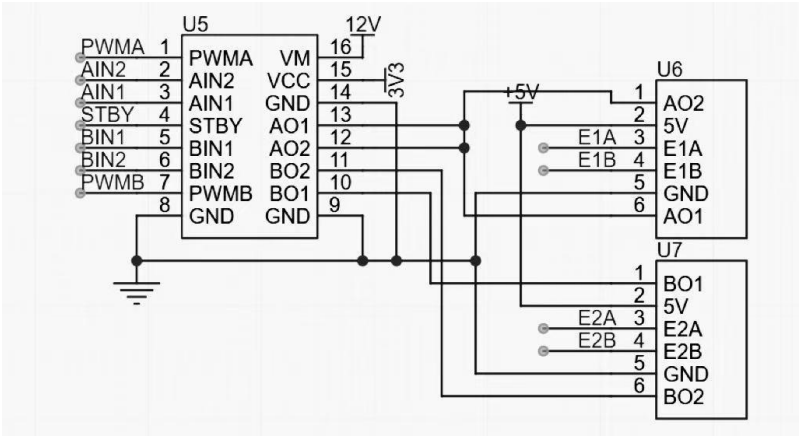

Figure 4. Motor Hardware Assembly Module.

3.2.3. Ultrasonic Sensing Module

The HC-SR04 ultrasonic sensor module is utilized for distance measurement (Figure 5). This module features a broad measurement range of 2 cm to 400 cm with an accuracy of ±0.3 cm, providing reliable data for precise obstacle avoidance and spatial awareness [8].

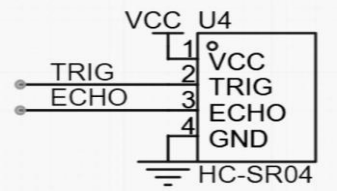

Figure 5. Ultrasonic Hardware Module Diagram.

3.2.4. Visual Sensing Module

In agricultural robot applications, the selection of a hardware visual module must consider multiple factors to ensure efficient, stable, and accurate visual perception tasks. The visual module is based on a lightweight improved YOLOv5 model deployed on the Jia Chuang K230 processor. This module serves dual functions:

- Early identification and classification of medium to long-distance (>50cm) obstacles, distinguishing targets such as plants, personnel, and equipment, providing the robot with lead time for smooth deceleration or advance trajectory planning, minimizing crop damage during operations.
- Detection of targets within the ultrasonic blind zone, such as low-lying tools and short seedlings, compensating for the limitations of ultrasonic detection dimensions.

The K230 processor integrates dual RISC-V cores and a dedicated AI acceleration engine, supporting efficient inference of the lightweight YOLOv5 model through hardware-level optimizations, achieving a good balance between power consumption and computational capability.

Practical tests demonstrate that the processor runs the lightweight YOLOv5 model with a single-frame recognition time of approximately 8ms, maintaining high recognition accuracy under complex lighting conditions such as strong sunlight and backlighting (over 92% in midday strong sunlight and over 88% in backlighting), meeting the real-time requirements of agricultural robots.

In terms of the overall solution, the integrated visual module combines the core processor, image sensor, storage device, etc., into a single unit, featuring compact size, high integration, and ease of integration into existing systems.

4. System Software Design

4.1. System Working Principle and Flow

Upon system startup, the robot maintains straight travel at a constant speed, continuously perceiving the environment through ultrasonic and visual sensors. In the absence of obstacles, it continues straight travel; when obstacles are detected, it enters the fusion decision-making module, calculates the obstacle's position, type, and distance, generates a safe path, and executes obstacle avoidance maneuvers before continuing cyclic perception to ensure stable operation in complex environments. The working principle flow is depicted in Figure 6.

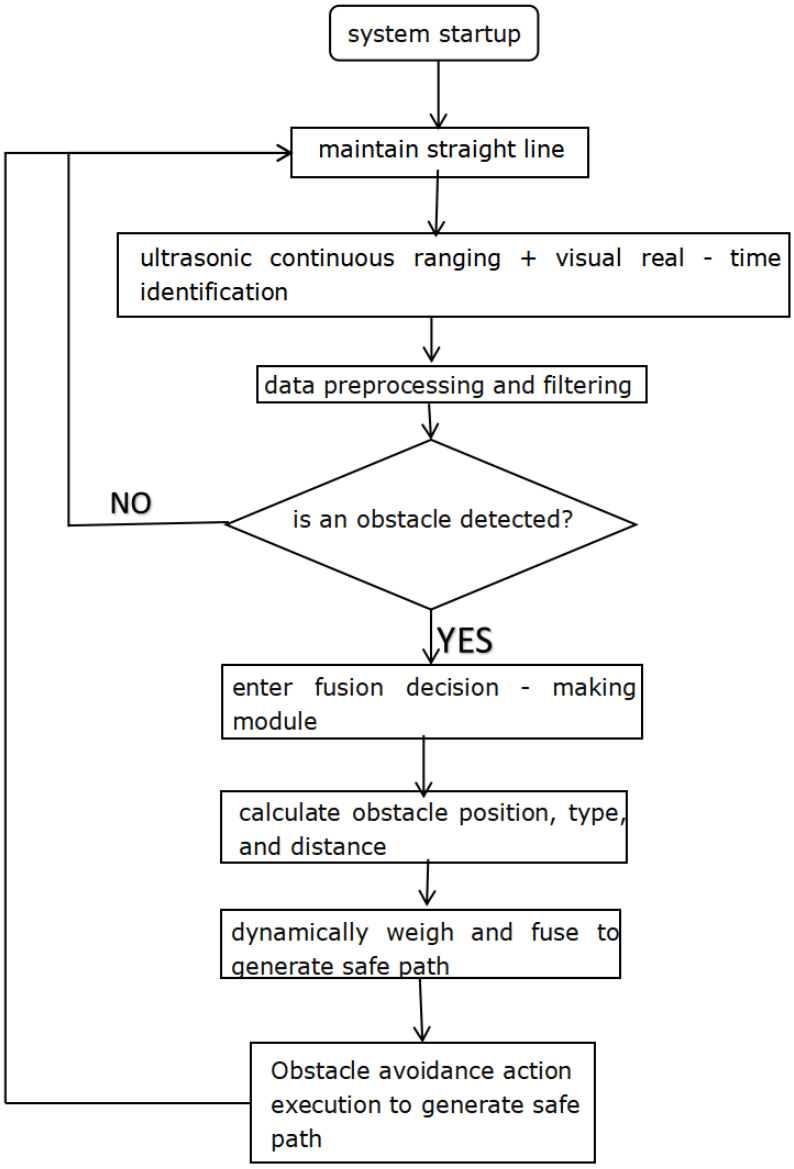

Figure 6. Obstacle Avoidance Principle Flow Diagram.

4.2. Ultrasonic Ranging

The STM32 triggers a measurement by sending a high-level pulse greater than 10μs to the Trig pin of the ultrasonic module; the module then emits ultrasonic waves.

When the ultrasonic waves encounter an obstacle and reflect, the Echo pin of the module outputs a high-level signal, with the duration corresponding to the round-trip time of the ultrasonic waves. The STM32 utilizes the timer's input capture function to record the count values of the Echo rising edge (ccr_2) and falling edge (ccr_1), calculates the difference between the two, and then converts Δccr into the actual time Δt based on the timer resolution. Finally, the obstacle distance is calculated using the formula:

$$Distance = \Delta t \times Vsound/2 = 2(ccr_1 - ccr_2) \times resolution \times Vsound/2 \quad (1)$$

4.3. Visual Recognition Algorithm Selection

This study primarily focuses on recognizing common obstacles in greenhouses and calculating their coordinates. In the field of target detection, common algorithms include Faster R-CNN, SSD, and YOLO series, among others. By comparing these algorithms (Table 2) and considering the real-time and accuracy requirements for obstacle detection in agricultural greenhouse environments, the YOLOv5 algorithm was ultimately selected. YOLOv5 offers fast detection speeds, high precision, and excellent performance in detecting small targets, with high deployment efficiency, meeting the needs for rapid and accurate identification of various obstacles in complex agricultural greenhouse scenarios.

Table 2. Algorithm Comparison.

Algorithm	Ecological Maturity	Inference Speed (FPS)	Deployment Support
YOLOv4	Medium	50+ (V100)	Requires manual optimization
YOLOv5	Extremely High	140+ (T4GPU)	One-click export to ONNX/TensorRT/CoreML
YOLOv8	High	213 (+52.1%)	Partial frameworks require adaptation

In the specific implementation, this study trained and pruned-optimized the YOLOv5s (a smaller version of YOLOv5 that maintains high precision while offering faster inference speeds) on a GPU-equipped PC using the VSCode integrated development environment and PyTorch deep learning framework.

To better adapt the model to the agricultural greenhouse environment, an existing greenhouse obstacle dataset was utilized for iterative training [9]. This dataset includes photos of various common obstacles in agricultural greenhouses, such as plants, farmers, agricultural equipment, and tools left in low-lying areas, covering different states, angles, lighting conditions, and background environments of the obstacles, with precise annotations of the obstacles in the images. After training, the model can recognize various common obstacles in greenhouses. The optimized model was deployed on the K230 processor, and practical tests confirm its ability to recognize multiple greenhouse obstacles in real-time, outputting detailed information such as target category, bounding box center coordinates, and width (Figure 7). Dataset and Training Configuration the model was trained on a dataset comprising 33,400 images covering 8 common greenhouse obstacle categories: plants (10,424 images), personnel (3,020 images), agricultural equipment (5,002 images), tools (4,300 images), low-growing seedlings (3,038 images), irrigation pipes (2,292 images), structural components (3,831 images), and shovels

(1,493 images). Data collection spanned diverse lighting conditions including midday glare (40%), backlighting (25%), and shaded areas (35%), with obstacle orientations randomly rotated from 0° to 360° to simulate multi-angle robot observations. The training process completed 300 iterations (about 12 hours on NVIDIA RTX 3090) using AdamW optimization (initial learning $rate = 1e^{-4}$ with cosine decay) and $batch\ size = 16$, with early stopping activated after 20 epochs without validation, improvement (triggered at epoch 287). Data augmentation strategies included random horizontal flipping (50% probability), brightness adjustment (±30% intensity), Gaussian blur ($\sigma = 0.5 - 1.5$), and synthetic occlusion (10%–30% area blocking) to enhance model robustness against real-world variations. The held-out validation set achieved mAP@0.5:0.95 of 92.3% (14.7% improvement over baseline), with category-wise AP ranging from 88.7% (low-growing seedlings) to 95.1% (personnel), directly supporting the obstacle avoidance system's real-time capability (32 FPS inference) and safety requirements (<0.7% miss rate) in dynamic greenhouse environments.

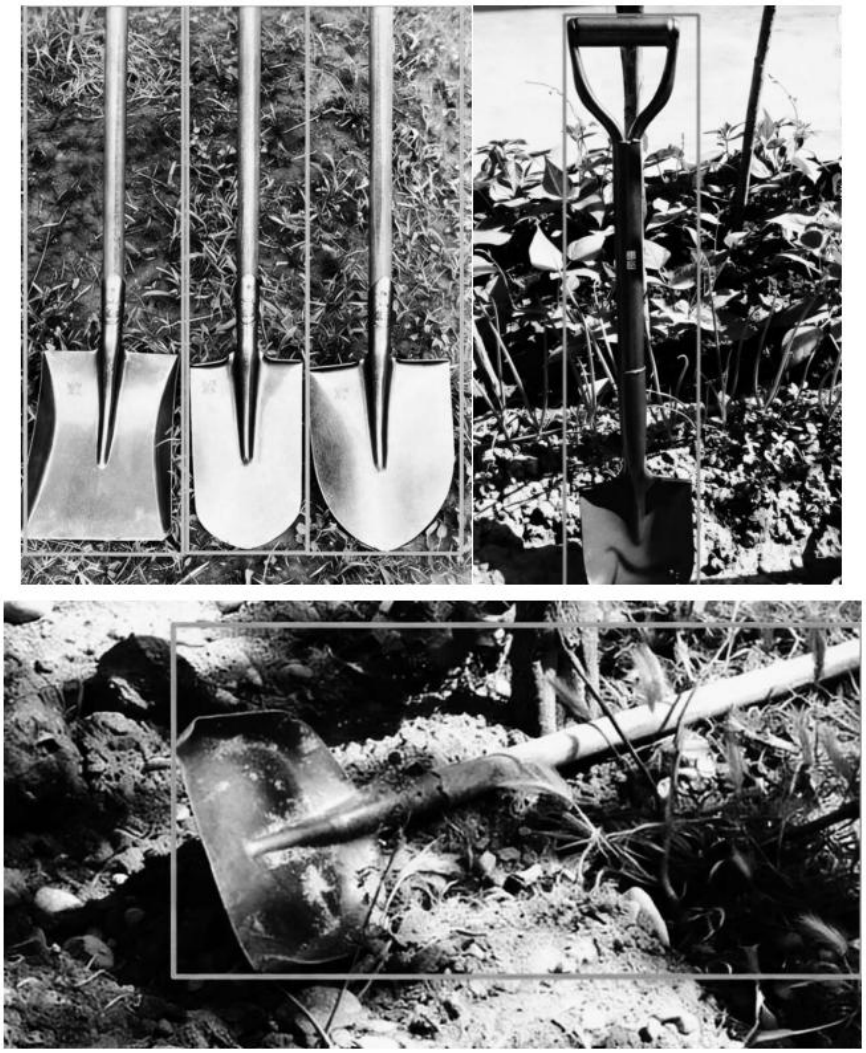

Figure 7. Visual Training Diagram.

4.4. Hierarchical Dynamic Weight Fusion Decision-Making Algorithm

The STM32 achieves precise motion control by fusing visual data from the K230 with ultrasonic ranging data, significantly enhancing the robot's accuracy in complex settings [10]. The K230 sends target information (category, coordinates, confidence) to the STM32 via UART. Ultrasonic distance data is received independently on dedicated pins. The fusion algorithm employs a dynamic weighting strategy: visual data is weighted more heavily for long-range obstacles to enable early warning and path pre-planning, while ultrasonic data is prioritized for close-range obstacles to ensure precise avoidance. This decision logic is summarized in Figure 8.

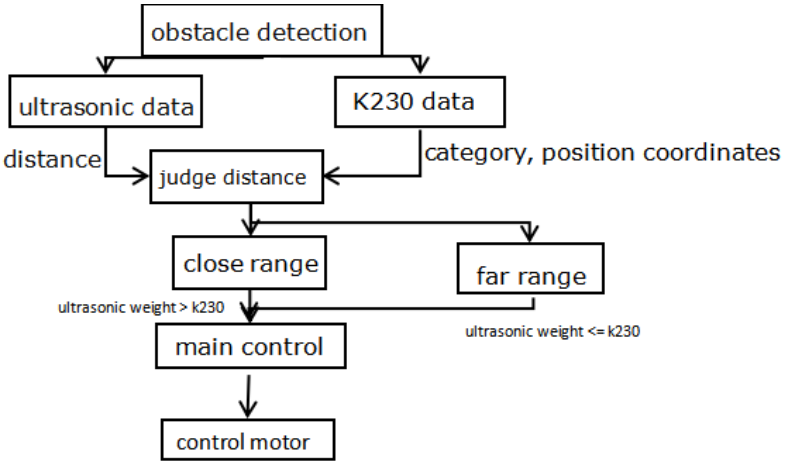

Figure 8. Obstacle Avoidance Principle Diagram.

4.5. System UI Design

The UI design primarily facilitates monitoring and parameter adjustment of the agricultural robot's operations within the greenhouse by the control center, as well as understanding the robot's specific working status. The upper computer design includes modules such as sensor data, visual recognition system, robot control system, data analysis, alarm management, and system settings. The UI interface is shown in Figure 9, with partial program code as follows:

```
function updateSensorData()
{
   console.log('Updating sensor data');
}
setInterval(updateSensorData, 5000);
// Initialization after page load
window.addEventListener('DOMContentLoaded',function()
{
   console.log('Agricultural Robot Monitoring System Started');
});
```

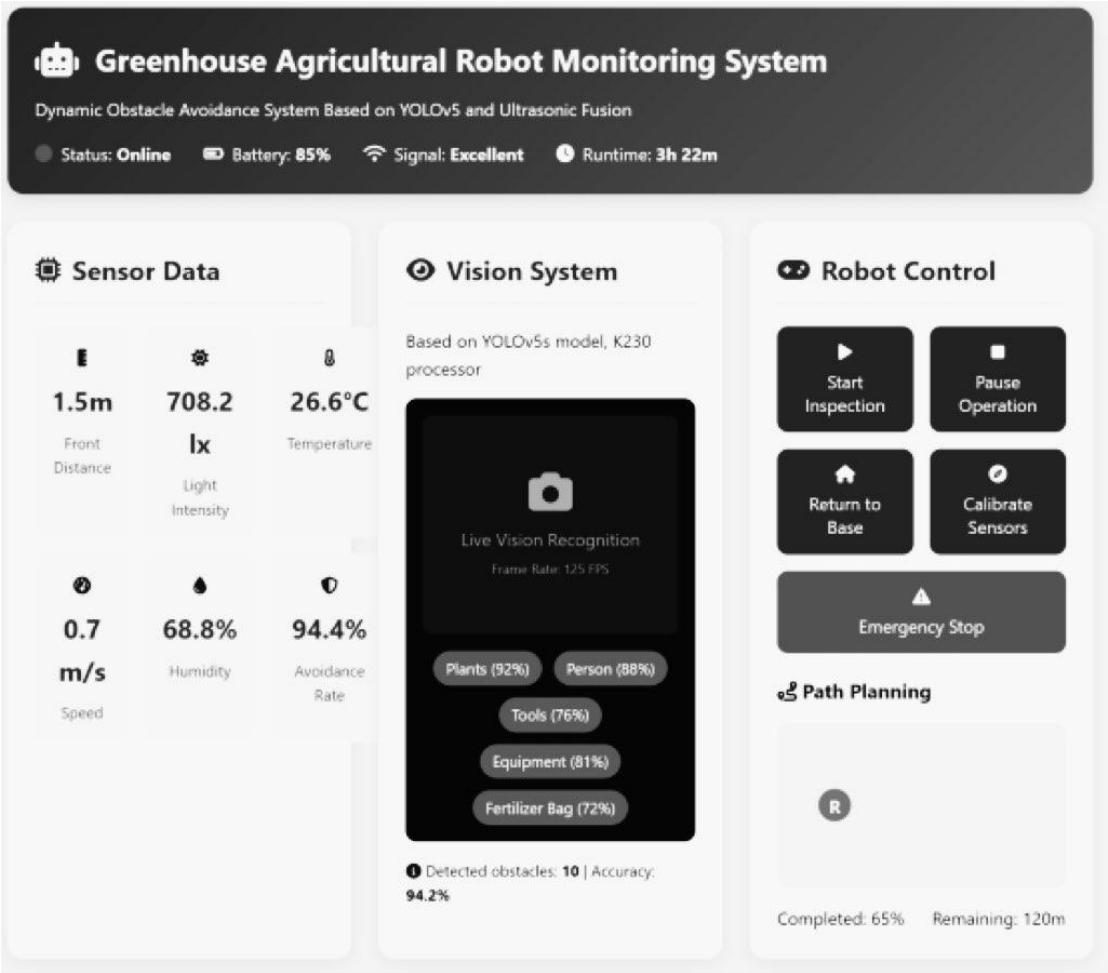

Figure 9. UI Interface.

5. Experiment and Results

To validate the performance of the proposed fusion perception system in complex greenhouse environments, a mobile robot platform was deployed in a real greenhouse for comparative experiments. The experiments were divided into two categories to test the adaptability of different perception schemes under dynamic obstacles and varying lighting conditions.

5.1. First Category of Experiments: Performance Comparison of Different Perception Schemes

Robot speed: 0.5 m/s (constant).

Test environment: Real greenhouse (containing typical obstacles such as fixed metal frames, suspended ropes, and moving personnel).

Evaluation metrics: Obstacle avoidance success rate (successful avoidance attempts / total avoidance attempts × 100%). Average response delay (time difference from obstacle detection to avoidance initiation). Trajectory smoothness index (quantified by acceleration standard

deviation; lower values indicate smoother trajectories)

The following three schemes were tested:

Group A: The proposed visual-ultrasonic fusion perception scheme

Group B: Pure visual perception scheme

Group C: Pure ultrasonic perception scheme

The performance comparison in the greenhouse is shown with specific metrics as follows:

Table 3. Test Results.

scheme	obstacle avoidance success rate	average response delay (ms)	trajectory smoothness index
A	94%	305	1.85
B	89%	420	1.18
C	82%	380	2.03

Obstacle Avoidance Success Rate: The fusion scheme improved the success rate by 8% compared to the pure visual scheme and by 15% compared to the pure ultrasonic scheme. This was primarily due to the complementary advantages of visual detection for low-reflectivity obstacles (e.g., ropes) and ultrasonic penetration for transparent objects (e.g., films).

Response Delay: The fusion scheme reduced response delay by 115 ms

compared to the pure visual scheme and by 75 ms compared to the pure ultrasonic scheme, thanks to parallel processing of multi-sensor data.

Trajectory Smoothness: The fusion scheme reduced trajectory fluctuations by 42% compared to the pure ultrasonic scheme, achieved through Kalman filtering of visual and ultrasonic data.

5.2. Second Category of Experiments: Adaptability Testing Under Different Lighting Conditions

Two lighting scenarios were established:

Scenario 1: Nighttime artificial lighting (LED source, illuminance 400 ± 50 lux)

Scenario 2: Daytime natural light (with shaded areas, illuminance 1200 ± 300 lux)

5.2.1. Experimental Conditions

Robot speed: 0.5 m/s (constant)

Test schemes: Group A (fusion scheme), Group B (pure visual scheme), Group C (pure ultrasonic scheme)

Evaluation metrics: Same as the first category of experiments

Table 4. Test Results.

scenario	scheme	obstacle avoidance success rate	average response delay (ms)	trajectory smoothness index
nighttime	A	92%	320	1.25
nighttime	B	78%	510	2.10
nighttime	C	80%	390	2.08
daytime	A	94%	290	1.12
daytime	B	92%	380	1.70
daytime	C	84%	370	1.95

5.2.2. Results Analysis

When illuminance decreased by 67% at night, the fusion scheme maintained an obstacle avoidance success rate with only a 2% drop, while the pure visual scheme experienced a 13% decline. This demonstrated the significant compensatory effect of ultrasonic sensors under low-light conditions.

Overall Adaptability: The fusion scheme exhibited a trajectory smoothness index standard deviation of 0.07 across both lighting conditions, showing greater stability compared to the pure visual scheme (0.20) and pure ultrasonic scheme (0.09).

5.2.3. Experimental Conclusions

The fusion perception scheme achieved a 94% obstacle avoidance success rate in complex greenhouse environments, improving performance by 5%–12% compared to single-sensor schemes through the complementary advantages of visual and ultrasonic sensors. The system demonstrated significantly enhanced adaptability to lighting variations, with nighttime performance degradation controlled within 2%, enabling 24-hour operation. Multisensory fusion improved trajectory smoothness by over 40%, effectively reducing robot vibrations that could impact crops.

5.3. Results

This paper presents the design of a dynamic obstacle avoidance system for greenhouse robots based on YOLOv5 and ultrasonic fusion. By utilizing the STM32F407 main control chip and K230 visual recognition, visual identification is successfully and deeply fused with traditional ultrasonic ranging on an embedded platform, effectively addressing the challenges posed by typical greenhouse obstacles. Experimental results demonstrate that the fused system significantly outperforms single-sensor solutions in terms of obstacle avoidance success rate and trajectory smoothness, meeting the real-time requirements for low-speed operations of

greenhouse robots and validating its practical value as a low-cost, high-efficiency solution.

Acknowledgements

This work was supported by the Chengdu Science and Technology Bureau Technological Innovation Project (No. 2019-YF05-00203-SN).

References

[1] Kumar, S., Mohan, S., & Skitova, V. (2023). Designing and implementing a versatile agricultural robot: A vehicle manipulator system for efficient multitasking in farming operations. *Machines*, 11(8), 776.

[2] Belforte, G., Deboli, R., Gay, P., Piccarolo, P., & Aimonino, D. R. (2006). Robot design and testing for greenhouse applications. *Biosystems Engineering*, 95(3), 309-321.

[3] Miao, C., Chen, G., Yan, C., & Wu, Y. (2021). Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm. *Computers & Industrial Engineering*, 156, 107230.

[4] Luo, Q., Wang, H., Zheng, Y., & He, J. (2020). Research on path planning of mobile robot based on improved ant colony algorithm. *Neural Computing and Applications*, 32(6), 1555-1566.

[5] Wei, B., Liu, J., Li, A., Cao, H., Wang, C., Shen, C., & Tang, J. (2024). Remote distance binocular vision ranging method based on improved YOLOv5. *IEEE Sensors Journal*, 24(7), 11328-11341.

[6] Yihai, Z., Kai, Z., Xianyi, K., Jiawen, P., Yipeng, Z., & Yujun, H. (2023). Design And Research Of Underwater Pipeline Cleaning Robot [J]. *Mechanical Management Development*, 38(2), 110-112.

[7] Zhou, B. F., & Zhang, J. L. (2020). Design of DC Motor PID Control System Based on STM32 Single Chip Microcomputer. *International Core Journal of Engineering*, 6(7), 62-67.

[8] Jin, Y., Li, S., Li, J., Sun, H., & Wu, Y. (2018, July). Design of an intelligent active obstacle avoidance car based on rotating ultrasonic sensors. *In 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)* (pp. 753-757). IEEE.

[9] Zhu, X. (2021, August). Design of barcode recognition system based on YOLOV5. *In Journal of physics: conference series* (Vol. 1995, No. 1, p. 012052). IOP Publishing.

[10] Wang, Z., Wu, Y., & Niu, Q. (2019). Multi-sensor fusion in automated driving: A survey. *Ieee Access*, 8, 2847-2868.