International Core Journal of Engineering Volume 12 Issue 2, 2026
ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

Machine Vision-Based Laser-Powered Unmanned Aerial Vehicle System

Boyu Shan, Zhigang Di*, Shuolin Xu, Zengshuai Song, Shengbo Zhang,
Yuhan Chen, Zengqi Zhang

North China University of Technology, Tangshan, Hebei, 063000, China

*Corresponding author: Zhigang Di (Email: dzg0512@126.com)

Abstract

To address the issues of short flight endurance and lengthy charging times for drones, a laser-powered drone system based on machine vision has been designed.This system centers on machine vision technology, integrating laser wireless power transmission. It employs the YOLOv11 object detection algorithm enhanced with an attention mechanism to improve detection accuracy for photovoltaic cells on drones, while utilizing PID control to optimize laser targeting precision.Experimental testing indicates that the system achieves over 95% confidence in detecting photovoltaic cell targets, with an inference time of 10.9 milliseconds per frame.This system features high recognition accuracy, rapid identification speed, and precise tracking, making it a portable and efficient laser-powered unmanned aerial vehicle system.

Keywords

Machine Vision; Attention Mechanism; YOLOv11 Algorithm; PID Control.

1. Introduction

In recent years, the cross-disciplinary integration of artificial intelligence and aviation technology has driven rapid advancements in the development and application of unmanned aerial vehicles (UAVs), expanding their reach across industrial, military, and civilian sectors[1]. Small UAVs predominantly utilize electric propulsion systems. However, the energy density of batteries powering electric UAVs remains significantly lower than that of fuel-based systems, resulting in limited flight range and endurance. This severely constrains the payload capacity of electric UAVs[2] and significantly impedes operational efficiency. Consequently, traditional power supply methods have become a limiting factor for the high-quality development of the UAV industry.Currently, laser wireless power transmission technology represents one of the most effective solutions to address the endurance challenges of unmanned aerial vehicles (UAVs)[3]. The introduction and application of laser wireless power transfer technology can significantly enhance the endurance of UAVs[4] and reduce charging time. This technology utilizes laser beams as the energy carrier, with photovoltaic receivers capturing laser energy for photoelectric conversion, thereby enabling long-distance wireless power transmission. Compared to UAVs powered by traditional methods, laser wireless power transfer technology increases UAV endurance by 24 times[5].The laser power supply system proposed in this paper integrates machine vision with laser wireless power transfer technology. Its functional design is as follows: (1) Target Detection: Computer vision is employed to process target detection of photovoltaic cells on drones. (2) Laser Tracking: A high-precision stepper motor gimbal paired with a fuzzy adaptive PID motion control algorithm ensures laser aiming accuracy. The system enables precise identification of the drone's photovoltaic panel target. Combined with the high-precision stepper motor gimbal, the laser achieves accurate tracking of the target.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

2. YOLOv11 Object Detection Mode

The object detection model adopted in this paper is YOLOv11. By integrating the Convolutional Block Attention Module (CBAM) into the YOLOv11 network architecture, the model enhances its ability to extract image features[6]. This facilitates ground cameras in capturing photovoltaic cell targets from drones, enabling precise identification and tracking of photovoltaic cells. The model used in this paper offers strong compatibility and can be directly deployed on RDK X5. The YOLOv11 model architecture comprises three components: a backbone network, a neck architecture, and a head network, collectively enabling efficient and accurate object detection. The core of YOLOv11's backbone network is the C3k2 module, which optimizes information flow within the network, resulting in faster image processing and lower computational costs. The neck architecture incorporates components such as the C3k2 module, SPPF module, and C2PSA mechanism. The SPPF module pools features from different image regions at varying scales, enhancing the network's ability to detect objects of varying sizes, particularly small ones [7]. The head network utilizes feature maps from the neck to perform object classification and bounding box regression. The YOLOv11 architecture is illustrated in Fig. 1.

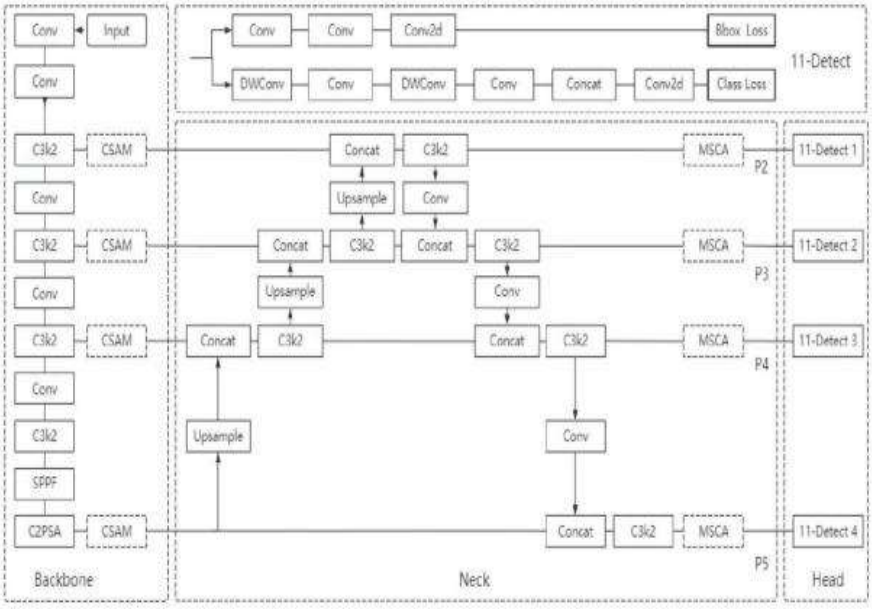

Fig. 1 YOLOv11 Architecture Diagram

3. Hardware Design

The system's hardware comprises the CH32V307VCT6 chip, RDK X5, camera module, and 2D stepper motor gimbal, with its block diagram shown in Fig. 2.

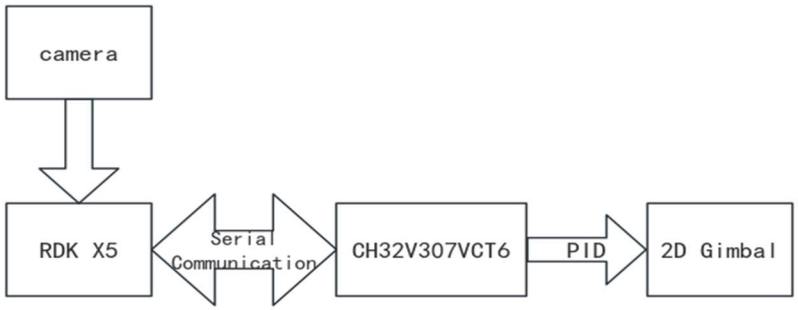

Fig. 2 Hardware Architecture Diagram

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

CH32V307VCT6 Chip: A high-performance RISC-V architecture MCU with a maximum operating frequency of 144MHz. with 256kB FLASH and 64kB SRAM on-chip. It integrates USB 2.0 Full-Speed, Ethernet MAC, and interfaces like SPI, I2C, and UART. eliminating the need for external memory chips. Ideal for industrial control, smart hardware, and IoT terminal applications requiring multi-interface coordination and large-data-volume processing.

RDK X5： The RDK X5 is a high-performance AI development core board based on Horizon's Journey 5 chip. It integrates a BPU deep learning processor with up to 10TOPS computing power, efficiently handling edge AI inference tasks such as image recognition and object detection. Equipped with rich peripheral interfaces including Gigabit Ethernet, HDMI, USB 3.0, and MIPI-CSI, it pairs with high-speed storage and memory to ensure efficient data read/write and multi-task parallel processing, significantly reducing hardware extension design complexity.

Lens: Features a 2-megapixel industrial CMOS lens with a USB 2.0 interface, offering high light sensitivity and low power consumption. It delivers excellent low-light performance, enabling stable and clear image capture under varying lighting conditions.

2D Stepper Motor Gimbal: Equipped with a 42 closed-loop stepper motor, this gimbal features an industrial-grade high-precision magnetic encoder and dual FOC closed-loop algorithms. It offers both pulse and serial interfaces, ensuring smooth, stable motor operation with rapid response. It enables precise adjustment of the laser's azimuth and elevation angles, providing stable motion control support for laser tracking and aiming.

Table 1. Three Scheme comparing

Numble	Scheme 1	Scheme 2	Scheme 3
1	456	456	123
2	789	213	644
3	213	654	649

4. Software Design

4.1 Experimental Environment and Parameter Settings

The parameters for the experimental environment configuration are shown in Table 2.

Table 2. Experimental Environment Setup

Parameter Settings	Configuration
CPU	Intel(R) core(TM)i9-12900HX，2300 Mhz
GPU	NVIDIA GeForce RTX 4060
Accelerated Environment	Cuda 11.8
Environment Setup	Torch2.0.0 ， Torchvision 0.15.1
Operating System	Ubuntu22.04.1
Deep learning framework	PyTorch
Programming language	Python3.10

4.2 Dataset Construction

The experimental dataset was constructed using a combination of indoor/outdoor field collection and publicly available datasets. The field collection component employed a 2-megapixel USB industrial camera to capture 200 images of drones and their onboard photovoltaic cells under various

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

conditions-including sunny, cloudy, and dusk scenarios-and at distances of 0.5m, 1m, and 2m. This collection encompassed conventional drone types such as multirotors and fixed-wing aircraft. Public datasets supplemented the collection with the VisDrone2021 drone detection dataset and a self-built photovoltaic cell dataset, totaling 400 images. Each image was labeled with "drone" and "photovoltaic cell" tags. Data augmentation techniques-including random scaling, random cropping, random arrangement [8], and noise addition-expanded the dataset to 1,000 images. The dataset was then divided into training, validation, and test sets at an 8:1:1 ratio.

4.3 Model Training

The training set provides a rich collection of data samples, enabling the model to learn more effective features during training. The validation set monitors training effectiveness and adjusts the model's hyperparameters based on its performance to optimize model capabilities. The test set is used to comprehensively evaluate the fully trained and optimized model, measuring its generalization ability and practical application effectiveness. During training, both stages employ stochastic gradient descent with momentum optimization, using a learning rate of 0.0001–0.01 and a momentum coefficient of 0.9. When detecting model training stagnation, the batch size is set to 16, with a total of 80 epochs trained.

4.4 Programming Process

The overall programming workflow involves initializing hardware first, followed by a laser power control process that forms a closed-loop feedback system encompassing data processing and workflow tasks such as model loading, image acquisition, target recognition, and command execution. The RDK X5 program flow is illustrated in Fig. 3.

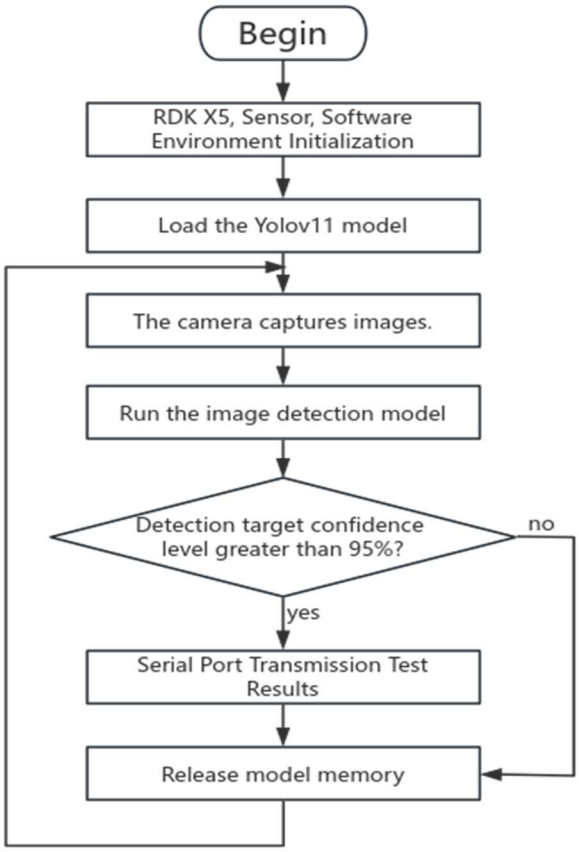

Fig. 3 RDK X5 Program Flowchart

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

5. Experiments and Analysis of Results

5.1 Experimental Environment and Parameter Settings

The data used for model training and evaluation, as shown in Fig. 4, depicts the variation of different losses and several evaluation metrics across the training set (train) and validation set (val).

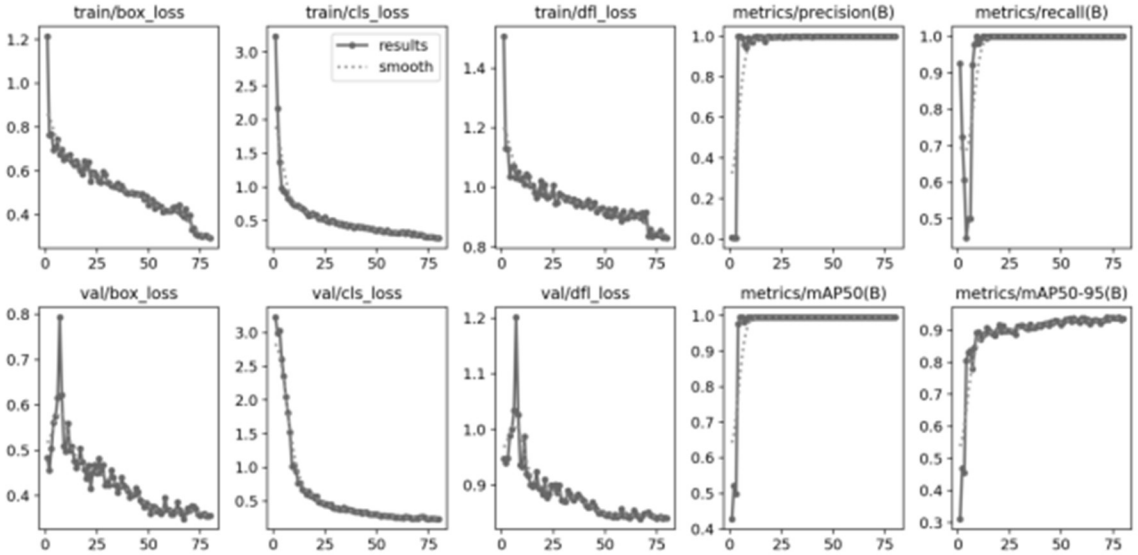

Fig. 4 Model Training Evaluation Metrics

As shown in Fig. 4, during the model training and validation phases, the loss on the training set gradually decreases while metrics such as precision and recall steadily increase. Concurrently, the mean average precision (mAP) on the validation set also continues to rise. This indicates that the model is progressively learning and improving, demonstrating strong performance and potential.

Table 3 and Table 4 present the specific values for each evaluation metric at 80 epochs.

Table 3. Training Set Model Training Evaluation Metrics (80 epochs)

Evaluation Indicators	Result value
Training Set Bounding Box Loss	0.293
Training Set Feature Point Loss	0.829
Training Set Classification Loss	0.244
Accuracy	0.999
Recall rate	1.000

Table 4. Model Training Evaluation Metrics on the Validation Set (80 epochs)

Evaluation Indicators	Result value
Validation Set Batch Loss	0.355
Validation Set Feature Point Loss	0.841
Validation Set Classification Loss	0.236
mAP@0.5	0.995
MAP@0.5:0.95	0.935

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

As shown in Tables 3 and 4, the model achieves favorable metrics on both the training and validation sets, with relatively consistent performance across metrics. This indicates that the model did not exhibit significant overfitting during training and demonstrates good generalization capability. However, further improvement is needed at higher IoU thresholds.

The experiment compares the YOLOv11 model incorporating a convolutional attention mechanism with the original YOLOv11 model. Both models were trained and validated using the same dataset. Table 5 presents the comparison results for the two models in terms of mAP, Recall, and inference time.

Table 5. Experimental Comparison Results

Model	Weight/MB	Average Detection Accuracy/%	Recall Rate/%	Inference time per frame/ms
YOLOv11 (with CBAM)	6.0	96.4	93.9	10.9
YOLOv11	4.5	90.8	89.4	7.5

As shown in Table 5, the YOLOv11 model incorporating the convolutional attention mechanism adds only 1.5MB to the weight size compared to the original YOLOv11 model. Although its inference time is 3.4ms slower than the original YOLOv11 model, its average detection accuracy and recall rate improve by 5.6 and 4.5 percentage points, respectively.

Experimental results demonstrate that the YOLOv11 model incorporating a convolutional attention mechanism exhibits superior performance compared to the original YOLOv11 model. It achieves better detection results for small objects and demonstrates enhanced robustness[9].

5.2 Experimental Environment and Parameter Settings

To evaluate the performance of YOLOv11 in detecting photovoltaic cell targets on drones of different categories, this paper employs precision (P), recall (R), average precision (AP), and mean average precision (mAP) as key metrics for assessing object detection models. Precision approaches 1 as accuracy increases; recall approaches 1 as accuracy increases; mAP approaches 1 as accuracy increases. The formulas for calculating P, R, AP, and mAP are as follows[10]:

$$P = \frac{TP}{TP+F} \quad (1)$$

$$R = \frac{TP}{TP+F} \quad (2)$$

$$AP = \int_0^1 P(R)dR. \quad (3)$$

$$mAP = \frac{AP_{mask}+AP_{nomask}}{2}. \quad (4)$$

Among these, TP denotes the number of positive samples predicted as positive by the model; FP denotes the number of negative samples predicted as positive by the model; FN denotes the number of positive samples predicted as negative by the model.

Experimental results indicate that when the drone detection model achieves 96.4% precision and 93.9% recall, mAP@0.5 reach 93.1%; for the photovoltaic cell target model with 95.1% precision and 92.7%

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

recall, mAP@0.5 attain 92.2%. Even when confronted with varying degrees of object occlusion and environmental interference, the model maintained detection accuracy above 95%, demonstrating strong robustness.

5.3 Testing Effect

The detection performance of drones and airborne photovoltaic panels under different environmental conditions is shown in Fig. 5.

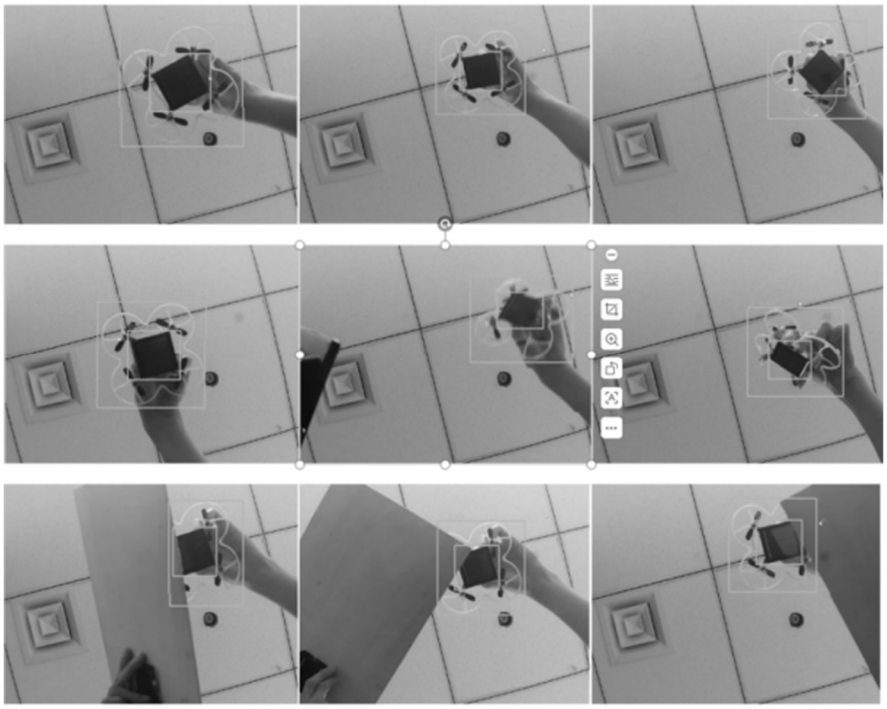

Fig. 5 Rendering of Drone and Airborne Solar Panel Inspection

As demonstrated by the detection results in various scenarios shown in Fig. 5, this system accurately identifies drones and their onboard photovoltaic cells under complex conditions such as indoor/outdoor lighting variations and partial target occlusion. Detection bounding boxes closely align with actual target areas, exhibiting no significant missed detections, false positives, or boundary shifts. Target detection confidence remains consistently above 92% across all scenarios, reaching over 97% in standard environments. This detection performance provides reliable target coordinate support for subsequent laser tracking, ensuring stable tracking based on precise target positioning even in complex outdoor operational conditions. This lays the foundation for the efficiency and continuity of wireless laser energy transfer.

5.4 System Integration

System integration serves as the core component enabling the closed-loop operation of the entire process chain-"object detection-command transmission-laser tracking." Through deep adaptation and functional coordination of hardware and software modules, the improved YOLOv11 object detection model, core control chip, image acquisition devices, and actuators are integrated into a unified operational system. This ensures efficient interconnection among all components, meeting the real-time and precision requirements for laser power delivery. The specific integration process is as follows: First, the YOLOv11 model is redesigned and trained by incorporating the Convolutional Attention Mechanism (CBAM) module and a small object detection layer to generate target detection weight files suitable for drones and photovoltaic cells. Second, the RDK X5 model deployment tool is invoked to convert the YOLOv11 model into an inference model executable on the development board, which is then saved to the board's storage unit. Third, establish physical connection channels

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

for camera image acquisition, data processing, and control execution by wiring the CH32V307VCT6 chip, RDK X5, 2-megapixel USB industrial camera, and 2D stepper motor gimbal. Develop a Linux-based program to implement all functionalities: model loading, camera image acquisition and preprocessing, target recognition, and laser tracking control command output.

During model inference, the target inference module feeds preprocessed image data captured by the camera into the YOLOv11 model installed on the RDK X5. Leveraging the BPU's powerful computational capabilities, it performs object detection to extract class labels and coordinate information for drones and photovoltaic cells. The results are then output via serial port to the CH32V307VCT6 chip. This chip invokes the fuzzy adaptive PID algorithm, integrating feedback from the two-dimensional stepper motor gimbal's encoder-specifically the azimuth and pitch angles-as output parameters for motor drive commands. This controls the stepper motor gimbal's movement, adjusting the laser's direction to enable laser tracking and energy transmission to the photovoltaic panels.

6. Conclusion

This paper investigates and implements a laser-powered unmanned aerial vehicle (UAV) system. Utilizing the YOLOv11 object detection algorithm and fuzzy adaptive PID control technology, it achieves target detection of photovoltaic panels on the UAV and laser tracking for power supply, providing a practical solution to the UAV endurance challenge. Experimental testing demonstrated that the system achieves target detection confidence exceeding 95% for both the UAV and its onboard photovoltaic panels, with laser aiming errors below 0.1°. This enables precise detection of the UAV and photovoltaic panels alongside laser-guided power supply, thereby advancing UAV development. By employing laser wireless power transfer technology, the system reduces lithium battery waste pollution, aligning with the “dual carbon” strategy. It also demonstrates strong adaptability, enabling transferability to other types of laser-based long-range power supply equipment. Future research will focus on further optimizing the model structure to reduce its size while maintaining detection accuracy. Concurrently, exploring multi-sensor fusion tracking solutions will enhance the system's environmental adaptability.

Acknowledgments

Hebei Provincial Department of Education Scientific Research Project (ZD2021332).

References

[1] Kai Huang. Research on the Application of Laser Power Supply Technology in UAV Endurance[D].Three Gorges University,2020.DOI:10.27270/d.cnki.gsxau.2020.000488.(In Chinese)

[2] Husheng Tang.Laser-Powered Unmanned Aerial Vehicle Systems and Their Application Prospects[J].China Science and Technology Information,2021,(12):24-27.(In Chinese)

[3] Xiaoyang C ,Yanji H ,Xing J .Energy management strategy for laser-powered near space aircraft[C]//IETP-Association.Abstracts of 2015 International Conference on Structural,Mechanical and Materials Engineering(ICSMME 2015).State Key Laboratory of Laser Propulsion &Application, Academy of Equipment;,2015:26.

[4] He X ,Zhong Y ,Li H .Distributed Robust Routing Optimization for Laser-Powered UAV ClusterwithTemporaryParkingCharging[J].AppliedSciences,2025,15(17):9676-9676.DOI:10.3390/APP15179676.

[5] Jianhua Yuan,Kai Huang,Husheng Hong,et al.An Optimized Tracking Algorithm for Laser-Powered Unmanned Aerial Vehicles[J].Applied Optics,2020,41(01):194-201.(In Chinese)

[6] Yao Y ,Qiao G ,Zhang S , et al.Research on Morphometric Methods for Larimichthys crocea BasedonYOLOv11-CBAMX-RayImaging[J].Fishes,2025,10(12):641-641.DOI:10.3390/FISHES10120641.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202602_12(2).0002

[7] Song L ,Tao Y.Occluded Small Target Detection for UAVs Based on YOLOv11[J].Engineering Letters,2026,34(1).

[8] Linlin Ma,Jianxin Ma,Jiafang Han,et al.Research on Object Detection Algorithms Based on YOLOv5s[J].ComputerKnowledgeandTechnology,2021,17(23):100-103.DOI:10.14004/j.cnki.ckt.2021.2402.(In Chinese)

[9] Adli T ,Bujaković M D ,Bondžulić P B , et al.Robustness of YOLO models for object detectioninremotesensingimages[J].JournalofElectricalEngineering,2025,76(5):429-442.DOI:10.2478/JEE-2025-0045.

[10] Xiaobo Li ,Yanggui Li ,Ning Guo , et al.YOLOv5 Mask Detection Algorithm with Integrated Attention Mechanism[J].Journal of Graphics,2023,44(01):16-25.(In Chinese)