ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

Research on Real-Time Microalgae Detection Based on Deep Learning and a Portable Dual-Mode Imaging System

Xijun Gao[1,*,#], Dairan Li[1,#], Yutong Li[1,#], Jiahong Yin[2,#]

[1]*School of Railway Intelligent Engineering, Dalian Jiaotong University, Dalian, China*
[2]*School of Materials Science and Engineering, Dalian Jiaotong University, Dalian, China*
**Corresponding author: 15242003178@163.com*
[#]*These authors contributed equally to this work*

Abstract: *As an important component of aquatic ecosystems, the population distribution and activity status of microalgae are key indicators for water quality monitoring and ecological research. However, traditional microalgae detection methods mainly rely on manual microscopic counting or flow cytometry, which suffer from limitations such as expensive equipment, cumbersome operations, and excessive dependence on laboratory environments and professional personnel, making it difficult to meet the needs of real-time on-site monitoring in the field. Addressing the aforementioned issues, this study proposes a low-cost, portable, automated microalgae detection system featuring software-hardware synergy. In terms of hardware, a visible light and fluorescence dual-mode microscopic imaging device based on 3D printing technology was designed, achieving synchronized acquisition of microalgae morphology and activity as well as device miniaturization. In terms of algorithms, an improved GIFF-AlgaeDet deep learning object detection network is proposed. By introducing the CAGS attention mechanism, a lightweight re-parameterized dual detection head, and the SIoU loss function, it effectively solves the problems of missed detection and false detection caused by the minute size of microalgae targets and complex backgrounds, while significantly reducing model parameters. Experimental results show that the average detection precision of the system on a self-built microalgae dataset reaches 96.2%, with a single-frame inference time of only 0.02 seconds. It is significantly superior to current mainstream detection algorithms and possesses the advantages of high precision, low cost, and portability, providing an efficient and intelligent solution for water environment ecological monitoring and aquaculture management.*

Keywords: *Microalgae detection, Deep learning, Portable microscopic imaging, Attention mechanism, Lightweight network*

1. Introduction

Microalgae play a crucial role in the biosphere's carbon cycle, aquaculture feed supply, and bioenergy development[1]. However, their explosive proliferation can also trigger ecological disasters such as red tides or algal blooms. Therefore, realizing rapid and accurate detection of microalgae species and biomass has important scientific and application value[2], [3]. For a long time, microalgae detection has mainly followed traditional laboratory analysis workflows, such as manual morphological identification under optical microscopes, PCR gene sequencing technology, and flow cytometry. Although these methods can provide relatively detailed biological information, their drawbacks are obvious: manual microscopy is extremely time-consuming and highly dependent on the subjective experience of taxonomic experts, making it difficult to process large-scale samples; while molecular biology and flow cytometry technologies offer higher precision, the equipment is bulky, acquisition and maintenance costs are high, and the sample pretreatment process is complex, making them incapable of performing immediate detection tasks in field environments such as rivers, lakes, and seas. Therefore, developing an automated microalgae detection device that combines portability, economy, and intelligence has become a technical bottleneck urgently needing solution in the current environmental monitoring field.

With the rapid development of computer vision technology[4], [5], object detection algorithms based on deep learning have provided new ideas for microalgae identification. In recent years, detection models represented by convolutional neural networks have achieved significant results in general object recognition tasks, and some researchers have begun to attempt to apply them to plankton image

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

analysis[6], [7]. However, directly migrating general object detection algorithms to microalgae detection scenarios still faces many challenges. First, microalgae individuals are usually extremely small, often occupying less than 5% of the pixel area in the microscopic field of view, and the morphological similarity between different algae is extremely high, easily causing insufficient feature extraction. Second, existing high-performance detection models are often accompanied by huge parameter counts and computational overheads, making them difficult to deploy on portable or embedded terminals with limited computing resources.

To break through the aforementioned technical limitations, this study constructed an end-to-end detection system integrating miniaturized optical hardware and lightweight intelligent algorithms. At the hardware level, we abandoned the heavy body design of traditional microscopes and used modular and 3D printing technology to build a visible light-fluorescence dual-mode imaging system. While ensuring micron-level imaging resolution, this greatly compressed the device volume, enabling it to adapt to complex and changing field operation environments. At the software algorithm level, addressing the difficulties of small target microalgae detection, we designed the GIFF-AlgaeDet network based on Global Information and Feature Fusion. This network enhances the model's spatial perception ability for minute features through an innovative attention mechanism, and utilizes structural re-parameterization technology and optimized loss functions to strengthen feature learning during the training phase and simplify the network structure during the inference phase, thereby achieving real-time inference speed while ensuring high detection precision. This study aims to explore a new path for low-cost, high-efficiency intelligent field detection of microalgae through the deep integration of software and hardware.

2. Methodology

This study aims to solve problems existing in traditional microalgae detection such as bulky equipment, reliance on manual labor, and difficulty in identifying small targets. To this end, we constructed an automated detection system integrating portable optical hardware and lightweight deep learning algorithms. This chapter will provide detailed mathematical elaborations on the system's hardware imaging mechanism, the mathematical definition of the GIFF-AlgaeDet network model, the tensor operation process of the CAGS attention mechanism[8], the inference acceleration principle of the re-parameterized detection head, and the geometric definition of the SIoU loss function[9], [10].

2.1 Optical and Physical Constraints of the Portable Dual-Mode Imaging System

To achieve high-precision microalgae detection in the field, the design of the hardware system follows the optimization goal of minimizing volume and maximizing imaging quality. The system adopts a modular design, mainly including a light source module, a sample module, and an imaging module, and its overall structure is shown in Figure 1. We define the design space of the imaging system as the physical constraint set C_{phys} and the optical constraint set C_{opt}. The physical constraints mainly target the portability of the device. Let the volume of the system be V and the mass be W; the design goal needs to satisfy $V \leq 790\ \text{cm}^3$ and $W \leq 221$ g to adapt to mobile operation requirements.

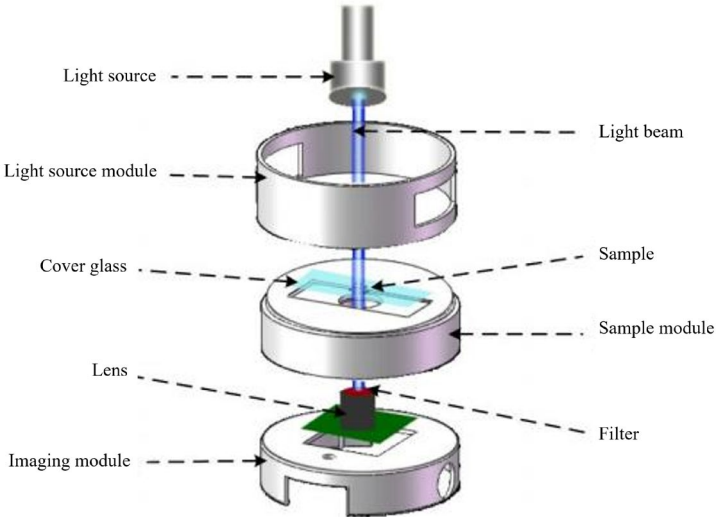

Fig. 1 Schematic diagram of the imaging system structure

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

At the optical imaging level, the system adopts a visible light and fluorescence dual-mode imaging mechanism. Let the spectral response function of the input sample be $S(\lambda)$. In the brightfield mode, the light source module provides broadband white light illumination $I_{bright}(\lambda)$, and the light intensity signal E_{bright} received by the imaging sensor can be expressed as the convolution of the sample transmittance $T(x,y)$ and the light source intensity. In the fluorescence mode, based on the physical characteristics of microalgae chlorophyll, the system selects an excitation wavelength $\lambda_{ex} \approx 460$ nm. At this time, the intensity of the fluorescence signal E_{fluo} depends on the excitation light intensity I_{ex}, the fluorescence quantum yield of the microalgae Φ, and the chlorophyll concentration $C(x,y)$. Its imaging model can be approximated as $E_{fluo}(x,y) \propto I_{ex}(\lambda_{ex}) \cdot \Phi \cdot C(x,y)$. This fluorescence signal is collected by the CMOS sensor after processing by a filter, thereby achieving quantitative characterization of microalgae activity information. By adjusting the distance d between the lens and the sample plane through a precise 3D printed mechanical structure, the system can resolve microalgae targets with size $s \geq 1\mu m$, meeting the microscopic imaging requirements under the Rayleigh criterion.

2.2 Topological Definition of the GIFF-AlgaeDet Network Model

Addressing the characteristics of minute microalgae target sizes and complex backgrounds, this paper proposes a detection model based on Global Information and Feature Fusion (GIFF). As shown in Figure 2, this model can be defined as an end-to-end mapping function $F: \mathbb{R}^{H\times W\times 3} \rightarrow \mathbb{R}^{N\times(4+C)}$, where the input is an image with resolution $H \times W$, and the output comprises boundary box coordinates and class probabilities for N predicted targets.

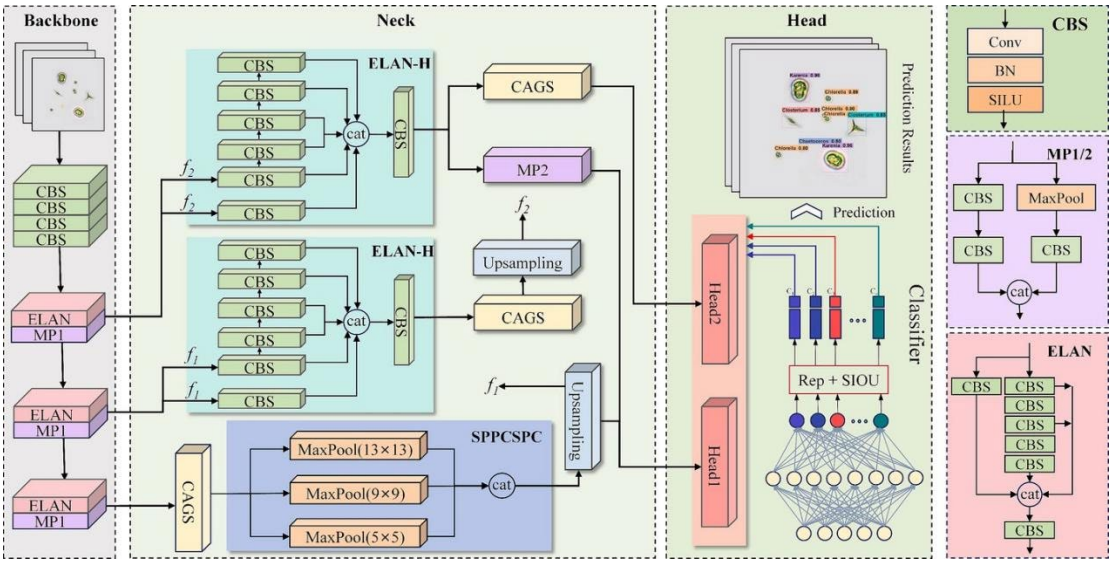

Fig. 2 Main framework of GIFF

The backbone network of the model is mainly responsible for feature extraction. Let the input feature map be X_{in}. The backbone network generates a multi-scale feature pyramid by stacking CBS modules (Convolution-Normalization-SiLU Activation), ELAN modules, and MP1 down-sampling modules. The feature map F_i at level i can be expressed as $F_i = \Phi_{ELAN}(\Phi_{MP}(\dots(X_{in})))$, where Φ represents the non-linear transformation operator of each module. To solve the problem of positional information loss in deep networks, the feature fusion network (Neck) introduces the CAGS module into the FPN structure. Let the input of the Neck layer be the feature set $\{F_3, F_4, F_5\}$ extracted from the backbone network. The fusion process includes the up-sampling operator U and the concatenation operator Cat. The output P_i after feature fusion can be described as $P_i = \text{CAGS}(Cat(F_i, U(P_{i+1})))$. This structural design ensures effective interaction between deep semantic information and shallow texture information, specifically enhancing the model's ability to capture minute targets.

2.3 Feature Enhancement Operation Based on CAGS Attention Mechanism

To overcome the limitations of traditional attention mechanisms in capturing long-range dependencies, this paper proposes the CAGS (Coordinate Attention with Group Shuffle) mechanism. As shown in Figure 3, this module enhances feature expression by embedding positional information into

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

feature maps and introducing channel shuffling. Let the input feature tensor be $X \in \mathbb{R}^{C\times H\times W}$.

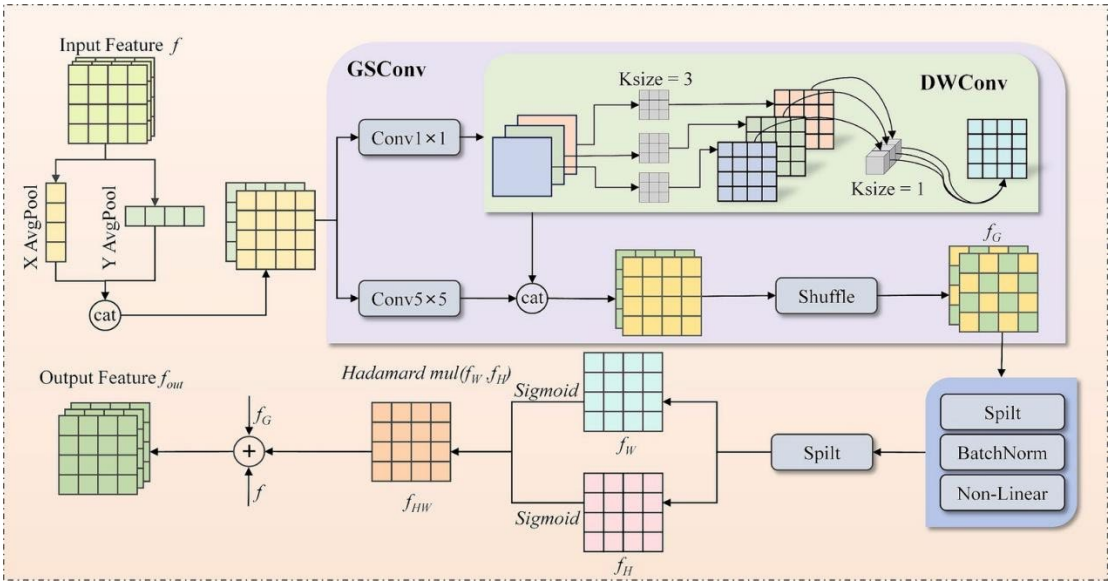

Fig. 3 Network structure of CAGS

First, to capture spatial positional information, CAGS uses global average pooling kernels $(H, 1)$ and $(1, W)$ to encode each channel along the horizontal direction (X-axis) and vertical direction (Y-axis), respectively. For the c-th channel, the output $z_h^c(h)$ at height h and the output $z_w^c(w)$ at width w are calculated as:

$$z_h^c(h) = \frac{1}{W}\sum_{0\le i<W} X_c(h, i) \tag{1}$$

$$z_w^c(w) = \frac{1}{H}\sum_{0\le j<H} X_c(j, w) \tag{2}$$

The above operations generate a pair of direction-aware feature maps. Subsequently, these two feature maps are concatenated and sent to the GSConv module for processing. GSConv utilizes a hybrid structure of depthwise separable convolution (DWConv) and standard convolution to reduce computational load, combined with a Channel Shuffle operation $S(\cdot)$ to promote information flow between channels. Let the transformed intermediate feature be $F_{inter} = S(\text{GSConv}([z_h, z_w]))$.

Finally, the feature map F_{inter} is split again into two independent tensors $f^h \in \mathbb{R}^{C\times H\times 1}$ and $f^w \in \mathbb{R}^{C\times 1\times W}$. The Sigmoid activation function $\sigma(\cdot)$ is used to generate the final attention weight vectors g^h and g^w. The final output Y of the CAGS module can be expressed as the element-wise product of the original features and the attention weights:

$$Y_c(i, j) = X_c(i, j) \times \sigma(f_c^h(i)) \times \sigma(f_c^w(j)) \tag{3}$$

Through the above formulas, the CAGS mechanism effectively encodes spatial coordinate information into weighted feature maps, significantly improving the model's spatial localization sensitivity to microalgae targets.

2.4 Re-parameterized Inference of the Lightweight Detection Head

To improve inference speed while ensuring detection precision, the detection head (Head) part adopts a re-parameterized design based on the RepVGG architecture. As shown in Figure 4, this design abandons the redundant three-head structure, retaining only two detection heads specifically for small and medium targets, and each detection head has different topological structures during training and inference phases.

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

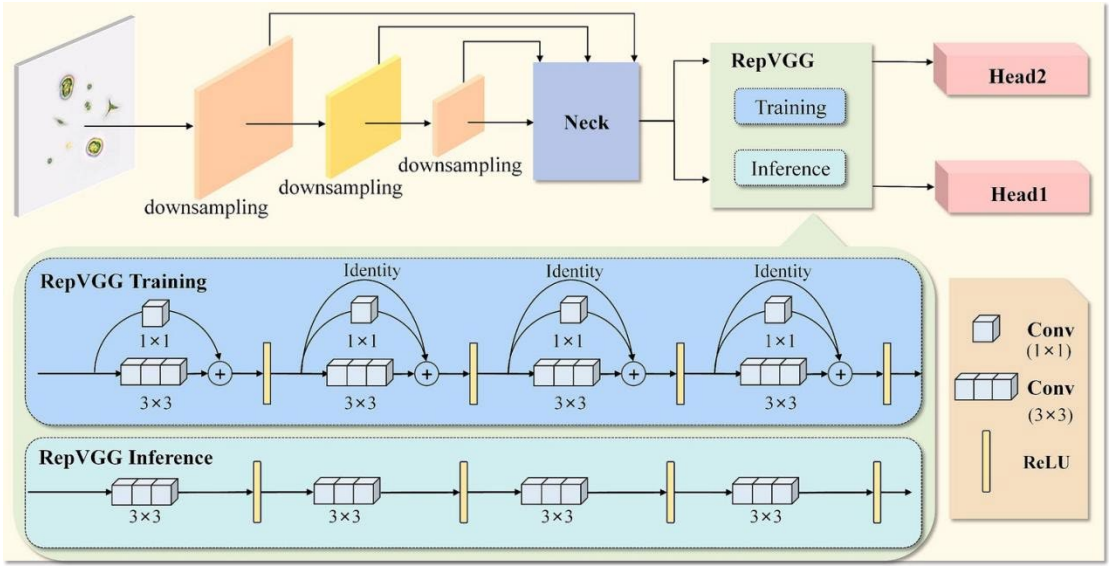

Fig. 4 Structure of the Lightweight Detection Head

In the training phase, the convolutional layer of the detection head consists of three parallel branches: a 3×3 convolution branch, a 1×1 convolution branch, and an identity mapping branch. Let the input feature be M_{in}, the output feature be M_{out}, the weight of the j-th convolution kernel be $W^{(j)}$, the bias be $b^{(j)}$, and the parameters of the Batch Normalization layer be $\{\mu, \sigma, \gamma, \beta\}$. The output during training can be expressed as the linear superposition of the outputs of the three branches:

$$M_{out} = \mathrm{BN}(W_{3\times3} * M_{in}) + \mathrm{BN}(W_{1\times1} * M_{in}) + \mathrm{BN}(M_{in}) \tag{4}$$

In the inference phase, to eliminate the memory access and computational overhead brought by the multi-branch structure, the additivity principle of convolution is used to equivalently transform the above structure into a single 3×3 convolutional layer. First, the BN layer parameters are fused into the convolution kernel. For any convolution branch, the fused weight W' and bias b' calculation formulas are:

$$W' = W \cdot \frac{\gamma}{\sigma}, b' = \beta - \mu \cdot \frac{\gamma}{\sigma} \tag{5}$$

By zero-padding the 1×1 convolution kernel to 3×3, and regarding the identity mapping as a 1×1 convolution with an identity matrix as weight, the final re-parameterized convolution kernel W_{rep} and bias b_{rep} can be expressed as the algebraic sum of the fused parameters of each branch:

$$W_{rep} = W'_{3\times3} + \mathrm{pad}(W'_{1\times1}) + W'_{id} \tag{6}$$

$$b_{rep} = b'_{3\times3} + b'_{1\times1} + b'_{id} \tag{7}$$

After re-parameterization, the model in the inference phase only contains a single 3×3 convolution operation, thereby achieving high-speed detection of 0.0219 seconds/image with a 29% reduction in parameters.

2.5 Definition of Angle-Aware SIoU Loss Function

To further optimize the stability of boundary box regression, especially to solve the problem of direction mismatch between the predicted box and the ground truth box in microalgae detection, this study introduces the SIoU loss function. This loss function consists of a weighted combination of angle cost, distance cost, and shape cost.

First, the angle cost Λ is defined to minimize the angle between the connection line of the predicted box center and the ground truth box center and the coordinate axes. Let the predicted box center be b, the ground truth box center be b^{gt}, the distance between the two in the height direction be c_h, and the Euclidean distance be σ. The angle cost is defined as:

$$\Lambda = 1 - 2\sin^2\left(\arcsin\left(\frac{c_h}{\sigma}\right) - \frac{\pi}{4}\right) \tag{8}$$

This formula prompts the predicted box to quickly regress to the nearest coordinate axis (X-axis or

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

Y-axis). Based on the angle cost, the distance cost Δ is defined. Let ρ_x, ρ_y be the normalized distances of the two box centers in the x and y axis directions respectively, and $\gamma = 2 - \Lambda$ be the angle-adaptive weight, then:

$$\Delta = \sum_{t\in\{x,y\}} (1 - e^{-\gamma\rho_t}) \tag{9}$$

The shape cost Ω is used to constrain the consistency of the aspect ratio. Let w, h and w^{gt}, h^{gt} be the width and height of the predicted box and the ground truth box respectively, then the shape cost is calculated as follows:

$$\Omega = (1 - e^{-\omega_w})^{\theta} + (1 - e^{-\omega_h})^{\theta} \tag{10}$$

Where $\omega_w = \frac{|w-w^{gt}|}{\max(w,w^{gt})}$, $\omega_h = \frac{|h-h^{gt}|}{\max(h,h^{gt})}$. Finally, combined with the traditional IoU metric, the total regression loss function $Loss_{SIoU}$ is defined as:

$$Loss_{SIoU} = 1 - IoU + \frac{\Delta+\Omega}{2} \tag{11}$$

The introduction of the SIoU loss function effectively solves the problem of unclear gradients when the IoU value is the same but the direction is misaligned, making the model converge more rapidly and localize more accurately in scenarios with densely distributed microalgae.

3. Experiments and Results

This chapter will describe in detail the construction process of the microalgae dataset used in the experiments, and comprehensively evaluate the performance of the proposed GIFF-AlgaeDet model through a series of ablation studies and comparative experiments. The experiments focus on verifying the impact of the CAGS attention mechanism, the re-parameterized dual detection head, and the SIoU loss function on the model's detection precision, inference speed, and parameter count.

3.1 Data Description

To train and evaluate the effectiveness of the deep learning model in microalgae detection tasks, this study constructed a high-quality dataset containing various common marine microalgae. The experimental samples were obtained from the Liaoning Institute of Marine Technology, covering six representative algae species: Chlorella, Karenia, Nitzschia closterium, Cryptomonas, Chrysophyta, and Chaetoceros. The raw image data was acquired using an Olympus CKX53 microscope, capturing a total of 1800 high-resolution microalgae micrographs.

Considering the dependence of deep learning models on data scale and the high cost of acquiring microalgae samples, to prevent the model from overfitting during training and to improve its generalization ability, this study employed various data augmentation techniques to expand the original dataset. Specific augmentation strategies included introducing salt-and-pepper noise to simulate sensor noise, random rotation and cropping to simulate different shooting angles and fields of view, and scaling operations to adapt to targets of different sizes. Through the above processing, the total volume of images in the dataset was expanded to 5400, and the number of samples for each category maintained a good balance, ensuring fair representation of each algae species during training. All images were finely annotated manually using the online annotation tool Make Sense, and the annotation files were uniformly formatted into the YOLO standard format. Detailed attributes of the dataset are shown in Table 1, and some constructed microalgae image samples and their augmentation effects are shown in Figure 5.

Table 1 Dataset related description

Data Attribute	Description
Data Source	6 common algae species provided by Liaoning Institute of Marine Technology (Chlorella, Karenia, etc.)
Data Format	RGB images taken by microscope (.jpg), annotated in YOLO format (.txt)
Data Scale	1800 original images, expanded to 5400 after augmentation (rotation, noise, cropping)
Data Characteristics	Small targets (microalgae occupy <5% image area), multi-scale, complex background

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

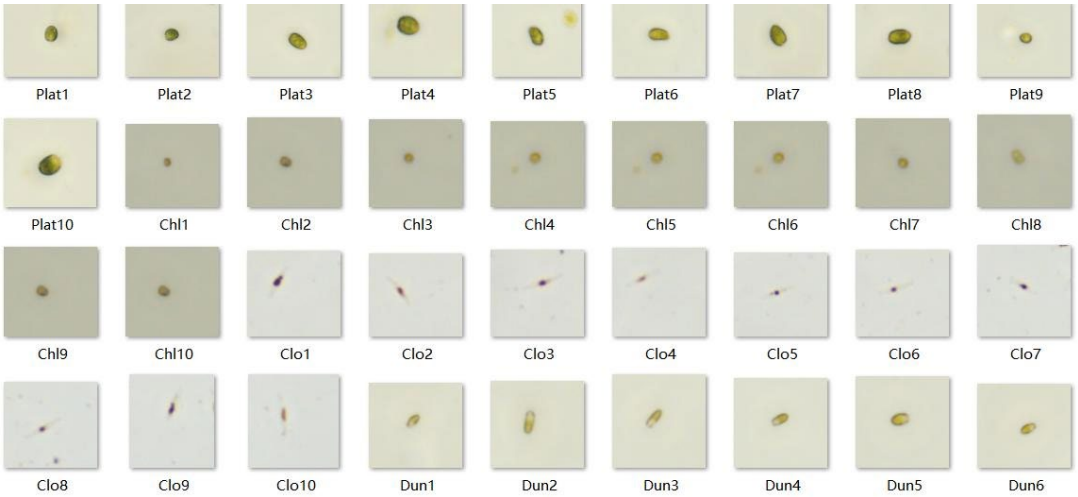

Fig. 5 Partial schematic diagram of the construction of the dataset

In terms of experimental settings, to scientifically evaluate model performance, the dataset was strictly divided into a training set, a validation set, and a test set in a ratio of 8:1:1. This division ensures sufficient data for model parameter learning while reserving independent data for hyperparameter tuning and final performance testing. In addition, the dataset possesses significant small target characteristics (microalgae occupy less than 5% of the image area), multi-scale distribution, and complex background interference, posing high challenges to detection algorithms.

3.2 Result Analysis

To verify the superiority of the GIFF-AlgaeDet model, this section first analyzes the specific contributions of each improved module to model performance through ablation studies, and then compares it horizontally with current mainstream object detection algorithms.

3.2.1 Ablation Study Analysis

This study used YOLOv7 as the baseline model (Baseline) and introduced the CAGS attention mechanism, SIoU loss function, and lightweight dual detection heads (TwoHeads) step by step in the same experimental environment to investigate the effectiveness of each component. Detailed ablation experiment results are recorded in Table 2.

Table 2 Ablation experiment comparison results

Model	CAGS	SIoU	TwoHeads	AP	mAP@0.5	mAP@0.95	Speed	Parameters
Baseline	×	×	×	90.6%	93.2%	56.3%	0.0501 s	36.5M
Ablation1	√	×	×	93.5%	96.2%	63.7%	0.0500 s	37.4M
Ablation2	√	√	×	93.5%	96.0%	64.7%	0.0455 s	37.4M
GIFF	√	√	√	93.7%	95.2%	66.1%	0.0219 s	26.3M

The experimental results show that each improved module had a positive impact on model performance. First, after introducing the CAGS attention mechanism into the baseline model (Ablation 1), the model improved by 3.0% and 7.4% in mAP@0.5 and mAP@0.95 metrics respectively, confirming the effectiveness of the CAGS module in capturing spatial positional information of microalgae and enhancing feature expression. Subsequently, after replacing the loss function with SIoU (Ablation 2), mAP@0.95 further increased to 64.7%, and the convergence speed during training was significantly accelerated, benefiting from the effective penalty of SIoU on boundary box angle deviation. Finally, after introducing lightweight dual detection heads combined with re-parameterization technology, the complete GIFF model was formed. The results show that the GIFF model not only further increased mAP@0.95 to 66.1%, but more critically, its model parameter count dropped significantly from 36.5M in the baseline to 26.3M, and the inference time for a single image was shortened from 0.0501 seconds to 0.0219 seconds. This indicates that the dual detection head design successfully removed redundant calculations for large targets, achieving a double breakthrough in precision and speed.

To demonstrate the improvement effect more intuitively, Figure 6 plots the mAP@0.95 change curve and Training Loss decline curve of each model within 120 training epochs. It can be seen that the GIFF model (red line) showed a faster convergence trend in the early stage of training, and the final convergence value was significantly higher than that of the baseline model.

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

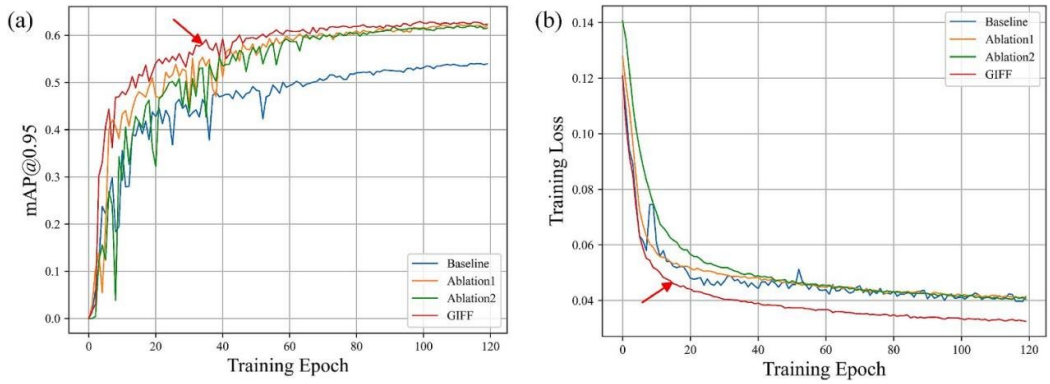

Fig. 6 Model training performance progress based on 120 epochs

In addition, the heat map visualization analysis in Figure 7 shows that compared to the baseline model focusing on background noise, the GIFF model can focus more precisely on the microalgae target area, significantly reducing background interference.

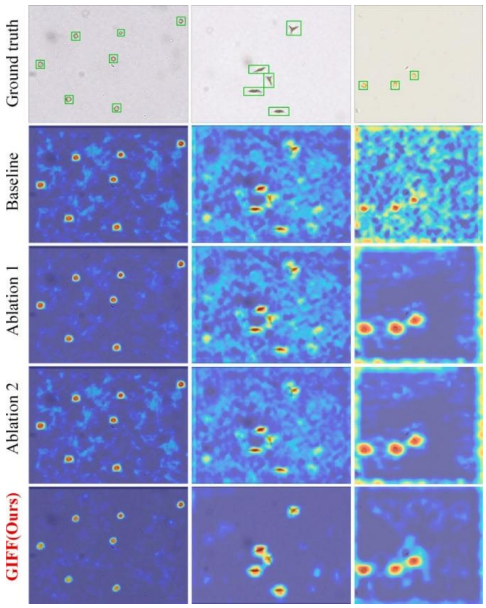

Fig. 7 Visualization analysis of ablation experiments

3.2.2 Comparative Experiment Analysis

To evaluate the advanced nature of GIFF-AlgaeDet in the field of microalgae detection, we comprehensively compared this model with mainstream object detection algorithms such as Faster R-CNN, DETR, and the YOLO series (v4, v5, v6, v7, v8). All models were trained and tested on the same microalgae dataset, and detailed performance comparison data is shown in Table 3.

Table 3 Comparison results with other models

Model	R	mAP@0.5	mAP@0.95	GFOPS	Parameters
Faster R-CNN	85.0%	84.9%	41.6%	207	40M
DETR	90.5%	92.2%	51.3%	225	41M
YOLOv4	92.1%	83.4%	38.1%	119	52.5M
YOLOv5l	94.7%	95.0%	65.0%	107.7	46.5M
YOLOv5x	93.3%	93.9%	65.4%	203.9	86.7M
YOLOv6m	94.0%	95.6%	64.8%	85.8	34.9M
YOLOv6l	94.5%	94.9%	63.8%	150.7	59.6M
YOLOv7	94.2%	93.2%	56.3%	103.2	36.5M
YOLOv8n	72.7%	84.8%	62.4%	8.1	30.0M
YOLOv8l	72.8%	84.7%	64.2%	164.8	43.6M
YOLOv8x	72.6%	84.6%	65.0%	257.4	68.1M
GIFF	96.2%	95.2%	66.1%	98.7	26.3M

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

In terms of detection precision, the GIFF model reached 66.1% on the mAP@0.95 metric, significantly outperforming YOLOv7 (56.3%) and YOLOv5l (65.0%), and even surpassing the huge-parameter YOLOv8x (65.0%). On the mAP@0.5 metric, the GIFF model also reached an extremely high level of 96.2%, proving its stable detection capability under different IoU thresholds. In terms of computational efficiency, the parameter count of the GIFF model is only 26.3M, far lower than YOLOv5x (86.7M) and YOLOv8x (68.1M), and GFOPS is controlled at 98.7, only about one-third of YOLOv8x. This indicates that while maintaining SOTA-level precision, the GIFF model greatly reduces computational resource consumption, making it more suitable for deployment on portable embedded devices.

Figure 8 provides a comparison of visual detection results of different models on actual test samples. It can be clearly observed that in cases of dense targets or occlusion, Faster R-CNN and early YOLO versions are prone to missed detections or false detections (as shown by the dotted circles in the figure), while the GIFF model can accurately identify all microalgae targets and give higher confidence scores. These results fully prove the superior performance of the GIFF-AlgaeDet algorithm in solving the difficulties of small microalgae target detection.

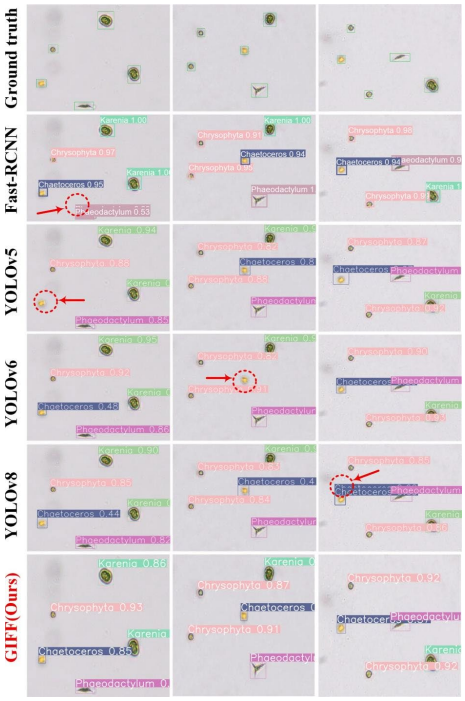

Fig. 8 Schematic diagram of detection results of different methods

4. Conclusion

This study successfully developed a portable microalgae detection and identification system based on deep learning, addressing the deficiencies of traditional microalgae detection technologies in timeliness, portability, and automation. By deeply integrating optical imaging technology with artificial intelligence algorithms, the system achieved full automation from sample collection to result output. In terms of hardware design, the innovative dual-mode imaging module not only possesses micron-level resolution capable of clearly capturing microalgae morphological details and fluorescence activity features, but also achieves high device integration and lightweighting through 3D printing technology, completely breaking away from dependence on large laboratory equipment, greatly reducing hardware costs and usage thresholds, and making real-time on-site monitoring in the field possible.

At the core algorithm level, the GIFF-AlgaeDet model proposed in this study effectively solves the recognition difficulties caused by dense small targets, strong background interference, and high feature similarity in microalgae images. By introducing the CAGS attention mechanism, the model can more precisely focus on target areas of interest and suppress background noise; the combination of lightweight dual detection heads and re-parameterized design successfully improves inference speed while reducing redundant model parameters, balancing the contradiction between detection precision and computational efficiency; and the application of the SIoU loss function further optimizes the regression quality of

ISSN 2616-5767 Vol.9, Issue 2: 27-36, DOI: 10.25236/AJETS.2026.090204

boundary boxes. Experimental data shows that the system performs excellently in key indicators such as detection precision, recall rate, and running speed, especially demonstrating robustness surpassing mainstream algorithms when dealing with minute targets. In summary, the results of this study not only provide a powerful technical tool for rapid classification and counting of microalgae, but also offer important reference value and broad application prospects for intelligent upgrades in fields such as environmental monitoring, marine ecological research, and aquaculture.

References

[1] L. Junfei, C. Ying, S. Yu, and C. Ting, "An approach to microalgae identification based on joint optimization model of slicing and detection," Expert Systems with Applications, vol. 280, no. Compendex, 2025, doi: 10.1016/j.eswa.2025.127565.

[2] J. Liu, Y. Chen, C. Wang, W. Li, and D. Zhu, "Motif field combined with two-stream feature fusion network and double detection head for identification and prediction of microalgae in seawater," Water Research, vol. 288, no. Compendex, 2026, doi: 10.1016/j.watres.2025.124574.

[3] J. Scarfe, D. Kosmutzky, and R. E. R. Nisbet, "A game of tag: A review of protein tags for the successful detection, purification and fluorescence labelling of proteins expressed in microalgae," Plant Journal, vol. 122, no. Compendex, 2025, doi: 10.1111/tpj.70272.

[4] J. Borisova, I. V. Morshchinin, V. I. Nazarova, N. Molodkina, and N. O. Nikitin, "Low-Cost Microalgae Cell Concentration Estimation in Hydrochemistry Applications Using Computer Vision," Sensors, vol. 25, no. Compendex, 2025, doi: 10.3390/s25154651.

[5] Zhou S , Chen T , Fu E S ,et al. A microfluidic microalgae detection system for cellular physiological response based on an object detection algorithm. Lab on a Chip, 2024, 24(10):12. DOI:10.1039/D3LC00941F.

[6] T. Eerola et al., "Survey of automatic plankton image recognition: challenges, existing solutions and future perspectives," Artificial Intelligence Review, vol. 57, no. Compendex, 2024, doi: 10.1007/s10462-024-10745-y.

[7] J. Luo, "Principal Component Analysis–Based Feature Recognition Technology for Marine Plankton Images in IoT Systems," Journal of Testing and Evaluation, vol. 52, no. Compendex, 2024, doi: 10.1520/JTE20230079.

[8] Q. Ye, Y. Shi, and S. Guo, "CAGs-Net: A Novel Adjacent-Context Network With Channel Attention Gate for 3D Brain Tumor Image Segmentation," International Journal of Biomedical Imaging, vol. 2025, no. Compendex, 2025, doi: 10.1155/ijbi/6656059.

[9] G. Li et al., "LR-Net: Lossless Feature Fusion and Revised SIoU for Small Object Detection," Computers, Materials and Continua, vol. 85, no. Compendex, pp. 3267–3288, 2025, doi: 10.32604/cmc.2025.067763.

[10] F. Ren, J. Fei, H. Li, and B. T. Doma, "Steel Surface Defect Detection Using Improved Deep Learning Algorithm: ECA-SimSPPF-SIoU-Yolov5," IEEE Access, vol. 12, no. Compendex, pp. 32545–32553, 2024, doi: 10.1109/ACCESS.2024.3371584.