



## OPEN ACCESS

## EDITED BY

Zhao Yunpeng,  
Dalian University of Technology, China

## REVIEWED BY

Houjie Li,  
Dalian Minzu University, China  
Jianchang Liu,  
Northeastern University, China  
Ning Wang,  
Dalian Maritime University, China

## \*CORRESPONDENCE

Yuan Cheng  
cheny\_nbi@dlut.edu.cn

## SPECIALTY SECTION

This article was submitted to  
Marine Fisheries, Aquaculture and  
Living Resources,  
a section of the journal  
Frontiers in Marine Science

RECEIVED 15 October 2022

ACCEPTED 02 November 2022

PUBLISHED 21 November 2022

## CITATION

Liu D, Wang P, Cheng Y and Bi H  
(2022) An improved algae-YOLO  
model based on deep learning for  
object detection of ocean  
microalgae considering aquacultural  
lightweight deployment.  
*Front. Mar. Sci.* 9:1070638.  
doi: 10.3389/fmars.2022.1070638

## COPYRIGHT

© 2022 Liu, Wang, Cheng and Bi. This is  
an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use,  
distribution or reproduction is  
permitted which does not comply with  
these terms.

# An improved algae-YOLO model based on deep learning for object detection of ocean microalgae considering aquacultural lightweight deployment

Dan Liu<sup>1,2</sup>, Pengqi Wang<sup>3</sup>, Yuan Cheng<sup>4\*</sup> and Hai Bi<sup>5</sup>

<sup>1</sup>College of Information Engineering, Dalian Ocean University, Dalian, China, <sup>2</sup>Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education, Dalian, China, <sup>3</sup>College of Mechanical and Power Engineering, Dalian Ocean University, Dalian, China, <sup>4</sup>Ningbo Institute of Dalian University of Technology, Ningbo, China, <sup>5</sup>Research and Development (R&D) Department, Hangzhou Yunxi Smart Vision Technology Co., LTD, Dalian, China

Algae are widely distributed and have a considerable impact on water quality. Harmful algae can degrade water quality and be detrimental to aquaculture, while beneficial algae are widely used. The accuracy and speed of existing intelligent algae detection methods are available, but the size of parameters of models is large, the equipment requirements are high, the deployment costs are high, and there is still little research on lightweight detection methods in the area of algae detection. In this paper, we propose an improved Algae-YOLO object detection approach, which is based on ShuffleNetV2 as the YOLO backbone network to reduce the parameter space, adapting the ECA attention mechanism to improve detection accuracy, and redesigning the neck structure replacing the neck structure with ghost convolution module for reducing the size of parameters, finally the method achieved the comparable accuracy. Experiments showed that the Algal-YOLO approach in this paper reduces the size of parameters by 82.3%, and the computation (FLOPs) is decreased from 16G to 2.9G with less loss of accuracy, and mAP by only 0.007 when compared to the original YOLOv5s. With high accuracy, the smaller model size are achieved, which reduces the equipment cost during actual deployment and helps to promote the practical application of algae detection.

## KEYWORDS

object detection, lightweight, ShuffleNetV2, ECA attention, algae

## Introduction

Algae are large and intricate biomes found in the ocean and rivers, and they have a substantial impact on the natural environment and water quality. Harmful algal blooms (HABs) can wreak in the water environment, affecting water quality and biological survival, as well as easily resulting in the loss or death of organisms in aquaculture zones. HABs can cover a large area of the water surface in a short time, blocking sunlight and thus affecting plant photosynthesis. In addition, algal blooms consume a large amount of oxygen in the water body, releasing undesirable gases and toxic substances and damaging the seawater and river water environment, which is not conducive to aquatic organism survival. Microalgae can also act as clean energy sources and purify wastewater, and some algae can be used as food for humans. The various uses of microalgae contribute significantly to reducing the greenhouse effect, saving energy and protecting the environment. Therefore, deploying algae monitoring systems to manage the abundance and types of algal populations in the water environment is necessary for aquaculture and the water environment.

One of the most frequent conventional techniques for detecting algal cells is manual microscopy using a microscope, which is not only time-consuming but also requires specialized knowledge of algae. As processing power has increased, computer vision has evolved, and support vector machines (SVMs) are increasingly employed for algal detection. For example, Xu Y (Xu et al., 2014), Babu M J (Babu et al., 2016) and Shan S (Shan et al., 2020) investigated the use of SVM in forecasting red tide, while Tao J (Tao et al., 2010) and Göröcs Z (Göröcs et al., 2018) combined a support vector machine with flow cytometry, and its accuracy reached 90%. Artificial neural networks (ANNs) have a better solution to this problem with significantly less manual involvement and generally higher accuracy. Mosleh M A A (Mosleh et al., 2012) and Park J (Park et al., 2019) used artificial neural networks for algal feature extraction with 93% accuracy, while Medina E (Medina et al., 2017) and Schaap A (Schaap et al., 2012) compared convolutional neural networks with MLP and discriminant analysis classification methods in artificial neural networks and found that convolutional neural networks possess greater performance.

Deep learning research has advanced with the development of computer image processing technology. Ning Wang et al (Wang et al., 2022) also prove that deep learning can solve a series of marine problems. In these years, Two-stage object detection approaches have swiftly progressed from R-CNN (Girshick et al., 2014) to FastRCNN (Girshick, 2015) and FasterRCNN (Ren et al., 2015), with advances in model design and additional increases in detection speed and accuracy, although their speed is not yet sufficient for real-time detection. In algae research, Samantaray A (Samantaray et al., 2018) proposed a computer vision system based on deep learning for algae monitoring with a wide range of applicable platforms but only 82% accuracy. Cho S (Cho et al., 2021) and Deglint J L (Deglint et al., 2019) explored the potential application

of deep learning on algae in conjunction with 3D printing and other devices with sufficient accuracy, but its cost is high. Hayashi K (Hayashi et al., 2018) investigated the detection of cell division in algae and separated mitosis from interphase division with a 92% degree of accuracy. The mean average precision (mAP) was between 74% and 81% in the deep learning framework proposed by Qian P (Qian et al., 2020). Ruiz-Santaquiteria J (Ruiz-Santaquiteria et al., 2020) compared instance segmentation with semantic segmentation for diatoms versus nondiatoms and found instance segmentation to be more accurate by 85%. The preceding study is significantly weak in detection speed and detection accuracy, as well as identification that is quick and accurate.

One-stage object detection approaches emphasize the speed of recognition. YOLO (Redmon et al., 2016) was the first one-stage object detection method, which is quicker than the two-stage approach but with the loss of precision and has undergone iterative development with several variants. Using YOLOv3 for algae detection, Park J (Park et al., 2021) obtained a mAP of 81 for five algal species. Tingkai Chen et al (Chen et al., 2021) proposed one-stage CNN detector-based benthonic organisms detection scheme, which used the generalized intersection over union (GIoU) for localizing benthonic organisms easily. Cao M (Cao et al., 2021) used MobileNet to enhance the backbone network of YOLOv3, hence reducing the weight of the model. Ali S (Ali et al., 2022) investigated YOLOv3, YOLOv4, and YOLOv5 and demonstrated that YOLOv5 produced the best outcomes. Park J (Park et al., 2022) studied YOLOv3, YOLOv3-small, YOLOv4 and YOLOv4-tiny, and the findings showed that the micro (tiny) model was superior in terms of detection speed and precision. Based on the algae detection advancements for YOLOv3, YOLOv4, and YOLOv5, the YOLO framework may possibly be used for algae detection, and the tiny model is enhanced with YOLOv5; however, there is a dearth of study on the improved approaches for algae identification in lightweight YOLOv5.

Based on the current status of the research literature, the research on algae has been constantly developing. From the improvement in traditional manual detection methods to the application of traditional computer vision, it can be seen that research on algae detection is constantly increasing. After the emergence of artificial intelligence technology, a large number of deep learning approaches have been used for algae detection, which is enough to prove the importance of microalgae detection. However, the aforementioned neural network technique for detecting algae has complex deep neural network architecture and a large number of parameters, which require high computational power and high deployment costs in a realistic setting, and there are no lightweight solutions for YOLOv5 algae identification. From the standpoint of research images, the majority of study objects are high-magnification microscopic images of algae, and there is a dearth of research on low-magnification microscopic photographs of algae. Although the pictures of algae at high magnification are clear, the number of cells in the field of view is much smaller than that of low-magnification microscopy. Furthermore, the high-magnification

microscope is more costly. In the low-magnification micrographs of algae, the cells are closely packed and difficult to identify with the naked eye, and low-magnification microscopes is cheaper.

Due to the high cost of high magnification microscopy acquisition, we used low magnification microscopy to acquire algae microscopic images to reduce the cost in the data collection. For the practical deployment of algae detection, we propose a lightweight algae object detection network, named Algae-YOLO, based on YOLOv5, by replacing original backbone network with lightweight module of ShuffleNetV2, thereby significantly reducing the model size and computation, and by adding an ECA attention mechanism to ShuffleNetV2, the detection accuracy significantly improves without increasing computational cost. In the neck structure, a lightweight structure combined with ghost convolution is designed, and the depth of the network and the width of the feature layer input are increased to ensure the balance between model accuracy and computational effort, allowing for low-cost deployment while achieving accuracy comparable to that of YOLOv5s.

## Related work

### Low-magnification microscopic image acquisition of algae

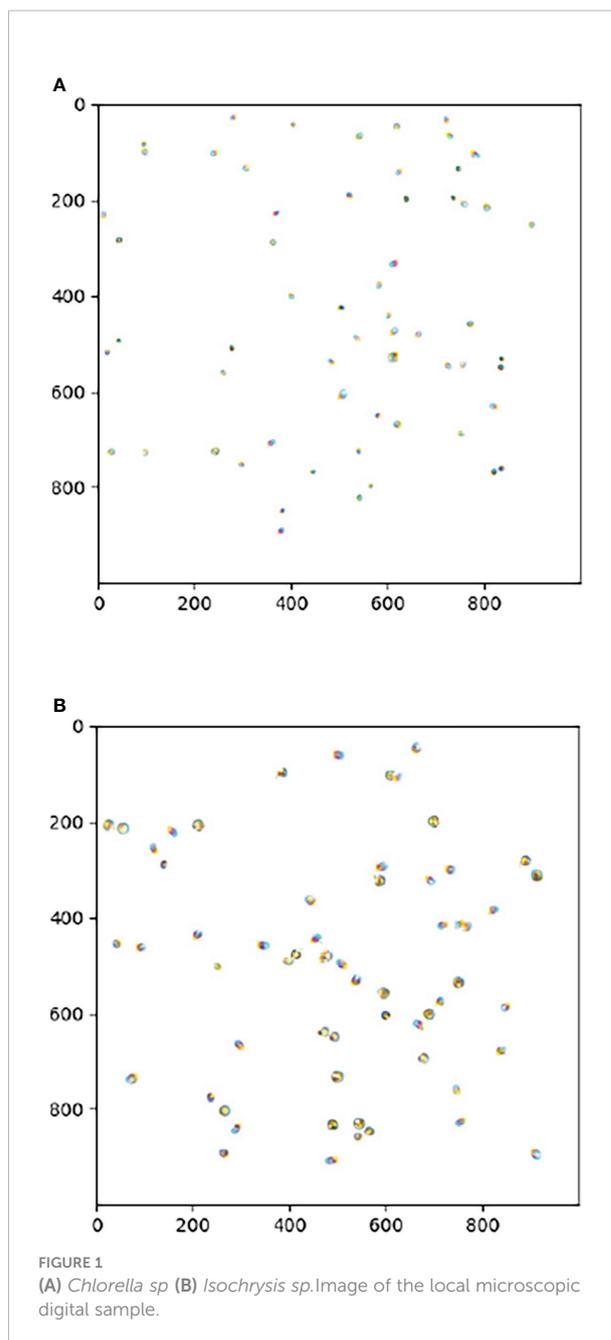
Microalgae strains (laboratory grown) of *Chlorella* sp. and *Isochrysis* sp. samples were photographed in full color using a Leica DM4 B digital microscope with a low-magnification. Sixty images of each of the two kinds of low-magnification algae were gathered, the number of algae in each sample photograph was between 100-200, and local pictures of the samples are shown in Figure 1.

### Data preprocessing

The original algal microscopic photos were few, their data were in TIF format with 1920\*1200 resolution, and the algal dispersion was dense and abundant. To minimize the difficulty of time-consuming manual annotation, data preprocessing was performed to obtain algal independent cells, and then the dataset was automatically constructed utilizing independent cells.

In the original image preprocessing, the image was read by OpenCV and then threshold and grayed out using a binarization method. The image was then stretched and eroded to complete the algae morphology, contour detection was conducted, and the set of algae border coordinates was located and saved as a distinct cell image.

After obtaining the independent cells, they were randomly pasted into blank pictures, and the location data were processed according to the YOLO dataset format. Figure 2 depicts the resulting photos of *Chlorella* sp., *Isochrysis* sp., and a mixture of



both. Some portions of the illustration are exhibited with local magnification for observational ease.

### Data enhancement

Considering the issue of algal cell occlusion in microscopic images, this study employs the mask technique to produce a training dataset of algal occlusion images. The previous work created distinct photographs of *Chlorella* sp. and *Isochrysis* sp.; therefore, the masking technique may be utilized to generate

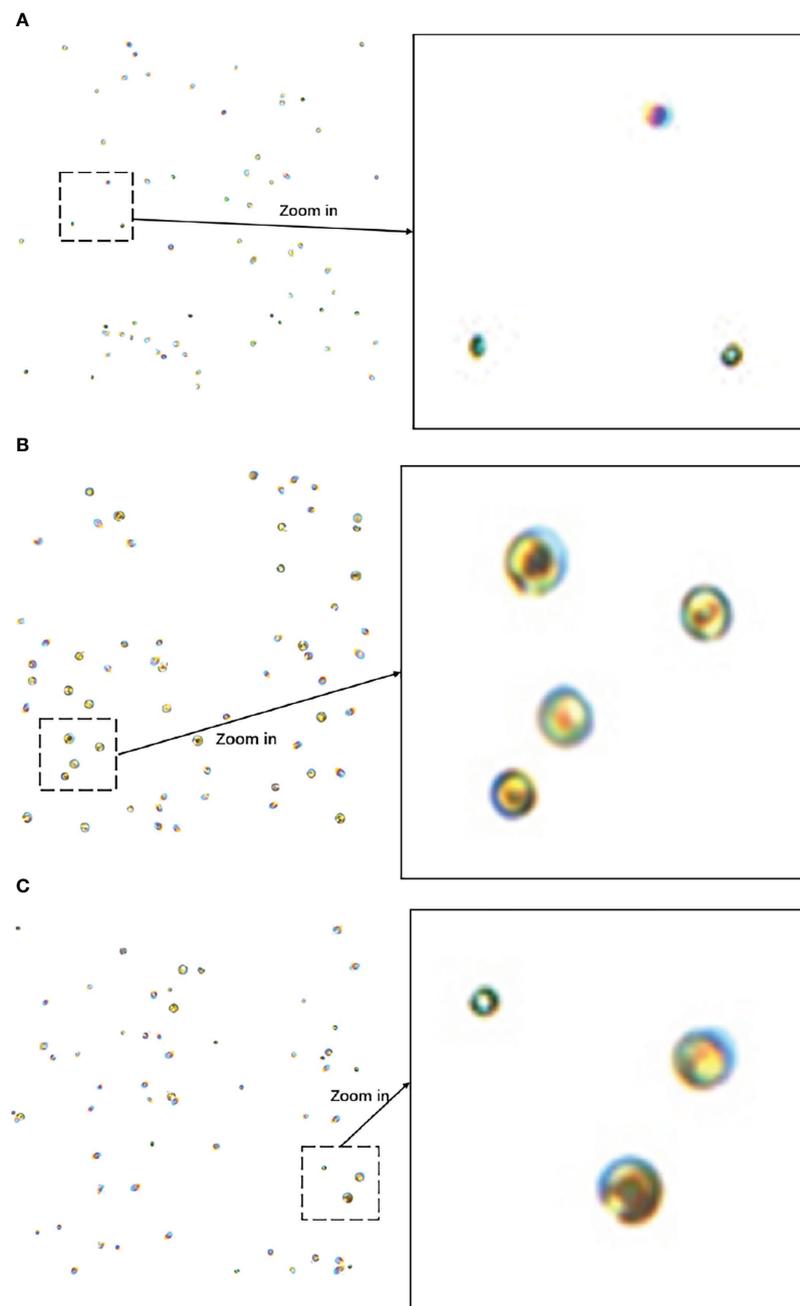


FIGURE 2

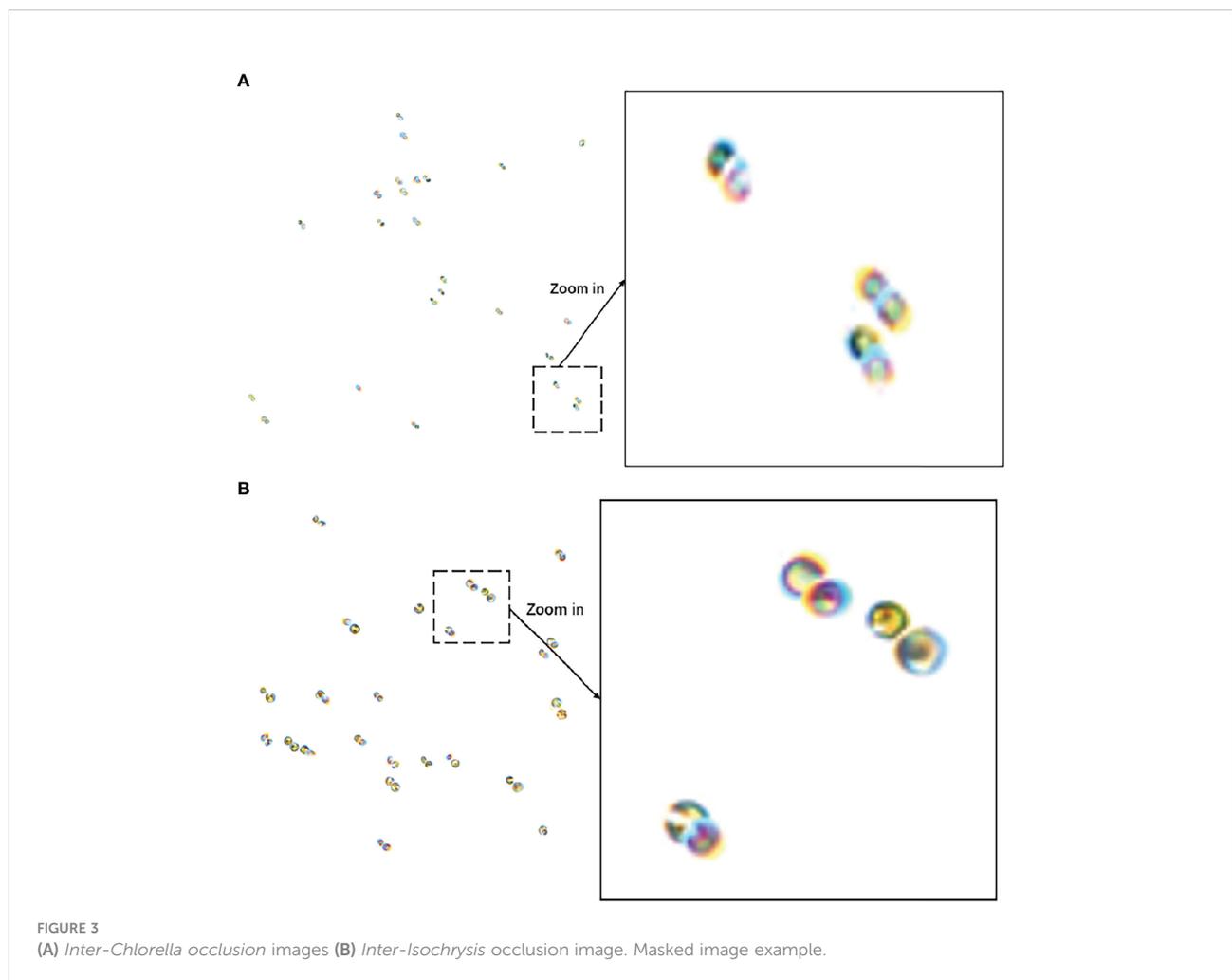
(A) Image of *Chlorella* sp. (B) *Isochrysis* sp. image (C) Mixed image of *Chlorella* sp. and *Isochrysis* sp. Example images of the algae dataset.

masked data images during the production of dataset images. The main process is to paste the first algal cell at random onto a blank image and then superimpose the second algal cell on top of the first, with the overlap area between the two set at 20%. An illustration of masked image data is shown in Figure 3.

To improve the generalization performance of the dataset, this research darkens the processed algae dataset photos and uses

them to replicate algae microscopic images in low-light circumstances. Its darkened image is shown in Figure 4.

2,000 photographs of *Chlorella* sp., 2,000 images of *Isochrysis* sp., 3,000 images of mixed algae, and 3,000 images of masked images were generated for this study. The dataset contains a total of 10,000 images, and the ratio of the training set, validation set, and test set is 8:1:1. In the process of the picture input model,



mosaic data enhancement is also activated, and four images are stitched together by random cropping and scaling so that the actual number of images involved in training is increased and the sample set is more diverse.

## Method introduction

### YOLOv5

Since its release in 2020, YOLOv5 has been one of the most popular object detection approaches by combining the latest research in a variety of fields with other approaches to make YOLO lighter and more efficient. YOLOv5 is composed of a backbone network (Backbone), a neck structure (Neck) and a prediction head (Head). Figure 5 depicts the backbone network structural diagram, which comprises focus, CSP1, CSP2, SPPF, and other structures.

After focus segmentation, the input image data are sent to the backbone network for feature extraction, processed by spatial

pyramid pooling-fast (SPPF), sent to the neck structure for the feature fusion operation, and finally predicted and output in the prediction head. Figure 6 illustrates the overall YOLOv5 flowchart.

### Attention Module

Efficient channel attention (ECA) is a novel lightweight attention mechanism (Hayashi et al., 2018). Compared to SE attention, ECA is effective at modifying feature weights while avoiding frequent dimensionality reduction operations. ECA acquires the interaction information of each channel and its  $k$  neighbors with a fast 1D convolution with convolution kernel  $k$  after global average pooling, where the convolution kernel  $k$  is proportionate to the number of channel dimensions to adaptive acquisition, the relationship between  $k$  and  $c$  is as follows:

$$k = \varphi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (1)$$

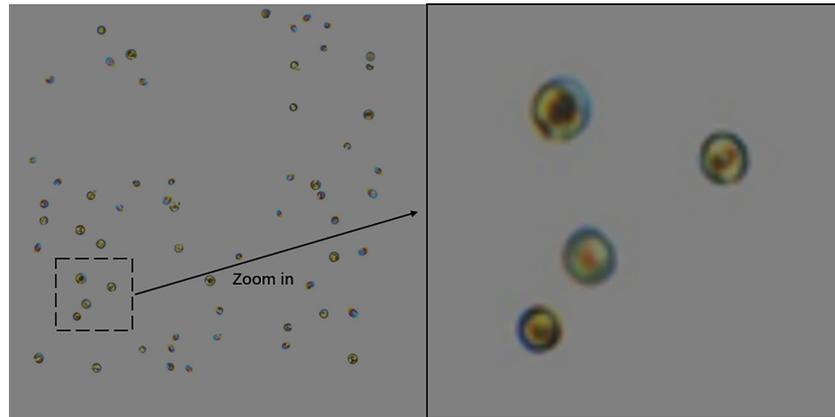


FIGURE 4 Example of darkened *Isochrysis* sp. image.

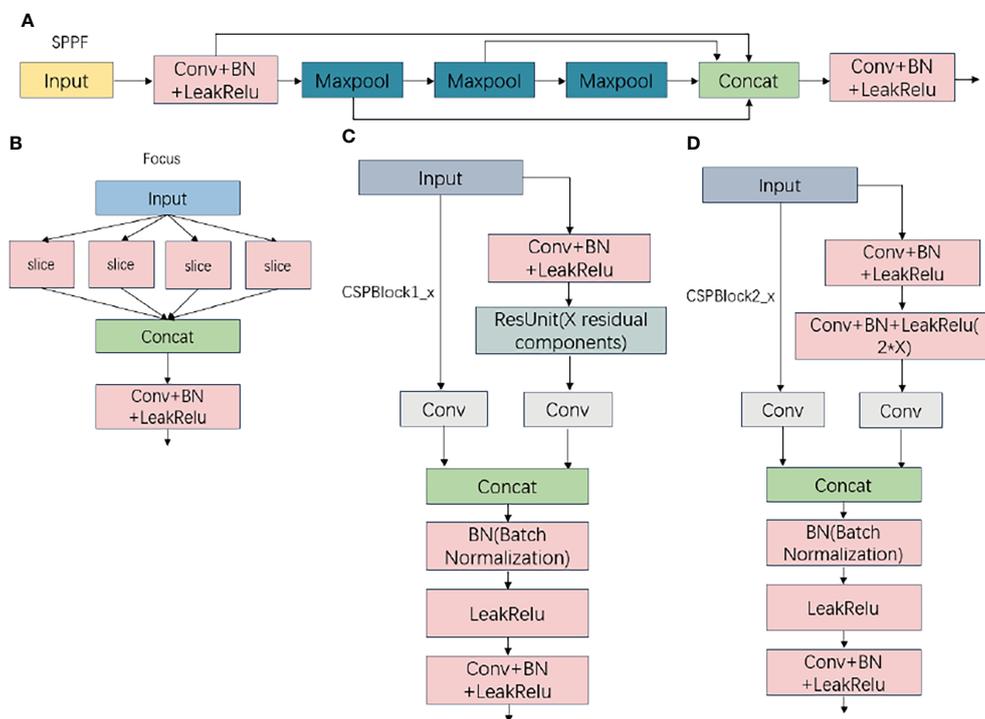


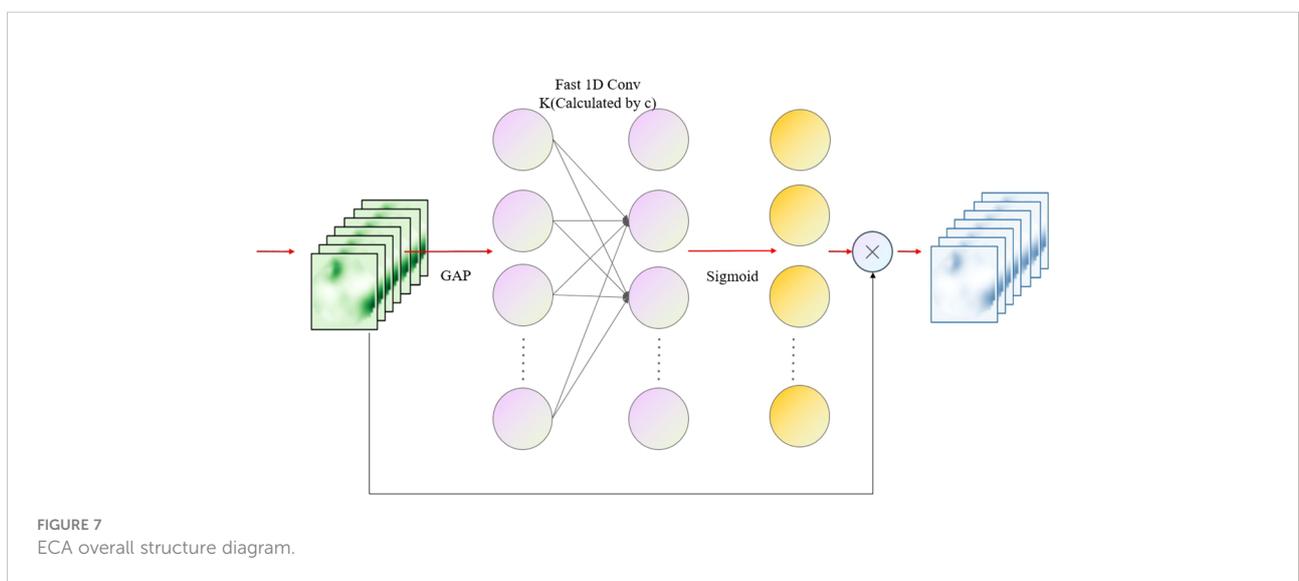
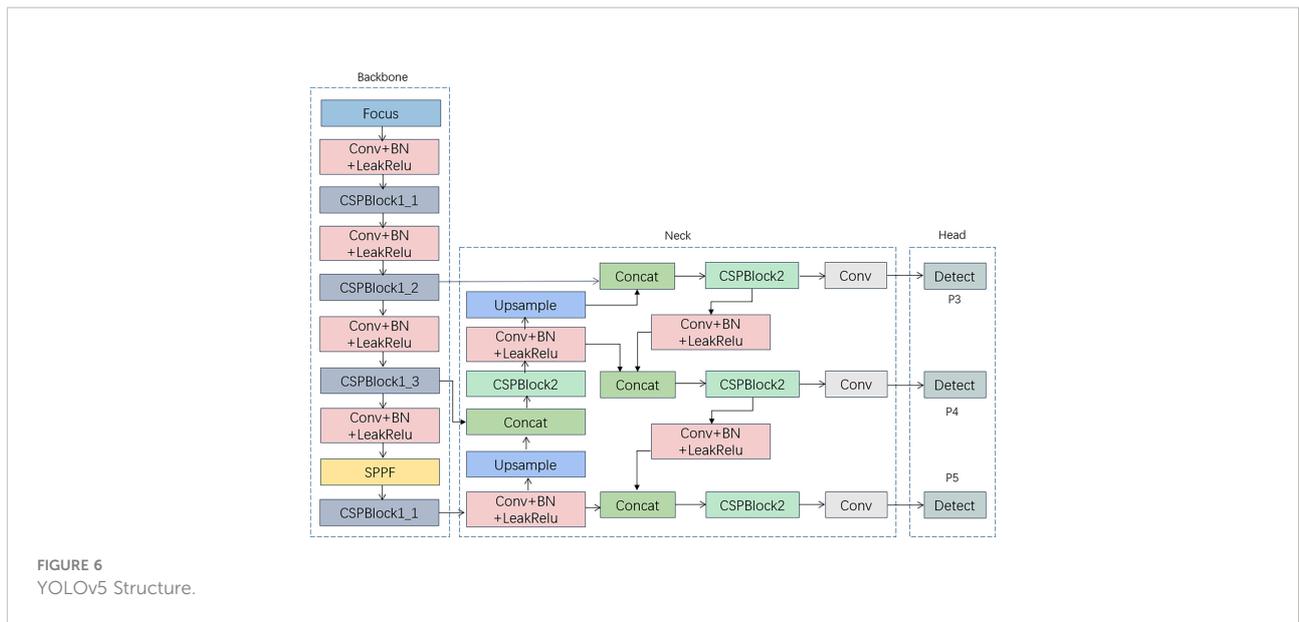
FIGURE 5 (A) SPPF Structure (B) Focus structure (C) CSPBlock1\_x Structure (D) CSPBlock2\_x Structure Images of Focus, CSP1, CSP2, and SPPF structures.

where  $\gamma$  and  $b$  are taken to be 2 and 1 respectively, and  $|t|_{odd}$  means the oddest number closest to  $|t|$ . To achieve the parameter complexity of tillage, the weights of all channels are shared. In this calculation, the  $w_k$  array containing  $K \times C$  parameters is utilized to learn the attention, and the weight  $\omega_i$  of channel  $y_i$  is calculated by considering only the interaction information

between  $y_i$  and its  $k$  neighbors, which is expressed as:

$$\omega_i = \sigma \left( \sum_{j=1}^k w_i^j y_i^j \right), y_i^j \in \Omega_i^k \quad (2)$$

where  $\Omega_i^k$  denotes the set of  $k$  neighboring channels of  $y_i$  and  $\sigma$  is the sigmoid activation function. The information interaction is



followed by weight sharing, which is given by:

$$\omega_i = \sigma \left( \sum_{j=1}^k w^j y_i^j \right), y_i^j \in \Omega_i^k \quad (3)$$

The entire calculation procedure is dependent on the fast 1D convolution with k convolution kernels (adaptive size of k) in the ECA structure, which is denoted by:

$$\omega = \sigma(C1D_k(y)) \quad (4)$$

where C1D is the fast 1D convolution. This cross-channel information interaction attention may effectively increase the effectiveness of attention without introducing additional

computing factors. The overall structure of ECA is shown in Figure 7.

### ShuffleNetV2

The ShuffleNetV2 network is based on the ShuffleNet network, and the network building blocks of V1 are modified in accordance with the four guidelines for designing efficient networks to reduce the memory access cost (MemoryAccessCost, MAC) of the network convolution process. Figure 8 depicts the structures of the two new ShuffleNetV2 units. In ShuffleNetV2 cell 1, the feature map of the input C-channel number is divided into two copies, C1 and C2, and

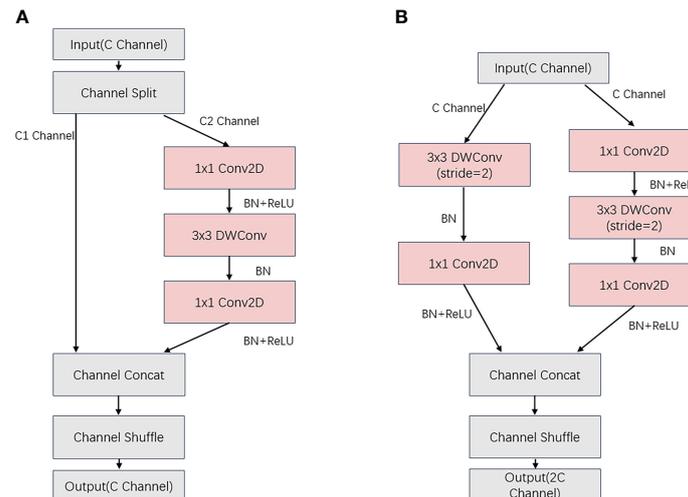


FIGURE 8  
(A) ShuffleNetV2 Unit1 (B) ShuffleNetV2 Unit2.

the feature map of the C1 channel number is not modified, while the feature map of the C2 channel number is passed through three convolution layers with the same input and output channels, including two 1X1 ordinary convolutions (Conv2D) and one 3X3 depth-separate convolution (DWConv). After the two branches are completed, the Concat operation is executed to concatenate the channels, which are fused into an initial C-channel feature map, and interact with the information data *via* channel blending. In ShuffleNetV2 cell 2, the input C-channel count feature map is no longer subjected to the channel division operation and is instead divided into two separate convolution operations. The same Concat operation is used after the convolution operation, and the 2C-channel count feature map is output after channel blending so that the network width is increased with fewer parameters added.

## GhostConv module

GhostConv is a type of convolution proposed by the GhostNet neural network that employs fewer parameters to obtain more target features during computation. GhostConv is separated into two portions of normal convolution at the beginning of the convolution procedure. To build a portion of the feature map, the first step is a small amount of convolution using a convolution kernel with fewer channels than standard convolution. The objective of this convolution is to integrate the features, generalize and condense the input feature map, and generate a new feature map. The second step is to execute a reduced computational operation on the feature map obtained in the first stage using deep separable convolution to build further feature maps. The two created feature maps are then stitched together using the Concat technique to create the ghost feature

map. An example of the ghost convolution process is shown in Figure 9.

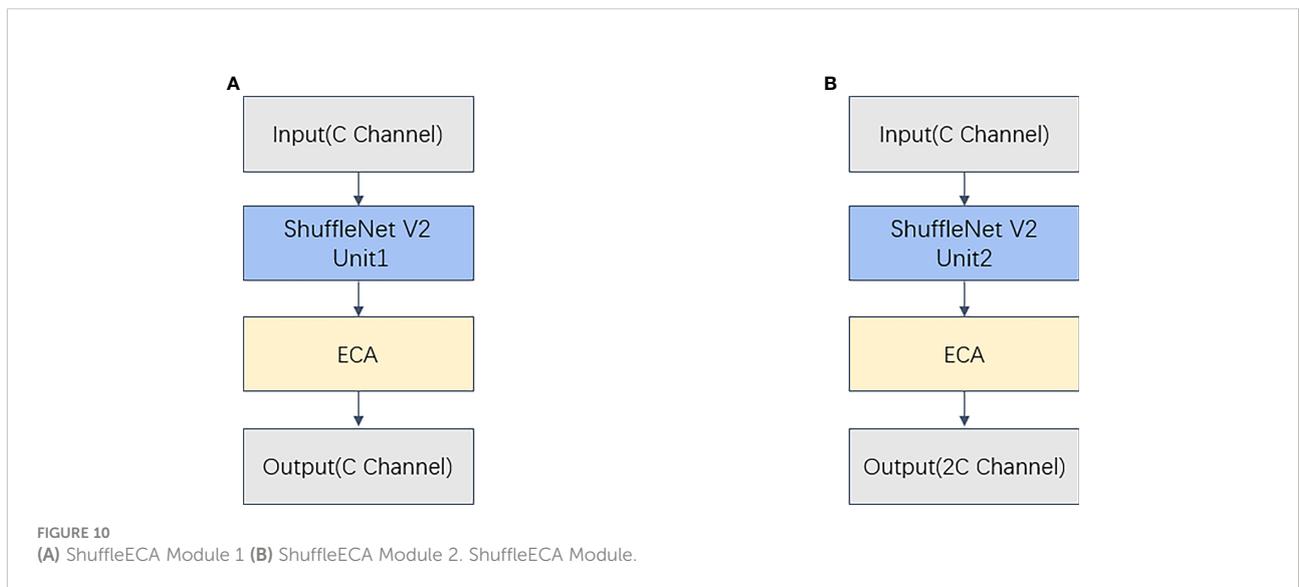
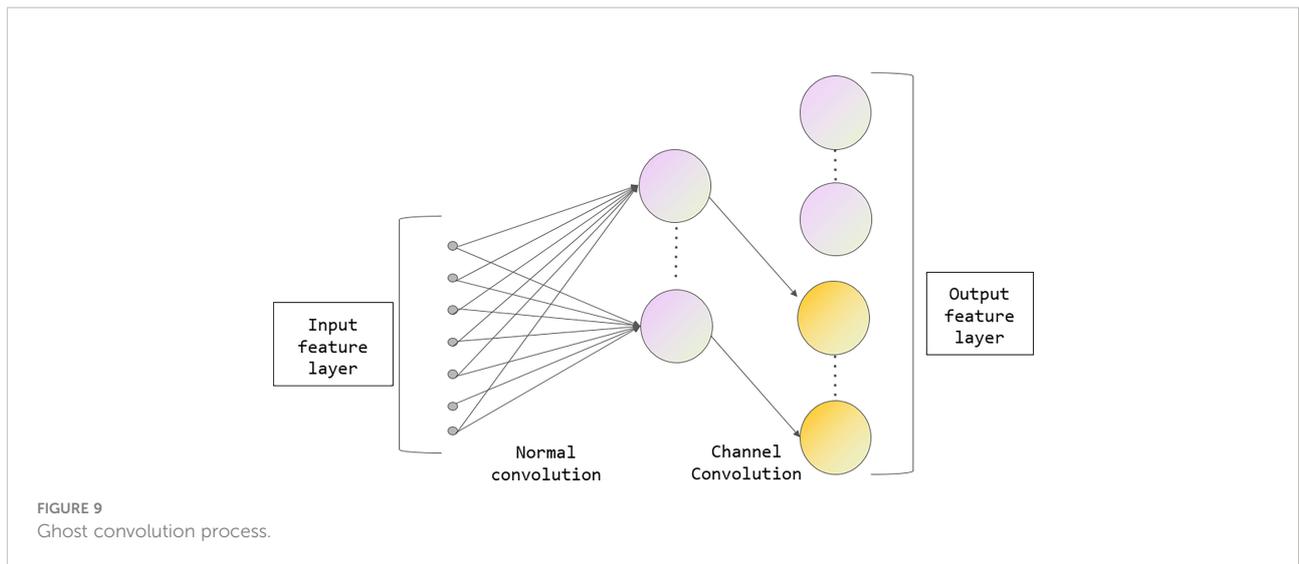
## Improvement

### ShuffleNetV2-ECA backbone network

The original backbone network (Backbone) component of YOLOv5 employs CSPDarknet53, which offers excellent feature extraction capabilities but cannot be implemented in a lightweight and cost-effective way due to its complicated network design. Therefore, this study chooses the lightweight network structure ShuffleNetV2 and enhances it based on its design principles, and all improvements in the paper are based on YOLOv5s.

In this study, ShuffleNet modules are constructed utilizing ShuffleNetV2 units 1 and 2 as a lightweight ShuffleNet backbone network for YOLOv5. To improve the feature extraction capability of the backbone network, ECA attention is combined with ShuffleNetV2 unit 1 and unit 2 to form ShuffleECA module 1 and module 2, respectively, which further improves the cross-channel capability and feature extraction capability with only a small number of parameters involved. ShuffleECA module 1 and module 2 are shown in Figure 10.

Using the ShuffleECA module, the ShuffleNetV2-ECA backbone network is then established. Table 1 displays the network structure of its upgraded backbone. ShuffleECA module\_1 indicates that the layer is ShuffleECA module 1, whereas ShuffleECA module\_2 indicates that ShuffleECA module 2 is in use. The three parameters in the argument



represent the number of input channels, the number of output channels, and the step size. Parameters indicate the number of layer structure parameters.

After the feature map is input to the backbone network, the original YOLOv5 first performs the focus module, and after multiple slicing together in the way of channel expansion, which creates more memory overhead and violates the ShuffleNetV2 design principle, the focus is replaced with a convolutional kernel of 3\*3 ordinary convolution with a BN layer, activation function, and maximum pooling convolutional combination, i.e., Conv\_BN\_ReLU\_Maxpool in Table 1, to keep the network light and efficient while reducing the degree of network fragmentation. In addition, the primary function of the SPPF structure at the end of the original backbone network is to expand the network perceptual field through the use of multiple

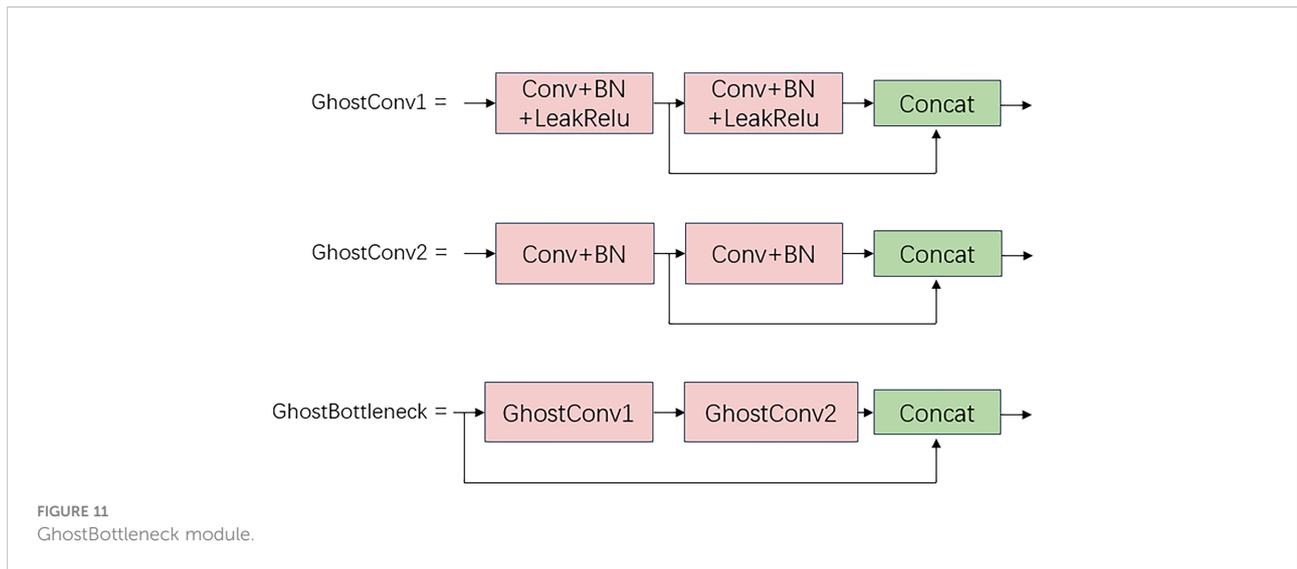
maximum pooling operations for enhanced feature extraction. However, since the algal object targets studied in this paper are tiny, such operations not only create a greater computational burden but also increase the possibility of losing algal targets, so it was decided to remove the SPPF structure from the improved ShuffleNetV2-ECA backbone network.

### Lightweight YOLOv5Neck

To further lower the computing needs of the model deployment, a lighter GhostBottleneck module is built in the neck section by adapting the GhostNet architecture. GhostBottleneck comprises two types of GhostConv and residual structures: GhostConv1 comprises GhostConv and

TABLE 1 ShuffleNetV2-ECA Backbone Network Structure Table.

Serial Number	N	Module	Parameters	Argument
0	1	Conv_BN_Rel_Maxpool	464	[3, 16, 2]
1	1	ShuffleECA Module_2	2,768	[16, 64, 2]
2	3	ShuffleECA Module_1	2,528	[64, 64, 1]
3	1	ShuffleECA Module_2	14,080	[64, 128, 2]
4	7	ShuffleECA Module_1	18,304	[128, 128, 1]
5	1	ShuffleECA Module_2	52,736	[128, 256, 2]
6	3	ShuffleECA Module_1	34,688	[256, 256, 1]



BN layers and activation functions, whereas GhostConv2 consists of only GhostConv and BN layers and can successfully extract image feature information in a lighter structure. The structure diagram of the three is shown in Figure 11. The CSPBlock2 structure of the whole neck component is replaced with GhostBottleneck, and the conventional convolutional CBL (Conv+BN+LeakRelu) is replaced with GhostConv1, resulting in a network with increased feature fusion that is lightweight and efficient.

### Improved lightweight algae-YOLO

YOLOv5 is divided into five models: n, s, m, l and x. The structure remains the same while the model depth and channel width increase in order, with YOLOv5m corresponding to values of 0.67 and 0.75. In the YAML file, the ratio of the input model depth to the channel width of the feature layer is changed from 0.33 vs. 0.5 to 0.67 vs. 0.75, which increases the number of runtime parameters but improves operational precision. The total network architecture is shown in Figure 12.

The algae target images are input from the ShuffleECA backbone network after one convolutional kernel size of 3\*3 ordinary convolution, BN layer normalization, activation function and maximum pooling and thus sent into the two stacked ShuffleECA modules, where N in the figure represents the number of ShuffleECA module repetitions on the right side, which are required to pass into the neck. Greater repetitions of the three ShuffleECA modules are present in the improved feature network. In the neck structure, GhostBottleneck substitutes the original CSPBlock\_2, and GhostConv\_1 replaces the original regular convolution process.

## Experiment and evaluation

### Evaluation metrics

To evaluate the model, research compares a variety of assessment measures. The mAP is the most commonly used evaluation metric in the field of object detection and is calculated from P (Precision) and R (Recall), with P and R being calculated as:

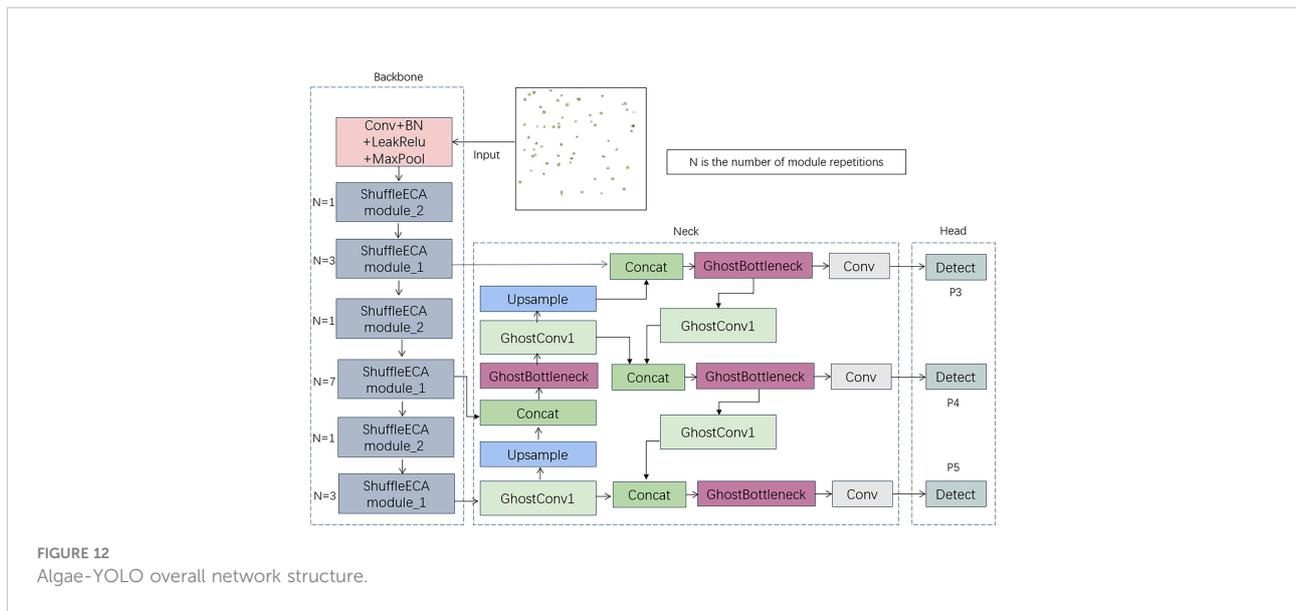


FIGURE 12 Algae-YOLO overall network structure.

$$P = \frac{TP}{TP + FP} \times 100\% \tag{5}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

where TP is the number of algal targets accurately recognized, FP represents the number of erroneous detections, and FN represents the number of targets that went unnoticed. The mAP is calculated using Equations (7) and (8), where n is the number of algal categories. The mAP is derived from P and R after computing the AP (average precision) of each category.

$$AP = \int_0^1 P(R)dR \tag{7}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \tag{8}$$

### Data comparison

Multiple object detection models are selected for comparison to confirm the benefits of the approaches described in this study. MobileNetV3 and GhostNet are both classic lightweight networks that have been frequently used in lightweight improvements and are good objects for comparison; therefore, this paper implements the backbone network replacement improvement of MobileNetV3 and GhostNet based on YOLOv5s, namely, MobileNet-YOLOv5s and GhostNet-YOLOv5s, respectively. In addition, the traditional approaches YOLOv4 and SSD are also chosen, and then the improved Algae-YOLO is added for training comparison, with all other hyper parameters remaining unaltered throughout the training process.

In this paper, the approach runs on Windows 10, the CPU is an Intel i5-12600KF, the GPU is GTX3060, the training framework is PyTorch version 1.7.1, the BatchSize is 16, and each approach is trained with 150 epochs. The Binary Cross-Entropy loss function is used to calculate the loss of category probabilities and target confidence scores during the training process, and CIOU Loss is used as the loss of the bounding box regression. After training completed, the data-enhanced test set is used to test and assess each model in turn. After training, each model is tested and assessed using the data-enhanced test set, and then the key evaluation parameters for each model are created, as shown in Table 2.

In Table 2, the model size is displayed. The mAP0.5 is the mAP size when the threshold is set to 0.5, whereas mAP0.5:0.95 is the average of all threshold values between 0.5 and 0.95. GPU Memory specifies the size of the video memory necessary to run the model. The deployment cost of a graphics card is proportional to the video memory size. Weigh shows the size of the model-generated parameter file. FLOPs represent the number of floating point operations per second, while the lower the value is, the less computations are performed per second and the fewer resources are used. FPS (frames per second) is the detection speed of the model. In this paper, we investigate FPS on CPU devices.

In Table 2, the indices of SSD and YOLOv4 are not dominant, and the sizes of MobileNet-YOLOv5s and GhostNet-YOLOv5s, which are improved and compared, have decreased by approximately half compared with the original YOLOv5s. However, the size in the improved Algae-YOLO approach in this paper has decreased by 82.3% compared with YOLOv5s, which is the lowest among all approaches. In terms of GPU Memory utilization during training runs, Algae-YOLO is only 0.37 G higher than

TABLE 2 Comparison of the parameters of the improved approach with other approaches.

Network	Parameters (M)	mAP 0.5	mAP 0.5:0.95	GPU Memory	Weight	FLOPs	FPS On CPU
YOLO5s	7.01	0.988	0.859	6.73G	14.3MB	16.0G	14.49
MobileNetYOLOv5s	3.54	0.982	0.799	4.83G	7.4MB	6.3G	20
GhostNetYOLOv5s	3.67	0.986	0.84	3.55G	7.7MB	8.0G	15.15
YOLOv4	64.36	0.955	0.738	9.6G	244MB	30.17G	5.64
SSD	26.15	0.856	0.631	6.6G	91.1MB	31.39G	3.92
Algae_ YOLO	<b>1.23</b>	0.981	0.79	3.92G	<b>2.8MB</b>	<b>2.9G</b>	19.61

The bold values indicate that some performance of AlgaeYOLO is better.

TABLE 3 Datasheet of ablation experiment results.

Network	Parameters (M)	mAP 0.5	mAP 0.5:0.95	GPUMemory	Weight	FLOPs	FPS On CPU
Model 0	7,01	0.988	0.859	6.73G	14.3 MB	16.0 G	14.49
Model 1	0.84	0.970	0.743	2.43G	1.9 MB	1.8 G	29.41
Model 2	0.84	0.977	0.768	2.55G	2.0MB	1.8 G	27.77
Model 3	0.49	0.975	0.756	2.43G	1.3 MB	1.2 G	27.02
Model 4	3.00	0.982	0.80	4.22G	6.3MB	5.2G	19.80
Model 5	1,23	0.981	0.79	3.92G	2.8 MB	2.9 G	19.61

GhostNet-YOLOv5s and 2.81 G lower than YOLOv5s; hence, the dependence of Algae-YOLO on memory capacity is significantly smaller than that of the original YOLOv5s. On FLOPs, the improved approach Algae-YOLO is just 2.9G, making it the most efficient of all models and requiring the fewest calculations per second. Finally, in terms of FPS metrics, Algae-YOLO performs second best.

## Ablation study

Ablation experiments are required to demonstrate the efficacy of the approach improvements. The original version of YOLOv5s is labeled as Model 0, YOLOv5s with ShuffleNetV2 as the backbone network labeled as Model 1, ShuffleNetV2-ECA-YOLOv5s with improved ECA attention labeled as Model 2, replaced the neck network on top of ShuffleNetV2-ECA labeled as Model 3, adjusted the depth and width of the model directly on top of ShuffleNetV2-ECA labeled as Model 4, and adjusted the network depth and feature layer channel width on top of Model 3 labeled as Model 5. Table 3 reveals the results of the ablation trials.

In Table 3, after replacing the backbone network with ShuffleNetV2 in Model 1, the model size of the network decreased significantly, from seven million to eighty-four million, and after adding the ECA attention mechanism, the sizes increased by only 30 from Model 1 to Model 2, which is almost negligible, but its mAP0.5 increased by 0.007 and the

mAP0.5:0.95 increased by 0.025. It is demonstrated that ECA attention has a positive effect on increasing the network's precision. Model 3 employs a lighter neck structure based on Model 2, and the FLOPs are lowered by an additional 0.6 G; however, it also results in a 0.002 accuracy loss. The network model is light, but its accuracy is too low, which impairs the detection results. The mAP value of Model 5 (Algae-YOLO) is 0.981, which is closer to Model 0 and has 2.9 GFLOPs, which means that the number of operations per second is much smaller than that of Model 0, i.e., with less computational overhead than Model 0 and Model 5 has a closer operational accuracy, indicating that the model proposed in this paper has good performance in terms of computation. The Algae-YOLO approach investigated is proven to have an optimal balance between computational overhead and precision. Based on the comparison between Model 4 and Model 5, it can be deduced that an increase in network depth and breadth as a result of improvements to ShuffleNet and ECA will result in a considerable increase in size of network parameters, as well as FLOPs, memory utilization, and weight file size. However, the mAP of Model 4 is enhanced by 0.1 compared to Model 5. Because of the dramatic increase in parameters, carrying out such a procedure is completely unnecessary. It also illustrates the need for making the neck section lightweight before expanding the network depth and channel width. In the model improvement, the FPS of Model 5 (AlgaeYOLO) decreased compared to Model 1, but it was still higher than that of Model 0 (YOLOv5s).

## Conclusion

In this research, to address the current algae detection problems, we propose a lightweight Algae-YOLO object detection network employing low-magnification microscopy images. The approach exhibits comparable detection accuracy while drastically decreasing computational overhead, reduces the computational cost, and serves as a benchmark for algal lightweight detection and deployment. The improved Algae-YOLO object detection approach reduces the detection accuracy in low-magnification microscopic images by 0.007 compared to the original YOLOv5s, assuming that parameter size is decreased by 82.3%, the GPU memory footprint is reduced by 2.81G, and the FLOPs are decreased by 13.1 GFLOPs. Ablation experiments demonstrate that the introduction of ECA attention considerably increases the network accuracy while essentially little increase in network size and improves detection accuracy. Some parameters increased due to the modified model depth and channel width, but they are significantly smaller than size of YOLOv5, and its mAP also increased. Overall, the improved model provides a tradeoff between parameters size and detection accuracy and can achieve good performance when deployed in low-cost devices, which is advantageous for mobile and embedded system platform deployment in aquatic work, as well as for reducing the incidence of harmful algae in aquaculture.

There is a wide variety of algal species, and the current study does not cover a broad enough sample, requiring further expansion of the training data. In future research, the species of single-celled algae might be enlarged even more, which will help improve the classification of different phyla and orders of algae.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## References

- Ali, S., Khan, Z., Hussain, A., Athar, A., and Kim, H.-C. (2022). Computer vision based deep learning approach for the detection and classification of algae species using microscopic images. *Water* 14 (14), 2219. doi: 10.3390/w14142219
- Babu, M. J., Geetha, P., and Soman, K. P. (2016). Classification of remotely sensed algae blooms along the coast of india using support vector machines and regularized least squares. *Ind. J. Sci. Technol.* 9, 30. doi: 10.17485/ijst/2016/v9i30/99001
- Cao, M., Wang, J., Chen, Y., and Wang, Y. (2021). Detection of microalgae objects based on the improved YOLOv3 model. *Environ. Sci.: Proc. Impacts* 23 (10), 1516–1530. doi: 10.1039/D1EM00159K
- Chen, T., Wang, N., Wang, R., Zhao, H., and Zhang, G. (2021). One-stage CNN detector-based benthonic organisms detection with limited training dataset. *Neural Networks* 144, 247–259. doi: 10.1016/j.neunet.2021.08.014
- Cho, S., Shin, S., Sim, J., and Lee, J. (2021). Development of microfluidic green algae cell counter based on deep learning. *J. Korean Soc. Visualization* 19 (2), 41–47. doi: 10.5407/jksv.2021.19.2.041
- Deglint, J. L., Jin, C., and Wong, A. (2019). Investigating the automatic classification of algae using the spectral and morphological characteristics via deep residual learning. *International Conference on Image Analysis and Recognition*. 269–280. (Cham: Springer).
- Göröcs, Z., Tamamitsu, M., Bianco, V., Wolf, P., Roy, S., Shindo, K., et al. (2018). A deep learning-enabled portable imaging flow cytometer for cost-effective, high-throughput, and label-free analysis of natural water samples. *Light: Sci. Appl.* 7 (1), 1–12. doi: 10.1038/s41377-018-0067-0
- Girshick, R. (2015). Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*. 1440–1448. (Washington, DC, USA: IEEE).

## Author contributions

DL, PW, YC conceived and designed the research. PW conducted the experiment. HB and YC analysed the data. DL and PW wrote the paper. All authors contributed to the article and approved the submitted version.

## Funding

Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education (202213); Dalian Program on Key Science and Technology R&D Project (2021YF16SN013); Basic Scientific Research Projects for Higher Education Institutions of the Education Department of Liaoning Province.

## Conflict of interest

Author HB is employed by Hangzhou Yunxi Smart Vision Technology Co., LTD.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor ZY declared a shared parent affiliation with the author CY at the time of review.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 580–587. (Washington, DC, USA: IEEE).
- Hayashi, K., Kato, S., and Matsunaga, S. (2018). Convolutional neural network-based automatic classification for algae morphogenesis. *Cytologia* 83 (3), 301–305. doi: 10.1508/cytologia.83.301
- Medina, E., Petraglia, M. R., Gomes, J. G. R., and Petraglia, A. (2017). Comparison of CNN and MLP classifiers for algae detection in underwater pipelines. In *2017 seventh international conference on image processing theory, tools and applications (IPTA)*. 1–6 (Washington, DC, USA: IEEE).
- Mosleh, M. A. A., Manssor, H., Malek, S., Milow, P., and Salleh, A. (2012). A preliminary study on automated freshwater algae recognition and classification system[C]//BMC bioinformatics. *BioMed. Cent.* 13 (17), 1–13. doi: 10.1186/1471-2105-13-S17-S25
- Park, J., Baek, J., Kim, J., You, K., and Kim, K. (2022). Deep learning-based algae detection model development considering field application. *Water* 14 (8), 1275. doi: 10.3390/w14081275
- Park, J., Baek, J., You, K., Nam, S. W., and Kim, J. (2021). Microalgae detection using a deep learning object detection algorithm, YOLOv3. *J. Korean Soc. Water Environ.* 37 (4), 275–285. doi: 10.15681/KSWE.2021.37.4.275
- Park, J., Lee, H., Park, C. Y., Hasan, S., Heo, T. Y., and Lee, W. H. (2019). Algae morphological identification in watersheds for drinking water supply using neural architecture search for convolutional neural network. *Water* 11 (7), 1338. doi: 10.3390/w11071338
- Qian, P., Zhao, Z., Liu, H., Wang, Y., Peng, Y., Hu, S., et al. (2020). Multi-target deep learning for algae detection and classification. In *2020 42nd annual international conference of the IEEE engineering in medicine & biology society (EMBC)*. 1954–1957 (Washington, DC, USA: IEEE).
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 779–788 (Las Vegas, NV, USA: IEEE).
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Adv. Neural Inf. Process. Syst.*, 28, 91–99. Available at: <https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf> (accessed on 2 May 2022).
- Ruiz-Santaquiteria, J., Bueno, G., Deniz, O., et al. (2020). Semantic versus instance segmentation in microscopic algae detection. *Eng. Appl. Artif. Intell.* 87, 103271. doi: 10.1016/j.engappai.2019.103271
- Samantaray, A., Yang, B., Dietz, J. E., and Min, B. C. (2018). *Algae detection using computer vision and deep learning* (Ithaca, NY, USA: arXiv).
- Schaap, A., Rohrlack, T., and Bellouard, Y. (2012). Optofluidic microdevice for algae classification: a comparison of results from discriminant analysis and neural network pattern recognition. In *Microfluidics, BioMEMS, and Medical Microsystems X. SPIE*, (San Francisco, California, United States: Microfluidics, BioMEMS, and Medical Microsystems X. SPIE) 8251, 31–40.
- Shan, S., Zhang, W., Wang, X., and Tong, M. (2020). Automated red tide algae recognition by the color microscopic image. in *2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. 852–861 (Washington, DC, USA: IEEE).
- Tao, J., Cheng, W., Boliang, W., Jiezheng, X., Nianzhi, J., and Tingwei, L. (2010). Real-time red tide algae recognition using SVM and SVDD. In *2010 IEEE international conference on intelligent computing and intelligent systems*, 1, 602–606 (Washington, DC, USA: IEEE).
- Wang, N., Wang, Y., and Er, M. J. (2022). Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Eng. Pract.* 118, 104458. doi: 10.1016/j.conengprac.2020.104458
- Xu, Y., Cheng, C., Zhang, Y., and Zhang, D. (2014). Identification of algae blooms based on support vector machine classification in haizhou bay, East China Sea. *Environ. Earth Sci.* 71 (1), 475–482. doi: 10.1007/s12665-013-2455-3