Check for updates

Vision meets algae: A novel way for microalgae recognition and health monitor

Shizheng Zhou[1], Juntao Jiang[2], Xiaohan Hong[3], Pengcheng Fu[1] and Hong Yan[1]*

[1]State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou, China, [2]College of Control Science and Engineering, Zhejiang University, Hangzhou, China, [3]College of Engineering and Applied Sciences, The State University of New York at Stony Brook, Stony Brook, NY, United States

Marine microalgae are widespread in the ocean and play a crucial role in the ecosystem. Automatic identification and location of marine microalgae in microscopy images would help establish marine ecological environment monitoring and water quality evaluation system. We proposed a new dataset for the detection of marine microalgae and a range of detection methods, the dataset including images of different genus of algae and the same genus in different states. We set the number of unbalanced classes in the data set and added images of mixed water samples in the test set to simulate the actual situation in the field. Then we trained, validated and tested the, TOOD, YOLOv5, YOLOv8 and variants of RCNN algorithms on this dataset. The results showed both one-stage and two-stage object detection models can achieve high mean average precision, which proves the ability of computer vision in multi-object detection of microalgae, and provides basic data and models for real-time detection of microalgal cells.

KEYWORDS

marine microalgae, object detection, data annotation, microscopic imaging, computer vision

1 Introduction

Marine microalgae are widely distributed in the ocean (Lu et al., 2021a) as part of the "blue carbon sink" that uses solar energy and dissolves CO_2 to produce oxygen as well as carbohydrates through photosynthesis (Lu et al., 2021b). They are thus involved in the global ocean–atmosphere carbon cycle to mitigate anthropogenic CO_2 emission, which is the leading cause of escalating climate change. Healthy and viable algal growth is critical to the prosperity of diverse marine ecosystems in the ocean and carbon capture, utilization, and storage. Also, species and quantities of microalgae are widely used as indicators for marine ecological environment monitoring and water quality evaluation worldwide (Gilbert et al., 1992; Domenighini and Giordano, 2009; Gordon and Leggat, 2010; Parmar et al., 2016; Naughton et al., 2020; Peter et al., 2021).

At present, the method of identifying and counting algal species is mainly through manual observation under a microscope, which is time-consuming and laborious and relies on their expertise, and algal samples are very likely to die in the process of collection and transportation, as shown in Figure 1A. The methods to monitor the health of algal cells at the single-cell level include chlorophyll determination, flow cytometry, and molecular biology. However, these methods rely on complex sensors and large instruments, which are not easily popularized in the general laboratory. It is challenging to conduct real-time detection when sampling in the field, as shown in Figure 1B.

In order to monitor and determine the health status of marine microalgae *in situ*, hardware and software integration is needed to enable real-time algal image acquisition, data analysis, and implementation of machine learning algorithms to recognize and classify the cells of marine microalgae. Microfluidic technology is a promising approach to recognizing microalgae and monitoring the health of microalgae at the single-cell level. It can achieve high throughput, has good biocompatibility, provides the ability to integrate with other methods, and requires a small sample size. Combining microfluidic technology with microscopic imaging can avoid the potential damage of labeled cells and enable dynamic cell detection, as shown in Figure 2. The method would produce images of algal cells or videos of algal flowing, just enough to provide data for computer vision methods.

Compared to current manual microscopic identification, which has disadvantages such as high professional level requirements, discontinuity of classifiers, and being time-consuming, automatic marine microalgae identification by using computer vision methods (Coltelli et al., 2014; Promdaen et al., 2014; Giraldo-Zuluaga et al., 2018; Reimann et al., 2020; Barsanti et al., 2021; Cao et al., 2021; Park et al., 2021; Xu et al., 2021) can meet the needs of rapid monitoring and provide convenience for researchers in marine and environmental science. In this analysis of microalgal images, automatic localization and identification are expected to be achieved simultaneously, facilitating the downstream cell analysis. As the joint tasks of classification and localization, object detection can provide the basis for algae identification based on image information combined with biomorphological features.

Object detection is the task of precisely estimating the concepts and locations of objects in each image (Felzenszwalb et al., 2010; Zhao et al., 2019). Traditional object detectors (Viola and Jones, 2001a; Viola and Jones, 2001b; Dalal and Triggs, 2005; Felzenszwalb et al., 2008) based on sophisticated handcrafted features used to be mainstream methods before the popularity of deep learning. Thanks to the capability of learning robust and high-level feature representations of images, convolutional neural networks (CNNs) have been widely used in object detection. Two-stage detectors (Girshick, 2015; Girshick et al., 2015; He et al., 2015; Ren et al., 2017; Cai and Vasconcelos, 2018; Sun et al., 2020; Zhang et al., 2020a) achieved object region selection and detection in two steps while one-stage detectors (Liu et al., 2016; Redmon et al., 2016; Lin et al., 2017; Redmon and Farhadi, 2017; Redmon and Farhadi, 2018; Bochkovskiy et al., 2020; Law and Deng, 2020; Feng et al., 2021; Ge et al., 2021; Jocher, 2021) gave the class probabilities and position coordinate values of objects directly. Recently, some transformer-based methods (Carion et al., 2020) were proposed, showing their strong ability in object detection. Commonly used object datasets include Microsoft COCO (Lin et al., 2014) and Pascal VOC (Everingham et al., 2010).

Some existing related works on microalgal microscopy images applied such a deep learning-based method. Cao et al. (2021) used MobileNet (Howard et al., 2017) and SPP (He et al., 2015) to improve YOLOv3 while Park et al. (2021) trained the YOLOv3

FIGURE 1
Detection and identification of microalgal cells. **(A)** Algal samples awaiting identification. Manual identification of algal water samples is time-consuming and laborious. **(B)** Algal water samples are collected in the field. The collected algal samples generally need to be fixed and concentrated, stored in cold storage, and then transported to scientific research institutions.
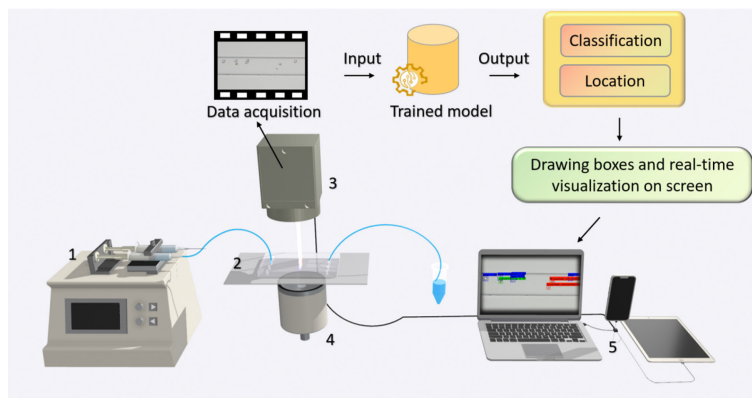
FIGURE 2
Microscopic imaging combined with computer vision and microfluidic technology is used for real-time detection and health monitor of algae. (1) Syringe pump; (2) microfluidic chip; (3) CMOS microscope; (4) light source; (5) computer.

model with Darknet53 backbone (Redmon and Farhadi, 2018). Cao acquired microalgal images using a Nikon Ts2-FL/TS2 fluorescence microscope and then used the data augmentation methods to obtain many images (Cao et al., 2021). The microalgal species in this dataset include *Prorocentrum lima* and *Karenia mikimotoi*. Park et al. (Redmon and Farhadi, 2018) collected 1,114 algal images with 3,663 objects collected by a microscope (Eclipse Ni, NIKON, Japan). These images were split into 10, 20, or 30 genera to compare the model's performance with different numbers of objects for classification.

In this study, we have taken samples of six genera of microalgae commonly found in the ocean (*Pinnularia*, *Chlorella*, *Platymonas*, *Dunaliella salina*, *Isochrysis*, and *Symbiodinium*) and created an image dataset for them. We also collected the images of *Symbiodinium* in different physiological states known as normal, bleaching, and translating, and an image set mixing all microalgal genera described above. We provided many details about this dataset, such as producing process, the original resolution, and the total number of images and objects. Since we will subsequently apply this dataset to the competition, the number of images containing *Pinnularia* was intentionally set lower than the others to examine how well the participants handled the sample imbalance problem in the dataset. In addition, we added an image set of mixed samples to the test set, which was designed better to fit the algal samples in the field environment.

We hope to discover minimal differences between cells from a computer vision perspective and understand life's heterogeneity, randomness, and synergy at the single-cell level. We gave a brief description of the dataset information and the annotation process.

Additionally, we trained some classical or state-of-the-art object detection methods as baselines on the training set and got a test result on the test set. Researchers are welcome to develop their own marine microalgal object detection algorithms on this dataset and compare their results with these baselines. We sincerely hope that this dataset of microscopy images for marine microalga detection will boost related research on marine biology and help establish future real-time monitoring and water quality evaluation system of the marine environment.

2 Materials and methods

This section specifically describes the workflow of building this dataset. The workflow of the dataset production includes data (water samples containing different genera of microalgae) collection, microscopic imaging, image annotation, annotation proofreading, and splitting of the dataset into the training set and the test set, as shown in Figure 3. In addition, we also described the training environment and relevant evaluation metrics of the object detection models.

2.1 Data collection

Microalgal samples collected in this study are common genera in seawater, including *Platymonas*, *Isochrysis*, *Chlorella*, *D. salina*, *Pinnularia*, and *Symbiodinium*, as shown in Figure 4. The microscopic magnification of these cell images is 40×. The algal
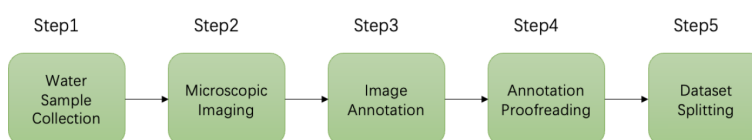

FIGURE 3
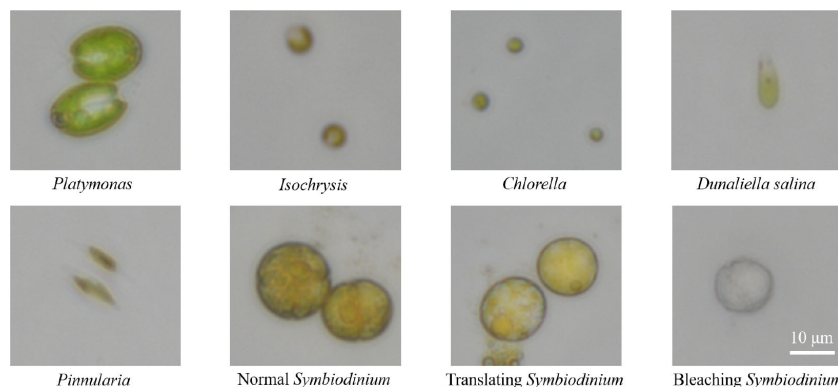The workflow of producing the dataset for marine microalgae detection.

FIGURE 4

The images of six genera of algae under a 40× microscope lens, including three physiological states of *Symbiodinium*. In this figure, all algae images are at a scale of 10 μm.

samples were taken from collected fresh seawater and then cultured in the medium.

- *Platymonas* is a genus of green algae in the family *Chlamydomonadaceae*, whose cell length is approximately 12 μm, and the algal body is flat and compressed. The cells are broadly ovate in front view, with a broad front end and a concave front of the top.
- *Isochrysis*, also known as golden algae, is a kind of algae whose active cells are 4.4–7.1 μm long and 2.7–4.4 μm wide, with a dark red, oval eye spot located in the center of the cell, occasionally near the front of the cell.
- *Chlorella* is a kind of spherical unicellular algae with a diameter of 3–8 μm, and without flagella.
- *D. salina* is a type of halophile green unicellular microalgae that is mainly found in hypersaline environments whose algal body is oval or pear-shaped, 18–28 μm long, and 9.5–14 μm wide. Without a cell wall, the front of the cell is generally concave, and there is a cup-shaped chromophore in the algae.
- *Pinnularia* is a unicellular planktonic diatom with a siliceous cell wall. The cells are spindle-shaped, with an enlarged central part and acuminate at both ends, 12–23 μm long and 2–3 μm wide. Two yellow-brown pigment bodies flank the nucleus in the center of the cell.

The algal genera mentioned above were collected from the Bohai Sea (Shangdong, China), isolated and purified by Fengyun Algae Co., Ltd., and cultured in an f/2 medium.

- *Symbiodinium*, also known as *Zooxanthella*, is a kind of algae establishing intracellular symbioses with organisms such as corals, anemones, jellyfish, nudibranchs, *Ciliophora*, *Foraminifera*, zoanthids, and sponges (Gordon and Leggat, 2010). It enters host cells by phagocytosis, persists as intracellular symbionts, multiplies, and disperses into the environment.

Symbiodinium used to produce this dataset was kindly donated by a scholar and isolated from the Scleractinia coral *Galaxea fascicularis* in West Island (18°14′16′′, 109°21′54′′, Sanya, Hainan, China). There were two culture mediums for *Symbiodinium*. The first culture medium was *standard* f/2 artificial seawater medium and the other in another 1-L shaker's flask was stressed by heat shock at 40°C for approximately 2 h to make the cells bleached.

Observed with a microscope, it was seen that the cells of normal *Symbiodinium* were yellowish brown, while the cells of bleached *Symbiodinium* were white, and the color of the translating cell was in between. There are minor differences in shape, size, and structure between the three groups of *Symbiodinium* cells.

2.2 Image acquisition and annotation

Image acquisition and processing are based on an inverted microscopic imaging platform. In our work, a high-definition camera (DS-FI3, Nikon, Japan) with a viewing field of 7.18×5.32 mm and a 5.9-megapixel complementary metal-oxide-semiconductor (CMOS) image sensor are attached to the inverted microscope. It can capture images with a resolution of up to 2,880×2,048 pixels and transfer them *via* a USB 3.0 port. The exposure time is set to 1 μs. More than 967 images of *Platymonas*, *Isochrysis*, *Chlorella*, *D. salina*, *Pinnularia*, and *Symbiodinium* have been captured individually by using a 40× microscope objective (CFI PlanFluor, NA = 0.45).

The LabelMe software, a commonly used graphical image annotation tool, was used to annotate the total dataset manually. The labeling results of different genera can be seen in Figure 5. After the annotation for the first time, another researcher proofread them using LabelMe. The original labeling results were saved as files in JSON format and were converted into txt files in YOLO annotation format for the convenience of implementing object detection algorithms. The original images were in TIFF format and were converted into PNG format.
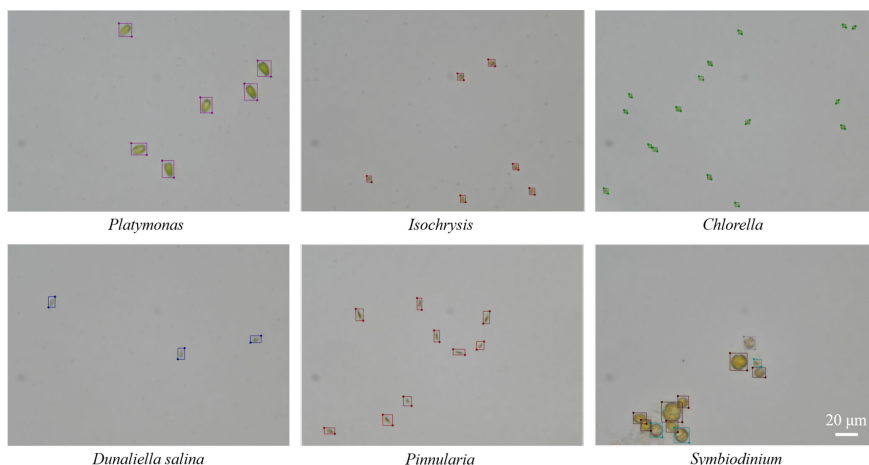
FIGURE 5
Examples of the labeling results for microalgal objects by using LabelMe software in the training set.

2.3 Dataset splitting

The total number of images in the dataset is 967 and all the objects in these images were annotated. The total number of annotated things is 3,915. The annotated dataset contains 150 microscopic images of each of the five genera of algae (*Platymonas*, *Isochrysis*, *Chlorella*, *D. salina*, and *Symbiodinium*), 40 images of *Pinnularia*, and 197 images of mixed samples. We divided the dataset, which included only a single class of objects per image, into a training set and a test set at a ratio of 0.7. We supplemented the test set with an image set that included multiple classes of objects per image. The total training set contained 537 images and the test set contained 430 images with an image resolution of 2,880 × 2,048, as shown in Figure 6A. The total number of microalgae objects in the training set and the test set was 2,085 and 1,830, respectively. The exact number of objects in different classes in the training and test sets is shown in Figure 6B.

Samples collected in the field often contain multiple types of microalgae; thus, we made all the mixed sample images in the test set more consistent with the actual detection conditions. In addition, the size of our test set is close to that of the training set. This is because image acquisition and labeling are time-consuming, and we hope that by training a deep learning model that performs well on a relatively small training set, we can significantly reduce the time and labor costs for researchers.

We make this dataset public with the following access address: https://github.com/Heyimace/Marine-MicroalgaeDetection-Dataset. The previous version of the dataset was used in the competition for IEEE UV2022 at https://tianchi.aliyun.com/competition/entrance/532036/introduction. The license of this dataset is GPL-3.0.

2.4 Implementation details

We trained and tested our dataset's classical and state-of-the-art object detection models. We hope that these experiments can also offer researchers some hints for continuing research. All experiments were done on NVIDIA GeForce RTX 3090 platforms with 24 GB RAM. The deep learning framework used is PyTorch (Paszke et al., 2017; Paszke et al., 2019). The training set in our dataset was split into two sets for training and validation. The new

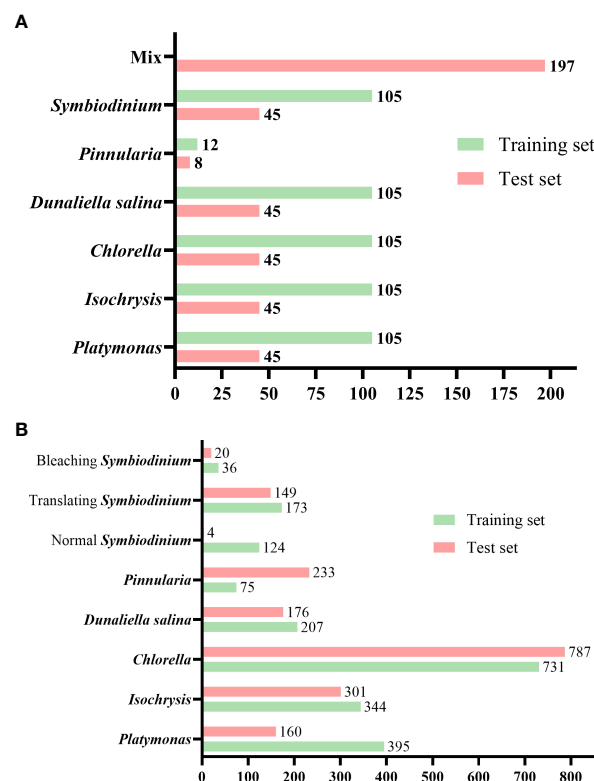

FIGURE 6
Number of images and labeled objects. **(A)** Number of images of various classes of algal cells in the test and training sets. **(B)** Number of annotated objects in the test set and training set for various classes of algal cells.

training set contains 430 images, and the validation set contains 107 images.

For Faster RCNN (Ren et al., 2017), Casacade R-CNN (Cai and Vasconcelos, 2018), Dynamic RCNN (Zhang et al., 2020b), and TOOD (Feng et al., 2021) models, we trained, validated, and tested models based on the mmdetection toolbox (Chen et al., 2019). We only used Random Flip with 0.5 probability for augmentation in the training process and Multi-Scale Flipping and Random Flipping for augmentation in the testing process. The sizes of input images in training, validation, and testing are 800×800. The stochastic gradient descent (SGD) optimizer was used in experiments. The learning rate was 0.01, the momentum was 0.9, and the weight decay is 0.0005. The total training epochs are 100.

For YOLOv5 and YOLOv8 models, we trained, validated, and tested models based on Ultralytics's framework (Jocher et al., 2023). The total training epochs are 100 and the training batch size is 16. The maximum detection in validation and testing is 300. The initial learning rate is 0.01, the final OneCycleLR learning rate is 0.0001, the momentum is 0.937, and mosaic augmentation is used. The fraction of image HSV-Hue augmentation is 0.015, the fraction of image HSV-Saturation augmentation is 0.7, and the fraction of image HSV-Value augmentation is 0.4. The fractions for image translation and scaling are 0.1 and 0.5, respectively. The probability of an image flipping horizontally is 0.5.

2.5 Evaluation metrics

Intersection over union (IoU) is a common metric in object detection. It refers to the intersection ratio between the bounding box predicted by the model and the ground truth, which is defined as:

$$IoU(A, B) = \frac{A \cap B}{A \cup B} \qquad (1)$$

where A and B represent the area of the bounding box and ground truth predicted by the model, respectively, and the detection is considered correct when the IoU is more significant than a certain threshold.

TP means True Positive prediction, TN means True Negative prediction, FP means False Positive prediction, and FN means False Negative prediction. Then, the performance of the model can be evaluated by Precision (P), Recall (R), and Average Precision (AP), as in:

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

AP is the area under the precision–recall curve. The mean average precision (mAP) is calculated by finding the AP for each class and then the average over all classes:

$$mAP = \frac{1}{N}\sum_{i=1}^{N}AP_i \qquad (4)$$

The evaluation metric in validation is mAP(0.50:0.95) to select the best model, and we gave the mAP(0.50:0.95) as well as mAP (0.50) results on the test set. The mAP(0.50:0.95) represents the average mAP over different IoU thresholds (from 0.5 to 0.95 in steps of 0.05) and the mAP(0.50) represents the average mAP over 0.5.

3 Results

Object detection is a computer vision task that involves locating and identifying objects in an image or video. Two main object detection methods exist: single- and two-stage object detection algorithms. One-stage object detection methods perform object classification and bounding-box regression directly without using pre-generated region proposals. Two-stage object detection methods first generate region proposals using a separate network or algorithm, such as selective search or region proposal network (RPN), and then classify and refine each proposal using another network. In this study, we trained one-stage object detection algorithms including YOLOv5, YOLOv8, and TOOD, and two-stage object detection algorithms including Faster-RCNN, Cascade-RCNN, and Dynamic-RCNN.

The detection results based on mmdetection on the test set are shown in Table 1. The evaluation metrics of the one-stage object detection algorithms are higher than the two-stage object detection algorithms. The indices of mAP (IoU = 0.5) of the Faster-RCNN series are only approximately 0.692 to 0.756. The Cascade-RCNN series, which uses cascade regression as a resampling mechanism to increase the IoU value of the proposal stage by stage so that the proposals resampled by the previous stage can adapt to the next stage with a higher threshold, have a higher detection accuracy. However, Cascade-RCNN with a larger backbone has a lower

TABLE 1 Results of different object detection methods on this dataset.

Method	Backbone	mAP@0.5	mAP@0.5:0.95
Faster-RCNN (Ren et al., 2017)	ResNet50 (He et al., 2016)	0.692	0.341
Faster-RCNN	ResNet101 (He et al., 2016)	0.756	0.407
Cascade-RCNN (Cai and Vasconcelos, 2018)	ResNet50	0.812	0.484
Cascade-RCNN	ResNet101	0.797	0.481
Dynamic-RCNN (Zhang et al., 2020b)	ResNet50	0.804	0.478
TOOD (Feng et al., 2021)	ResNet50	0.842	0.539
TOOD	ResNet101	0.950	0.571

detection accuracy than the one with a smaller backbone. This may be because the larger model causes overfitting more easily. Dynamic-RCNN, which continuously adaptively increases the positive sample threshold and adaptively modifies the SmoothL1 Loss parameter, also achieves better results than Faster-RCNN. TOOD, a one-stage detection method that uses Task-aligned head and Task Alignment Learning to solve the problem of classification and positioning misalignment, achieves better results than two-stage methods.

We also trained YOLOv5 and YOLOv8 models, two of the most advanced one-stage object detection algorithms, and explored the detection performance of models with different model sizes. The detection results of YOLOv5 and YOLOv8 on the test set are shown in Table 2, and the visualizations of detection results on the test set are shown in Figure 7. The detection results of YOLOv8m for different classes is shown in Table 3 and Figure 8. The mean mAP (IoU = 0.5:0.95) of YOLO series is higher than the other algorithms, which indicates that its prediction boxes can locate and match objects more precisely at progressively higher IoU thresholds. YOLOv8 replaces the C3 structure of YOLOv5 with the C2f structure and adjusts the number of channels for different scale models. It separates the classification and detection heads, and changes from Anchor-Based to Anchor-Free, using the TaskAlignedAssigner positive sample distribution strategy and introducing Distribution Focal Loss. Compared to YOLOv5,

YOLOv8 models lead to some improvement in our dataset. We can also find that larger models do not perform better than smaller models. In addition to the characteristics of the dataset image itself, this may also be related to the small amount of data.

4 Conclusion

In this paper, we apply algorithms in computer vision to the multiple object detection and physiological state assessment of marine microalgae. We introduced a new dataset for marine microalgae detection in microscopy images and the building process, including data collection, microscopic imaging, image annotation, proofreading, and splitting. This dataset contains microalgal images of six genera, namely, *Pinnularia*, *Chlorella*, *Platymonas*, *D. salina*, *Isochrysis*, and *Symbiodinium*. Moreover, we also collected the images of *Symbiodinium* in different physiological states, namely, normal, bleaching, and translating, which can be an indicator of the situation of coral and its water environment. The dataset contains 967 images, among which 537 are in the training set and 430 are in the test set.

We trained and tested a number of classical or state-of-the-art models in our dataset, including one-stage object detection algorithms and two-stage object detection algorithms. Among these algorithms, YOLOv5 and YOLOv8 were able to perform the

TABLE 2 Results of YOLOv5 and YOLOv8 models on this dataset.

Method	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5s (Jocher, 2020)	0.839	0.754	0.832	0.569
YOLOv5m (Jocher, 2020)	0.763	0.775	0.785	0.533
YOLOv8s (Jocher et al., 2023)	0.825	0.758	0.833	0.579
YOLOv8s (Jocher et al., 2023)	0.792	0.764	0.826	0.571

FIGURE 7
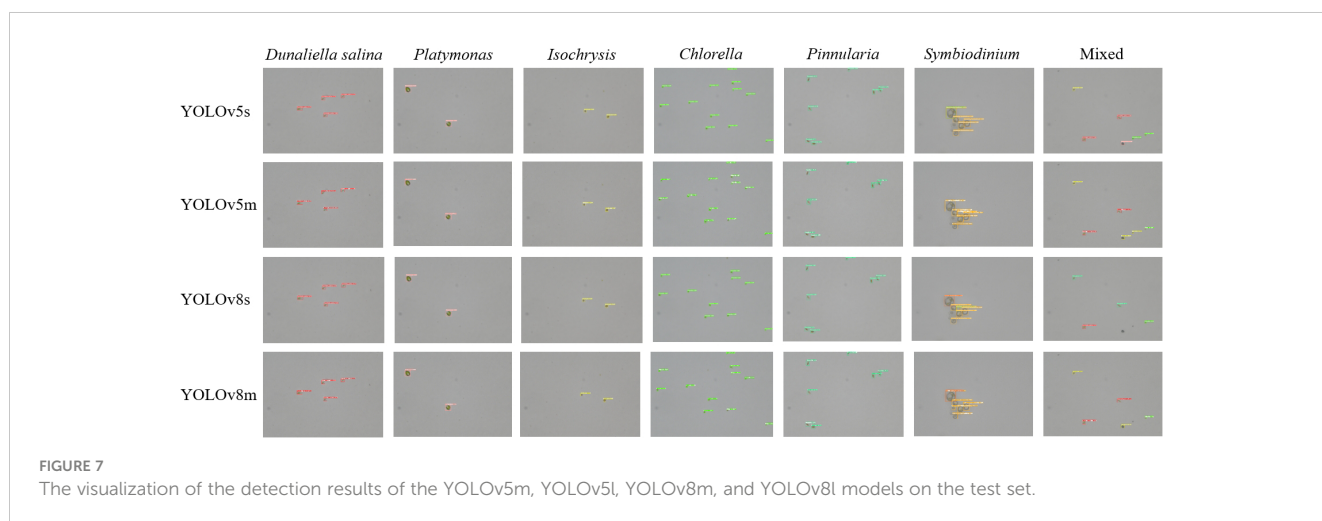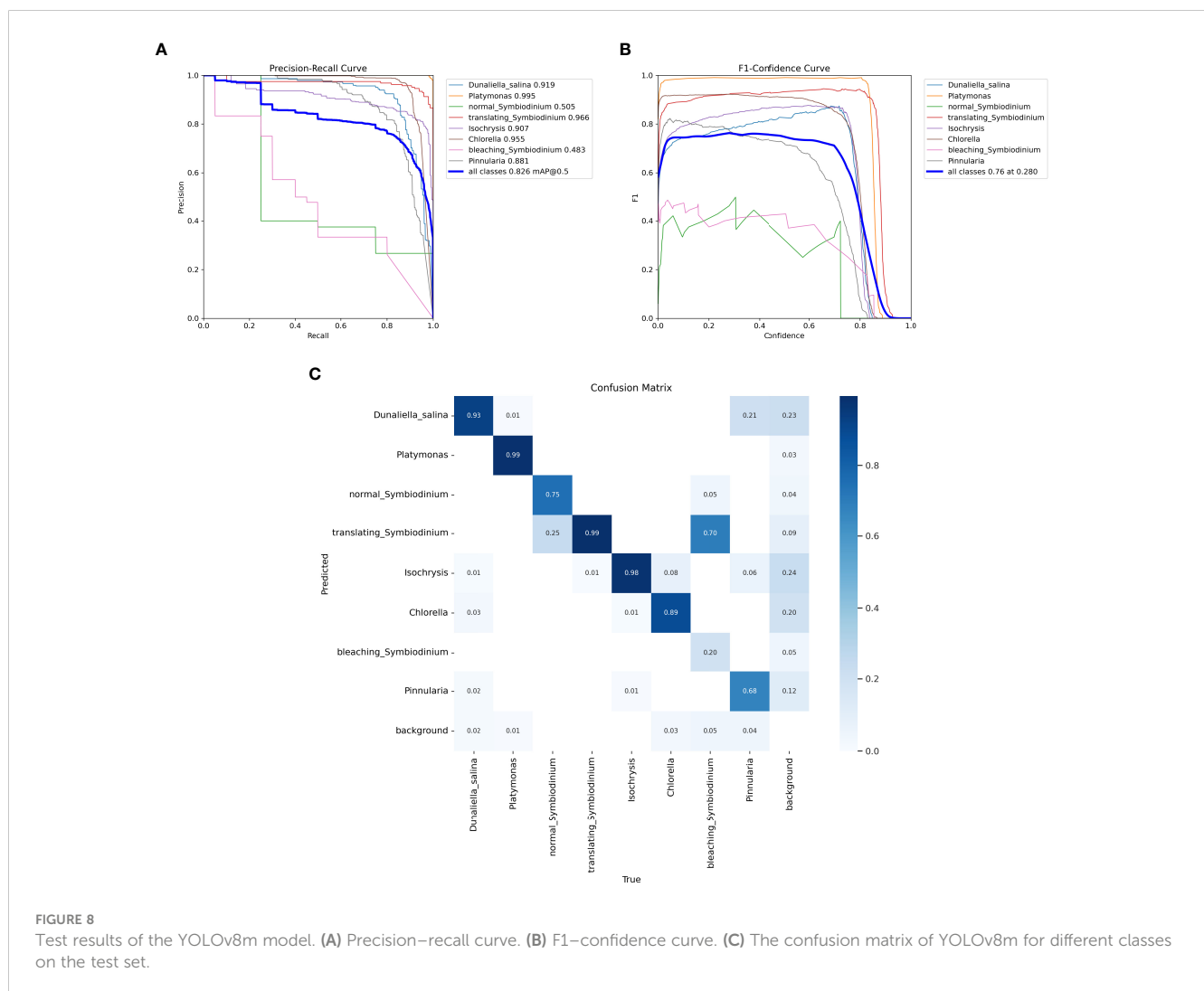The visualization of the detection results of the YOLOv5m, YOLOv5l, YOLOv8m, and YOLOv8l models on the test set.

TABLE 3 Detection results of YOLOv8m for different classes on this dataset.

Class	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Dunaliella salina	0.729	0.903	0.919	0.541
Platymonas	0.981	0.995	0.995	0.820
Isochrysis	0.766	0.977	0.907	0.596
Chlorella	0.970	0.863	0.955	0.609
Pinnularia	0.940	0.605	0.881	0.497
Normal *Symbiodinium*	0.388	0.481	0.505	0.408
Translating *Symbiodinium*	0.865	0.989	0.966	0.778
Bleaching *Symbiodinium*	0.697	0.300	0.483	0.328

FIGURE 8

Test results of the YOLOv8m model. **(A)** Precision–recall curve. **(B)** F1–confidence curve. **(C)** The confusion matrix of YOLOv8m for different classes on the test set.

task with high accuracy for the classification and localization of microalgal cells. The main challenges of this task can be concluded as follows:

- High-resolution images and tiny objects
- Small training set size that is close to the test set

- Multiple species in some images in the test set
- Sample imbalance in the training set

For future research, we will make a larger dataset as the next version and continue to develop stronger baselines for the microalgal detection tasks.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

SZ: Conceptualization, Methodology, Resources, Visualization, and Writing—Original Draft; JJ: Investigation, Software, and Writing—Original Draft; XH: Data Curation; PF: Project Administration and Funding Acquisition; HY: Writing—Review and Editing, Supervision, and Funding Acquisition. All authors contributed to the article and approved the submitted version.

Conflict of interest

The 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.

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.

References

Barsanti, L., Birindelli, L., and Gualtieri, P. (2021). Water monitoring by means of digital microscopy identification and classification of microalgae. *Environ. Sci.: Process. Impacts.* 23, 1443–1457. doi: 10.1039/d1em00258a

Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv* doi. doi: 10.48550/arXiv.2004.10934

Cai, Z., and Vasconcelos, N. (2018). "Cascade r-cnn: Delving into high quality object detection," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA. pp. 6154–6162. doi: 10.1109/CVPR.2018.00644

Cao, M., Wang, J., Chen, Y., and Wang, Y. (2021). Detection of microalgae objects based on the improved yolov3 model. *Environ. Sci.: Process. Impacts.* 23, 1516–1530. doi: 10.1039/D1EM00159K

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S. (2020). End-to-end object detection with transformers. arXiv. 213–229. doi: 10.48550/arXiv.2005.12872

Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., et al. (2019). Mmdetection: Open mmlab detection toolbox and benchmark. *arXiv.* doi: 10.48550/arXiv.1906.07155

Coltelli, P., Barsanti, L., Evangelista, V., Frassanito, A. M., and Gualtieri, P. (2014). Water monitoring: automated and real time identification and classification of algae using digital microscopy. *Environ. Sci.: Process. Impacts.* 16, 2656–2665. doi: 10.1039/C4EM00451E

Dalal, N., and Triggs, B. (2005). "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05).* San Diego, CA, USA. pp. 886–893. doi: 10.1109/CVPR.2005.177

Domenighini, A., and Giordano, M. (2009). Fourier Transform infrared spectroscopy of microalgae as a novel tool for biodiversity studies, species identification, and the assessment of water quality. *J. phycol.* 45, 522–531. doi: 10.1111/j.1529-8817.2009.00662.x

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *Int. J. Comput. vision.* 88, 303–338. doi: 10.1007/11736790_8

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D. (2010). Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Anal. Mach. Intelligence.* 32 (9), 1627–1645. doi: 10.1109/TPAMI.2009.167

Felzenszwalb, P., McAllester, D., and Ramanan, D. (2008). "A discriminatively trained, multiscale, deformable part model," in *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1–8. doi: 10.1109/CVPR.2008.4587597

Feng, C., Zhong, Y., Gao, Y., Scott, M. R., and Huang, W. (2021). "Tood: Task-aligned one-stage object detection," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV).* Montreal, QC, Canada. pp. 3490–3499. doi: 10.1109/ICCV48922.2021.00349

Ge, Z., Liu, S., Wang, F., Li, Z., and Sun, J. (2021). Yolox: Exceeding yolo series in 2021. *arXiv.* doi: 10.48550/arXiv.2107.08430

Gilbert, F., Galgani, F., and Cadiou, Y. (1992). Rapid assessment of metabolic activity in marine microalgae: application in ecotoxicological tests and evaluation of water quality. *Mar. Biol.* 112, 199–205. doi: 10.1007/BF00702462

Giraldo-Zuluaga, J.-H., Salazar, A., Diez, G., Gomez, A., Martínez, T., Vargas, J. F., et al. (2018). Automatic identification of scenedesmus polymorphic microalgae from microscopic images. *Pattern Anal. Applications.* 21, 601–612. doi: 10.1007/s10044-017-0662-3

Girshick, R. (2015). "Fast r-cnn," in *2015 IEEE International Conference on Computer Vision.* 1440–1448 (IEEE). doi: 10.1109/ICCV.2015.169

Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2015). Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intelligence.* 38 (1), 142–158. doi: 10.1109/TPAMI.2015.2437384

Gordon, B. R., and Leggat, W. (2010). *Symbiodinium*–invertebrate symbioses and the role of metabolomics. *Mar. Drugs* 8, 2546–2568. doi: 10.3390/md8102546

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. intelligence.* 37, 1904–1916. doi: 10.1109/TPAMI.2015.2389824

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* Las Vegas, NV, USA. pp. 770–778. doi: 10.1109/CVPR.2016.90

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv.* doi: 10.48550/arXiv.1704

Jocher, G. (2020). YOLOv5 by ultralytics. doi: 10.5281/zenodo.3908559

Jocher, G. (2021) *yolov5. github repository.* Available at: https://github.com/ultralytics/yolov5.

Jocher, G., Chaurasia, A., and Qiu, J. (2023). *YOLO by ultralytics.*

Law, H., and Deng, J. (2020). Cornernet: Detecting objects as paired keypoints. *Int. J. Comput. Vision.* 128, 642–656. doi: 10.1007/s11263-019-01204-1

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). "Feature pyramid networks for object detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. pp. 2117–2125. doi: 10.1109/CVPR.2017.106

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). "Microsoft Coco: Commonobjects in context," in *Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science*. Eds. D. Fleet, T. Pajdla, B. Schiele and T. Tuytelaars. (Cham: Springer) 8693. doi: 10.1007/978-3-319-10602-1_48

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*. Eds. B. Leibe, J. Matas, N. Sebe and M. Welling. (Cham: Springer), 9905. doi: 10.1007/978-3-319-46448-0_2

Lu, X., Cui, Y., Chen, Y., Xiao, Y., Song, X., Gao, F., et al. (2021a). Sustainable development of microalgal biotechnology in coastal zone for aquaculture and food. *Sci. Total Environ.* 780, 146369. doi: 10.1016/j.scitotenv.2021.146369

Lu, Y., Gan, Q., Iwai, M., Alboresi, A., Burlacot, A., Dautermann, O., et al. (2021b). Role of an ancient light-harvesting protein of psi in light absorption and photoprotection. *Nat. Commun.* 12, 679. doi: 10.1038/s41467-021-20967-1

Naughton, S., Kavanagh, S., Lynch, M., and Rowan, N. J. (2020). Synchronizing use of sophisticated wet-laboratory and in-field handheld technologies for real-time monitoring of key microalgae, bacteria and physicochemical parameters influencing efficacy of water quality in a freshwater aquaculture recirculation system: A case study from the republic of ireland. *Aquaculture* 526, 735377. doi: 10.1016/j.aquaculture.2020.735377

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, 275–285. doi: 10.15681/KSWE.2021.37.4.275

Parmar, T. K., Rawtani, D., and Agrawal, Y. (2016). Bioindicators: the natural indicator of environmental pollution. *Front. Life science.* 9, 110–118. doi: 10.1080/21553769.2016.1162753

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in pytorch. In *NIPS 2017 Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques*. Long Beach, 2017, 1–4. Available at: https://openreview.net/forum?id=BJJsrmfCZ.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: An imperative style, high-performance deep learning library. *arXiv*. doi: 10.48550/arXiv.1912.01703

Peter, A. P., Khoo, K. S., Chew, K. W., Ling, T. C., Ho, S.-H., Chang, J.-S., et al. (2021). Microalgae for biofuels, wastewater treatment and environmental monitoring. *Environ. Chem. Letters.* 19, 2891–2904. doi: 10.1007/s10311-021-01219-6

Promdaen, S., Wattuya, P., and Sanevas, N. (2014). Automated microalgae image classification. *Proc. Comput. Science.* 29, 1981–1992. doi: 10.1016/j.procs.2014.05.182

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA. pp. 779–788. doi: 10.1109/CVPR.2016.91

Redmon, J., and Farhadi, A. (2017). "Yolo9000: better, faster, stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. pp. 6517–6525. doi: 10.1109/CVPR.2017.690

Redmon, J., and Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv*. doi: 10.48550/arXiv.1804.02767

Reimann, R., Zeng, B., Jakopec, M., Burdukiewicz, M., Petrick, I., Schierack, P., et al. (2020). Classification of dead and living microalgae chlorella vulgaris by bioimage informatics and machine learning. *Algal. Res.* 48, 101908. doi: 10.1016/j.algal.2020.101908

Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi: 10.1109/TPAMI.2016.2577031

Sun, P., Zhang, R., Jiang, Y., Kong, T., Xu, C., Zhan, W., et al. (2020). "Sparse R-CNN: End-to-End Object Detection with Learnable Proposals," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (Nashville, TN, USA), 14449–58. doi: 10.1109/CVPR46437.2021.01422

Viola, P., and Jones, M. (2001a). "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CPVR)*. Kauai, HI, USA. pp. I–I. doi: 10.1109/CVPR.2001.990517

Viola, P., and Jones, M. (2001b). Robust real-time face detection. *Int. J. Comput. Vision.* 2, 747–747. doi: 10.1109/ICCV.2001.937709

Xu, M., Harmon, J., Yuan, D., Yan, S., Lei, C., Hiramatsu, K., et al. (2021). Morphological indicator for directed evolution of euglena gracilis with a high heavy metal removal efficiency. *Environ. Sci. &Technol.* 55, 7880–7889. doi: 10.1021/acs.est.0c05278

Zhang, H., Chang, H., Ma, B., Wang, N., and Chen, X. (2020a). Dynamic r-CNN: Towards high quality object detection *via* dynamic training, *arXiv*. doi: 10.48550/arXiv.2004.06002

Zhang, H., Chang, H., Ma, B., Wang, N., and Chen, X. (2020b). "Dynamic r-cnn: Towards high quality object detection *via* dynamic training," in *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*. Eds. A. Vedaldi, H. Bischof, T. Brox and J. M. Frahm (Cham: Springer) 12360. doi: 10.1007/978-3-030-58555-6_16

Zhao, Z.-Q., Zheng, P., Xu, S.-t., and Wu, X. (2019). Object detection with deep learning: A review. *IEEE Trans. Neural Networks Learn. Syst.* 30, 3212–3232. doi: 10.1109/TNNLS.2018.2876865