Recognition of the gonad of Pacific oysters via object detection

Yifei Chen[1], Jun Yue[1*], Weijun Wang[2], Jianmin Yang[2], Zhenbo Li[3]

(1. *College of Information and Electrical Engineering, Ludong University, Yantai 264025, Shandong, China*;
2. *College of Agricultural Sciences, Ludong University, Yantai 264025, Shandong, China*;
3. *College of Information and Electrical Engineering, China Agricultural University, Beijing 100091, China*)

Abstract: Oyster is the largest cultured shellfish in the world, and it has high economic value. The plumpness of the Pacific oyster gonad has important implications for the quality and breeding of subsequent parents. At present, only the conventional method of breaking their shells allows for the observation and study of the interior tissues of Pacific oysters. It is an important task to use computer technology for non-destructive sex detection of oysters and to select mature and full oysters for breeding. In this study, based on the multi-effect feature fusion network R-SINet algorithm, a CF-Net algorithm was designed through a boundary enhancement algorithm to detect inconspicuous objects that appear to be seamlessly embedded in the surrounding environment in nuclear magnetic resonance (NMR) images, effectively solving the problem of difficulty in distinguishing Pacific oyster gonads from background images. In addition, calculations were performed on the segmented gonadal regions to obtain a grayscale value difference map between male and female oysters. It was found that there were significant differences in grayscale values between females and males. This task allows for non-destructive detection of the gender of oysters. Firstly, a small animal magnetic resonance imaging (MRI) system was used to perform MRI on Pacific oysters, and a dataset of oyster gonads was established. Secondly, a gonadal segmentation model was created, and the Compact Pyramid Refinement Module and Switchable Excitation Model were applied to the R-SINet algorithm model to achieve multi-effect feature fusion. Then, the Convformer encoder, Token Reinforcement Module, and Adjacent Transfer Module were used together to form the CF-Net network algorithm, further improving the segmentation accuracy. The experimental results on the oyster gonad dataset have demonstrated the effectiveness of this method. Based on the segmentation results, it is possible to calculate the grayscale values of the gonadal region and obtain the distribution map of the grayscale value difference between male and female oysters. The results can provide a technical methodology for the non-destructive discrimination of oyster gender and later reproduction.

Keywords: Pacific oyster gonad, unapparent object detection, target segmentation, deep learning, MRI, R-SINet
DOI: 10.25165/j.ijabe.20241706.8478

Citation: Chen Y F, Yue J, Wang W J, Yang J M, Li Z B. Recognition of the gonad of Pacific oysters via object detection. Int J Agric & Biol Eng, 2024; 17(6): 230–237.

1 Introduction

This study aimed to determine the sex of oysters visually in order to ensure the selection of specific-gender oysters with ideal commercial characteristics for breeding purposes. This will enable to continuously improve the commercial characteristics of the oyster population in a non-destructive and sustainable manner[1]. Using the Pacific oyster as an example, we select Pacific oysters with plump gonads for parental breeding. With the rapid development of small animal imaging technology and convolutional neural networks, we can now use a small animal imaging system to obtain Pacific oyster magnetic resonance imaging (MRI) images in which we can clearly and intuitively observe the gonad part of oysters without harm to live Pacific oysters. This can solve the problem of high similarity between organs and tissues and insignificant color differences in MRI images when segmenting the gonads by detecting unapparent objects in MRI images with relatively complex backgrounds, which is important for improving the integrity and accuracy of segmenting the gonads of Pacific oysters.

In the field of agriculture, Xia et al.[2] used spatial resolution spectroscopy to detect the internal information of agricultural products, making certain contributions to the sustainable development of agriculture. Small animal MRI is a branch of magnetic resonance imaging, which has greater magnetic field strength, stronger gradient intensity, and can achieve higher spatial and temporal resolution, allowing MRI to reach the level of molecular imaging. This technique enables in vivo imaging of the target animal to capture the changes in biological information such as tissue structure, function, and metabolism under various conditions, which is becoming an important tool for studying the internal structure of small animals[3]. In 2001, the first clinical study of canine intervertebral disc disease was conducted using MRI technology in China. Small animal MRI technology has been developed in China since 2009, leading to the emergence of powerful and versatile small animal imaging systems. In 2019, Zhang et al.[4] applied small animal MRI systems to study Alzheimer's disease (AD), providing multimodal imaging techniques to help diagnose early AD. In 2022, Hang et al.[5] used 7.0T small animal MRI equipment to noninvasively observe brain injury in a rat model of classic heat stroke. Small animal MRI techniques started earlier abroad. In 1990, Button et al. used 0.35T small animal MRI to observe the growth and morphology of tumors in mice. In 2021,

Received date: 2023-08-10 **Accepted date:** 2024-09-08
Biographies: Yifei Chen, MS, research interest: artificial intelligence technology, Email: chenyifei98@126.com; **Weijun Wang**, PhD, Professor, research interest: shellfish breeding, Email: wwj2530616@163.com; **Jianmin Yang**, PhD, Professor, research interest: shellfish breeding, Email: ladderup@126.com; **Zhenbo Li**, PhD, Professor, research interest: computer graphics, Email: lizb@cau.edu.cn.
***Corresponding author: Jun Yue**, PhD, Professor, research interest: artificial intelligence technology. College of Information and Electrical Engineering, Ludong University, Yantai 264025, Shandong, China. Tel: +86-13562559603, Email: yuejuncn@126.com.

Gilchrist et al.[6] designed a gating unit for synchronized control of the small animal heart and respiration using small animal MRI, while in 2022, Liu et al.[7] conducted a multimodal animal MRI study for memory generalization in mice. In 2024, Baskaya et al.[8] used longitudinal multi-parameter MRI to perform pathophysiological mapping of chronic liver disease in an animal model. The above-mentioned studies show that specialized small animal MRI techniques provide technical support for the study of small animal organ tissues. However, nowadays, while small animal MRI techniques mainly focus on the detection and study of terrestrial animals, there are few studies in the field of marine organisms, especially shellfish organisms.

In recent years, with the iterations and advances in computer vision, unapparent object detection identification techniques have developed rapidly. In 2020, Fan et al.[9] first proposed a camouflaged object detection technique, designed the SINet network architecture aimed at identifying small target objects in complex backgrounds, and constructed a COD10K dataset for their task. In 2021, Lyu et al.[10] proposed a hierarchical localization of target regions, introducing reverse attention[11] to capture more details of the spatial structure, and designed the LSR algorithm model. Zhai et al.[12] built an edge-shrinkage graph inference module to guide the learning of feature representations of camouflaged objects. Fan et al.[13] proposed the SINet_v2 network architecture with optimization improvements based on SINet. In 2022, Jia et al.[14] performed both amplification and repetition operations for camouflage object segmentation to achieve accurate localization of camouflage objects by iteration and target object amplification and proposed the SegMaR algorithm model. Li et al.[15] performed work to identify fish in complex backgrounds. In 2024, Xia et al.[16] used improved YOLOv5 to detect surface defects of maize seeds. However, when these existing detection and segmentation algorithms for unapparent object detection are used to train MRI grayscale images, the local feature extractions in all of these algorithms are coarse and the global feature fusion is ignored. This results in inadequate local feature extractions and loss of global information in the segmented grayscale images, which in turn cause incomplete segmentation targets and unclear boundaries, and the overall level of the evaluation index of segmentation decreases.

In summary, in response to the shortcomings of small animal MRI technology in aquatic shellfish applications and the inability of existing unclear object detection algorithms to accurately segment grayscale images, this study proposed a multi-effect feature fusion network algorithm R-SINet and a boundary refinement segmentation algorithm CF-Net, which are presented progressively. Therefore, the main contributions of this study are five aspects:

1) An unprecedented dataset of oyster gonads was established;

2) A new network algorithm R-SINet was proposed, which can effectively improve the integrity of feature extraction from Pacific oyster grayscale images;

3) A boundary refinement algorithm CF-Net was proposed, which can ensure the integrity of the extracted gonad region while finely segmenting the gonad edge region;

4) A large number of experimental results on the self-built oyster gonad dataset show that CF-Net is more effective than other state-of-the-art methods;

5) The grayscale histogram was used to obtain the grayscale value difference between male and female gonads, laying the foundation for non-destructive testing of the gender of Pacific oysters in the future.

2 Materials and methods

The technology roadmap of this study is shown in Figure 1. Firstly, a dataset of Pacific oyster gonads is established. On this basis, a gonadal region segmentation model is created, which includes two algorithms. The SINet algorithm initially segments the oyster gonadal region completely and then uses the CF-Net algorithm to refine the segmentation of the gonadal edge region through boundary refinement operations. In this algorithm, the focus is on extracting adjacent features and fusing them for reinforcement. Finally, the image is segmented to obtain a grayscale histogram of the oyster gonads.

2.1 Compact Pyramid Refinement Module

To solve the large semantic generation gap that appears after the target object is subsampled, pyramid models are usually used to improve the accuracy of small target detection. The common pyramid models[17,18] at this stage have problems such as large computation, large memory consumption, and slow inference speed. To solve such problems, this study added a lightweight feature pyramid module after effectively fusing neighboring features, deeply fused high-level and low-level features, and proposed a Compact Pyramid Refinement Module, which improves efficiency while ensuring accuracy.

First, we use Res2Net50[19] network for image $I \in R^{W\times H\times 3}$ extracting features f_k ($k \in \{1,2,3,4,5\}$), obtaining five features with a resolution of $f_k = \frac{H}{2^k} + \frac{W}{2^k}$, and then expand the perceptual field by the texture enhancement module (TEM). We only use f_3, f_4, and f_5 as the feature images $f_k^{\mathrm{nc}} = F_{\mathrm{NC}}\left(f_k'; W_{\mathrm{NC}}^u\right)$, $u \in \{1,2,3\}$,

The compact pyramid refinement uses the idea of depth direction separable convolution[20]. The specific equation is shown below:

$$\begin{aligned} f_1 &= \mathrm{Conv}_{1\times 1}(f) \\ f_2^{d_i} &= \mathrm{Conv}_{3\times 3}^{d_i}(f_1),\ \ i = 2,4,8 \\ f_2 &= \mathrm{ReLU}\left(\mathrm{BN}\left(f_2^{d_1} + f_2^{d_2} + f_2^{d_3}\right)\right) \\ f_3 &= \mathrm{Conv}_{1\times 1}(f_2) + f \end{aligned} \tag{1}$$

where, f denotes the input of the feature image; $\mathrm{Conv}_{n\times n}$ denotes the input for the $n\times n$ convolution operation; ReLU denotes the nonlinear activation function; BN denotes the Batch Normalization; and d_n denotes the corresponding expansion rate size.

In this way, the lightweight decoder with feature pyramid refinement proposed in this study can aggregate multi-level features from top to bottom and achieve efficient feature capture at all levels. The structure diagram of the Compact Pyramid Refinement Model is shown in Figure 2.

2.2 Switchable Excitation Model

The predicted target regions are erased by a reverse attention mechanism (GRA) as a way to mine the detailed content in the complementary regions. First, the candidate feature maps are treated $\left\{p_i^k \in R^{H/2^k\times W/2^k\times C}, k = 3,4,5\right\}$, grouped by dimensional channels, and divided into $\{m_i = C/g_i, i = 1,2,3\}$, where g_i denotes the grouping size of processed features. Additionally, the inversion map obtained under the inversion guidance operation r_1^k, regularly inserted into the processed grouped features, can be finally represented as $\left\{\{p_{i,1}^k, r_1^k\}, \ldots, \{p_{i,m_i}^k, r_1^k\}, p_{i,j}^k \in R^{H/2^K\times W/2^K\times g_i}\right\}$. Finally, an iterative refinement operation is performed on multiple combined GRA modules. By grouping inversion attention, more attention is paid to the local feature information of the target edges, but attention based on the global scope is still lacking.

In this study, a switchable excitation module was proposed that

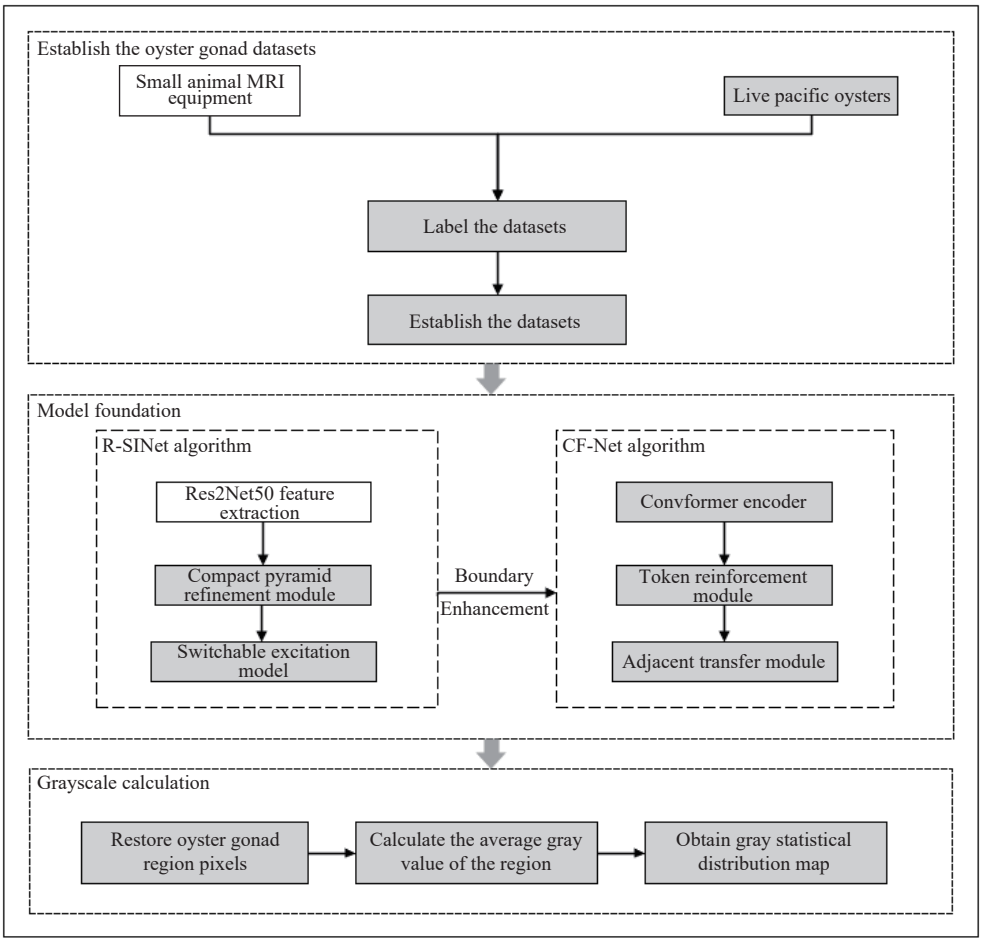

Figure 1 Technology roadmap of Pacific oyster gonad recognition

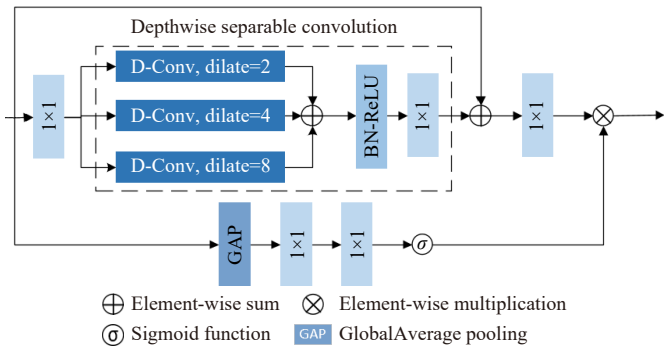

Figure 2 Compact pyramid refinement diagram

automatically decides to select and integrate attention operators to compute attention graphs. The type of incentive operator was adjusted according to the network layer and the scenario to achieve the combination of different incentive operators in different network layers. The switchable excitation module proposed in this study was added before the iterative refinement operation so that it trains the Sigmoid values of each channel to obtain the corresponding weights for each channel, and finally gives more attention to the channels with larger weights, while suppressing the channels with smaller weight values.

2.3 Convformer encoder

To address the issue that the global attention mechanism of the transformer may be weakened in the context of CNN, we propose a Convformer encoder. Convformer is an encoder model that combines CNN style with transformer thinking. It mainly consists of two parts: CNN-style Convolutional Self-Attention (CSA) and Convolutional Feedforward Network (CFFN).

CNN-style Convolutional Self-Attention (CSA) draws inspiration from the convolutional operation in CNN and simulates the computation of query (Q), key (K), and value (V) in the self-attention mechanism through local convolution. This not only maintains the ability to extract local features but also introduces a global attention mechanism, enabling the model to better handle global information. The calculation formulas of Q and K are shown in Equation (2).

$$Q_{i,j} = \sum_{l=-1}^{l}\sum_{g=-1}^{l} E^{q}_{2+l,2+g} x_{i+l,j+g}$$
$$K_{i,j} = \sum_{l=-1}^{l}\sum_{g=-1}^{l} E^{k}_{2+l,2+g} x_{i+l,j+g} \quad (2)$$

where, Q represents the query; K represents the key value; V represents the attribute value; and E is a learnable projection matrix.

The specific process of the CSA module is shown in Figure 3.

The Convolutional Feedforward Network (CFFN) plays one of the main roles in Convformer encoders. As a convolutional neural network, its main goal is to extract local features from images and integrate spatial information. This design concept aims to ensure that Convformer can maintain similar performance to traditional CNN when processing local information. The main structure of CFFN consists of convolutional layers, Batch Normalization (BN), and ReLU activation function, which work together to effectively extract and integrate local features in images. In addition, CFFN also plays a role in improving the features generated by Convolutional Self-Attention. The structural flow of CFFN is shown in Figure 4, consisting of two CBRs linearly connected.

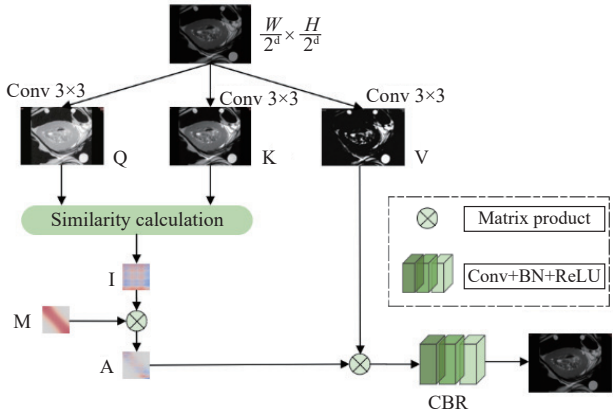

Figure 3 Convolutional Self-Attention diagram

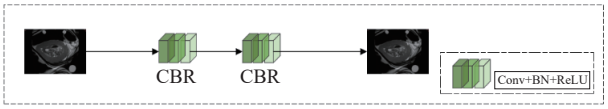

Figure 4 Convolutional Feedforward Network diagram

2.4 Token Reinforcement Module

The main function of the Token Reinforcement Module is to interact and explore features between tokens in adjacent local regions. Through this approach, it is hoped that the model can pay more attention to local features and enhance its ability to extract local features. Specifically, the Token Reinforcement Module introduces a local feature interaction mechanism to enable adjacent tokens to better learn and understand each other's features. In this way, the model can not only utilize global information but also more effectively capture and utilize local features, thereby improving the accuracy of boundary recognition in grayscale image segmentation tasks. The structure diagram of the Token Reinforcement Module is shown in Figure 5.

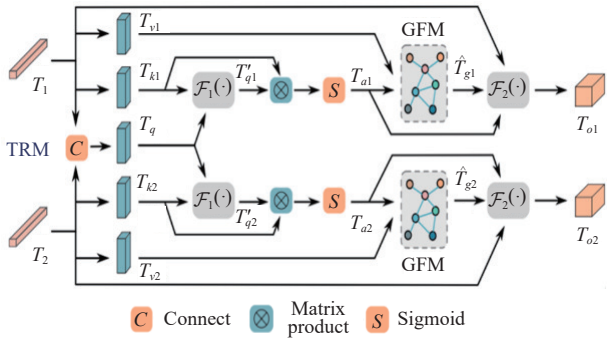

Figure 5 Token Reinforcement Module diagram

2.5 Adjacent Transfer Module

At this stage, the adjacency propagation aggregation module was proposed. It will fuse the current feature with the adjacent feature of the previous one and put the fused feature into the following transmission. It can be seen that the Adjacent Transfer Module plays a bridging role in the decoder for adjacent feature fusion and information transfer (both within the same layer and across layers). The module architecture is shown in Figure 6.

F_p is the feature passed to the Adjacent Transfer Module of the next layer, and the adjacent transfer aggregation module can be specifically expressed as the following equation:

$$F_p = CBR(C(CBR(C(F_{p-1}, F_i)), F_{i-1})) \tag{3}$$

where, F_i and F_{i-1} are the adjacent feature pairs of the current feature layer; F_{p-1} is the output aggregate feature of the Adjacent Transfer Module of the previous layer; $C(\cdot)$ represents the join operation; and CBR is the set of convolution, Batch Normalization, and modified linear unit functions.

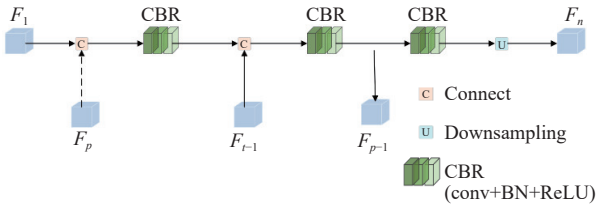

Figure 6 Adjacent Transfer Module diagram

3 Experiments and analysis of results

3.1 Establishment of the datasets

To balance the differences between winter and summer growth of Pacific oysters and ensure that oysters of any season can be used for sex identification, 300 Pacific oysters of similar shape and individual size were selected in the same culture environment. The 7.0 T high field intensity small animal magnetic resonance imaging system equipment was used to obtain MRI images, as shown in Figure 7. Figure 7a represents a small animal MRI instrument using Bruker BioSpec 70/20 USR, manufactured by Bruker, Germany. Figure 7b represents individual images of oysters of similar size selected during the experiment, in which larger individuals cannot be selected due to the limited aperture of the equipment. Figure 7c represents the NMR experiment process of a single Pacific oyster. The main technical specifications of the equipment were: 7.0 T magnet, aperture width of 20 cm, 660 mT/m gradient intensity, 7 groups of high-order uniform field coils, gradient power supply of 500 V/300 A, and the highest image pixel resolution of 10 μm. As for the main parameter settings, the longitudinal slice length was set to 2 mm, the number of slices per Pacific oyster was 20, and the echo time TE of transverse (T_2) relaxation was set to 30 ms.

Figure 7 Photographs taken with NMR equipment

A total of 6000 slice images were obtained from 300 oysters, and only the parts of the MRI images containing the gonads were retained. 3500 Pacific oyster MRI images were obtained after screening, and the gonadal boundaries of the original images were labeled with labelme software to build the oyster gonad datasets. The annotated images were randomly divided into training and test sets in the ratio of 6:1, of which 3000 were used for training the segmentation model and 500 were used for the tested model. An example map of oyster gonad dataset annotation is shown in Figure 8. Labelme software was used to label the gonad boundary of the original image, the gonad area was completely depicted, and

then the .json file was uniformly converted into binary ground truth image format. The top row in the image represents the original NMR image, and the bottom corresponds to the binary ground truth image of the gonadal region.

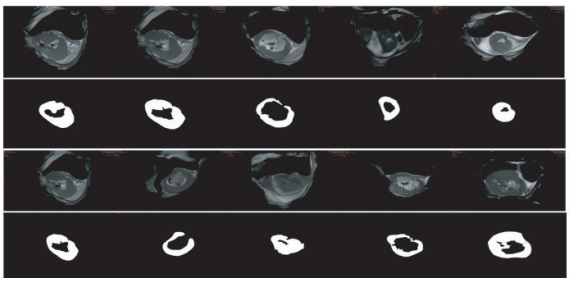

Figure 8 Example of Pacific oyster gonad dataset annotation

3.2 Experimental environment and evaluation index

The hardware and software parameters used in this study are configured as listed in Table 1.

Table 1 Software and hardware parameter configuration

Software and hardware environment	Configuration
Small animal MRI system	Bruker BioSpec 70/20 USR
Processor	Intel(R) Core(TM) i9-9820X
Graphics processor	NVIDIA Corporation GP100GL
Graphics processor computing platform	CUDA 10.2, cuDNN 7.4
Compile the program	Pycharm, Anaconda
Frame	Pytorch
Programming languages	Python 3.8

The input image size of the experiments in this study is 352×352, the epoch size is 50 during training, and the batch size is 8. The specific hyperparameter settings of the algorithm model in this study are listed in Table 2. The training is performed using Adam optimizer[21], and the whole training process takes about 70 min.

Table 2 R-SINet & CF-Net algorithm hyperparameter settings

Parameters	Numerical value
Input size	352×352
Learning rate	0.0001
Batch size	8
Epoch	30
Number of iterations	3
Optimizer	Adam

The model's evaluation metrics include *S*-measure ($S\alpha$)[22], enhanced-matching evaluation metrics *E*-measure ($E\Phi$)[23], weighted *F*-measure (ωF)[24], and mean absolute error (MAE)[25] . $S\alpha$ is an evaluation of the perceived structural similarity of the object, and a larger value indicates a better match of structural similarity. $E\Phi$ evaluation index is based on the human visual perception mechanism and is used to evaluate pixel-level similarity and image-level similarity, which is suitable for evaluating the overall and local accuracy of unapparent object detection, and a larger value indicates higher detection accuracy. ωF is shown through experiments to be more reliable than the traditional *F*-measure method[26], and its role is to carry out a comprehensive measure of both accuracy and recall; the larger its value, the better the segmentation effect. MAE is used to evaluate the accuracy difference between the prediction map and the true image pixel level, which can effectively evaluate the occurrence and number of errors; the smaller its value, the smaller the error.

3.3 Ablation experiments

To verify whether the two models proposed in R-SINet algorithm, the Compact Pyramid Refinement Module (CPR) and the Switchable Excitation Model (SEM), are effective in this segmentation task, corresponding ablation experiments were conducted for both of them. The ablation experiments were tested in the oyster gonad datasets using the same hyperparameters, and the test results are listed in Table 3. From the data in the table, it can be seen that the benchmark module, SINet_v2, improves all four evaluation indices after adding the Compact Pyramid Refinement Model, among which $S\alpha$ has a large improvement of 1.8%, reaching a similarity rate of 88.3%, indicating that the model has a greater effect on the feature extraction of the spatial structure in the algorithm. When only the Switchable Excitation Model is added, $E\Phi$ and ωF evaluation indices improve relatively more, by 1.3% and 1.5%, respectively, indicating that this model can obtain global and local information more accurately. The experimental results prove that the addition of the two models has a positive effect on the results of the algorithm, which performs well in all four evaluation indices, and shows the best results after incorporating the two models into the overall framework at the same time, verifying the effectiveness of the two models.

Table 3 Impact of the two models proposed in this study on the R-SINet algorithm

Group	SINet_v2	CPR	SEM	$S\alpha\uparrow$	$E\Phi\uparrow$	$\omega F\uparrow$	MAE↓
No.1	√			0.865	0.910	0.863	0.038
No.2	√	√		0.883	0.911	0.865	0.033
No.3	√		√	0.866	0.923	0.878	0.039
No.4	√	√	√	**0.888**	**0.928**	**0.879**	**0.032**

Note: CPR: Compact Pyramid Refinement Module; SEM: Switchable Excitation Model; $S\alpha$ represents the *S*-measure; $E\Phi$ represents the enhanced-matching evaluation metrics *E*-measure; ωF represents the weighted *F*-measure; MAE: Mean absolute error.

In order to verify the effectiveness of the three boundary enhancement methods proposed by the CF-Net algorithm, Convformer encoder (CE), Token Enhancement Module (TEM), and Adjacent Transfer Module (ATM) in this segmentation task, corresponding ablation experiments were conducted on them. The same hyperparameters were used to test the ablation experiment in the oyster gonad dataset, and the test results are listed in Table 4. As can be seen from the data in the table, the three methods have been successively added to the benchmark module (BM), all of which have been improved. The similarity rate of $E\Phi$ increased by 4.2%-90.4%, indicating that the model has a great effect on the feature extraction of spatial structure in the algorithm. When only the marker-enhanced model is added, the $S\alpha$ evaluation index increases by 8.3%, indicating that this model can obtain global and local information more accurately. When the Adjacent Transfer Module is added, the ωF performance is effectively improved. Therefore, the experimental results show that the addition of the three modules has a positive effect on the results of the algorithm, which performs well in the four evaluation indicators, and shows the best effect after integrating them into the overall framework at the same time, increasing by 10.5%, 6.8%, 3.6%, and 2.4%, respectively, verifying the effectiveness of the optimized algorithm in this study.

3.4 Comparative experiments and analysis of results

To verify the performance level of the proposed R-SINet and CF-Net, they were tested with the SINet[9], LSR[10], SINet_v2[13], and SegMaR[14] in the oyster gonad datasets and the same running environment, respectively. These four comparison algorithms were

selected from the current excellent papers on the segmentation of unapparent objects, and the comparison was authentic and effective. The quantitative evaluation results of the experimental comparison are listed in Table 5.

Table 4 Impact of the two models proposed in this study on the CF-Net algorithm

Group	BM	CE	TEM	ATM	$S\alpha\uparrow$	$E\Phi\uparrow$	$\omega F\uparrow$	MAE↓
No.1	√				0.803	0.862	0.883	0.047
No.2	√	√			0.834	0.904	0.889	0.040
No.3	√		√		0.886	0.890	0.890	0.035
No.4	√			√	0.885	0.891	0.898	0.038
No.5	√	√	√	√	**0.908**	**0.930**	**0.919**	**0.023**

Note: BM: benchmark module ; CE: Convformer encoder; TEM: Token Enhancement Module; ATM: Adjacent Transfer Module; $S\alpha$ represents the S-measure; $E\Phi$ represents the enhanced-matching evaluation metrics E-measure; ωF represents the weighted F-measure; MAE: Mean absolute error.

Table 5 Evaluation results of different algorithms in the four evaluation indices

Algorithm	$S\alpha\uparrow$	$E\Phi\uparrow$	$\omega F\uparrow$	MAE ↓
SINet[9]	0.662	0.887	0.476	0.068
LSR[10]	0.658	0.917	0.475	0.058
SINet_v2[13]	0.865	0.910	0.863	0.038
SegMaR[14]	0.653	0.918	0.474	0.059
R-SINet (the method of this study)	**0.888**	**0.928**	**0.879**	**0.032**
CF-Net (the method of this study)	**0.908**	**0.930**	**0.919**	**0.023**

SINet algorithm[9], which first introduced the concept of camouflaged object detection, has a network structure that includes the acquisition of sensory field search target objects and the extension of the Partial Decoder Component (PDC) with a dense connection to aggregate the remaining three layers of features to obtain the final target object. The network framework is simple and cannot achieve precise segmentation of the Pacific oyster gonads.

The LSR algorithm[10] adds the inverse attention module to the SINet algorithm to obtain more spatial attention features, and its method slightly outperforms the SINet algorithm in two evaluation metrics, $E\Phi$ and MAE, by 3.0% and 0.1%, respectively. However, its performance is inadequate compared with the algorithm proposed in this study.

The SINet_v2 algorithm[13], which is the benchmark model of the algorithm in this study, contributes by improving the sensory field module and PDC module based on SINet while adding the GRA module in the identification of the target object stage to segment the gonadal part more finely. According to the quantitative results, its $S\alpha$, ωF, and MAE evaluation indices have been significantly improved. However, its performance is still inferior to the algorithm proposed in this study when all evaluation indices are combined.

SegMaR algorithm[14], based on recognition segmentation, adds the step of zooming in on the local area of the image and operates iterative refinement. $E\Phi$ detection accuracy is improved, reaching an accuracy of 91.8%, but it performs poorly in the remaining three indices. The main problem is that the zooming process focuses on local details and tends to ignore the overall differences in spatial perceptual structure.

From the evaluation results, it can be seen that the segmentation performance of the methods proposed in this study was higher than that of the other four algorithms. Compared with the above four algorithms, the R-SINet algorithm model proposed in this study can simultaneously take into account the search and localization of the target object and the recognition segmentation task. By obtaining larger sensory fields, adding compact feature pyramids for efficient feature fusion, and introducing Switchable Excitation Models after grouping embedded inverted feature maps to make the network pay more attention to channels with larger weights, the segmented image edge features are refined to achieve the effect of accurate segmentation. By combining CNN and transformer methods to optimize the segmentation of oyster gonad boundaries, the CF-Net network architecture was also proposed. In the feature extraction stage, a Convformer encoder is designed to promote attention convergence, and a transformer is used to extract global information features. Convolutional Self-Attention and Convolutional Feedforward Networks are introduced to achieve better segmentation performance. Then, a marker enhancement module was proposed, which acts on adjacent local regions to enhance the ability to extract local features. In the decoder stage, an Adjacent Transfer Module is used to form a feature shrinkage decoder, which fuses target feature information to complete the refinement task of oyster gonad boundaries. Finally, the model achieved the best test results in all four evaluation metrics, and the visualized images also showed that the segmentation images obtained by the method had clearer and more accurate boundaries, proving the good practical performance of the method.

3.5 Visualization results

Some of the visualized segmentation results are shown in Figure 9a, which shows the Pacific oyster MRI map. Figure 9b shows the ground-truth map of the Pacific oyster gonad segmentation. Figures 9c and 9d show the test result images of oyster gonad segmentation obtained using the methods of this study, in which it can be seen that the boundaries are clearer and the segmentation results are closer to the ground-truth map. Figure 9e shows the test result images of oyster gonad segmentation using the SegMaR algorithm, in which some areas are inaccurate compared to the ground-truth map. Figure 9f shows the test result images of oyster gonad segmentation using the SINet_v2 algorithm, where the boundaries in the resulting images are blurred. Figure 9g shows the test result images of oyster gonad segmentation using the LSR algorithm, in which it can be seen that the boundaries are not precise enough. Figure 9h shows the test results of oyster gonad segmentation using the SINet algorithm, where the visual segmentation results are somewhat vague.

Figure 9 shows that the contour structure of the segmented gonads of Pacific oysters obtained by the two methods of this study is closer to the true value map, and the gonad edges are clearer than the other four algorithms, which confirms the effectiveness of these methods.

3.6 Gray value calculation

Based on the method proposed in this study for optimizing the boundary segmentation of the Pacific oyster's gonads, the segmented gonad is mapped onto the original image to obtain a grayscale image of the region. Then, the average gray value in the irregular enclosed area in the image is calculated. Through the above operations, the average gray value of the gonadal region of a single image can be obtained. This study randomly selected 2000 images after model testing to calculate the average gray value of the gonadal region and finally used statistical graphs to represent the distribution of their gray values, as shown in Figure 10.

By analyzing these numerical differences, we can find the difference rule of the gray value range of the gonads of male and female individuals. This can not only help this study better understand the degree of gonadal development in oysters, but also

provide technical support for subsequent non-destructive discrimination of oyster sex, provide the best pairing scheme for their reproduction and breeding, improve breeding efficiency and success rate, and provide technical assistance for subsequent non-destructive discrimination of oyster sex in parent breeding and other applications.

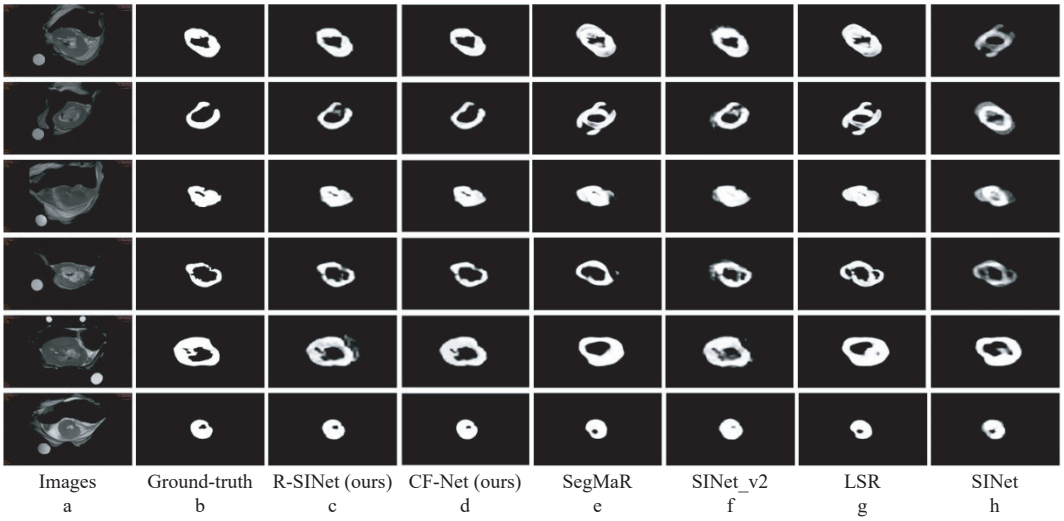

Figure 9 Visual comparison of gonad segmentation results

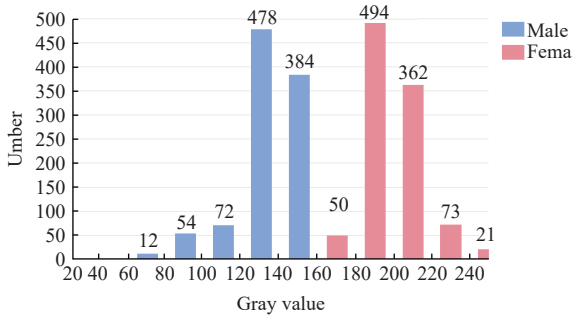

Figure 10 Gray histogram curve of male and female oysters

4 Conclusions

This study is grounded in the magnetic resonance imaging of Pacific oysters and employs the approach of unobvious target detection to identify and segment the gonad images of Pacific oysters. Additionally, the R-SINet network architecture was proposed to achieve a complete segmentation of the oyster gonad area, and the CF-Net boundary optimization algorithm was utilized to enhance the precision of oyster gonad segmentation. Firstly, an oyster gonad dataset consisting of 3500 images was established. Secondly, two models were created. For the first model, a Compact Pyramid Refinement Module was designed to fuse adjacent semantic features, and a Switchable Excitation Model capable of adaptive recalibration was proposed. This model automatically adjusts the size of the excitation operator based on the significance of different excitation operators, effectively enhancing the interaction among attention channels, and ultimately achieving multi-effective feature fusion for non-significant object recognition and detection of oyster gonads. For the second model, in the feature extraction stage, a Convformer encoder was designed to facilitate attention convergence, and Convolutional Self-Attention and Convolutional Feedforward Networks were introduced to obtain better segmentation performance. The designed label enhancement module operates in adjacent local regions, enhancing the extraction ability of local features. Moreover, in the decoder stage, a feature contraction decoder composed of Adjacent Transfer Modules was utilized to fuse target feature information and complete the refinement task of oyster gonad boundaries. The models proposed in this study achieved the optimal test results in all four evaluation indicators, demonstrating the effectiveness of the proposed approach. Furthermore, through the grayscale histogram, it was visually discovered that there is a significant difference in the grayscale values of the female and male oyster gonads, providing new techniques and means for the non-destructive selection of oyster gender in subsequent studies.

This provides empirical evidence to support the continued development of gonad recognition to achieve high-quality trait inheritance in a range of shellfish and contributes to the continued development of deep learning-based recognition of complex backgrounds in marine organisms. In the current work, the large and tedious dataset labeling took a lot of time and effort. In future research, semi-supervised or self-supervised image segmentation should be tried in order to remove the problem of a large image labeling workload in the early stage.

Acknowledgements

The authors acknowledge that this work was financially supported by the Shandong Province Key R&D Program Project (Grant 2022TZXD005), Shandong Province Science and Technology SMEs Innovation Capacity Enhancement Project (Grant 2021TSGC1003), and Yantai City Key R&D Program Project (Grant 2022XCZX079).

[References]

[1] Purdon A, Mole M A, Selier J, Kruger J, Mafumo H, Olivier P I. Using the Rao's Q diversity index as an indicator of protected area effectiveness in conserving biodiversity. Ecological Informatics, 2022; 72: 101920.

[2] Xia Y, Liu W X, Meng J W, Hu J H, Liu W B, Kang J, et al. Principles, developments, and applications of spatially resolved spectroscopy in agriculture: A review. Frontiers in Plant Science, 2024; 14: 1324881.

[3] Yang L. Introduction to the management of hospital Bruker BioSpec94/30 USR type small animal MRI research equipment. China Equipment Engineering, 2022; 4: 51–52. (in Chinese)

[4] Zhang Z N, Zheng Y, Wang X M. Application of 7.0T small animal MRI to study the progress of Alzheimer's disease. Chinese Journal of Medical Imaging Technology, 2019; 35(6): 930–933. (in Chinese)

[5] Hang K B, Su W W, Huang J, Bao G J, Liu W H, Li S P. 7.0T small animal

MR instrumentation to observe brain injury in a rat model of classic pyrexia. Chinese Journal of Medical Imaging Technology, 2022; 38(4): 481–485. (in Chinese)

[6] Gilchrist S, Kinchesh P, Kersemans V, Beech J, Allen D, Brady M, et al. A simple, open and extensible gating control unit for cardiac and respiratory synchronisation control in small animal MRI and demonstration of its robust performance in steady-state maintained CINE-MRI. Magnetic Resonance Imaging, 2021; 81: 1–9.

[7] Liu W L, Li J H, Li L, Zhang Y H, Yang M G, Liang S C, et al. Enhanced medial prefrontal cortex and hippocampal activity improves memory generalization in APP/PS1 mice: A multimodal animal MRI study. Frontiers in Cellular Neuroscience, 2022; 16: 848967.

[8] Baskaya F, Lemainque T, Klinkhammer B, Koletnik S, von Stillfried S, Talbot S R, et al. Pathophysiologic mapping of chronic liver diseases with longitudinal multiparametric MRI in animal models. Investigative Radiology, 2024; 59(10): 699–710.

[9] Fan D P, Ji G P, Sun G L, Cheng M M, Shen J B, Shao L. Camouflaged object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle: IEEE, 2020; pp.2774–2784. doi: 10.1109/CVPR42600.2020.00285.

[10] Lv Y Q, Zhang J, Dai Y C, Li A X, Liu B W, Barnes N, et al. Simultaneously localize, segment and rank the camouflaged objects. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville: IEEE, 2021; pp.11591–11601. doi: 10.1109/CVPR46437.2021.01142.

[11] Chen S H, Tan X L, Wang B, Hu X L. Reverse attention for salient object detection. In: Computer Vision – ECCV 2018, 2018; 11213: 236–252.

[12] Zhai Q, Li X, Yang F, Jiao Z C, Luo P, Cheng H, et al. MGL: Mutual graph learning for camouflaged object detection. IEEE Transactions on Image Processing, 2023; 32: 1897–1910.

[13] Fan D P, Ji G P, Cheng M M, Shao L. Concealed object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021; 44(10): 6024–6042. https://arxiv.org/pdf/2102.10274.

[14] Jia Q, Yao S L, Liu Y, Fan X, Liu R S, Luo Z X. Segment, magnify and reiterate: Detecting camouflaged objects the hard way. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans: IEEE, 2022; pp.4703-4712. doi: 10.1109/CVPR52688.2022.00467.

[15] Li L P, Shi F P, Wang C X. Fish image recognition method based on multi-layer feature fusion convolutional network. Ecological Informatics, 2022; 72: 101873.

[16] Xia Y, Che T C, Meng J W, Hu J H, Qiao G L, Liu W B, et al. Detection of surface defects for maize seeds based on YOLOv5. Journal of Stored Products Research, 2024; 105: 102242.

[17] Fu K R, Fan D P, Ji G P, Zhao Q J, Shen J B, Zhu C. Siamese network for RGB-D salient object detection and beyond. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021; 44(9): 5541–5559.

[18] Li C Y, Cong R M, Piao Y R, Xu Q Q, Loy C C. RGB-D salient object detection with cross-modality modulation and selection. In: Computer Vision - ECCV, 2020; 12353: 225–241.

[19] Gao S H, Cheng M M, Zhao K, Zhang X Y, Yang M H, Torr P. Res2Net: A new multi-scale backbone architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019; 43(2): 652–662.

[20] Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint, 2017; arXiv: 1704.04861.

[21] Kingma D P, Ba J L. Adam: A method for stochastic optimization. In: ICLR 2015. 2015.

[22] Fan D P, Cheng M M, Liu Y, Li T, Borji A. Structure-measure: A new way to evaluate foreground maps. In: 2017 IEEE International Conference on Computer Vision (ICCV), 2021; pp.4558–4567. doi: 10.1109/ICCV.2017.487.

[23] Fan D P, Gong C, Cao Y, Ren B, Cheng M M, Borji A, et al. Enhanced alignment measure for binary foreground map evaluation. In: IJCAI, 2018: Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018; pp.698–704. doi: 10.24963/ijcai.2018/97.

[24] Margolin R, Zelinik-Manor L, Tal A. How to evaluate foreground maps. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus: IEEE, 2014; pp.248–255. doi: 10.1109/CVPR.2014.39.

[25] Perazzi F, Krähenbühl P, Pritch Y, Hornung A. Saliency filters: Contrast based filtering for salient region detection. 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012: 733–740. doi: 10.1109/CVPR.2012.6247743.

[26] Hand D, Christen P. A note on using the *F*-measure for evaluating record linkage algorithms. Statistics and Computing, 2018; 28: 539–547.