Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-6, Issue-9, September 2024
DOI: 10.53469/jrse.2024.06(09).04

Comparison and Analysis of CNN based Underwater Aquatic Products Recognition Methods

Liqiong Lu*, Dong Wu, Liuyin Wang

School of Computer Science and Intelligence Education, Lingnan Normal University, Zhanjiang 524048, China
**Correspondence author, hb_llq@126.com*

Abstract: ***Underwater aquatic products are naturally cultured in seawater exceeding two meters to obtain excellent quality. The catching of these aquatic products mostly relies on professional fishermen, which consumes a lot of manpower and material resources. In recent years, underwater fishing robots have emerged, but due to inaccurate positioning of Underwater aquatic products, the fishing efficiency is not satisfactory. Based on CNN, underwater aquatic product recognition methods were researched. Firstly, an underwater aquatic products recognition dataset containing 5443 aquatic product images was constructed based on the training data provided by the National Underwater Robot Competition - Underwater Object Detection Competition. Subsequently, SSD, Faster RCNN, YOLO V5, and YOLO V8 were used to recognize underwater aquatic products on the above dataset and the recognition performance of various methods was compared. The experimental results show that YOLO V8 has the most ideal recognition performance, with an mAP value of 0.862.***

Keywords: Underwater images, Underwater aquatic products recognition, CNN.

1. Introduction

There are many ancient and beautiful coastal cities in China, in which a considerable number of fishermen are born to the sea. These fishermen live by cultivating aquatic products such as sea cucumbers and sea urchins on the seabed. In order to obtain the excellent quality close to wild aquatic products, the seabed cultivation sites often choose the depths of more than two meters. During the annual seafood harvest season, it is necessary to rely on professional ginseng collectors who specialize in underwater fishing to go deep into the sea to catch aquatic products. The working environment of artificial underwater fishing is harsh, and the risk factors of underwater depth are uncontrollable, making it a high-risk industry. In recent years, with the rapid development of robot and artificial intelligence technology, submarine working robots came into being. However, due to the limited recognition performance of submarine aquatic products, the fishing efficiency is not high. The main reason for the limited recognition performance of underwater aquatic products is due to the limitations of the special imaging environment underwater. Underwater images often have many problems such as noise interference, blurred texture features, low contrast, and color distortion, which make underwater image feature extraction very difficult.

The recognition method of underwater aquatic products belongs to the category of underwater object recognition methods, which are mainly divided into two types from technical perspective. The first kind is to combine the characteristics of underwater images with traditional machine learning methods to identify underwater objects. Chen et al. [1] proposed to combine wavelet transform with multi-layer perceptron, Guo et al. [2] combined texture features extracted from gray level co-occurrence matrix with Support Vector Machine, SVM), and Zhu et al. [3] used independent component analysis to combine fusion features with improved Gentle AdaBoost model for underwater object recognition. The recognition accuracy of these methods depends on the degree of matching between image feature extraction and recognition tasks. The proposed methods often have limited generalization ability and are prone to overfitting. The inherent flaw of such methods, namely the tedious manual feature extraction, limits the further development of this technology in the field of underwater object recognition. Manually extracting feature engineering is often too singular and has poor generalization ability, resulting in poor accuracy and universality of such methods.

In recent years, Convolution Neural Networks (CNNs) have been widely used in various fields of computer vision and have achieved remarkable results. Therefore, the second type of underwater target recognition method based on CNN has emerged in the field of underwater object recognition. Most CNN based underwater object recognition methods are based on existing excellent methods in the field of object recognition, and combined with the characteristics of underwater objects, these methods are improved to obtain new methods suitable for underwater object recognition. Dong et al. [4] first summarized the datasets of underwater targets, including fish and benthic organisms, and gave the corresponding links. Then, based on Faster RCNN [5] and YOLO V3 [6], three typical methods for improving the detection performance of targets were compared and analyzed, and it was concluded that the application of high-resolution network could obviously improve the performance of object recognition. Huang et al. [7] combined the coordinate attention mechanism with YOLO V5 [8], and modified the arc tangent function in the penalty term of the traditional loss function CIoU in YOLO V5 to Sigmoid function to propose an underwater object recognition method. Qiang et al. [9] replaced the VGG network [11] in SSD [10] with the ResNet network [12] and proposed an underwater object recognition method using deep separation deformable convolution modules for feature extraction. Li et al. [13] first constructed a deep sea cold spring biological image object detection dataset, and then combined the R-FCN framework with ResNet network and anchor reset to propose a deep sea cold spring biological object recognition method. Generally speaking, the underwater object recognition method based on CNN automatically extracts features from the original signal without professional knowledge in related fields, so as to avoid the loss of features in the manual extraction process and improve the generalization ability, and it can also improve the efficiency and accuracy of automatic recognition in the process of continuously optimizing the model, and its performance is superior to the traditional machine learning method. Therefore, the academic and industrial circles focus on the second method of underwater target recognition.

In this paper, firstly, a dataset of underwater aquatic products identification is constructed, which contains 5443 aquatic products images. The aquatic products in these underwater images include sea cucumber, sea urchin, starfish and scallop, and they are taken by underwater fishing robots in real time. It is of great practical and guiding significance to use these images as the data set for identifying underwater aquatic products. Subsequently, based on the deep learning platform, the general methods in the field of object recognition are realized, including SSD [10], Faster RCNN [5], YOLO V5 [8] and YOLO V8 [14]. These methods are applied to the constructed underwater aquatic product recognition data set, and the recognition performance is compared and analyzed. The experimental results show that YOLO V8 has the best recognition performance, with a score of 0.862 in *mAP*.

2. Underwater Aquatic Product Recognition Dataset

A dataset of underwater aquatic products was constructed using the certain filtering test images provided by the National Underwater Robot Competition - Underwater Target Detection Competition as raw data. The training set in this dataset contains 4988 images, while the test set contains 554 images. There are four types of aquatic products involved in this dataset, namely sea cucumber, sea urchin, starfish, and scallop. In addition, the underwater aquatic product images in the dataset are obtained from real-time shooting by underwater operation robots. Due to the complex underwater environment, the quality of underwater aquatic product images is poor. For example, uneven imaging light in the deep sea leads to severe color distortion, low contrast, and overall green color bias in the images of these underwater aquatic products; The continuous movement of underwater currents often results in blurry images captured in real-time; Due to the difficulty in effectively adjusting the underwater environment, there may be an imbalance in the total number of different categories of aquatic products captured on the seabed.

Figure 1 lists the sample images of aquatic products in this paper, and the presentation of these images can basically prove the characteristics of color cast, blur and low contrast of aquatic products. It can be seen from Figure 1 that the overall color of most underwater images is distorted to some extent, and the main color is mainly green. Some underwater images are full of green, so it is difficult to see the existence of aquatic products; Some images of underwater aquatic products are dull in overall tone and uneven in overall light. In addition, most of the real-time images of the seabed are blurred and have low contrast, some aquatic products are covered by fishing nets, and some aquatic products blend seamlessly with the background, which makes the recognition of aquatic products on the seabed quite difficult.

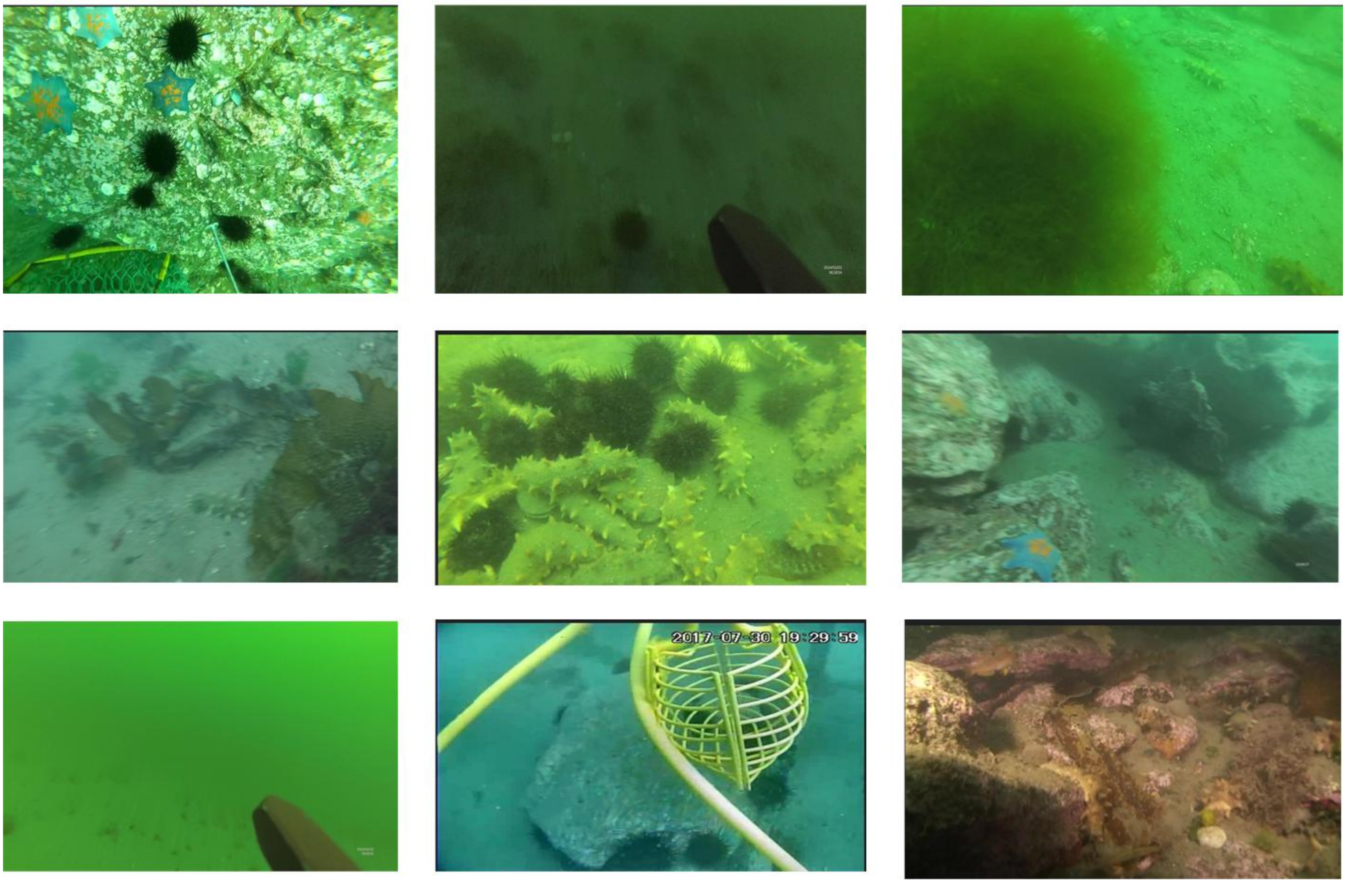

Figure 1: Sample images of seafloor aquatic products

This paper calculates the total number of samples of various aquatic products in the seabed images in the training set, and the statistical results are shown in Figure 2. From Figure 2, it can be seen that sea urchins have the highest sample size, reaching nearly 20000, while sea cucumbers have the lowest sample size, approaching 5000. The number of starfish and scallops is basically close to 6000. Due to the difficulty in predicting the quantity of different types of underwater aquatic products during filming, there is inevitably an imbalance in the dataset. Therefore, when designing underwater aquatic product recognition algorithms, it is necessary to accurately mine the characteristics of different aquatic products in order to cope with the imbalance in sample size.

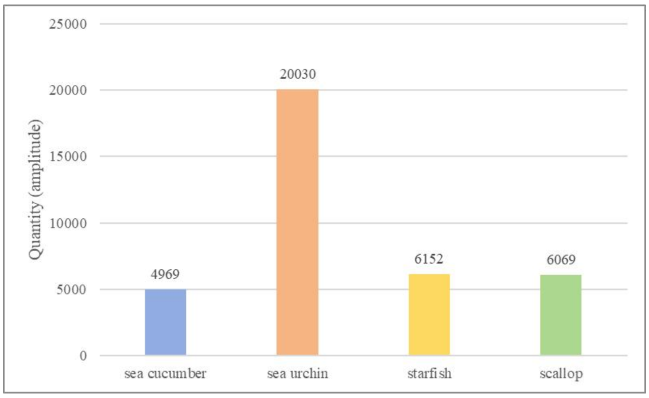

Figure 2: Statistics of underwater aquatic product categories in the training set

3. Comparison and Analysis of Identification Methods of Seafood Based on CNN

3.1 Experimental Set

In this paper, SSD and Faster RCNN methods are implemented based on Tensorflow 2.4, and YOLO V5 and YOLO V8 methods are implemented based on Pytorch 1.9. The above four methods are all trained and tested using Geforce 3060 graphics card. Because the underwater aquatic products belong to small target detection in the image, this paper resets the size of the default anchors when implementing SSD and Faster RCNN methods. The anchor_size is set to [21, 45, 99, 153, 207, 261, 315] in SSD and [32, 256, 512] in Faster RCNN. Other parameters in the experiment process, such as learning rate, are consistent with the initial setting of the method, and the value of bathsize is adjusted according to the size of GPU memory. YOLO V5 and YOLO V8 use yolo5s.pt and yolo8s.pt to initialize the parameters respectively.

3.2 Performance Evaluation Indicators for Underwater Aquatic Products Recognition

Similar to the performance indicators of general object recognition, this paper uses accuracy (Precision, P), regression rate (Recall, R), and comprehensive indicators mAP to evaluate the recognition performance of underwater aquatic products.

$$P = \frac{FP}{FP+TP};$$

$$R = \frac{TP}{TP+FN};$$

Among them, TP (true positive) represents the number of correctly identified samples; FP (false positive) represents the number of error identified samples; FN (false negative) represents the number of undetected samples. Here, when the IOU between the recognition box and the ground truth box exceeds 50%, it is considered correct recognition. mAP(mean Average Precision) is the average of the Average Precision values for all categories, calculated based on the Precision Recall curve. It interpolates Precision at different recall rates and calculates the area under the interpolation curve.

3.3 Comparative Analysis of Identification Methods of Submarine Aquatic Products

Table 1 lists the recognition results of different methods on the dataset of submarine aquatic products constructed in this paper. From the experimental results listed in Table 1, YOLO V8 has the highest value, reaching 0.862, followed by YOLO V5 with a mAP value of 0.858, followed by SSD and Faster RCNN. Moreover, the values of SSD and Faster RCNN are almost less than half of those of YOLO V5 and YOLO V8, which shows that YOLO V5 and YOLO V8 are more suitable for the recognition of underwater aquatic products. In addition, from the accuracy point of view, YOLO V5 has the highest accuracy, but its recall rate is lower than YOLO V8. On the whole, the recognition performance of YOLO series methods is obviously higher than SSD and Faster RCNN, and the recognition performance index of YOLO V8 is slightly better than YOLO V5.

Table 1: Comparison of recognition performance

Method	*P*	*R*	*mAP*
SSD	0.708	0.318	0.397
Faster CNN	0.659	0.276	0.376
YOLO V5	0.853	0.781	0.858
YOLO V8	0.832	0.794	0.862

YOLO V5 and YOLO V8 have excellent performance in recognizing underwater aquatic products. The reason is that the YOLO series methods are superior in design, mainly reflected in the following aspects: (1) YOLO V5 and YOLO V8 are more optimized in network design, introducing CSPDarknet to integrate the network structure of cross stage parts, achieving a certain balance between computational cost and accuracy. In addition, YOLO V8 further enhances CSPDarknet and optimizes feature extraction ability and channel configuration; (2) Multi scale fusion strategy optimization, YOLO V5 utilizes Path Aggregation Network (PANet) to optimize multi-scale feature fusion, and YOLO V8 adds skip connections and feature aggregation methods on this basis; (3) YOLO V8 adopts an Anchor free method, which does not require pre-set anchor sizes and has advantages in detecting small targets. In addition, YOLO V8 uses the Swish activation function, which does not use the Leaky ReLU function in YOLO V5.

From the demonstration and analysis of the above experimental results, it can be seen that the YOLO series methods are more advantageous as a basic method in the field of underwater aquatic product identification. Researchers in this field can combine the YOLO series methods with the feature extraction of underwater aquatic products to further improve the accuracy, regression rate and of underwater aquatic product identification.

3.4 Sample Analysis of Underwater Aquatic Product Recognition

Four sample images of the same image recognized by different methods are listed in Figure 3. It can be seen from Figure 3 that YOLO V5 and YOLO V8 can recognize more underwater aquatic products for the same image, especially YOLO V8, which can more accurately detect the positions of underwater aquatic products mixed with the background and blurred, and recognize the corresponding categories.

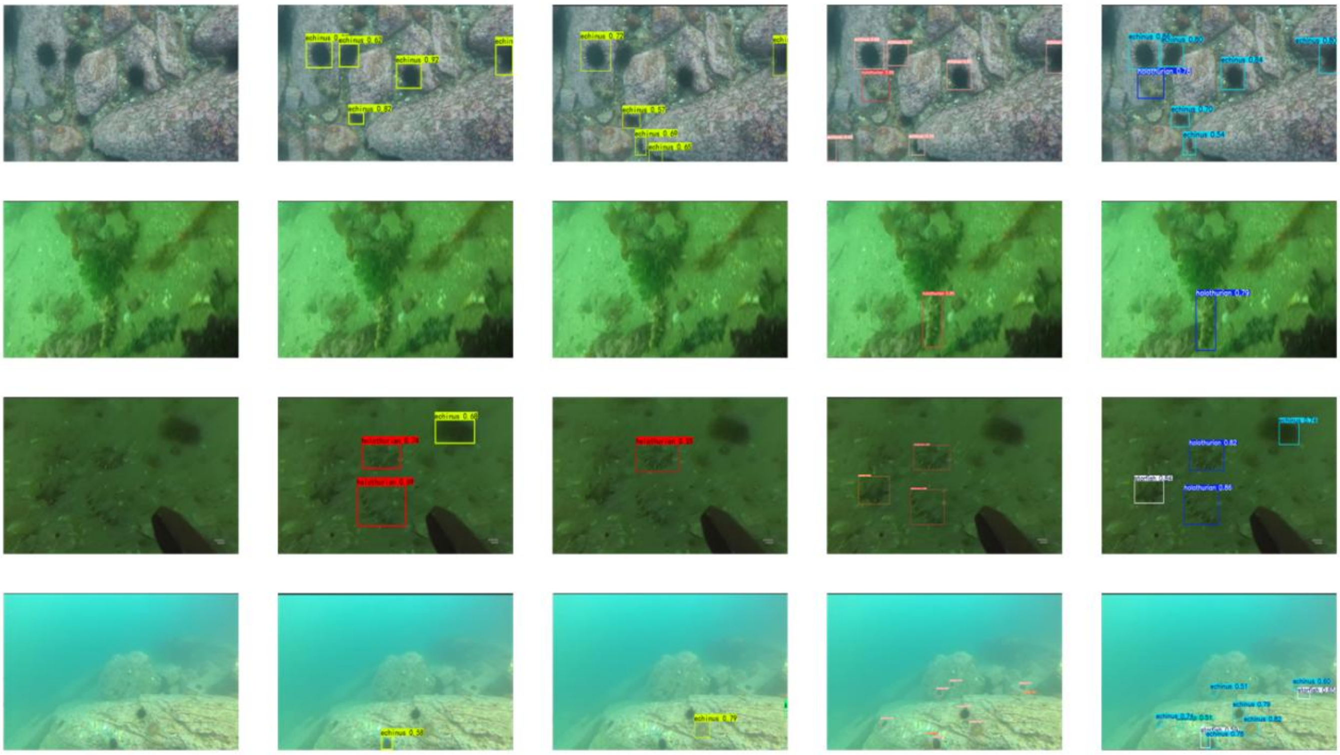

Figure 3: Original image and Recognition results of SSD, Faster RCNN, YOLO V5 and YOLO V8.

4. Conclusion

This paper focuses on the recognition of underwater aquatic products. Firstly, a dataset containing 5542 images was constructed for underwater aquatic product recognition. Then, four methods including SSD, Faster RCNN, YOLO V5, and YOLO V8 were used to compare and analyze the recognition performance of underwater aquatic products on this dataset. The experimental results show that YOLO V8 has the best performance, with a *mAP* value of 0.862, followed by YOLO V5 with a *mAP* value of 0.858. The recognition performance of SSD and Faster RCNN is not ideal. Based on the YOLO series methods, we plan to further analyze the characteristics of underwater aquatic products, such as fuzziness and small size, and design appropriate feature extraction mechanisms to improve the performance of underwater aquatic product recognition.

Acknowledgments

The authors are grateful to the anonymous reviewers and the helpful suggestion given by the partners. The research was supported by the Technology Project of Zhanjiang (no. 2022A01005), the Guangdong province philosophy and social science planning project (no. GD24CJY21).

References

[1] CHEN C H, LEE J D, LIN M C. Classification of underwater signals using wavelet transforms and neural networks[J]. Mathematical and Computer Modelling, 1998, 27(2):47-60.

[2] GUO J, MA J F, WANG A. Research on classification of side-scan sonar images based on SVM algorithm and GLCM[J]. Surveying and Mapping and Spatial Geographic Information, 2015, 38(3):60-63.

[3] ZHU B Y, WANG X, CHU Z, et al. Active learning for recognition of shipwreck target in side-scan sonar image[J]. Remote Sensing, 2019, 11(3):1030243.

[4] DONG J N, YANG M, XIE Z R, et al. Overview of underwater image target detection data set and detection algorithm [J]. Journal of marine technology, 2022, 41(5): 60-72.

[5] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6):1137-1149.

[6] Joseph Redmon, Ali Farhadi. YOLOv3: An Incremental Improvement [EB/OL]. [2018], https://alumni.soe.ucsc.edu/~czczycz/src/YOLOv3.pdf.

[7] HUANG C X, LIU P T. Research on underwater target recognition method based on CA-YOLOv5s [J]. Journal of marine technology, 2024, 43(03):17-24.

[8] Joseph Redmon. YOLO V5[EB/OL]. [2020-05-18]. https://github.com/ultralytics/yolov5

[9] QIANG W, HE R Y, GUO Y J, et al. Research on Underwater Target Detection Algorithm Based on Improved SSD [J]. Journal of northwestern polytechnical university, 2020, 38(4): 747-754.

[10] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[J]. 14th European Conference on Computer Vision (ECCV), 2015, (9905):21-37.

[11] Simonyan K, Zisserman A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science.

[12] He K, Zhang X, Ren S, et al. 2016. Deep Residual Learning for Image Recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[13] LI H J, LIAN C, GUAN L C, et al. Study on identification and distribution of organisms in deep-sea cold spring area based on R-FCN_ResNet [J]. Marine science, 2020, 44(9): 54-62.
[14] Ultralytics. YOLO V8 [EB/OL]. [2023-02-03]. https://github.com/ultralytics/ultralytics.