International Core Journal of Engineering
Volume 7 Issue 5, 2021
ISSN: 2414-1895
DOI: 10.6919/ICJE.202105_7(5).0028

A Comparative Study of Underwater Marine Products Detection based on YOLOv5 and Underwater Image Enhancement

Guanxi Huang

Shanghai Maritime University, Shanghai 201306, China.

353735042@qq.com

Abstract

The application of target detection algorithms in underwater images has not been very effective due to the quality problems such as image blurring and color incongruity that often exist in underwater images. In this paper, the YOLOv5 algorithm is used as a target detection network model, which is trained by using underwater marine products dataset and combined with six underwater image enhancement recovery algorithms to enhance and recover the images before they are detected, an attempt is made to improve the deep learning based target detection method at the input side. Finally, the impact of different underwater image enhancement algorithms on the YOLOv5-based target detection algorithm is compared and summarized through experiments, and the specific image enhancement methods can effectively improve the detection performance of the algorithm under different image background environments.

Keywords

Underwater Image Enhancement and Recovery; YOLO; Target Detection; Deep Learning.

1. Introduction

The oceans occupy 71% of the earth's surface area and contain a large amount of biological and mineral resources, which have always been an important source of food and biological research objects. With the development of underwater detection operation technology and computer vision technology, the detection technology of underwater organisms has been greatly improved. Underwater target detection is roughly the same as that on land, but the unique environmental characteristics make it difficult to detect targets in underwater environments.

Seawater itself is a chemical mixture with color, and there are also a large number of plankton, mineral particles, bacteria and other suspended matter, resulting in turbid seawater. And because light propagation in the water will be absorbed by the body of water, as well as scattered by the suspended matter in the water caused by light attenuation, the underwater image imaging difficulties, the image is often accompanied by low contrast, blurred, inconsistent color and noise and other quality problems, which increases the difficulty of target detection.

The existing underwater target detection methods are mainly classified into traditional algorithm-based detection methods and deep learning-based detection methods. The current traditional underwater target algorithms include Detecting man-made objects in unconstrained subsea videos[1], Real-world underwater fish recognition and identification, using sparse representation[2], and Fish Population Estimation and Species Classification from Underwater Video Sequences Using Blob Counting and Shape Analysis[3], but these traditional algorithms have obvious drawbacks, such as slow speed, low accuracy, and targeting, which require manual feature extraction, high workload, and lack of practicality and universality.

ISSN: 2414-1895
DOI: 10.6919/ICJE.202105_7(5).0028

With the rapid development in the field of artificial intelligence in recent years, methods for target detection using deep learning neural networks have been applied in many ways. The most typical neural network models used for target detection in recent years are RCNN[4], Fast-RCNN[5], Faster-RCNN[6], YOLO[7], and SSD[8]. This target detection method based on deep learning by training neural network models using data has the features of high universality, wide application scenarios, and can achieve detection of different targets by training using different datasets, and also has the advantages of high accuracy and speed, and has been widely used in target detection. In underwater target detection, deep learning-based target detection methods have been effectively applied in underwater environments through the improvement and optimization of neural networks, such as Research on underwater object recognition based on YOLOv3[9], Underwater Object Recognition Based on Deep Encoding-Decoding Network[10], A Deep Convolutional Neural Network Inspired by Auditory Perception for Underwater Acoustic Target Recognition[11], etc. These improved neural network detection methods have achieved high detection accuracy and speed in the application of underwater target detection. Both traditional target detection algorithms and deep learning-based target detection algorithms with better detection performance still have many limitations in the detection of underwater images with blurred images and complex backgrounds. As for underwater image enhancement, it is described in great detail in the review of experiment-based underwater image enhancement and recovery methods proposed by Yan et al[12].

Therefore, in order to further improve the detection accuracy of target detection performed underwater, this paper takes the detection of underwater marine products as an example and proposes that the underwater marine products images to be detected are first subjected to image restoration and enhancement processing, combined with deep learning-based target detection methods to improve the detection accuracy of target detection algorithms in underwater environments at the input side, and to compare each underwater image enhancement method.

2. Underwater marine products detection method based on YOLOv5 and underwater image enhancementt

2.1 YOLOv5

The YOLO algorithm is an end-to-end neural network model that uses only one convolutional neural network to segment the whole image and predicts the bounding box and the category to which each grid belongs, thus ensuring a high accuracy while detecting the image quickly. The YOLO network consists of three main components: Backbone, Neck and Head[13].

1.Backbone:Convolutional neural networks for aggregating and forming image features at different image granularity.

2.Neck:A series of network layers that mix and combine image features and pass the image features to the prediction layer.

3.Head:Prediction of image features, generation of bounding boxes and prediction of categories.

YOLOv5 is a latest version released in June 2020, so YOLOv5 is selected as the target detection algorithm in this paper.

2.1.1 Model Selection

YOLOv5 provides four models, which are divided into YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x according to the width of the feature map and the depth of the network. The deeper the network, the longer the training and detection time, but the better the training results. In this paper, YOLOv5l is chosen as the training and detecting model, the training time of YOLOv5l model is longer, but it can guarantee a good training result.

2.1.2 Evaluating indicator

For a target detection model based on deep learning methods, there are usually loss function metrics to evaluate the training degree of the model, and detection metrics to evaluate the detection effectiveness of the model. YOLOv5 uses GIOU (Generalized Intersection Over Union) as the

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028

network loss function. For an image, let the presence position of the object be A and B be the detection box of the model, and draw a minimum box C that can contain A and B. Then the GIOU is calculated as follows:

$$IOU = \frac{|A \cap B|}{|A \cup B|}$$

$$GIOU = IOU - \frac{|C \backslash (A \cup B)|}{|C|}$$

The metrics used to assess the effectiveness of the test are usually Precision, Recall, and AP.

1.Precision:Percentage of correct targets among detected targets.

2.Recall:Ratio of the number of correct targets detected to the number of all targets in the label.

3.AP:Using Recall value as the horizontal axis and Precision value as the vertical axis, the PR curve is obtained, and AP is the area contained in the curve, which is the most commonly used evaluation indicator in target detection.

2.2 Underwater image enhancement and recovery algorithm

In this paper, six typical underwater image enhancement and recovery algorithms mentioned in the literature [12] are selected to enhance underwater marine products images.

2.2.1 DCP algorithm

The DCP (Dark Channel Prior) algorithm was originally proposed by He et al[14] for the recovery of fogged images on the ground, and based on the similar characteristics of light propagation in water and in fog, image recovery can be achieved to some extent by performing a defogging operation on underwater images.

2.2.2 UDCP algorithm

On the basis of DCP, Drews et al[15] proposed UDCP (Underwater Dark Channel Prior) specifically for underwater images, which takes into account the fact that red light decays much faster than blue and green light when light propagates in water. It eliminates the effect of red light, and further enhances the effectiveness of the DCP algorithm.

2.2.3 RGHS algorithm

RGHS (Relative Global Histogram Stretching) is an underwater image enhancement method based on relative global histogram stretching with adaptive parameter acquisition proposed by Huang et al[16]. The algorithm first corrects the input image for contrast by relative global histogram stretching, and then transforms the image color model to correct the color by adaptive stretching, thus achieving underwater image enhancement.

2.2.4 UCM algorithm

UCM (Unsupervised color Correction Method) was proposed by Kashif et al[17],which is an unsupervised color correction algorithm that can effectively remove blue shadows from images, solve the problem of severe red light attenuation and low illumination, and thus complete the enhancement of underwater images. The algorithm is divided into three stages:

1.Regenerate the intensity values of the light in the scene by calculating the losses in the red, green and blue color channels as the light propagates through the water and compensating for them, assuming that the objects in the image are all at the same depth.

2.Contrast correction of the RGB (Red, Green, Blue) color model of the image.

3.Contrast correction of the HSI (Hue, Saturation, Intensity) color model of the image.

2.2.5 ULAP algorithm

The ULAP (Underwater Light Attenuation Prior) algorithm is a fast depth-of-field estimation model algorithm for underwater images based on the underwater light attenuation prior proposed by Song et al[18]. The algorithm calculates the depth-of-field map of an image based on the principle that red

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028

light decays fastest underwater and blue-green decays slowly. Then the background light and transmission maps of the image are estimated from this to recover the underwater real scene image.

2.2.6 RD algorithm

The RD (underwater image enhancement method based on Rayleigh Distribution) algorithm is an underwater image enhancement method based on Rayleigh distribution stretching proposed by Abdul et al[19], which enhances the contrast and visibility of the image by quality of underwater images and also reduces the noise in the images. The algorithm is divided into two stages: contrast correction and color correction.

3. Underwater marine products image enhancement detection and comparison

3.1 Environment of the experiment

The software environment used in this experiment includes Windows 10 operating system, CUDA 10.0, CUDNN 7.0, Pytorch 1.6, Python 3.8; the hardware environment includes Intel core i7-8700 CPU with 3.2GHz, six cores and twelve threads, NVIDIA GeForce RTX2080 (8G memory) graphics card, 32G of memory space.

3.2 Model training

In this paper, we use YOLOv5l as a training and detecting model, our dataset comes from the Underwater Robot Professional Contest (URPC) official website. We used the videos and images provided by the website to manually annotate the images into a dataset with four categories of marine targets, namely echinus, holothurian, scallop and starfish. The dataset contains a total of 2901 images and corresponding annotation files, which are randomly divided into training set and validation set according to the ratio of 8:2.

In this paper, the model was trained with 1200 epochs using the above dataset, that is, all the training samples were forward and backward propagated 1200 times in the neural network, which took a total of 29.1 hours. From the loss function GIOU of the network in Figure 1, it can be seen that the neural network has basically converged.

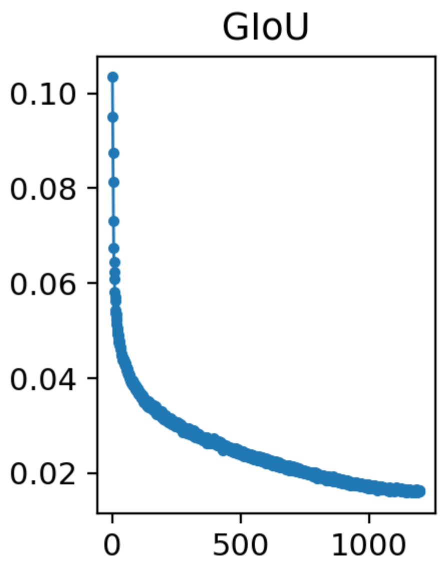
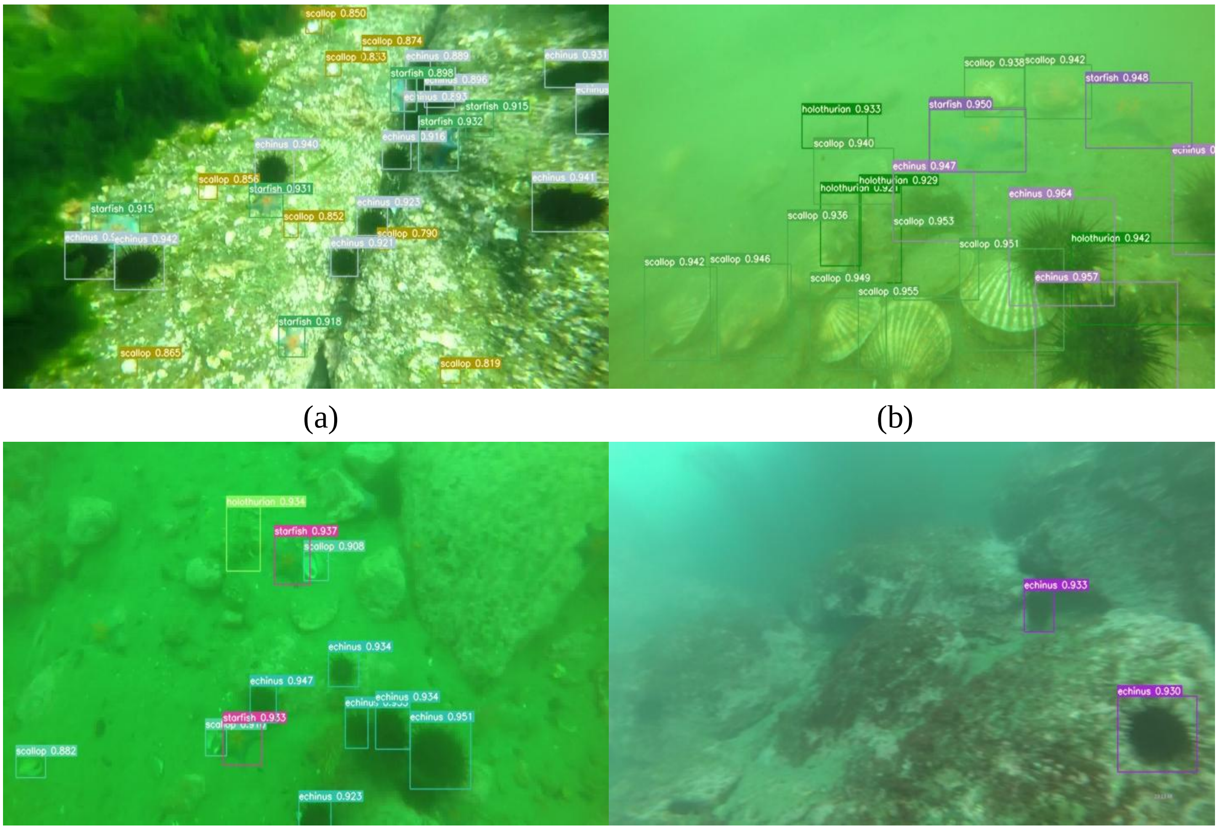


Figure 1. GIOU Loss

Figure 2 shows the detection results of some underwater marine products video images using this neural network model. It can be seen that the model has good detection effect and high confidence for clear images, but for some fuzzy images, there will be some missed detection and wrong detection.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028



(a) (b) (c) (d)

Figure 2. Detection effect of the model

3.3 Image enhancement detection and comparison

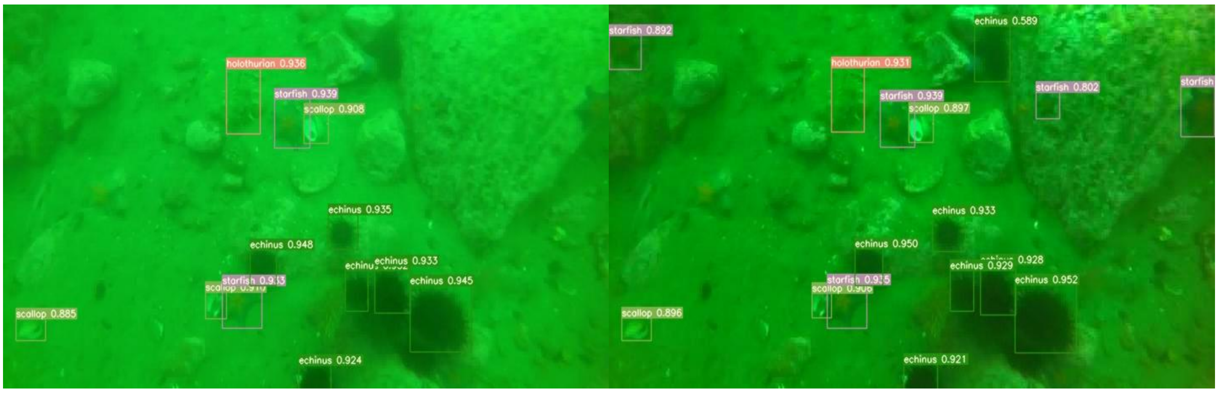
For some blurred images, this paper takes Figure 2(c) with green as the main background color as an example and performs the six image enhancement processes introduced above for them respectively, and uses the trained YOLOv5 model to detect the single image, and the results are shown in Figure 3. By using different methods to enhance the image, the comparison reveals that each enhancement method has a different focus on the image processing. In general, the clarity of the images were all improved, and the background information and marine products in the figure became more obvious and easy to identify. The detection models all have different detection effects on the images, and the performance indicators of the detection effects of single images are shown in Table 1.

As can be seen from Figure 3 and Table 1, the images processed by the UDCP and ULAP algorithms become clear, the detection model can detect more targets, and the mAP is effectively improved; for the images processed by the DCP, RGHS and UCM algorithms, the detection algorithm can detect more targets in some images, but it will miss some of the original targets, which has little impact on the performance of the detection model. And for the images processed by the RD algorithm have produced color distortion, which instead reduces the effectiveness of the detection model. Obviously, for this image environment, UDCP and ULAP algorithms can effectively improve the detection effect of the detection algorithm.

For effective comparison, the same image enhancement process was again performed and detected in this paper for Figure 2(d) with a predominantly blue background, and the results are shown in Figure 4 and Table 2.

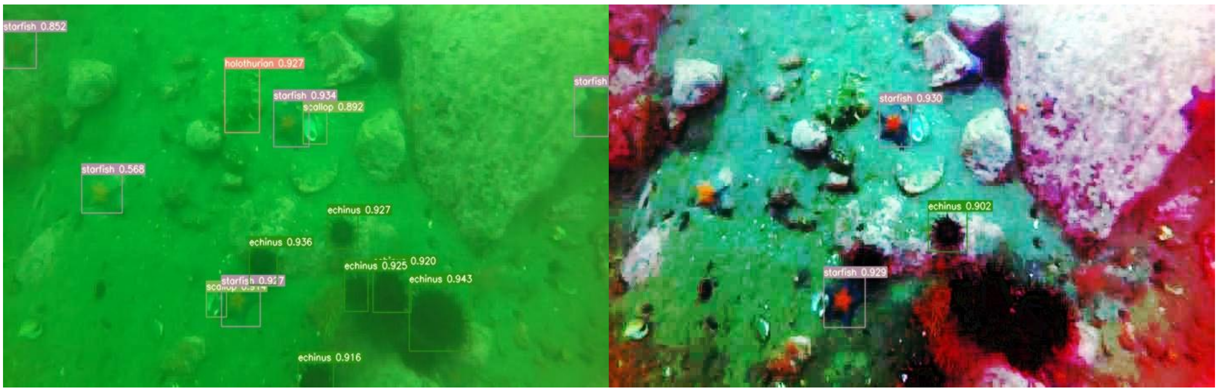
By observing Figure 4 and Table 2, it can be concluded that in this one image environment, the detection model is able to detect more targets for the images processed by DCP, RGHS, ULAP and RD algorithms, with RGHS showing the best results. While for the other two algorithms, UDCP and UCM, the detection performance of the detection model is reduced.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028



(a) DCP (b) UDCP

(c) RGHS (d) UCM



(e) ULAP (f) RD

Figure 3. Comparison of the detection effect of enhanced underwater image with green background

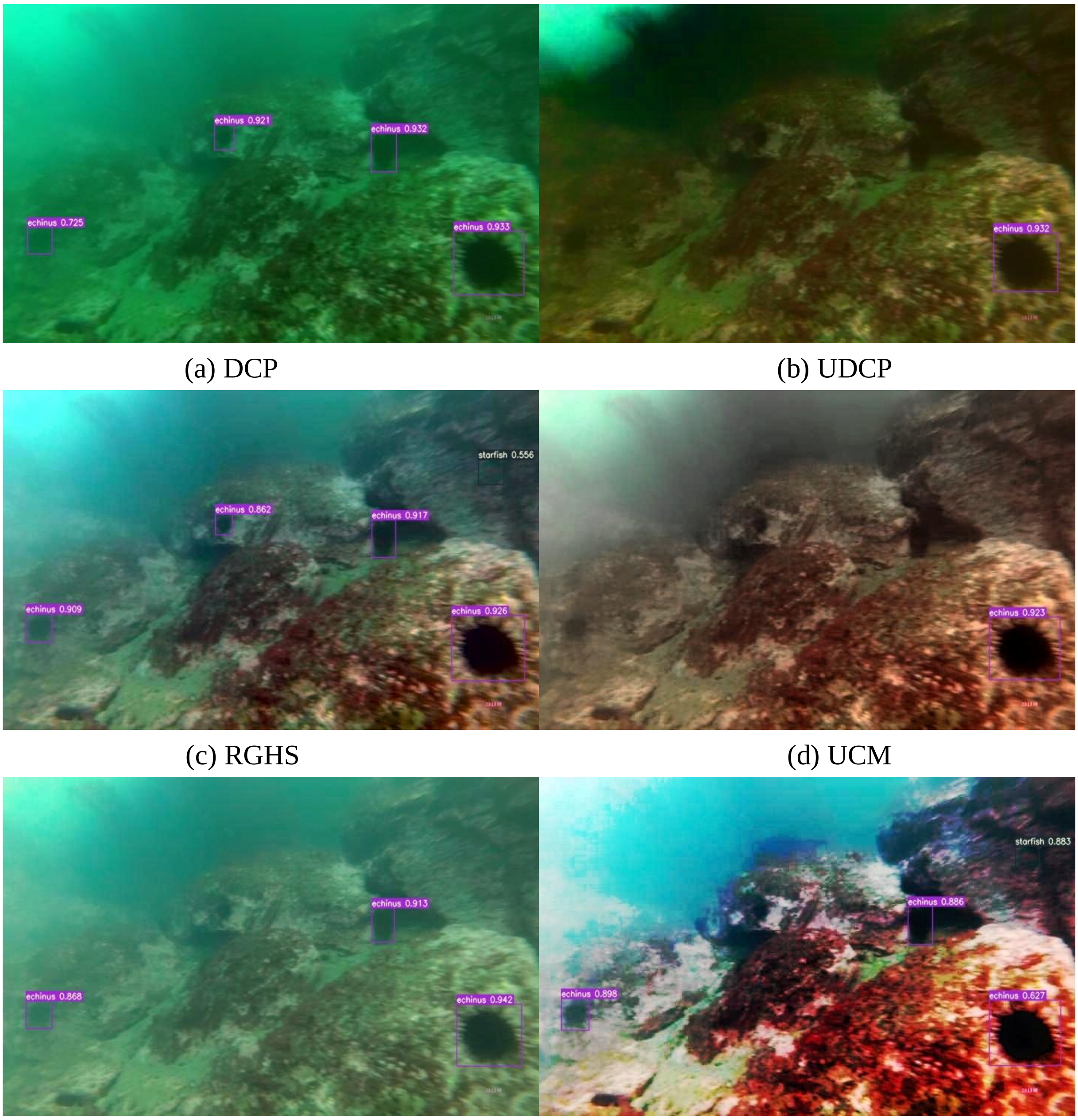
Table 1. The AP of model detection

Category	Original	DCP	UDCP	RGHS	UCM	ULAP	RD
Echinus	0.667	0.667	0.667	0.444	0.556	0.667	0.111
Holothurian	0.500	0.500	0.500	-	0.250	0.500	-
Scollop	0.750	0.750	0.750	-	0.250	0.500	-
Starfish	0.286	0.286	0.714	0.714	0.571	0.714	0.286
mAP	0.551	0.551	0.658	0.579	0.407	0.595	0.198

From the above, it can be concluded that the detection model is able to detect more targets effectively for the image processed by using the UDCP algorithm for Figure 2(c), while for Figure 2(d), the detection performance of the detection model is reduced for the image processed by the UDCP algorithm instead. Similarly, the detection performance of the detection model is degraded for Figure

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028

2(c) after the RD algorithm processing, while in Figure 2(d), the image processed by the RD algorithm is able to be detected with more targets. Therefore, the same image enhancement algorithm cannot be applied to all images with different background environments, but in general, there is always an image enhancement algorithm that can effectively improve the detection algorithm for images with different backgrounds, and Figure 5 and Table 3 show the comparison of the detection results before and after the processing of three consecutive frames.



(a) DCP (b) UDCP

(c) RGHS (d) UCM

(e) ULAP (f) RD

Figure 4. Comparison of the detection effect of enhanced underwater image with blue background

Table 2. The AP of model detection

Category	Original	DCP	UDCP	RGHS	UCM	ULAP	RD
Echinus	0.4	0.8	0.2	0.8	0.2	0.6	0.6
Starfish	-	-	-	1	-	-	1
mAP	0.4	0.8	0.2	0.9	0.2	0.6	0.8

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028

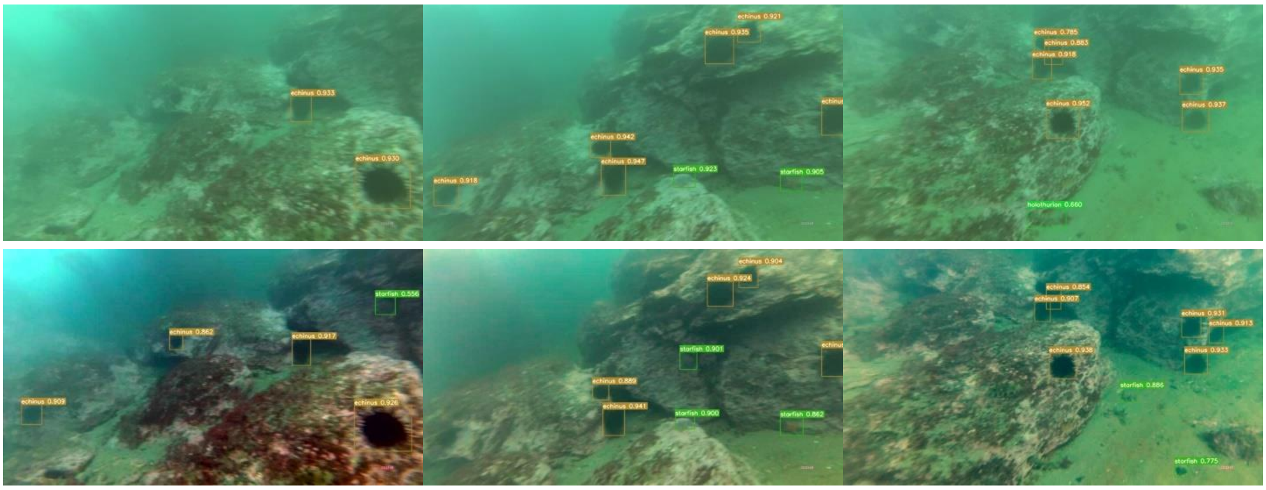


Figure 5. Comparison of the detection effect of three consecutive frames

Table 3. The AP of model detection

	Original	Improved
mAP	0.612	0.845

4. Conclusion

In this paper, six typical underwater image enhancement and recovery algorithms are used to enhance and recover underwater marine products images in different color background, and the images processed by the algorithms are used as inputs for underwater marine products detection using the trained YOLOv5 neural network model. Overall, for the same image, the images obtained by using different underwater image enhancement algorithms have different effects on the target detection model, some algorithms can significantly and effectively improve the performance of the target detection model; and for the same underwater image enhancement algorithm, the enhanced images with different background environments have different effects on the target detection model. Generally, use the suitable image enhancement recovery algorithms can effectively improve the target detection accuracy based on deep learning models, but a universal and unified underwater image enhancement algorithm is lacking. In practical applications, the impact of multiple underwater image enhancement methods on the target detection algorithm needs to be compared in order to obtain the best detection results.

References

[1] OLMOS A, TRUCCO E. Detecting man-made objects in unconstrained subsea videos[C]. IEEE International Conference on British Machine Vision. British Publisher: BMVA Press, 2002:517-526

[2] Yi-Hao Hsiao, Chaur-Chin Chen, Sun-In Lin, et al. Real-world underwater fish recognition and identification, using sparse representation[J]. Ecological Informatics (S1574-9541), 2014, 23:13-21.

[3] FABIC J N, TURLA I E, CAPACILLO J A, et al. Fish Population Estimation and Species Classification from Underwater Video Sequences Using Blob Counting and Shape Analysis[C]. Underwater Technology Symposium, 2013: 1-6

[4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE conference on Computer Vision and Pattern Recognition, 2014: 580-587.

[5] Girshick R. Fast r-cnn[C]. IEEE International Conference on Computer Vision (ICCV), 2015: 1440-1448.

[6] 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 and Machine Intelligence (S0162-8828), 2017, 39(6): 1137-1149.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0028

[7] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[8] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]. European Conference on Computer Vision, 2016: 21-37.

[9] Yang, H., Liu, P., Hu, Y. et al. Research on underwater object recognition based on YOLOv3[J]. Microsyst Technol (S1432-1858), 2020.

[10] Wang, X., Ouyang, J., Li, D. et al. Underwater Object Recognition Based on Deep Encoding-Decoding Network[J]. Journal of Ocean University of China (S1993-5021), 2019, 18(2):376–382.

[11] Yang Honghui, Li Junhao, Shen Sheng, et al. A Deep Convolutional Neural Network Inspired by Auditory Perception for Underwater Acoustic Target Recognition[J]. Sensors (1424-8220), 2019, 19(5).

[12] Y. Wang, W. Song, G. Fortino, L. Qi, W. Zhang and A. Liotta, An Experimental-Based Review of Image Enhancement and Image Restoration Methods for Underwater Imaging [J], IEEE Access (S2169-3536), 2019, 7: 140233-140251.

[13] William. Understand Yolo V5 and Yolo V4 in one article. (2020-07-28)[EB/OL]. https://zhuanlan. zhihu. com/p/161083602.

[14] He K, Sun J, Tang X. Single Image Haze Removal Using Dark Channel Prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S1063-6919), 2011, 33(12): 2341-2353.

[15] P. Drews Jr, E. do Nascimento, F. Moraes, S. Botelho and M. Campos. Transmission Estimation in Underwater Single Images[C]. 2013 IEEE International Conference on Computer Vision Workshops, Sydney, NSW, 2013, pp. 825-830.

[16] D. Huang, Y. Wang, W. Song, J. Sequeira, and S. Mavromatis. Shallow water image enhancement using relative global histogram stretching based on adaptive parameter acquisition[C], MultiMedia Modeling. 2018, pp. 453–465.

[17] K. Iqbal, M. Odetayo, A. James, Rosalina Abdul Salam and Abdullah Zawawi Hj Talib. Enhancing the low quality images using Unsupervised Colour Correction Method[C]. 2010 IEEE International Conference on Systems, Man and Cybernetics, Istanbul, 2010, pp. 1703-1709.

[18] Song, Wei, Wang, Yan, Tjondronegoro, Dian. A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration[C]. Advances in Multimedia Information Processing – PCM 2018, pp. 678–688.

[19] A. S. A. Ghani and N. A. M. Isa. Underwater image quality enhancement through composition of dual-intensity images and Rayleigh-stretching[J]. SpringerPlus (S2193-1801), 2014,3(1):757.