2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP) | 978-1-6654-7857-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSP54964.2022.9778430

COTS recognition and detection based on Improved YOLO v5 model

Yufeng Jiang*,
Hong Kong Metropolitan University,
Hong Kong, China,
jyf05152018@163.com,

Abstract—**At present, the ecological environment of the Great Barrier Reef is becoming more and more fragile. Stopping the propagation and spread of COTS is an important part of protecting the environment of the Great Barrier Reef. It is becoming more and more important to identify and detect the distribution of COTS. With the development of computer science, deep learning technology has been widely used in the field of image recognition. Based on YOLOv5 algorithm and WBF model, this paper constructs a more accurate and efficient detection model to frame the distribution position of COTS. Our algorithm has been verified on the KAGGLE platform. The results show that our algorithm has great advantages in detection performance compared with other detection models. Quantitatively, the F_2 value of our model is 39.241% and 5.263% higher than that of Faster R-CNN and YOLO v5 algorithms, respectively, which provides a certain reference for the protection of the Great Barrier Reef.**

Index Terms—COTS, YOLO v5, WBF, Image detection

I. INTRODUCTION

The Great Barrier Reef is the largest coral reef group in the world，which has been listed as a natural heritage in 1981. There are not only 400 kinds of corals, 1500 kinds of fish and 4000 kinds of molluscs, but also a variety of endangered animals. However, in recent decades, many species and ecosystems of the Great Barrier Reef have suffered serious damage from human environmental destruction, including pollution from coastal development, agricultural run-off and ocean acidification. Fragile ecosystems could lead to the extinction of this species, coral reefs [1]. In order to protect the long-term existence of the Great Barrier Reef, some scholars pointed out that stopping the propagation and spread of crown-of-thorns starfish (or COTS for short) is an important part of restoring the health of the Great Barrier Reef. Scientists and relevant government personnel have formulated a series of intervention plans to control COTS outbreaks at the level of sustainable development [2].

In order to effectively control the number of COTS and its impact on the Great Barrier Reef, it is very important to accurately detect the number and location of starfish. As a traditional method of underwater target recognition and detection, Manta tow is mainly used by divers to evaluate the number and distribution of starfish subjectively and visually under the traction of the ship, which has obvious limitations, including data resolution, data reliability and so on. Sonar image is also a common method to detect underwater targets [3], but there are still some defects, such as cumbersome steps, low efficiency and relying on artificial experience, leading to considerable errors in the detection of small starfish [4]. Therefore, it is still difficult to accurately detect the distribution and number of starfish.

With the continuous development of science and technology, the application of large depth and high-definition underwater computer enables us to obtain enough clear underwater pictures to detect the number and distribution of COTS [5]. At the same time, combined with the deep learning algorithm, the original image data is used for training, and the complex original data features are automatically extracted, which avoids the dependence of traditional methods on subjective experience and solves the difficulty of processing a large amount of data [6]. Based on the image data provided by KAGGLE platform, this paper uses YOLOv5 target detection technology and WBF fusion model to detect the number and distribution of starfish in real time and accurately. You only look once (YOLO) is a state-of-the-art, real-time object detection system. The results show that YOLOv5 algorithm can accurately detect the position of starfish, and the WBF model can further improve the accuracy of the algorithm. Besides, compared with other algorithms, our algorithm has the highest accuracy, such as faster R-CNN.

The main innovations and contributions of this paper can be summarized as follows:

- Aiming at the problem of underwater target recognition and detection, we proposed a YOLO V5 algorithm, and used WBF technology to optimize and improve the model, which can effectively detect the distribution of COTS.

- We verified our proposed model based on underwater videos of core reefs provided by the KAGGLE platform. The results show that our algorithm has better performance than Faster RCNN algorithm, and WBF fusion technology also improves the detection accuracy of YOLO V5 algorithm. It also provides some valuable references for researchers to protect the Great Barrier Reef.

A. Related Work

In this paper, the detection of COTS can be regarded as the application of image target recognition technology. Generally speaking, traditional image recognition technology mainly includes three parts: image preprocessing, feature extraction and classification and recognition. Usually, researchers preprocess the image by filtering and denoising, then extract the important features, and finally classify and recognize the content of the image by using machine learning algorithms such as decision tree, support vector machine and neural network. For example, Pan *et al* [7] used simulated annealing algorithm to optimize machine parameters, improved decision tree SVM algorithm to a certain extent, and obtained good classification accuracy, but with longer training time. Chen extracted the typical features of fish and frogman in sonar images, and used BP neural network to classify and recognize them. The results show that BP neural network has high accuracy in the classification of typical features [8].

The proposal of object detection method based on candidate box has made a milestone leap in the development of object detection, such as R-CNN, Fast R-CNN and so on [9], due to their good performance. R-CNN algorithm draws lessons from the idea of sliding window, uses selective search to divide the picture into a large number of candidate regions, and then trains and classifies each region through CNN and SVM algorithm. The algorithm uses the depth neural network to extract the features in the image instead of the traditional feature extraction, so as to obtain more accurate detection results [10]. However, there are still some shortcomings, such as low efficiency and long time-consuming. Fast R-CNN algorithm is an improvement of R-CNN algorithm, which improves the speed of target detection, but it is still difficult to achieve real-time detection [11]. The Faster R-CNN algorithm is further optimized. The RPN network is used to replace the selective search, which greatly improves the detection efficiency and reaches 17fps per second [12].

Some researchers believe that YOLOv5 model is a promising target detection model, both in running speed and detection performance [13]. Ma et al. [14] constructed a lightweight target detection model based on YOLOv5 algorithm, which can well complete the task of underwater treasure distribution detection. In addition, some studies have shown that YOLOv5 algorithm has the advantages of high accuracy and fast speed in small target detection, and there are few problems of missed detection or poor detection effect [15]. Therefore, this paper also constructs a COTS detection model based on YOLO v5 algorithm, and compares it with faster R-CNN and other algorithms to verify the effectiveness and superiority of our model.

The contents of other chapters of this paper are arranged as follows. In the following section, the algorithm used in this paper is introduced in detail, and the setting of relevant parameters of the model is given; In the third chapter, the data set in our experiment is introduced and the detection results are deeply analyzed; Finally, the main conclusions of this paper and the prospect of future work are described.

II. YOLO v5 algorithm

YOLO is an acronym for 'You only look once', which shows the ambition of the authors. As a family of object detection architectures and models, YOLO has important significance in both theory and practice. YOLO v5 has some similarities with its predessors. For instance, the architecture of it can still be devided into four parts, which are input, backbone, neck, and prediction. The main improvements of YOLO v5 lie in these four parts. In the input part, Mosiac data augmentation is used. Meanwhile, operations for data augmentation like random scaling, random cropping, and random arrangement show good benefits for the final results. Besides, YOLO v5 adds auto anchor box algorithm and adaptive black-edge padding in input part. The most important improvements of YOLO v5 lie in the backbone part. The focus structure and CSP are added to improve the feature extraction. In the neck part, the FPN and PANet are used. And the CSP2 structure are used in the neck part of YOLO v5, which is inspired by CSPnet. The training loss function of YOLO v5 is consist of CIOU loss for bounding box, classification loss and object detection loss. CIOU loss is written as:

$$CIOU_{loss} = 1 - \left(IOU - \frac{d_2^2}{d_c^2} - \frac{v^2}{1-IOU+v}\right), \quad (1)$$

where v is written as:

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w_{gt}}{h_{gt}} - \arctan\frac{w_p}{h_p}\right) \quad (2)$$

where w_{gt} and h_{gt} represent the width and height of the ground truth bounding box. w_p and h_p mean the width and height of the predicted bounding box. The overview of YOLO v5 is shown in Figure 1.

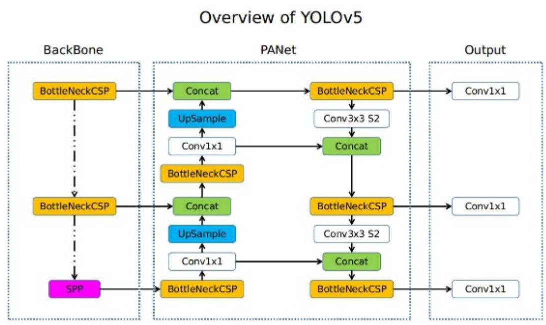

Figure 1. The overview of YOLO v5.

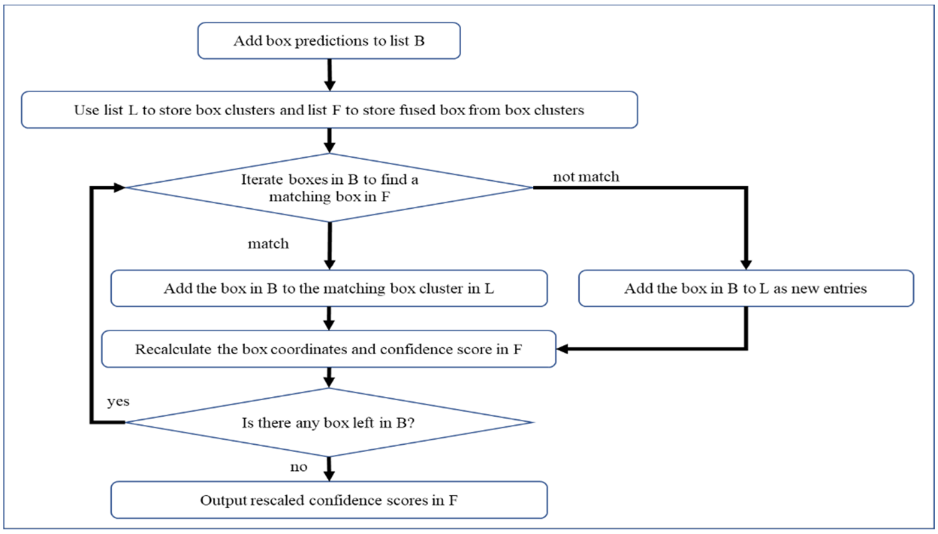

Figure 2. Overview of WBF.

WBF, standing for weighted boxes fusion, is a novel method to predict the final object box using multiple box predictions. Unlike non-maximum suppression and soft non-maximum suppression, WBF uses all box predictions rather than excluding some boxes. The overview of WBF is shown in Figure 2.

III. EXPERIMENTS

In this section, we gave the results of COTS detection and compared the performance of several different algorithms in the task of detecting COTS distribution.

First, let's introduce the data set used in this paper.

a) data set

The data set used in this paper is provided by the KAGGLE platform and contains more than 35000 underwater images taken at different times and places around the Great Barrier Reef. These images are divided into training set and test set, which randomly contains about 23,500 images and 13,000 images respectively. The difference is that there are some special annotations in the data of training set, "x", "y", "width" and "height", which are used to frame the position of COTS. Then, the deep learning method is used to extract and learn the features of these positions, and used to detect the distribution of COTS in the test set data.

b) detection results

In this section, we first give the schematic diagram of detection results of different algorithms, and qualitatively compare the performance of several different algorithms，as shown in Figure 3 and Figure 4.

WBF is significantly higher than that of YOLOv5 algorithm and fast algorithm alone, and almost all positions of starfish are framed, which verifies the superiority of our algorithm.

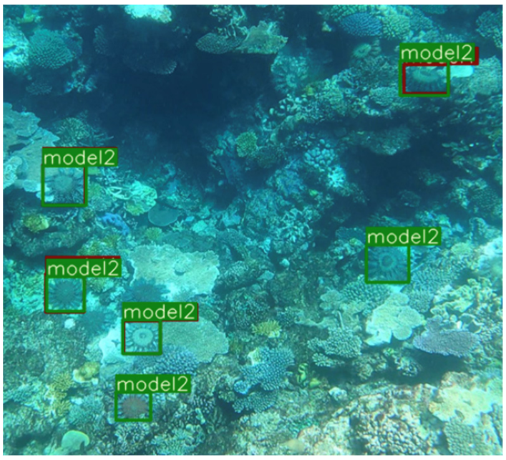

Figure 3. Comparison of YOLOv5 and Faster R-CNN detection results

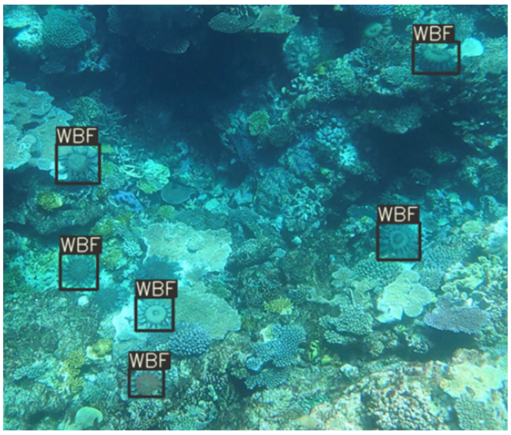

Figure 4. Detection results of YOLOv5 with WBF

From the above two figures, we can see that it is difficult to completely detect the distribution of indistinguishable COTS by YOLOv5 algorithm or Faster R-CNN algorithm alone. Evidently, the detection accuracy of YOLOv5 model optimized by *c)* Quantitative analysis of experimental results

For the model of image recognition and detection, Precision and recall rates are two commonly used performance indicators. In general, the higher the two indicators, the better the detection effect. But in most cases, these two parameters restrict each other. Therefore, we need to comprehensively weigh these two indicators to get more objective evaluation. At the same time, COTS detection belongs to general search task, and the importance of precision is greater than that of recall. So, in this paper, we choose F_2 index to comprehensively measure the performance of our algorithms, which can be written as:

$$F_2 = \frac{5}{\frac{4}{Precision} + \frac{1}{Recall}} \quad (3)$$

The higher the F_2 value, the better the detection performance of the model. Then, we give the F_2 values of several algorithms, as shown in Table 1.

Table 1. F_2 values of different algorithms

Models	F2
YOLO v5	0.627
Faster RCNN	0.474
YOLO v5 with WBF	0.660

Obviously, our algorithm has the highest F_2 value, which shows that our model has the best performance than other algorithms. Specifically, the F_2 value of YOLOv5 algorithm with WBF model is 39.241% and 5.263% higher than that of Faster R-CNN algorithm and YOLOv5 algorithm, which can enable us to detect the distribution of COTS more effectively and provide some constructive suggestions.

IV. CONCLUSIONS

Stopping the further propagation and spread of COTS is an important part of restoring the health of the Great Barrier Reef. Identifying and detecting the distribution of COTS has become a top priority. Based on a large number of underwater images provided by KAGGLE platform, combined with depth learning algorithm, in this paper, we realize the high-precision and efficient detection of COTS. In this paper, YOLOv5 algorithm optimized by WBF model is used to detect the distribution of COTS. Compared with other algorithms, the superiority of our proposed model is verified. Specifically, the F_2 value of our proposed model is 39.241% and 5.263% higher than that of Faster R-CNN and YOLOv5 algorithms, respectively. Then, we will collect more data and continue to optimize the algorithm to achieve better detection performance.

REFERENCES

[1] Bellwood D R , Hughes T P , Folke C , et al. Confronting the coral reef crisis[J]. Nature, 2004, 429(6994):827-833.

[2] Randall R H . Coral reef recovery following extensive damage by the Crown-of-thorns starfish, Acanthaster plancii (L.)[J]. Publications of the Seto Marine Biological Laboratory, 1973, 20:469-489.

[3] Ma Y , Mao Z Y , Qin J , et al. A Quick Deployment Method for Sonar Buoy Detection Under the Overview Situation of Underwater Cluster Targets[J]. IEEE Access, 2019, PP(99):1-1.

[4] Kumar R , Thakur N , Thakur V . An Overview Of Sonar Andacoustic Underwater Communication[J]. International Journal of Advanced Research in Electrical Electronics & Instrumentation Engineering, 2013, 2(5).

[5] Rowshyra A. Castañeda and Olaf L.F. Weyl and Nicholas E. Mandrak. Using occupancy models to assess the effectiveness of underwater cameras to detect rare stream fishes[J]. Aquatic Conservation: Marine and Freshwater Ecosystems, 2020, 30(3) : 565-576.

[6] Zeng W-G. Sonar image target recognition and detection based on convolutional neural network[D]. DaLian University of Technology, 2019.

[7] Pan Y. Image Recognition Method Based on Improved Decision Tree SVM[J]. Journal of Xinxiang University, 2018.

[8] Chen Q. Research-based Underwater Acoustic Images Underwater Target Recognition[D]. Harbin Engineering University, 2012.

[9] Sreya R C , Kumar V V . A Review on Instance Segmentation Using Mask R-CNN[J]. Social Science Electronic Publishing.

[10] Girshick R , Donahue J , Darrell T , et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[J]. IEEE Computer Society, 2013.

[11] Girshick R . Fast R-CNN[J]. Computer Science, 2015.

[12] 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, 2017, 39(6):1137-1149.

[13] He Wang, Song Zhang, Shili Zhao, et al. Real-time detection and tracking of fish abnormal behavior based on improved YOLOV5 and SiamRPN++ - ScienceDirect[J]. 2021.

[14] Ma Z Q. Sea Treasure Target Detection Algorithm Based on Improved YOLOv5[J]. Modern Informationn Technology, 2021, 5(18):80-85.

[15] Jiang L, Cui Y. Small Target Detection Based on YOLOv5[J]. Computer Knowledge and Technology, 2021, 17(26):131-133.

[16] Roman S., Weimin W., Tatianna G. Weighted boxes fusion: Ensembling boxes from different object detection models. Arxiv, 20