Abnormal Behavior Fish and Population Detection Method based on Deep Learning

Zexin Zhao *

College of Electronic and Information Engineering, Guangdong Ocean University, Zhanjiang 524000, China
* **Corresponding author Email**: 2112210024@stu.gdou.edu.cn

Abstract: This paper presents a detection model of fish with abnormal behavior and their number based on YOLO v8 and Deep Sort algorithm. The method firstly uses computer and acquisition system to monitor and analyze the fish behavior in real time, and can effectively detect the abnormal behavior of fish, such as abnormal swimming trajectory and abnormal residence time. The main work of this paper is to preprocess fish behavior videos, including video segmentation, data enhancement and other operations, and use data enhancement technology to improve the problem of fish occlusion in data set, which is easy to cause model false detection. Then, YOLO v8 and Deep Sort algorithm were used for multi-target tracking and target detection to extract the key information of fish behavior. Finally, through the analysis and comparison of the extracted information, the detection of fish with abnormal behavior and its quantity are realized. The experimental results show that the method proposed in this paper can effectively detect the abnormal behavior of fish, has high accuracy and real-time, and has certain application and popularization value.

Keywords: Deep Learning; Target Detection; Deep Sort; YOLO v8; Fish Abnormal Behavior Recognition; Computer Vision.

1. Introduction

With the destruction of natural environment and the continuous development of fishery production, the detection of abnormal behavior of fish becomes more and more important. Abnormal fish behavior reflects fish disease and changes in water environment. Counting fish with abnormal behavior can accurately judge hatching rate and mortality. It plays an important role in promoting aquaculture and water environment monitoring. Traditional fish abnormal behavior detection methods mainly rely on manual labeling and manual feature extraction, which has some problems such as inaccurate labeling and difficult feature extraction.

Since machine learning methods for fish detection in underwater images are inexpensive and easy to operate, Therefore, it has been widely used in fish species detection [1-5], fish monitoring [6-7], fish sorting [8], fish behavior research [9], fish habitat investigation [10], fish school analysis [11-12] and fish population statistics [13-14]. Therefore, in recent years, fish abnormal behavior detection methods based on deep learning have received extensive attention and research.

Object detection is the detection of the location and number of specific objects in an image or video. Object detection has a broad application prospect in industrial detection, unmanned driving, medical image processing and other fields [15]. Deep learning can achieve efficient and accurate target detection by constructing deep neural network models. At present, the main method of deep learning in the field of object detection is the method based on convolutional neural network (CNN). CNN is a special kind of neural network that can automatically learn features in images and use them for target detection. Through multi-layer convolution and pooling operation, CNN can gradually extract high-level features in the image, so as to achieve target detection. In object detection, common deep learning models include Faster R-CNN, YOLO, DETR, SSD, etc. These models are all based on CNN, but their structures and algorithms are different. At present, single-stage detection method and two-stage detection method are the two most widely used methods. Faster R-CNN is a two-stage object detection method that first generates candidate regions through a region extraction network (RPN), and then classifiers and regressions these candidate regions. YOLO and SSD are single-stage object detection methods that directly classify and regression the entire image to achieve object detection.

Researchers at home and abroad have done a lot of meaningful work in this area, but there are still many challenges and difficulties. Redmon et al proposed the YOLO algorithm [16] in 2016, which transformed target detection into a regression problem and improved the detection speed [17]. Wang et al. [18] introduced Dense Net into YOLO v5, increased the number of Bottleneck structures, compressed the number of parameters in the model, and significantly improved the detection speed [19]. In order to better solve these problems, this paper will introduce a fish abnormal behavior detection method based on YOLOv8 algorithm. YOLOv8 algorithm is an efficient target detection algorithm with high accuracy and fast speed, which can be effectively applied in the field of fish abnormal behavior detection. This paper will introduce the principle and implementation method of YOLOv8 algorithm in detail, and verify its effectiveness and superiority in fish abnormal behavior detection through experiments.

2. Model Introduction and Related Work

2.1. Algorithm Introduction

2.1.1. Yolo v8

YOLOv8 algorithm is an object detection algorithm based on deep learning. Compared with the previous version, Yolov8 has significantly improved in speed and accuracy.

YOLOv8 is a SOTA model that builds on the success of previous YOLO versions and introduces new features and improvements to further improve performance and flexibility.

Specific innovations include a new backbone network, a new Ancher-Free detection header, and a new loss function that can run on a variety of hardware platforms from cpus to Gpus. Backbone still uses the CSP idea, but the C3 module in YOLOv5 has been replaced with the C2f module to achieve further lightweight, while YOLOv8 still uses the SPPF module used in YOLOv5 and other architectures. YOLOv8 still uses the idea of PAN, but by comparing the structure diagram of YOLOv5 and YOLOv8, it can be seen that YOLOv8 deleted the convolution structure of CBS 1*1 in the up-sampling stage of PAN-FPN in YOLOv5 and replaced the C3 module with the C2f module. YOLOv8 uses VFL Loss as classification Loss and DFL Loss + CIOU Loss as classification loss. YOLOv8 uses Task-Aligned Assigner matching instead of IOU matching or single-sided proportional allocation. Compared with spp, simple parallel max pooling is changed to serial + parallel. Compare Figure 1 (SPP on the left, SPPF on the right):

In general, Yolov8 is a very excellent target detection algorithm, which has the advantages of fast speed, high precision and strong robustness, and has been widely used in many practical applications. Therefore, this paper chooses YOLOv8 as the algorithm of fish abnormal behavior detection model.
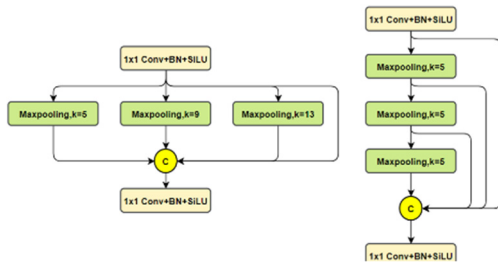

Figure 1. Comparison of SPP and SPPF structures

2.2. Deep Sort Multi-Target Tracking Algorithm

Multi-target tracking is the task of tracking multiple targets simultaneously in a video. At present, deep sort algorithm is mainly popular in industry, which uses the principles of Mahalanobis distance, Hungarian algorithm and Kalman filter. Mahalanobis distance: A method used to calculate the similarity of two unknown sample sets. Hungarian algorithm: The target is associated with the detection frame based on the similarity matrix of the two frames. Kalman filtering: predicts and follows the position of new targets.

The sort tracking algorithm flow is as follows:

1) The detector obtains the check object in the current frame. 2) Using Kalman filter to predict tracking objects in past frames. 3) Calculate the IOU similarity matrix of the predicted bbox of the tracking object and the detected object box. 4) Using the Hungarian algorithm to find the best match. 5) Kalman filter update.

The optimization of deep sort compared to sort is as follows: The appearance matching degree and distance matching degree are increased instead of the IOU matching degree in sort algorithm. The matching cascade method is introduced to carry out feature matching, which improves the tracking effect of targets with occlusions. Reduced target id hopping issues. In deep sort, when frame t is detected, a reid extraction network is used to extract reid features from the detection target obtained by the target detection algorithm. Then, the cascaded matching between the confirmed trace object and

the detection object is performed to obtain the reid feature matching result. The unmatched time_since_update value is 1 and the unconfirmed trace object is matched with the unmatched detection object. The final matching result is obtained, and the matched object is updated by Kalman to get the final bbox.

2.3. Improvement of IOU Loss Function

In the original Yolov8 model, CIOU was used as the loss function of detection frame regression, and the overlapping area, center point distance and aspect ratio of prediction frame and real frame in boundary frame regression were considered

$$\{(W = kWgt, H = kHgt) \ k \ \in R +\}. \qquad (1)$$

Where: W and H are the width and height of the prediction box respectively; Wgt and Hgt are the width and height of the real box, respectively. The aspect ratio in formula (1) only reflects the ratio relationship between the width or height of the prediction box and the real box. When the aspect ratio between different prediction boxes and the real box is the same, the calculation result of CIOU is the same.

To solve this problem, SIOU [23] was used as the detection frame regression loss function in this study. SIOU redefines the penalty index considering the vector Angle between the required regression. SIOU loss (USIO) includes distance loss Δ, shape loss Ω and IOU loss (UIO), and its calculation formula is as follows

$$U_{SIO} = U_{IO} - (\Delta+\Omega) / 2. \qquad (2)$$

3. Experiment

3.1. Prepare the Required Materials

The experimental animals selected for this experiment have strong vitality and can live in Zhanjiang Bay laboratory for half a year (no more than half a month before and after) to ensure the activity and health of the experimental fish. Will not die during the experiment, resulting in the loss of the experiment.

The experiment platform consists of two parts. The first part is a fish pond with water filtration function and oxygen supply function. The water is the tap water of Zhanjiang Bay experiment, and the water quality is in accordance with the municipal water standard. The water temperature is controlled at 19 degrees Celsius (up or down no more than 1 degree Celsius) and the pH value is 7.1 (up or down no more than 0.2). The second part is the information acquisition system, using a network camera, when recording video, the frame rate is 24FPS (1280*720), is fixed in front of the fish pond to shoot the fish pond, the distance from the fish tank is not more than 50cm.

3.2. Data Acquisition

An underwater camera is a device specifically designed for underwater shooting, which can capture high-quality images and videos in underwater environments. For this study, we chose a Hikvision HD underwater camera to ensure the quality of the data we collected. Prior to data collection, we selected an area where fish congregreed to ensure that the data collected was representative. In this study, we selected 120 minutes of continuous video collection and frame extraction, and finally 669 pictures were formed after data cleaning and manual screening.

3.3. Data Preprocessing

For the problem of irregular image shooting, the automatic rotation and flip function of MixUp data enhancement method has been adjusted to achieve the purpose of enhancing image readability. Then, the fragments from the data set were imported into the model training. In order to select some fragments for accuracy evaluation, the experiment divided the data set, validation and test set in a ratio of 7:2:1.

3.4. Data Annotation

Data annotation is also an important step in deep learning object detection. In this study, Labelimg was used for manual annotation. The problem with the experiment is that the fish are very fast and move a lot. So how to understand the behavior of fish is the key in the future data set processing. Therefore, in order to accurately mark which behaviors are abnormal behaviors of fish, teachers and students from the College of Fisheries are invited to give advice to help label behavioral data in the recorded fish behavior snippets. After labeling the fish in each picture and marking the abnormal behavior state of the fish, it is expected that the abnormal swimming of the fish can be more accurately identified. Through the above steps, 669 high-quality fish pictures were successfully selected and labeled. This data will serve as a training set for our deep learning object detection model to enable the identification of anomalous fish behavior.

3.5. Experimental Platform and Model Training Parameters

The experiment adopts GPU acceleration for training. In order to obtain the detection speed, the experiment uses the pre-training weight yolov8s.pt trained by YOLOv8 network on ImageNet. In order to obtain the detection accuracy, the experiment increases the default iteration number to 500 times, and the learning rate is 0.01. The hardware configuration is shown in Table 1 and the software running platform is shown in Table 2.

Table 1. The hardware platform used in the experiment

CPU	Intel（R）Core™ i5-12500 3GHz
GPU	NVIDIA（R）GeForce RTX™ 4060 Ti
RAM	DDR4 16GB
SSD	PCI 512GB

Table 2. Software platform for the experiment

Models	Python 3.9
Experimental running platform	PyCharm
Cuda	12.2
PyTorch framework	2.0.1
Operating system	Windows 10

3.6. Evaluation Index

The computation is used to measure the execution time of the network model, and the unit is 1 billion floating-point operations per second (GFLOP). The larger the value is, the greater the computing power required by the computer. The average precision mean (mAP) was used to evaluate the accuracy of the model

Performance, which is an important performance measurement indicator, that is, to average the average accuracy value of all categories, its calculation formula is

$$P_{mA} = \sum P_A/N \tag{3}$$

In the formula: P_{mA} is mAP; N is the total number of categories; PA is the area enclosed by a curve with recall rate as the horizontal axis and accuracy as the vertical axis. mAP@0.5 is the mean precision for the IOU threshold of 0.5.

3.7. Comparative Test of Detection and Recognition of Different Models

Comparing the improved Yolo v8 model with the original Yolo v8 model and Yolo v5 model, the model proposed in this paper has obvious advantages in terms of detection accuracy and computation amount under this data set. As shown in Table 3

Table 3. Comparative test of detection and recognition of different models

Model	mAP@0.5/%	Calculated amount
Yolo v5	66.93	7.1
Yolo v8	71.36	6.85
Yolo v8-Deep Sort-SIOU	73.49	6.52

3.8. Experimental Result

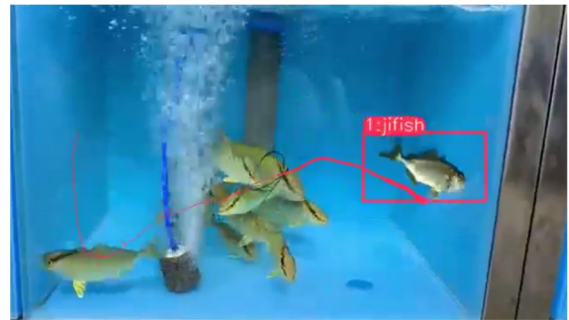**Figure 2.** Original Image

Figure 3. Detection effect picture

Figure 4. Original Image

The abnormal behavior fish detection and counting model based on Deep Sort and Yolo v8 has good detection accuracy and speed, and can accurately and quickly detect fish with abnormal behavior, and can display its number, while normal fish will not be mis detected. For the detection under

occlusion, the detection effect of this model is still good. For both local occlusion and large-area occlusion, this model can detect accurately, as shown in Figure 2-5.
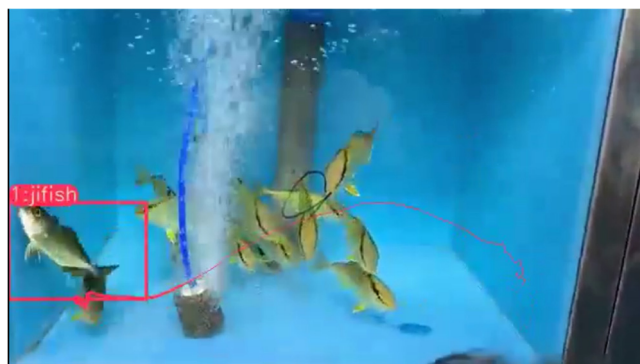

Figure 5. Detection effect picture

4. Conclusion

This paper proposes a fish abnormal behavior recognition method based on deep learning object detection technology. The method uses YOLOv8 algorithm for real-time monitoring and detection of fish behavior, and can quickly and accurately find fish with abnormal behavior. The experimental results show that the detection speed, accuracy and real-time performance of this method are better than the traditional manual recognition method, and it has certain practical value.

However, there are still some shortcomings in this study that need to be improved. First of all, the data set size in this paper is small, so it is necessary to further expand the data set size to improve the generalization ability of the model. Secondly, the impact of the algorithm on environmental factors such as light and water quality needs to be further studied and optimized. In addition, the algorithm in this paper needs to be further improved to recognize the behavior of different kinds of fish.

The future can be improved from the following aspects: First, the algorithm can be further optimized to improve the detection accuracy and robustness. Secondly, it can be combined with other technical means, such as sound recognition, water quality monitoring, optical flow method, etc., to improve the recognition ability of abnormal behavior of fish. In addition, more efficient and refined data acquisition and annotation methods are explored to improve the quality and size of data sets. Finally, the method can be applied to the actual production to realize real-time monitoring and early warning of the smart Marine pasture, and improve the efficiency and economic benefits of aquaculture. The method is even applied to a wider range of fields and scenarios, such as aquatic protection, to meet the needs of different fields.

Finally, the method of fish abnormal behavior recognition proposed in this paper has certain application value and research significance, but it still needs to be further perfected and optimized.

References

[1] STORBECK F, DAAN B. Fish species recognition using computer vision and a neural network[J]. Fisheries Research, 2001, 51(1):11-15.

[2] Yuan Hongchun, Zhang Shuo. Underwater fish target detection method based on Faster R-CNN and image enhancement [J]. Journal of Dalian Ocean University, 2020,35(4):612-619. YUAN Hongchun, ZHANG Shuo. Detection of underwater fish based on Faster R-CNN and image enhancement[J]. Journal of Dalian Fisheries University, 2019, 35 (4):612-619. (In Chinese)

[3] GRAY P C, FLEISHMAN A B, KLEIN D J, et al. A convolutional neural network for detecting sea turtles in drone imagery[J]. Methods in Ecology and Evolution, 2018, 10 (3) : 345-355.

[4] ALSMADI M K, OMAR K B, NOAH S A, et al. Fish recognition based on robust features extraction from size and shape measurements using neural network[J]. Journal of Computer Science, 2010,6(10):1059-1065.

[5] Sun Longqing, Liu Ting, Chen Shuaihua, et al. Research on Multi-target fish object proposal detection Algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(12):260-267. SUN Longqing, LIU Ting, CHEN Shuaihua, et al. Multi-target fish detection algorithm based on object proposals[J].Transactions of the Chinese Society for Agricultural Machinery,2019,50(12):260-267.(in Chinese)

[6] SABERIOON M M,CISAR P. Automated multiple fish tracking in three-dimension using a structured light sensor [J]. Computers & Electronics in Agriculture, 2016, 121 (1): 215-221. (in Chinese)

[7] Yu Bing, Fu Zetian, Zhang Xiaoshuan, et al. Research on Water quality and Environment monitoring system of ornamental fish culture based on WSN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2016,47(1):392-400. HAN Yubing, FU Zetian, ZHANG Xiaoshuan, et al. Monitoring system of ornamental fish culturing water quality based on wireless sensor network[J]. Transactions of the Chinese Society for Agricultural Machinery,2016,47(1):392-400. (in Chinese)

[8] Liu Hu, Gong Yu, Zhang Biao, et al. Design and Optimization of Working Parameters of Fish fry sorting Machine based on EDEM [J]. Transactions of the Chinese Society for Agricultural Machinery, 20,51(1): 114-121.LIU Hu, GONG Yu, ZHANG Biao, et al. Design and optimization of working parameters of fry sorting machine based on EDEM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,51(1):114-121.(in Chinese)

[9] COLAS F, TARDIVEL M, PERCHOC J, et al. The ZooCAM, a new in-flow imaging system for fast onboard counting, sizing and classification of fish eggs and meta zooplankton[J]. Progress in Oceanography,2017,166(1):54-65.

[10] HSIAO Y H, CHEN C C, LIN S I, et al. Real-world underwater fish recognition and identification, using sparse representation [J]. Ecological Informatics,2014,23(1):13-21.

[11] STOBART B, GARCIACHARTON J A, ESPEJO C, et al. A baited underwater video technique to assess shallow-water mediterranean fish assemblages: methodological evaluation [J]. Journal of Experimental Marine Biology & Ecology, 2007345, (2) : 158-174.

[12] SCHMID K, REISFILHO J A, HARVEY E, et al. Baited remote underwater video as a promising nondestructive tool to assess fish assemblages in clearwater Amazonian rivers: testing the effect of bait and habitat type[J]. Hydrobiologia, 2017, 784(1):93-109. (in Chinese)

[13] Zhao Dean, Liu Xiaoyang, Sun Yueping, et al. Underwater Mitten crab recognition Method based on Machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(3):151-158.ZHAO Dean, LIU Xiaoyang, SUN Yueping, et al. Detection of underwater crabs based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery,2019,50(3):151-158.(in Chinese)

[14] QIN Xuebiao, Huang Dongmei, Song Wei et al. Fish detection method based on multiple enhancement and output of fuzzy underwater images [J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(07):243-249.

[15] Dong Gang, Xie Weicheng, Huang Xiaolong et al. Overview of small target Detection algorithms in Deep Learning [J]. Computer Engineering and Applications,2023,59(11):16-27. (In Chinese)

[16] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 779–788.

[17] Research review on YOLO based autonomous driving target detection. Computer Engineering and Applications, 2022, 58 (15): 68-77. (in Chinese) [doi: 10.3778/j.issn.1002-8331. 2203-0310].

[18] WANG C, WANG H, YU F, et al. A high-precision fast smoky vehicle detection method based on improved YOLOv8 network [C]// IEEE International Conference on Artificial Intelligence and Industrial Design (AIID), 2022:255-259.

[19] Yang Feng, Ding Zhitong, Xing Mengmeng et al. An overview of object detection algorithm improvement in Deep Learning [J]. Computer Engineering and Applications, 2023, 59 (11):1-15. (in Chinese)