Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 3, No. 1, 2023

Research on Water Garbage Detection Algorithm Based on GFL Network

Nengmin Yi *, Weilin Luo

Southwest Minzu University, Chengdu 610200, China
* **Corresponding author**: Nengmin Yi

Abstract: Aiming at the problem that existing water surface garbage detection algorithms cannot meet the tradeoff between real-time and accuracy in natural scenes, we propose a GFL (Generalized Focal Loss) based GFL_ HAM algorithm to improve the accuracy of water surface garbage detection, and better apply to water surface garbage detection application scenarios. We have added the HAM (Hybrid Attention Model) module to the feature extraction module (resnet50) of the GFL network to improve the feature extraction capability of the backbone network and the network representation capability. And use the self-constructed garbage detection datasets under natural scenes for training and testing. After analyzing the training process and test results, our network has higher reliability compared to the current mainstream single stage network. Our network can achieve 60.12% mAP on the water surface garbage detection datasets, which is 1.09%, 1.36%, and 1.61% higher than YOLOv3, SSD, and GFL, respectively.

Keywords: Garbage detection; HAM attention mechanism; GFL (Generalized Focal Loss).

1. Introduction

With the development of economy and society, people's living standards are increasingly improving, and intelligent and modern lifestyles have replaced traditional production and life. Diversified production and lifestyle is a double-edged sword, and the rising quality of life has brought serious environmental problems. Changes in production and lifestyle have brought more serious environmental pollution. The Global Environment Outlook 6 [1] mentions that the Earth on which people depend for their survival has suffered extremely serious environmental pollution and damage. If effective and more vigorous actions are not taken to protect the ecological environment, the Earth's ecosystem and the sustainable development of mankind will inevitably be more severely threatened. In the face of the destruction of the ecological environment and the increasing production of garbage, how to effectively and quickly resolve garbage classification, monitoring, and recycling, and improve the ecological environment, is one of the urgent issues that need to be addressed throughout the world and all mankind. From May 1, 2020, the "Beijing Municipal Regulations on the Administration of Domestic Waste" [2] proposed the concept of domestic waste classification, requiring the public to establish a sense of waste classification and actively classify domestic waste, which is conducive to the classified treatment and recycling of waste. It can be seen that the task of garbage classification has become an integral part of garbage treatment and environmental protection. Due to the insufficient awareness of environmental protection among some people, some garbage has been dumped into the water. Failure to clean it up in a timely manner will have a serious impact on people's health and production. The traditional method of water surface garbage removal is manual salvage. This traditional method of water surface garbage removal not only cannot ensure safety and trash removal, but also has high salvage time and economic cost, which is not sufficient to meet the requirements of safety and timeliness of water surface garbage removal. The emergence of surface garbage salvage robots has solved the problems of safety and timeliness in surface garbage salvage the water surface garbage detection algorithm is one of the core technologies of the water surface garbage salvage robot [3], and the research and development of the water surface garbage detection algorithm is closely related to the water surface garbage salvage robot.

In recent years, the research on water surface target detection algorithms has developed vigorously, and a large number of water surface target detection algorithms have emerged. In 2005, Hou [4] et al. proposed a method for detecting significant water surface targets based on change detection background modeling. When the difference in target background changes is small, the detection accuracy of the algorithm will decrease. In 2013, Huang et al[5]. proposed an algorithm for extracting and detecting water objects in natural scenes based on water color and texture features. In 2016, Yang et al. [6] proposed a garbage classification system based on Support Vector Machine (SVM), which uses a sliding window of constant size to extract features from the input image, and then uses a SVM classifier to achieve garbage classification. This achieved a classification accuracy of 63% on a self-built dataset, but it can only achieve garbage image classification and cannot detect the specific location of garbage in the image. In 2021, Wang et al. [7] proposed a garbage detection method based on the two-stage target detection algorithm Faster-rcnn, which can achieve 83.4% mAP (average accuracy) at 6.3FPS and 85.5% mAP (average accuracy) at 6.8FPS on their own dry and wet garbage datasets, respectively. Although the accuracy has met the needs of intelligent garbage detection, timeliness is still a major pain point. In 2022, Xu et al. [8] proposed a lightweight garbage target detection algorithm Ghost-YOLO based on low-power devices. This algorithm ensures lightweight while maintaining high garbage detection accuracy, but while meeting the lightweight requirements, there are certain errors in accuracy, which affects the garbage detection effect. To meet both accuracy and real-time requirements.

Most existing water surface garbage detection algorithms focus on the accuracy or timeliness of detection. However, for water surface garbage detection tasks, it is necessary to achieve a trade-off between accuracy and real-time performance, and to address issues such as low photo clarity and reflection caused by different shooting angles and shooting environments. Therefore, the algorithm not only requires real-time detection, but also needs to be able to eliminate the impact of water surface interference factors to improve detection accuracy. In summary, this study proposes a water surface garbage detection algorithm based on GFL (Generalized Focal Loss) [9] for complex scenes to meet the real-time and effective requirements of water surface garbage detection in complex scenes.

2. Methodology

2.1. Overall construction process of surface waste detection

In order to meet the real-time and accuracy requirements of water surface garbage detection tasks, we propose a knowledge based GFL (Generalized Focal Loss) model based GFL_ HAM water surface garbage detection algorithm. First, construct a water surface garbage detection dataset. Then, using rotation, changing brightness, and changing contrast to simulate the possible impact of the construction process of the water surface garbage datasets, the datasets is enhanced. Next, use the surface garbage detection datasets to train GFL_ HAM network. Finally, we tested our model using real water surface garbage images from natural scenes. The overall construction process of the algorithm is shown in Figure 1.

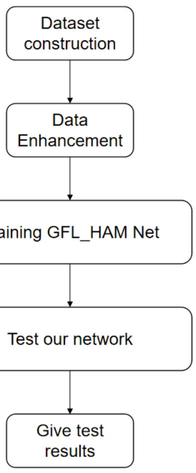

Figure 1. Overall Algorithm Construction Process

2.2. GFL_ HAM detection network structure

In order to improve the detection accuracy of the model to meet the requirements of water surface garbage detection application scenarios, we introduced the HAM attention mechanism based on the GFL (Generalized Focal Loss) model to enhance the network feature extraction ability and improve network accuracy. GFL_ The entire network of HAM is divided into four parts: the backbone, the attention mechanism of HAM (Hybrid Attention Model), the neck, and the head, as shown in Figure 2. The backbone network of the model is a ResNet50 [10] structure; The neck network is an FPN [11] structure; The header network is composed of GFL_ Head [9] structure composition. The input image undergoes feature extraction and fusion operations through the backbone network and neck network, and then inputs the extracted features into the GFL_ Head obtains the predicted results of the model.

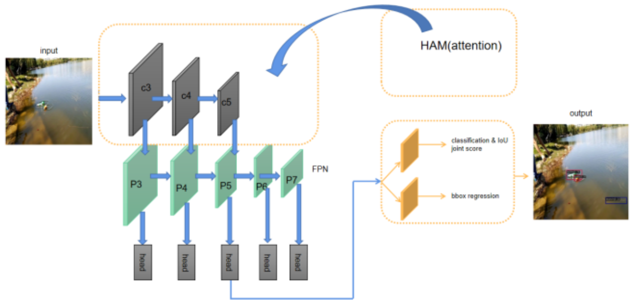

Figure 2. GFL_ HAM network structure diagram

The GFL model mainly re represents the physical objects at the end of the head. In previous work, training and reasoning were inconsistent in terms of quality estimation and classification scores. During the training process, quality estimation and classification scores are trained separately. In the reasoning process, quality estimates and classification scores are used together. The product of quality estimation and classification score is used to act on the NMS score of non-maximum suppression post processing [12], which is unreasonable. Quality estimation only performs position monitoring training on positive samples, resulting in a gap between training and testing, which reduces model performance. The GFL model proposes a new solution that combines classification scores and quality estimation by introducing Quality Focal Loss (QFL):

$$QFL(\sigma) = -|y-\sigma|^{\beta}((1-y)\log(1-\sigma)+y\log(\sigma)) \quad (1)$$

(Where y is a quality label of 0-1 and sigma is a prediction. Note that the global minimum solution for QFL is sigma=y. In this way, the cross entropy part becomes a complete cross entropy, while the adjustment factor becomes a power function of the absolute value of the distance. Similar to Focal Loss [13], it is generally optimal to take beta=2.) Instead of our usual Focal Loss:

$$FL(p) = -(1-p_t)^{\gamma}\log(p_t),\ p_t = \begin{cases} p, & y=1 \\ 1-p, & y=0 \end{cases} \quad (2)$$

To solve the problem of inconsistency in the quality estimation and classification score training reasoning process. In complex scenes, targets can have many uncertainties, such as surfboards blocked by water and highly overlapping elephants. In previous work, boundary box regression was modeled as a random distribution (Dirac or Gaussian distribution, etc.), resulting in inaccurate boundary box regression in complex scenes. Research has found that too random distribution is not conducive to learning network regression, resulting in a decline in network accuracy. Instead of using a loss function similar to GIOU-Loss [14], the GFL model redesigns the Distribution Focal Loss (DFL) DFL:

$$DFL(S_i, S_{i+1}) = \Big|^{\beta} \quad ((1-y)\,log(1-\sigma) + y\,log(\sigma))$$

Realize prediction of edge distribution, predict sharp distribution in clear boundary regions, predict smooth distribution at fuzzy locations, and improve model accuracy.

In order to improve the detection ability of the model and more accurately detect surface garbage, we introduce the HAM attention mechanism into the backbone network of the GFL model to improve the network's feature extraction ability, thereby improving network accuracy. HAM (Hybrid Attention Model) is a channel spatial attention mechanism, which consists of two parts: the first part is the channel attention mechanism, and the other part is the spatial attention mechanism. As shown in Figure 3:

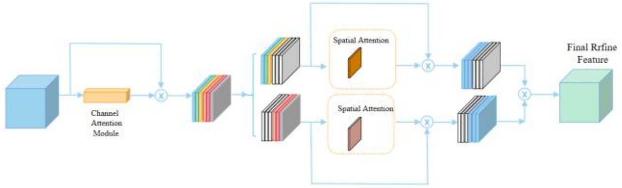

Figure 3. HAM (Hybrid Attention Model) Attention Mechanism Structure Diagram

First, the input feature is passed through the channel attention module (shown in the following formula) to obtain a channel weighted feature.

$$w = a[CA(F_{in})] \tag{3}$$

$$F_{out} = F_{in} * w \tag{4}$$

Where w represents the channel weighting coefficient, CA represents the channel attention mechanism of fast one-dimensional convolution, and a represents the adaptive mechanism, represents F_{in} input characteristics, and F_{out} represents output characteristics. Then, the obtained channel weighted features are input into the spatial attention mechanism,

$$C_{im} = \lambda C_{in} \tag{5}$$

$$f_{im} = f_{in} * C_{im} \tag{6}$$

$$f_{nim} = f_{in} * (1 - C_{im}) \tag{7}$$

$$f_{out} = SA * f_{im} + SA * f_{nim} \tag{8}$$

Among them, C_{im} 、 C_{in} represent important and input feature channels, and, f_{im} 、 f_{nim} 、 f_{in} represent important, non-important, and input features, λ represent channel importance separation coefficients, and SA represent spatial attention mechanisms based on one-dimensional convolution.

The HAM attention mechanism alleviates the burden of the channel attention mechanism through fast one-dimensional convolution, and introduces channel separation technology into the spatial attention mechanism to adaptively emphasize important features, reducing the amount of parameters while increasing the representation ability of the model, It solves the problem that existing attention mechanisms such as CBAM (Convolutional Block Attention Module) [15] attention mechanisms are difficult to achieve a good balance between performance and model complexity.

3. Results and discussion

In this section, a series of experiments have been designed to evaluate the GFL (Generalized Focal Loss) improvement proposed in this article_ The performance of HAM model water surface garbage mark detection algorithm.

3.1. Introduction to dataset preparation and experimental platform

The dataset used in this experiment is a water surface garbage dataset in a natural scene. This dataset contains a total of 2400 images, 1200 for training, 600 for validation, and 600 for testing. It includes eight categories: bottom, grass, branch, milk-box, plastic-bag, plastic-garbage, ball, and leaf.

3.2. Experimental setup

All water surface garbage data sets are divided into 80% of the data sets into training sets according to the principle of scene co distribution, 10% of the data sets are used as verification sets, and the remaining 10% are used as test sets. The super parameters set are as follows: The total training round is 12 rounds, using a random gradient descent strategy[16], and the initial learning rate is set to 0 001, momentum and weight attenuation are set to 0 937 and 0 0005。 Perform training using a single GPU with a batch size of 2.

3.3. Analysis of experimental results

This experiment was improved on GFL (Generalized Focal Loss). In order to verify the effectiveness of the HAM (Hbrid Attention Model) module, this article designed a group of experiments to verify.

In order to verify the feasibility of the proposed method, a series of ablation experiments were conducted on our proposed garbage detection dataset under natural scenarios, as shown in Table 1. In Table 1, "√" indicates adding the module, otherwise, it is not added. The impact of HAM on the experimental results was considered in the experiment, and the test set image size was 512 × 512, with a training cycle of 12 epochs. In this experiment, GFL (Generalized Focal Loss) is used as the basic model, and the results of ablation experiments are shown in Table 1.

Table.1 Ablation experiment

index	*HAM*	*mAP / %*	*FPS*
1	√	60.12	45
2		58.51	38

The results show that compared to GFL (Generalized Focal Loss), GFL_ The HAM model mAP has been improved by 1.61%. From the experimental results, it can be seen that the HAM (Hbrid Attention Model) module is effective, improving the detection accuracy of the model, making the model more suitable for target detection in natural scenes.

To verify the superiority of the algorithm, GFL_ HAM is experimentally compared with common target detection algorithms YOLOv3 [17], SSD [18], and GFL on a water surface garbage detection dataset. The framework used is pythoch, with a default input image size of 512x512, and a training cycle of 12 epochs. Table 2 shows the comparison of our method with other algorithms.

Table. 2 Comparison with different target detection algorithms

index	*mAP / %*
SSD	58.76
YOLOv3	59.03
GFL	58.51
GFL_HAM (OURS)	60.12

As shown in Table 2, the experimental results show that our method has a significant advantage in accuracy compared to current mainstream target detection algorithms, with mAP increasing 1.09%, 1.36%, and 1.61% compared to YOLOv3, SSD, and GFL, respectively. Figure 4 shows our model GFL_ Visualization of garbage detection in natural scenes using HAM.

Figure 4. Visualization Results

From the visualization result graph, it can be seen that our model can accurately detect most objects. However, due to the small number of data sets, uneven distribution, image shooting angle, and light reflection, some problems such as target misdetection and overlapping detection frames have resulted.

4. Conclusion

In order to improve the accuracy of water surface garbage detection, this paper proposes a garbage detection algorithm based on improved GFL (Generalized Focal Loss) due to the small number of data sets, uneven distribution, image shooting angle, and light reflection in water surface garbage data sets. This algorithm introduces the HAM (Hybrid Attention Model) attention mechanism into the feature extraction network Resnet [10], greatly improving the detection accuracy at the expense of a small amount of detection speed. On the surface garbage datasets, GFL_ The accuracy and speed of the HAM algorithm are better than other current mainstream single stage algorithms, but based on the visualization results, there are still some problems. Due to insufficient quantity and uneven data, there are some cases of false detection and missed detection of targets. Therefore, in the future work, how to achieve the balance between accuracy and speed of water surface garbage detection in natural scenes remains the focus of work.

Acknowledgments

This project is supported by the Southwest University for Nationalities Graduate Innovative Scientific Research Project (Project No. YB2022758)

References

[1] Duan, Ling. "Global Environment Outlook 6" was released after 5 years: The Earth has been severely damaged [J]. The world environment,2020(2):3.

[2] Regulations of Beijing Municipality on the Administration of Domestic Waste [J]. Communiqu é of the Standing Committee of the Beijing Municipal People's Congress, 2020, 000(005):P.83-95.

[3] Meinecke G, R atmeyer V, Renken J, HYBRID-ROV-Development of a new underwater vehicle for high-risk areas [C]/ /OCEANS11-MTS /IEEE Kona, 2011: 1-6.

[4] Hou Z Q, Han C Z, A Background reconstruction algorithm based on pixel intensity classification [J], Journal of Software, 2005, 16(9) : 1568-1576,

[5] Huang, Tao, Research on Water Object Extraction Algorithms in Natural Scenes[J]. Communication Technology,2013(3) : 80-82.

[6] YANG M, THUNG G. Classification of trash for recyclability status[R].Stanford CS229 Project Report,2016:832-841.

[7] Wang, Xie, Yang, et al. A garbage classification and detection method based on deep learning [J] .Modern electronic technology, 2021, 44(21):4.

[8] Xu, Xiong. An improved lightweight garbage target detection algorithm [J]. Computer Technology and Development, 2022(002):032.

[9] Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J., & Yang, J. (2020). Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection. ArXiv, abs/2006.04388.

[10] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[J]. IEEE, 2016.

[11] LIN T Y, DOLLARP, GIR SHICKR, et al, Feature pyramid networks for object detection[C]/ /2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , Piscataway: IEEE, 2017: 2117-2125,

[12] Neubeck A , Gool LJV . Efficient Non-Maximum Suppression[C] // 18th International Conference on Pattern Recognition (ICPR 2006), 20-24 August 2006, Hong Kong, China. IEEE Computer Society, 2006.

[13] Lin T Y , Goyal P , Girshick R , et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):2999-3007.

[14] Rezatofighi, Hamid, et al. "Generalized intersection over union: A metric and a loss for bounding box regression." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

[15] Woo, Sanghyun et al.“CBAM: Convolutional Block Attention Module.”European Conference on Computer Vision (2018).

[16] NTY,DOLLARP, GIRSHICKR, et al, Feature pyramid networks for object detection [C]/ /2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway : IEEE , 2017: 2117-2125.

[17] Redmon J , Farhadi A . YOLOv3: An Incremental Improvement [J]. arXiv e-prints, 2018.

[18] Berg AC , Fu CY , Szegedy C , et al. SSD: Single Shot Multi Box Detector:, 10.1007/978-3-319-46448-0_2[P]. 2015.