ISSN: 2710-0723

Lightweight YOLOv8 Mine Target Detection Algorithm for Embedded Deployment

Shixian Wei, Jiaqi Shu, Kejie Zhao, Siqi Zhu, Wenjie Su, Lu Zhang

College of Electrical and Information Engineering, Quzhou University, Quzhou, Zhejiang 324000, China

Abstract

Mine safety monitoring systems are usually deployed on embedded devices with limited computing, storage and energy resources. Traditional YOLOv8 models have large number of parameters and high computational complexity, which are difficult to realize real-time inference on these devices. In order to solve this problem, this study proposes a lightweight YOLOv8 mine target detection algorithm. The original PAN-FPN structure is replaced with BiFPN bidirectional feature pyramid to improve multi-scale feature reuse efficiency and reduce invalid computation through adaptive weighted fusion. A GCHead lightweight structure based on Group Convolution is introduced into the detection head to reduce parameter dimension and computational cost while maintaining detection accuracy. Experimental results show that compared with the original YOLOv8, the number of parameters and computational complexity of YOLOv8 are reduced by 46.5% and 37.8% respectively, while mAP@0.5 only decreases by 0.5%. It achieves the optimal balance between detection accuracy and lightweight, and can be efficiently deployed on mine embedded monitoring terminals. This research provides a useful reference for the application of target detection technology in resource-constrained environments.

Keywords

Lightweight YOLOv8; BiFPN; Group Convolution; Embedded Deployment.

1. Research Status and Significance

With the continuous advancement of intelligent mine construction, more and more monitoring devices are deployed in underground mines[4]. However, due to the harsh underground environment and limited power supply conditions, most monitoring devices use embedded terminals with strict restrictions on computing, storage and energy consumption. Traditional deep learning models such as YOLOv8 have high detection accuracy, but their large number of parameters and computational complexity make them unable to realize real-time inference on embedded devices.

At present, most lightweight methods realize model compression by directly cutting network channels or using depthwise separable convolution, which often sacrifice detection accuracy and are difficult to adapt to complex mine detection scenes. The traditional Kalman filter algorithm used in some positioning systems also has limitations in dealing with nonlinear problems, which further affects the real-time performance of the system.

The research of this project aims to meet the market demand for lightweight mine detection systems and realize real-time target detection on embedded devices. Aiming at the limitations of traditional YOLOv8 model in resource-constrained environments, this study adopts the method of combining BiFPN feature fusion and group convolution detection head to compress the model while maintaining detection accuracy. This method can effectively solve the problem

ISSN: 2710-0723

that the model cannot run in real time on embedded devices, and provide technical support for the popularization of intelligent mine monitoring systems[4,5].

2. Lightweight YOLOv8 Model Design

When the mobile robot captures dynamic targets, it needs to optimize the path planning to reduce the number of action steps. Similarly, when designing a lightweight target detection model, it needs to optimize the network structure to reduce the computational complexity while maintaining detection performance. The specific research contents of this study include the improvement of BiFPN lightweight feature fusion, the design of GCHead lightweight detection head, and the construction of YOLOv8 overall model.

2.1. BiFPN Lightweight Feature Fusion

The original YOLOv8 adopts PAN-FPN feature fusion structure, which has the problems of redundant connections, large computation and low feature reuse efficiency. BiFPN realizes lightweight feature fusion through the following optimizations[1]:

Simplify connection paths: Remove redundant edges in PAN-FPN that only connect single input nodes, and only retain key paths that contribute to feature fusion;

Adaptive weighted fusion: Assign learnable weights to input features of different scales, and dynamically adjust the fusion ratio according to feature importance;

Repeated stacking structure: Stack BiFPN modules multiple times to further enhance multi-scale feature fusion ability.

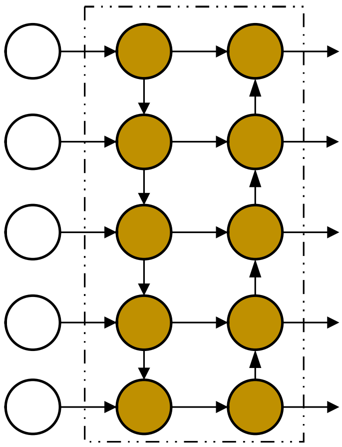

Figure 1. Schematic of the Original PAN-FPN Structure

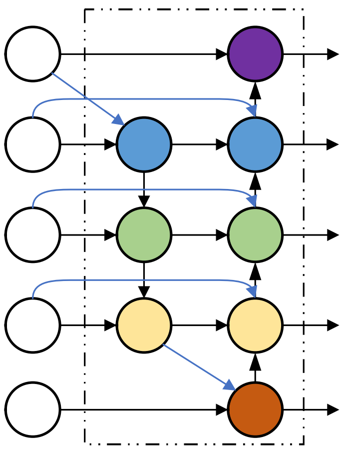

Figure 2. Schematic of the Introduced BiFPN Structure

ISSN: 2710-0723

In this study, the overall channel number of the neck network is first reduced to directly reduce the number of model parameters and computation; then the original PAN-FPN structure is replaced with BiFPN to compensate for the accuracy loss caused by channel reduction while reducing computation, and improve the detection ability of small targets.

2.2. GCHead Lightweight Detection Head

The original YOLOv8 detection head uses standard convolution to realize classification and regression tasks, and the number of parameters and computation account for a high proportion of the total model overhead. Group Convolution divides the input channels into multiple independent groups, and each convolution kernel only performs convolution operations with the channels of the corresponding group, which greatly reduces the computational complexity[2].

The computational complexity of standard convolution is:

$$C_{in} \times C_{out} \times k^2 \times H \times W$$

The computational complexity of group convolution is:

$$\frac{C_{in} \times C_{out} \times k^2 \times H \times W}{G}$$

where H and W are the feature map size, C_{in} and C_{out} are the number of input and output channels, K is the convolution kernel size, and G is the number of groups.

In this study, a GCHead lightweight detection head is constructed, and two layers of group convolution are used to replace the standard convolution in the original detection head, which significantly reduces the number of parameters and floating-point operations while maintaining classification and regression accuracy, and adapts to the inference requirements of embedded devices.

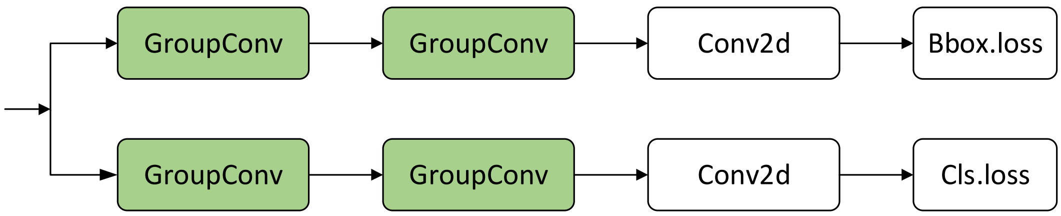

Figure 3. Schematic of GCHead

2.3. Overall Structure of YOLOv8 Model

The YOLOv8 model takes YOLOv8n as the benchmark, and its overall structure is divided into four parts:

Lightweight backbone network: Retain the CSPDarknet structure of YOLOv8n and appropriately adjust the number of channels;

Channel-reduced C2f module: Reduce the number of channels of the C2f module to further compress the model size;

BiFPN neck network: Adopt BiFPN bidirectional feature pyramid to realize multi-scale feature fusion;

GCHead detection head: Lightweight detection head based on group convolution to output target detection results[2].

ISSN: 2710-0723

Through the three-level lightweight design of channel reduction, BiFPN fusion and group convolution detection head, the optimal balance between accuracy and inference efficiency is achieved.

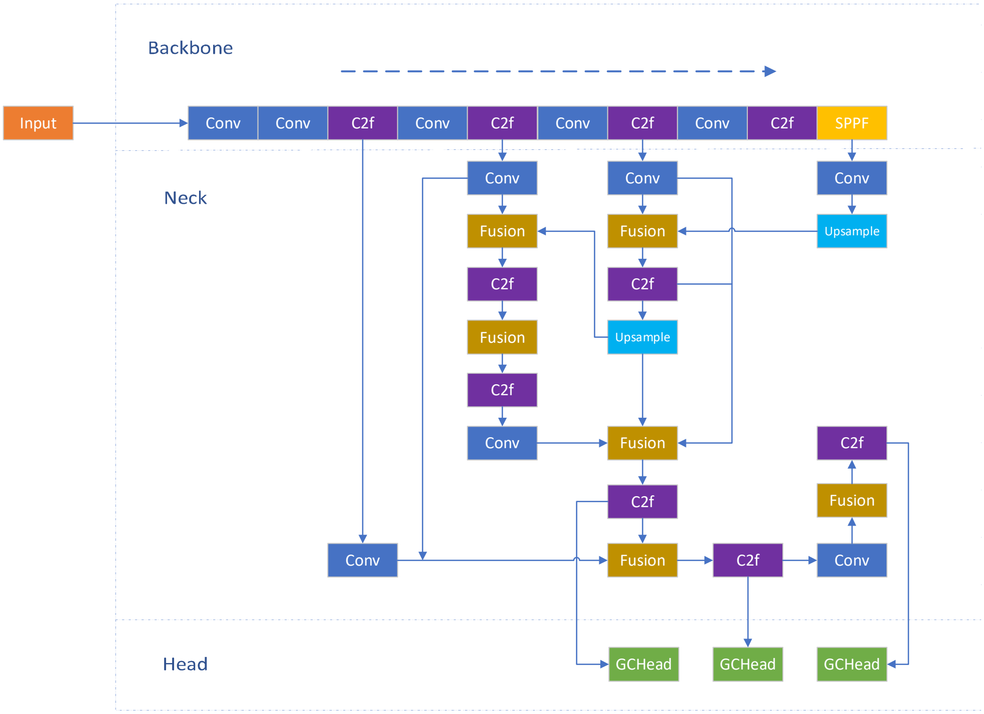

Figure 4. Structure Schematic of Lightweight YOLOv8

3. Lightweight Performance Experiment

Path planning can make the mobile robot complete the task with the least number of steps. Similarly, reasonable lightweight design can make the target detection model run in real time on embedded devices with the least computational cost. For lightweight mine target detection, this study designs a series of experiments to verify the effectiveness of the proposed method.

3.1. Experimental Setting

The experimental dataset, hardware environment, software environment and training parameters are completely consistent with those in Paper 1. The evaluation indicators are mAP@0.5, number of parameters and GFLOPs.

3.2. Ablation Experiment

In order to verify the effectiveness of each lightweight module, ablation experiments are carried out on the self-built mine dataset, and the results are shown in Table 1.

Table 1. Ablation experiment results on mine dataset

Model	Image Size	Parameters	GFLOPs	mAP@0.5/%
baseline	640*640	3.01M	8.2	0.852
+BiFPN	640*640	1.99M	7.3	0.860
+GCHead	640*640	2.42M	5.8	0.859
+BiFPN+GCHead	**640*640**	**1.61M**	**5.1**	**0.857**

It can be seen from Table 1 that:

ISSN: 2710-0723

Introducing BiFPN reduces the number of parameters by 33.9% and the computation by 11.0%, while increasing mAP@0.5 by 0.8%;

Replacing GCHead reduces the computation by 29.3%, and mAP@0.5 increases by 0.7%;

The combination of the two reduces the number of parameters by 46.5% and the computation by 37.8%, while mAP@0.5 only decreases by 0.5%, achieving a perfect balance between accuracy and lightweight.

3.3. Feature Fusion Mechanism Comparison Experiment

On the VOC2012 dataset, compare BiFPN with mainstream feature fusion mechanisms such as Weighted Fusion, Adaptive Fusion and Concat Fusion, and the results are shown in Table2[1].

Table 2. Comparison results of different feature fusion mechanisms

Model	Image Size	Parameters	GFLOPs	mAP@0.5/%
baseline	640*640	3.01M	8.2	0.642
BiFPN Fusion	**640*640**	**1.99M**	**7.3**	**0.641**
Weighted Fusion	640*640	2.05M	7.6	0.639
Adaptive Fusion	640*640	2.05M	7.6	0.637
Concat Fusion	640*640	2.08M	7.5	0.643

It can be seen from Table 2 that BiFPN has the lowest computation and the accuracy is basically the same as the baseline model, with the most significant lightweight advantage.

3.4. Lightweight Detection Head Comparison Experiment

On the VOC2012 dataset, compare GCHead with mainstream lightweight detection heads such as SCConv, RepConv and PartialConv, and the results are shown in Table 3[2,3].

Table 3. Comparison results of different lightweight detection heads

Model	Image Size	Parameters	GFLOPs	mAP@0.5/%
baseline	640*640	3.01M	8.2	0.642
GCHead	**640*640**	**2.42M**	**5.8**	**0.651**
SCConv	640*640	2.54M	5.8	0.621
RepConv	640*640	2.43M	5.6	0.642
PartialConv	640*640	2.43M	5.6	0.647

It can be seen from Table 3 that GCHead increases mAP@0.5 by 0.9% while reducing the computation, and the balance effect is better than all other comparison methods.

3.5. Generalization Performance Experiment

Ablation experiments are carried out on the VOC2012 dataset to verify the generalization performance of YOLOv8, and the results are shown in Table 4.

Table 4. Ablation experiment results on VOC2012 dataset

Model	Image Size	Parameters	GFLOPs	mAP@0.5/%
baseline	640*640	3.01M	8.2	0.642
+BiFPN	640*640	1.99M	7.3	0.641
+GCHead	640*640	2.42M	5.8	0.651
+BiFPN+GCHead	**640*640**	**1.61M**	**5.1**	**0.646**

ISSN: 2710-0723

The experimental results show that YOLOv8 reduces the number of parameters by 46.5% and the computation by 37.8% on the general dataset, while mAP@0.5 only decreases by 0.4%. It has excellent generalization performance and is suitable for embedded deployment in multiple scenarios.

4. Conclusion and Prospect

In the lightweight design of mine target detection models, comprehensive consideration of detection accuracy, computational complexity and deployment feasibility are key factors to ensure the real-time performance of the system. The comparison and experimental verification of different lightweight methods can provide effective guidance and support for the application of target detection technology in resource-constrained environments. As an efficient model compression method, group convolution is of great significance to improve the inference speed of the model. However, in practical applications, it is necessary to fully consider the characteristics and requirements of the embedded device and choose the appropriate lightweight strategy.

Future research directions include further optimizing the deployment of the model on edge devices to realize end-to-end real-time inference, and combining multi-sensor fusion technology to improve the perception ability of the system. In addition, the combination of other lightweight algorithms and deep learning technologies can further reduce the computational complexity of the model, so as to achieve more accurate and efficient mine target detection on low-power embedded devices.

Acknowledgments

This work was partially supported by National innovation and entrepreneurship training program for College Students (No. 202511488031).

References

[1] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 10781-10790.

[2] Ioannou Y, Robertson D, Cipolla R, et al. Deep roots: Improving cnn efficiency with hierarchical filter groups[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1231-1240.

[3] Dong W X, Liang H T, Liu G Z. A review of deep convolution applied to target detection algorithms[J]. Journal of Computer Science and Exploration, 2022, 16(05): 1025-1042.

[4] Liu Y, Ding Z. Discussion on the construction of coal mine safety intelligent system[J]. Industry and Mine Automation, 2023, 49(S1): 18-20.

[5] Fan W Q, Wang X J, Zhang Y H, et al. Underground personnel detection and tracking algorithm based on improved YOLOv7 and DeepSORT[J]. Coal Science and Technology, 2024, 52(S2): 343-355.