Physicochem. Probl. Miner. Process., 60(2), 2024, 187975 | Physicochemical Problems of Mineral Processing

http://www.journalssystem.com/ppmp | ISSN 1643-1049

Research on light weight intelligent identification method of coal and gangue

Deyong Li [1], Guofa Wang [1,2], Yongcun Guo [1], Yuhao Yang [1], Shuang Wang [1], Wenshan Wang[1]

[1] State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines, Anhui University of Science and Technology, Huainan 232001, china

[2] China Coal Technology Engineering Group Coal Mining Research Institute, Beijing 100013, China

Corresponding author: guofawang@126.com (Guofa Wang)

Abstract: Accurate identification of coal and gangue is essential for clean and efficient use of coal. Existing target detection algorithms are ineffective in detecting small-target and overlapping gangue, and contain complex network structure and large parameter volume, which cannot meet the demand of real-time detection of edge devices. To address the above problems, a lightweight detection and identification approach of coal gangue based on improved YOLOv5s is proposed. The depth-separable convolutions are used to replace ordinary convolutions, and the C3 (Concentrated-Comprehensive Convolution Block) Ghost module is constructed to replace all C3 modules in the YOLOv5s to reduce model computation and parameters. The CA (Coordinate Attention) attention mechanism is introduced to strengthen the attention to the target to be detected, suppress irrelevant background interference, and improve the detection accuracy of the model. The Focal- EIOU (Focal and Efficient Intersection Over Union) loss function was introduced to replace the original CIOU. Extensive experiments substantiated the proposed approach can effectively and quickly detect the small-target and overlapping coal gangue accurately, and the mAP (mean Average Presicion) reaches 97.7%. Compared with the original YOLOv5s, the proposed approach reduces the number of parameters and the amount of computation by 48.5% and 43%, respectively, under the premise of maintaining the same detection accuracy.

Keywords: intelligent sorting of coal gangue, image identification of coal gangue, deep learning, machine vision, depth-separable convolution

1. Introduction

Coal is the main source of energy in today's world, and it is the most economical and safe energy that can be utilized in a clean and efficient way (Wang and Meng, 2023; Li and Wang, 2019). Coal is mixed with gangue in the process of mining, which will cause environmental pollution and reduce the quality of combustion, so it is need to carry out coal gangue sorting (Yan et al. 2024; Lv et al. 2023). At present, China's coal gangue sorting has developed from manual gangue selection to mechanized gangue selection, automated gangue selection, and then to the direction of intelligent gangue selection (Shang et al. 2024). The key of intelligent separation of coal gangue is to identify coal gangue accurately and quickly.

At present, the methods of coal and gangue identification are ray method, multi-spectral identification, signal processing and image recognition method, but there are certain limitations, and failed to promote the use of large-scale (Xu et al. 2023; Wang et al. 2023). In recent years, deep learning technology through the convolutional neural network to automatically obtain and learn the image features, which can quickly extract and detect the feature information in the image of coal and gangue (Iwaszenko et al. 2021; Gao et al. 2024; Liu et al. 2024). Lv et al. (2022) propose an oversized gangue segmentation network model based on multi-task learning theory, which realizes the effective segmentation and identification of oversized gangue. Lai et al. (2022) improve Mask R-CNN combined with multispectral imaging segmentation of coal gangue instances, which is able to accurately locate coal and gangue.

DOI: 10.37190/ppmp/187975

With the requirement of detection efficiency and mobile deployment, more and more scholars have started to research on lightweight gangue detection network. Pan et al. (2022) propose a fast gangue recognition model based on improved YOLOv3-tiny, which improves the detection speed of the model. Wen et al. (2023) establish a lightweight gangue detection network based on CNN and Swell Transformer by introducing Swin Transformer module, a lightweight gangue detection network based on CNN and Transformer is established to achieve accurate location and recognition. Shang et al. (2023) propose an improved lightweight gangue recognition algorithm based on YOLOv5s, optimizing the feature extraction through the introduction of GhostNet and reducing the model parameters. Wei et al. (2023) propose a coal gangue image recognition model based on CSPNet-YOLOv7 target detection algorithm, which improves the recognition accuracy and speed by migration training.

To summarize, experts and scholars at home and abroad have proposed different improved algorithms for gangue detection, which improve the accuracy of gangue recognition while reducing the amount of computation. However, in the process of coal gangue detection, coal and gangue have different degrees of mutual obscuration, and at the same time mixed with small particle size of coal and gangue, these conditions to a certain extent affect the detection effect of the model, and at the same time, the complexity of the model still needs to be further reduced. Aiming at the above problems, this paper proposes a lightweight recognition method of coal gangue based on CCF-YOLOv5s. The experimental results show that the lightweight identification model can better take into account the two indicators of detection accuracy and model complexity, and is more suitable for real-time detection of coal gangue targets.

2. Lightweight gangue detection model design

2.1. Subsection

The YOLOv5 algorithm consists of Input, Backbone, Neck, and Head. On the input side, the optimal anchor frame parameters are calculated by the Adaptive Training Sample Selection (ATSS) method; Mosaic data enhancement is used for image splicing, enriching the background information of the images and adaptive image scaling to adapt to the detection of different sizes of coal gangue. In versions after YOLOv5-6.0 Backbone backbone network consists of standard convolutional layers Conv module and C3 module and SPPF for feature extraction work. The Neck network layer performs multi-scale feature fusion between Feature Pyramid Network (FPN) and Path Aggregation Network (PAN), so that the output features contain strong localization information of shallow features and advanced semantic information of deep features. The Head section performs multi-scale target detection on the image based on the fused feature information.

YOLOv5 network is divided into five versions, n, s, m, l and x, according to the model depth multiplier and layer channel multiplier, with the deepening of the network structure, the number of parameters of the model, the computational complexity and the recognition accuracy will be improved, but the inference speed of the model will be decreased. Among them, YOLOv5s has relatively fast detection speed and small memory occupation under the precondition of maintaining high recognition accuracy, so this paper adopts the YOLOv5s model as the basic framework and optimizes and improves the backbone network and the neck network to improve the detection performance of coal gangue targets and reduce the model complexity.

The structure of the CCF-YOLOv5s model proposed in this paper is shown in Fig. 1.Firstly, some ordinary convolutions are replaced with depth-separated convolutions, the BottleNeck module in C3 is replaced with the GhostBottleNeck module, so as to construct the C3Ghost module, and all the C3 modules in the neck network of YOLOv5s are replaced with the C3Ghost module, so as to effectively improve the performance of the target detection in the coal gangue and reduce the complexity of the model and at the same time enhance the generalization ability of the model, which makes the model more suitable to be applied to a variety of different scenarios and devices. Second, by using the CA attention mechanism added to the backbone and neck networks of YOLOv5s, features relevant to the target detection task can be captured more efficiently to improve the detection performance of the coal gangue target, and at the same time, the network can be more concerned about the important target areas, thus improving the target localization accuracy. Finally, the Focal-EIOU loss function is used instead of the original CIOU loss function. By combining the advantages of Focal Loss and EIOU Loss, the Focal-EIOU loss function increases the model's focus on difficult samples and improves the localization accuracy of the bounding box. This combination allows the model to improve both the handling of category imbalance and the improvement of localization accuracy, thus enhancing the performance of target detection in general.

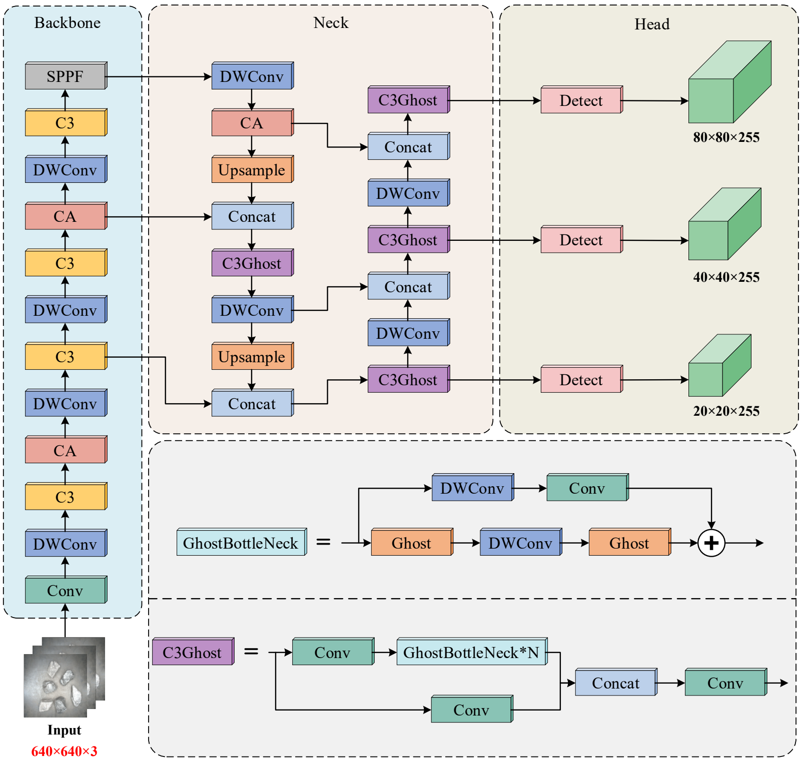


Fig. 1. CCF-YOLOv5s model structure

2.2. C3Ghost module

Ghost convolution is a lightweight feature extraction module proposed by Han et al. (2020), which generates feature maps using less number of parameters.The GhostBottleNeck module is a backbone network constructed on the basis of Ghost convolution, and there are two structures, as shown in Fig. 2. Both structures with step size 1 and step size 2 use residual structure and Fig. 2(b) has an extra depth convolution in the middle of the Ghost convolution.

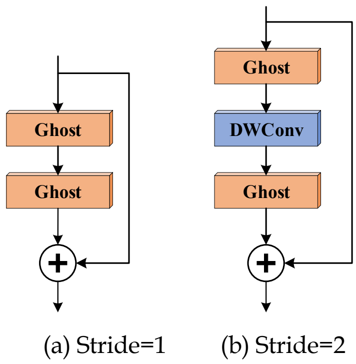


Fig. 2. GhostBottleNeck network structure diagram

This paper replaces the BottleNeck module in C3 with GhostBottleNeck module with step 2, and constructs the C3Ghost structure. C3Ghos are used to replace all the C3 modules in the YOLOv5s neck network, thus achieving the goal of reducing the amount of computation and the number of parameters. Fig. 3 shows the GhostBottleNeck structure used in this paper and Fig. 4 shows the C3Ghost network structure.

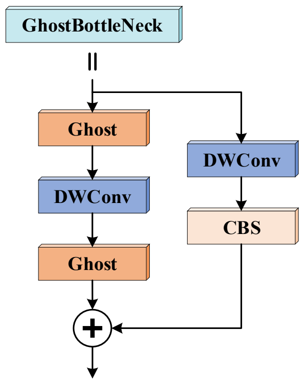


Fig. 3. The GhostBottleNeck network architecture used in this article

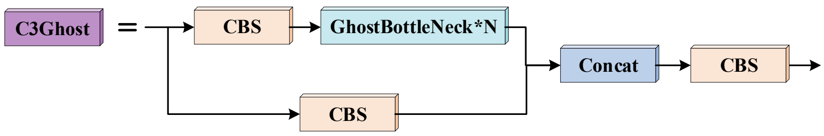


Fig. 4. C3Ghost network structure diagram

2.3. CA attention mechanism

The network structure diagram of CA attention mechanism is shown in Fig. 5 (Hou et al. 2021), from which it can be found that CA attention mechanism is by averaging pooling in horizontal and vertical directions respectively, and then encoding the spatial information using converters, and finally fusing the spatial information into the channel by weighting, which will be beneficial for CA attention mechanism to comprehensively consider spatial information and channel information, and to enhance the feature extraction capability. In deep networks, the number of channels of the feature graph may increase as the level of information transfer deepens, but not every channel is useful for the final task. the CA attention mechanism can help the network selectively transfer information that is useful for the task, optimize the efficiency of information transfer, and improve the performance of the network.

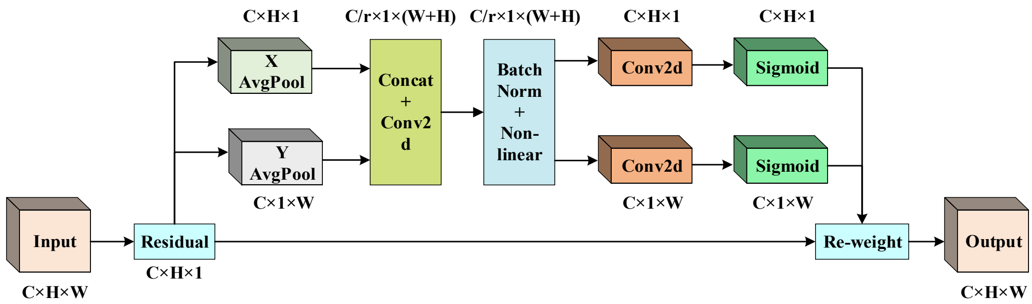


Fig. 5. CA attention mechanism structure diagram

2.4. Focal-EIOU loss function

CIOU (Computer Intersection Over Union) is based on the loss function of DIOU (Distance intersection over union) considering the aspect ratio of the prediction box (Bounding box) (Zheng et al. 2020), which further improves the model accuracy, both CIOU and DIOU are from the same Both CIOU and DIOU are from the same literature, and their mathematical expressions are shown in Eqs. (1) to (4):

$$Loss = L_{box} + L_{obj} + L_{cls} \tag{1}$$

$$L_{\mathrm{CIoU}} = 1\text{-IoU} + \frac{\rho^2(\mathrm{b},\mathrm{b}^{\mathrm{gt}})}{\mathrm{c}^2} + \alpha v \tag{2}$$

$$\alpha = \frac{v}{(1\text{-IoU})+v} \tag{3}$$

$$v = \frac{4}{\pi^2}(\arctan\frac{\omega^{\mathrm{gt}}}{\mathrm{h}^{\mathrm{gt}}}\text{-}\arctan\frac{\omega}{\mathrm{h}})^2 \tag{4}$$

In the formula, IoU denotes the intersection and concurrency ratio of the prediction frame to gt. b and b^{gt} then denote the coordinates of the centers of the prediction frame and gt. $\rho^2(\bullet)$ denotes the Euclidean distance between two coordinates. C denotes the diagonal length of the smallest outer rectangle of the prediction frame with gt.

Because CIOU is based on the difference in the aspect ratio of υ reflected in its formula, rather than the true difference in the confidence of width and height, it sometimes impedes effective optimization of similarity. Therefore, to solve this problem, Focal-EIOU separated the aspect ratio on the basis of CIOU and added Focal clustering high-quality anchor boxes (Zhang et al. 2022), whose mathematical Equations are shown as (5) and (6):

$$L_{\mathrm{EIoU}} = 1\text{-IoU} + \frac{\rho^2(\mathrm{b},\mathrm{b^{gt}})}{\mathrm{c}^2} + \frac{\rho^2(\omega,\omega^{\mathrm{gt}})}{\mathrm{C}_\omega^2} + \frac{\rho^2(\mathrm{h},\mathrm{h^{gt}})}{\mathrm{C}_\mathrm{h}^2} \tag{5}$$

$$L_{\text{Focal-EIoU}} = \mathrm{IoU}^\gamma \mathrm{L}_{\mathrm{EIoU}} \tag{6}$$

3. Experiment and result analysis

3.1. Data set construction and processing

Through the coal gangue image acquisition experimental platform to collect the resolution of 2448×2048 of different sizes and morphology of the coal gangue combination of images, coal gangue image acquisition experimental platform as shown in Fig. 6, the experimental device mainly includes the feeder, conveyor, CMOS industrial surface array camera, computer and light source controller.The CMOS industrial surface array camera model is Havel VisionMVCA050-11 UM/UC. The data acquisition system uses a frame rate of 35fps, and a light source controller is used to ensure that the light intensity in the acquisition area is stable at 1800 (±10) Lux, and two additional LED strips are added as auxiliary light sources. In addition, the acquisition system communicates with the camera using a USB 3.0 interface.

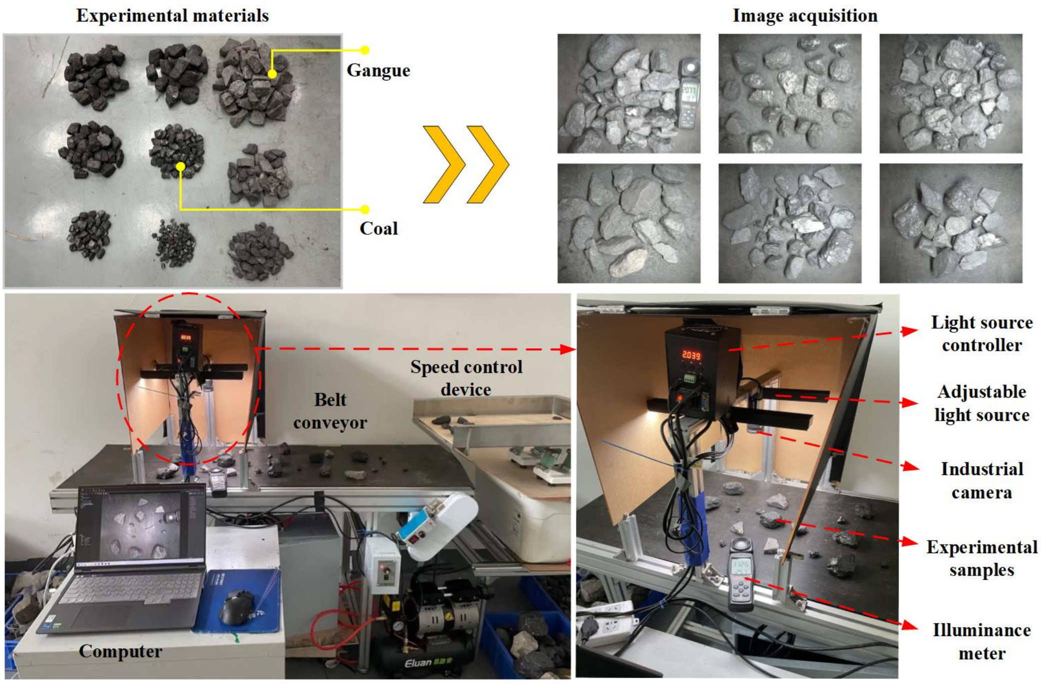


Fig. 6. Experimental platform for coal gangue image acquisition

The laboratory self-manufactured coal gangue data set through the CMOS industrial surface array camera in accordance with the ratio of coal and gangue 1:1 using different placement, photo distance, light intensity and combination of a total of more than 1,500 images, and through the data enhancement method to expand to 6,865 images. The gangue dataset produced in the laboratory was labeled using Labelimg image annotation software, and the dataset was divided into training and testing sets according to 9:1.

In order to expand the dataset size and enhance the dataset features, data enhancement methods are used to expand the laboratory homemade coal gangue dataset. The commonly used methods of data enhancement methods are image equal scale scaling, unequal scale scaling, horizontal flipping, vertical flipping, random contrast, random brightness, adding noise, such as Gaussian noise and pretzel noise, and adding blur, such as motion blur and Gaussian blur, etc., to enhance the data of coal gangue images, as shown in Fig. 7.

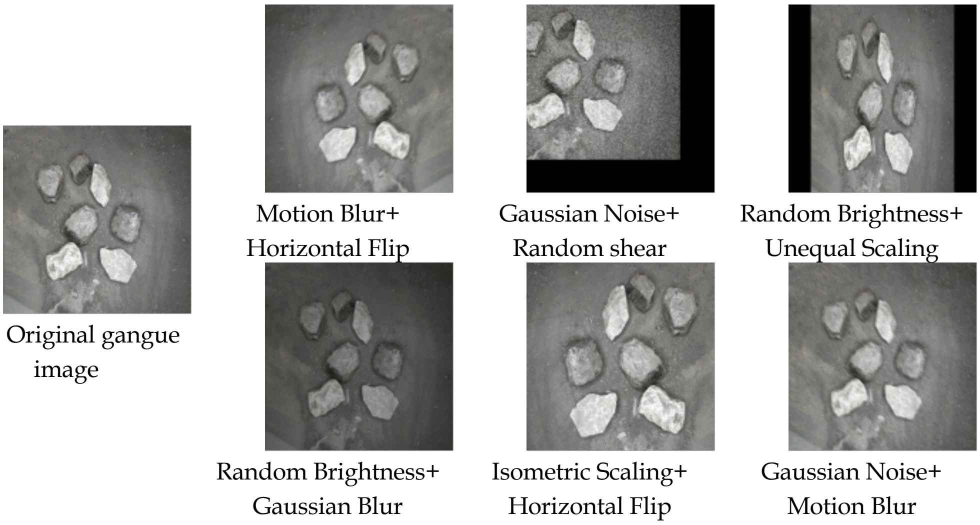


Fig. 7. Sample image of the data enhancement section

3.2. Model training and evaluation

This study was trained and tested on a computer with an Intel(R) Core (TM) i7 CPU @ 2.90GHz and an NVIDIA GeForce RTX2060 GPU. It was also configured with CUDA version 10.0 parallel computing meter framework and Cudnn version 7.3 deep learning acceleration library. Python 3.6 was used for programming implementation under the PyTorch deep learning framework. Before training the model for gangue images, the training hyperparameters are adjusted to obtain the optimal model, and the specific parameters are shown in Table 1.

Table 1. Model training hyper-parameter setting

Hyperparameterization	Parameter value
Image size	640×640
Weight decay	0.0005
Learning rate	0.01
Hue augmentation	0.015
Momentum	0.937
Saturation augmentation	0.7

In order to evaluate the performance of the improved YOLO v5s model through the detection results, Accuracy (Precision, P), Recall (Recall, R), mean average precision (mAP), and average precision (AP) are chosen as the evaluation indexes.

$$P = \frac{TP}{TP+FP} \tag{7}$$

$$R = \frac{TP}{TP+FP} \tag{8}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_{(i)} \tag{9}$$

$$AP = \int_0^1 P(r)dr \tag{10}$$

where TP is the number of correctly detected positive samples, FP is the number of negative samples detected positive samples, FN is the number of undetected positive samples, N_{pred} is the number of predicted detected frames, and N_{GT} is the number of true frames (ground truth of manually labeled frames).

3.3. Adding different improvement strategies to ablation experiments

In order to visualize the impact of different improvement measures on the detection performance of the network model, ablation experiments are conducted, with the same training environments, the three improvement measures are stacked sequentially to improve and train the network model respectively, and the results are in Table 2 and Fig. 8.

Table 2. Ablation experimental results

Mould	Deeply separable convolution + C3Ghost	CA attention mechanism	Focal-EIoU loss function	mAP@0.5/%	Parameters	GFLOPs
A				97.6	7015519	15.8
B	√			97.0	3598255	8.9
C	√	√		97.5	3613303	9.0
D	√	√	√	97.7	3613303	9.0

3.3.1. Model A → Model B

By replacing part of the ordinary convolution with depth-divisible convolution, and at the same time replacing all the C3 modules with C3Ghost modules in the neck network of the YOLOv5s model, we can reduce the number of parameters and computation volume of the model, improve the speed and efficiency of the model, and at the same time, reduce the risk of overfitting while maintaining a high detection accuracy, so as to make the model lighter and more suitable for mobile This makes the model more lightweight and suitable for mobile and other resource-constrained scenarios. From the experimental results, it can be seen that, compared with Model A, there is a 47.9% reduction in the number of parameters and a 43.7% reduction in the amount of computation. Although the number of parameters and the amount of computation have decreased dramatically, the consequent detection accuracy has also decreased slightly, with the mAP decreasing from 97.6% to 97.0%.

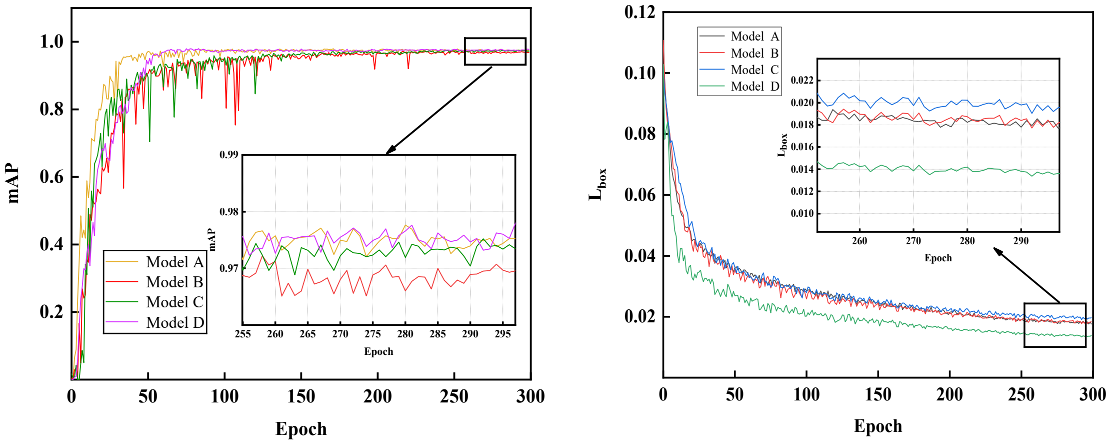
3.3.2. Model B → Model C

To compensate for the decrease in detection accuracy due to lightweighting, the addition of the CA attention mechanism allows the model to more effectively utilize the correlation between different channels in the feature map, thus improving the accuracy of target detection. From the experimental results, it can be seen that mAP increased from 97.0% to 97.5% with the addition of CA attention mechanism compared to model B. Additionally, in order to show more intuitively the degree of attention to gangue features in Model B and Model C, the output layers of the two models were visualized and analyzed by employing class-activated heat maps. The result of its class activation heat map comparison is shown in Fig. 9, from which it can be seen that the addition of the CA attention mechanism allows the network to focus more on the feature channels that are useful for the target detection task and inhibit the channels that are not relevant to the task, thus improving the ability of feature representation.

3.3.3. Model C → Model D

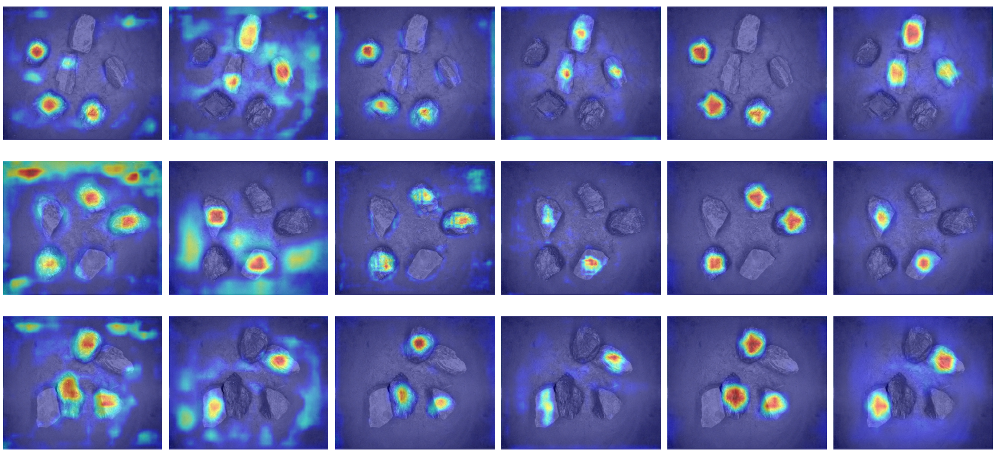
The Focal-EIoU loss function is used instead of the original CIOU loss function. By combining the advantages of both Focal Loss and EIOU Loss, the Focal-EIOU loss function increases the model's focus on difficult samples and improves the localization accuracy of the bounding box. This combination allows the model to improve both the handling of category imbalance and the improvement of localization accuracy, thus enhancing the performance of target detection in general. From the experimental results, it can be seen that the mAP with the Focal-EIoU loss function improves from 97.5%

to 97.7% compared to Model C. The introduction of the Focal-EIoU loss function allows the model to show better adaptability and robustness in the face of small-target detection and detection under occlusion.



(a) mAP@0.5 Results of indicator experiments (b) Boundary box loss Lbox comparison results

Fig. 8. Ablation experimental results



Coal Gangue Coal Gangue Coal Gangue

(a) Original figure (b) Model B (c) Model C

Fig. 9. Class activation heatmap

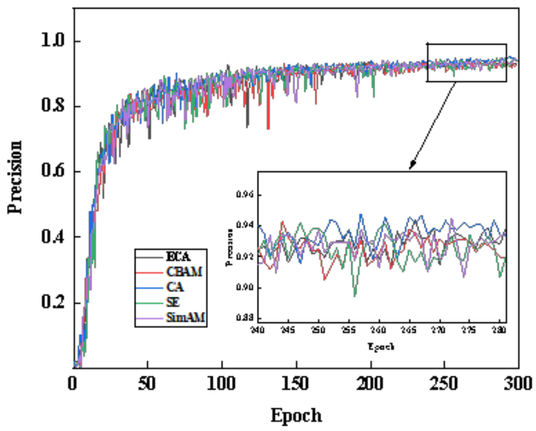
3.4. Comparative experiments adding different attention mechanisms

In this paper, by adding CA attention mechanism to the backbone network part and neck part of the YOLOv5s model, we improve the feature discriminative and information transfer efficiency by adaptively learning the weights of the channels in the feature graph, which significantly enhances the target detection accuracy and model generalization ability.To further verify the comprehensive detection performance of CA attention mechanism, CBAM attention mechanism (Woo et al. 2018), SE attention mechanism (Hu et al. 2018), SimAM attention mechanism (Yang et al. 2021) and ECA attention mechanism (Wang et al. 2020) were added at the same position of model B in the ablation experiments for the comparison experiments under the condition of ensuring the same conditions of other tests, and the results of their addition of different attention mechanisms are shown in Table 3 and Fig. 10.From the experimental results in Table 3 and Fig. 10, it can be seen that the five attention mechanisms do not fluctuate significantly in terms of the number of model parameters and the amount of computation, all of which are around 3600000, and still maintain the advantage of lightweight. Specifically, adding CA attention mechanism increases mAP the most, which is 0.5 percentage points, while adding SE attention mechanism, SimAM attention mechanism and ECA attention mechanism mAP improves by 0.3, 0.4 and 0.1 percentage points, respectively, and adding CBAM attention mechanism mAP remains unchanged.

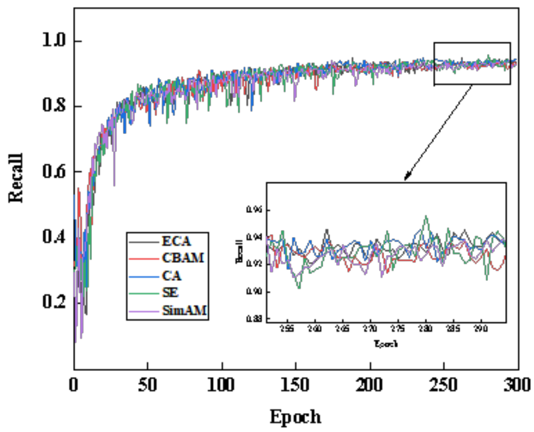
In summary, the model works optimally with the addition of the CA unparticipated attention mechanism.

Table 3. results of adding different attention mechanisms

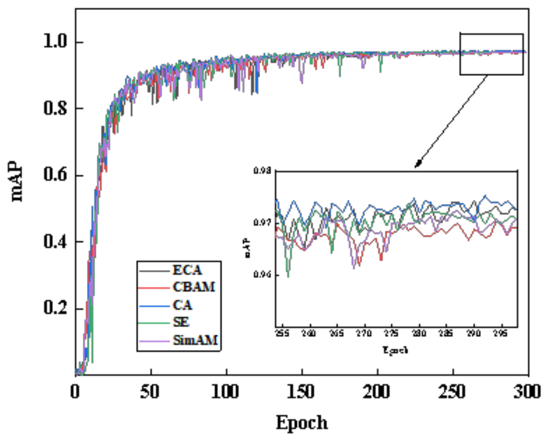
Mould	mAP@0.5/ %	Parameters	GFLOPs
Mould B + CA	97.5	3613303	9.0
Mould B + CBAM	97.0	3616060	9.0
Mould B + SE	97.3	3599986	8.9
Mould B + SimAM	97.1	3598255	8.9
Mould B + ECA	97.4	3598264	8.9



(a) Precision experiment results



（b）Recall experiment results

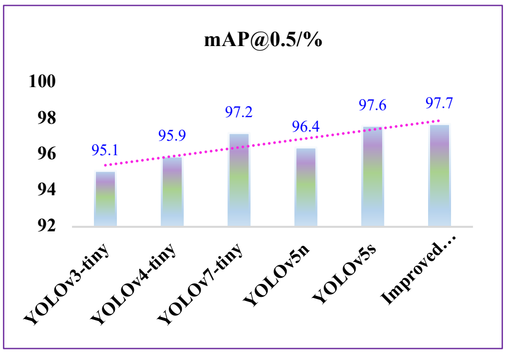


(c) mAP@0.5 experiment results

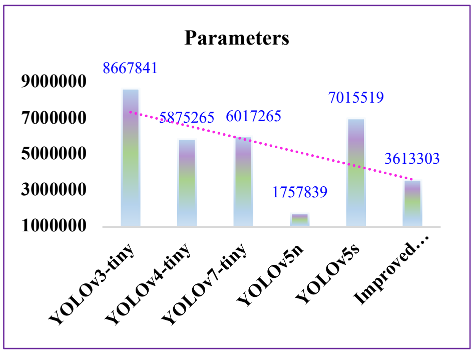
Fig. 10. Results of adding different attention mechanisms

3.5. Comparative experiments of different target detection algorithms

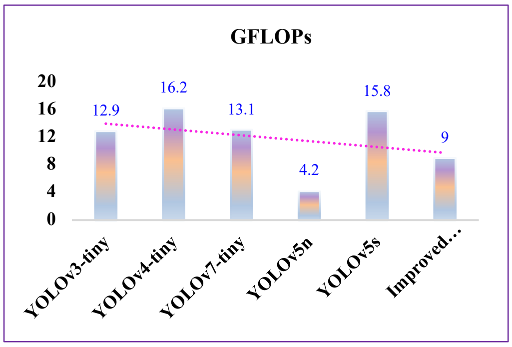
To further validate the detection performance of CCF-YOLOv5s model, this paper compares it with other YOLO series detection algorithms, and the experimental results are shown in Fig. 11. From the figure, it can be seen that although YOLOv5n performs well in terms of the number of parameters and calculation indicators. However, in terms of accuracy, mAP@0.5 is 1.3% lower than that of the improved YOLOv5s model. Compared with YOLOv3-tiny (Adarsh et al. 2020), YOLOv4-tiny (Wang et al. 2021), YOLOv5s and YOLOv7-tiny (Wang et al. 2023), the improved YOLOv5s model has advantages in three indicators. The mAP@0.5 index increased by 2.6%, 1.8%, 0.5% and 0.1%, respectively. The number of parameters decreased by 58.3%, 38.5%, 39.9% and 48.5%, respectively. The amount of computation decreased by 30.2%, 44.4%, 31.3% and 43%, respectively. Therefore, the CCF-YOLOv5s lightweight coal gangue target detection model proposed in this paper has obvious advantages over mainstream target detection algorithms in terms of the number of parameters and the lightweight computation.



(a) mAP@0.5 indicator experiment results



(b) Recall indicator experiment results



(c) GFLOPs indicator experiment results

Fig. 11. Comparative experimental results of algorithms

3.6. Visualization of test results

To validate the superiority of the comprehensive detection performance of the CCF-YOLOv5s model proposed in this paper, 687 test set gangue pictures which have been divided are tested and 8 pictures are randomly selected from the test set to be displayed, and their detection results are shown in Fig. 12.From the figure, it can be seen that the coal and gangue have different degrees of mutual obscuration, while mixed with small particle size of coal and gangue, these conditions affect the detection effect of the original model to a certain extent. However, the CCF-YOLOv5s lightweight model proposed in this paper can quickly and accurately recognize and annotate the position and confidence level of all the coal and gangue targets, while there is no misdetection or omission.Therefore, CCF-YOLOv5s model has the best effect on target detection for mutual occlusion gangue and small particle size gangue, and it is a great progress in lightweighting, and compared with mainstream algorithms, its advantages are more obvious, which better meets the deployment needs of edge devices.

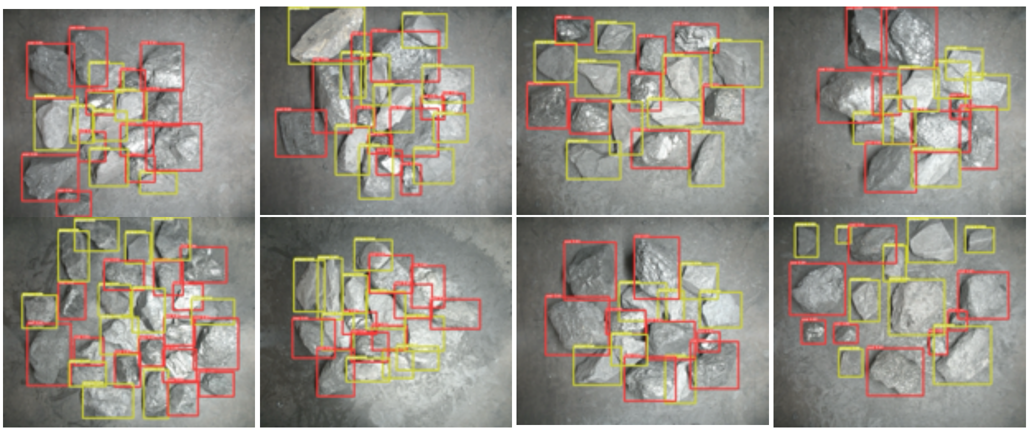


Fig. 12. Improved YOLOv5s model detection results

4. Conclusions

Existing target detection algorithms are ineffective in detecting small target gangue and occluded gangue, and the algorithms contain complex network structure and large parameter volume, which can not reach the demand of real-time detection of edge devices. To address the above problems, a lightweight detection approach of coal gangue based on improved YOLOv5s model is proposed. The experimental setup was built to construct multi-conditional coal and gangue datasets, train the models, and the experimental results showed:

(1) The proposed algorithm can effectively and rapidly detect small target gangue and obscuring gangue accurately with 97.7% mAP, which is 0.1% higher than the original YOLOv5s algorithm, and the amount of parameter is reduced by 48.5% and the amount of computation is reduced by 43%.

(2) Comparing with other YOLO series algorithms, the improved YOLOv5s algorithm has more obvious comprehensive performance improvement, better environmental robustness and practicality, and provides theoretical and technical references for coal gangue detection and recognition.

(3) In the study, this paper does not consider the effect of image quality in the dataset on model accuracy. In underground environments, the effects of dust and light result in gangue images that usually have low contrast and severe color deviations, which negatively affect gangue detection. In order to further improve the model's detection accuracy in complex situations, image enhancement algorithms will be used in future research to improve the dataset quality.

Acknowledgments

This work is supported by SKLMRDPC23KF21, No. 2023yjrc56, No.2308085Y37, MHTE23-R07, No.52274152.

References

ADARSH, P., RATHI, P., KUMAR, M. 2020. *YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In 2020 6th international conference on advanced computing and communication systems* (ICACCS) (pp. 687-694). IEEE.

GAO, A., LU, S., XU, R., LI, Z., WANG, B., ZHU, S., GAO Y., PAN, B. 2024. *Deep reinforcement learning based planning method in state space for lunar rovers*. Engineering Applications of Artificial Intelligence, 127, 107287.

HAN, K., WANG, Y., TIAN, Q., GUO, J., XU, C., XU, C. 2020. *Ghostnet: More features from cheap operations*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 1580-1589).

HOU, Q., ZHOU, D., FENG, J. 2021. *Coordinate attention for efficient mobile network design*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 13713-13722).

HU, J., SHEN, L., SUN, G. 2018. *Squeeze-and-excitation networks*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7132-7141).

IWASZENKO, S., ROG, L. 2021. *Application of deep learning in petrographic coal images segmentation*. Minerals, 11(11), 1265.

LAI, W., HU, F., KONG, X., YAN, P., BIAN, K., DAI, X. 2022. *The study of coal gangue segmentation for location and shape predicts based on multispectral and improved Mask R-CNN*. Powder Technology, 407, 117655.

LI, J., WANG, J. 2019. *Comprehensive utilization and environmental risks of coal gangue: A review*. Journal of Cleaner Production, 239, 117946.

LIU, T., ZHANG, Z., LEI, Z., HUO, Y., WANG, S., ZHAO, J., ZHANG J., JIN X., Zhang, X. 2024. *An approach to ship target detection based on combined optimization model of dehazing and detection*. Engineering Applications of Artificial Intelligence, 127, 107332.

LV, Z., CUI, Y., ZHANG, K., SUN, M., LI, H., WANG, W. 2023. *Investigating comparisons on the coal and gangue in various scenarios using multidimensional image features.* Minerals Engineering, 204, 108450.

LV, Z., WANG, W., ZHANG, K., LI, W., FENG, J., XU, Z. 2022. *A synchronous detection-segmentation method for oversized gangue on a coal preparation plant based on multi-task learning.* Minerals Engineering, 187, 107806.

PAN, H., SHI, Y., LEI, X., WANG, Z., XIN, F. 2022. *Fast identification model for coal and gangue based on the improved tiny YOLO v3.* Journal of Real-Time Image Processing, 19(3), 687-701.

SHANG, D., HUANG, Y., HUANG, X., ZHANG, T., NIU, Y. 2024. *Fuzzy adaptive control of coal gangue sorting parallel robot with variable load.* International Journal of Coal Preparation and Utilization, 1-22.

SHANG, D., YANG, Z., LV, Z. 2023. *Recognition of coal and gangue under low illumination based on SG-YOLO model.* International Journal of Coal Preparation and Utilization, 1-16.

WANG, C. Y., BOCHKOVSKIY, A., LIAO, H. Y. M. 2023. YOLOv7: *Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors.* In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475).

WANG, C. Y., BOCHKOVSKIY, A., LIAO, H. Y. M. 2021. *Scaled-yolov4: Scaling cross stage partial network.* In Proceedings of the IEEE/cvf conference on computer vision and pattern recognition (pp. 13029-13038).

WANG, G., MENG, L., 2023. *Development of coal mine intelligence and its technical equipment*. China Coal. 50,1-27.

WANG, L., WANG, X., LI, B. 2023. *Data-driven model SSD-BSP for multi-target coal-gangue detection*. Measurement, 219, 113244.

WANG, Q., WU, B., ZHU, P., LI, P., ZUO, W., HU, Q. 2020. *ECA-Net: Efficient channel attention for deep convolutional neural networks*. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11534-11542).

WEN, X., LI, B., WANG, X., LI, J., WEI, D., GAO, J., ZHANG, J. 2023. *A Swin transformer-functionalized lightweight YOLOv5s for real-time coal–gangue detection.* Journal of Real-Time Image Processing, 20(3), 47.

WEI, X., WANG, F., LIU, C., HE, D., XU, D. 2023. *Coal gangue image recognition model based on CSPNet-YOLOv7 target detection algorithm.* Coal Science and Technology, 1-13.

WOO, S., PARK, J., LEE, J. Y., KWEON, I. S. 2018. *Cbam: Convolutional block attention module*. In Proceedings of the European conference on computer vision (ECCV) (pp. 3-19).

XU, M., MAO, Y., YAN, Z., ZHANG, M., Xiao, D. 2023. *Coal and Gangue Classification Based on Laser-Induced Breakdown Spectroscopy and Deep Learning*. ACS omega, 8(50), 47646-47657.

YANG, L., ZHANG, R. Y., LI, L., XIE, X. 2021. Simam: *A simple, parameter-free attention module for convolutional neural networks.* In International conference on machine learning (pp. 11863-11874). PMLR.

YAN, P., WANG, W., LI, G., ZHAO, Y., WANG, J., WEN, Z. 2024. *Detection of coal gangue based on spectral technology and enhanced lightweight YOLOv7-tiny.* International Journal of Coal Preparation and Utilization, 1-21.

ZHANG, Y. F., REN, W., ZHANG, Z., JIA, Z., WANG, L., TAN, T. 2022. *Focal and efficient IOU loss for accurate bounding box regression.* Neurocomputing, 506, 146-157.

ZHENG, Z., WANG, P., LIU, W., LI, J., YE, R., REN, D. 2020. *Distance-IoU loss: Faster and better learning for bounding box regression*. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 07, pp. 12993-13000).