processes

MDPI

Article

Research and Application of Coal Gangue Detection Method Based on Improved YOLOv7-Tiny

Shenglei Hao [1,*], Jian Ma [2], Zhenyang Zhang [1], Yong Liu [1], Dongxu Wu [1], Lehua Zhao [2], Peng Zhang [2], Kun Zhang [2] and Mingchao Du [2]

[1] Yankuang Energy Group Co., Ltd., No. 3 Coal Mine, Jining 273500, China; 15621356152@163.com (Z.Z.); 19560737012@163.com (Y.L.); 17685538572@163.com (D.W.)
[2] Shandong Provincial Key Laboratory of Robotics and Intelligent Technology, Shandong University of Science and Technology, Qingdao 266590, China; sdwfmj@gmail.com (J.M.); zhaolehua112@163.com (L.Z.); sdustzhangpeng@163.com (P.Z.); zhangkunliaoning@163.com (K.Z.); dumingchao@sdust.edu.cn (M.D.)
* Correspondence: haoshengleijisan@163.com; Tel.: +86-053-7262-8588

Abstract

Coal gangue sorting is crucial for improving coal quality and reducing environmental pollution; however, traditional methods suffer from resource wastage, high cost, and intensive labor demands. To address these challenges, this paper investigates an image recognition-based coal gangue sorting technique and proposes an improved YOLOv7-tiny detection model tailored for edge GPU devices with limited computational power and memory. YOLOv7-tiny is selected as the baseline due to its balanced performance in detection accuracy, architectural maturity, and deployment stability on edge GPUs. Compared to newer lightweight detectors such as YOLOv8-N and YOLOv6-N, YOLOv7-tiny adopts an ELAN-based modular design, which facilitates structural optimization without relying on anchor-free reconstruction or complex post-training strategies, making it particularly suitable for engineering enhancements in real-time industrial sorting under resource constraints. To tackle the limitations in computing and storage, we first introduce an ELAN-PC feature extraction module based on partial convolution and ELAN. Secondly, a GhostCSP module is proposed by integrating cross-stage aggregation and Ghost bottleneck concepts. These modules replace the original ELAN structures in the backbone and neck networks, significantly reducing floating-point operations (FLOPs) and the number of parameters. Furthermore, the SIoU loss function is adopted to replace the original bounding box loss, enhancing detection accuracy. Experimental results demonstrate that compared with the baseline YOLOv7-tiny, the improved model increases mAP0.5 from 86.9% to 88.7% (a gain of 1.8%), reduces FLOPs from 13.2 G to 9.2 G (a decrease of 30%), and cuts parameters from 6.0 M to 4.3 M (a reduction of 28%). In dynamic sorting tests, the model achieves a coal gangue sorting rate of 82.2% with a misclassification rate of 8.1%, indicating promising practical applicability.

Keywords: deep learning; object detection; YOLOv7-tiny; coal gangue identification; belt conveyor

Check for updates

Academic Editor: Dan Zhang

Received: 16 December 2025
Revised: 22 January 2026
Accepted: 26 January 2026
Published: 30 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Coal gangue sorting promotes the efficient utilization of coal resources while mitigating negative environmental impacts [1,2]. Through precise sorting technology, impurities are removed from coal, enhancing its overall quality to better meet the stringent demands of industrial and energy production. This process conserves valuable coal resources, prevents

waste, and reduces pollution by lowering emissions that contaminate air, soil, and water bodies. Raw coal extracted from underground mines undergoes multiple sorting stages at coal preparation plants to separate gangue and enhance the quality of the raw coal.

Among various sorting methods, image-based coal gangue separation has emerged as a research focus in the field due to its advantages over physical separation techniques [3,4], including low cost, absence of radiation hazards, and minimal resource waste. Consequently, coal gangue detection methods, as its core technology, have also been extensively explored. With the advancement of deep learning, object detection models based on deep learning have gradually become the mainstream approach in the field of object detection due to their advantages of strong robustness, good generalization, excellent detection performance, and high recognition accuracy [5,6]. Chien-Yao Wan et al. [7] conducted research on scaling strategies for networks of different sizes, introducing a lightweight model of YOLOv4 named YOLOv4-tiny. Following YOLOv4 and YOLOv4-tiny, the YOLO series established a model architecture comprising the input, backbone network, neck network, and detection head. Through YOLOv9 [8–10], YOLOv5 built upon this framework, integrating various research advancements to enhance model performance. Each iteration provided multiple models [11–13] tailored to diverse application scenarios. Ultralytics [14] introduced the YOLOv5 series models, employing CSPDarknet53 [15,16] as the backbone network and proposing the C3 feature extraction module. Five models—N, S, M, L, and X—were provided based on scaling factors, with increasing floating-point operations and parameter counts, corresponding to progressively better detection performance. Chuyi Li et al. [17] proposed the YOLOv6 series, comprising four scales (N, S, M, L). It employs EfficientRep [18]—built on structural reparameterization—as the backbone network, introduces RepBlock and CSPStackRepBlock [19] to refine the neck network, utilizes a decoupled head to separate classification and localization tasks, and adopts anchor-free detection. Chien-Yao Wang et al. [20] introduced the YOLOv7 series, comprising three scales: YOLOv7-tiny, YOLOv7, and YOLOv7-X. They proposed an efficient layer aggregation network module [21] for constructing the backbone and neck networks, further enhancing model performance through techniques like structural reparameterization [22], auxiliary head detection [23,24], and dynamic label allocation strategies [25,26]. Deep learning methods demonstrate significant advantages in coal gangue detection [27,28] but also exhibit certain limitations. These approaches demand substantial computational resources, leading to deployment challenges that hinder their application in mining environments. Concurrently, detection methods proposed in some studies [29] fail to sufficiently align with the specific operational contexts of coal gangue detection, resulting in limited practical implementation during actual detection processes.

This paper investigates coal gangue sorting technology based on image recognition and proposes an improved YOLOv7-tiny coal gangue detection method tailored for GPU devices with limited computational power and memory. The approach introduces the ELAN-PC feature extraction module [30] to enhance the ELAN backbone network and the GhostCSP feature extraction module to improve the ELAN neck network. These modifications significantly reduce the model's floating-point operations and parameter count. Additionally, the SIOU bounding box loss function is adopted to refine YOLOv7-tiny's bounding box loss function, thereby enhancing detection accuracy. Experimental results demonstrate that for input images at 640 × 640 resolution, the improved YOLOv7-tiny model achieves an mAP0.5 of 88.7%, representing a 1.8% increase over the baseline YOLOv7-tiny (86.9%). Floating-point operations decreased from 13.2 G to 9.2 G (a 30% reduction), while the number of parameters decreased from 6.0 M to 4.3 M, a reduction of 28%. When deployed as a detection model in coal gangue sorting experiments, it achieved an 82.2% sorting rate for coal gangue and an 8.1% false selection rate for coal. This

demonstrates the reliability of the designed coal gangue sorting system, laying a foundation for subsequent model optimization and system upgrades.

Based on the above analysis, the main contributions of this work can be summarized as follows:

(1) A lightweight coal gangue detection framework is proposed based on YOLOv7-tiny, in which the backbone network is structurally optimized for industrial conveyor-belt scenarios with strict real-time and computational constraints.
(2) An ELAN-PC module is designed by introducing partial convolution into the ELAN structure, significantly reducing redundant computation while preserving fine-grained feature extraction capability for irregular coal gangue targets.
(3) A GhostCSP module is incorporated to enhance cross-stage feature fusion efficiency and robustness under dense stacking and partial occlusion conditions.
(4) The SIoU loss function is employed for bounding box regression, enabling joint optimization of distance, direction, shape, and overlap, which improves localization accuracy for irregularly shaped objects.
(5) Extensive experiments, including extended ablation studies and comparisons with recent lightweight detectors, are conducted to validate the effectiveness, efficiency, and practical applicability of the proposed method.

2. Coal Gangue Detection Based on an Improved YOLOv7-Tiny Approach

2.1. Improvements to the YOLOv7-Tiny Model

To address the computational and memory constraints of edge GPU devices, the ELAN-PC feature extraction module (the module is implemented using PyTorch 1.12.0) was proposed based on partial convolutions and ELAN. Inspired by the cross-stage aggregation network concept, the GhostCSP feature extraction module [31] was developed using the Ghost bottleneck architecture. These two modules were applied to refine the ELAN backbone and neck networks, respectively, thereby reducing the model's floating-point operations and parameter count. The bounding box loss function of YOLOv7-tiny is enhanced using the SIOU bounding box loss function [32,33]. This optimizes the adjustment process of bounding boxes during training, improving model convergence and detection accuracy. Based on these improvements, an enhanced YOLOv7-tiny fly ash detection model tailored for edge GPU devices is proposed, with its structure illustrated in Figure 1.

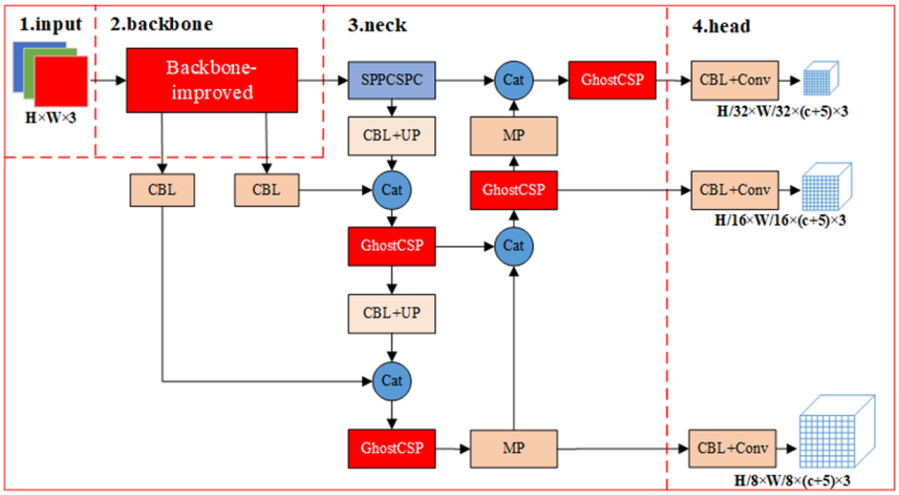


Figure 1. Structure of the improved YOLOv7-tiny model.

2.2. Improvements to the Backbone Network

To achieve lightweight backbone networks without compromising detection accuracy, this study designed the ELAN-PC module by integrating partial convolutions with the ELAN architecture. Addressing the limitations of direct partial convolutions, the module adopts the FasterNetBlock design philosophy [34] by incorporating pointwise convolutions at the end of partial convolutions to form the partial convolution block (PConvBlock). This module first processes feature maps via partial convolution and then utilizes pointwise convolution to enhance information exchange between feature map channels while adjusting the number of output feature map channels. Integrating the strengths of the partial convolution module and ELAN, the final ELAN-PC feature extraction module is formed, with its specific structure shown in Figure 2.

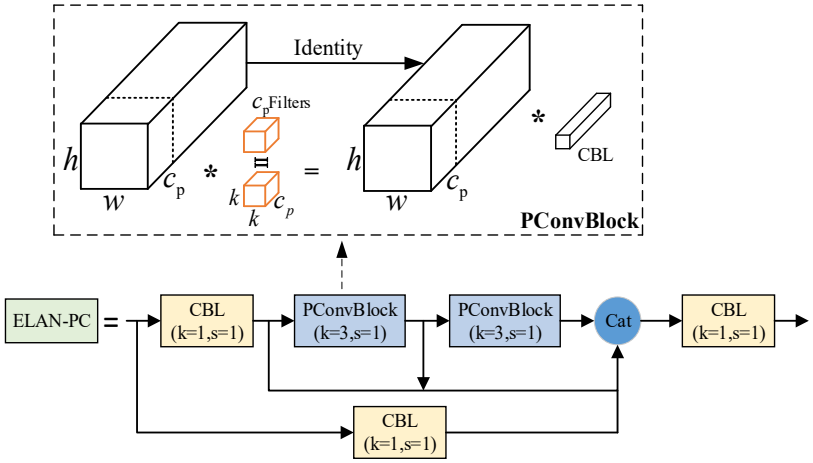


Figure 2. ELAN-PC structure.

To further clarify the computational advantage of the proposed ELAN-PC module, its mathematical formulation and complexity are analyzed. For a standard convolution with kernel size $k \times k$, input channels C_{in}, output channels C_{out}, and feature map size $H \times W$, the computational cost can be expressed as follows:

$$FLOPs_{\text{conv}} = H \times W \times C_{\text{in}} \times C_{\text{out}} \times k^2 \quad (1)$$

In contrast, partial convolution operates only on a fraction of the input channels. Let α ($0 < \alpha < 1$) denote the proportion of channels involved in convolution; then, the computational cost of partial convolution can be written as

$$FLOPs_{\text{conv}} = H \times W \times \alpha C_{\text{in}} \times C_{\text{out}} \times k^2 \quad (2)$$

By introducing pointwise convolution after partial convolution, the ELAN-PC module preserves inter-channel information exchange while significantly reducing redundant computations. When replacing the original ELAN module with ELAN-PC, the backbone network achieves a substantial reduction in floating-point operations and parameter count, while maintaining sufficient feature aggregation capability. This design enables efficient extraction of fine-grained texture and boundary features that are critical for distinguishing coal from gangue.

By replacing the original ELAN module in the YOLOv7-tiny backbone network with the ELAN-PC module, an improved network architecture is constructed as shown in Figure 3. The ELAN-PC module inherits ELAN's efficient feature fusion mechanism, effectively integrating feature information extracted from various convolutional layers.

Simultaneously, it utilizes partial convolutions to reduce redundancy among feature map channels, significantly enhancing feature extraction capabilities. The optimized network maintains feature layer output channel counts of 512, 256, and 128, ensuring sufficient capacity to learn diverse features of coal and coal gangue [27]. This effectively resolves the issue of inadequate feature learning caused by insufficient channel counts.

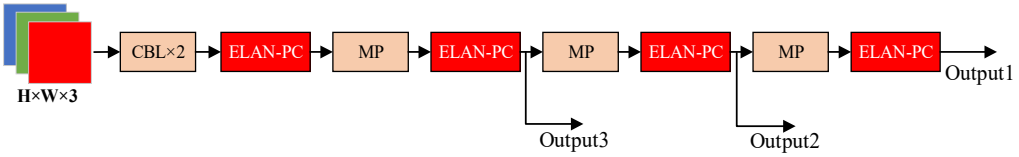


Figure 3. Backbone network structure of improved YOLOv7-tiny.

2.3. Improvements to the Neck Network

Based on the design philosophy of cross-stage local networks, the GhostCSP feature extraction module was constructed by integrating the Ghost bottleneck, with its specific structure shown in Figure 4. This module employs cross-stage local connections [35] to fuse initial feature maps with final feature maps, effectively enhancing the module's feature learning capability. This enables the model to better recognize the characteristic differences between coal and coal gangue across varying network depth levels. The Ghost bottleneck employed within the module fully leverages the advantages of phantom convolutions [36], achieving richer feature representations with fewer parameters. This significantly enhances the feature characterization capabilities for coal and gangue, enabling the model to comprehensively capture their distinctive characteristics.

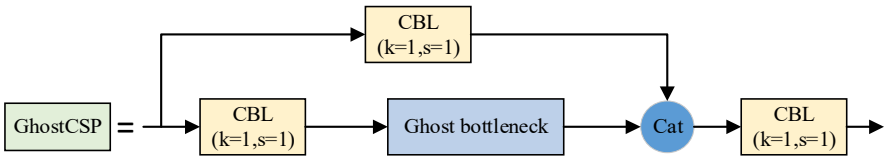


Figure 4. Structure diagram of GhostCSP.

2.4. Improvements to the Loss Function

Using L_{SIOU} as the loss function for bounding box regression, as shown in Equation (3). The improved YOLOv7-tiny loss function is defined in Equation (4).

$$\mathcal{L}_{\text{bbox_SIOU}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{I}_{ij}^{\text{obj}} \mathcal{L}_{\text{SIOU}} \tag{3}$$

where S denotes the number of grid cells along one spatial dimension of the feature map, resulting in a total of $S \times S$ grid regions, and B represents the number of candidate bounding boxes predicted at each grid cell. The indicator function I_{ij}^{obj} takes the value of 1 when the j-th predicted bounding box in the i-th grid cell is responsible for object regression, and 0 otherwise.

$$\mathcal{L} = \lambda_{\text{obj}} \mathcal{L}_{\text{obj}} + \lambda_{\text{cls}} \mathcal{L}_{\text{cls}} + \lambda_{\text{bbox}} \mathcal{L}_{\text{bbox_SIOU}} \tag{4}$$

Here, $\mathcal{L}$ denotes the total loss function used for model optimization. $\mathcal{L}_{\text{obj}}$ represents the objectness loss, which measures the confidence of target presence in each predicted bounding box. $\mathcal{L}_{\text{cls}}$ denotes the classification loss, which is responsible for distinguishing coal from gangue. $\mathcal{L}_{\text{bbox_SIOU}}$ is the bounding box regression loss based on the SIoU metric. The coefficients λ_{obj}, λ_{cls}, and λ_{bbox} are weighting factors used to balance the contributions of different loss components during training.

Figure 5 illustrates the spatial relationship between ground truth and predicted boxes in the SIOU loss function. When optimizing bounding boxes, this loss function not only considers the spatial distance, geometric shape, and overlap between ground truth and predicted boxes but also introduces an angular deviation factor between them. This multidimensional optimization strategy [37] enables the model to learn the geometric features and spatial positions of target objects more precisely from ground truth boxes. It significantly enhances training efficiency, allowing the model to achieve highly accurate predictions with fewer training iterations, thereby effectively improving object detection accuracy.

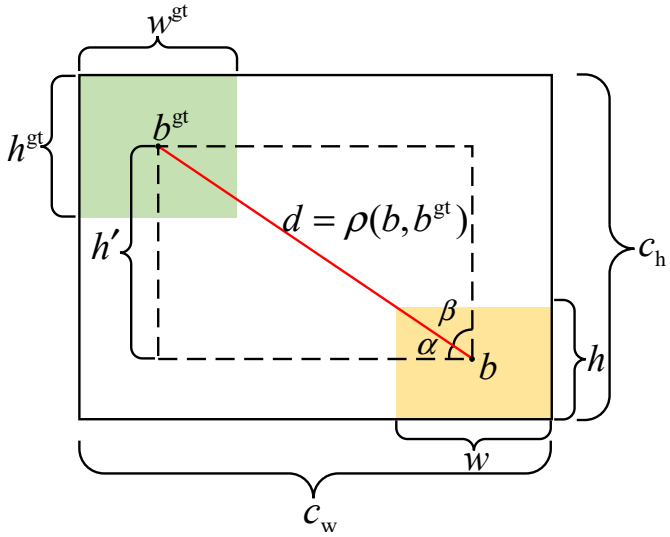


Figure 5. Schematic representation of the real and projected frames (SIOU).

3. Testing and Results Analysis

3.1. Experimental Data and Training Platform

The dataset used in this study was constructed from images collected on an experimental coal gangue sorting platform, covering diverse operating conditions. The original dataset consists of 3742 raw images, which were manually annotated following the YOLO format.

To improve model robustness and generalization ability, data augmentation techniques including random rotation, brightness adjustment, Gaussian noise injection, and motion blur simulation were applied. These augmentation strategies were selected to simulate common industrial disturbances such as belt vibration, illumination variation, and camera noise.

In addition, a subset of challenging samples was constructed to evaluate model robustness under adverse conditions. Challenging samples are defined as images containing at least one of the following characteristics: (1) severe target occlusion caused by high-density stacking, (2) motion blur due to conveyor belt movement, (3) uneven illumination or shadows, and (4) surface contamination such as moisture or mud. These samples are representative of real-world coal preparation environments and are used for targeted performance evaluation.

Model training equipment and parameters used in the experiment: Operating System: Ubuntu 18.04 GPU: RTX 3090 CPU: Intel® Xeon® Gold 5218 CPU CUDA: CUDA 11.3 cudnn: cudnn 8.2.

3.2. Experimental Results and Analysis

3.2.1. Ablation Study and Mechanism Analysis

To further clarify the contribution and underlying mechanism of each proposed improvement, a mechanism-oriented ablation study was conducted. Beyond the overall detection accuracy (mAP@0.5), additional evaluation metrics were introduced, including small-object detection performance (AP_{small}), recall under occlusion scenarios ($Recall_{occ}$), and localization accuracy (Mean IoU). The experimental results are summarized in Table 1.

Table 1. Mechanism-oriented ablation study of the improved YOLOv7-tiny.

Number	ELAN-PC	GhostCSP	SIOU	AP_{small} (%)	mAP0.5 (%)	$Recall_{occ}$ (%)	Mean IoU	FLOPs (G)
M1				81.3	86.9	78.6	0.684	13.2
M2	√			82.6	87.2	79.1	0.689	10.1
M3	√	√		83.4	87.5	80.5	0.692	9.2
M4	√	√	√	85.1	88.7	83.2	0.716	9.2

Compared with the baseline YOLOv7-tiny (M1), introducing the ELAN-PC module (M2) leads to a noticeable improvement in AP_{small} while significantly reducing computational complexity. This indicates that partial convolution effectively alleviates channel redundancy and enhances the representation of fine-grained texture and boundary features, which are critical for identifying small-sized coal gangue fragments.

By further incorporating the GhostCSP module into the neck network (M3), the model achieves a higher $Recall_{occ}$, demonstrating improved robustness in densely stacked and partially occluded scenarios. This improvement can be attributed to the cross-stage feature aggregation mechanism of GhostCSP, which strengthens semantic consistency between shallow and deep feature maps and enhances feature reuse across network stages.

Finally, replacing the original bounding box regression loss with the SIOU loss function (M4) results in a significant increase in Mean IoU, while maintaining the same computational cost. The introduction of distance, angle, and shape constraints in SIOU enables more accurate geometric alignment between predicted and ground-truth bounding boxes, leading to improved localization precision and overall detection performance.

3.2.2. Model Comparison Analysis

To further validate the advantages of the proposed model, a comparative analysis was performed between our model and several popular object detection models currently in use. The detailed parameters of each compared model are provided in Table 2. Experimental results indicate that the SSD [38] and YOLOv8-N [39] models, both employing a MobileNet backbone network, exhibit excellent computational efficiency and parameter compactness. However, their mAP0.5 scores remain relatively low, making it difficult to meet the accuracy requirements for coal gangue detection tasks. The YOLOv4-tiny model delivered generally moderate overall performance. Among the five compared models, YOLOv5-S achieved the highest mAP0.5, yet it incurs a relatively high computational load and parameter count, which may hinder its deployment efficiency on edge devices. The YOLOv7-tiny model showed high detection accuracy, but its computational demand remains considerable.

Table 2. Comprehensive information of comparative experiment (YOLOv7-tiny).

Model	Dimensions	mAP0.5(%)	FPS (Frames per Second)	FLOPs (G)	Params (M)
SSD-mobilenet	640^2	86.7	33	6.1	3.7
YOLOv4-tiny	640^2	86.4	48	16.2	5.9
YOLOv5-S	640^2	87.2	34	16.5	7.1
YOLOv7-tiny	640^2	86.9	41	13.2	6.0
YOLOv8-N	640^2	86.5	50	8.2	3.0
YOLOv7-tinyimproved	640^2	88.7	43	9.2	4.3

In comparison, the improved YOLOv7-tiny model not only attains the best mAP0.5 but also maintains low computational cost and parameter volume. For applications on edge GPU devices, model selection must comprehensively consider detection accuracy and device capability, with mAP0.5, computational complexity, and parameter count serving as key evaluation metrics. The experimental findings demonstrate that the improved YOLOv7-tiny model achieves lightweight optimization without compromising detection accuracy, showing significant advantages in edge-device deployment for coal gangue detection tasks.

3.3. The Engineering Correlation of the Comparative Experiment

In addition to detection accuracy, computational efficiency and memory footprint are critical factors for deployment on edge GPU devices. The proposed model achieves a favorable trade-off between accuracy and efficiency by significantly reducing FLOPs and parameter size compared with standard YOLO-based detectors.

Compared with other lightweight models such as YOLOv8-N and MobileNet-based detectors, the improved YOLOv7-tiny model demonstrates superior robustness under complex industrial conditions while maintaining real-time inference capability on resource-constrained devices. This balance is essential for practical deployment in coal preparation plants, where both detection reliability and hardware limitations must be considered.

4. Application Testing of Coal Gangue Sorting System

The dynamic coal gangue sorting experiments in this study were conducted on a laboratory-scale intelligent sorting platform, as shown in Figure 6, with a conveyor belt speed of 0.6 m/s. At the current stage, the proposed detection algorithm has not yet been deployed on an actual production conveyor in a coal preparation plant. This experimental setting was primarily adopted to enable controlled evaluation of the detection accuracy, sorting effectiveness, and system stability. Large-scale field testing under full industrial operating conditions is planned as future work.

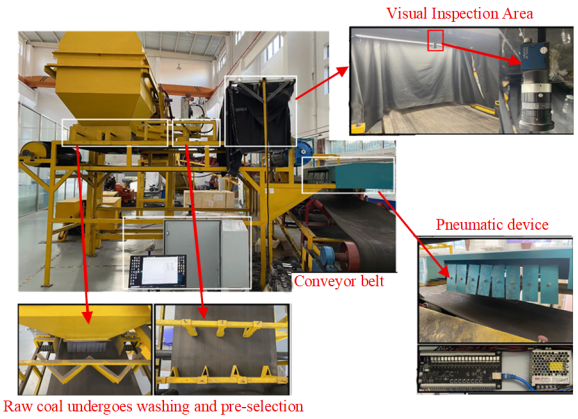


Figure 6. Physical diagram of the experimental platform for the sorting system.

In real coal preparation plants, the visual detection of coal gangue is often challenged by complex environmental factors, among which target occlusion caused by high-density stacking and image quality degradation due to surface moisture or mud contamination are the most critical. To evaluate the robustness of the proposed method under such conditions, additional analyses were conducted on test samples exhibiting partial occlusion and dense stacking. The results indicate that the improved YOLOv7-tiny model achieves higher recall in occluded scenarios compared with the baseline model. This improvement can be attributed to the enhanced cross-stage feature aggregation capability introduced by the GhostCSP module, which strengthens semantic consistency across feature layers and improves discrimination under overlapping conditions.

Regarding surface moisture and mud interference, the training process incorporated targeted data augmentation strategies, including random brightness adjustment, contrast variation, and blur perturbation, to simulate image degradation commonly observed in industrial environments. Experimental results demonstrate that the proposed model maintains stable detection performance under these simulated adverse visual conditions, indicating improved robustness to illumination variations and surface contamination.

Dynamic sorting experiments for coal gangue were conducted on the laboratory-built coal gangue sorting system test platform. A total of 300 samples were tested, comprising 210 coal samples and 90 coal gangue samples, with a belt speed of 0.6 m/s.

Figure 7a–f illustrate the complete process of dynamic coal gangue sorting. Blue boxes denote coal, red boxes indicate coal gangue, and white boxes represent the cylinder assembly equipped with impact plates. The specific sequence is as follows: Figure 7a shows coal and gangue falling sequentially; Figure 7b depicts gangue approaching the cylinder's impact range as the cylinder prepares to strike; Figure 7c shows the cylinder rapidly actuating upon the gangue entering the effective impact zone, accurately ejecting it into the designated hopper; Figure 7d depicts the device resetting after completing the impact; Figure 7e demonstrates the device reactivating as new gangue enters the impact zone while coal continues normal conveyance; Figure 7f shows the device in its reset standby state. Specific data from relevant sorting experiments are detailed in Table 3.

Table 3. Results of the sorting experiment.

Model	Number of Coal	Coal Gangue Count	Number of Incorrect Selections (Coal)	Number of Selected Pieces (Scrap)	Sorting Efficiency (%)	Wrong Selection Rate (%)
YOLOv7-tiny improved	210 data	90 data	17	74	82.2	8.1

In coal gangue sorting experiments, experimental data showed that when using YOLOv7-tinyimproved as the detection model, the gangue sorting accuracy reached 82.2%, with a coal misclassification rate of 8.1%. However, the model still exhibited some missed detections and misclassifications during actual sorting operations, as illustrated in Figure 8. These primarily manifested in three types: failure to identify gangue, misclassifying gangue as coal, and misclassifying coal as gangue. Despite these recognition errors, the sorting system built upon this improved model demonstrated reliable performance, providing crucial reference data for subsequent model optimization and system enhancement.

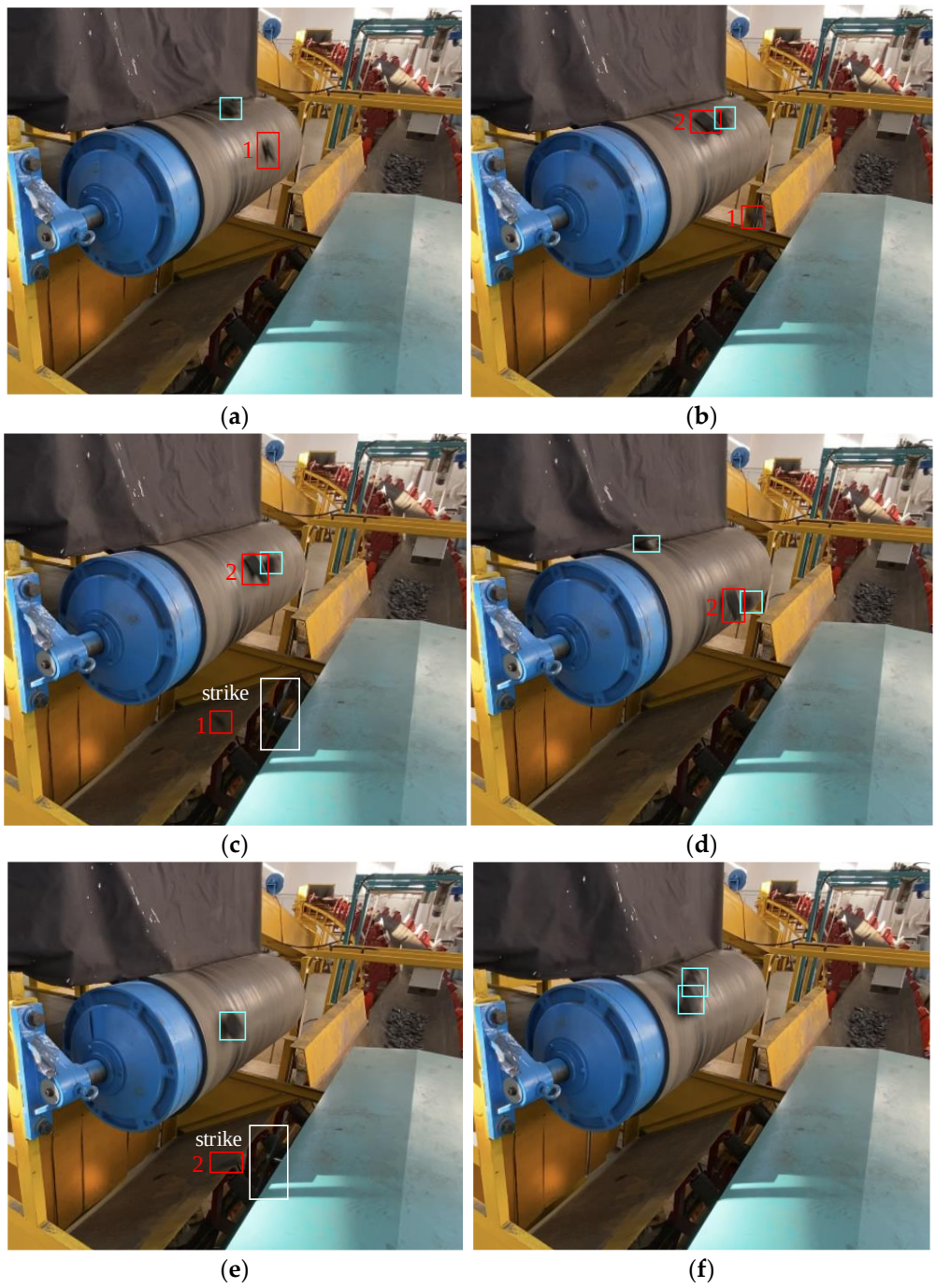


Figure 7. Dynamic process of sorting (**a**) coal gangue and coal falling process-1; (**b**) coal gangue enters the cylinder impact zone; (**c**) cylinder impacting coal gangue-1; (**d**) coal gangue and coal falling process-2; (**e**) cylinder impacting coal gangue-2; (**f**) waiting for the coal gangue to fall.

Although the proposed method demonstrates strong performance on the experimental platform, several practical considerations should be noted for real-world deployment. On edge GPU devices, inference speed and memory usage remain key constraints, especially under high-throughput conveyor conditions.

While the lightweight design significantly reduces computational overhead, extreme cases such as severe occlusion, heavy surface contamination, or abrupt illumination changes may still affect detection accuracy. Future work will focus on large-scale field testing and further optimization to enhance robustness under fully industrial operating conditions.

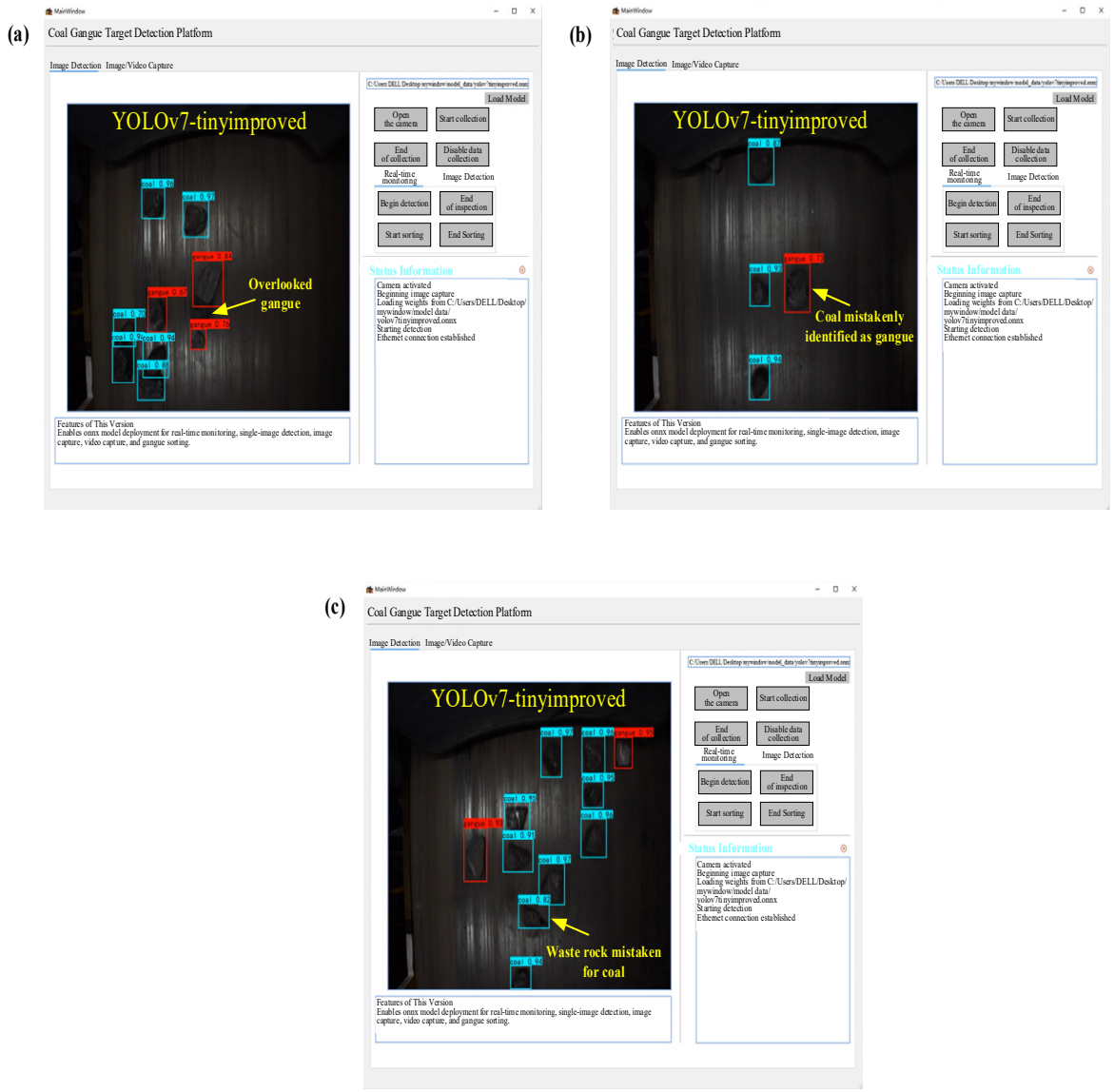


Figure 8. Sorting system experimental platform physical diagram: (**a**) YOLOv7-tinyimproved omission of inspection; (**b**) YOLOv7-tinyimproved omission of inspection-1; (**c**) YOLOv7-tinyimproved omission of inspection-2.

5. Conclusions

This paper builds upon the YOLOv7-tiny model, integrating recent advancements in computer vision to enhance the architecture. It proposes an improved YOLOv7-tiny method for coal gangue detection tailored for low-computational-power and low-memory GPU devices. Based on this model, a coal gangue sorting system was constructed, and dynamic sorting experiments were conducted, validating the model's feasibility. The model introduces the ELAN-PC feature extraction module for partial convolutions and ELAN. Subsequently, the GhostCSP feature extraction module is developed based on the Ghost bottleneck concept, inspired by cross-stage aggregation networks. Finally, the bounding box loss function of YOLOv7-tiny was enhanced using the SIOU bounding box loss function. Training results demonstrate that for input images at 640 × 640 resolution, the improved YOLOv7-tiny model achieved an mAP0.5 of 88.7%, representing a 1.8% increase over the baseline YOLOv7-tiny (86.9%). Floating-point operations decreased from 13.2 G to 9.2 G (30% reduction), while the number of parameters decreased from 6.0 M to 4.3 M, a reduction of 28%. Experimental results demonstrate that the coal gangue sorting system experimental

platform based on the improved YOLOv7-tiny model achieved a coal gangue sorting rate of 82.2% and a coal misclassification rate of 8.1%. These results lay the foundation for subsequent model refinements and system upgrades.

Author Contributions: Conceptualization, S.H. and J.M.; methodology, Z.Z.; validation, Y.L.; formal analysis, D.W.; investigation, L.Z.; resources, P.Z.; writing—original draft preparation, K.Z. and M.D.; writing—review and editing, J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by The 7th Young Talent Support Program (Project No. 2021QNRC001); The Shandong Provincial University Youth Innovation and Technology Support Program (Project No.: 2023KG304).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Shenglei Hao, Zhenyang Zhang, Yong Liu and Dongxu Wu were employed by the company Yankuang Energy Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Li, J.; Wang, J. Comprehensive utilization and environmental risks of coal gangue: A review. *J. Clean. Prod.* **2019**, *239*, 117946. [CrossRef]
2. Zheng, Q.; Zhou, Y.; Liu, X.; Liu, M.; Liao, L.; Lv, G. Environmental hazards and comprehensive utilization of solid waste coal gangue. *Prog. Nat. Sci. Mater. Int.* **2024**, *34*, 223–239. [CrossRef]
3. Jiang, J.; Han, Y.; Zhao, H.; Suo, J.; Cao, Q. Recognition and sorting of coal and gangue based on image process and multilayer perceptron. *Int. J. Coal Prep. Util.* **2023**, *43*, 54–72. [CrossRef]
4. Fu, C.; Lu, F.; Zhang, G. Gradient-enhanced waterpixels clustering for coal gangue image segmentation. *Int. J. Coal Prep. Util.* **2023**, *43*, 677–690. [CrossRef]
5. Zhang, F.; Luo, C.; Xu, J.; Luo, Y.; Zheng, F.-C. Deep learning based automatic modulation recognition: Models, datasets, and challenges. *Digit. Signal Process.* **2022**, *129*, 103650. [CrossRef]
6. Sun, Y.; Huang, J.; Cheng, Y.; Zhang, J.; Shi, Y.; Pan, L. High-accuracy dynamic gesture recognition: A universal and self-adaptive deep-learning-assisted system leveraging high-performance ionogels-based strain sensors. *SmartMat* **2024**, *5*, e1269. [CrossRef]
7. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. [CrossRef]
8. Li, R.; Wu, Y. Improved YOLO v5 wheat ear detection algorithm based on attention mechanism. *Electronics* **2022**, *11*, 1673. [CrossRef]
9. Xu, L.; Dong, S.; Wei, H.; Ren, Q.; Huang, J.; Liu, J. Defect signal intelligent recognition of weld radiographs based on YOLO V5-IMPROVEMENT. *J. Manuf. Process.* **2023**, *99*, 373–381. [CrossRef]
10. Kim, J.H.; Kim, N.; Park, Y.W.; Won, C.S. Object detection and classification based on YOLO-V5 with improved maritime dataset. *J. Mar. Sci. Eng.* **2022**, *10*, 377. [CrossRef]
11. Argyros, I.K. *The Theory and Applications of Iteration Methods*; CRC Press: Boca Raton, FL, USA, 2022. [CrossRef]
12. Shen, J.; Chen, Y.; Liu, Y.; Zuo, X.; Fan, H.; Yang, W. ICAFusion: Iterative cross-attention guided feature fusion for multispectral object detection. *Pattern Recognit.* **2024**, *145*, 109913. [CrossRef]
13. Dong, X.; Xiong, G.; Guan, X.; Zhang, C. Constitutive Modeling of Coal Gangue Concrete with Integrated Global–Local Explainable AI and Finite Element Validation. *Buildings* **2025**, *15*, 3007. [CrossRef]
14. Kamal Al-anni, M.; Abdullah, A.A.; Drap, P. Automatic Deep-Sea Amphorae Detection Using Optimal 2D Ultralytics Deep Learning. *Int. J. Comput.* **2025**, *18*, 1–12. [CrossRef]
15. Mahasin, M.; Dewi, I.A. Comparison of cspdarknet53, cspresnext-50, and efficientnet-b0 backbones on yolo v4 as object detector. *Int. J. Eng. Sci. Inf. Technol.* **2022**, *2*, 64–72. [CrossRef]
16. Wang, J.; Zhang, Z.; Dai, B.; Zhao, K.; Shen, W.; Yin, Y.; Li, Y. Cow-YOLO: Automatic cow mounting detection based on non-local CSPDarknet53 and multiscale neck. *Int. J. Agric. Biol. Eng.* **2024**, *17*, 193–202. [CrossRef]
17. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W. YOLOv6: A single-stage object detection framework for industrial applications. *arXiv* **2022**, arXiv:2209.02976. [CrossRef]

18. Weng, K.; Chu, X.; Xu, X.; Huang, J.; Wei, X. Efficientrep: An efficient repvgg-style convnets with hardware-aware neural network design. *arXiv* **2023**, arXiv:2302.00386. [CrossRef]
19. Nandi, D.; Roy, S.; Prasad, A.; Patra, S.N. High precision automatic coronal hole detection from January 2019 to July 2023 using the AIA 193 Å data obtained by solar dynamic observatory. *Serbian Astron. J.* **2025**, *3*, 71–90. [CrossRef]
20. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. You only learn one representation: Unified network for multiple tasks. *arXiv* **2021**, arXiv:2105.04206. [CrossRef]
21. Li, S.; Wang, Z.; Liu, Z.; Tan, C.; Lin, H.; Wu, D.; Chen, Z.; Zheng, J.; Li, S.Z. Moganet: Multi-order gated aggregation network. *arXiv* **2022**, arXiv:2211.03295. [CrossRef]
22. Hoyer, S.; Sohl-Dickstein, J.; Greydanus, S. Neural reparameterization improves structural optimization. *arXiv* **2019**, arXiv:1909.04240. [CrossRef]
23. Zhang, Y.; Quan, S.; Xiao, H.; Liu, J.; Shao, Z.; Wang, Z.; Peng, Y.; Li, H. A YOLO-based Polymerized Head-auxiliary Structures for Target Detection in Remote Sensing Images. *Pattern Recognit.* **2025**, *174*, 112961. [CrossRef]
24. Chen, J.; Wang, G.; Liu, W.; Zhong, X.; Tian, Y.; Wu, Z. Perception reinforcement using auxiliary learning feature fusion: A modified yolov8 for head detection. In *Proceedings of the 2023 China Automation Congress (CAC), Chongqing, China, 17–19 November 2023*; IEEE: New York, NY, USA, 2023; pp. 4709–4714. [CrossRef]
25. Zhang, Y.; Luo, C. A dynamic label assignment strategy for one-stage detectors. *Neurocomputing* **2024**, *577*, 127383. [CrossRef]
26. Yuan, M.; Zhang, L.; Li, X.Y.; Yang, L.-Z.; Xiong, H. Adaptive model scheduling for resource-efficient data labeling. *ACM Trans. Knowl. Discov. Data (TKDD)* **2022**, *16*, 71. [CrossRef]
27. Liu, Q.; Li, J.; Li, Y.; Gao, M. Recognition methods for coal and coal gangue based on deep learning. *IEEE Access* **2021**, *9*, 77599–77610. [CrossRef]
28. Zhang, Y.; Wang, J.; Yu, Z.; Zhao, S.; Bei, G. Research on intelligent detection of coal gangue based on deep learning. *Measurement* **2022**, *198*, 111415. [CrossRef]
29. Korlapati, N.V.S.; Khan, F.; Noor, Q.; Mirza, S.; Vaddiraju, S. Review and analysis of pipeline leak detection methods. *J. Pipeline Sci. Eng.* **2022**, *2*, 100074. [CrossRef]
30. Chen, L.; Sun, R.; Chen, D.; Tang, W.; Feng, H. Small Target Detection on Water Based on Improved YOLOv7 Unmanned Vessel. In *Proceedings of the 2024 36th Chinese Control and Decision Conference (CCDC), Xi'an, China, 25–27 May 2024*; IEEE: New York, NY, USA, 2024; pp. 5607–5613. [CrossRef]
31. Li, S.; Fu, X.; Dong, J. Improved ship detection algorithm based on YOLOX for SAR outline enhancement image. *Remote Sens.* **2022**, *14*, 4070. [CrossRef]
32. Gevorgyan, Z. SIoU loss: More powerful learning for bounding box regression. *arXiv* **2022**, arXiv:2205.12740. [CrossRef]
33. Ren, F.; Fei, J.; Li, H.; Doma, B.T. Steel surface defect detection using improved deep learning algorithm: ECA-SimSPPF-SIoU-Yolov5. *IEEE Access* **2024**, *12*, 32545–32553. [CrossRef]
34. Qin, C.; Zhou, Z. YOLO-FGD: A fast lightweight PCB defect method based on FasterNet and the Gather-and-Distribute mechanism. *J. Real-Time Image Process.* **2024**, *21*, 122. [CrossRef]
35. Lu, Y.F.; Yu, Q.; Gao, J.W.; Li, Y.; Zou, J.-C.; Qiao, H. Cross stage partial connections based weighted bi-directional feature pyramid and enhanced spatial transformation network for robust object detection. *Neurocomputing* **2022**, *513*, 70–82. [CrossRef]
36. Liu, Y.; Tang, K.; Cai, W.; Chen, A.; Zhou, G.; Li, L.; Liu, R. MPC-STANet: Alzheimer's disease recognition method based on multiple phantom convolution and spatial transformation attention mechanism. *Front. Aging Neurosci.* **2022**, *14*, 918462. [CrossRef]
37. Wang, X.; Li, Y.; Yang, K.; Xu, Z.; Zhang, J. A regionally coordinated allocation strategy for medical resources based on multidimensional uncertain information. *Inf. Sci.* **2024**, *666*, 120384. [CrossRef]
38. Zhai, S.; Shang, D.; Wang, S.; Dong, S. DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion. *IEEE Access* **2020**, *8*, 24344–24357. [CrossRef]
39. Wang, Z.; Chen, Y.; Ding, H.; Liu, Z.; Xie, Y.; Zhou, K.; Feng, P.; Yu, Y.; Li, C.; Zhang, D.; et al. YOLOv8-DBW: An Improved High-Accuracy Fast Landslide Detection Model. *Trans. GIS* **2025**, *29*, e70021. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.