energies

MDPI

Article

Coal Mine Belt Conveyor Foreign Objects Recognition Method of Improved YOLOv5 Algorithm with Defogging and Deblurring

Qinghua Mao [1,2], Shikun Li [1,2,*], Xin Hu [1,2] and Xusheng Xue [1,2]

[1] School of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an 710054, China
[2] Shaanxi Key Laboratory of Mine Electromechanical Equipment Intelligent Detection and Control, Xi'an 710054, China
* Correspondence: 20205224051@stu.xust.edu.cn

Abstract: The belt conveyor is the main equipment for underground coal transportation. Its coal flow is mixed with large coal, gangue, anchor rods, wooden strips, and other foreign objects, which easily causes failure of the conveyor belt, such as scratching, tearing, and even broken belts. Aiming at the problem that it was difficult to accurately identify the foreign objects of underground belt conveyors due to the influence of fog, high-speed operation, and obscuration, the coal mine belt conveyor foreign object recognition method of improved YOLOv5 algorithm with defogging and deblurring was proposed. In order to improve the clarity of the monitoring video of the belt conveyor, the dark channel priori defogging algorithm is applied to reduce the impact of fog on the clarity of the monitoring video, and the image is sharpened by user-defined convolution method to reduce the blurring effect on the image in high-speed operation condition. In order to improve the precision of foreign object identification, the convolution block attention module is used to improve the feature expression ability of the foreign object in the complex background. Through adaptive spatial feature fusion, the multi-layer feature information of the foreign object image is more fully fused so as to achieve the goal of accurate recognition of foreign objects. In order to verify the recognition effect of the improved YOLOv5 algorithm, a comparative test is conducted with self-built data set and a public data set. The results show that the performance of the improved YOLOv5 algorithm is better than SSD, YOLOv3, and YOLOv5. The belt conveyor monitoring video of resolution for 1920 × 1080 in Huangling Coal Mine is used for identification verification, the recognition accuracy can reach 95.09%, and the recognition frame rate is 56.50 FPS. The improved YOLOv5 algorithm can provide a reference for the accurate recognition of targets in a complex underground environment.

Keywords: belt conveyor; foreign objects recognition; YOLOv5; attention mechanism; adaptive spatial feature fusion; deep learning

Citation: Mao, Q.; Li, S.; Hu, X.; Xue, X. Coal Mine Belt Conveyor Foreign Objects Recognition Method of Improved YOLOv5 Algorithm with Defogging and Deblurring. *Energies* **2022**, *15*, 9504. https://doi.org/10.3390/en15249504

Academic Editor: Pavel A. Strizhak

Received: 13 October 2022
Accepted: 12 December 2022
Published: 14 December 2022

1. Introduction

The belt conveyor is an important transportation equipment in the coal mine. The environment of a coal mine underground is complex, and the coal flow of the belt conveyor is mixed with large coals, gangues, anchor rods, angle irons, wood, and other foreign objects [1,2]. If these foreign objects are not cleaned in time, they will easily get stuck and even cause a tear in the conveyor belt, affecting normal mining of coal mines [3]. In recent years, the "Guidance on Accelerating the Development of Intelligent Coal Mine" and the "Guide to Intelligent Coal Mine Construction 2021 Edition" were published in china, which requires vigorously promoting intelligent development of coal mine transportation and realizing intelligent identification of foreign objects on belt conveyor [4,5]. Therefore, it is of great significance for ensuring the normal operation of belt conveyors and efficient mining of coal mines to study intelligent identification methods of foreign objects on belt conveyors.

At present, the common methods for foreign object identification of belt conveyors in coal mines are manual sorting, iron remover method, spectral detection, and image identification [6]. Although the manual sorting method is widely used, it is labor-intensive and inefficient, which does not meet the requirements of intelligent and unmanned. The iron remover method cannot be applicable to non-metal foreign objects such as gangue and wood. The spectral detection method has low efficiency. Compared with traditional image feature recognition algorithm, target recognition algorithm based on deep learning has stronger robustness and better generalization ability in extracting abstract features and has better applicability in high-speed recognition.

In recent years, many scholars worked on using machine vision and deep learning methods to achieve the identification of foreign objects on belt conveyors. Wu Shoupeng et al. [7] proposed a foreign objects identification model based on Faster-RCNN and a two-feature network pyramid to achieve accurate identification of gangues and anchor rods on belt conveyors. LV Zhiqiang [8] used improved Faster R-CNN with VGG16 as a feature extraction network to realize foreign object recognition such as gangue and iron. Ren Zhiling et al. [9] proposed a foreign objects recognition algorithm based on improved CenterNet in view of a large difference in size and uneven distribution of foreign objects, which improved the recognition precision of the network and reduced false detection rate and missing detection rate, but the detection speed is slow. Guanchao Ma et al. [10] proposed an improved CenterNet center algorithm for the fast-running speed of the coal conveying belt and the influence of background light source on the detected object. Hu Jinghao et al. [11] used the focal loss function to replace the cross-entropy loss function to improve the YOLOv3 model, which improves the prediction performance of the model for anchor rods, angle iron, and nuts, and the confidence are above 94%. Wang et al. [12] put forward a video recognition method based on SSD for foreign objects, the depth separable convolution and giou loss function are used to improve detection speed to 90.2%. Du Jingyi [13] lightened the YOLOv3 model and deployed it on Jetson Xavier NX and increased the detection speed to 30.7 FPS. Hao Shuai et al. [2] proposed a YOLOv5 recognition algorithm based on the convolution block attention model (CBAM) for coal dust interference and high-speed operation, and the recognition precision can reach 94.7%. Cheng Deqiang et al. [14] proposed a lightweight network integrating residual information, which improves recognition precision and recognition speed of gangues and anchor rods. Xiao Dong et al. [15] proposed a foreign object detection method based on YOLOv3. After the implementation of channel pruning and layer pruning strategies, the processing speed of the model is faster. To sum up, there are many pieces of research on the identification of large gangue and anchor bolts on belt conveyors using the depth learning method, but there are many kinds of foreign matters in the existing belt conveyor, and the identification method is affected by underground fog, high-speed operation of belt conveyor, and coal block shielding, which leads to the problems of low accuracy and low recall in the video identification of a variety of foreign matters.

Regarding the issue above, we proposed a coal mine belt conveyor foreign objects recognition method of improved YOLOv5 algorithm with defogging and deblurring. The improved method improves the clarity of the image and enhances the contour features of foreign objects through the dark channel prior to defogging and user-defined convolution sharpening methods. The improved method by integrating CBAM into the C3 module to improve the ability of the C3 module to extract key features of foreign objects under complex backgrounds. The adaptively spatial feature fusion (ASFF) method is used to improve the ability to fuse multi-dimensional information of foreign object images, so as to finally improve the precision of foreign object recognition. It is of great significance to the safe and efficient operation of belt conveyors.

2. Algorithm Improvement

2.1. Video Image Preprocessing

The monitoring image of the belt conveyor in the Huangling coal mine is presented in Figure 1. It can be seen from Figure 1 that the image is affected by underground fog and high-speed operation, and monitoring video quality of the belt conveyor is low, which will directly affect the labeling of the data set and identification of foreign objects. Defogging and motion blur removal of monitoring image of belt conveyor can effectively improve the clarity of the image, make the contour, color, and other features of foreign objects more obvious, and even improve the quality of data set and recognition precision of foreign objects. Therefore, it is necessary to pre-process the monitoring video of the belt conveyor before foreign object recognition. The preprocessing methods for monitoring the image of the belt conveyor in the coal mine are as follows.

Figure 1. Monitoring image of belt conveyor in the coal mine.

2.1.1. Image Defogging Method Based on Dark Channel Prior

A priori defogging method based on the dark channel is used to reduce the influence of fog for monitoring video clarity [16,17], because the underground is not a closed indoor environment, which communicates with the ground atmosphere environment through the ventilation system. So, the atmospheric scattering is similar to the atmospheric environment. Dark channel priori is a simple and effective image priori rule obtained from outdoor fog-free image database. In each local area of most outdoor fog-free images, there is at least one-color channel with low intensity. Therefore, these dark channels can be used to evaluate the influence of fog on the image. According to McCarney's atmospheric scattering model, a better defogging image can be obtained according to the physical characteristics of light transmission on foggy days. McCarney's atmospheric scattering model is presented in Equation (1).

$$I(x) = t(x)J(x) + (1 - t(x))A \tag{1}$$

where $I(x)$ is the light intensity of the input image, $J(x)$ is clear images, $t(x)$ represents light transmittance, A is atmospheric light intensity, $(1 - t(x))A$ is the atmospheric light term. Priori defogging of the dark channel is to estimate atmospheric light intensity A and transmittance $t(x)$, and get restored image $J(x)$ with known fog images $I(x)$. The prior calculation formula of the dark channel is presented in Equation (2).

$$J^{dark}(x) = \min_{c \in \{r,g,b\}} \left(\min_{y \in \Omega(x)} (J_c(y)) \right) \tag{2}$$

where $\Omega(x)$ is a square neighborhood centered on x, J_c is a channel for three primary colors of J, $J^{dark}(x)$ is the dark channel of image J in the adjacent area.

In this way, images with defogging can be output by combining the results of dark channel calculation with the atmospheric scattering model.

2.1.2. Image Enhancement Method of User-Defined Convolution Kernel

High-speed operation of the belt conveyor will lead to motion blur of objects in the monitoring screen. In order to reduce the blurring of foreign objects' images, the image is sharpened with a user-defined 3 × 3 convolution kernel [18,19]. Enhanced images can be obtained by superimposing the original image and the Laplace image. The second-order differential value changes greatly for the edge of foreign objects. Therefore, the isotropic

Laplace differential operator can be used to detect gray level mutation of the image and finally, achieve the effect of image clarity enhancement. The discrete form transformation of the second-order difference of image by the Laplace operator is presented in Equation (3).

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \tag{3}$$

where $\nabla^2 f$ is Laplace transform.

2.2. Foreign Objects Recognition Method of Improved YOLOv5 Algorithm

2.2.1. Improved YOLOv5 Algorithm

Compared with multi-stage target recognition algorithms such as Faster R-CNN [20] and Cascade R-CNN [21], YOLOv5 [22] has a faster recognition speed, but recognition precision is lower. Therefore, it is necessary to improve the recognition precision of YOLOv5 with the condition of maintaining real-time. The improved YOLOv5 algorithm structure is presented in Figure 2. The improvement of the YOLOv5 algorithm structure has the following two parts. The improved parts are presented in red modules in Figure 2.

(1) CBAM is integrated into the C3 module of the backbone network to form CBAM-C3.
(2) The adaptive spatial feature fusion is added to the neck.

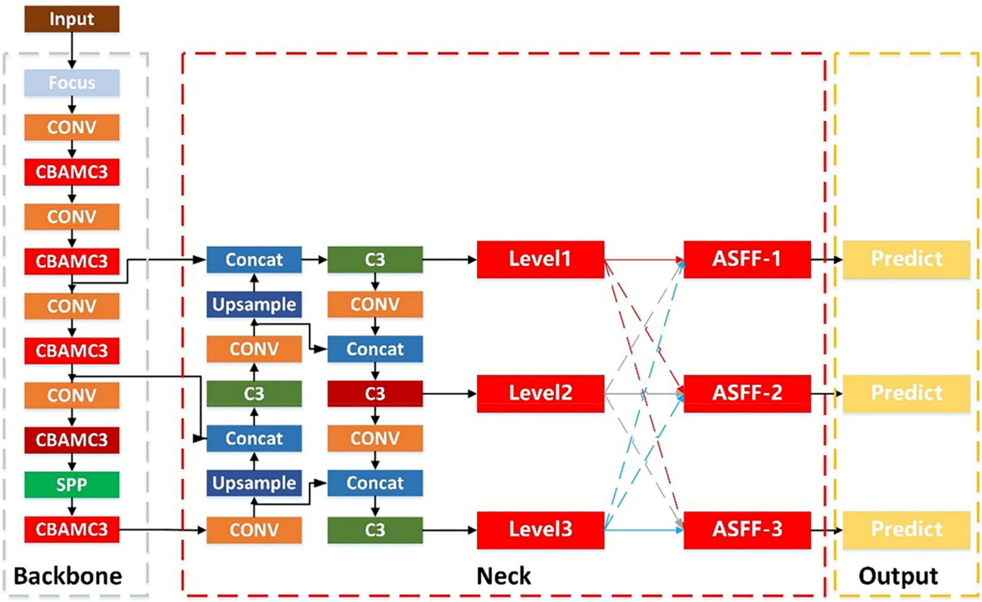


Figure 2. Structure of improved YOLOv5 algorithm.

2.2.2. Method of Improving Precision of Foreign Object Identification

The underground environment of the coal mine is special. The complex background of the belt conveyor makes it difficult to extract the key features of foreign objects, which greatly affects the precision of foreign object recognition. CBAM can reduce interference of complex backgrounds in coal mines to foreign object recognition, highlight key features of foreign objects and enhance the ability to extract key features of foreign objects. CBAM [23] was integrated into the C3 module. CBAM-C3 is used to better extract features of foreign objects and fuse feature information of different scales.

The spatial attention module (SAM) treats features in each channel equally, so it ignores the information interaction between channels. Channel attention module (CAM) processes information in a channel globally. It ignores the information interaction in the channel. The structure of CBAM-C3 is presented in Figure 3. It can be seen from Figure 4 that CBAM is composed of CAM and SAM, which avoids the disadvantages of single spatial attention or single channel attention.

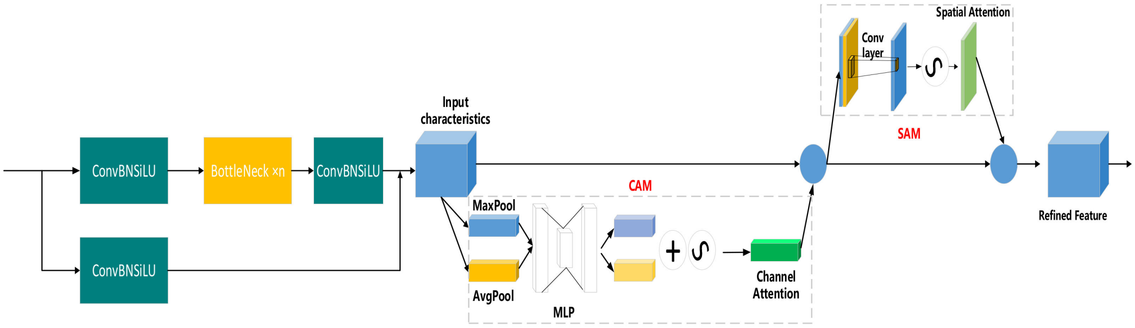


Figure 3. Structure of CBAM.

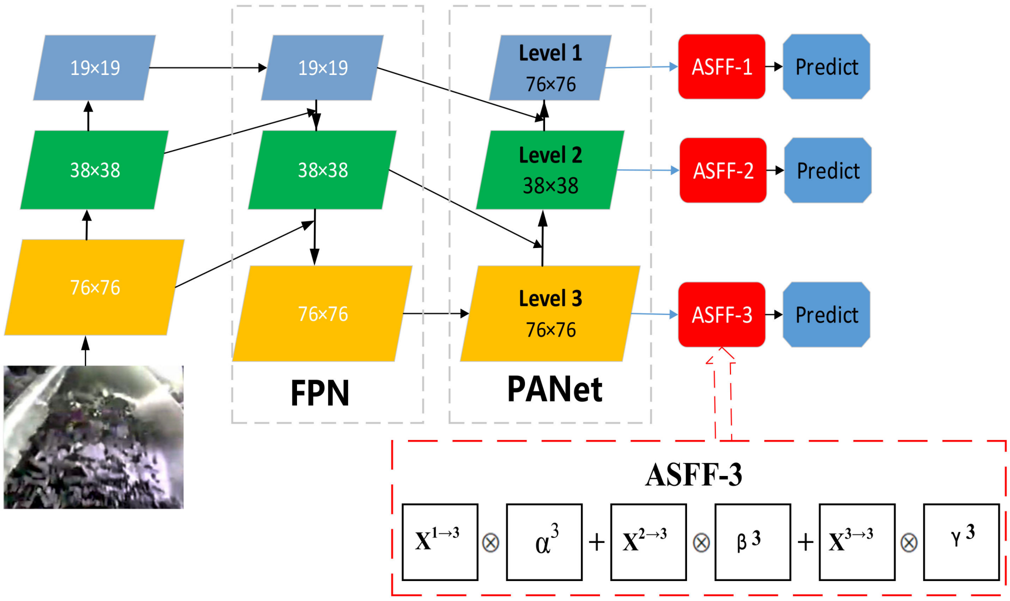


Figure 4. Structure diagram after adding ASFF.

The purpose of CAM is to present the correlation between different channels, automatically acquire the importance of each feature channel in network learning, and automatically assign different weight coefficients to each channel. In this way, it can strengthen important features and suppress non-important ones. Global maximum pooling and global average pooling based on width and height are performed for the input feature map $F(H \times W \times C)$ respectively. The global average pooling can better preserve the local details of foreign objects, which can help us better extract key features of foreign objects. The global average pool is more sensitive to background information, which can better help us classify foreign objects. The characteristic graphs are obtained and sent to the MLP neural network with shared parameters. Two features of the MLP output are added element-wise and then activated by a sigmoid function to generate the final channel attention feature $M_c(F)$. Its expression formula is presented in Equation (4). Finally, $M_c(F)$ and the input feature map $F(H \times W \times C)$ are multiplied element-wise, generating input features required by the spatial attention module.

$$\begin{aligned} M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ &= \sigma\left(W_1\left(W_0\left(F^c_{avg}\right)\right) + W_1(W_0(F^c_{\max}))\right) \end{aligned} \tag{4}$$

where σ denotes the sigmoid function. W_0 and W_1 are the shared parameters of perceptron network. $F^c_{\max}$ is the global average pooling operation for the channel attention mechanism. F^c_{avg} is the maximum average pooling operation of the channel attention mechanism.

The SAM can improve feature expression of key areas. By generating a weight mask for each position and adding weight for output, SAM can enhance the interest of the algorithm model in a specific target area and weaken interference of irrelevant background. The feature map of foreign objects output by CAM is used as the input feature map of SAM. Firstly, maximum global pooling and global average pooling-based channels are performed for $M_c(F)$, and characteristic graphs $F^s_{\max}$ and F^s_{avg} are obtained. Then, the two feature maps are spliced based on the channel, 7×7 convolution is better than 3×3 convolutions after testing, so use a 7×7 convolution to make it become a one-dimensional channel. Secondly, $M_s(F)$ are generated by sigmoid. Its expression formula is presented in Equation (5). Finally, the feature is multiplied with the input feature of the module to obtain the refined feature.

$$M_s(F) = \sigma\Big(f^{7\times 7}([AvgPool(F); MaxPool(F)])\Big) = \sigma\Big(f^{7*7}\Big(\Big[F^s_{avg}; F^s_{\max}\Big]\Big)\Big) \tag{5}$$

where $f^{7\times 7}$ is a convolution kernel of 7×7 general convolution operation. [ ] is concat operation. $F^s_{\max}$ is a global average pooling operation for spatial attention mechanism. F^s_{avg} is the maximum average pooling operation of the spatial attention mechanism.

2.2.3. Method of Improving the Ability of Feature Fusion

The FPN + PANet [24] module is the neck network of YOLOv5. This fusion mechanism strengthens network feature fusion ability and improves the diversity and robustness of features. However, FPN + PANet does not make full use of features of different scales when fusing features of objects in the image. The features of different sizes in the image are simply transformed into the same size and then added, so some targets will be missed.

Due to the complexity of the underground mining environment of the coal mine, foreign objects on the belt conveyor will be obscured by coal blocks, leading to foreign objects missing detection. At the same time, in order to make better use of the key features of foreign matters extracted by CBAM in the backbone, ASFF [25] is used to make better use of features of foreign objects in different scales and more fully integrate the high-level semantic information and the low-level contour information and color features of a foreign object. Therefore, ASFF is integrated into the neck to improve recall of the network model.

The structure diagram after adding ASFF is presented in Figure 4. The improved parts are presented in red modules in Figure.

The key idea of ASFF is to adaptively learn the spatial fusion weight of each scale feature map by the two steps for constant scaling and adaptive fusion.

Constant scaling. For upsampling, we first apply a 1×1 convolution layer to compress the number of channels to features to that in level l and then upscale the resolutions respectively with interpolation. For down-sampling with a $1/2$ ratio, a 3×3 convolution layer with a stride of two is used to modify the number of channels and the resolution simultaneously. For the scale ratio of $1/4$, we add a 2-stride max pooling layer before the 2-stride convolution.

Adaptive fusion. $x^{n\to l}_{ij}$ represents the feature vector at the position (i, j) of feature map resized from the nth level to the l level. The l level features is presented in Equation (6).

$$y^l_{ij} = \alpha^l_{ij} \cdot x^{1\to l}_{ij} + \beta^l_{ij} \cdot x^{2\to l}_{ij} + \gamma^l_{ij} \cdot x^{3\to l}_{ij} \tag{6}$$

where y^l_{ij} implies the (i,j)-th vector of the output feature maps y^l among channels. α^l_{ij}, β^l_{ij} and γ^l_{ij} refer to the spatial importance weights for feature maps at three different levels to level l. We force $\alpha^l_{ij} + \beta^l_{ij} + \gamma^l_{ij} = 1$ and $\alpha^l_{ij}, \beta^l_{ij}, \gamma^l_{ij} \in [0, 1]$.

3. Foreign Objects Identification Process of Belt Conveyor

The foreign objects recognition process of the belt conveyor is presented in Figure 5. Firstly, collected data sets are preprocessed by defogging and deblurring, and label the dataset. Then, the whole data set is randomly divided into a training set and a test set with a ratio of 8:2. The training set is used for model training and learning, and the test set is used to test the detection ability of the model. The foreign objects data set is introduced into the improved YOLOv5 algorithm model for training to get the foreign objects recognition model. Finally, the foreign objects recognition model is used to recognize monitoring video of belt conveyors in real-time.

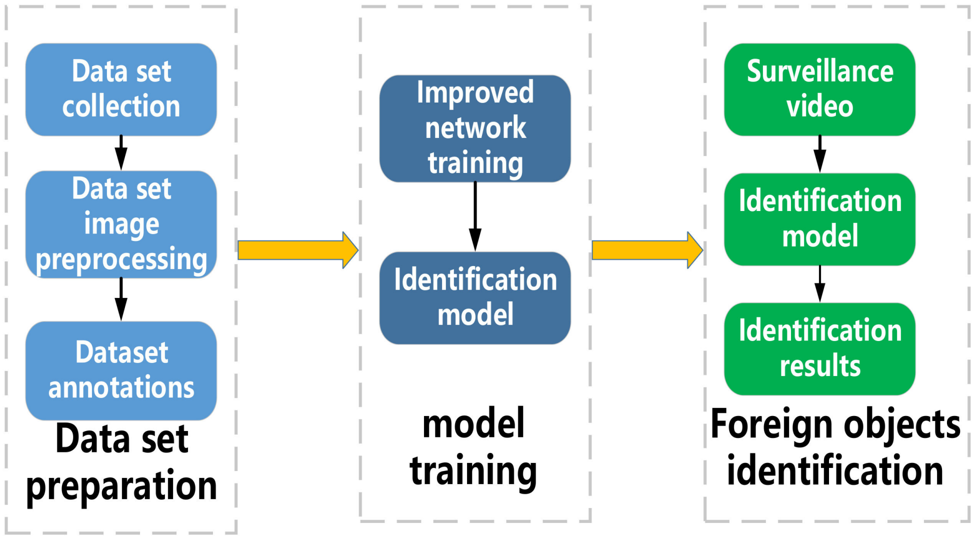


Figure 5. Identification process of foreign objects.

4. Experimental Results and Analysis

4.1. Experimental Equipment

The hardware parameters and software parameters of the equipment used in the experiment are presented in Table 1.

Table 1. Parameters of Software equipment and hardware equipment.

Configuration Name	Parameter
Operating system	Windows10
GPU	NVIDIA GeForce RTX 3080
CPU	12thGen Intel(R) Core(TM)i7-12700K 3.61 GHZ
Deep learning framework	Pytorch
Monitor camera frame rate	25 FPS
Belt conveyor running speed	3.5 m/s

The belt conveyor in the laboratory is presented in Figure 6a, and camera arrangement for data collection in the laboratory is presented in Figure 6b, and a light shielding device is added beside camera collection to simulate uneven illumination underground.

(a) (b)

Figure 6. Layout of belt conveyor and other devices in the laboratory. (**a**) Laboratory belt conveyor. (**b**) Camera arrangement.

4.2. Dataset Production

Part of the dataset comes from the monitoring video of the belt conveyor in Huangling Coal Mine, and the remaining part of the dataset comes from the monitoring image of the laboratory belt conveyor. The dataset not only contains foreign objects images with different degrees of brightness and darkness but also contains images of shielded foreign objects. All foreign objects' data images are rotated to simulate different angles of foreign objects' appearance to improve recognize ability of foreign objects model in different conditions.

There are 5060 foreign body images in the foreign objects data set. 4048 pieces were randomly selected as training samples, and the remaining 1012 pieces were used as test samples. The data set contains five kinds of foreign objects anchor rods, angle iron, wood strip, gangue, and large coal. The partial image display of the data set is presented in Figure 7. In order to facilitate subsequent identification and recognition, FB1, FB2, FB4, FB6, and FB7, respectively, represent anchor rods, angle iron, wood, gangue, and large coal when labeling data and displaying recognition results.

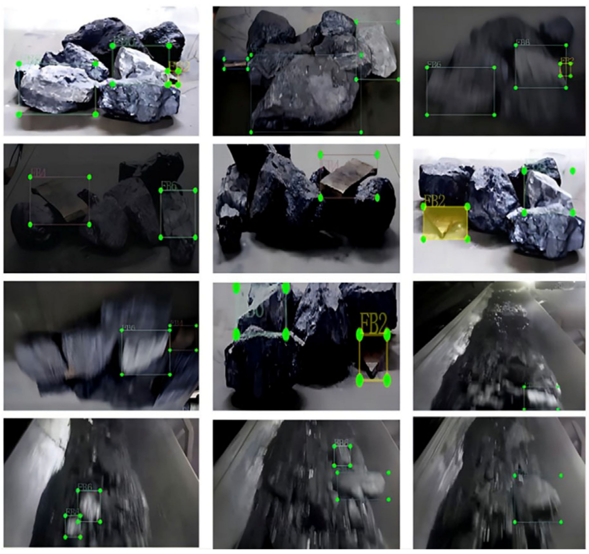
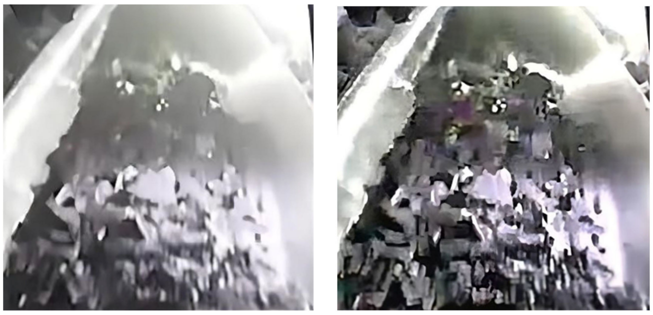


Figure 7. Partial images of the dataset.

4.3. Analysis of Experimental Results

4.3.1. Analysis of Image Preprocessing Results

The defogging algorithm based on the dark channel priori is used to reduce the influence of fog on the clarity of video images, and the image deblurring method based on user-defined convolution is used to reduce the blurring effect caused by the high-speed operation. The image before preprocessing is presented in Figure 8a and the image after preprocessing is presented in Figure 8b. It can be seen that the clarity of the image is better, and the outlines of foreign objects are more obvious.



(a) (b)

Figure 8. Comparison before and after preprocessing. (**a**) Image before preprocessing. (**b**) Image after preprocessing.

In order to measure the effectiveness of the image preprocessing method, we select the Vollath function and information entropy to measure the change of clarity before and after preprocessing. The value of Vollath is proportional to the clarity of the image. Information entropy can reflect the image of the diversity of gray values and indicates the aggregation characteristics of the image grayscale distribution. When the entropy value is larger, the color of the image is brighter, and the outline is clearer [26]. The calculation formula of vollath value and information entropy are presented in Equations (7) and (8).

$$D(f) = \sum_{y}\sum_{x} f(x,y) \times f(x+1,y) - M \times N \times \mu^2 \quad (7)$$

$$H = \sum_{i=0}^{255} P_i \log P_i \quad (8)$$

where $D(f)$ is the result of the calculation of the vollath value. μ is the average gray value of the whole image. M and N are image width and height, respectively. H is the information entropy calculation result. P_i is the probability that a certain gray level appears in the image, which can be obtained from the gray level histogram.

There were 50 processed images that were selected to calculate their Vollath value and information entropy, and their average values are presented in Table 2. The results indicate that the vollath value and information entropy of the image after preprocessing are larger than the original image, which indicates that the image after preprocessing is clearer.

Table 2. Image definition and quality evaluation.

	Vollath Value	**Information Entropy**
Original image	82.19	7.6755
Image after preprocessing	90.35	7.8305

4.3.2. Analysis of Recognition Results for Improved YOLOv5 Algorithm

(1) Comparative analysis of recognition results before and after the improvement

In order to verify the superiority of the improved algorithm, the improved YOLOv5 algorithm is compared with the YOLOv5 algorithm. Precision rate changes are presented in Figure 9. From Figure 9, it can be seen that the improved YOLOv5 algorithm has a faster iteration speed and can achieve a higher precision rate with fewer iterations. The calculation formula of the precision (P) and the recall rate (R) are presented in Equations (9) and (10), respectively.

$$P = N_{TP}/(N_{TP} + N_{FP}) \tag{9}$$

$$R = N_{TP}/(N_{TP} + N_{FN}) \tag{10}$$

where N_{TP} represents the number of true samples, N_{FP} represents the number of false positive samples, and N_{FN} represents the number of true negative samples.

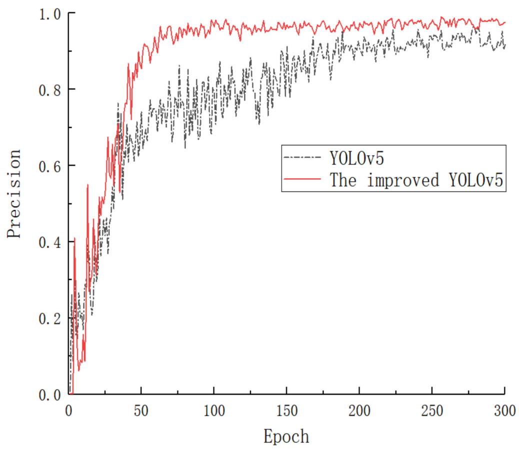


Figure 9. Change curve of recognition precision of YOLOv5 before and after improvement.

Comparison of recognition precision of various foreign objects of YOLOv5 before and after improvement are presented in Table 3, and a comparison of recall rate is presented in Table 4. It can be seen from Tables 3 and 4 that the recognition precision of anchor rod, angle iron, wood, gangue, and large coal are increased by 4.4%, 4.2%, 4.6%, 5.8%, and 5.2%, respectively, and the recall rate are increased by 3.6%, 7.2%, 3.8%, 6.1%, and 8.4% respectively.

Table 3. Comparison of foreign objects recognition precision before and after improvement.

Category	Recognition Precision before Improvement/%	Recognition Precision after Improvement/%
Anchor rod	93.1	97.5
Angle iron	92.7	96.9
wood	92.0	96.6
Gangue	91.0	96.8
Large coal	90.2	95.4

Table 4. Comparison of foreign objects recalls rate before and after improvement.

Category	Recall Rate before Improvement/%	Recall Rate after Improvement/%
Anchor rod	94.7	98.3
Angle iron	90.4	97.6
wood	93.4	97.2
Gangue	90.1	96.2
Large coal	88.2	96.6

(2) Analysis of experimental results of each improved part

In order to verify the influence of each improved part on the overall algorithm, ablation experiments were carried out. The same data set and software and the same hardware equipment are used in the experiment. Recognition precision and processing time of each improved module in the experiment are presented in Table 5.

Table 5. Ablation experimental results.

Preprocessing	CBAM	ASFF	Precision/%	t/s
			91.8	0.0090
√			92.6	0.0134
	√		96.2	0.0120
		√	92.8	0.0110
√	√	√	96.6	0.0157

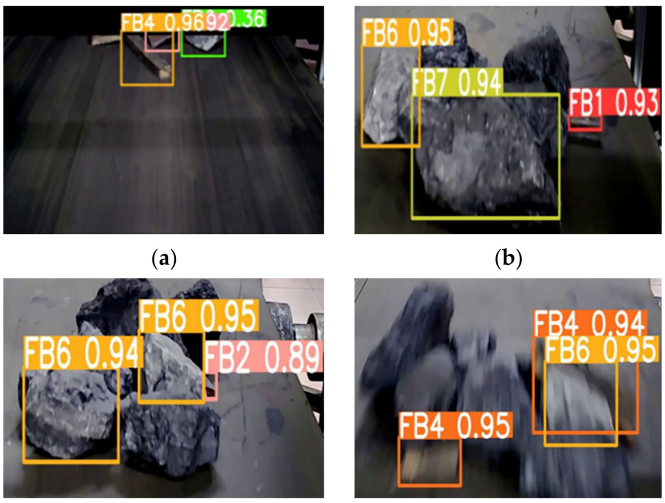
The first line in Table 5 presents the recognition results of the YOLOv5 network. After image preprocessing, the recognition precision of YOLOv5 is improved by 0.8%. After CBAM is incorporated into the neck, the feature expression ability of foreign objects in the complex environment is improved, and recognition precision is increased by 4.4%. When ASFF is added, recognition precision is improved by 1%. After integrating the advantages of each improved module, the recognition precision rate can reach 96.6%, which is 4.8% higher than the recognition precision rate of the original YOLOv5 algorithm. Moreover, it has a fast recognition speed. It takes 0.0157s to detect each frame of images. Therefore, the improved YOLOv5 algorithm can accurately recognize foreign objects of the belt conveyor in real-time.

4.4. Analysis of Laboratory and Coal Mine Field Test Results

4.4.1. Analysis of Laboratory Test Results

In order to verify the effect of the improved YOLOv5 algorithm, 10 rounds of detection experiments are carried out in the laboratory. The foreign objects used in experiments are 10 anchor rods, 10 angle irons, 16 wooden strips, 24 gangues, and 16 large coals. Nine anchor rods, 10 angle irons, 16 wooden strips, 22 gangues, and 15 large coals are correctly identified. The identification accuracy is 94.73%, and the average recognition frame rate is 63.29 FPS.

As presented in Figure 10, the test results of the belt conveyor in the laboratory can also be identified by shielded wood, angle iron, and anchor rods. In Figure 10a, wooden strips, angle iron, and gangue are identified. In Figure 10b, gangue, large coal, and shielded anchor rods are identified. In Figure 10c, two gangues and shielded angle iron are identified. In Figure 10d, a gangue and two blocked wooden strips are identified.



(a) (b) (c) (d)

Figure 10. Test results of laboratory. (**a**) Recognition results of FB2, FB4 and FB6 in the laboratory. (**b**) Recognition results of FB1, FB6 and FB7 in the laboratory. (**c**) Recognition results of FB2 and FB6 in the laboratory. (**d**) Recognition results of FB4 and FB6 in the laboratory.

4.4.2. Analysis of Test Results in Coal Mine

In order to verify the effect of this improved YOLOv5 algorithm in the actual coal mine field, multi-belt conveyor monitoring videos with a resolution of 1920 × 1080 in the Huangling coal mine are detected. In the video of the belt conveyor of Huangling coal mine, there are 143 gangues, 17 large coals, one piece of wood, and two anchor rods. Moreover, 136 gangues, 16 large coals, one wooden strip, and two anchor rods are successfully identified. The recognition accuracy is 95.09%, and the average recognition speed is 56.50 FPS.

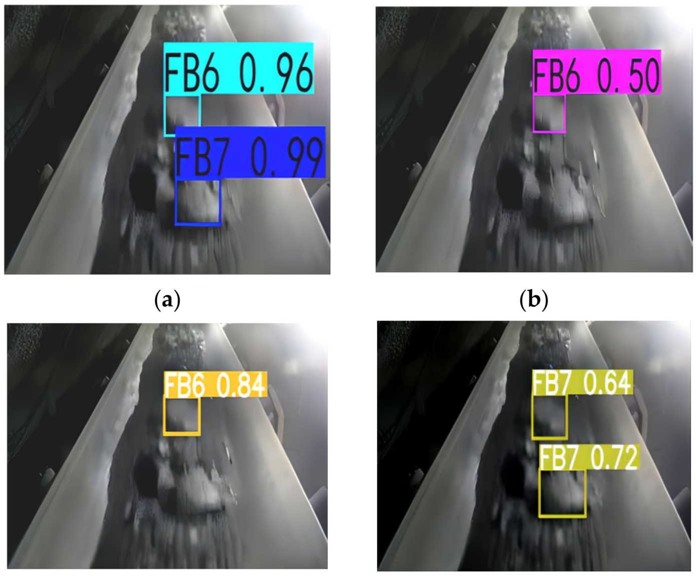
In order to verify the effectiveness of the improved YOLOv5 algorithm, SSD, YOLOv3, and YOLOv5 algorithms are selected for comparison, and all algorithms use the same dataset and training environment. The recognition results are as follows.

Figure 11a is the recognition result of YOLOv5. Due to the influence of fog and high-speed motion blur, the monitoring image is low in clarity, the characteristics of the target object are fuzzy and difficult to distinguish, and the anchor rod is missed. Figure 11b is the recognition results of the improved YOLOv5 algorithm. It can be seen from a comparison of recognition results the clarity of the image is improved, and the anchor rod is finally correctly recognized.

(**a**) (**b**)

Figure 11. Comparison of recognition results of YOLOv5 before and after improvement. (**a**) Recognition results of the original YOLOv5 algorithm (**b**) Recognition results of the improved YOLOv5 algorithm.

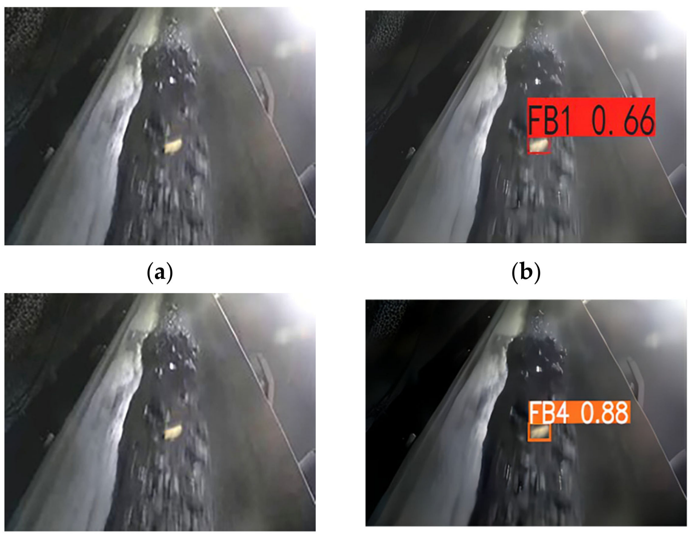
Figure 12a is the identification result of SSD, which identifies a large coal, and another large coal is mistakenly identified as gangue due to the influence of dust mist. Figure 12b,c are the recognition results of YOLOv3 and YOLOv5, respectively. Under the influence of dust fog and high-speed motion, one large coal is mistakenly recognized as gangue, and the other large coal is missed due to mixed background and nonprominent features. Figure 12d is the recognition results of the improved YOLOv5 algorithm. It can be seen that the results of the improved YOLOv5 algorithm are right. Because the improved YOLOv5 algorithm with preprocessing and CBAM makes channel attention automatically assigns different weight coefficients to each channel, the improved YOLOv5 algorithm has been enhanced to extract important features and suppress nonimportant features. In addition, a weight mask is automatically generated for each position and weighted output through CBAM, which weakens irrelevant background areas and enhances specific target areas of interest. It makes the recognition network model biased toward the foreign object target and further improves the feature expression ability of foreign objects. The two large coals are successfully identified in the end.



(**a**) (**b**) (**c**) (**d**)

Figure 12. Recognition results of four algorithms. (**a**) Recognition results of SSD. (**b**) Recognition results of YOLOv3. (**c**) Recognition results of YOLOv5. (**d**) Recognition results of the improved YOLOv5 algorithm.

Figure 13a,c are the identification results of SSD and YOLOv5, respectively. The wooden strip is partially shielded by the coal block, and only part of it is exposed outside, so it is missed. Figure 13b presents the recognition result of YOLOv3, and the blocked wooden strip is mistakenly recognized as an anchor rod. Figure 13d is the recognition results of the improved YOLOv5 algorithm. Since ASFF is added, the semantic information of the upper layer and the contour and shape information of the lower layer of the image can be more fully fused. In addition, image preprocessing and CBAM make the foreign object's features of the target more prominent and easier to be recognized, which better addresses the problem of missed detection or false detection caused by the obscuration of the target. Finally, the wooden strip is identified.



(**a**) (**b**) (**c**) (**d**)

Figure 13. Recognition results of various algorithms. (**a**) Recognition results of SSD. (**b**) Recognition results of YOLOv3. (**c**) Recognition results of YOLOv5. (**d**) Recognition results of the improved YOLOv5 algorithm.

It can see the focus of the model in detecting and identifying foreign objects by adding a class activation heatmap. Figure 14a,b presents the activation heatmap of the anchor before and after the YOLOv5 algorithm improvement, and Figure 14c,d presents the activation heatmap of the large coal before and after the YOLOv5 algorithm improvement. From Figure 14, it can be seen that the improved YOLOv5 algorithm has a stronger ability to capture foreign objects, and the confidence in identification has also been improved.

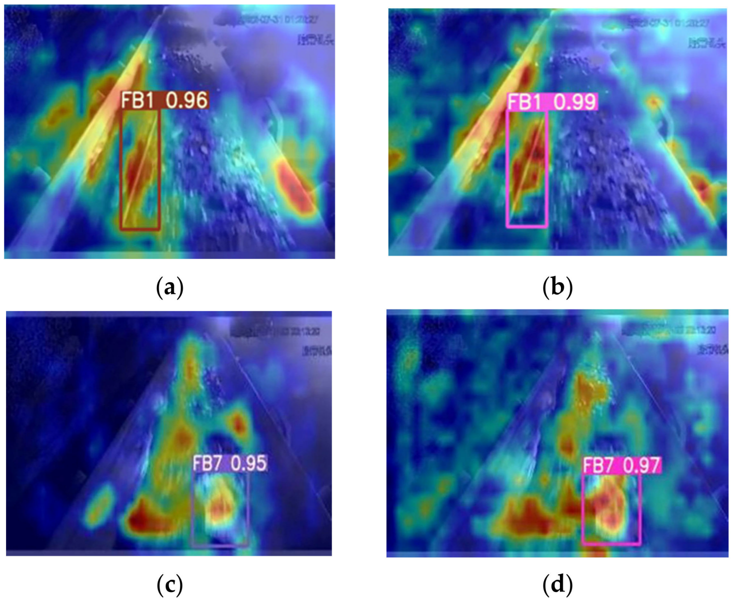


Figure 14. Activation heatmap of foreign objects of YOLOv5 before and after improvement. (**a**) Heatmap of the anchor with YOLOv5. (**b**) Heatmap of the anchor with the improved YOLOv5 algorithm. (**c**) Heatmap of the gangue with YOLOv5. (**d**) Heatmap of the gangue with the improved YOLOv5 algorithm.

4.4.3. Comparison of Recognition Results of the Improved YOLOv5 Algorithm

In order to verify the effectiveness of the improved YOLOv5 algorithm, several common YOLOv5 improvements are selected for comparison, and all algorithms use the same dataset and training environment. The results are presented in Table 6. In addition, we also selected the public data sets VOC2007 and COCO128 to verify the performance of the improved algorithm. The results are presented in Tables 7 and 8.

Table 6. Comparison of recognition results of four Algorithms.

Algorithm	Average Precision/%	t/s
YOLOv5	91.8	0.0100
YOLOv5-SE-BIFPN	93.4	0.0110
YOLOV5-CA-ASFF	96.1	0.0150
Our improved YOLOv5 algorithm	96.6	0.0157

Table 7. Recognition results of improved YOLOv5 Algorithms with VOC2007.

Algorithm	Average Precision/%	Average Recall/%
YOLOv5	74.2	57.3
The improved YOLOv5 algorithm	76.4	61.0

Table 8. Recognition results of improved YOLOv5 Algorithms with COCO128.

Algorithm	Average Precision/%	Average Recall/%
YOLOv5	93.6	92.5
The improved YOLOv5 algorithm	95.6	93.6

It can be seen from Table 6 that our improved YOLOv5 algorithm has the highest recognition accuracy while maintaining real-time performance. It can be seen from Tables 7 and 8 that the average recognition accuracy and recall of the improved YOLOv5 algorithm has improved in public datasets, so the improved YOLOv5 algorithm is effective.

5. Conclusions

Aiming at the problem that it was difficult to accurately identify the foreign objects of belt conveyors in a coal mine due to the influence of fog, high-speed operation, and obscuration, the coal mine belt conveyor foreign objects recognition method of Improved YOLOv5 algorithm with defogging and deblurring is proposed. In order to verify the practical application effect of the improved YOLOv5 algorithm, the monitoring videos of the belt conveyor in the laboratory and Huangling coal mine are used for testing. The conclusions are as follows.

(1) The dark channel priori defogging algorithm can reduce the impact of fog on the clarity of monitoring video, and the user-defined convolution method can reduce the blurring effect on the image in high-speed operating conditions. Through verification, the image quality after preprocessing has been significantly improved. It even improves the precision of foreign object identification.

(2) The CBAM-C3 improves the feature expression ability of the foreign object in the complex background. Through adaptive spatial feature fusion, the multi-layer feature information of the foreign object image is more fully fused. Compared with the YOLOv5 algorithm, the average recognition precision of the improved YOLOv5 algorithm has increased by 4.8%.

(3) The foreign objects recognition test is carried out by monitoring video of a belt conveyor in the Huangling coal mine. The test results presented that the recognition accuracy is 95.09%, and the average recognition rate is 56.50 FPS. So, the improved YOLOv5 algorithm addresses problem that foreign objects of the belt conveyor are difficult to be accurately identified due to dust, high-speed operation, and obscuration.

Author Contributions: For this research article, Q.M. was in charge of the methodology; S.L. was in charge of the writing, review, and data analysis; X.H. and X.X. was in charge of the data curation and field inspection. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the General project of Shaanxi coal joint fund of Shaanxi Provincial Department of Science and Technology (2019JLM-39) and the Shaanxi Science and Technology Innovation Team Project (2018TD-032).

Data Availability Statement: Not applicable.

Acknowledgments: The study was approved by the Xi'an University of Science and Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, J.; Huang, Z.; Wang, M.; Fu, L.; Liu, A.; Li, Y.; Yv, Z.; Wang, Z.; Huang, L. Unwanted object recognition based on the pyramid convolution model of the machine vision grayscale for the coal flow monitoring. *Coal Sci. Technol.* **2021**, 1–9. [CrossRef]
2. Hao, S.; Zhang, X.; Ma, X.; Sun, S.; Wen, H.; Wang, J.; Bai, Q. Foreign objects detection in coal mine conveyor belt based on CBAM-YOLOv5. *J. China Coal Soc.* **2022**, 1–11. [CrossRef]
3. Gao, R.; Miao, C.; Miao, D.; Li, X. Multi-view image adaptive enhancement method for conveyor belt fault detection. *J. China Coal Soc.* **2017**, *42*, 594–602. [CrossRef]

4. Wang, G.; Zhao, G.; Ren, H. Analysis on key technologies of intelligent coal mine and intelligent mining. *J. China Coal Soc.* **2019**, *44*, 34–41. [CrossRef]
5. Wang, G.; Ren, H.; Zhao, G.; Du, Y.; Pang, Y.; Xv, Y.; Zhang, D. Analysis and countermeasures of ten'pain points' of intelligent coal mine. *Ind. Mine Autom.* **2021**, *47*, 1–11. [CrossRef]
6. Wu, S. Rescarch on Detcction Mcthod of Forcign Object on Coal Conveyor Belt Based on Computer Vision. Ph.D. Thesis, China University of Mining and Technology, Xuzhou, China, 2019.
7. Wu, S.; Ding, E.; Yu, X. Foreign Body ldentification of Belt Based on Improved FPN. *Saf. Coal Mines* **2019**, *50*, 127–130. [CrossRef]
8. Lv, Z. Research on Imagc Recognition of Foreign Bodies in the Process of Coal Minc Belt Transportation in Complex Environmcnt. Ph.D. Thesis, China University of Mining and Technology, Xuzhou, China, 2020.
9. Ren, Z.; Zhu, Y. Research on foreign objects Recognition of Coal Mine Belt Transportation with Improved CenterNet Algorithm. *Control. Eng. China* **2021**, 1–8. [CrossRef]
10. Ma, G.; Wang, X.; Liu, J.; Chen, W.; Niu, Q.; Liu, Y.; Gao, X. Intelligent Detection of Foreign Matter in Coal Mine Transportation Belt Based on Convolution Neural Network. *Sci. Program.* **2022**, *2022*, 9740622. [CrossRef]
11. Hu, J.; Gao, Y.; Zhang, H.; Jin, B. Research on the identification method of non-coal foreign object of belt conveyor based on deep learning. *Ind. Mine Autom.* **2021**, *47*, 57–62+90. [CrossRef]
12. Wang, Y.; Wang, Y.; Dang, L. Video detection of foreign objects on the surface of belt conveyor underground coal mine based on improved SSD. *J. Ambient. Intell. Humaniz. Comput.* **2020**, 1–10. [CrossRef]
13. Du, J.; Chen, R.; Hao, L.; Shi, Z. Coal mine belt conveyor foreign objects detection. *Ind. Mine Autom.* **2021**, *47*, 77–83. [CrossRef]
14. Cheng, D.; Xu, J.; Kou, Q.; Zhang, H.; Han, C.; Yv, B.; Qian, J. Lightweight network based on residual information for foreign body classification on coal conveyor belt. *J. China Coal Soc.* **2022**, *47*, 1361–1369. [CrossRef]
15. Xiao, D.; Kang, Z.; Yu, H.; Wan, L. Research on belt foreign body detection method based on deep learning. *Trans. Inst. Meas. Control.* **2022**, *44*, 2919–2927. [CrossRef]
16. Wang, R. The Research of Real-Time Image Dehazing Based on Dark Channel Prior. Ph.D. Thesis, Xi'an University of Electronic Science and Technology, Xi'an, China, 2018; pp. 1–54.
17. Wang, Y.; Wei, S.; Duan, Y.; Wu, H. Defogging algorithm of underground coal mine image based on adaptive dual-channel prior. *J. Mine Autom.* **2022**, *48*, 46–51+84. [CrossRef]
18. Wei, L.; Wang, J. Simulation research on edge sharpening enhancement of motion blurred digital image. *Comput. Simul.* **2020**, *37*, 459–462.
19. Hu, D.; Tan, J.; Zhang, L.; Ge, X. Image deblurring based on enhanced salient edge selection. *Vis. Comput.* **2021**, 1–16. [CrossRef]
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *28*, 1137–1149. [CrossRef] [PubMed]
21. Cai, Z.; Vasconcelos, N. Cascade R-CNN: Delving into high quality object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6154–6162. [CrossRef]
22. Yan, P.; Sun, Q.; Yin, N.; Hua, L.; Shang, S.; Zhang, C. Detection of coal and gangue based on improved YOLOv5. 1 which embedded scSE module. *Measurement* **2022**, *188*, 110530. [CrossRef]
23. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional block attention module. In *European Conference on Computer Vision (ECCV), Proceedings of the 15th European Conference, Munich, Germany, 8–14 September 2018, Proceedings, Part VII*; Springer: Cham, Switzerland, 2018; pp. 3–19.
24. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
25. Liu, S.; Huang, D.; Wang, Y. Learning spatial fusion for single-shot object detection. *arXiv* **2019**, arXiv:1911.09516.
26. Mao, Q.; Wang, Y.; Zhang, X.; Zhao, X.; Zhang, G.; Mushayi, K. Clarity method of fog and dust image in fully mechanized mining face. *Mach. Vis. Appl.* **2022**, *33*, 30. [CrossRef]