Article

Not peer-reviewed version

# A Foreign Object Detection Method for Belt Conveyor Based on Improved YOLOX Model

Rongbin Yao , Peng Qi , Dezheng Hua , Xu Zhang , He Lu , Xinhua Liu *

Posted Date: 21 July 2023

doi: 10.20944/preprints202307.1534.v1

Keywords: Belt conveyor; foreign object detection; YOLOX; image enhancement; rotation detection

 doi:10.20944/preprints202307.1534.v1



*Article*

# A Foreign Object Detection Method for Belt Conveyor Based on Improved YOLOX Model

**Rongbin Yao [1], Peng Qi [2], Dezheng Hua [2], Xu Zhang [2], He Lu [1] and Xinhua Liu [2,*]**

[1] Lianyungang Normal College, Lianyungang 222006, China
[2] School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, China
* Correspondence: liuxinhua@cumt.edu.cn

**Abstract:** As one of the main equipment of coal transportation, the belt conveyor with detection system is an important direction for the development of intelligent mine. The occurrences of non-coal foreign objects making contact with belts are common phenomenon in complex production environments and improper human operations. In order to avoid major safety accidents caused by scratches, deviation and breakage of the belt, a foreign object detection method is proposed for belt conveyor in this work. Firstly, a foreign object image dataset is collected and established, and IAT image enhancement module and attention mechanism of CBAM are proposed to enhance image data sample. Moreover, to predict the angle information of foreign objects with large aspect ratios, a rotating decoupling head is designed, and a MO-YOLOX network structure is constructed. Some experiments are carried out with the belt conveyor in the mine intelligent mining equipment laboratory, and the detections of different foreign objects are analyzed. Experimental results show that the accuracy, recall, and $mAP^{50}$ of the proposed rotating frame foreign object detection method reaches 93.87%, 93.69%, and 93.72%, and the average inference time of foreign object detection is 25 ms.

**Keywords:** belt conveyor; foreign object detection; YOLOX; image enhancement; rotation detection

## 1. Introduction

With the development of intelligent mine and the improvement of machine vision technology, the real-time detection system of belt conveyor has become an important research content in recent years. [1]. The transportation belt plays a pivotal role, and is prone to severe accidents such as deviation, slipping, belt breakage, and longitudinal belt tearing during production [2]. In-depth analysis reveals that non-coal foreign objects entering the belt conveyor system are accountable for 61% of belt tearing and breakage incidents, amounting to a total of 21 cases. Accurate and rapid identification of foreign objects in the belt transportation system, followed by removal, can substantially mitigate damage to the belt system and ensure the safe and stable operation of the belt transportation system.

Intelligent detection for safety production of underground coal mines has become a hot research topic [3, 4]. By utilizing video surveillance image, combined with image processing and machine vision related technologies, the theory of mine image monitoring has been applied to multiple aspects of automatic safety detection in coal mines, such as automatic identification of spontaneous combustion fires [5], coal production monitoring [6], face detection and recognition methods for underground miners [7], automatic recognition of coal rock interfaces in coal face [8], etc. However, the traditional belt conveyor foreign object system still relies on cameras to transmit the collected video data to the central control room, where the staff can monitor the coal transportation area and the surrounding environment in real time. This practice is associated with significant drawbacks, including duplication of work and fatigue-induced misjudgments. At the same time, staff cannot address foreign objects in a timely manner, which can easily cause foreign objects to block the transportation belt or sharp parts to scratch the belt, resulting in belt tearing and causing major safety accidents.

In this work, a foreign object detection method for belt conveyors is proposed, and the remainder of this paper is organized as follows. In section 2, the status and deficiencies of current foreign objects

 doi:10.20944/preprints202307.1534.v1



detection methods are introduced. In section 3, the improved algorithm and model architecture are proposed. In section 4, comparative experiments of proposed model are carried out, and the effectiveness of the improved algorithm is verified in the self-made foreign object dataset. Conclusions and future works are summarized in section 5.

## 2. Literature review

### *2.1. Foreign object detection methods based on image processing*

The foreign object detection methods for belt conveyors based on image are to deal with images of coal and non-coal foreign objects to obtain shallow or deep abstract features of the object, and uses image processing to detect foreign objects. It has the advantages of simple installation and maintenance, low application cost, and has become one of the research focuses of foreign object detection for belt conveyors. Due to the diversity of types of foreign objects (such as anchor rods and wood) causing belt tearing, many scholars have begun to extract the color, texture, shape, spatial relationship and other features of foreign objects from the image features [9], and achieve automatic detection of foreign objects through image processing. Jiang et al. [10] used extreme median filtering to perform image noise processing and improve the traditional Canny edge detection algorithm to obtain an improved edge detection method for Canny operator. The algorithm is used to perform image edge detection, and the image gray histogram is used to enhance the foreign object image processing. Donskoi et al. [11] can classify iron ore by image processing and extracting texture features, and can predict the recovery rate and grade of iron ore products. Chatterjee et al. [12] processed the limestone image, used a variety of methods to segment the image, extracted a total of 189 eigenvalues such as shape, color and texture, and analyzed the eigenvalues by principal component analysis, extracted five principal component features with a contribution rate of 95%, trained with neural network, and realized on-line monitoring of limestone ore grade. Zhang et al. [13] proposed a new image segmentation algorithm for belt conveying. A multi-scale Linear filter composed of Hessian matrix and Gaussian function forms the core of the algorithm, which can effectively obtain the edge intensity image, form a good seed area for watershed segmentation, and effectively segment the background between the coal pile and foreign objects. Saran et al. [14] developed an image processing based foreign object detection solution to detect foreign objects such as concrete boulders and iron bars that often occur in the conveyor belt of G furnace raw coal. The solution uses a multi-mode imaging (polarization camera) based system to distinguish foreign objects. Tu et al. [15] proposed a new moving target detection method to solve the difficulties caused by intermittent motion, temperature and dynamic background sequence of moving targets. By further comparing the similarity of edge images, ghosts and real static objects can be classified. Lins et al. [16] developed a system based on the concept of machine vision, which aims to realize the automation of crack measurement process. Using this method, a series of images can be processed and the crack size can be estimated as long as a camera is installed on a truck or robot.

### *2.2. Foreign object detection method based on deep learning*

With the rapid development of deep learning, using the data learning method of deep learning to learn image data features and perceive the surrounding environment, to obtain a foreign object detection model that is more adaptable to the complex and changeable environment has good research value in foreign object detection [17, 18]. Deep learning is achieved by establishing and simulating the information processing neural structure of the human brain to extract low-level to high-level features from external input data, enabling machines to understand the learning data and obtain useful information [19]. Pu et al. [20] used CNN to identify coal and gangue images and help to separate coal and gangue, and introduced migration learning to solve the problem of massive trainable parameters and limited computing power faced by the model. In order to apply CNN to the field of target detection, Ren et al. [21] put forward RCNN method, which uses Selective Search to obtain pre-selected regions, and completes image recognition through CNN combined with SVM. Because the multi-stage implementation of the algorithm leads to its huge time cost, Girshick et al.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



[22] further put forward the concept of ROI (region of interest) pooling layer, and replaced SVM with fully connected neural network, and proposed Fast-RCNN algorithm. In order to solve the problem that foreign objects on the belt conveyor in coal mine damage the belt conveyor, Wang et al. [23] proposed a video detection method of foreign objects on the surface of the belt conveyor based on SSD. Firstly, the deep separable convolution method was adopted to reduce the number of parameters of SSD algorithm and improve the calculation speed. Then, the GIOU loss function is used to replace the position loss function in the original SSD, which improves the detection accuracy. Finally, the extraction position of feature map and the proportion of default frame are optimized, which improves the detection accuracy. Chen et al. [24] used the KinD++ low light image enhancement algorithm to improve the quality of captured low-quality images through feature processing. Considering the fast-running speed of the belt and the influence of background and light source on foreign object targets, Ma et al. [25] proposed an improved Center-Net algorithm, which improved detection efficiency. The normalization method was optimized to reduce computer memory consumption, and a weighted feature fusion method was added to fully utilize the features of each layer, improving detection accuracy. In the experimental environment, the average detection rate was about 20FPS, Meet the demand for real-time detection of foreign objects. Xiao et al. [26] used median filtering method to preprocess images with foreign objects, remove the influence of dust, improve the clarity of ore edges, establish a dataset to train the YOLOv3 belt foreign object detection algorithm, detect belt foreign objects. Finally, after sparse training based on BN layer, the YOLOv3 model was lightweight, and parameters were fine tuned. Compared with the original YOLOv3 model, the model achieved smaller calculations, faster processing, and smaller size.

### *2.3. Discussion*

Due to the special coal mine environment, with a lot of dust, noise, complex background of foreign objects, it is difficult to achieve accurate detection of belt foreign objects. At the same time, the robustness of the traditional foreign object detection algorithm is poor, and the extraction of foreign object feature requires a wealth of experience. So, it cannot flexibly adapt to different scenes in different mining areas. According to the existing dataset, the foreign object detection model of belt conveyor is updated in real time, and the network model learns the actual scene information to adapt to the foreign object detection task in different environment.

## 3. The proposed foreign object detection method

### *3.1. Target detection of YOLO model*

YOLO series target detection algorithm is a supervised learning target detection algorithm [27]. Its basic principle is to divide the input image into several grids, then extract the features of each part of the image through the convolution neural network, and finally output the predicted boundary box, which are the center coordinates of the predicted object, the length and width of the detected object, and the confidence of the object category.

As shown in the Figure 1, the input foreign object image is divided into S×S squares, and features are extracted from each grid through convolutional neural network, and then features are fused and analyzed to output the confidence degree of foreign object target, boundary box coordinate information and foreign object category. In order to improve the accuracy of foreign object detection, a fixed number of anchor boxes are used for each grid to assist in learning position information. Clustering analysis is performed on the known labels of the target detection object in the image to obtain the initial size of the anchor box.

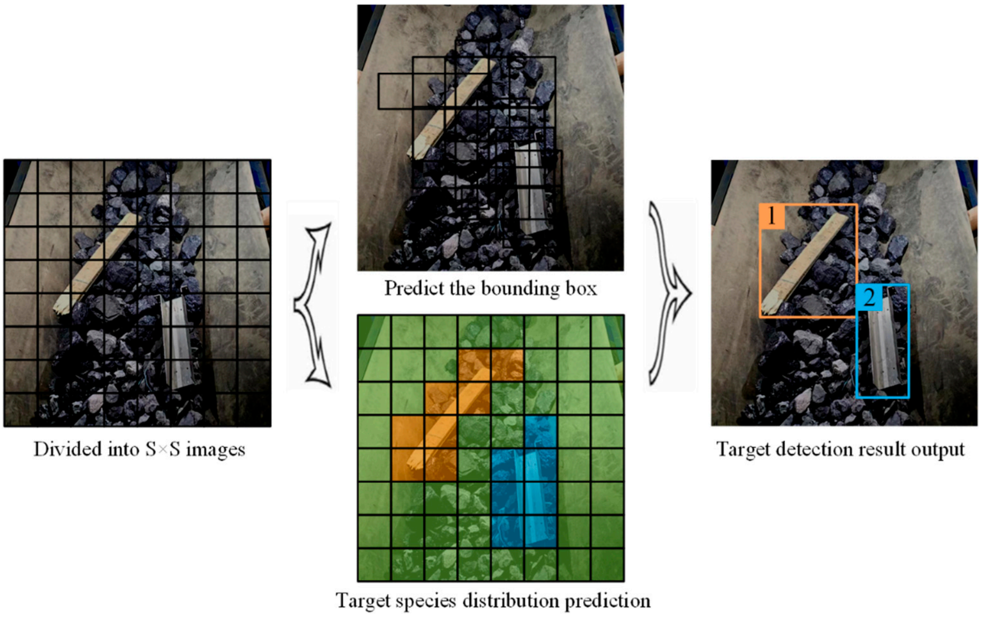


**Figure 1.** YOLO target detection process.

The framework of YOLO series detection models has always been composed of three parts: backbone feature extraction network, feature fusion layer, and detection decoupling head, as shown in the Figure 2. The feature extraction network mainly extracts features from the input image data, then the feature fusion layer fuses the low-dimensional and high-dimensional features of the image to provide richer image information. Finally, the detection decoupling head outputs and predicts the position and category information of objects of interest. The YOLO series of algorithms all use a three branch detection head algorithm to predict objects of different scales, such as large, medium, and small.

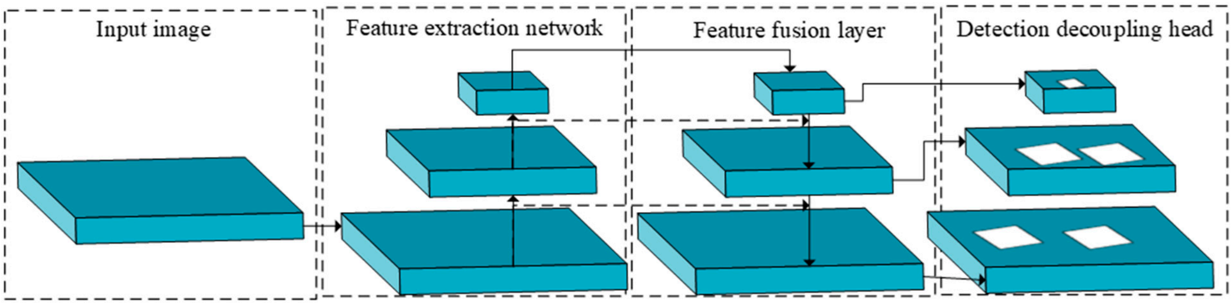


**Figure 2.** YOLO series network architecture.

In the actual target detection process, directly predicting the central coordinates, width and height of the bounding box will result in too large solution space for the predicted target, which will seriously waste computing resources. Therefore, an anchor frame mechanism is designed to accelerate the convergence of the model and improve the target detection accuracy, and the prediction principle of the boundary box is shown in Figure 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



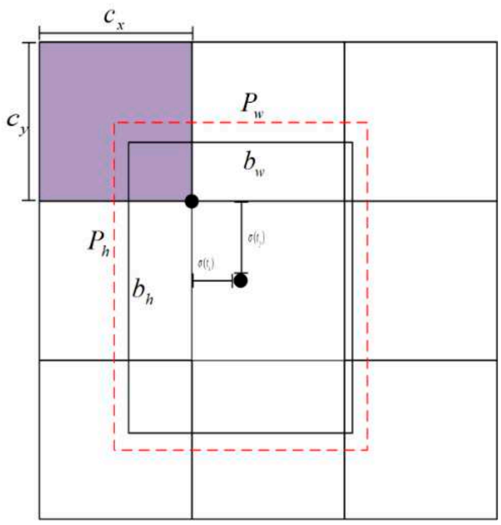


**Figure 3.** The prediction principle of YOLO's bounding box.

$$b_x = \sigma(t_x) + c_x \tag{1}$$

$$b_y = \sigma(t_y) + c_y \tag{2}$$

$$b_w = p_w e^{t_w} \tag{3}$$

$$b_h = p_h e^{t_h} \tag{4}$$

where $p_w$ and $p_h$ are the width and height of the anchor frame, $b_w$ and $b_h$ are the width and height of the prediction box, $t_x$ and $t_y$ are the offset from the anchor frame to the center of the prediction box, $c_x$ and $c_y$ are the coordinates of the upper left corner of the bounding box, σ() is the normalized function.

*3.2. Established foreign object image dataset*

At present, the publicly available large-scale datasets do not include non-coal foreign objects. Therefore, it is necessary to establish an actual foreign object engineering dataset for belt conveyors to solve the problem of foreign object detection in practical engineering. The self-made dataset is named the Belt Conveyor Foreign Object Detection Dataset, and the sample categories of the dataset mainly include the following three types of foreign objects: iron, wood, and large gangue. As showed in Figure 4.


**Figure 4.** Different kinds of foreign object samples.

This work selects laboratory and belt conveyor work scenarios for foreign object image collection. At the same time, foreign object image datasets are captured under different natural light conditions, and directional foreign objects such as iron and wood are offset to increase the information of image angles. In the laboratory environment, for the same foreign object, a foreign object dataset can be established that includes images of areas without coal flow, areas with coal flow, and areas obstructed by coal flow. Photos of foreign objects in different directions are collected in the laboratory environment to increase the diversity of foreign object dataset samples, as shown in Figures 5.

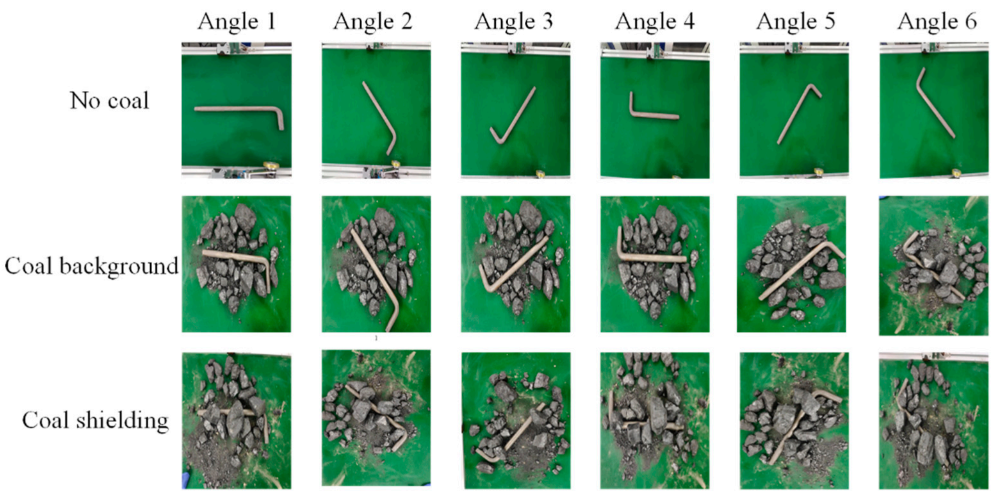


**Figure 5.** Single foreign material data sample.

In order to ensure the diversity of perspective in the collected dataset and better simulate the different shooting angles of cameras installed in actual working conditions, the top view images are collected by using a DJI drone with a pan tilt camera. The heights from the ground during the collection are 1 m, 2 m, and 4 m, respectively, to ensure the diversity of perspective in the collected data, as shown in Figure 6.

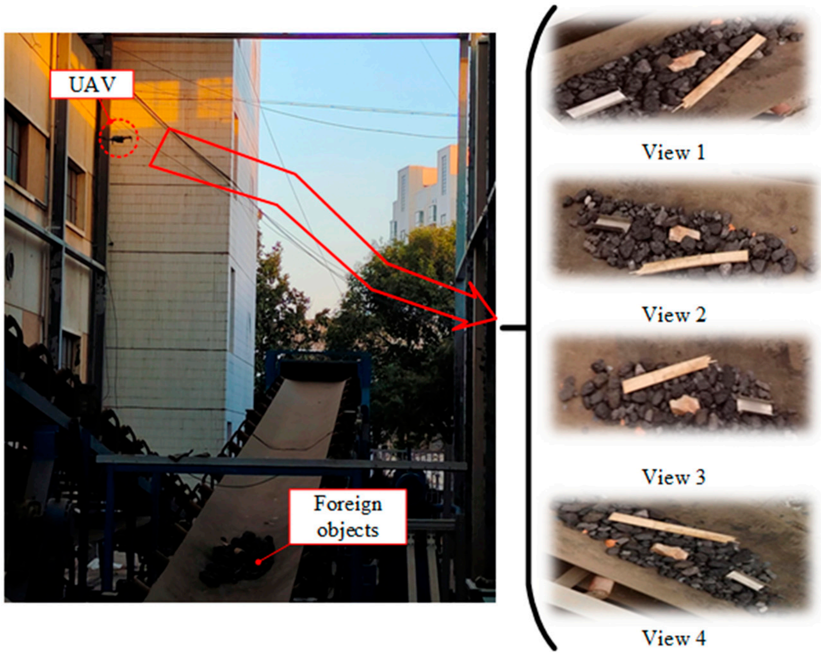


**Figure 6.** Multi-view image acquisition by UAV.

In order to improve the robustness and generalization of the model, the Mosaic multi-samples data augmentation method proposed by YOLOv4 is adopted. During the training process, four images in the training set are randomly selected, and the images are randomly scaled, cropped, and arranged for image combination. The sample size of the images during the training process is expanded, as shown in Figures 7.

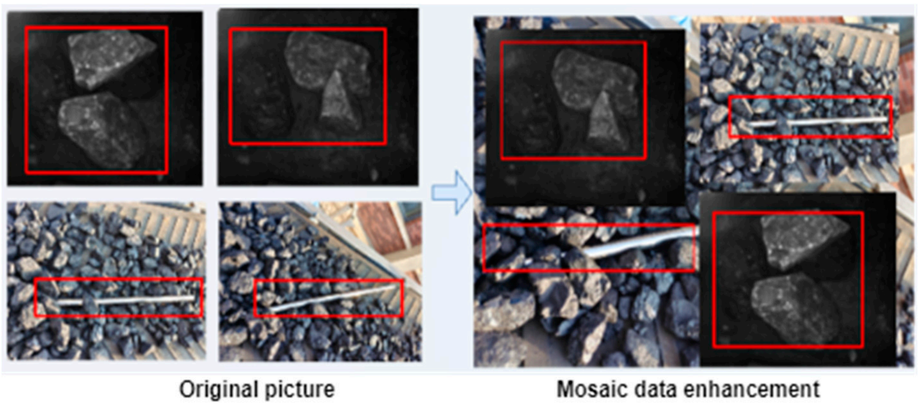


**Figure 7.** Mosaic multi-sample data enhancement.

As shown in Figure 8, the images are expanded by means of horizontal flipping, random occlusion, random scaling, motion blur, random scaling and filling, and salt and pepper noise. A total of 1105 foreign object image datasets were collected for the belt conveyor foreign object image dataset, including 303 large gangue datasets, 401 iron tools datasets, 301 wood datasets, and 100 mixed target images. The belt conveyor foreign object image dataset was labeled with horizontal and rotating boxes, and the horizontal and rotating box foreign object detection datasets were constructed. Finally, the dataset was expanded to 8100 datasets through geometric expansion, A complete dataset of foreign object images for belt conveyors has been constructed.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



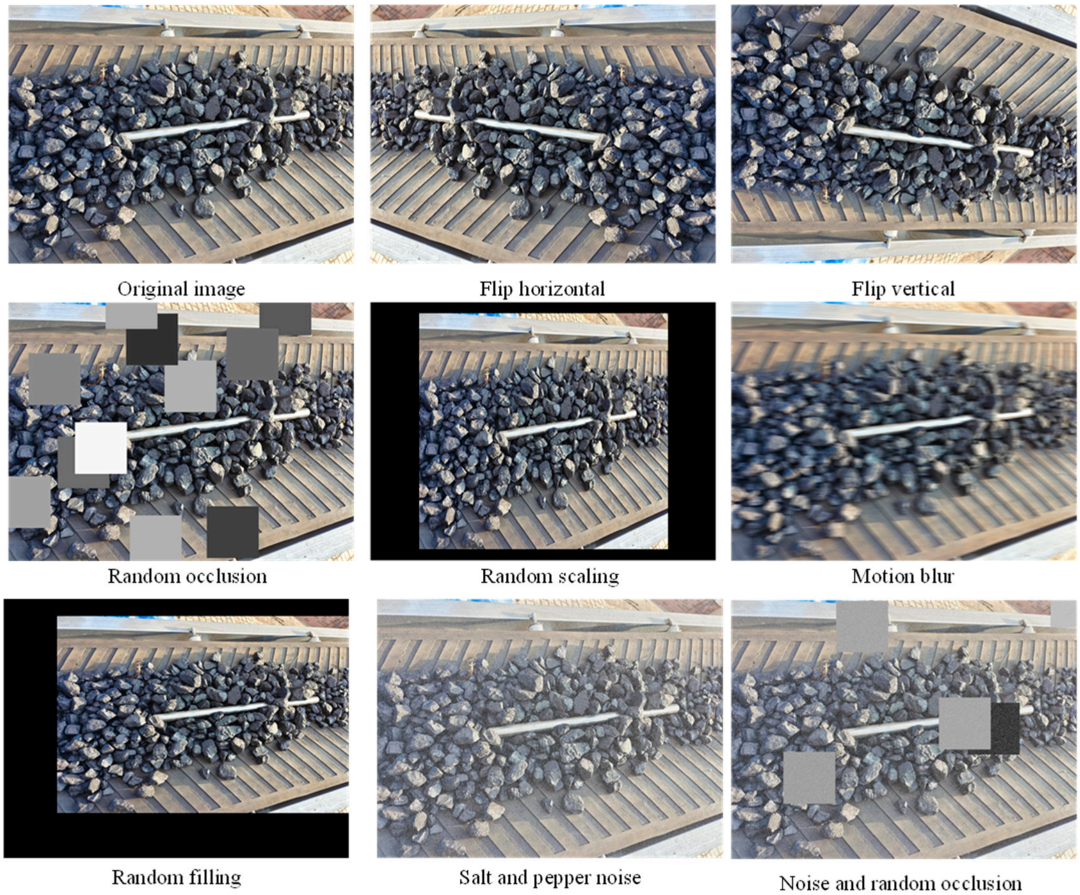


**Figure 8.** Geometry data enhancement.

*3.3. Improved depthwise separable convolution block*

Depthwise separable convolution decomposes the operations of standard convolution into deep convolution and point by point convolution [28]. Deep convolution performs separate spatial convolutions on each input channel, while point by point convolution combines the convolution results of each channel, which can greatly reduce the size and complexity of the model while maintaining high accuracy. The specific operations are shown in Figure 9.

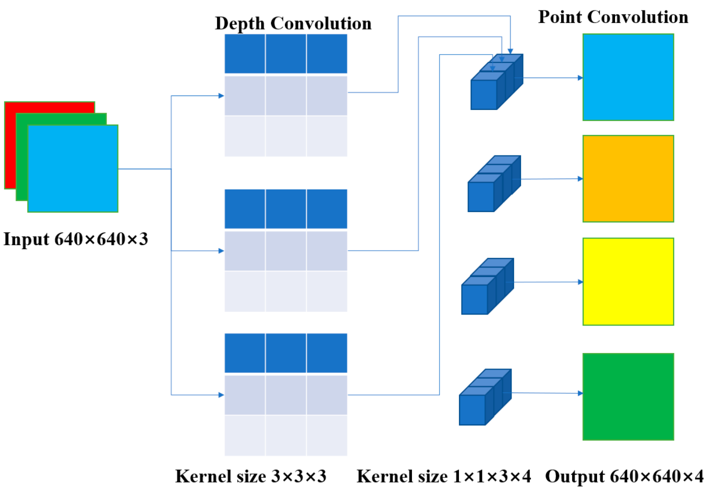


**Figure 9.** Deep separable convolution.

Using deep separable convolution for convolution operations can effectively reduce the number of parameters in the model, ensuring the feature extraction ability of the convolution and facilitating

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



the lightweight of the model. In addition, the Hard Swish activation function is selected as the activation function of the belt conveyor foreign object detection model, as shown in the Figure 10.

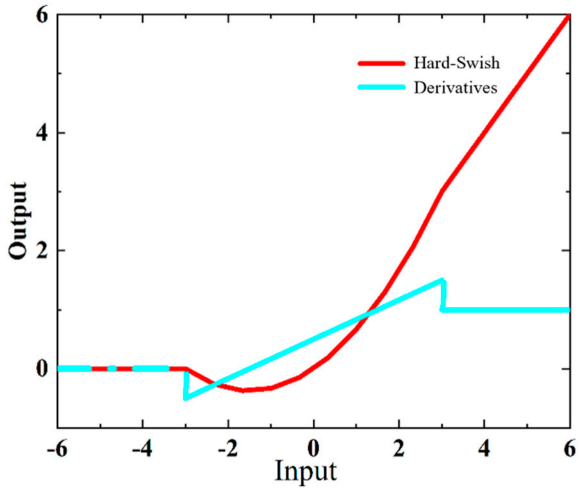


**Figure 10.** Hard-Swish activation function and its derivative.

The basic module of the improved foreign object detection model is shown in the Figure 11. Replace the ordinary convolution at the end of the merge channel in the CSP1_ X and CSP2_X module with a depthwise separatable convolution to reduce the number of parameters in the convolution process and accelerate the inference speed of the model.

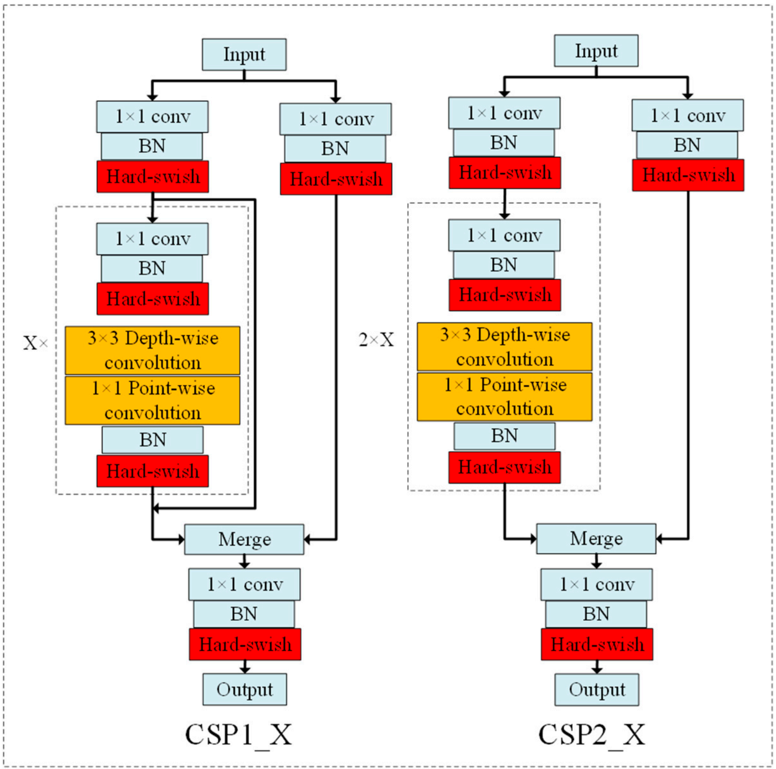


**Figure 11.** Improved CSP1_ X and CSP2_ X Structural block.

*3.4. IAT image enhancement module*

In order to ensure the end-to-end output characteristics of depth learning, the IAT image enhancement module [29] is introduced to achieve image enhancement, and the network structure is shown in Figure 12. The IAT module can enhance the brightness of the image, restore the relevant details, improve the image quality, reduce the noise and enhance the image contrast.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



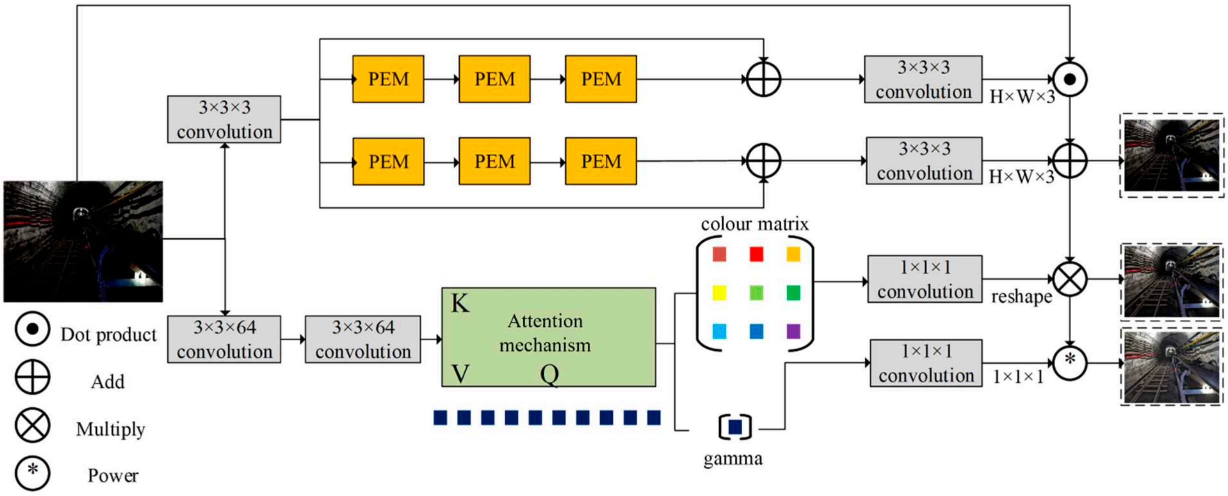


**Figure 12.** IAT image enhancement structure.

At the same time, the objective evaluation index Peak Signal to Noise Ratio (PSNR) for image enhancement is used as the specific evaluation index for image enhancement, and the formula is as follows:

$$PSNR = 10\log_{10}\left[\frac{(2^n - 1)^2}{MSE}\right] \tag{5}$$

$$MSE = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W}(X(i,j) - Y(i,j))^2 \tag{6}$$

where $X(i, j)$ is the pixel values of the original image, $Y(i, j)$ is the pixel values of the enhanced image, $H$ and $W$ are the length and width of the image, respectively.

*3.5. Improved CBAM attention block*

CBAM [30] is a convolutional neural network module based on attention mechanism, which is used to improve the overall performance of the model. Its essence is to inhibit the expression of redundant features by increasing the weight of non-redundant features. It is composed of channel attention module (CAM) and spatial attention module (SAM), and the specific network structure is shown in Figure 13.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



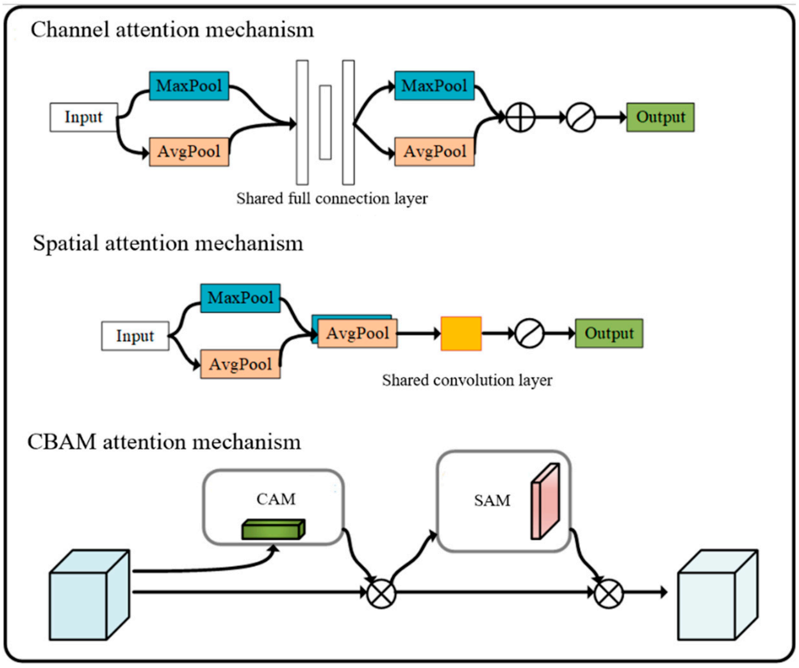


**Figure 13.** CBAM attention mechanism.

Therefore, in order to suppress redundant features and obtain attention feature maps that pay more attention to channels and spaces, the CBAM attention mechanism is introduced into the network structure, and the specific location of the addition is shown in the Figure 14.

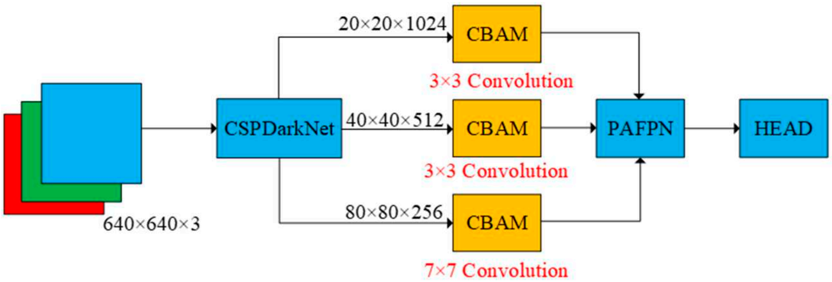


**Figure 14.** Improve the adding location of CBAM in the network.

*3.6. Designed rotating decoupling head and MO-YOLOX network*

The detection boxes of YOLO series object detection algorithms are all horizontal boxes, which is not conducive to the detection of foreign objects with diverse distribution directions such as ironware. Therefore, angle regression prediction is added to head network of YOLOX, and a branch decoupling head based on angle regression is constructed to accurately locate directional foreign objects. The structure of the rotary decoupling head is shown in figure 15, and the overall network structure of MO-YOLOX is shown in Figure 16.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



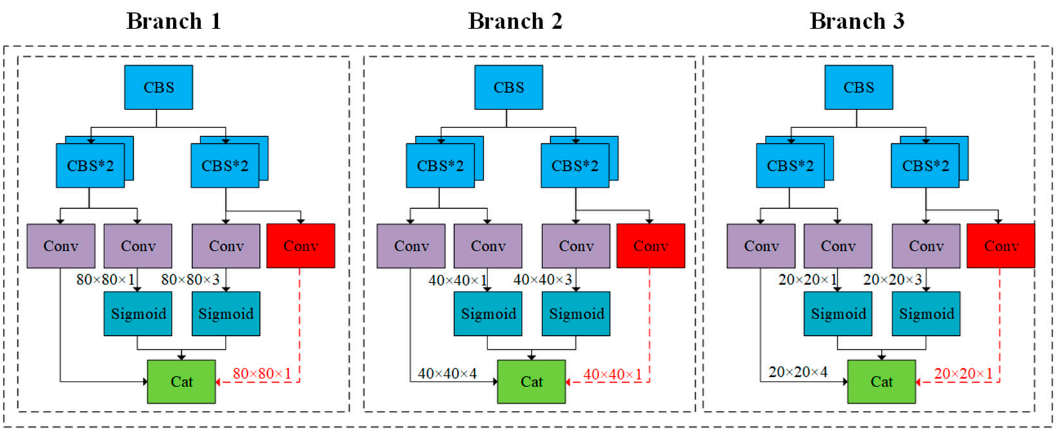


**Figure 15.** MO-YOLOX rotary decoupling head structure.

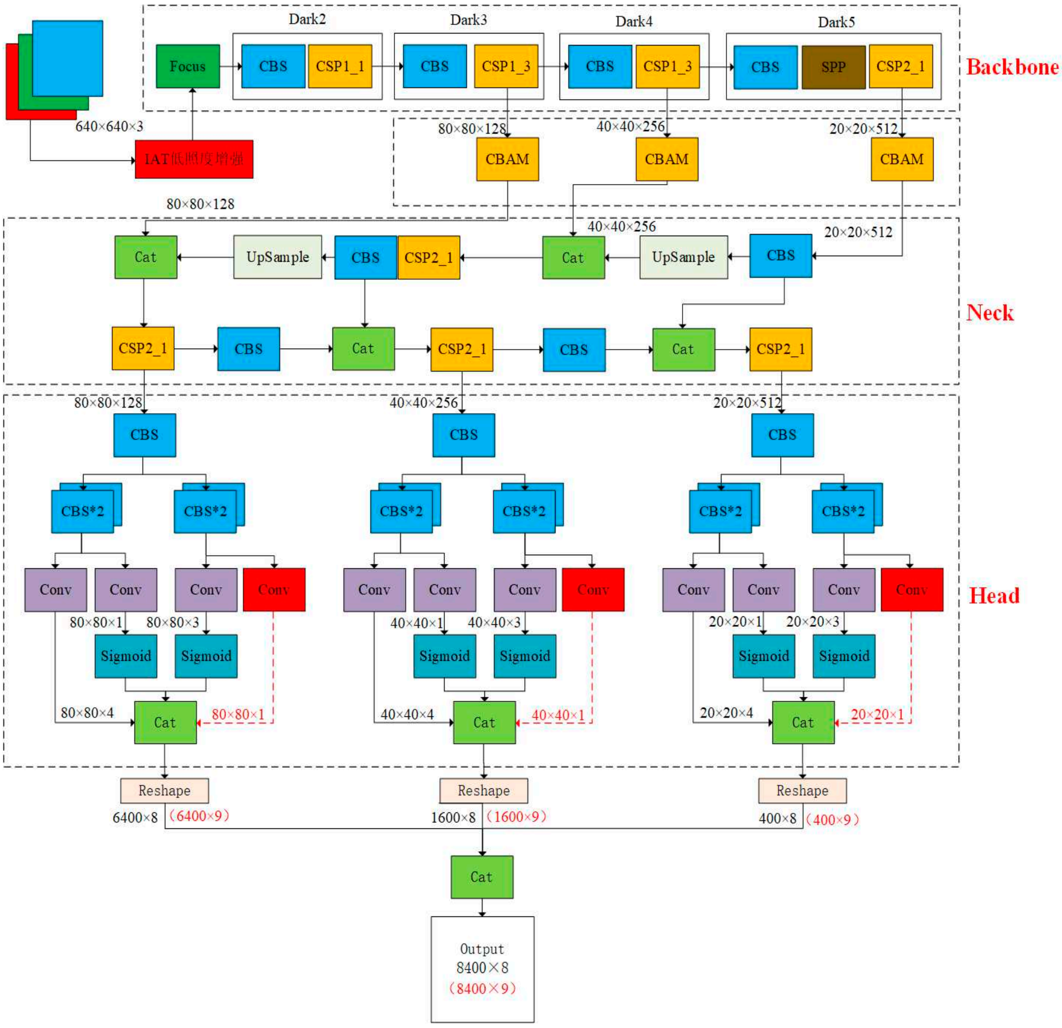


**Figure 16.** MO-YOLOX network model.

## 4. Experimental example and discussion

### *4.1. Experimental platform*

The proposed MO-YOLOX network model is trained in the GPU environment, and the environment configuration is shown in Table 1 below.

**Table 1.** Model training environment configuration.

 doi:10.20944/preprints202307.1534.v1



| name | parameter |
|---|---|
| CPU | Intel Core i9-10980XE |
| Hard disk | 2T |
| GPU | NVIDIA RTX A4000 |
| Memory | 16G |
| Deep learning framework | Pytorch1.8.0 |
| OS | Window10 |
| Programming Language | Python3.8 |
| CUDA | 11.2 |

*4.2. Experimental tests*

The training dataset is the foreign object detection dataset of belt conveyor, including horizontal frame labeling dataset and rotating frame marking dataset, and the relevant parameters are shown in Table 2. After 300 rounds of model training iterations, the proposed model can converge to relatively stable positions, and the loss values during the training process are shown in Figure 17. The model obtained from the above training is used as the optimal model for experimental comparison, and comparison experiments are conducted.

**Table 2.** Model training parameters.

| Training parameters | Setting values |
|---|---|
| activation function | Hard-Swish |
| Pooling method | Max-Pooling |
| optimization algorithm | Adams，Batch-size=8， |
| loss function | Cross entropy Loss function, KLD |
| Epoch | 300 |
| data enhancement | Mosaic |
| Learning rate | Initial Learning rate $\alpha_0 = 0.01$, Nature Index attenuation |
| Dataset partitioning ratio | Training set: Verification set: Test set=0.6:0.3:0.1 |

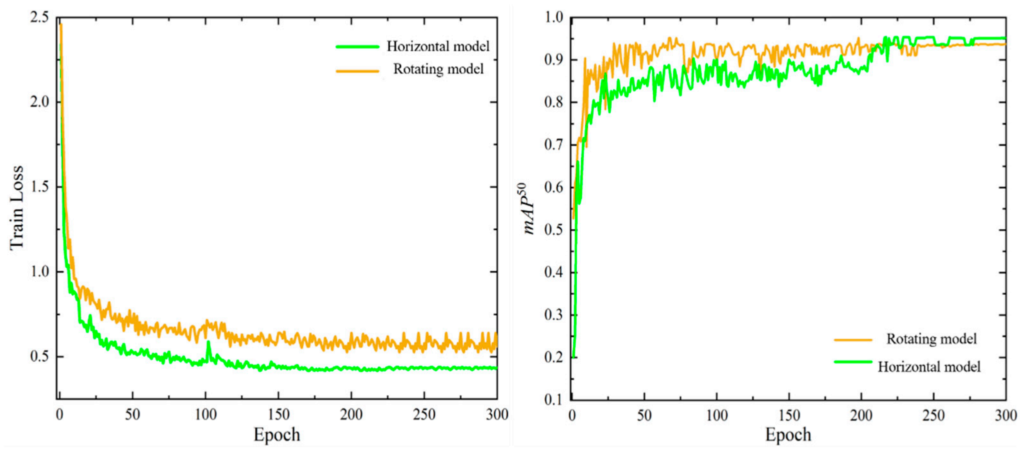


**Figure 17.** Model training curve.

 doi:10.20944/preprints202307.1534.v1



In order to verify the detection performance of the foreign object detection model obtained in this article on the dataset, a self-made belt conveyor foreign object detection dataset is used for testing, with a total of 1070 images (including 253 background images). The confusion matrixes during the foreign object detection process are obtained, as shown in Figure 18. Ten-fold cross-validation is used to comprehensively evaluate the performance of the model, and the results of the cross-validation is shown in Figure 19.

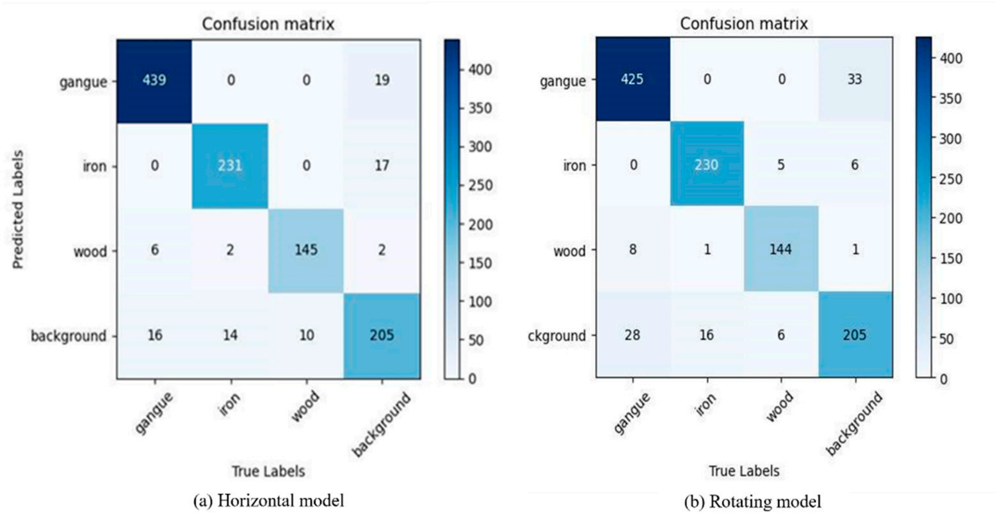


**Figure 18.** Confusion matrix result of self-made dataset.

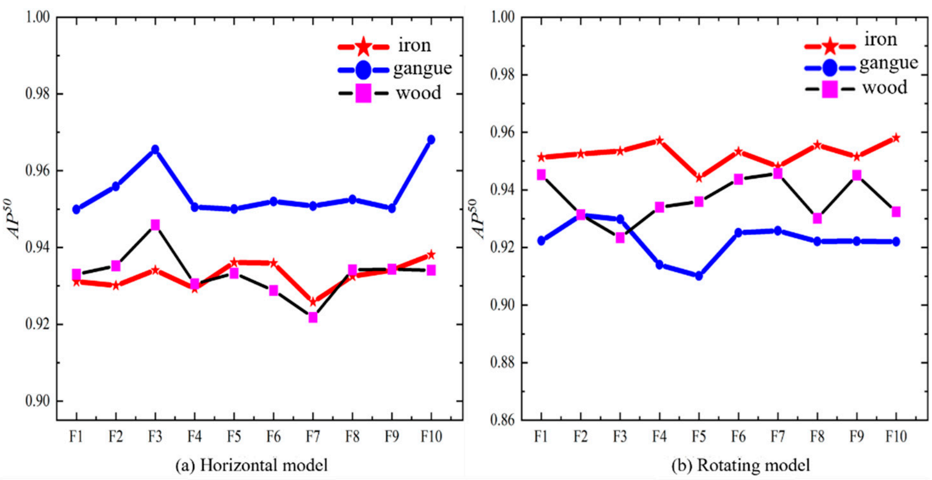


**Figure 19.** Cross validation results of foreign object detection models.

It can be seen from Figures 20 and 21 that the proposed foreign object detection model can effectively detect foreign objects in the case of coal flow background. The rectangle in figures is the target result predicted by the foreign object detection model, and different colors represent different categories. In Figure 21, the predicted angle information is represented by the long side of the rotating rectangular box, with angle values of 36.8, -30.3, and 65.1, which can verify the effectiveness of the rotation decoupling head in angle regression prediction. Figures 22 and 23 show the results of foreign object detection under coal flow occlusion and the multi-angle detection results of the same foreign object, respectively. The proposed model can locate the foreign object in the image more accurately, and the performance indicators of the foreign object detection model are shown in Tables 3 and 4.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 July 2023 doi:10.20944/preprints202307.1534.v1



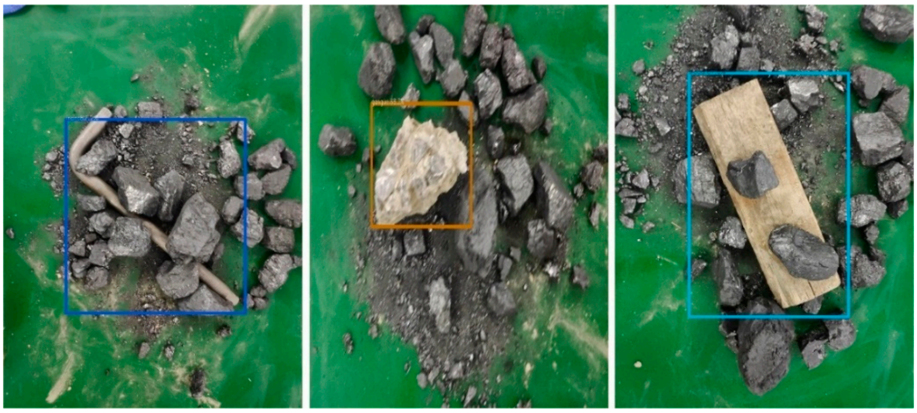

**Figure 20.** Test results of foreign object detection of horizontal frame.

**Figure 21.** Test results of foreign object detection of rotating frame.

（a）Horizontal box detection （b）Rotation box detection

**Figure 22.** Detection of foreign object under the shelter of coal flow.

 doi:10.20944/preprints202307.1534.v1



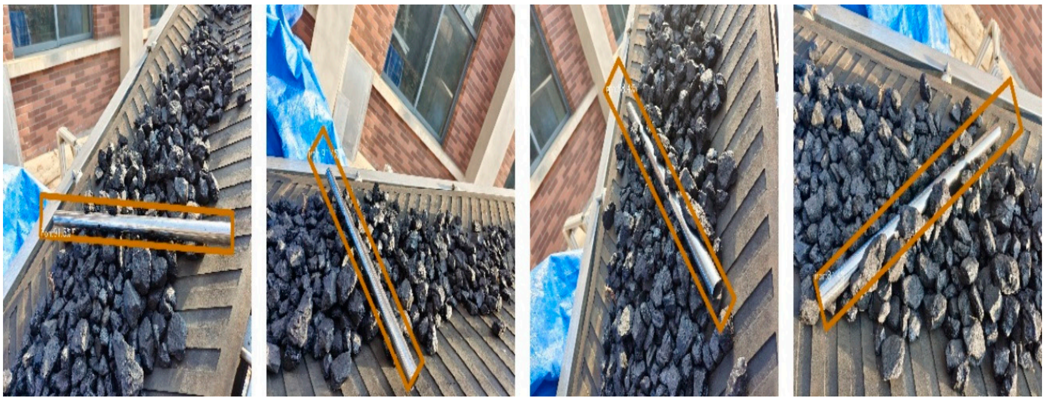

**Figure 23.** Multi-angle foreign object detection under the same foreign object sample.

**Table 3.** MO-YOLOX horizontal foreign object detection performance index parameters.

| | Precision | Recall | $AP^{50}$ | *F2-score* | Inference time /ms |
|---|---|---|---|---|---|
| iron | 93.71% | 93.20% | 93.27% | 93.30% | 21 |
| wood | 93.12% | 93.62% | 93.30% | 95.80% | 23 |
| Bulk gangue | 95.32% | 95.92% | 95.45% | 93.52% | 22 |
| average value | 94.05% | 94.25% | 94.01% | 94.20% | 22 |

**Table 4.** MO-YOLOX rotating frame foreign object detection performance index parameters.

| | Precision | Recall | $AP^{50}$ | *F2-score* | Inference time /ms |
|---|---|---|---|---|---|
| iron | 93.71% | 93.20% | 93.27% | 95.28% | 28 |
| wood | 93.12% | 93.62% | 93.30% | 93.46% | 26 |
| large gangue | 95.32% | 95.92% | 95.45% | 92.44% | 29 |
| average value | 94.05% | 94.25% | 94.01% | 93.73% | 27.7 |

From the experimental results, it can be seen that the proposed foreign object detection model of the belt conveyor can accurately detect foreign objects on both the horizontal foreign object dataset and the rotating foreign object dataset. The accuracy rate and recall rate of large gangue in the horizontal frame detection are higher, but detection effect of the iron is poor when the length width ratio of foreign object detection samples is obvious. The rotating frame foreign object detection has high accuracy on targets with large length width ratio, and the reasoning time is correspondingly increased. Iron and wood with obvious length width ratio can be effectively detected, but for gangue, the characteristics are quite different, and the angle information of data labels is distributed irregularly, which brings great difficulties to the angle regression prediction of the network. Moreover, in the presence of slight coal background occlusion, both horizontal frame foreign object detection and rotating foreign object detection can accurately detect foreign objects and determine the category of foreign objects.

## 5. Conclusions and future works

In this paper, a foreign object image dataset for the belt conveyor is collected and established, and the IAT image enhancement module and CBAM attention mechanism are introduced. Secondly, a novel rotating decoupling head is designed to predict the angle information of foreign objects, and a MO-YOLOX network structure is constructed. The experimental results show that the proposed algorithm has a performance of 71.9% and 73.2% on the VOC and DOTA test datasets, respectively, with an average inference time of around 26ms, which can meet the requirements of real-time inference. Ten-fold cross validation is conducted on the self-built foreign object dataset of the belt conveyor, and the accuracy, recall, and $mAP^{50}$ of horizontal frame foreign object detection are 94.05%, 94.25%, and 94.01%, respectively. Moreover, the accuracy, recall, and $mAP^{50}$ of the rotating frame foreign object detection reaches 93.87%, 93.69%, and 93.72%, and the average inference time of foreign object detection is 25ms.

However, the proposed foreign object detection method for belt conveyors designed has not yet considered the foreign object removal method in the three-dimensional coordinate system. In the future, further researches are needed on the foreign object removal method in the three-dimensional coordinate system.

**A**cknowledgments：The support of the Lianyungang 521 High-level Talent Training Project (LYG06521202203), Qing Lan project for excellent teaching team of Jiangsu province (2022), National Natural Science Foundation of China (51975568) and the Natural Science Foundation of Jiangsu Province (BK20191341) in carrying out this research are gratefully acknowledged.

**Conflicts of Interest：** The authors declare that there is no conflict of interests regarding the publication of this article.

**E**thics statement：N/A.

**Data access statement：** All data that produce the results in this work can be requested from the corresponding author.

## References

1. Petrikova I, Marvalova B, Samal S, et al. Digital image correlation as a measurement tool for large deformations of a Conveyor Belt. Applied Mechanics and Materials, 2015, 732: 77-80.
2. Zimroz R, Stefaniak P K, Bartelmus W, et al. Novel techniques of diagnostic data processing for belt conveyor maintenance. Proceedings of the 12th International Symposium Continuous Surface Mining-Aachen 2014. Springer International Publishing, 2015: 31-40.
3. Chen W, Wang X. Coal mine safety intelligent monitoring based on wireless sensor network. IEEE sensors journal, 2020, 21(22): 25465-25471.
4. Gao D, Li W, Dai K. Design of coal mine intelligent monitoring system based on ZigBee wireless sensor network. International Conference on Mechanics, Materials and Structural Engineering (ICMMSE 2016). Atlantis Press, 2016: 182-187.
5. Norkobil Saydirasulovich S, Abdusalomov A, Jamil M K, et al. A YOLOv6-based improved fire detection approach for smart city environments. Sensors, 2023, 23(6): 3161.
6. Xu P, Zhou Z, Geng Z. Safety monitoring method of moving target in underground coal mine based on computer vision processing. Scientific reports, 2022, 12(1): 17899.
7. Dai L, Qi P, Lu H, et al. Image enhancement method in underground coal mines based on an improved particle swarm optimization algorithm. Applied Sciences, 2023, 13(5): 3254.
8. Wei W, Li L, Shi W, et al. Ultrasonic imaging recognition of coal-rock interface based on the improved variational mode decomposition. Measurement, 2021, 170: 108728.
9. Radke R J, Andra S, Al-Kofahi O, et al. Image change detection algorithms: a systematic survey. IEEE transactions on image processing, 2005, 14(3): 294-307.
10. Jiang L, Peng G, Xu B, et al. Foreign object recognition technology for port transportation channel based on automatic image recognition. EURASIP Journal on Image and Video Processing, 2018, 2018: 1-9.
11. Donskoi E, Suthers S P, Campbell J J, et al. Modelling and optimization of hydrocyclone for iron ore fines beneficiation-using optical image analysis and iron ore texture classification. International Journal of Mineral Processing, 2008, 87(3-4): 106-119.
12. Chatterjee S, Bandopadhyay S, Machuca D. Ore grade prediction using a genetic algorithm and clustering based ensemble neural network model. Mathematical Geosciences, 2010, 42: 309-326.
13. Zhang Z, Su X, Ding L, et al. Multi-scale image segmentation of coal piles on a belt based on the Hessian matrix. Particuology, 2013, 11(5): 549-555.
14. Saran G, Ganguly A, Tripathi V, et al. Multi-modal imaging-based foreign particle detection system on coal conveyor belt. Transactions of the Indian Institute of Metals, 2022, 75(9): 2231-2240.
15. Tu L, Zhong S, Peng Q. Moving object detection method based on complementary multi resolution background models. Journal of Central South University, 2014, 21(6): 2306-2314.
16. Lins R G, Givigi S N. Automatic crack detection and measurement based on image analysis. IEEE Transactions on Instrumentation and Measurement, 2016, 65(3): 583-590.
17. Ghasemi Y, Jeong H, Choi S H, et al. Deep learning-based object detection in augmented reality: A systematic review. Computers in Industry, 2022, 139: 103661.
18. Kaur J, Singh W. Tools, techniques, datasets and application areas for object detection in an image: a review. Multimedia Tools and Applications, 2022, 81(27): 38297-38351.
19. LeCun Y, Bengio Y, Hinton G. Deep learning. nature, 2015, 521(7553): 436-444.

 doi:10.20944/preprints202307.1534.v1




20. Pu Y, Apel D B, Szmigiel A, et al. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning. Energies, 2019, 12(9): 1735.
21. Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
22. Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 2015, 28.
23. Wang Y, Wang Y, Dang L. Video detection of foreign objects on the surface of belt conveyor underground coal mine based on improved SSD. Journal of Ambient Intelligence and Humanized Computing, 2020: 1-10.
24. Chen Y, Sun X, Xu L, et al. Application of YOLOv4 algorithm for foreign object detection on a belt conveyor in a low-illumination environment. Sensors, 2022, 22(18): 6851.
25. Ma G, Wang X, Liu J, et al. Intelligent detection of foreign matter in coal mine transportation belt based on convolution neural network. Scientific Programming, 2022, 2022.
26. Xiao D, Kang Z, Yu H, et al. Research on belt foreign body detection method based on deep learning. Transactions of the Institute of Measurement and Control, 2022, 44(15): 2919-2927.
27. Ge Z, Liu S, Wang F, et al. Yolox: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430, 2021.
28. Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.
29. Cui Z, Li K, Gu L, et al. You Only Need 90K Parameters to Adapt Light: A Light Weight Transformer for Image Enhancement and Exposure Correction. BMVC. 2022: 238.
30. Yang X, Zhang G, Yang X, et al. Detecting rotated objects as gaussian distributions and its 3-d generalization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(4): 4335-4354.