Defect Detection Study of Connecting Buckles for Mining Powerful Transportation Belts with Deep Learning

Xiaoqing Yu, Hongen Wu, Hu Cui, Yongding Xie, Yongquan Wang, Xiliang Ma

***Abstract*—The connecting buckles of underground transportation belts are a critical component. These components are susceptible to damage from material impact and pulley wear-induced damage, which can compromise industrial safety and transportation efficiency. While the importance of detecting defects in connecting buckles is self-evident, effective methods to solve this problem are currently lacking. To fill this gap, this study proposed a defect detection method using an enhanced version of the YOLOv7 model. The process incorporated a target detection head, an attention mechanism, and a loss function to enhance the model's generalization ability and robustness under complex working environments and diverse defect types. Through experimental validation using a self-constructed U38 buckle dataset, the method achieved an 83.2% detection rate for various defect types, with critical defects identified with an accuracy of up to 98.1%. The enhanced YOLOv7_m model exhibited strong performance in defect detection, providing valuable technical support for improving the safety and efficiency of coal mine transportation. This research provides a valuable reference for further optimizing the diagnosis of connection buckle defects based on deep learning.**

***Index Terms*—Deep learning; Defect detection; Connecting Buckles; Machine vision**

I. INTRODUCTION

The powerful belt conveyor (see Fig. 1(a)) is widely used for bulk material transportation in mining systems [1] and plays a crucial role [2]. It relies on connection devices to link various components. The operating space in the mine is more confined compared to ground-level operations (see Fig. 1(b)). Traditional vulcanized splicing is greatly restricted, and this method is only employed in fixed transportation scenarios. For frequently moving working surfaces and changing occasions, the conveyor belt needs to be disassembled and reassembled frequently, and the conveyor belt connecting buckles have gradually become the main connecting devices. However, damage to the connecting buckles may cause the entire conveyor belt to tear and disconnect, or even lead to system failure [3]. Therefore, the detection and maintenance of the connecting buckles are crucial for the normal operation of coal mine production.

Defect detection methods can be broadly categorized into manual detection methods [4] and non-destructive testing methods [5]. Manual detection primarily relies on human touch and visual inspection. Non-destructive testing methods offer high accuracy and real-time capabilities, making them crucial for preventing conveyor belt tearing and enabling real-time monitoring. Machine vision methods are the predominant technology in this field. Detecting damage to buckles is essentially a matter of target detection. In particular, deep learning-based target detection methods exhibit superior accuracy and speed compared to traditional approaches. The R-CNNs (Region-based Convolutional Neural Networks) [6] algorithm is a pioneering network architecture that applies deep learning to target detection. However, this algorithm suffers from a large number of candidate frames, making it challenging to meet real-time requirements. Since then, single-stage target detection models, including SSD (Single Shot MultiBox Detector) [7], CenterNet [8], and YOLO (You Only Look Once) [9], have been introduced. Given the belt speed ($\geq$1.5m/s), and the detection efficiency of the end device [10], the detection model must balance accuracy and speed. Single-stage models are better suited to meet this requirement. M. Y. Chen et al. [11] proposed an enhanced YOLOv4-tiny algorithm for defect identification on conveyor belts. They incorporated image segmentation techniques to accurately identify defects. T. Z. Liu et al. [12] improved the YOLOv4 model for the problem of jujube fruit classification by combining the loss function to enhance the detection accuracy. Y. Z. Fu et al. [13] improved the YOLOv5-based method for excellent defect detection in car door trim panels. Y. Zhang et al. [14] proposed a new heavy-necked object detection model based on YOLOv5s that optimizes the feature fusion capability of pyramid networks.

Manuscript received May 2, 2024; revised September 25, 2024.

Hongen Wu is an associate professor of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China (corresponding author, phone: +86+13969603895; e-mail: whenli@sina.com)

Xiaoqing Yu is a postgraduate student of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China. (e-mail: 1468915885@qq.com)

Hu Cui is a postgraduate student of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China. (e-mail: 1343522438@qq.com)

Yongding Xie is a postgraduate student of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China. (e-mail: 3340568505@qq.com)

Yongquan Wang is a postgraduate student of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China. (e-mail: wangyq990524@163.com)

Xiliang Ma is a postgraduate student of College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, SD, China. (e-mail: 2831915166@qq.com)

(a) The powerful belt conveyor (b) The operating space in the mine

Fig. 1. The mining conveyor.

The main contributions of this paper are as follows: (1) A unique dataset of connection buckles of the U38 model of mining powerful transportation belts was created through on-site data collection and expansion. (2) To tackle the challenges of complex backgrounds [15] and varying defect scales during defect detection, the enhanced model in this paper addressed localization and classification issues in defect detection. To address the issue of feature information loss caused by multiple downsampling of the feature map, the enhanced model emphasized focusing on identified targets. Furthermore, to address inaccuracies in the original model's computation of the loss function, the enhanced model optimized the loss function for bounding box regression. This optimization enhanced the detection of small target defects on the connecting buckles. (3) The proposed detection method is straightforward, does not require additional equipment, and can uniformly recognize and detect multiple types of defects.

II. Data collection and processing

A. Preparation of data

Computing power, algorithms, and data are the three fundamental elements of artificial intelligence [16][17], while the quantity and quality of data are crucial factors affecting the detection effectiveness in deep learning for damage detection. Due to the unique characteristics of the mining connection buckle, neither the current publicly available datasets nor the galleries on the web provide the datasets used in this paper. Therefore, this paper utilized on-site collection and subsequent manual labeling methods to create a new dataset for training in damage detection. A portion of the dataset is illustrated in Fig. 2.

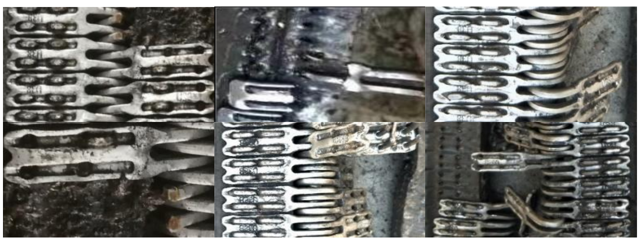
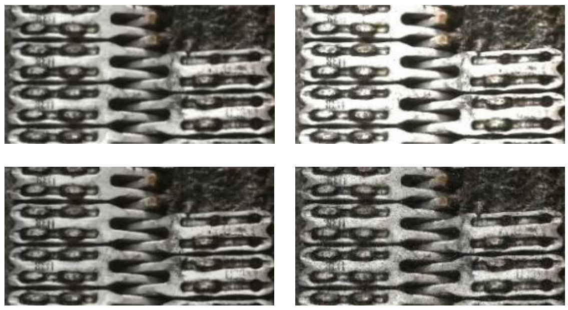


Fig. 2. The data set of defective connection buckles for some U38 models.

The dataset in this paper was divided into only two categories: normal and damaged. This categorization is based on the understanding that any damage can affect the entire system. Therefore, detect the damage and sound an alarm, allowing the field technician to decide whether the replacement should be halted.

Due to the specificity of the research object, only 961 (561 damage images + 400 normal connection buckle images) original images (resolution: 3060 × 4080) were obtained from on-site shooting, but the accuracy of damage detection requires a large amount of data support for connecting buckle defects, so here the original images were data enhanced and expanded [18]. After the expansion to 3366 data sets, some of the data enhancement effects are shown in Fig. 3, plus the normal connection buckle image a total of 3766 data sets of images, the specific types and quantities of expansion are shown in TABLE I below. Because a larger proportion of the training set can improve the model's ability to learn and generalize the defects of the connecting buckle and reduce the risk of overfitting, the image samples in this paper were roughly divided into the training set, validation set, and test set in the ratio of 8:1:1.

(a) Original data



(b) Post-enhancement data

Fig. 3. Data enhancement effects.

TABLE I. Number of expanded classifications for specific datasets.

Typology	Revolve	Noises	Glare	Darkness	Zoom	Contrast enhancement	Total
Training sets	356	511	383	524	441	488	2703
Validation sets	40	75	50	53	56	52	326
Test sets	48	78	36	53	57	65	337
Total	444	664	469	630	554	605	3366

B. Sample labeling

In this paper, Make Sense software was used to annotate the images, as illustrated in Fig. 4. The dataset labels were stored in .txt format for training in this paper. The labeling samples in this paper were classified with the following labels: the normal connection buckling category was 1, and the label was DefectsWC (Defects Will Occur); the damage category was 0, and the label was DefectsOC (Defects Occurred).

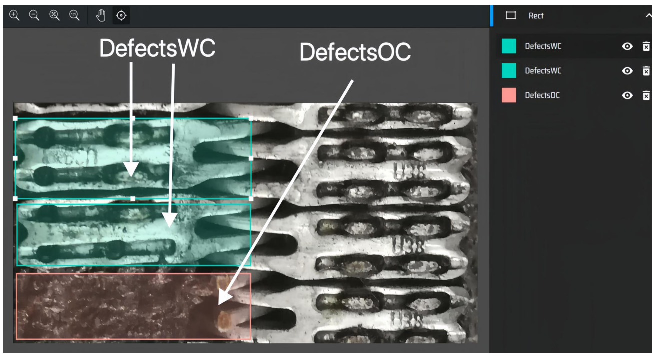


Fig. 4. Image annotation process.

III. YOLOV7 MODEL IMPROVEMENTS

A. *YOLOv7 General Structure*

The YOLOv7 [26] architecture described in this article consists of four primary components: Input, Backbone, Neck, and Head. The network structure is illustrated in Fig 5. To enhance the data volume and increase the batch size, the Mosaic data augmentation technique is applied to the Input. The Backbone network consists of several layers, including CBS, ELAN, and MPConv. ELAN consists of multiple CBS modules that maintain the initial gradient path, thereby enhancing the network's learning potential and enabling better acquisition of diversified features. The MPConv layer, which consists of Maxpool and convolutional layers, incorporates two branches that combine features extracted from upper and lower layers using the Concat module. This fusion enhances the model's overall feature extraction capability. The Neck network consists of a structural connection composed of CBS convolution modules, MP modules, E-ELAN, and SPPCSPC. In the Neck section, high-level features are combined with features from the lower layer, thereby enhancing the model's feature extraction capability.

B. *YOLOv7 Algorithm Improvement*

Add Dynamic Head

Conveyor belt buckle defect detection does not require consideration of the classification problem; therefore, it is essential to combine localization and classification effectively. DyHead (Dynamic Head) [19] combines scale-awareness, task-awareness, and spatial-awareness.

The principle is as follows:

The output of L different levels of feature maps from the backbone is:

$$F_{in} = \{F_i\}_{i=1}^{L} \tag{1}$$

The feature pyramid is rescaled to obtain a four-dimensional feature map:

$$F \in R^{L\times H\times W\times C} \tag{2}$$

Where L is the number of levels in the pyramid, H is the height of the median level feature, W is the width of the median level feature, and C is the channel of the median level feature.

The four-dimensional feature map is further reshaped into a three-dimensional tensor:

$$F \in R^{L\times S\times C}\ (S = H\times W) \tag{3}$$

The traditional method of directly learning the attention function on all dimensions is computationally large. So it is split into three serial attentions and each of them focuses on a single dimension only.

$$W(F)=\Pi_C\left(\Pi_S\left(\Pi_L(F)\cdot F\right)\cdot F\right)\cdot F \tag{4}$$

where $\Pi(*)$ is the attention functions in dimensions C, S, and L, respectively.

The overall process is to carry out the operations of Eq.(4) multiple times by connecting the above three processes (1), (2), and (3) in series, and keep stacking multiple attention modules, as shown in Fig. 6 below, which shows the configuration of the DyHead module.

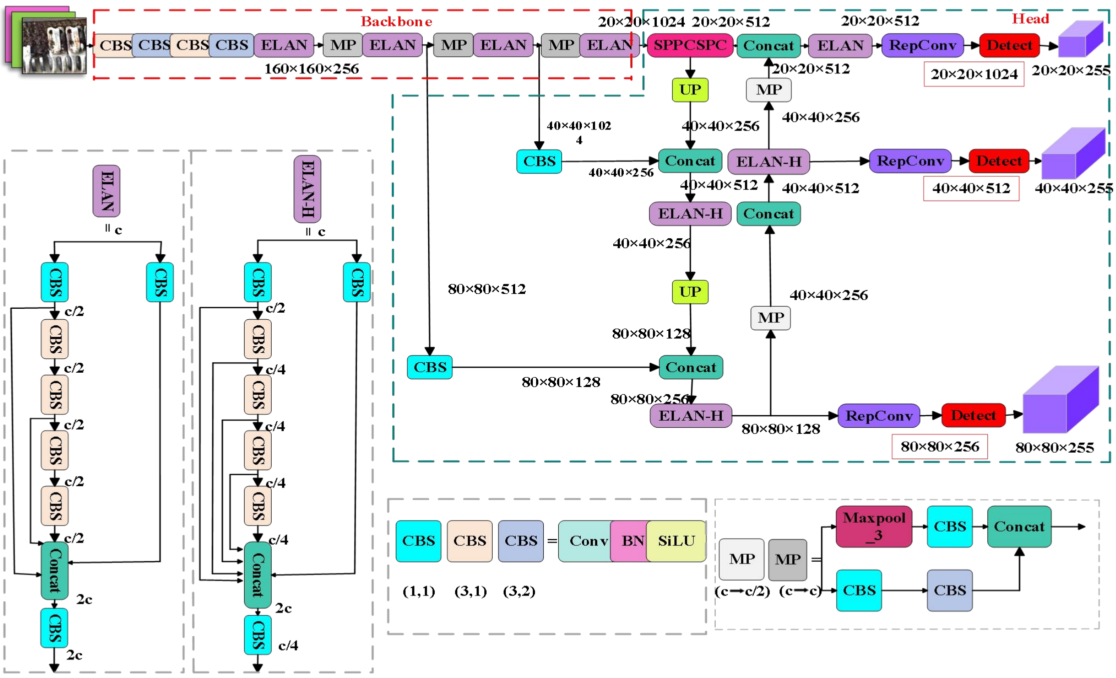
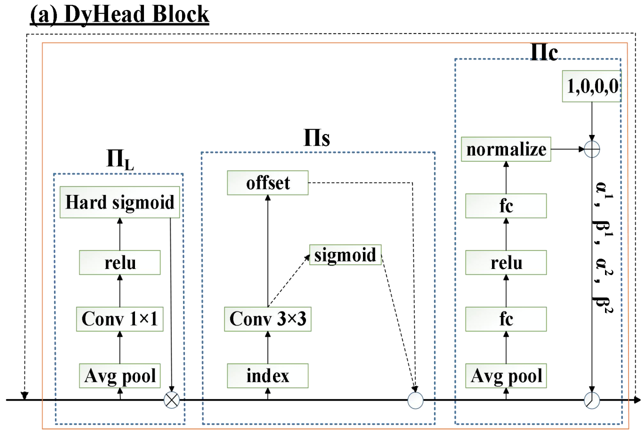
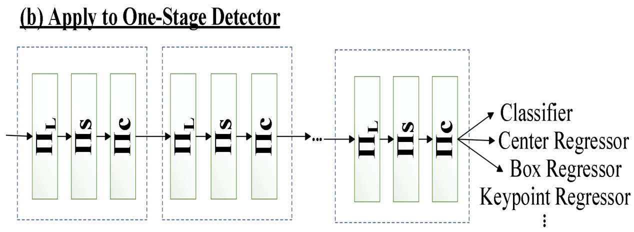


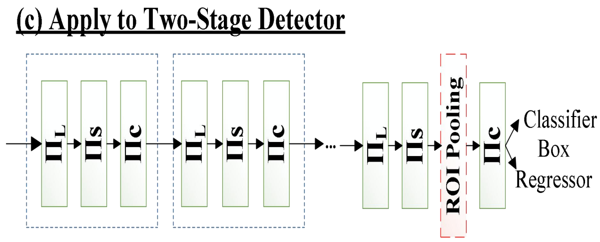
Fig. 5. YOLOv7 network structure diagram.



(a) Detailed implementation process for each attention module



(b) DyHead blocks applied to a primary target detector process



(c) DyHead blocks applied to the secondary target detector process

Fig. 6. DyHead detailed configuration.

Add Squeeze-and-Excitation Networks

For defect detection of conveyor belt buckles, it is essential to capture the most salient features. Incorporating a learning mechanism into neural networks is an important way to enhance the feature representation of convolutional neural networks. SENet (Squeeze-and-Excitation Networks) [20] is a mechanism that learns to selectively emphasize useful features and suppress irrelevant ones by utilizing global information

The SE block structure is schematically shown in Fig. 7:

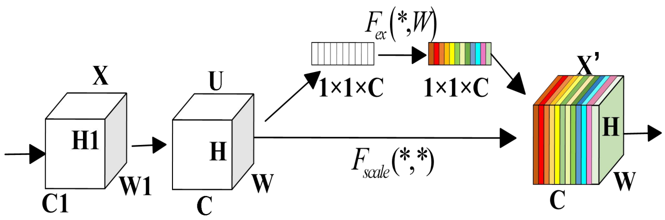
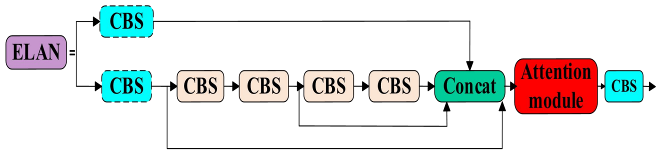
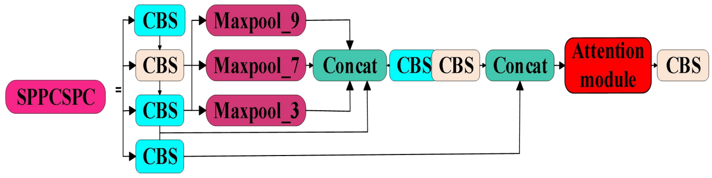


Fig. 7. SE block structure schematic.

The SENet module was added to the YOLOv7 model location as specified in Fig. 8 below.



(a) SENet was added to the ELAN module.



(b) SENet added to SPPCSPC module

Fig. 8. SENet was added to the YOLOv7-specific location.

Add the loss function MPDIoU

Current target detectors of the YOLO family [21-26] all rely on the Bounding Box Regression (BBR) module to determine object positions. Most of the current loss functions, however, cannot be optimized when the predicted box has the same aspect ratio as the actual labeled box but with different widths and heights. For this reason, the comparison metric MPDIoU[27], which incorporates all the correlations considered in the existing loss functions, is invoked, along with the MPDIoU-based bounding box loss function, called L_{MPDIoU} [27].

As in Fig. 9, MPDIoU serves as the boundary loss.

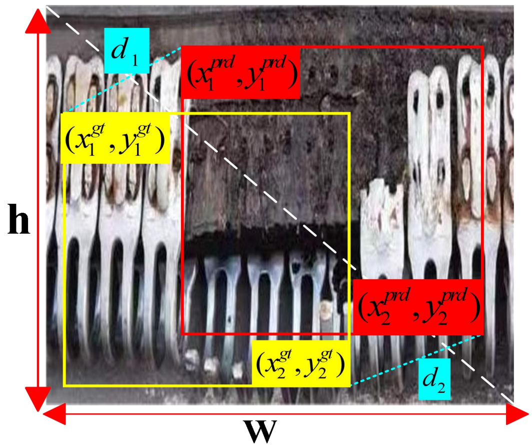


Fig. 9. MPDIoU serves as the boundary loss.

IV. Experimentation and Analysis

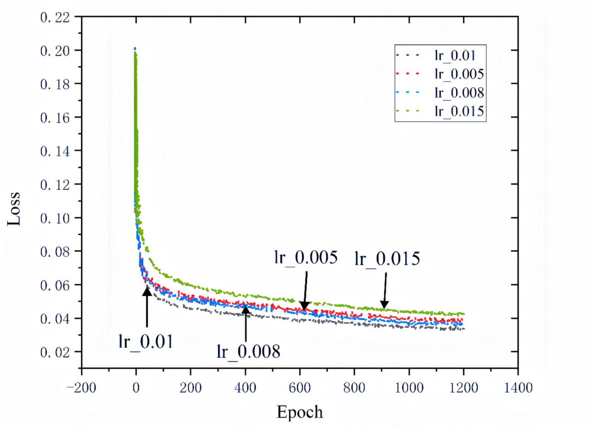
A. *Experimental environment and evaluation indicators*

The experimental environment is configured in TABLE II below:

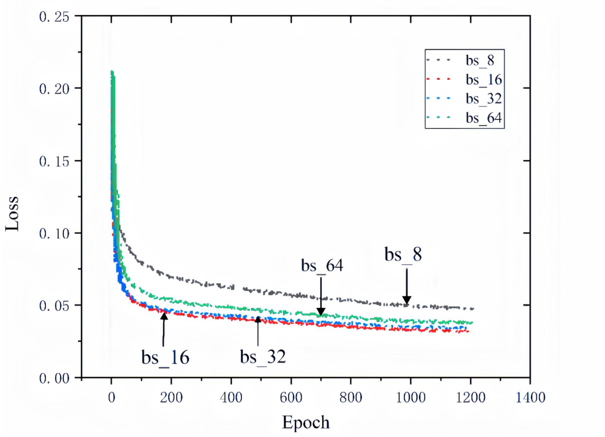
TABLE II. EXPERIMENTAL ENVIRONMENT CONFIGURATION.

OS	CPU	CPU clock speed	Graphics card	Video memory	Pytorch
Windows10	Inter(R) Coe i5-7300HQ	2.5GHz	NVIDIA GeForce GTX 1080ti	16GB	Pytorch1.11.1 +CUDA11.6

The optimizer Adam is used to update the network parameters iteratively. In this paper, after conducting several parameter-seeking optimization experiments, as illustrated in Fig. 10, the initial learning rate is set to 0.01, and the initial batch size is set to 16, resulting in the best loss function convergence. The final momentum is set to 0.937, and the Epochs are 400.



(a) Model loss function curve under different learning rates



(b) Model loss function curve under different batch_sizes

Fig. 10. Parameter-seeking optimization experiments.

In this experiment, the parameters Precision, Recall, mAP@0.5 (mean Average Precision, IOU=0.5), and the computational FLOPs, which reflect the complexity of the model, were used as the experimental evaluation metrics.

B. Modeling and experimental results

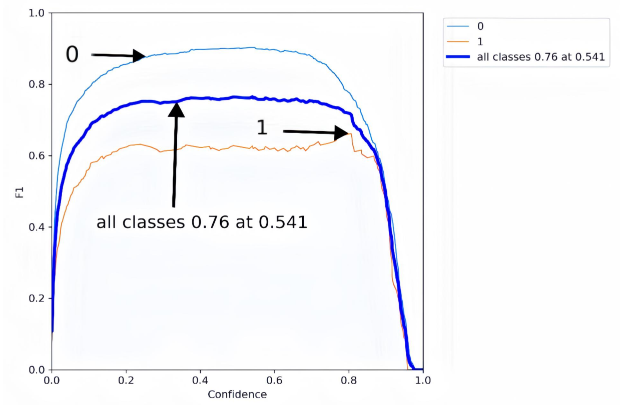
DyHead Module Comparison Experiment with Different Number of Modules

The number of control blocks is generally used to evaluate the efficiency of the Dynamic Head [19]. DyHead stacking counts comparison experiments are shown below in TABLE III.

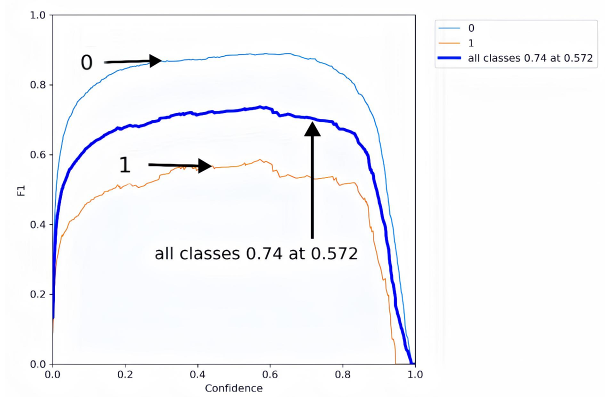
TABLE III. COMPARATIVE EXPERIMENTAL RESULTS OF DIFFERENT DYHEAD MODULES

Number	Precision(%)	Recall(%)	mAP@0.5(%)	FLOPs(G)
0	73.1	77.8	74.6	100.9
2	70.2	69.8	66.3	103.4
4	**73.2**	**78.2**	**74.7**	**105.9**
6	72.4	71.9	71.7	108.4

From the data in TABLE III, it can be seen that when the DyHead module is stacked 4 times and 6 times, the two metrics Precision and Recall conflict. In this case, the F1-score is used to judge, which is the harmonic mean of Precision and Recall. The results of the judgment are shown in Fig. 11.



(a) DyHead is stacked 4 times



(b) DyHead is stacked 6 times

Fig. 11. F1-score comparison.

From Fig. 11, it can be seen that when DyHead is stacked 4 times, the F1-score is 0.76, whereas when DyHead is stacked 6 times, the F1-score is only 0.74. Therefore, the model performs best when DyHead is stacked 4 times.

Comparative Experiments on Attentional Mechanisms

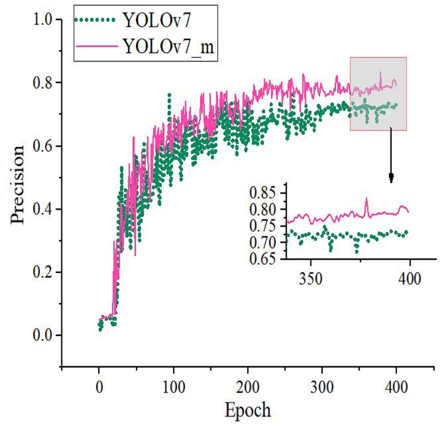
In the paper, to verify the effectiveness of the SENet [20] attention mechanism for detecting defects in the homemade connection buckling dataset, we integrated the CBAM (Channel Attention) [28], CA (Coord Attention) [29], and SA (Shuffle Attention) [30] attention mechanisms, which have demonstrated superior performance at this stage. The results of the comparative experiments are presented in TABLE IV.

TABLE IV. COMPARATIVE EXPERIMENTS ON ATTENTIONAL MECHANISMS.

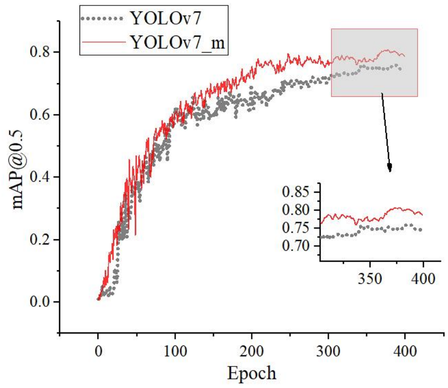
Model	Precision(%)	Recall(%)	mAP@0.5(%)	FLOPs(G)
YOLOv7	73.1	77.8	74.6	100.9
YOLOv7+SENet	**74.7**	**78.0**	**74.9**	**102.5**
YOLOv7+CBAM	74.3	76.6	73.3	105.1
YOLOv7+CA	72.3	73.9	72.5	101.1
YOLOv7+SA	73.6	75.0	74.8	105.1

Analysis of Experimental Results of Improved Model for Detection of Multiple Types of Defects

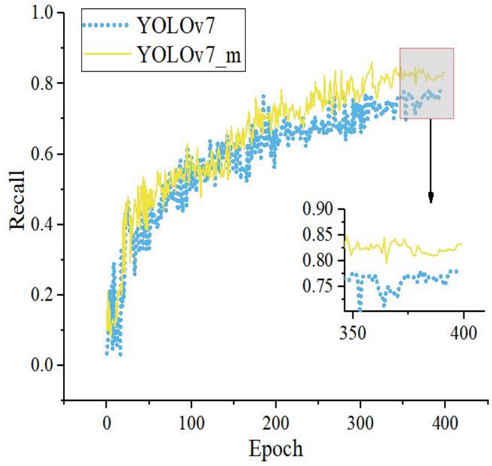
The loss function MPDIoU was added at the end of the enhanced model to optimize the boundary regression loss function. To distinguish it, the enhanced YOLOv7 model was referred to as YOLOv7_m in this paper. a comparison of the model before and after the enhancement is presented in Fig. 12 below.



(a) Precision curve



(b) mAP@0.5 curve



(c) Recall curve

Fig. 12. Comparison chart of experimental results.

From the above image, it can be observed that as the number of iterations increases, both Precision and Recall also increase. The fluctuations gradually decrease and eventually stabilize. No overfitting phenomenon is observed in either the original YOLOv7 model or the enhanced YOLOv7_m model. From the three figures in Fig. 12, it can be seen that the values of mAP, Precision, and Recall of the enhanced YOLOv7_m are higher than those of the original YOLOv7 model. As the model reaches saturation, the detection results stabilize, the convergence accelerates, and the performance improves.

To further validate the superiority of the enhanced algorithm, this paper also selected some of the mainstream defect detection algorithms at this stage and conducted comparative ablation experiments. The results of the experiments are presented in TABLE V.

TABLE V. COMPARISON OF ABLATION EXPERIMENTS BY DETECTION ALGORITHM.

Arithmetic	SENet	DyHead4	MPDIoU	P(%)	R(%)	mAP@0.5(%)
Faster R-CNN[31]				65.2	60.0	63.3
Faster R-CNN+SENet	√			65.8	61.9	64.2
Faster R-CNN+DyHead		√		69.3	66,4	64.5
Faster R-CNN_m	√	√	√	72.0	69.6	71.3
SSD [7]				66.1	62.4	66.0
SSD+SENet	√			66.9	62.3	66.2
SSD+DyHead		√		67.4	64.3	67.5
SSD_m	√	√	√	68.7	66.8	67.7
YOLOv4 [23]				69.1	70.2	71.5
YOLOv4+SENet	√			69.2	71.4	68.8
YOLOv4+DyHead		√		71.2	69.7	72.3
YOLOv4_m	√	√	√	70.1	70.4	69.2
YOLOv7 [26]				73.1	77.8	74.6
YOLOv7+SENet	√			74.7	78.0	74.9
YOLOv7+DyHead		√		73.2	78.2	74.7
YOLOv7_m	√	√	√	**79.3**	**83.2**	**78.8**
YOLOv8				69.8	71.4	70.9
YOLOv8+SENet	√			73.5	71.6	73.2
YOLOv8+DyHead		√		72.9	74.1	73.9
YOLOv8_m	√	√	√	73.4	78.8	74.7

(a) YOLOv7 test results

(b) YOLOv7_m test results

Fig. 14. Comparison chart of experimental results.

To enhance intuitiveness, the experimental results from different algorithms were represented as a line graph based on TABLE V, as shown in Fig. 13.

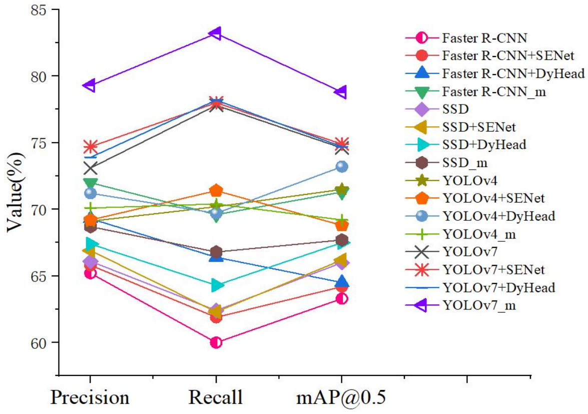


Fig. 13. Comparison chart of ablation experiments.

As demonstrated in TABLE V and Fig. 13, the SENet attention mechanism and DyHead can somewhat enhance the original model's detection performance. However, the final experimental results demonstrate that the YOLOv7_m algorithm significantly improves all aspects of defect detection compared to other mainstream algorithms. Compared with the original YOLOv7 model, the accuracy of the enhanced YOLOv7_m model in detecting defects in connecting buckles has increased by 6.2%, allowing for improved localization of faulty components. The Recall rate has improved by 5.4%, decreasing the defect leakage rate in the model. Additionally, the mean Average Precision (mAP) detection accuracy has improved by 4.2%, making it more suitable for detecting defects in connecting buckles.

To further verify that the improved YOLOv7_m model has a practical effect on the detection of connection buckles in the dataset, randomly selected dataset samples were detected, and the specific results are shown in Fig. 14.

Analysis of the Effectiveness of Improved Models for Critical Defect Detection

To more accurately assess the objectivity and operability of the improved model, this paper further considered the impact of the type and number of connection buckle defects on equipment performance. The study distinguished between the types of defects that need to be overhauled when the equipment is shut down and those that may lead to serious equipment failures and thus require immediate shutdown for treatment. This paper called the second type of defects as critical defects. In addition, three or more consecutive non-critical defect types were categorized as critical defects.

The precise identification of significant flaws in the connection buckle is essential for maintaining the safe functioning of the mine transportation system. To ensure the effectiveness of the detection model, the model must exhibit a high level of accuracy in identifying critical defects within the system. This accuracy is necessary for promptly detecting potential hazards and preventing the occurrence of serious accidents.

Hence, by the aforementioned classification criteria, a novel dataset comprising critical defects of connection buckles was established for experimental validation of the model, as illustrated in TABLE VI.

The YOLOv8 model, which had performed relatively well in comprehensive defect-type detection, was selected to conduct comparison experiments with YOLOv7 and the improved model YOLOv7_m. The experimental results are presented in Table VII.

TABLE VI. NUMBER OF EXPANDED CLASSIFICATIONS FOR SPECIFIC DATASETS.

Typology	Severe defects	High number of defects	Total
Training sets	457	366	823
Validation sets	53	47	100
Test sets	59	42	101
Total	569	455	1024

TABLE VII. COMPARATIVE EXPERIMENTS ON ATTENTIONAL MECHANISMS.

Arithmetic	Precision(%)	Recall(%)	mAP@0.5(%)
YOLOv7	86.9	90.0	88.2
YOLOv7_m	**98.1**	**97.7**	**97.8**
YOLOv8_m	92.4	92.3	90.1

On the critical defect dataset, the YOLOv7_m model proposed in this study demonstrates a detection accuracy of 98.1%, surpassing the original algorithm by 11.2 percentage points. This suggests that the enhanced model offers notable improvements in performance for critical defect detection, particularly in terms of accuracy.

The precision results are visually depicted in Fig. 15 for further comparison.

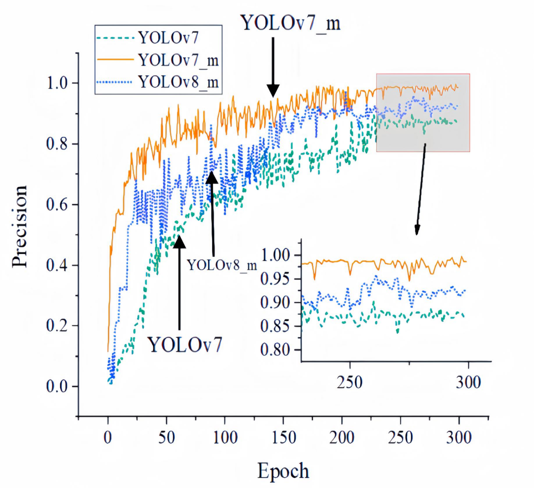


Fig. 15. Comparison chart of precision results.

Further observation of Fig. 15 reveals that the model tends to converge after 300 iterations when detecting critical defects on the critical defects dataset mentioned earlier due to fewer critical defects on the connection buckle, and the final detection rate of YOLOv7_m can be stably maintained at around 98%, with a faster convergence speed. Next, some key datasets were extracted for visualization to verify the actual effect of the model, as shown in Fig. 16 below.

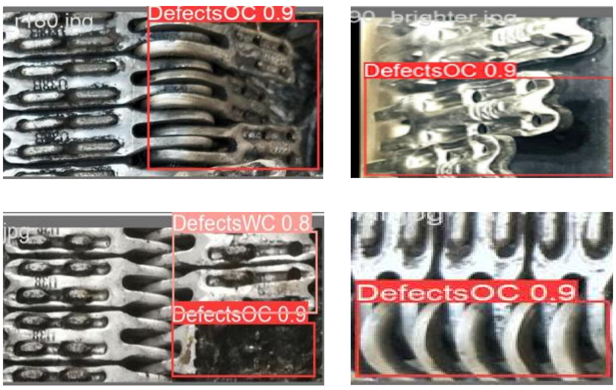


Fig. 16. Visualization of critical defect detection results.

The enhanced algorithm discussed in this study demonstrates a consistent detection accuracy exceeding 90% for critical defects. Furthermore, it reliably identifies connecting buckles with potential defects and exhibits strong generalization capabilities.

In brief, the enhanced YOLOv7 model demonstrates superior detection capabilities across all defect categories, particularly excelling in identifying critical defect types. Additionally, it exhibits faster convergence rates and enhanced overall detection performance, effectively aligning with the practical demands of the industry.

V. Conclusion and Future Work

Connection buckles are widely used in joining underground transportation belts, and it is essential to accurately identify any defects they may have. Despite the importance of this issue, there is a lack of research on it at both national and global levels. This study aims to address this gap by proposing an improved defect detection model and outlining its advantages.

The enhanced YOLOv7_m model demonstrates strong performance in detecting both complex and critical defects in connection buckles. This model showcases clear detection superiority compared to the original network model and prevailing defect detection models currently in use. Specifically:

1. The enhanced YOLOv7_m model demonstrates significant efficacy in identifying various types of complex defects in connection buckles. This model exhibits superior capabilities in accurately recognizing and detecting various types of defects with heightened efficiency, thereby largely fulfilling the industry's overarching criteria for connecting buckle defect detection.

2. The enhanced model demonstrates improved efficacy and increased precision in identifying critical defects in connecting buckles. This outcome serves to validate the viability of the enhanced model and establishes a theoretical and technological foundation for the future implementation of defect detection in connection buckles within industrial settings.

In future work, we will consider incorporating self-learning, semi-supervised, or unsupervised models based on the enhanced model to decrease the need for manual labeling, increase automation, and facilitate the application of the improved model to other types of connection buckle defect detection.

References

[1] Peng Zhang, Shaochuan Xu, and Minghao Ma, "Research and Design of Belt Deviation Detection System Based on Single-side Rollers," Engineering Letters, vol. 32, no. 4, pp871-879, 2024

[2] K. Chaabane, et al. Integrated imperfect multimission selective maintenance and repairpersons assignment problem, Reliab. Eng. Syst. Saf. 199 (2020). https://doi.org/10.1016/j.ress.2020.106895

[3] S. Jocelyn, et al. Application of logical analysis of data to machinery related accident prevention based on scarce data, Reliab. Eng. Syst. Saf. 159 (2017) 223–236. https://doi.org/10.1016/j.ress.2016.11.015

[4] H. J. Pan. Research on online skeleton detection and fault identification of large continuous conveyor belts[D].ShenYang: Dongbei University, 2012,5-6.

[5] Y. Pang, G. Lodewijks. A novel embedded conductive detection syste -m for intelligent conveyor belt monitoring, in 2006 IEEE Internation al conference on service operations and logistics, and informatics, Sha nghai, 2006, pp. 803–808. https://doi.org/10.1109/soli.2006.328958

[6] G. Ross, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of IEEE conference on computer vision and pattern recognition (CVPR), 2014: pp. 580-587. https://doi.org/10.1109/cvpr.2014.81

[7] Liu. W, et al. SSD: Single Shot MultiBox Detector, vol. 9905.2016, pp 21-37

[8] X. Zhou, D. Wang, and P. Kra¨henbu¨hl. Objects as Points, 2019.

[9] J. Redmon, et al. You Only Look Once: Unified, Real-Time Object Detection. 2016. 779-788. https://doi.org/10.1109/cvpr.2016.91

[10] X. Q. Guo, X. H. Liu, P. Gardoni, A. Glowacz, et al. Machine vision based damage detection for conveyor belt safety using Fusion knowledge distillation; Alexandria Engineering Journal (2023) 71, 161–172. https://doi.org/10.1016/j.aej.2023.03.034

[11] M. Y. Chen. Research and realization of conveyor belt defect detectio -n system based on machine vision[D].Yanshan: Yanshan University, 2022

[12] Tianzhen Liu, Yingchun Yuan, Guifa Teng, and Xi Meng, "Improved Deep Convolutional Neural Network-Based Method for Detecting Winter Jujube Fruit in Orchards," Engineering Letters, vol. 32, no. 3, pp569-578, 2024
[13] Yongzhong Fu, Liang Qiu, Xiao Kong, and Haifu Xu, "Deep Learning-Based Online Surface Defect Detection Method for Door Trim Panel," Engineering Letters, vol. 32, no. 5, pp939-948, 2024
[14] Yu Zhang, Ming Ma, Zhongxiang Wang, Jing Li, and Yan Sun, "POD-YOLO Object Detection Model Based on Bi-directional Dynamic Cross-level Pyramid Network," Engineering Letters, vol. 32, no. 5, pp995-1003, 2024
[15] Shiqin Li, and Weisheng Liu, "Small Target Detection Model in Aerial Images Based on YOLOv7X+," Engineering Letters, vol. 32, no. 2, pp436-443, 2024
[16] H. Lu, Y. Li, M. Chen, H. Kim, S. Serikawa, Brain intelligence: go beyond artificial intelligence, Mobile Netw. Appl. 23 (2) (2018) 368–375. https://doi.org/10.1007/s11036-017-0932-8
[17] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, J. Zhang. Edge Intelligence: Paving the last mile of artificial intelligence with edge computing, P I -EEE 107 (8) (2019) 1738–1762. https://doi.org/10.1109/jproc.2019.2918951
[18] Xin Wen, Yi Yao, Ying Cai, Zixing Zhao, Tianjiao Chen, Ziyu Zeng, Zhen Tang, "A Lightweight ST-YOLO Based Model for Detection of Tea Bud in Unstructured Natural Environments," IAENG International Journal of Applied Mathematics, vol. 54, no. 3, pp342-349, 2024
[19] X. Y. Dai, et al. Dynamic Head: Unifying Object Detection Heads with Attentions; Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR), Electr Network, F Jun 19-25, 2021[C].2021. https://doi.org/10.1109/cvpr46437.2021.00729
[20] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks [J].IEEE transactions on pattern analysis and machine intelligence, 2020, 42(8): 2011-23. https://doi.org/10.1109/cvpr.2018.00745
[21] J. Redmon, A. Farhadi. Yolo9000: Better, faster, stronger, in IEEE conference on computer vision & pattern recognition, 2017, pp. 6517–6525. https://doi.org/10.48550/arXiv.1612.08242
[22] A. F. Joseph Redmon. Yolov3: An incremental improvement. https://doi.org/10.48550/arXiv.1804.02767
[23] Bochkovskiy, C. Y. Wang. Yolov4: Optimal speed and accuracy of object detection. https://doi.org/10.48550/arXiv.2004.10934
[24] W. S. Mseddi, R. Ghali, M. Jmal, R. Attia. Fire detection and segmentation using yolov5 and u-net, in 2021 29th European Signal Processing Conference (EUSIPCO), 2021, pp. 741–745. https://doi.org/10.23919/eusipco54536.2021.9616026
[25] Q. Chen, Y. Wang, T. Yang, X. Zhang, J. Cheng, J. Sun. You only look at one-level feature, in 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR), 2021, pp. 13034–13043. https://doi.org/10.1109/cvpr46437.2021.01284
[26] C. Y. Wang, A. Bochkovskiy, H. Liao, Yolov7: Trainable bag-of-free -bies sets new state-of-the-art for real-time object detectors. https://doi.org/10.1109/cvpr52729.2023.00721
[27] Siliang Ma, Y. Xu. MPDIoU: A Loss for Efficient and Accurate Boun -ding Box Regression. 2011. https://doi.org/10.48550/arXiv.2307.07662
[28] H. X. Fu , G. Q. Song, Y. C Wang. Improved YOLOv4 marine target detection combined with CBAM[J].Symmetry, 2021, 13(4): 623. https://doi.org/10.3390/sym13040623
[29] Q. B. Hou, D. Q. Zhou, J. S. Feng. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 13713-13722. https://doi.org/10.48550/arXiv.2103.02907
[30] Q. L. Zhang, Y. B. Yang. SA-Net: Shuffle Attention for Deep Convolutional Neural Networks[Z].Secondary. 2021: 2235-2239. https://doi.org/10.1109/icassp39728.2021.9414568
[31] S. Q. Ren, K. M. He, et al. Faster R-CNN: Towards real-time object.