

Research Article

HRM: An Intelligent Helmet Recognition Model in Complex Scenes

Panbo He ¹, Chunxue Wu ¹, Rami Yared,² and Yuanhao Ma³

¹*School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*

²*School of Computer Science, Arab International University, Damascus 999099, Syria*

³*College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Chunxue Wu; wcx@usst.edu.cn

Received 14 December 2021; Revised 5 January 2022; Accepted 30 April 2022; Published 4 July 2022

Academic Editor: Antonio Scarfone

Copyright © 2022 Panbo He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents an intelligent helmet recognition model in complex scenes based on YOLOv5. Firstly, in construction site projects, consider that the photograph which needs to be identified has numerous problems. For example, helmet's pixels are too tiny to detect, or a large number of workers makes helmets appear densely. A SE-Net channel attention module is added in different parts of the network layer of the model, so that the improved model can pay more attention to the global variables and increase the detection performance of small target information and dense target information. In addition, this paper constructs a helmet data set based on projects and adds training samples of dense targets and long-range small targets. Finally, the modified mosaic data enhancement reduces the influence of redundant background on the model and improves the recognition accuracy of the tiny target. The experimental results show that in the project, the average accuracy of helmet detection reaches 92.82%. Compared with SSD, YOLOv3, and YOLOv5, the average accuracy of this algorithm is improved by 6.89%, 8.28%, and 2.44% and has strong generalization ability in dense scenes and small target scenes, which meets the accuracy requirements of helmet wearing detection in engineering applications.

1. Introduction

The working environment of the construction site is complex and full of danger, and the workers are very vulnerable to injury in the process of work. Therefore, it is essential for workers on the construction site to wear safety protection equipment. In actual work, manual supervision is usually used to judge whether workers wear safety helmets. Therefore, there are problems such as wide operation ranges of constructors and failure to track and manage all workers in time at the construction site [1–5]. Automatic monitoring based on video stream is conducive to on-site real-time monitoring.

Traditional target detection usually adopts manual feature selection and designs and trains classifiers according to specific detection objects. This method has strong subjectivity, complex design process, and poor generalization ability and has great limitations in engineering application.

Therefore, many researchers are committed to the possibility of combining a domain model with an actual construction site scene. In 2015, Redmon et al. [6] proposed a one-stage detection model YOLO (You Only Look Once), which abstracted the detection task as a regression problem for the first time. In 2016, Liu et al. [7] proposed an SSD (single shot multibox detector) detection algorithm and introduced a multiscale detection method, which can effectively detect groups of small targets. In 2018, Redmon and Farhadi [8] further proposed You Only Look Once version 3. The model uses the Feature Pyramid Network (FPN) method to integrate three feature maps of different sizes for detection tasks, which significantly improves the detection effect of small-size targets. In 2020, Bochkovskiy et al. [9] proposed YOLOv4 (You Only Look Once version 4). The model selects CSP (Cross Stage Partial) darknet-53 as the backbone network and uses PAnet (path aggregation network) method to replace the FPN algorithm in the YOLOv3 network,

which greatly improves the detection accuracy of the model. In 2020, Usman et al. [10] proposed YOLOv5 (You Only Look Once version 5), which adds a focus structure to the backbone network to achieve the best balance between speed and accuracy.

Nowadays, a large number of scholars have carried out a series of related research on helmet detection. In 2016, Rubaiyat et al. and Silva et al. [11–14] combined the frequency domain information in the image with the Histogram of Orientation Gradients (HOG) algorithm to detect the human body and then used the Circle Hough Transform (CHT) to detect the helmet. In 2017, Li [15] uses a visual background extractor (Vibe) algorithm to locate the human body, then used the convex algorithm to detect the head, and finally combined the HOG algorithm and SVM to realize the helmet wearing detection. In 2018, Hao and Jza [16] proposed a hybrid descriptor composed of Local Binary Patterns (LBP), Color Histograms (CH), and Hu Moment Invariants (HMI) to extract helmet features and then constructed Hierarchical Support Vector Machine (H-SVM) to classify helmets. Due to the complex environment, the detection accuracy of helmet wearing is low at this stage, which does not meet the monitoring requirements of the actual production environment.

In this paper, we use two kinds of targets, construction workers wearing safety helmets and construction workers without safety helmets, which are taken as the detection tasks. A total of 26491 pictures are collected from the online collection and operation site for preprocessing, the safety helmet detection data set is constructed, and the YOLOv5 network model is selected as the basic model. The main contributions of this paper are as follows:

- (1) A helmet data set based on an engineering project is constructed, and the training samples of dense targets and long-range small targets are added. The data-enhanced mosaic is optimized to reduce useless boundary information and improve the robustness of the model
- (2) After the C3 module of the YOLOv5 backbone network and neck layer, the SE-Net channel attention module is introduced to collect global information in the feature extraction stage to improve the effect of small target helmet detection
- (3) By introducing pixel IoU, the problem of an inaccurate positioning frame is solved, and the sensitivity of loss value to the processing dimension frame is improved

The experimental results show that the mean average precision (mAP) of the optimized model algorithm is significantly improved to meet the detection requirements in the construction scene.

2. Materials and Methods

2.1. Principle of the YOLOv5 Algorithm. The YOLOv5 network structure is divided into input, backbone, neck, and prediction according to the processing stages, as shown in

Figure 1. The input part completes basic processing tasks such as data enhancement, adaptive picture scaling, and anchor box calculation. As the backbone network, the backbone mainly uses CSP structure to extract the main information in the input samples for subsequent stages. The neck part uses FPN and pan structure and uses the information extracted from the backbone part to strengthen feature fusion. The prediction section makes a prediction and calculates the GIOU_Loss value such as loss (see Figure 1).

2.2. SE-Net Attention Module. An attention mechanism comes from the way the human brain processes visual information. By rapidly observing the global information of the image, human beings find out the candidate area that needs to be focused, that is, the location of the focus, and will focus on this area to extract more detailed information of the target [17]. Because of its powerful and effective expression, it has been widely used in deep learning, especially in deep-seated high-performance networks [18] (see Figure 2).

Firstly, for the feature map with channel number C , each channel contains different feature information. During feature extraction, the convolution layer mainly calculates the feature information of adjacent positions of each feature map without considering the correlation mapping between channel information [19]. Because the image resolution of small target helmet is low and the pixel value and channel characteristic information are limited, it is necessary to strengthen the training of relevant channel characteristic information in the training process. References [20–26] fully demonstrate that the attention module of the SE-Net channel can optimize the learning of specific categories of feature information in a deep-seated network. And the module is also a plug and play module, which is usually applied behind the convolution module. Therefore, we add the SE-Net channel attention module behind the C3 module of the neck detection layer of the YOLOv5 network; that is, after the detection layers of different scales, we add the SE-Net module, respectively, by establishing feature mapping relationship between channels; the network makes full use of these global information and gives higher weight to the channel feature information of small targets. So as to better fit the relevant characteristic information between small target channels, ignore and suppress useless information, and finally make the model focus on training the specific category of small target helmet.

2.3. Loss Function Improved. The types of loss functions commonly used in YOLO series algorithms are GIOU, DIOU, and CIOU. The evolution from GIOU to CIOU makes the regression loss more accurate and the target frame regression more stable. However, it is found that the above three types of loss functions will cause the problem of an inaccurate positioning frame for targets with a high aspect ratio and dense targets; in order to solve this problem, the pixel IoU (PIoU) function [27] is introduced.

By introducing a rotation parameter, the loss function can frame the target more compactly. In order to accurately calculate the target intersection union ratio, the loss function calculates the target IoU by pixel counting, which makes the

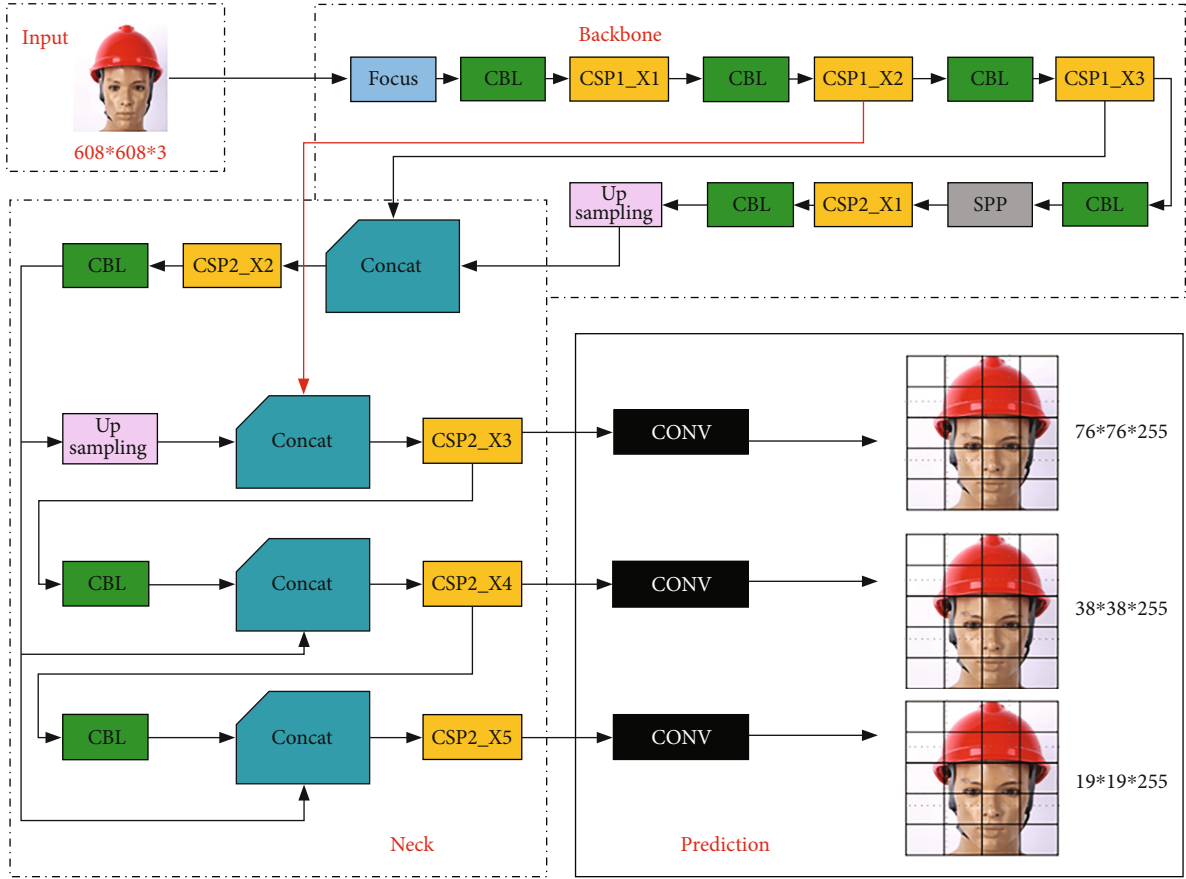
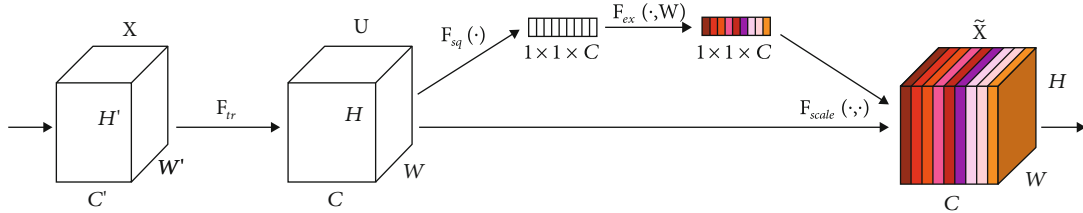


FIGURE 1: Basic framework of the YOLOv5 network structure, including input, backbone, neck, and prediction.


 FIGURE 2: SE-Net structure diagram. F_{sq} is the squeeze, F_{ex} is the excitation, and F_{scale} is the scale.

loss value sensitive to the size, position, and rotation of the processing label box. The calculation formula of PIoU loss is

$$L_{PIoU} = -\left(\frac{1}{|M|}\right) \sum_{(b,b') \in M} \ln F_{PIoU}(b, b'). \quad (1)$$

In the formula, M is the set of all positive samples and $|M|$ is the number of positive samples. b is the ground truth, and b' is the target box. The calculation formula of the PIoU function is

$$F_{PIoU}(b, b') = \frac{S_{b \cap b'}}{S_{b \cup b'}}. \quad (2)$$

In the formula, $S_{b \cap b'}$ and $S_{b \cup b'}$, respectively, represent the number of pixels in the intersection of target b and target frame b' and the number of pixels in the union after being processed by the loss function kernel function.

2.4. Network Layer Adds SE-Net. In the small target detection task, with the gradual increase in the number of network layers, the collectable small target feature information is also gradually weakened, so it is easy to cause false detection and missed detection of small targets by the network model. The SE attention module itself uses the global average pool and other frequency components to enhance the features in the feature map, so that the network can strengthen the easy view learning of the target features in the training process. However, at this stage, no research

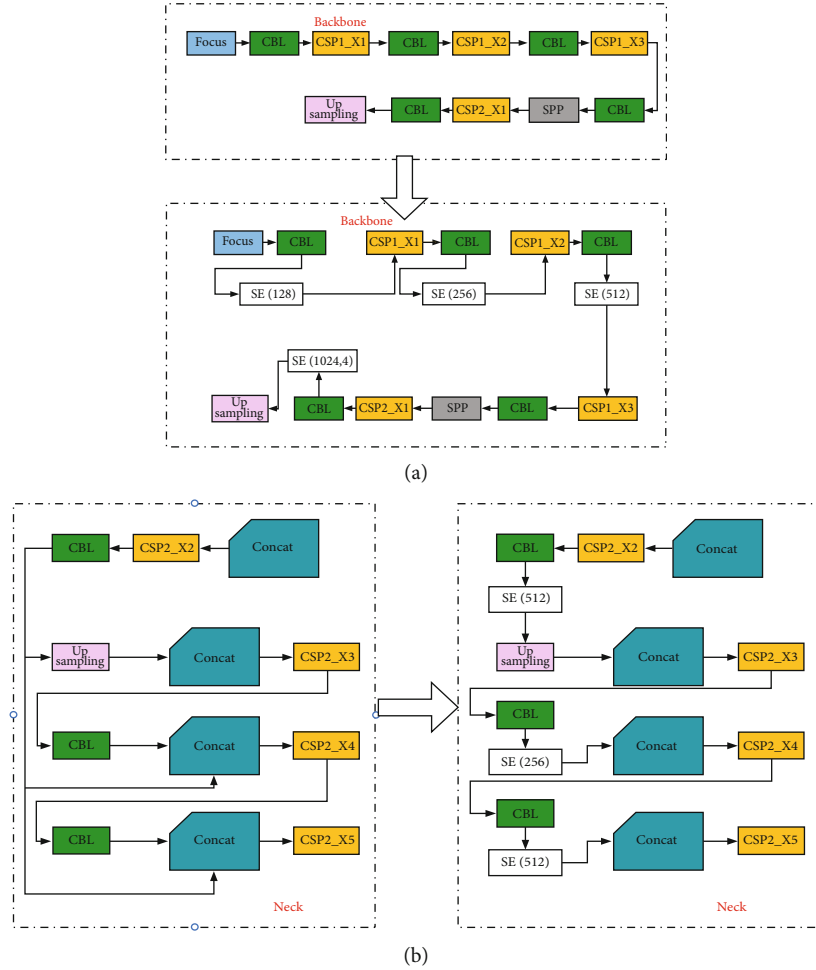


FIGURE 3: (a) Adding SE-Net in the YOLOv5 backbone layer. (b) Adding SE-Net in the YOLOv5 neck part. SE (128) means that this layer's input size is 128.

shows that the SE attention mechanism module should integrate which location of the network can effectively improve the detection efficiency.

Inspired by References [28–33], this paper integrates the SE attention module into different positions of the network model and studies the detection results. According to the structure of the YOLOv5 network model, the SE attention module is fused in the backbone network and neck of YOLOv5. Since the SE module performs feature enhancement in important channels and spatial locations, the SE attention module is fused to each feature fusion area in the above two parts, respectively, resulting in two new network models based on the YOLOv5 algorithm: SE-backbone and SE-neck. Figure 3 shows the specific location of the SE attention module fusion network.

The experimental comparison of adding the SE attention module in two different positions is shown in Table 1. It can be concluded that after the SE-Net module is integrated into the backbone network of YOLOv5, the detection accuracy of small targets is significantly improved, which can effectively improve the detection effect of the network on small target objects, and the mean average precision (mAP) is increased by 3.3 percentage points. After the SE-Net module is inte-

TABLE 1: Ablation results.

Name	P%	R%	mAP%
YOLOv5	91.613	91.494	89.813
SE-backbone	93.433	94.087	91.104
SE-neck	94.37	93.858	90.729
SE-neck+mosaic	95.023	93.522	91.761
SE-backbone+mosaic	95.518	94.776	92.827

grated into the neck module of YOLOv5, the performance of the model is not improved; on the contrary, the map is also reduced. This paper holds that the reason why the SE-Net module is fused to different positions in the model to produce different experimental results is that although the semantic information in the backbone network is not rich, it still implies the texture information and contour information that are easy to be ignored in the middle and low layers of the target. The fusion of the SE-Net module in the backbone network can better fuse the spatial features and channel features of small targets in the feature map, so as to enhance the feature information. In the deeper neck and

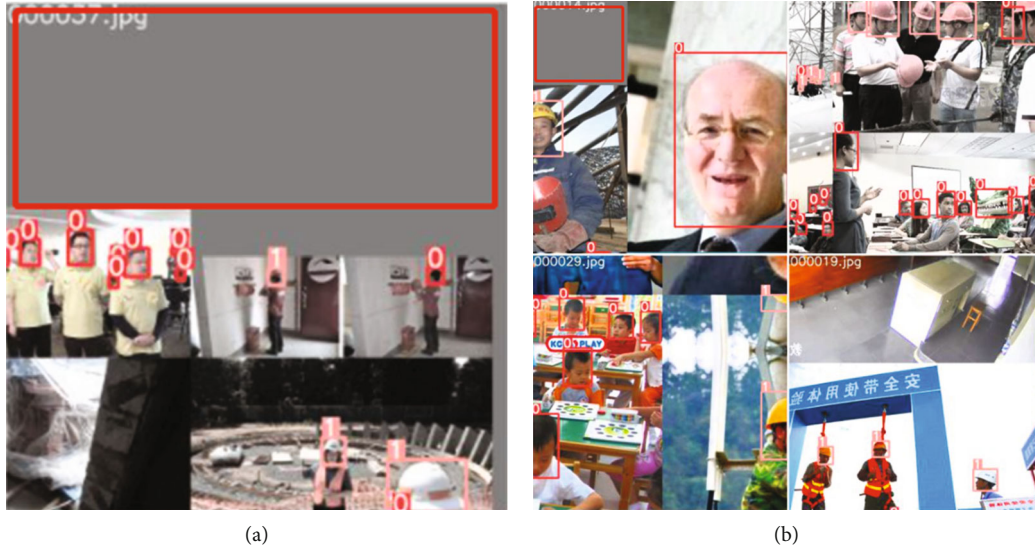


FIGURE 4: Compare two mosaics: (a) having huge blank filling; (b) improved picture having much small area of blank filling.



FIGURE 5: Improved mosaic result: a blocked helmet is detected in (b).

prediction module of the network, because its feature map has richer semantic features, a smaller feature map, and a huge receptive field, the SE net module has difficulty distinguishing important spatial features and channel features.

2.5. Mosaic Improvements. Mosaic, the data enhancement method in the YOLOv5 algorithm, is very practical. The basic principle is to randomly select four pictures: first cut them randomly, then splice them on one picture clockwise, and finally scale them to the set input size, which is introduced into the model as a new sample. This enriches the background of targets, increases the number of small targets, and achieves the balance between targets of different scales [34].

The data set in this paper is divided into two categories: helmet and head. The total amount of data is also relatively small compared with the coco data set, with only more than 20000 pieces. Due to the particularity of safety helmet pictures on the construction site, the targets to be identified are often not in the center of the picture. Random cutting has a high probability to cut the targets, so that there is only background in the input picture samples, and this will make the spliced pictures have more black-and-white boundaries,

which will lead to a large amount of useless feature information in model training and affect the convergence speed of the model [35].

Different from the coco data set, most of the images in the data set in this paper come from the frames taken from the video stream, and the size of the images is consistent with that of the pixels. Therefore, according to the characteristics of the data set, this paper improves the mosaic data enhancement method. Firstly, the number of mosaics is changed from 4 to 16, and then, the judgment conditions are added to ensure as little useless area as possible, and these filled black-and-white boundaries are at the edge of the image. As shown in the figure, the optimized original mosaic has a large area of blank filling, and the improved picture has a small area of blank filling (see Figure 4).

The performance comparison of the two methods is shown in Table 1. It can be seen that the accuracy of the improved model remains basically unchanged, but the recall and average accuracy of the improved model are increased by 2.08% and 1.71%, respectively.

The comparison of test results is shown in Figure 5. Figure 5(a) is the original model test diagram, and Figure 5(b) is the improved model test effect diagram.

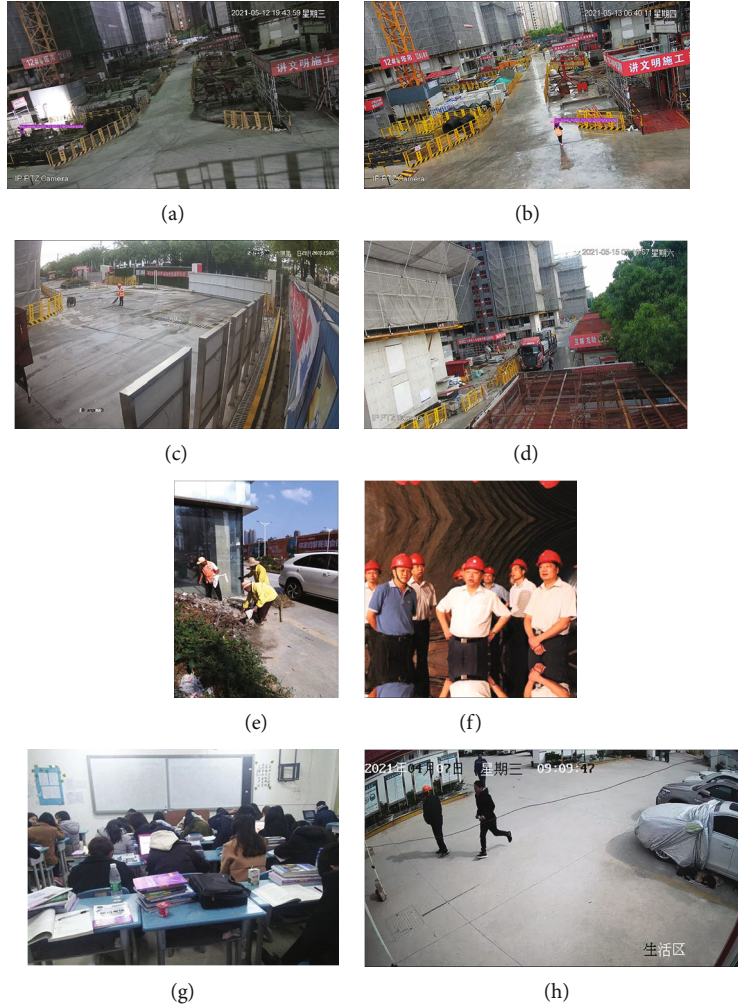


FIGURE 6: Helmet sample: (a) dark light; (b) small targets; (c, d) real environment; (e) interference term; (f) dense term; (g) dense interference term; (h) helmet and without helmet.

Figure 5 shows the helmet image of densely occluded small targets. It can be seen from the observation that compared with the original model, the improved model detects the targets wearing safety helmets. It shows that the improved model has strong robustness in these scenarios (see Figure 5).

3. Results and Discussion

3.1. Data and Training Setting. The results and discussion may be presented separately or in one combined section and may optionally be divided into headed subsections.

The main source of open-source helmet data set on the network is SHWD (Safety Helmet Wearing Dataset) [36], and most scenes are monitoring images of students in class. The data set deviates greatly from the standard site scene data set. Therefore, this paper expands the data set. The images mainly come from the site video stream frame cutting and handheld device shooting. The data collected include two types: workers wearing safety helmets and workers not wearing safety helmets in different site environments. In order to increase the diversity and robustness of

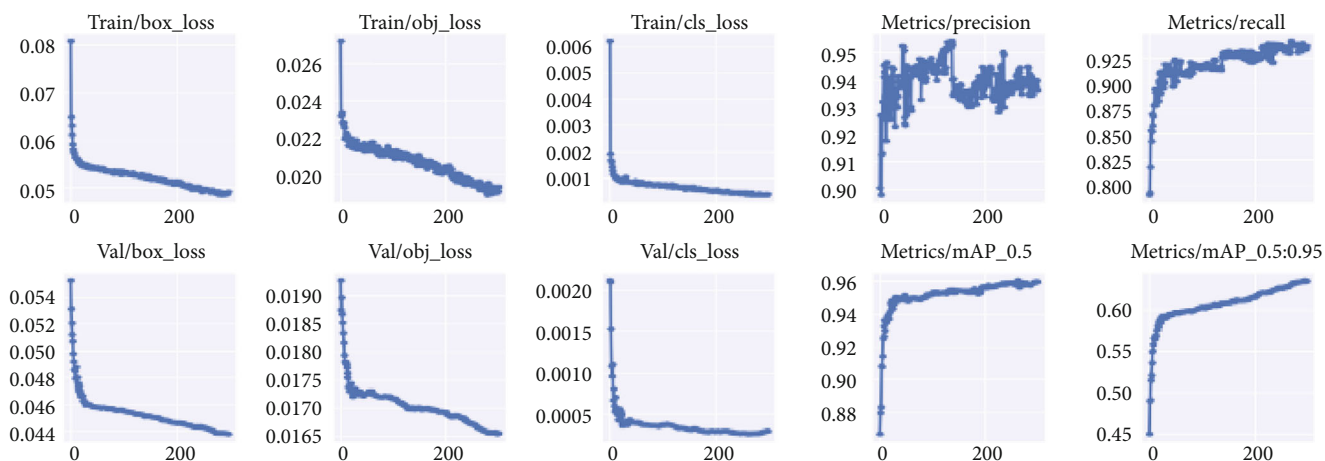
TABLE 2: Data set.

Type	Picture number	Box number
Hat	13428	45969
Head	14815	220453
Sum	26491	266422

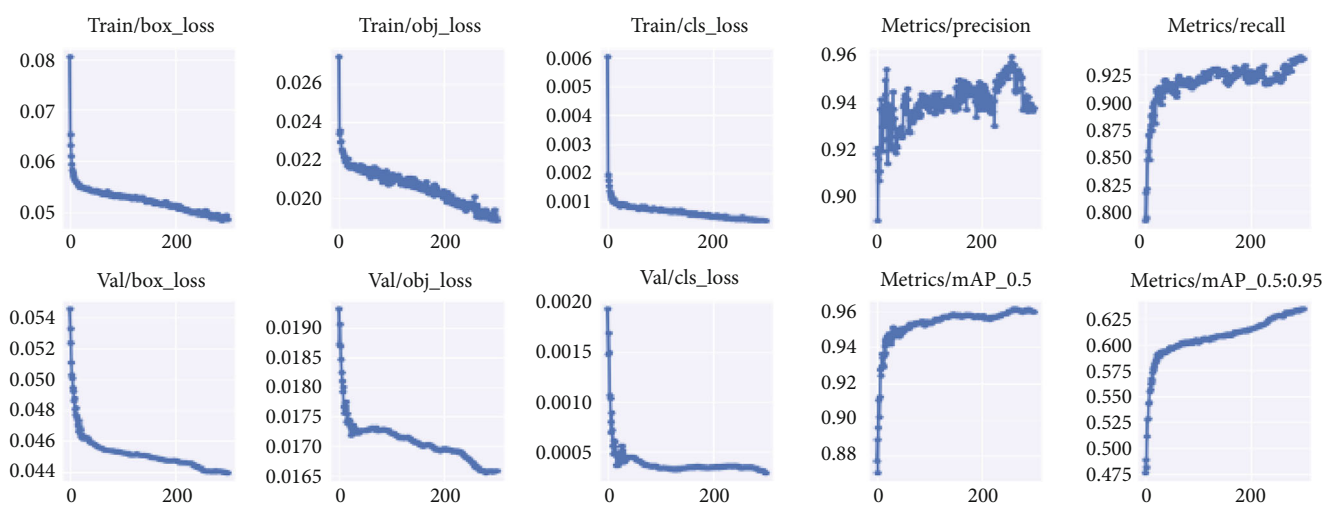
the training set, some baseball caps, hard hats, and other data are added to increase the generalization ability of the model (see Figure 6).

The data set finally obtained in this paper has a total of 26491 pictures, of which the specific information of the targets wearing and not wearing helmets in this data set is shown in Table 2. The data set contains a variety of construction scenes, which can fully reflect the real construction scenes.

This paper is based on Linux, Ubuntu 18.04, and GeForce 3090 with 24 GPU video memory, the CUDA version is 11.2, and pytorch is selected as the framework. The data set is randomly divided into a training set, verification set, and test set according to the ratio of 8 : 1 : 1, batch size = 32, epoch = 300, and LR = 0.01.



(a)



—●— Results

(b)

FIGURE 7: Model training loss diagram: (a) origin YOLOv5; (b) our model.

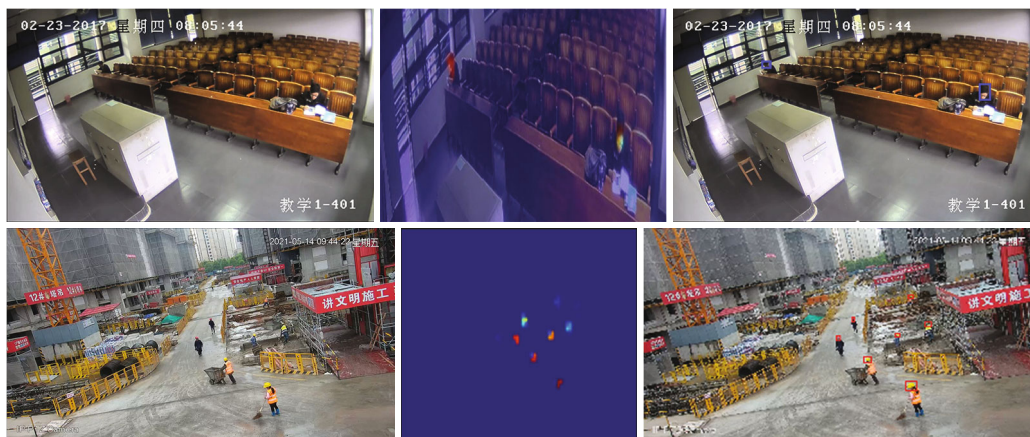


FIGURE 8: Characteristic thermodynamic diagram. Our model can capture small and dense targets.



FIGURE 9: Detection pictures. Use for different work environments. Good at detecting small and dense targets.

TABLE 3: Comparison of results of multiple detection models.

Name	mAP (%)	Speed (s)	Params (M)
SSD	85.39	0.11/picture	570.4
YOLOv3	84	0.06/picture	340.7
YOLOv5	89.84	0.03/picture	269.8
Ours	92.82	0.04/picture	266.1

3.2. Experimental Results. The model training loss diagram is compared as follows (see Figure 7), and the thermal diagram of our model is printed. It can be seen that the optimized model can put forward the characteristics of small goals and secret script goals (see Figure 8).

It can be seen from the figure that our model can capture the long-distance recognition targets in the image according to the characteristics of the pictures taken by the camera. The midpicture describes the feature map of the original figure. It can be concluded that our model well distinguishes between foreground and background and increases the feature capture of small and microtargets.

From the measured results of some site pictures, it can be seen that the optimized model in this paper improves the detection effect of small targets and dense targets and is suitable for many different actual scenes (see Figure 9).

At the same time, this paper also compares the optimized model with the mainstream target detection model. The results are shown in Table 3.

The experimental results show that this algorithm can effectively improve the detection accuracy of safety helmets and construction workers without safety helmets. The average detection accuracy of this algorithm for construction workers wearing safety helmets is 94.77%, and the mean average precision (mAP) is 92.82%, which is much higher than those of the YOLOv5 and SSD. Compared with SSD, YOLOv3, and original YOLOv5, this algorithm has a certain

improvement in accuracy and mAP. This shows that this algorithm performs well in the accuracy of helmet wearing detection and can meet the accuracy requirements of helmet detection in complex working environment.

4. Conclusions

In order to improve the shortcomings of existing helmet algorithms in dense target and small target scenarios, a YOLOv5 algorithm is proposed in this paper. Through the SE-Net attention module and improved data enhancement method, the detection effect of the model for small target helmets is improved, the loss function is optimized, and the generalization ability of the model in dense scenes is increased. Through comparative experiments, compared with the original YOLOv5 model, the improved model reduces the missing detection of helmet and improves the classification confidence score. The experimental results show that this algorithm can obtain better detection accuracy and basically meet the accuracy requirements of helmet wearing detection in complex construction scenes.

Data Availability

The data set includes corresponding pictures and labels, and the file size is 2 GB. Data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Nos. U1836208, 61811530332, and 61811540410).

References

- [1] J. Li, Y. F. Bai, and J. L. He, "Design and implementation of embedded video surveillance system based on human detection," *Applied Mechanics & Materials*, vol. 513-517, pp. 4535-4538, 2014.
- [2] Z. Wang, T. Li, N. Xiong, and Y. Pan, "Erratum to: a novel dynamic network data replication scheme based on historical access record and proactive deletion," *Journal of Supercomputing*, vol. 62, no. 1, p. 616, 2012.
- [3] W. Guo, N. Xiong, A. V. Vasilakos, G. Chen, and H. Cheng, "Multi-source temporal data aggregation in wireless sensor networks," *Wireless Personal Communications*, vol. 56, no. 3, pp. 359-370, 2011.
- [4] J. Yin, W. Lo, S. Deng, Y. Li, Z. Wu, and N. Xiong, "Colbar: a collaborative location-based regularization framework for QoS prediction," *Information Sciences*, vol. 265, pp. 68-84, 2014.
- [5] Z. Wan, N. Xiong, N. Ghani, A. V. Vasilakos, and L. Zhou, "Adaptive unequal protection for wireless video transmission over IEEE 802.11e networks," *Multimedia Tools & Applications*, vol. 72, no. 1, pp. 541-571, 2014.
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 2016.
- [7] W. Liu, D. Anguelov, D. Erhan et al., *SSD: single shot multi box detector*, Springer, Cham, 2016.
- [8] J. Redmon and A. Farhadi, "YOLOv3: an incremental improvement," 2018, <https://arxiv.org/abs/1804.02767>.
- [9] A. Bochkovskiy, C. Y. Wang, and H. Liao, "YOLOv4: optimal speed and accuracy of object detection," 2020, <https://arxiv.org/abs/2004.10934>.
- [10] S. M. Usman, G. R. Jocher, S. T. Dye, W. F. McDonough, and J. G. Learned, "AGM2015: Antineutrino Global Map 2015," *Scientific reports*, vol. 5, no. 1, article ???, 2015.
- [11] A. H. Rubaiyat, T. T. Toma, M. Kalantari-Khandani et al., "Automatic detection of helmet uses for construction safety," in *IEEE/WIC/ACM, International Conference on Web Intelligence Workshops*, Omaha, NE, USA, 2017.
- [12] R. Silva, K. Aires, and R. Veras, "Helmet detection on motorcyclists using image descriptors and classifiers," *2014 27th SIB-GRAPI Conference on Graphics, Patterns and Images*, pp. , 2014141-148, 2014.
- [13] Y. Qu and N. Xiong, "RFH: a resilient, fault-tolerant and high-efficient replication algorithm for distributed cloud storage," in *2012 41st International Conference on Parallel Processing*, Pittsburgh, PA, USA, 2012.
- [14] M. Wu, L. Tan, and N. Xiong, "A structure fidelity approach for big data collection in wireless sensor networks," *Sensors*, vol. 15, no. 1, pp. 248-273, 2015.
- [15] Q. R. Li, *A Research and Implementation of Safety-Helmet Video Detection System Based on Human Body Recognition*, University of Electronic Science and Technology of China, 2017.
- [16] W. A. Hao and B. Jza, "An intelligent vision-based approach for helmet identification for work safety," *Computers in Industry*, vol. 100, pp. 267-277, 2018.
- [17] X. Tang, B. Zhong, J. Peng, B. Hao, and J. Li, "Multi-scale channel importance sorting and spatial attention mechanism for retinal vessels segmentation," *Applied Soft Computing*, vol. 93, p. 106353, 2020.
- [18] G. Zhang, J. Yang, Y. Zheng, Y. Wang, Y. Wu, and S. Chen, "Hybrid-attention guided network with multiple resolution features for person re-identification," *Information Sciences*, vol. 578, pp. 525-538, 2021.
- [19] W. H. Chen, J. He, and G. Liu, "Convolutional neural network with attention mechanism for hyperspectral image classification," *Laser & Optoelectronics Progress*, vol. 59, article 0411002, 2022.
- [20] Z. Y. Zou, S. Y. Gai, F. P. Da, and L. L., "An occluded pedestrian detection algorithm based on attention mechanism," *Acta Optica Sinica*, vol. 41, pp. 149-157, 2021.
- [21] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 99, 2017.
- [22] R. He, N. Xiong, L. T. Yang, and J. H. Park, "Using multi-modal semantic association rules to fuse keywords and visual features automatically for web image retrieval," *Information Fusion*, vol. 12, no. 3, pp. 223-230, 2011.
- [23] Q. Zhang, C. Zhou, Y. C. Tian, N. Xiong, Y. Qin, and B. Hu, "A fuzzy probability Bayesian network approach for dynamic cybersecurity risk assessment in industrial control systems," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 2497-2506, 2017.
- [24] S. Huang, A. Liu, S. Zhang, T. Wang, and N. N. Xiong, "BD-VTE: a novel baseline data based verifiable trust evaluation scheme for smart network systems," *IEEE Transactions on Network Science and Engineering*, vol. 8, pp. 2087-2105, 2021.
- [25] H. Li, J. Liu, K. Wu, Z. Yang, R. W. Liu, and N. Xiong, "Spatio-temporal vessel trajectory clustering based on data mapping and density," *IEEE Access*, vol. 6, pp. 58939-58954, 2018.
- [26] Y. Yao, N. Xiong, J. H. Park, L. Ma, and J. Liu, "Privacy-preserving max/min query in two-tiered wireless sensor networks," *Computers & Mathematics with Applications*, vol. 65, no. 9, pp. 1318-1325, 2013.
- [27] Z. Cao, M. Shao, L. Xu, S. Mu, and H. Qu, "MaskHunter: real-time object detection of face masks during the COVID-19 pandemic," *IET Image Processing*, vol. 14, p. 10223, 2020.
- [28] Z. Zhang, X. Zhang, C. Peng, X. Xue, and J. Sun, "ExFuse: enhancing feature fusion for semantic segmentation," in *European Conference on Computer Vision Springer*, Cham, 2018.
- [29] S. Chaib, H. Liu, Y. Gu, and H. Yao, "Deep feature fusion for VHR remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4775-4784, 2017.
- [30] G. Ghiasi, T. Y. Lin, and Q. V. Le, "NAS-FPN: learning scalable feature pyramid architecture for object detection," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019.
- [31] F. Xia, R. Hao, J. Li, N. Xiong, L. T. Yang, and Y. Zhang, "Adaptive GTS allocation in IEEE 802.15.4 for real-time wireless sensor networks," *Journal of Systems Architecture*, vol. 59, no. 10, pp. 1231-1242, 2013.

- [32] K. Gao, F. Han, P. Dong, N. Xiong, and R. Du, "Connected vehicle as a mobile sensor for real time queue length at signalized intersections," *Sensors*, vol. 19, no. 9, p. 2059, 2019.
- [33] Y. Lu, S. Wu, Z. Fang, N. Xiong, S. Yoon, and D. S. Park, "Exploring finger vein based personal authentication for secure IoT," *Future Generation Computer Systems*, vol. 77, pp. 149–160, 2017.
- [34] L. Chengyue, Y. Jianmin, and L. Zhixian, "Object detection method based on improved YOLO lightweight network," *Laser & Optoelectronics Progress*, vol. 57, no. 14, p. 141003, 2020.
- [35] A. Gautier, R. Roussel, C. Lange et al., "Effects of genetic susceptibility for type 2 diabetes on the evolution of glucose homeostasis traits before and after diabetes diagnosis," *Diabetes*, vol. 60, no. 10, pp. 2654–2663, 2011.
- [36] D. Peng, Z. Sun, Z. Chen, Z. Cai, L. Xie, and L. Jin, "Detecting heads using feature refine net and cascaded multi-scale architecture," in *2018 24th International Conference on Pattern Recognition (ICPR)*, Beijing, China, 2018.