Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

Research on Improved Algorithm for Safety Helmet Detection Based on YOLOv5

School of Computer and Communication, Hunan Institute of Engineering Xiangtan 411104, China
Tang Zhi-hang
Email: zhtang@hnie.edu.cn
He Yi
Email:926250671@qq.com
Tang Jiale
Email:41423804@qq.com
Zhao Si Meng
Email:2760742243@qq.com

----------**ABSTRACT**----------

In order to improve the accuracy and robustness of helmet wearing detection, the algorithm was further optimized based on YOLOv5, combined with CBAM attention mechanism, BiFPN structure, Focal loss function and other technologies, significantly improving the accuracy and precision of helmet wearing detection and recognition, comprehensively enhancing the level of safety protection in the production process, and meeting the accuracy and real-time requirements of helmet wearing detection. The experimental results show that the improved helmet detection algorithm has a 3.2% increase in mAP compared to the original YOLOv5 algorithm.

Keywords- **Safety helmet Target detection, YOLOv5, Attention mechanism**

Date of Submission: February 03, 2025 Date of Acceptance : February 25, 2025

I. INTRODUCTION

The safety helmet wearing detection technology in chemical safety scenarios, as a key measure to ensure the safety of production personnel, is widely used in high-risk areas such as chemical plants, storage tank areas, and refineries. In these complex and potentially dangerous environments, the correct wearing of a safety helmet is crucial for preventing the consequences of accidents such as chemical spills, fires, and explosions. However, the wearing detection of safety helmets in chemical safety scenarios faces severe challenges such as harsh environmental conditions, complex background interference, dynamic changes in personnel, changes in lighting and shadows. YOLOv5 may face the risk of false or missed detections when dealing with blurry, small targets, or disturbed images in chemical safety scenarios.

In order to improve the accuracy and robustness of safety helmet wearing detection in chemical safety scenarios, the algorithm was further optimized based on YOLOv5, combined with CBAM attention mechanism, BiFPN structure, Focal loss function and other technologies, significantly enhancing the accuracy and precision of safety helmet wearing detection and recognition, comprehensively improving the safety protection level in chemical production processes, and meeting the accuracy and real-time requirements of safety helmet wearing detection in chemical safety scenarios.

In the rapidly developing chemical industry, safety is always a core element of business operations [1]. However, the chemical production environment is complex and changeable, involving high temperature, high pressure, flammable and explosive and other high-risk factors, which makes safety management a very

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

challenging task. Traditional safety management often relies on manual inspection and monitoring, which is not only inefficient, but also susceptible to the influence of human factors, making it difficult to detect and eliminate the hidden dangers of safety accidents in a timely manner. With the rapid development of computer vision and deep learning technology, target detection technology has gradually become an important tool in the field of chemical safety [2]. Target detection technology can realize automatic identification and tracking of specific targets through intelligent analysis of video images, providing a new solution for helmet wearing detection and flame detection in chemical safety scenarios.[3]

Earlier, researchers mainly used image processing techniques, such as color segmentation [4] and edge detection, to identify whether a worker is wearing a helmet or not[5]. However, these methods are highly affected by factors such as light, angle and occlusion, and have poor robustness. With the rise of smart wearable devices, some manufacturers have started to combine smart helmets with smart wearable devices to realize real-time monitoring of wearing compliance. However, smart helmets have drawbacks and limitations such as high cost, technical complexity, dependence on the network, privacy and data security issues. Initial research focused on feature extraction and classifier design, such as the use of HOG (Histogram of Orientation Gradient) features and SVM (Support Vector Machine) classifiers for helmet detection [6].

With the rise of deep learning technology, domestic research in the field of helmet wear detection has made breakthrough progress. In 2016, Jia Junsu et al [7] proposed a recognition method based on the composite of color features, HOG features and local binary pattern features. After entering 2018, domestic research has focused more on real-time and diversity processing. Huimin He et al [8] used CNN (Convolutional Neural Network) to determine the worker's position, and then detected helmet wearing by color space and CHT (Hough Transform). In 2019, Shi Hui et al [9] combined the image pyramid structure and the YOLOv3 model to further improve the real-time and accuracy of helmet wearing detection.

In recent years, with the continuous maturation of deep learning technology, the performance of helmet wearing detection algorithms has been significantly improved. Based on deep neural network models such as Faster R-CNN [10], SSD [11] and YOLOv3 [12], researchers have constructed a series of efficient and accurate helmet recognition systems. These researches not only improve the accuracy of detection, but also meet the real-time requirements, providing strong support for practical applications.

Driven by deep learning technology, the performance and real-time performance of helmet wear detection have been greatly improved. However, in the face of complex and changing real-world environments, it is still necessary to continuously optimize the algorithm and improve the robustness and diversity processing capability. In this paper, the algorithm is further optimized based on YOLOv5, combining the CBAM attention mechanism, BiFPN structure and Focal loss function and other technologies, which significantly improves the accuracy and precision of helmet wear detection and identification, comprehensively improves the level of safety protection in the chemical production process, and meets the accuracy and real-time demand of helmet wear detection in chemical safety scenarios.

II. DESIGN OF HELMET DETECTION ALGORITHMS FOR CHEMICAL SAFETY SCENARIOS

2.1 CBAM attention mechanism

CBAM (Convolutional Block Attention Module) Attention Mechanism is a lightweight and efficient attention module designed to improve the performance of a model by enhancing the feature representation capabilities of Convolutional Neural Networks (CNNs) in both channel and spatial dimensions. CBAM was developed by Sanghyun Woo et al [13] in 2018, which enables the model to capture and utilize key information more accurately by introducing both channel attention and spatial attention mechanisms. The two modules, Channel Attention Module (CAM) and Spatial Attention Module (SAM), can be operated independently but are usually used in tandem to sequentially enhance the feature map. The following is a detailed description of the CBAM attention mechanism.

CBAM attention mechanism is an effective deep learning

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

technique, which significantly improves the feature representation capability and model performance of convolutional neural networks by introducing two mechanisms, channel attention and spatial attention. CBAM module has the advantages of high adaptivity, high computational efficiency and plug-and-play, and has a wide range of application prospects in the fields of computer vision and natural language processing, etc. CBAM attention structure is shown in Figure 1.

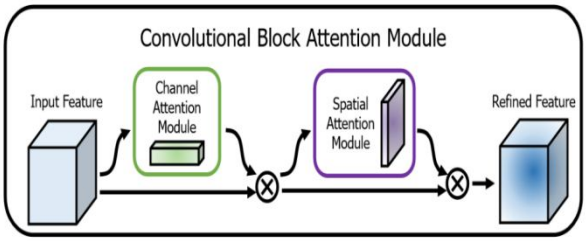

Figure 1 CBAM Attention Structure

2.2 BiFPN module

BiFPN (Bi-directional Feature Pyramid Network) module [14] is an efficient feature fusion network for computer vision tasks such as target detection and semantic segmentation. It significantly improves the accuracy and efficiency of target detection through multi-level feature fusion and dynamic feature weight assignment. BiFPN is an improved version of Feature Pyramid Network (FPN) [15], which aims to solve the information bottleneck and computational efficiency problems of FPN when dealing with multi-scale features.FPN is an network structure, which is usually used in combination with a backbone network (such as ResNet or EfficientNet) to generate feature pyramids with different resolutions to improve the performance of object detection and segmentation. BiFPN, on the other hand, is optimized on this basis by introducing mechanisms such as bi-directional connectivity and adaptive feature tuning for better fusion of multi-scale features and improved model performance.

As an efficient feature fusion network for target detection, the BiFPN module significantly improves the accuracy and efficiency of target detection through mechanisms such as bi-directional connectivity, adaptive feature tuning and modularized design. Its wide range of application scenarios and superior performance make BiFPN an important research value and application prospect in the field of computer vision.

2.3 Focal loss function

The category imbalance problem is a common challenge during helmet detection in chemical safety scenarios, and this imbalance may lead to poor detection performance of the model for a small number of categories because the model prefers to predict a larger number of categories. To address this problem, this paper adds the Focal loss function to the loss function of YOLOv5. Focal Loss is a loss function designed to solve the category imbalance problem, which was proposed by Kaiming He's team in 2017 [16], and is particularly suitable for object detection tasks such as the RetinaNet model. In tasks such as target detection, the number of positive and negative samples (the target object and the background) is often extremely unbalanced, causing the model to focus excessively on simple negative samples during training, which affects the performance of the model. Focal Loss effectively solves this problem by adjusting the weights of easy-to-classify and hard-to-classify samples.

In traditional classification tasks, the commonly used loss function is the cross-entropy loss, which treats the classification errors of all samples equally, without considering the difficulty of the samples or the imbalance of the class distribution. However, in practical applications such as target detection, this treatment can cause the model to pay too much attention to the simple negative samples and ignore the difficult to classify positive samples, which in turn affects the overall performance of the model. Focal Loss improves on the cross-entropy loss by introducing two new parameters, namely, the modulating factor and the scaling Focal Loss is an improvement on the cross-entropy loss by introducing two new parameters, modulating factor and scaling factor, which allow the model to focus more on the samples that are difficult to categorize. The specific form of Focal Loss is shown below:

$$FL(p_t) = -\alpha_t(1-p_t)^{\gamma}\log(p_t)$$

$$p_t\begin{cases} p, & if \quad y=1 \\ 1-p, & otherwise \end{cases} \tag{1}$$

where p_t is the predicted probability of the model for the target category, α_t is a balancing factor to balance the weights of positive and negative samples, and γ is a Focal Loss factor to regulate the weights of difficult and easy samples. When γ is 0, Focal Loss is degraded to Cross Entropy Loss.

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

The balancing factor α_t is used to adjust the contribution of positive and negative samples to the loss function. In the target detection task, since the number of negative samples (background) is usually much larger than the number of positive samples (target object), the direct use of Cross Entropy Loss will cause the model to focus too much on the negative samples. This imbalance can be mitigated by setting α_t , which weights the losses for positive and negative samples. Usually, for positive samples, α_t = α; for negative samples, α_t = 1-α. The value of α ranges from 0 to 1 and needs to be adjusted according to the sample distribution of the specific task. The focus factor γ is used to adjust the weights of difficult and easy samples. When the samples are correctly categorized with high confidence (p_t is close to 1), the value of $(1-p_t)^\gamma$ will be close to 0, which decreases the contribution of that sample to the loss; on the contrary, when the samples are incorrectly categorized or with low confidence (p_t is close to 0), the value of $(1-p_t)^\gamma$ will be close to 1, which increases the contribution of that sample to the loss. In this way, Focal Loss makes the model focus more on samples that are difficult to classify.

The working principle of Focal Loss can be summarized in two aspects: one is to adjust the weights of positive and negative samples through the balancing factor α_t, and the other is to adjust the weights of difficult samples through the focus factor γ. In the traditional cross-entropy loss, the positive and negative samples contribute equally to the loss. However, in tasks such as target detection, since the number of negative samples is much larger than the number of positive samples, the direct use of cross-entropy loss will cause the model to focus excessively on the negative samples. Focal Loss weights the losses of positive and negative samples by introducing a balancing factor, α_t , which enables the model to focus more on the positive samples, especially the difficult-to-detect positive samples, during the training process. During model training, easily categorized samples (samples with high confidence) tend to contribute less to the loss, while difficult to categorize samples (samples with low confidence) contribute more to the loss. However, since the number of easy-to-classify samples is usually much larger than the number of hard-to-classify samples, the sum of their losses dominates the total loss. Focal Loss downweights the losses of the easy-to-classify samples by introducing a focus factor γ, which makes the model focus more on the hard-to-classify samples. When γ takes a large value, the loss of easy-to-classify samples decreases significantly, while the loss of hard-to-classify samples remains relatively unchanged or increases slightly.

2.4 Improved method based on YOLOv5

Focusing on the task of helmet detection in chemical safety scenarios, this study proposes an innovative deep learning algorithm based on the YOLOv5 framework, as shown in Figs. 2. In view of the strict regulatory requirements for helmet wearing in chemical environments and the limitations of traditional detection methods in complex scenarios, a set of algorithms integrating an advanced hybrid attention mechanism and efficient feature fusion techniques is designed, aiming to significantly improve the accuracy and efficiency of helmet detection.

One of the core strategies is to seamlessly integrate the CBAM module into the neck part of the YOLOv5s model. This initiative not only enhances the model's sensitivity to critical regions in the image, but also enables the model to focus more accurately on the subtle features of the worker's head through the refined channel and spatial attention mechanisms, and effectively recognize helmet wearing even under challenging conditions such as light changes, occlusion, or long-distance shooting. The introduction of the CBAM module significantly enhances the model's complex chemical environment for helmet detection capability.

To further enhance the performance of the algorithm, especially for the possible target scale diversity and category imbalance problems in chemical scenes, the BiFPN module is innovatively introduced. BiFPN realizes top-down and bottom-up feature fusion by constructing bi-directional cross-scalar connectivity, which not only retains the rich details of the low-level features, but also fuses the semantic information of the high-level features. This efficient feature fusion mechanism enables the model to show strong adaptability and accuracy in detecting helmets of different sizes, shapes and positions. Combined with the multi-scale prediction capability of YOLOv5s, the addition of BiFPN further enhances the overall detection effect of the algorithm.

In addition, for the common category imbalance problem

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

in chemical safety detection, i.e., the difference in the number of samples between helmet wearers and non-wearers, we adopt the Focal Loss function as the loss function. By dynamically adjusting the weights of easy-to-classify samples and difficult-to-classify samples, the Focal Loss effectively mitigates the negative impact of the category imbalance, allowing the model to focus more on the samples that are difficult to differentiate during training, and thus improving the overall effectiveness of the algorithm. those difficult-to-distinguish samples, thus improving the detection accuracy for a few classes (e.g., the case of not wearing a helmet). This improvement ensures that the algorithm maintains a high degree of accuracy and robustness in chemical safety scenarios, both for the detection of correct helmet wearing and violations.

In summary, in this study, an efficient and accurate helmet detection algorithm for chemical safety scenarios is constructed by integrating the CBAM module, BiFPN module, and adopting the Focal Loss function to deeply optimize the YOLOv5s algorithm. The algorithm not only improves the accuracy of helmet detection, but also enhances its adaptability and robustness in complex environments, which provides a powerful technical support for safe production in the chemical industry.

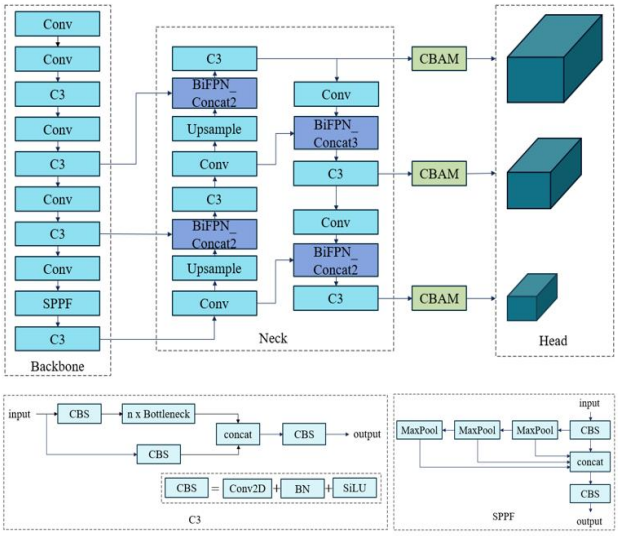

Figure 2 Improved YOLOv5s network structure diagram

III. EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Experimental evaluation indicators

In the field of target detection, evaluation metrics are key measures of model performance. These metrics not only help us understand how the model performs on a particular dataset, but also guide us on how to improve and optimize the model. The following is a detailed description of common evaluation metrics in target detection, including Accuracy, Precision, Recall, F1 value, Intersection over Union (IoU), and mean Average Precision (mAP). Precision (mAP).

Although accuracy is the most intuitive metric to evaluate in classification problems, its direct application in target detection is somewhat limited. The accuracy rate is defined as the ratio of the number of samples correctly predicted by the model to the total number of samples. However, in the target detection task, since the image may contain multiple targets and the background (negative samples) is much more than the target (positive samples), a simple accuracy rate calculation is difficult to fully reflect the performance of the model. Therefore, more detailed metrics such as precision rate and recall rate are more often used in target detection.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (2)$$

Where true (TP) indicates the number of positive samples that are correctly detected. True Negative (TN) denotes the number of negative samples that are correctly detected as negative samples, but TN is usually not of interest in target detection because the total number of negative samples is much larger than the number of positive samples and is difficult to define precisely. False Positive (FP) indicates the number of negative samples that were incorrectly detected as positive samples, also known as false positives. False Negative (FN) indicates the number of positive samples that were not detected, also known as underreporting.

The precision rate is the proportion of all samples predicted as positive by the model that are truly positive, that is, how many of the targets predicted by the model are correct. The precision rate reflects the accuracy of the model in identifying positive samples, i.e. the ability to “find the right one”. The formula is as follows:

$$Precision = \frac{TP}{FP + FP} \quad 3$$

Recall is the proportion of all real positive samples that are correctly recognized as positive by the model, that is, how many of the actual targets are correctly detected by the model. Recall reflects the degree of coverage of positive

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

samples by the model, i.e. the ability to “find all”. The formula is as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Precision and recall often have a contradictory relationship, and increasing one indicator may decrease the other. Therefore, in practical applications, it is necessary to find a balance between the two according to specific needs.

In order to comprehensively evaluate precision and recall, the F1 value is introduced. F1 value is the reconciled average of precision and recall, and its calculation formula is shown in (5). The higher F1 value indicates that the model performs better in both precision and recall. F1 value is a single value, which facilitates the direct comparison between different models.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

The intersection and concatenation ratio is a key metric to measure the localization accuracy of the target detection model. It calculates the ratio of intersection and concatenation between the model's predicted rectangular box (Detection Result, DR) and the real labeled rectangular box (Ground Truth, GT), and the formula is shown in (6). The higher value of IoU indicates that the model's predicted box overlaps with the real box more, and the more accurate the localization is. In target detection, an IoU threshold (e.g., 0.5) is usually set, and the prediction is considered correct only when the IoU value of the predicted frame and the true frame is greater than this threshold.

$$IoU = \frac{DR \cup GT}{DR \cap GT} \quad (6)$$

Average mean precision is an important measure of the overall performance of a target detection model over multiple categories. Since the target detection task usually involves the detection of multiple categories, the Average Precision (AP) needs to be calculated separately for each category and then averaged to obtain mAP. for a given category, the AP is the average of the precision rates at different recall rates. In order to calculate the AP, it is necessary to first sort the prediction frames from high to low confidence, then calculate the precision and recall corresponding to each prediction frame in turn, plot the PR curve, and calculate the Area Under Curve (AUC), which is the AP value. The AP values of all categories are averaged to obtain mAP, whose formula is shown in (7), where AP_i is the AP of the Kth category, and K is the total number of categories. mAP value is higher, which indicates that the overall performance of the model on multiple categories is better. mAP is a comprehensive indicator of the performance of the target detection model, which is widely used in a variety of target detection competitions and practical applications.

$$mAP = \frac{1}{K}\sum_{i=1}^{k} AP_i \quad (7)$$

In order to evaluate the model performance more comprehensively, the mAP under different IoU thresholds is sometimes calculated. for example, the COCO dataset provides mAP metrics under a total of 10 IoU thresholds ranging from 0.5 to 0.95 (with a step size of 0.05). For the small target detection problem, it is relatively difficult to detect small targets due to their small proportion in the image and their susceptibility to background interference and noise. In order to evaluate the performance of the model on small target detection, the mAP value of small targets is usually calculated separately. In addition to the detection accuracy, the real-time performance of the target detection model is also one of the important evaluation indexes. Real-time performance is usually measured by the number of image frames processed per second (Frames Per Second, FPS). In practical applications, the trade-off between accuracy and real-time performance needs to be made according to specific scenes. In summary, the evaluation metrics of target detection include accuracy, precision, recall, F1 value, intersection and concurrency ratio, and average mean precision. These metrics comprehensively evaluate the performance of the target detection model from different perspectives and provide strong support for model optimization and improvement.

3.2 Experimental environment

In this paper, the image data and labeled dataset of the custom test are divided into training set, validation set, and test set in the ratio of 96%:2%:2%. There are 5000 images in the dataset, of which 4800 images are in the training set, 100 images are in the validation set, 100 images are in the test set, and the training set and test set are strictly

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

independent. The batch size (Batch Size) is set to 32 and the training period (epoch) is set to 50 during the training process.

3.3 Analysis of experimental results

In this study, YOLOv5s is used as the base algorithm, which is trained and validated with the algorithm that combines the CBAM module alone and the improved algorithm that combines the CBAM module as well as the BiFPN module, respectively, in the chemical safety scenario helmet-wearing detection dataset of this paper, and a comparative analysis of the data is carried out. Among them, the improved training results are shown in Figure 3:

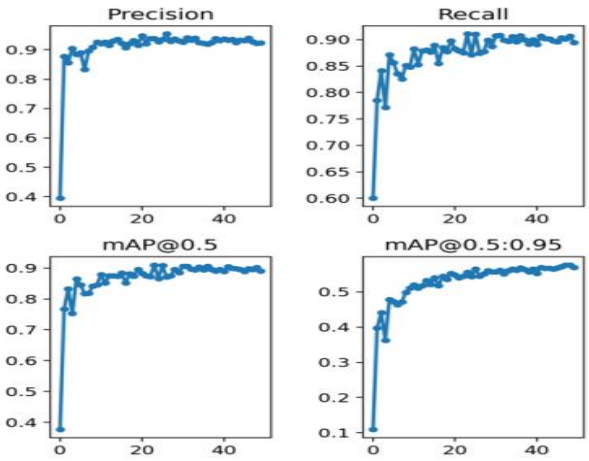

Figure 3 Improved algorithm training result graph

In this study, Precision, Recall value, and mAP were used as the main evaluation indexes. mAP@0.5 is the calculated value of mAP when the IoU threshold is 0.5, which is an important indicator to measure the performance of the object detection model. It reflects the average detection accuracy of the model on different categories, especially under the criterion that detection is considered correct when the IoU is at least 0.5. mAP@.5:.95 is a more stringent evaluation criterion that calculates the mAP at each threshold as the IoU threshold changes from 0.5 to 0.95 (including 0.05 steps) and averages these values. This type of evaluation provides a more comprehensive assessment of the model's performance under different accuracy requirements. The comparative experimental data are shown in Table 1:

Table 1 Comparative Experimental Data Table

Model	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@.5:.95 (%)
YOLOv5s	92.0	88.8	88.6	54.9
YOLOv5s+CBAM	92.3	89.5	89.0	57.0
Improved YOLOv5s	**95.3**	**90.9**	**91.8**	**57.4**

From the data in Table 1, it can be analyzed that different versions of the YOLOv5 model show different performance in the task of helmet detection in chemical safety scenarios. The base model, YOLOv5s, has shown quite good performance, with its precision (92.0%) and recall (88.8%) at a high level, which suggests that the model can effectively recognize helmets in images in most cases. However, 从 mAP@0.5 (88.6%) 和 mAP@.5:.95 (54.9%) these two indicators show that YOLOv5s may have some limitations when dealing with complex scenes or subtle differences, especially under the mAP@.5:.95 standard, which requires higher detection accuracy, its performance needs to be improved. The YOLOv5s+CBAM model after the introduction of the CBAM module shows that the attention mechanism helps the model to recognize helmets more accurately in terms of accuracy, The recall rate and mAP@0.5 have improved, especially in improving the detection accuracy, mAP@.5:.95 also increased to 57.0%, which suggests that the attention mechanism effectively enhances the model's ability to capture key in-image information capturing ability, thus improving the accuracy of helmet detection. The Improved YOLOv5s model, on the other hand, exhibits the highest precision (95.3%) and recall (90.9%), which suggests that the model greatly improves the accuracy of detection while maintaining a high recall, which is important for reducing missed and false detections. The significant improvement of the model on mAP@0.5 (91.8%) indicates that its detection ability of safety helmets in chemical safety scenarios is more comprehensive and stable, and it is suitable for safety monitoring tasks requiring high accuracy and high recall rate.

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

The improvement of mAP@.5:.95 shows that the model maintains better stability and robustness over a wider range of IoU thresholds, which is critical for practical applications in chemical safety scenarios.

In the task of target detection, Ground Truth Box and Prediction Box are two core concepts, where the Ground Truth Box is the bounding box of the target that actually exists in the image, which is usually determined by manual labeling or other accurate methods. It represents the region where the target really exists in the image, including information such as the target's location, size and category. The prediction box is the box derived by the model through computation and reasoning about the area where the target may be located. It is the result of the model's prediction of the location and shape of the target based on the input image data and learned features. The accuracy and effectiveness of the prediction frame is usually evaluated by calculating the IOU value between it and the true frame. The higher the IOU value, the higher the overlap between the prediction frame and the true frame, and the better the prediction effect of the model.

The real frame and the predicted frame represent the actual position of the target in the image and the position predicted by the model in the target detection task, respectively. The accuracy and effectiveness of target detection can be improved by continuously optimizing the model performance so that the prediction frame is as close as possible to the real frame.

IV. Conclusion

In this paper, introduced the CBAM module, BiFPN structure, Focal loss function, and improved YOLOv5 model, and constructed an efficient and accurate safety helmet detection algorithm in chemical safety scenarios. Next, the experimental evaluation indicators such as accuracy, precision, recall, and intersection to union ratio were described, and the experimental environment parameters were listed. Then, the training results of the improved algorithm for detecting the wearing of safety helmets in chemical safety scenarios were presented, and comparative experimental analysis was conducted. The results showed that it has a more comprehensive and stable detection ability for safety helmets in chemical safety scenarios, and is suitable for safety monitoring tasks that require high precision and high recall. The improved model maintains better stability and robustness in a wider range of IoU thresholds, which is suitable for practical applications in chemical safety scenarios. Finally, the box predicted by the model is basically consistent with the real box, indicating that the improved model has better accuracy and prediction performance.

ACKNOWLEDGEMENTS

Project supported by Provincial Natural Science Foundation of Hunan (2022JJ50120), innovation and entrepreneurship training program for college students in Hunan Province in 2023and 2024.

REFERENCES

[1]. Cui T, Wang Y, Xu G. The Storage Tank Explosion Damage and the Effectiveness of Control Measures in the Chemical Industrial Parks of Smart Cities. Electronics. 2024; 13(14):2757.

[2]. Tang Y, Wang B, He W, et al. PointDet++: an object detection framework based on human local features with transformer encoder[J]. Neural Computing and Applications, 2022:1-12.

[3]. Javeria Amin, Annam Zafar. A Survey: Content Based Image Retrieval [J].Int. J. Advanced Networking and Applications, 2014, 5(6): 2076-2083

[4]. Duffner S, Odobez J M. Leveraging Colour Segmentation for Upper-Body Detection[J]. Pattern Recognition, 2014, 47(6):2222-2230.

[5]. P.Manimegalai, Dr.K.Thanushkodi. Content-Based Color Image Retrieval Using Adaptive Lifting[J].Int. J. Advanced Networking and Applications, 2010, 1(6): 359-366

[6]. Park M W, Elsafty N, Zhu Z H. Hardhat-wearing detection for enhancing on-site safety of construction workers. J Constr Eng Manage, 2015, 141(9): 04015024

[7]. JIA J S, BAO Q J, TANG H M. Safety helmet wear detection based on deformable component

Volume: 16 Issue: 05 Pages: 6569-6577 (2025) ISSN: 0975-0290

modeling[J].Computer Applications Research,2016,33(03):953-956.

[8]. HO Hui-Min, LIN Chi-Hsien, KUO Tai-Liang. Automatic recognition of pedestrian helmet based on convolutional neural network[J]. Cable TV Technology,2018,(03):104-108

[9]. Shi H, Chen X Q, Yang Y. Safety helmet wearing detection method of improved YOLO v3. Comput Eng Appl, 2019, 55(11): 213

[10]. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal Mach Intell, 2017, 39(6):1137

[11]. Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector // European Conference on Computer Vision. Cham, 2016: 21

[12]. REDMON J, FARHADI A. Yolov3: an incremental improvement[C] //Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 89- 95.

[13]. Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.

[14]. Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10778

[15]. Lin T Y, Dollar P, Girshick R, et al. Feature Pyramid Networks for Object Detection[J]. IEEE Computer Society, 2017.

[16]. Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP(99):2999-3007.