A Review of YOLO Algorithm and The Research in Garbage Classification Target Detection

Juan Zhang
Mapua University
Makati City, Philippines
jzhang@mymail.mapua.edu.ph

Eric B. Blancaflor
Mapua University
Makati City, Philippines
EBBlancaflor@mapua.edu.ph

Abstract

Target detection is an important research direction in computer vision, how to real-time accurate object detection of objects is the focus of research. In recent years, deep learning technology has developed rapidly and are widely used in various fields. Firstly, this paper summarizes the idea of YOLO series detection algorithm and its improvement process, analyzes the advantages and disadvantages of YOLO series algorithm; then, expounds the application status of YOLO algorithm in garbage classification target detection, summarizes the public garbage classification dataset and commonly used evaluation indicators in target detection; finally, summarizes the problems in garbage classification identification, and forecast the future development direction.

CCS Concepts

• **General and reference** → Document types; Surveys and overviews.

Keywords

Target detection, Garbage classification, Real-time detection, YOLO algorithm

ACM Reference Format:
Juan Zhang and Eric B. Blancaflor. 2024. A Review of YOLO Algorithm and The Research in Garbage Classification Target Detection. In *2024 4th International Joint Conference on Robotics and Artificial Intelligence (JCRAI 2024), September 13–15, 2024, Shanghai, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3696474.3702878

1 Introduction

According to the authoritative release of the National Bureau of Statistics of China, by 2023, China's total population in China has exceeded 1.4 billion. With the rapid development of China's economy, the development of express delivery, takeout, electronic products, fast fashion and other industries have brought convenience to people's life, while producing more domestic garbage. If these household garbage cannot be timely and effectively classified and treated, it will cause serious environmental pollution and the spread of infectious diseases, and will bring about a serious waste of resources.

In recent years, the target detection technology based on deep learning has become increasingly mature, and more and more deep learning technology can be applied to the field of target detection. The two-stage algorithm extracts the Region Proposal network, and then regression the samples. The representative algorithm includes R-CNN, Fast R-CNN and Faster R CNN); the single-stage algorithm directly extracts the features for classification and regression, which greatly improves the efficiency of target detection [1]. The representative algorithm includes SSD and YOLO series algorithms. As the pioneer of the single stage algorithm, YOLOv1 creatively detection problems directly as a unified end-to-end regression problem, make it compared with the same period of R-CNN algorithm has faster detection speed, since then, YOLOv2-YOLOv9 algorithm optimization, in the case of guarantee accuracy, shortens the target detection time, and in the garbage classification, and other application scenarios show a strong ability to identify.

2 The YOLO algorithm development

2.1 *YOLOv1–YOLOv3*

Yolov1 Is an end-to-end object detection algorithm based on a single image. The core idea is to use a single convolutional neural network to predict the category and boundary boxes of objects. During training, the Yolo network borrows from the GoogLe Net classification network structure. Yolo uses 1x1 convolutional layers and 3x3 convolutional layers to replace the inception module, and the whole detection network includes 24 convolutional layers and 2 fully connected layers. The convolution layer is used to extract image features, and the fully connected layer is used to predict the image location and category probability values [2].

Yolov2 proposed in 2017, Darknet-19, including 19 convolution layers and 5 pooling layers, mainly using 3x3 convolution kernel and 1x1 convolution kernel.1x1 convolution can compress the number of feature graph channels to reduce the model calculation amount and parameters. After each convolution layer, BN layer is used to accelerate the model convergence while preventing overfitting. Finally, global avg pool is used for prediction. Yolov2 Support more object categories, can predict over 9,000 object categories, and simultaneously detect multiple target of different sizes.Yolov3 was added on the basis of Darknet-19, and further deepened the network, the improved network has 53 convolution layers, named Darknet-53, this version has been greatly improved in terms of speed and accuracy.

Yolov3 The multi-scale feature graph prediction strategy is adopted, and multiple convolution layers are added to expand the

JCRAI 2024, Shanghai, China

ACM ISBN 979-8-4007-1010-0/24/09
https://doi.org/10.1145/3696474.3702878

network receptive field block to enhance the network generalization ability.

2.2 *OLOv4–YOLOv6*

YOLOv4 Adding the backbone structure to the backbone network Darknet-53 of YOLOv3, including 5 CSP modules, increased the Droblock to alleviate the overfitting phenomenon, and used the Mish activation function to improve the accuracy. This algorithm further improves the detection speed and performance based on Yolov3, using CmBN instead of BN, SPP and PAN technology, making the average accuracy (AP) of YOLOv4 on the COCO dataset reaches 43.5%.

Yolov5 is a target detection algorithm based on PyTorch, which uses a detection method called ”anchor-free”, which does not need to define the box size and position of a prior in advance. It uses ”Swish” activation function and data enhancement technology for training, which improves the detection speed and accuracy.

YOLOX adds two 3x3 convolutions based on YOLOv3 network structure to decouple classification and regression tasks, achieving the effect of improving accuracy and convergence speed. For FPN features, a 1x1 convolution layer is used to reduce feature channels, and parallel branches are added for classification and regression tasks respectively.

Yolov6 Mainly optimizes the model structure and training strategy. The algorithm adopts a convolution operation called ”Asymmetric Convolution”, and uses the lightweight technology to reduce the model size and running time while ensuring the accuracy.

2.3 *YOLOv7–YOLOv9*

Yolov7 is the version of YOLOV4’s original team, which focuses on improvement and optimization on the ”model structure”, using ”autoML” technology, namely automatic machine learning technology.

Both YOLOv8 and YOLOv5 Ultralytics team release, provides a new State Of The Art model and based on the scaling coefficient provide N, S, M, L, X scale model of different size to meet the requirements of different scenarios, on the basis of YOLOv5 based on YOLOv7 E-ELAN module C2f module of C3 module in YOLOv5, add more gradient flow information, and refer to YOLOX, using the coupling head structure to decoupling the classification and regression. YOLOv8 No a priori box detection method and task alignment learning dynamic matching method. From the perspective of task alignment, the high-quality prior box is dynamically selected as the function of the positive sample integration loss design, and Distribution Focal Loss and CIoU Loss are used as the regression loss.

YOLOv9 is an advanced object detection model that represents a significant advance in computer vision technology, is currently the latest version and is known for its high speed and accuracy in detecting objects in images. YOLOv9 is unique in that it integrates the programmable gradient information PGI and introduces the generalized high-efficiency layer aggregation network GELAN, aiming to improve the model performance and efficiency. YOLOv9’s performance on the coco data set is shown in Fig1.

The difference in size of YOLOV9 models meets more application needs, from lightweight models on edge devices to more comprehensive models in high-performance computing environments.The performance metrics of different versions of the YOLOv9 model are shown in Fig2.

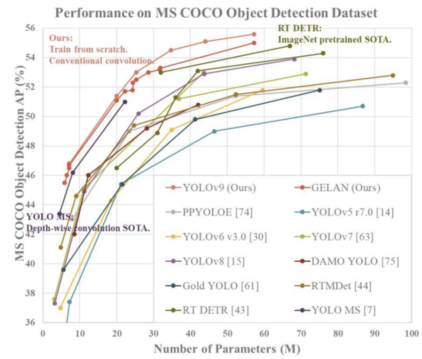

Figure 1: YOLOv9 Performance on the coco datasets.

Model	size (pixels)	AP^{val}	AP_{50}^{val}	AP_{75}^{val}	Params (M)	FLOPs (G)
YOLOv9-S	640	46.8%	63.4%	50.7%	7.1	26.3
YOLOv9-M	640	51.4%	68.1%	56.1%	20.0	76.3
YOLOv9-C	640	53.0%	70.2%	57.8%	25.3	102.1
YOLOv9-E	640	55.6%	72.8%	60.6%	57.3	189.0

Figure 2: Model performance parameters of the YOLOV9.

2.4 *Main improvement direction of the YOLO series algorithm*

YOLO series target detection algorithm focuses on balancing the detection speed and accuracy, aiming to improve the real-time detection without sacrificing the detection accuracy. Each version of YOLO improves the real-time performance while improving the accuracy. The model after YOLOv4 provides different proportions of models to adapt to the specific application environment, which is more convenient for users to choose.The development will aim to obtain faster and more efficient network architecture, more efficient feature integration methods, more accurate detection methods, more robust loss function, more effective label allocation methods, more effective training methods and better compatibility, etc.

3 The Research of garbage classification based on yolo algorithm

3.1 *Garbage classification image datasets*

The number, type, and size of image datasets significantly impact the robustness of a model’s performance[3]. For example, the public model training data is too little, most of which are not effective in practice, and some homemade data sets have too much emphasis on single feature.

In 2016, Yang et al. of Stanford University created the dataset TrashNet. TrashNet consists of 2527 household waste images divided into six categories, including 594 paper images, 501 glass images, 482 plastic images, 410 metal images, 403 cardboard images, and 137 other junk images[4]. TrashNet data set is an RGB image of garbage placed on a white poster and then taken in natural light or indoor lighting. The resolution of all the images was adjusted to 512x384, and each image contained only one type of garbage. Similar to TrashNet, Glassense-Vision are both datasets In which the objects are relatively separated from the background. Glassense-Vision collected 2,000 images covering seven categories (paper money, cereal, medicine, cans, ketchup, water bottles and taste sticks) and 136 sub-categories. The Drinking Waste dataset was produced by University College London using some of the images in TrashNet. Drinking Waste Taken with a 12 MP mobile phone camera, it contains about 10,000 images, divided into four categories: cans, glass bottles, plastic bottles and plastic milk boxes. The GINI data set is obtained by searching for junk images on the search engine Bing, so the data in GINI is divided into junk and non-junk two categories. Of these, there were 2,561 non-junk images, and 1,496 junk images marked by demarcation boxes. The UAVVaste dataset currently has 772 images and 3,716 annotations, and the main motivation for creating it is to compensate for the scarcity of drone junk detection datasets. But this open-source dataset is now also widely used for evaluation benchmarks for target detection. TACO contains junk images from tropical beaches to the streets of London taken in a variety of environments, containing 1,500 images annotated in the COCO-json format. TACO roughly divided the images into 28 categories, and the TACO boxes dataset manually annotated the images of TACO, dividing TACO into 7 major categories and 60 subcategories. Waste_pictures has over 24,000 images divided into 34 categories. TrashBox The data set has 17,785 images. The images were divided into seven major categories (drugs, e-waste, plastic, metal, paper, glass and cardboard) and 25 subcategories. Domestic Trash The data volume of the data set is also relatively large, it is taken in a variety of light conditions, weather, indoor and outdoor conditions, with about 9,000 common garbage images. Domestic Trash Contains a variety of junk categories and multiple network formats, along with material labels, but has only 250 free samples on Kaggle, requiring a license to download and use the full data set. WaDaBa Is a plastic junk image dataset. It placed the plastic waste on the background plate, then took 10 photos of the plastic intact, slightly damaged, moderately damaged and severely damaged, and took the amount of plastic waste to get the 4,000 images in the dataset.

The data set used in the 2019 China "Huawei Cloud Artificial Intelligence Competition · Garbage Classification Challenge Cup" has also been made public. According to the garbage classification standard of Shenzhen, China, the domestic waste is divided into recyclables, kitchen waste, hazardous waste and other categories; the four categories are divided into 44 categories according to common garbage categories in life, with a total of 14,964 Figure [5]. Garbage Classification Dataset Released by the Institute of Software of the Chinese Academy of Sciences, it contains about 5,000 pictures of garbage, divided into seven categories, including kitchen waste, hazardous waste, recyclable waste and others. This dataset also has some applications in garbage classification research. Tianjin University Trash Dataset: The garbage classification data set released by Tianjin University contains about 9,000 pictures and is divided into five categories: kitchen waste, hazardous waste, recyclable waste, other garbage and recyclable waste. This dataset is mainly used for the deep learning study of garbage classification. HUST-DeepGarbage Dataset: The data set was released by Huazhong University of Science and Technology and contains more than 50,000 images of garbage, covering 11 categories, including recyclable waste, hazardous waste, kitchen waste and others. This dataset is widely used in research in the fields of deep learning and image recognition. The above garbage classification datasets are shown in Table1.

The data set mentioned above is the public data set of garbage classification. Most of the tested garbage classification data sets are not yet open source, so the garbage target detection data set is still relatively scarce.

3.2 *Evaluation indicators and parameters*

The evaluation indexes of target detection mainly include Precision, Recall, all categories mean Average Precision, Frame Per Second, F1 score and so on. For the single-stage end-to-end detection algorithm YOLO, another important performance indicator is speed.

Accuracy (P) refers to the proportion of the correctly predicted positive samples to the total predicted positive samples, also known as the precision accuracy, which reflects the accuracy of the detection results.

Recall rate (R) refers to the proportion of the correctly predicted positive samples to all the actual positive samples, also known as the recall rate, which reflects the comprehensiveness of the test results.

The average accuracy (mAP) is the area enclosed by the P-R curve, measuring how well the model predicts on each category. The quality of the mAP measurement model in all categories is an important evaluation index of the target detection algorithm.

The F1 score, also known as the balanced F score, represents the harmonic average evaluation index of precision and recall, which considers two indicators to evaluate the performance of the classification algorithm.

FPS is the number of frames detected per second, which refers to the number of image frames that the algorithm can process in a unit of time, reflecting the running speed of the algorithm. It is calculated by multiplying the reciprocal of the total image preprocessing time, inference time and NMS time by 1000. A larger FPS value indicates faster detection and better model performance.

3.3 *Research of garbage classification target detection*

Wu Jianian[5] others balanced the detection accuracy and speed of garbage detection and designed an improved YOLOv4 detection algorithm. This algorithm replaces the backbone of the original network with MobileNetV3 to achieve lightweight. To improve detection accuracy, MobileNetV3 is used to lightweight HrNet, ensuring speed while improving detection speed. The experimental results show that the improved algorithm has improved detection speed and accuracy compared to the original baseline network.

Table 1: Garbage classification dataset

Data Set	Categories	Sub Classification	Data Set	Categories	Sub Classification
TrashNet	6	-	Waste_pictures	34	-
Glassense-Vision	7	136	TrashBox	7	25
Drinking Waste	4	-	Domestic Trash	10	-
GINI	2	-	WaDaBa	8	32
UAVVaste	-	-	HuaweiCloud classification dataset	4	44
TACO	28	-	Tianjin University Dataset:	5	-

Xu Wei [6] proposed a lightweight Ghost-YOLO garbage detection algorithm. The algorithm is aimed at low power consumption equipment and improved based on YOLOV3, which can guarantee the detection speed and have high detection accuracy. In terms of reducing the number of parameters, the Ghost bottle neck module was used to replace the original backbone bottleneck. In terms of improving the accuracy, the feature fusion part was improved, and the down sampling link was introduced for secondary feature fusion. The experiment proved that the number of parameters of the improved algorithm is greatly reduced, and the detection accuracy reaches 89.02%.

Dong Xiaoxiao [7] improved the YOLOV3 network to realize the garbage detection under the complex background. The improved algorithm combines YOLOV3 with the lightweight network to ensure the network detection accuracy and improve the recognition speed. Then, the feature pyramid of YOLOV3 is improved, the weighted bidirectional feature pyramid is introduced, and the empty space pyramid pooling is introduced to improve the detection accuracy. The experiment is verified on the data set, and the improved algorithm model is greatly improved in both accuracy and detection speed.

Wang Zao[8] designed a SPMYOLOv3 detection algorithm for water surface waste, which was improved based on YOLOV3. First made the surface garbage detection data set, and use the improved K-means algorithm to complete the data set clustering, generate a new anchor box, then to strengthen the feature extraction ability, introduce SE-PPM module in the backbone part, and then use the multi-directional pyramid network to replace YOLOV3 feature pyramid network, get more context information, finally introduce the focal loss loss function to solve the problem of imbalance between positive and negative samples. Experiments show that the improved algorithm accuracy improves by 3.96% over YOLOV3.

Xu Chuanyun[4] proposed a multi-level feature weighted fusion algorithm for garbage classification. This algorithm is based on ResNet and extracts and processes image features at different levels through a multi branch network. Based on this, a better weighted fusion method is selected for multi-level feature fusion to improve the accuracy of garbage classification. The experimental results show that compared to other algorithms, this algorithm has better overall performance and stronger practical performance.

Tian Manjun [9]proposed a garbage detection algorithm based on YOLOv4 for detecting underwater garbage. This algorithm takes YOLOv4 as the baseline network, adds a detection head to improve detection accuracy, and prunes the model to improve detection speed. Finally, the improved algorithm is applied to underwater robots to achieve automatic garbage detection underwater. The experiment shows that the improved algorithm has achieved significant improvements in both accuracy and speed.

Yao Lintao [10] proposed a lightweight network based on YOLOv4, aiming to deploy improved algorithms on mobile devices with limited resources. Replace the backbone of the original network with GhostNet, replace ordinary convolutions with Ghost modules, and finally introduce attention mechanisms. Experimental results have shown that the improved network model significantly reduces parameters and achieves an accuracy of 94.28%, successfully achieving deployment on Android mobile devices.

Xiao-yan wang and others studied the YOLOV3, RetinaNet and FasterRCNN application in the field of garbage detection identification effect, through the experiment platform, in turn on homemade data set design experiment, get the final experimental results, by comparison, Faster RCNN detection effect is best, followed by RetinaNet, finally is YOLOv3, so the Faster RCNN to achieve garbage detection, complete the recycling of resources.

Zhao He Yue [11] improved YOLOv5 for night style scene and poor small target detection effect, and proposed DST-YOLOv5 target detection algorithm. Firstly, the Backbone benchmark network of the model is improved by dual-channel feature extraction and weighted feature fusion to improve the detection effect of night style data set; then, Swin Transformer module and multi-scale detection head are introduced to enhance the global correlation of features and the acquisition of shallow features to improve the effect of small target detection; and finally, the YOLOv5 improves and is compared with the current popular target detection algorithm. The experimental results show that DST-YOLOv5 achieved mAP of 92.9 and FPS of 172.3.

Liu Jinjun[12] proposed an improved YOLOv7 algorithm to realize the classification and detection of construction waste. The improved algorithm adopts content content-aware reassembly of features, and the upper sampling operator replaces the nearest interpolation operator in YOLOv7, thus improving the target detection accuracy; introduces distribution shifting convolution module to replace some traditional convolution of the head network of YOLOv7, which realizes the lightweight of the model. The results show that the mAP value of the improved algorithm reaches 90.7%, and the model calculation is only 96G. It has a high application value in the actual scene of construction waste classification and detection.

Research fields in European and American countries, A Seredkin et.al[13] proposed a garbage detection robot sorting system for conveyor belts, and R. Laszlo et al[14]proposed a vision based waste recycling sorting system. In the field of application, ZenRobotics first developed an intelligent robot used in the field of garbage detection and classification. In 2019, ABB developed a new generation of intelligent robot used in garbage classification scenarios, which was displayed at the Domestic Expo. As people pay more attention to environmental issues, more and more companies are involved in research on efficient garbage sorting robots, including Max-AI robots developed by American companies and FANUC robots developed by Japanese companies. These robots are able to sort garbage quickly and accurately because their core is based on machine vision technology.

In general, using YOLO algorithm to study the garbage classification can achieve a fast, efficient and accurate classification effect, which plays a positive role in promoting the promotion of garbage classification in various countries.

4 Summary

Yolo algorithm is efficient, accurate and real-time in garbage classification and identification. By dividing images into multiple grids and classifying each grid, abnormal areas can be quickly detected, so as to quickly identify the categories of garbage. Yolo algorithm has reliability and versatility and can be applied to in different types of garbage target detection applications[15] . However, Yolo algorithm also has some shortcomings, such as the effect is often not ideal when dealing with small targets, and household garbage often has some small targets, resulting in low accuracy of the algorithm. Moreover, the identification of targets by Yolo algorithm depends on the quality of the dataset, and for large amounts of unknown data, Yolo algorithm is prone to misjudg[16]. Combined with the literature research, the direction of several future research is discussed.

Faster detection speed and higher positioning accuracy. With the development of Internet of Things and edge computing, more and more application algorithms need to meet the real-time requirements [17]. By studying the deeper and wider neural network structure, more effective feature extraction method, and more refined target classification method, the positioning accuracy and detection efficiency are improved.

Light-weight and small-scale target detection algorithms. Starting with YOLOv3, the demand for lightweight algorithms for the first time[18]. Application-level requirements such as target recognition and capture of mobile robots, agricultural applications, etc. Although the current detection algorithms have surpassed humans in terms of detection accuracy, they are still below the human eye in terms of details and functionality, especially those that are light enough and small enough.

Multimodal fusion. Target detection can not only be carried out through image and video, but also through sound, infrared, radar and other sensors[19]. By integrating data from different sensors to achieve more accurate and comprehensive detection of targets, we can make information complementary, and realize error correction, diversity enhancement and real-time improvement.

Scene generalization. The performance in different scenarios may be different[20]. Through adaptive learning and other methods, the accuracy and robustness of the target detection algorithm in different scenarios can be improved, and the application generalization performance of target detection in various industries can be improved.

Attention mechanism. The introduction of attention mechanism can greatly reduce the detection cost[21] , pay attention to the useful information of the image, and quickly improve the detection performance of the detection algorithm.

References

[1] Zhang Tengze, Li Xujun, Rao Liming. 2023.Improving the YOLOv5 small object detection algorithm for remote sensing images. Computer age, 2023,(12):89-95.DOI:10.16644/j.cnki.cn33-1094/tp.2023.12.020.

[2] Wang Wanting. 2023. Research and improvement of YOLO algorithm in the application of small object detection.Qilu University of Technology .DOI:10.27278/d.cnki.2023.000615.

[3] Niu Yitong, Meng Deyao, Wu Ronghai, *et al.*2022 Comparative study of the garbage image classification model based on transfer learning . Journal of Dali University, 2022,7 (12): 8-14.

[4] Xu Chuanyun, Wang Ying, Wang Wenmin, *et al.*2022.Multi-level feature-weighted fusion algorithm for household waste image classification . Journal of Chongqing University of Technology (Natural Science), 2022,36 (09): 146-155.;

[5] Wu Jialian, Lu Kang, Li Hong.2022. Based on an improved YOLOV4 multi-stage lightweight recyclable waste detection algorithm. Journal of Hefei University (comprehensive edition), 2022,39 (05): 60-66.

[6] Xu Wei.2021. Research on the lightweight garbage detection algorithm based on convolutional neural network. Zhejiang Sci-TechUniversity, 2021.DOI:10.27786/d.cnki.gzjlg.2021.000300.

[7] Dong Xiaoxiao, Liu Zhenguo, Song Tengteng, *et al.* 2021.Real-time waste detection algorithm based on RB-YOLO.Smart Computer and Application, 2021,11 (10): 5-11 + 19.

[8] Wang Zao, Ma Jiying, Luo Xing, et.al. 2023.Target detection of water surface waste based on SPMYOLOv3.Computer system application, 2023,32(03):163-170.DOI:10.15888/j.cnki.csa.009001.

[9] Tian Manjun,Li Chardonnay, Kong Shihan, *et al.*2022.The visual detection algorithm for underwater waste cleaning robot based on improved YOLOv4 .Frontiers of Information Technology Electronic Engineering,2022,23(08):1217-1229.

[10] Yao Lintao, Geng Zhiqing. Improved lightweight garbage detection system based on YOLOv4. Information and Computer (theoretical edition), 2022,34 (11): 111-114.

[11] Zhao He Yue. 2023.Research on household waste detection algorithm based on deep learning.Nanjing University of Posts and Telecommunications, 2023.DOI:10.27251/d.cnki.gnjdc.2023.001811

[12] Liu Jinjun, Li Leilei, Zhang Haoyu, *et al.*2024.Construction waste classification detection based on an improved YOLOv7 algorithm. Journal of Environmental Engineering, 2024,18 (01): 270-279.

[13] Girshick R, Donahue J, Darrell T, *et al.*2014.Rich feature hierarchies for accurate object detection and semantic segmentation.Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2014:580-587.

[14] Redmon, Joseph etal.2016.You only look once:unified,real-time object detection.Proceedings of the IEEE Conference on Computer Vision and Pat-tern Recognition,2016:779-788.

[15] WANG C Y,BOCHKOVSKIY A,LIAO H.YOLOv7.2023.trainable bag-of-freebies sets new state-of-the-art for real-time object detectors.Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023:7464-7475.

[16] Ye Zhaobing, Duan Xianhua, Zhao Chu. 2023. Improved YOLOv3-SPP underwater target detection study . Computer Engineering and Application, 2023,59(6):231-240.YE Z B, DUAN X H, ZHAO C.Research on underwater target detection by improved YOLOv3-SPP.Computer Engineering and Applications,2023,59(6):231-240.

[17] Wang Jianbo, Wu Youxin. 2023. Improved YOLOv4-tiny helmet wear detection algorithm . Computer Engineering and Applications, 2023,59 (4): 183-190. WANG J B, WU Y X.Helmet wearing detection algorithm of improved YOLOv4tiny.Computer Engineering and Applications, 2023,59(4):183-190.

[18] ZHANG D Y,CHEN X Y,et al.2021.Smart YOLO:a light-weight real-time object detection network.Journal of Physics:Conference Series,2021,1757(1):012096.

[19] LI J C,WANG H Z,XU Y,etc.2021.Road object detection of YOLO algorithm with attention mechanism.Frontiers in Signal Processing,2021,5(1):9-16.
[20] MA Y J, ZHANG S H.2021.Feature selection module for CNN based object detector.IEEE Access, 2021, 9:69456-69466.
[21] HUANG Z C,WANG J L,FU X S,et al.2020.DC-SPP-YOLO:dense connection and spatial pyramid pooling based YOLO for object detection.Information Sciences,2020,522:241-258.