ISSN 2959-6157

A review of YOLO-based traffic sign target detection

Siteng Liu

Northeastern University at Qinhuangdao, Qinhuangdao, Hebei, 066004, China
Email: 202011437@stu.neu.edu.cn

Abstract:

YOLO (You Only Look Once) as an efficient target detection algorithm has significant advantages in the field of image recognition and traffic sign detection. The continuous development of autonomous driving technology needs to be supported by an algorithm that can quickly and accurately identify traffic signs, vehicles, pedestrians and other important objects on the road. By using the YOLO algorithm, we can achieve fast and accurate recognition of traffic signs, which is of great significance for improving the safety of autonomous driving technology. This study firstly introduces the general framework of YOLO series algorithms, including the network structure, introduces the development history of YOLO series and analyses the characteristics of each generation of algorithms, then discusses the application of YOLO algorithms in the field of traffic sign recognition, and finally summarizes the existing problems and puts forward a few points of possible optimization directions in the future.

Keywords: Autonomous Driving, Real-time Detection, YOLO Algorithm, Traffic Signs

1. Introduction

In recent years, automated car driving technology has been developing rapidly, and both semi-autonomous and fully automated driving systems need to interact with the outside world and make appropriate decisions in real time. In order for autonomous vehicles to drive safely in a variable road environment, autonomous driving systems must follow the same traffic rules as human drivers. One of the most important pieces of information includes traffic sign information, and the continuous development of computer vision technology gives us a whole new approach to solving this problem based on artificial intelligence.

YOLO (You Only Look Once) is a deep neural network-based target detection algorithm, the core idea of which is to transform the target detection problem into a regression problem by directly predicting the class and location of the target through a single convolutional neural network.YOLO divides the image into multiple grids, each of which is responsible for predicting a fixed number of bounding boxes as well as the probability of the class of these bounding boxes . In the prediction phase, YOLO only needs to perform one forward propagation to get both the category and position information of the target, thus it has a fast processing speed and achieves real-time target detection. Since this is a vision-based solution, it encounters problems arising from visibility limitations such as light variations, occlusion and camera viewpoints, since the type of lens (wide-angle or telephoto lens) can deform the object differently than in real life.

In this study, we will review the development history of YOLO series, introduce the advantages and significance of YOLO technology in traffic information target detection, and look forward to the future development trend of this technology in this field by analysing the challenges and difficulties of current applications.

2. YOLO series development history and its characteristics

YOLO (You Only Look Once) is a target detection technology that identifies objects in an image by looking at them once. Since its debut, YOLO technology has evolved through several versions, each of which has been continuously optimised and improved to increase detection accuracy and speed. This section describes the network architecture of the different versions of YOLO over the generations and the optimisations of the current version with respect to the previous one.

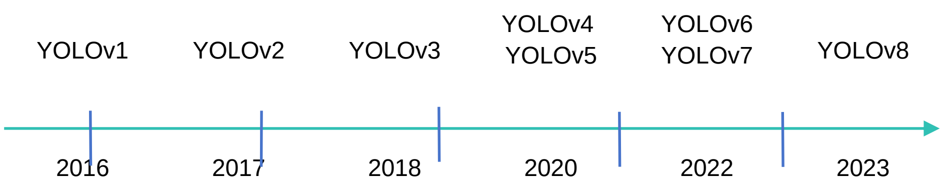

Figure 1 YOLO version timeline

2.1 *Overall YOLO framework*

As shown in the figure, the framework structure of YOLO is mainly based on Convolutional Neural Networks (CNNs) [1], and the overall framework structure can be divided into three phases: input, feature extraction and prediction.

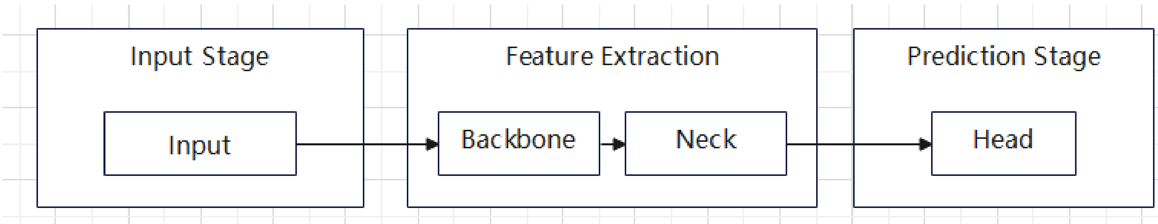

Fig.2 Architecture of modern object detectors

(1) Input Layer: receives images as input and usually uses standard size images for training and inference.
(2) Convolutional Layers: use convolutional operations to extract image features. These convolutional layers are usually used in conjunction with activation functions (e.g., ReLU) to introduce nonlinearities. In YOLOv3, DBL is the basic component which consists of convolution, bulk normalisation (BN) and Leaky ReLU activation functions [2].
(3) Pooling layer: reduce the spatial size of the feature map by downsampling operations, commonly used pooling methods are maximum pooling and average pooling. These operations help to extract higher level features and reduce the amount of computation.
(4) Darknet Layer: in YOLO, Darknet [3] is a series of modules consisting of convolutional and pooling layers for extracting multi-scale features. As YOLO versions evolve, the structure of Darknet is optimised, e.g. darknet-53 in YOLOv3.
(5) Fully Connected Layer: In earlier versions of YOLO, the fully connected layer was used to convert feature mappings into predicted values, including the bounding box coordinates and category probabilities of the target. However, as YOLO evolved, the fully connected layer was gradually replaced by convolutional layers for more efficient feature extraction and prediction.
(6) Output layer: outputs the detection results as bounding box coordinates and category probabilities of the target. In YOLO, the output processing consists of applying non-maximal suppression (NMS) to filter the overlapping bounding boxes [4] to obtain the final target detection results.

2.2 *YOLO development history*

Inspired by the structure of GoogLeNet [5] in 2016, Redmon et al [6] first proposed the YOLO target detection concept, where target detection can be accomplished through a single forward propagation for real-time processing.The YOLOv1 input divides the image into $s \times s$ grid regions, and 24 convolutional layers are used for extracting the image features, and 2 fully-connected prediction of bounding box and category probability layers are used to obtain $s \times s$ feature vectors. All vectors in the prediction stage are to be passed through the NMS and finally a box with the highest confidence is retained to generate the result.The YOLOv1 algorithm has a simple framework is fast and requires only one regression on the image and can meet the criteria for real-time detection. The loss function is simple and easy to train and adjust. However, for small and overlapping targets its accuracy is not high, the error is large, there is a leakage of false detection, the performance is not as good as advanced two-stage target detection algorithms such as R-CNN.

In 2017, based on YOLOv1 Redmon et al [7] proposed YOLOv2 that can detect more than 9000 categories, YOLOv2 improves on the problems of v1 version, and its performance is better in predicting small and overlapping targets.YOLOv2 is inspired by Visual Geometry Group [8] and Network -in-Network [9] to introduce Batch Normalisation and high resolution classifiers to improve the network structure. K-means clustering algorithm is added at the input side to determine the number of prediction frames, and the input image size is changed to make it perform better in small target detection, improving the detection accuracy and speed. The mean accuracy (mAP) on VOC12 was improved by 15.5%.

Since YOLOv2 a classifier can only detect a fixed number of objects and cannot handle variable number of objects, it requires a large amount of memory and computational resources, and it is difficult to train and tune the model. In 2018 Redmon et al [10] proposed YOLOv3 again with a deeper network structure Darknet-53.YOLOv3 was inspired by the Feature Pyramid Network (FPN) [11] inspired by adding a Neck layer between the backbone network and the prediction layer to avoid feature information loss. It also introduces multi-scale prediction that can detect three targets at the same time to improve the detection of different sized targets, and the average accuracy of YOLOv3 on VOC12 and VOC17 is 79.30% and 31.00%, respectively.

With the development of deep learning techniques, many new methods have been able to introduce the YOLO algorithm to improve the detection performance.In 2020, based on the YOLOv3 version, Bochkovskiy et al [12] built YOLOv4 [13].YOLOv4 adopts the CSP (cross stage partial connections) module [14] of CSPDarknet53 as the base network, introducing CIoU (Complete-IoU) loss function [15] and Mish activation function [16], by optimising the training process in order to improve the detection accuracy and speed. Meanwhile, the SPP (Spatial Pyramid Pooling) module [17] is introduced in the Neck layer, and the FPN+PAN module [18] is used instead of the FPN module to further improve the detection accuracy and speed. The combination of these techniques has enabled YOLOv4 to achieve excellent performance in several target detection tasks. The disadvantage is that it may lead to an increase in computational resource requirements due to the adoption of more complex network structures and algorithms.

YOLOv5 [19] adopts a lightweight network structure and a new training strategy to achieve faster detection and high accuracy target localisation. Mosaic data enhancement is introduced to strengthen the effect on small target detection.YOLOv5 pays special attention to the performance in real application scenarios, and therefore focuses on the real-time and ease of use of the model while maintaining high accuracy. The disadvantage of YOLOv5 is that it adopts a shallow network structure, which affects its detection accuracy in some complex scenes.

In 2022, Li et al [20] proposed YOLOv6, which is mainly oriented to industrial applications, and it focuses on both detection accuracy and inference efficiency to meet the industrial demand for real-time object detection.YOLOv6 optimally designs a more concise and effective Efficient Decoupled Head, which maintains the accuracy while further reducing the general decoupled head that imposes an additional latency overhead, adopting Anchor-free no longer using anchor frame assistance, while using SimOTA [21] label assignment strategy and SIoU [22] bounding box regression loss to further improve the learning effect. In the same year, Alexey et al. designed YOLOv7, which employs a lightweight optimisation strategy for better arrangement on mobile devices, and has been widely used in the field of autonomous driving and UAV applications. In terms of architecture, YOLOv7 proposes an extended E-ELAN based on ELAN (Efficient Layer Aggregation Network) [23]. The new E-ELAN does not change the gradient transfer path of the original architecture at all, and it uses group convolution to increase the base of added features and combines different groups of features by shuffle and merge cardinality, an operation that enhances the features learned from different feature maps.

Table 1 Detection results of different YOLO versions

YOLO version		dimension	FPS/s	mAP/%
YOLOv1	PASCAL VOC07	448		63.40
YOLOv2	PASCAL VOC07	416		78.60
YOLOv3	MS COCO17	416	35	31.00
YOLOv4	MS COCO17	608	65	43.50
YOLOv5	MS COCO17	640	170	41.20
YOLOv6	MS COCO17	640	98	52.80
YOLOv7	MS COCO17	1280	16	56.80

3. YOLO in the field of traffic sign recognition

YOLO target detection technology has been widely used in various fields, including construction, medical, agriculture, autonomous driving, drones, etc. The rise of YOLO technology in traffic sign recognition research mainly stems from the continuous development of autonomous driving technology, and YOLO technology enables real-time target detection, which is crucial for autonomous driving vehicles. With the increasing complexity of traffic networks, autonomous driving systems need to recognise traffic signs quickly and accurately in various situations to ensure driving safety. At the same time, traffic authorities need an efficient and reliable way to monitor and maintain traffic signs on the road.

3.1 *Early traffic sign target detection*

Early traffic sign target detection is mainly divided into colour-based detection algorithms and shape-based target detection algorithms. In the study of colour-based detection algorithms, how to separate the traffic signs from the background image is the key issue in the study.In 2003, Benallal et al [24] split each colour component in the

RGB colour space according to the gradient difference, so as to separate the traffic signs and the background colour. However, the image information in RGB colour space is easily affected by external lighting conditions, so in 2007 Zhu Shuangdong et al [25] shifted the research field from RGB colour space to HSI colour space, which is a colour saturation space that can effectively eliminate the influence of external ambient lighting image information [26], the disadvantage is that the HSI colour space is easily affected by noise. In 2012 Chang et al [27] introduced the method of backtracking intra-group standard deviation on the basis of this principle to improve the performance of this detection method.

Shape based traffic sign target detection is less likely to be affected by the lighting environment compared to colour based target detection.Hoffman Transform[28] is the most commonly used method for shape based target detection, which takes advantage of the fact that all traffic sign shapes are regular shapes such as triangles, quadrilaterals, etc.However, the method is computationally expensive.In 2007 Paulo et al[29] set up multiple controllable regions, using Harris point predictor to predict the shape features in the region and search for regular shape targets.In 2015 Tang Kai et al[30] proposed fusion multi-scale algorithm, which fuses colour information, shape information, size information and other aspects for target detection. Chang Faliang et al [31] integrated RGB colour space and HIS colour space for extracting traffic sign information from image information, this algorithm performs better than single feature target detection.

In 2014 Girshick et al [32] proposed a candidate target region based detection algorithm RCNN This algorithm can accurately detect the target object based on the features extracted from the candidate region, and at the same time can complete the semantic segmentation task based on the classification regression bounding box.He et al [33] optimised on the basis of RCNN by adding SPP layer space in the middle of convolutional and fully connected layers pyramid pooling structure, this algorithm no longer extracts features from each candidate frame, and its computing speed is accelerated by tens of times. In the same year Girshick et al [34] proposed Fast R-CNN on this basis to further optimise the algorithm.

3.2 *YOLO-based traffic sign target detection*

The rise of deep learning provides a completely new research direction for image recognition, deep learning has a strong feature learning ability, due to the internal structure contains multiple neural network layers, theoretically can be covered to any function, the larger and richer the dataset is, the better it performs, and at the same time, adjusting the internal parameters can also improve its recognition ability.

In 2016 Redom et al [35] designed the YOLO network, which requires only a single forward propagation to complete the target detection task.The YOLO network has a significant speed improvement compared to the RCNN network, but the first generation version is not effective in detecting small targets.In 2017 Zhang et al. applied the YOLO algorithm to traffic sign detection.Compared to other?target detection algorithm time was reduced by 0.017 seconds [36-38]. In the same year the team made public the CCTSDB traffic sign target detection dataset [39].Benjumea et al [40] modified the Neck layer and network structure to improve the accuracy of small target detection.In 2018 Yu et al [41] designed a fusion model based on the YOLOv3 and VGG19 networks to detect traffic signs in autonomous driving using the sequential relationship between the images, and its accuracy of detecting traffic signs is more than 90%.Dewi et al [42] augmented training of existing dataset by generating traffic sign images through generative adversarial network and the average accuracy on YOLOv3, YOLOv4 model is 84.9%, 89.33%.Gao et al [43] conducted test experiments on YOLOv4 for traffic sign sign signals and the results show that YOLOv4 has better image recognition performance. Pan Huiping et al [44] used YOLOv4 in combination with the SPP module to lift the original network's limitation on image size, modified the size of the source image captured by the car recorder in order to increase the receptive field, and greatly improved the detection of traffic signs with an average accuracy of 99.0%. In the same year, Natarajan et al [45] proposed a convolutional neural network consisting of four branched weights to identify traffic signs, this network structure has lower complexity and fewer parameters compared to the network structure used in previous studies, which accelerated the recognition speed.In 2019 Li et al [46] replaced the n × n convolution of YOLOv4 with n × 1 and 1 × n convolutions to reduce the convolutional layer complexity, reducing the computational cost, and this classifier has better performance in recognising traffic signs.Qian et al [47] used composite data enhancement to augment the model inputs, and designed a multi-scale YOLOv5 feature fusion network, which improves the model's performance of detecting and processing traffic signals.Mao et al [48] optimised the architecture of YOLOv7 based on the introduction of the SIoU loss function to improve YOLOv7's ability to detect traffic signs.2022 Omran Nacir et al [49] used a migration learning approach from classification to detection, pre-processed the training images with data enhancement, added luminance variations to simulate different lighting conditions, saturation to increase the effect of old and new signs, position and scale variations, and finally introduced clipping to simulate different levels of occlusion.2022 Aisha Batool et al [50] proposed an Extreme Learning Machine based traffic sign detection method, which set a new record for the detection accuracy of the GTSRB, advancing the progress of research in this field.

4. Summary and outlook

With the continuous development of YOLO series technology, the traffic sign target detection technology based on YOLO has been relatively mature and has been applied in real autonomous driving scenarios. However, vehicles equipped with ITS may encounter complex situations during road travelling, and there are still many difficulties that need to be solved in this field, among which several major difficulties need to be solved as follows:

(1). Impact of the natural environment on traffic signs

As most of the traffic signs placed outdoors, so it is very vulnerable to extreme weather and other impacts lead to traffic sign damage, on both sides of the road for a long time in the wind and rain, the sign itself will have varying degrees of discolouration and aging, which will inevitably increase the difficulty of detecting and identifying traffic signs, light irradiation varying in intensity will increase the difficulty of detecting and identifying traffic signs, when faced with the light when the traffic signs will be reflective, when the back of light when the When facing the light traffic signs will be reflective, when the back to the light when the colour will be dark, in the night driving will also encounter such problems.

(2). Effects of filming while the vehicle is in motion

Drivers driving vehicles, due to some sections of the road situation is unknown so driving speeds fast and slow, encountered a bumpy road when the vehicle-mounted recorders captured images may be blurred, or even deformation, the uncertainty of these images captured while driving the vehicle will increase the difficulty of detecting and identifying the traffic signs themselves may also be due to certain physical factors leading to the shape of the rotating or deformed, will affect the vehicle-mounted recorders to detect and identify the traffic signs. Recorder detects and identifies traffic signs.

(3). Shading and the impact of small targets

Vehicles travelling on the highway are likely to be affected by such as tree branches falling onto the traffic signs, traffic signs placed in a crowded mutual blocking, people walking to cover the traffic signs and other effects, and the traffic signs themselves in the shape of the volume is small, all of which will affect the effectiveness of traffic sign detection and recognition.

Possible future optimisation directions directions for this family of algorithms are:

(1). Higher recognition accuracy and real-time performance. With the deep development of deep learning technology, the performance of YOLO series algorithms is expected to be further improved. By optimising the network structure and improving the training strategy, the accuracy and real-time performance of traffic sign recognition can be further improved to meet more complex traffic environments and higher application requirements.

(2). Multimodal data fusion. In addition to image data, multimodal data fusion can be combined with other sensor data (e.g., radar, lidar, etc.) to improve the robustness and accuracy of traffic sign recognition. This multi-source information fusion approach can better cope with complex conditions such as bad weather and light changes.

(3). Optimisation for specific scenarios. The YOLO algorithm can be customised and optimised for the characteristics of different traffic scenarios (e.g., city roads, highways, country roads, etc.). By collecting traffic sign data for a specific scene, targeted training and tuning can improve the recognition performance of the algorithm in that scene.

(4). Lightweight model design. With the development of IoT and edge computing, there is an increasing demand for lightweight and efficient traffic sign recognition models. Future research can be devoted to designing more lightweight YOLO models to reduce model size and computational complexity, making them more suitable for running on resource-limited devices.

(5). Interactive and dynamic traffic sign recognition. Future traffic systems may introduce more interactive and dynamic traffic signs whose information may change over time. Therefore, the YOLO algorithm needs to be able to adapt to this dynamism by recognising and interpreting these changing sign messages in real time.

REFERENCES

[1] CAO J L, LI Y L, SUN H Q, et al. A review of visual object detection technology based on deep learning [J]. Chinese Journal of Image and Graphics,2022, 27(06):1680-1730

[2] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models[C]//Proc. icml. 2013, 30(1): 3.

[3] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[4] NEUBECK,GOOL L.Efficient non-maximum suppression[C]//International Conference on Pattern Recognition,2006:840-852.

[5] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 2-10.

[6] REDMON J,DIVVALA S,GIRSHICK R,et al.You only

look once:unified,real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016:780-790.

[7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 6520-6525.

[8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-5-1) [2023-5-29].https://ar xiv.org/pdf/1409.1556.pdf.

[9] LIN M, CHEN Q, YAN S C. Network in network [EB/OL]. (2013-12-16) [2023-5-29]. https://arxiv.org/pdf/1312.4400.pdf.

[10] REDMON J,FARHADI A.Yolov3:an incremental improvement[J].arXiv:1804.02767,2018.

[11] LIN T Y, DOLLÁR P, CIRSHICK R, et al. Feature pyramid networks for object detection [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017:940-944.

[12] BOCHKOVSKIY A,WANG C Y,LIAO H Y M.Yolov4:optimal speed and accuracy of object detection[J]. arXiv:2004.10934,2020.

[13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. (2020-4-23) [2023-5-29]. https: //arxiv.org/pdf/2004.10934.pdf.

[14] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C]// Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2020: 1571-1580.

[15] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [EB/OL]. (2019-11-19) [2023-8-7]. https://arxiv.org/pdf/1911.08287.pdf.

[16] MISRA D. Mish: a self regularized non-monotonic activation function [EB/OL]. (2019-8-23) [2023-5-29].https://arxiv.org/pdf/1908.08681.pdf.

[17] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(9): 1904-1916.

[18] LIN T Y, DOLLÁR P, CIRSHICK R, et al. Feature pyramid networks for object detection [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017:1036-1044.

[19] JOCHERG.Yolov5[EB/OL].[2023-03-20].http//github.com/ultralytics/yolov5.

[20] LI C Y,LI L L,JIANG H L,et al.Yolov6:a single stage object detection framework for industrial applications[J]. arXiv:2209.02976,2022.

[21] GE Z, LIU S T, LI Z M, et al. OTA: optimal transport assignment for object detection [C]// Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 305-310.

[22] GEVORGYAN Z. SIoU loss: more powerful learning for bounding box regression[EB/OL].(2022-5-25)[2023-6-1].https://arxiv.org/pdf/2205.12740.pdf

[23] Li C,Zhou A, Yao A.Omni-dimensional dynamic convolution[J].Computer Vision and Pattern Recognition,2022.

[24] Benallal M, Meunier J. Real-time color segmentation of road signs[C]. Electrical and Computer Engineering. Toronto, Canada, 2003: 1823-1826.

[25] Zhu S D,Liu L L,LU X F, et al. Color-geometric model for traffic sign recognition[J].Chinese Journal of Scientific Instrument.2007,28(5):954-958.

[26] VAZQUEZ-CORRAL J, GALDRAN A, CYRIAC P. A fast image dehazing method that does not introduce color artifacts[J]. Journal of RealTime Image Processing, 2020, 17(3): 607-622

[27] SHEN Z H,FU M Y,YANG Y, et al. Traffic sign image segmentation method based on background pixels mutation analysis[J]. Application Research of Computers. 2012, 29(9): 3531-3535.

[28] Loy G, Barnes N, Shaw D, et al. Regular polygon detection[C]. IEEE International Conference on Computer Vision. San Diego. CA, USA , 2005: 778-785.

[29] Paulo C F,Correia P L.Automatic detection and classification of traffic signs[C].Eighth International Workshop on Image Analysis for Multimedia Interactive Services. Santorini,Greece,2007: 11-14.

[30] Tang K,Li Shiying,Liu Juan,et al. Traffic Sign Detection Based on Multiple Features Cooperation[J]. Computer Engineering,2015,41(3):211-217.

[31] Chang Faliang1,Huang Cui1,Liu Chengyun1,et al. Traffic sign detection based on Gaussian color model and SVM[J]. Chinese Journal of Scientific Instrument. 2014, 35(1): 43-49.

[32] Girshick R,Donahue J,Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C].IEEE conference on computer vision and pattern recognition. Columbus,OH,USA,2014: 580-590.

[33] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence. 2015, 37(9): 1904- 1916.

[34] R. Girshick. Fast R-CNN[C]. IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440-1450.

[35] Redmon J, Divvala S, Girshick, R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 780–790.

[36] Zhang J, Xie Z, Sun J, et al. A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection[J]. IEEE access, 2020, 8: 29740-29760.

[37] Zhang J, Wang W, Lu C, et al. Lightweight deep network for traffic sign classification[J]. Annals of Telecommunications, 2020, 75(7): 369-379.

[38] Li Xudong, Zhang. Jianming,et al.A Fast Traffic Sign Detection Algorithm Based on Three-Scale Nested Residual Structure[J].Journal of Computer Research and Developme nt,2020,57(05):1022-1036.
[39] Zhang J, Huang M, Jin X, et al. A real-time chinese traffic sign detection algorithm based on modified YOLOv2[J]. Algorithms, 2017, 10(4): 127.
[40] BENJUMEA A, TEETI I, CUZZOLIN F, et al. YOLO-Z: improving small object detection in YOLOv5 for autonomous vehicles [EB/OL]. (2021-12-22) [2023-6-3]. https://arxiv.org/pdf/2112.11798.pdf.
[41] YU J, YE X, TU Q. Traffic sign detection and recognition in multiimages using a fusion model with YOLO and VGG network [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(9):16630-16642.
[42] DEWI C, CHEN R C, LIU Y T, et al. Yolo V4 for Advanced Traffic Sign Recognition with Synthetic Training Data Generated by Various GAN[J]. IEEE Access, 2021, 9: 97228-97242.
[43] GAO H, WANG W H, YANG C J, et al. Traffic Signal Image Detection Technology Based on YOLO[C]// Journal of Physics: Conference Series, 2021, 1961(1).
[44] PAN H P, WANG M Q, ZHANG F Q. Traffic Sign Detection and Recognition Method Based on Optimized YOLO-V4[J]. Computer Science, 2022, 49(11): 179-184.
[45] Natarajan S, Annamraju A K, Baradkar C S, Traffic signrecognition using weighted multiconvolutional neuralnetwork[J]. IET Intelligent Transport Systems. 2018, 12(10): 1396-1405.
[46] Li Jia, Wang Zengfu. Real-time traffic sign recognitionbased on efficient CNNs in the wild[J].IEEE Transac-tions on Intelligent Transportation Systems. 2019, 20(3): 975-984.
[47] QIAN W, WANG G Z, LI G P. Traffic Light Detection Algorithm Based on Multi Scale YOLOv5[J]. Software Guide, 2022, 21(9): 19-25.
[48] MAO K J, JIN R H, YING L K, et al. SC-YOLO: provide application level recognition and perception capabilities for smart city industrial cyber-physical system [J]. IEEE Systems Journal, 2023:1-12.
[49] O.Nacir, M.Amna, W.Imen and B.Hamdi,"Yolo V5 for Traffic Sign Recognition and Detection Using Transfer Learning," 2022 IEEE International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), Tunis, Tunisia, 2022:1-4.
[50] Batool A, Nisar M W, Shah J H, et al. iELMNet: integrating novel improved extreme learning machine and convolutional neural network model for traffic sign detection[J]. Big data, 2022.