Paper ID: CS852 The 4th International Conference on Applied Mathematics Modeling and Computer Simulation (AMMCS 2024)

Truck recognition under different lighting conditions based on the YOLOv7-SCI

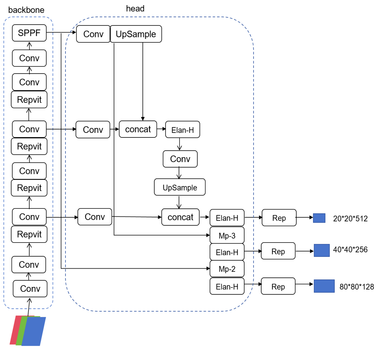
Li Ruixi[a] *, **Li Haibo**[a]

[a] School of Mechanical Engineering, Shanghai, 201620, China

*Corresponding author's email: 409162674@qq.com

Introduction

Yolov7-SCI general framework



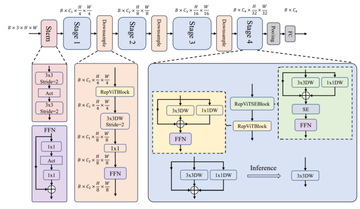
Comparison between original image and SCI data enhanced image

a.Original Image b.Enhanced Image

Research objectives

- How to improve the recognition of large trucks based on different lighting conditions
- The RepVit module is added to the Backbone of YOLOv7 to improve the detection accuracy of the model a nd enhance the generalization ability and adaptability of the model

Neural network module

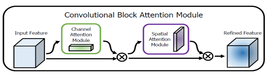


The structure diagram of Repvit module

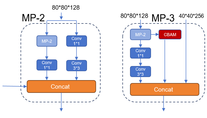
The RepVit module, through its ability to dynamically adjust the size of the conv olution kernel, can effectively capture multi-scale features in images. Adding Rep Vit to the Backbone of YOLOv7 can make the model more sensitive to object dete ction at different scales, which is particularly important for handling objects with multiple sizes. This not only helps improve the detection accuracy of small targets, but also enhances the understanding ability of dense target scenes.

The introduction of RepVit module helps alleviate the gradient vanishing or expl oding problem in YOLOv7 during training, as dynamically adjusted convolution k ernels can better adapt to data distribution, thereby improving training stability and convergence speed. This enables YOLOv7 to achieve higher detection accuracy m ore quickly when processing complex datasets.

Attention mechanism



The structure diagram of CBAM module



Structure comparison between the original MP-2 and the softer CBAM module MP-3

Activation function

$$L_{wIoU} = re^{\frac{(x-x_{gt})^2+(y-y_{gt})^2}{W_g^2+H_g^2}} L_{IoU}$$

$$L_{IoU} = 1 - \frac{W_i H_i}{wh + w_{gt}h_{gt} - W_i H_i}$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}}$$

$$\beta = \frac{L_{IoU}}{\overline{L_{IoU}}} \in [0, +\infty)$$

Formulas in the derivation of the wiou loss function

Result

Comparison of results of different target detection

Model	Recall rate	AP50
Yolov5_7.0	90.98	93.8
YOLOv7	86.9	89.5
Improved YOLOv7	96.5	98.9
Yolov8	88.7	95.0
Fast_rcnn	83.47	85.36
Yolov9	90.2	93.2

Comparison of ablation experiment results

	YOLOv7	RepViT	WIoU	CBAM	Number of parameters (in ten thousands)	MAP0.5%	Recall%	Precision%
1	✓				376	89.5	86.9	80.5
2	✓	✓			539	90.4	87.6	83.6
3	✓	✓	✓		539	91.3	86.8	82.7
4	✓	✓		✓	539	92.5	89.2	85.6
5	✓	✓	✓	✓	530	98.9	96.5	97.4

yolov7 yolov7-SCI

Conclusions

- A YOLOV7-SCI model based on YOLOv7 network structure is proposed.the SCI self-calibrated lighting frame is used to pre-process the pictures for the road section with low light.
- The RepViT module was used as the backbone layer feature extraction network to replace the original ELAN module, and the CBMA attention mechanism was integra ted into the neck network to replace the loss function in the original network module with Wise_iou.Compared with the original yolov7 model, the average accuracy (AP) of truck detection is increased by 8.2%, and the recall rate is increased by 6.1%.

Reference

[1] C. P. Papageorgiou, M. Oren and T. Poggio, "A general framework for object detection," Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), B ombay,India, 1998, pp. 555-562, doi: 10.1109/ICCV.1998.710772.

[2] Harris, Christopher G. and M.J. Stephens. A Combined Corner and Edge Detector.Alvey Vision Conference(1988).

[3] REN Jie,LI Gang,ZHAO Yan-Jiao.et al.Detection of Urban Trucks Based on Improved Faster RCNN[J]. ComputerSystems&Applications,2022,31(12):316-321.

[4]Wang A, Chen H, Lin Z, et al.Repvit: Revisiting mobile cnn from vit perspective. arXiv preprint arXiv:2307.09283(2023)

[5] 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.

[6]Vaswani A,Shazeer N,Parmar N,etal.Attention Is All You eed[J].arXiv,2017.DOI:10.48550/arXiv.1706.03762.