

基于 YOLOv5 的交通标志识别^①

郑红彬, 宋晓茹, 刘 康

(西安工业大学 电子信息工程学院, 西安 710021)
通信作者: 宋晓茹, E-mail: masha0422@163.com



摘 要: 交通标志识别是自动驾驶技术中的关键一部分. 针对交通标志在道路场景中目标较小且识别精度较低的问题, 提出一种改进的 YOLOv5 算法. 首先在 YOLOv5 模型中引入全局注意力机制 (GAM), 提高网络捕获不同尺度交通标志特征的能力; 其次将 YOLOv5 算法中使用的 GIoU 损失函数更换为更具回归特性的 CIoU 损失函数来优化模型, 提高对交通标志的识别精度. 最后在 Tsinghua-Tencent 100K 数据集上进行训练, 实验结果表明, 改进后的 YOLOv5 算法对交通标志识别的平均精度均值为 93.00%, 相比于原算法提升了 5.72%, 具有更好的识别性能.

关键词: YOLOv5; 交通标志识别; GAM; CIoU

引用格式: 郑红彬, 宋晓茹, 刘康. 基于 YOLOv5 的交通标志识别. 计算机系统应用, 2023, 32(8): 230-237. <http://www.c-s-a.org.cn/1003-3254/9204.html>

Traffic Sign Recognition Based on YOLOv5

ZHENG Hong-Bin, SONG Xiao-Ru, LIU Kang

(School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China)

Abstract: Traffic sign recognition is a key part of autonomous driving technology. Given the problems of small targets and low recognition accuracy of traffic signs in road scenes, an improved YOLOv5 algorithm is proposed. First, the global attention mechanism (GAM) is introduced into the YOLOv5 model to improve the network's ability to capture traffic sign features of different scales. Second, the GIoU loss function used in the YOLOv5 algorithm is replaced with the CIoU loss function which is more regressive to optimize the model and improve the recognition accuracy of traffic signs. Finally, the training is carried out on the Tsinghua-Tencent 100K dataset. The experimental results show that the average accuracy of the improved YOLOv5 algorithm for traffic sign recognition is 93.00%, which is 5.72% higher than that of the original one, indicating that the improved algorithm has better recognition performance.

Key words: YOLOv5; traffic sign recognition; global attention mechanism (GAM); CIoU

近年来, 为了节能减排、减少环境污染和一些不必要的人为失误, 对自动驾驶的研究受到了众多研究学者的青睐, 而对交通标志识别的研究是自动驾驶中的关键一部分, 在自动驾驶和智能交通系统等方面有着重要的研究价值和意义^[1]. 随着计算机技术和深度学习理论的发展, 涌现出了很多交通标志识别算法.

基于传统的交通标志识别方法通常以手工提取特征为主, 如 Sobel 边缘检测特征^[2]、Haar 特征^[3]、Hog

特征^[4]等, 此类算法泛化能力较弱, 提取效果欠佳, 时间复杂度较高, 不适合在对实时性要求高的交通标志识别场景中应用.

近些年, 基于深度学习的方法在大多数先进的目标检测算法都得到了使用, 并在目标检测任务中取得了丰硕的成果, 其中 YOLO 算法因其良好的识别性能在交通标志识别领域备受瞩目. Redmon 等人在 2016 年提出 YOLO 算法^[5], 将目标检测转化为回归问题, 使得

① 基金项目: 陕西省重点研发计划 (2021GY-287)

收稿时间: 2023-02-09; 修改时间: 2023-03-08; 采用时间: 2023-03-21; csa 在线出版时间: 2023-06-09

CNKI 网络首发时间: 2023-06-09

检测速度得到提升^[6]。之后两年里 Redmon 等人针对 YOLO 算法加以改进, 将其不断更新至 YOLOv3^[7], YOLOv3 采用 FPN^[8] 结构实现了多尺度预测, 在各项性能上均有了较大的提升, 但是检测精度与实时性方面仍有所欠缺。王浩等人^[9] 提出了基于 YOLOv3 改进的实时交通标志检测算法, 采用跨阶段局部网络作为特征提取模块且以路径聚合网络替代特征金字塔网络, 以此来提升实时检测要求和检测精度。Wan 等人^[10] 在 YOLOv3 基础上进行网络剪枝和损失函数改进等优化策略来解决小交通标志检测问题。2020 年, Bochkovskiy 等人提出了 YOLOv4 模型^[11], YOLOv4 在特征提取上采用 CSPDarknet53 网络^[12], 同时应用 SPP 模块增强特征表达能力, 使其在检测的实时性和精度方面均有很高的性能, 但模型的复杂度也随之增加。郭继峰等人^[13] 提出的改进 YOLOv4 的交通标志识别算法, 使用深度可分离卷积与 focal 损失函数来解决模型参数量大和正负样本数量不均衡的问题, 在 TT100K 数据集上测试的平均精度均值为 87.5%, 相比 YOLOv4 提升了 3.9%。尹宋麟等人^[14] 等在 YOLOv4 主干网络特征层后加入 ECA 模块并且利用 K-means 聚类算法重新生成网络先验框, 这些优化策略在 TT100K 数据集上的平均精

度均值达到了 84.95%, 比 YOLOv4 提升了 4.58%。

以上算法及其改进在一定程度上对交通标志的识别精度均有所提高, 但不可避免的引入了过高的计算复杂度与参数量, 使其在交通标志实时识别方面仍有欠缺, 然而 YOLOv5 在 YOLO 系列中识别速度与识别精度均为最优且更适合实际工程的应用, 所以本文使用 YOLOv5 模型作为基准模型。考虑到交通标志图像主要来自于车载摄像头, 交通标志在图像中的像素占比较小, 因此本文在 YOLOv5 的基础上引入全局注意力机制来提升对不同尺度目标的识别能力, 同时采用 CIoU 损失函数来解决 GIoU 损失函数收敛速度较慢的问题, 提升了对交通标志的识别能力, 最后在 TT100K 数据集上进行实验验证本文改进的 YOLOv5 算法的有效性。

1 YOLOv5 模型结构分析

YOLOv5s 是 YOLO 系列中最小的型号, 模型体积小节约部署时间, 识别速度快实时性好, 且其使用 PyTorch 框架更容易投入生产, 所以本文以 YOLOv5s 为基准网络进行改进。图 1 所示为 YOLOv5 网络结构。

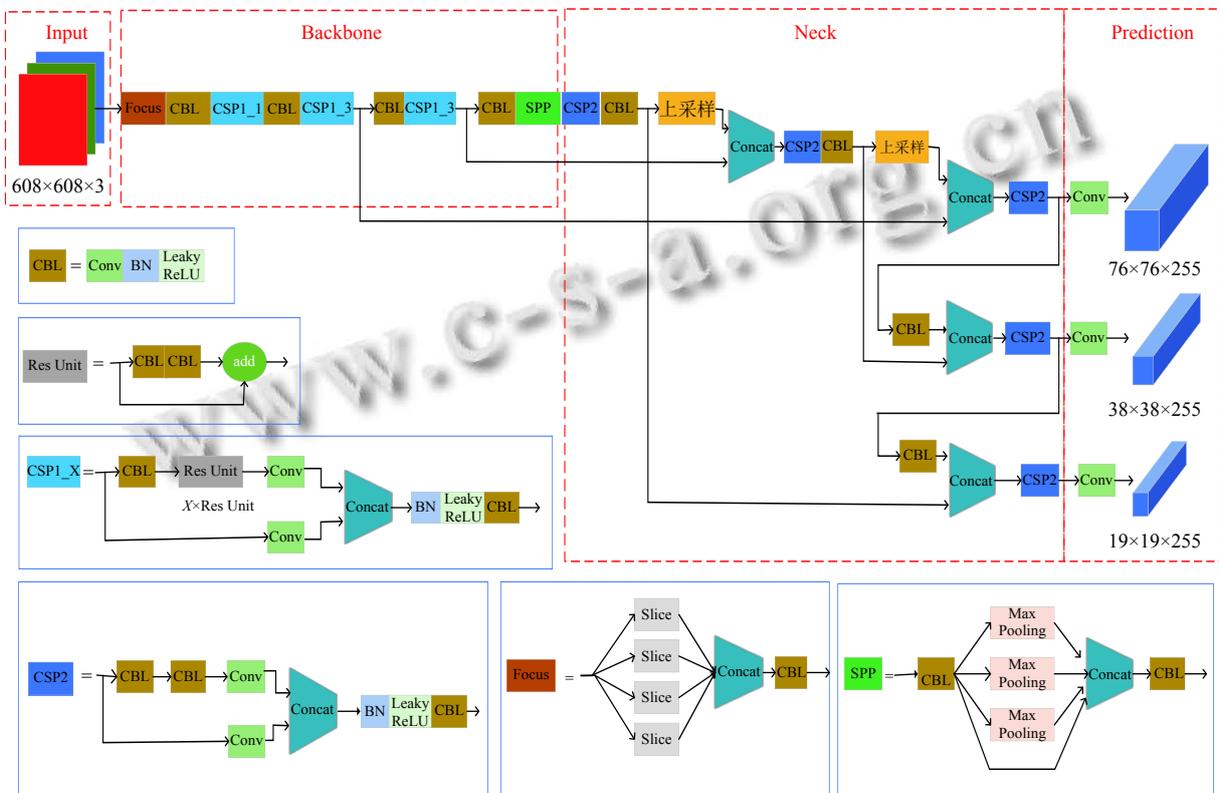


图 1 YOLOv5 网络结构

YOLOv5 的框架由 Input、Backbone、Neck 和 Prediction 这 4 个部分组成^[15]。Input 部分通过使用 Mosaic 数据增强方法来扩充数据集和自适应锚框计算、自适应图片缩放加强对图像的处理能力。Backbone 部分由 Focus 结构、CSP 结构和 SPP 模块构成。其中 Focus 的作用是图片进入 Backbone 时对其使用切片操作，最终得到没有信息丢失的特征图；CSP 结构将原输入一分为二，一个分支进行卷积操作，另一个分支经过多个残差结构，使得模型学习更多的特征；SPP 模块的作用是将全局信息和局部信息进行融合从而丰富特征图。Neck 部分由 FPN+PAN^[16] 结构组成，主要用于聚合图像的高低层特征，提升目标检测的效果。最后，网络通过 Prediction 进行目标预测和输出。

2 YOLOv5 模型的改进

针对传统 YOLOv5 算法，提出以下策略来优化其对交通标志的检测性能。注意力机制可以有效关注到图像中包含交通标志的区域，有助于提升对较小交通

标志的识别性能；损失函数的改进使得模型收敛速度加快，识别精度也得到一定提升。

2.1 注意力机制的引入

由于交通标志在道路中分布比较分散且目标较小，为了更好地识别不同尺度的小目标特征，选择在对交通标志的识别中引入全局注意力机制 (global attention mechanism, GAM)^[17]，使其对存在交通标志的区域有选择性的关注，同时也可以抑制一些无关信息，提高对远距离交通标志的识别性能，从而做出更加准确的判断。改进后的 YOLOv5 网络结构如图 2 所示。

GAM 与 CBAM 相似，采用通道-空间注意力机制，并在此基础上重新设计子模块，整体模块结构如图 3 所示。其中， F_1 为输入， F_2 为中间状态， F_3 为输出， F_2 与 F_3 的定义如式 (1) 和式 (2) 所示， M_c 为通道注意力模块， M_s 为空间注意力模块， \otimes 表示逐元素乘法。

$$F_2 = M_c(F_1) \otimes F_1 \tag{1}$$

$$F_3 = M_s(F_2) \otimes F_2 \tag{2}$$

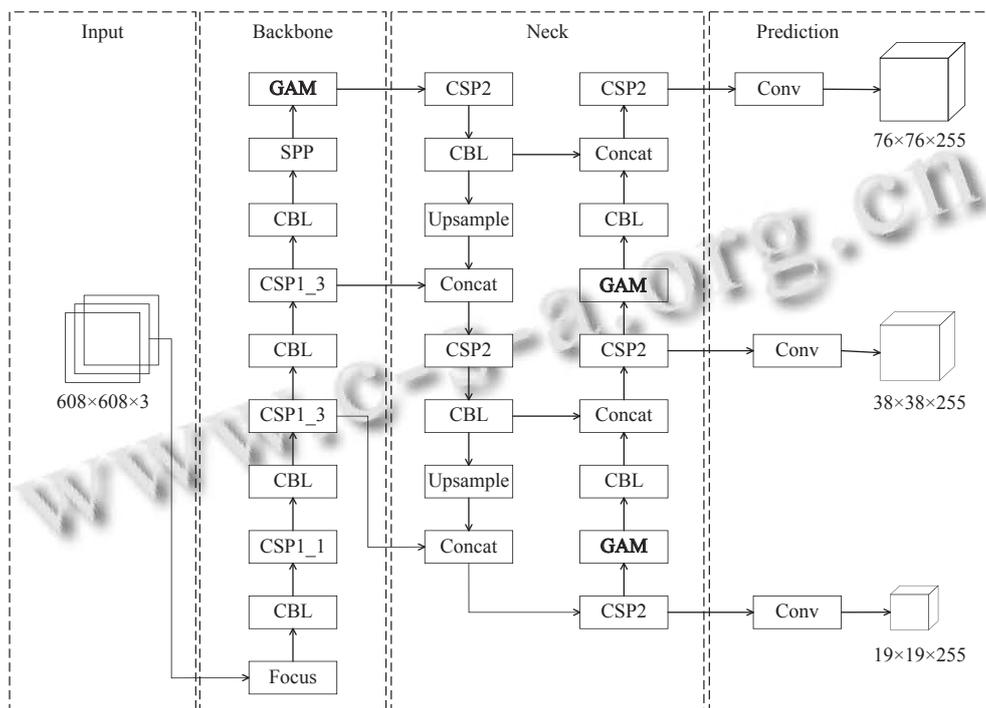


图 2 改进后的 YOLOv5 网络结构

(1) 通道注意力子模块

如图 4 所示，通道注意力子模块先使用三维通道置换来保留信息，然后使用两层多层感知器 (MLP)

进行维度转换，最后进行 Sigmoid 处理输出。

(2) 空间注意力子模块

如图 5 所示，在空间注意力子模块中，使用两个卷

积层进行空间信息融合,先使用 7×7 卷积缩减通道数(缩减率 r),再经过 7×7 卷积增加通道数量,保持通道数量的一致,最后进行 Sigmoid 处理输出.

2.2 损失函数的改进

损失函数作为衡量模型预测值与真实值之间差异程度的运算函数,其值越小,表明模型的鲁棒性越好,所以在训练过程中损失函数选择的合适与否尤为重要.

在传统 YOLOv5 的 prediction 部分中,使用 $GIoU^{[18]}$ 作为其损失函数, $IoU^{[19]}$ 表示的是预测框和真实框的交并比,是目标检测中常用的损失函数,其计算公式如下:

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (3)$$

$$L_{IoU} = 1 - IoU \quad (4)$$

其中, B 表示预测框, B^{gt} 表示真实框.

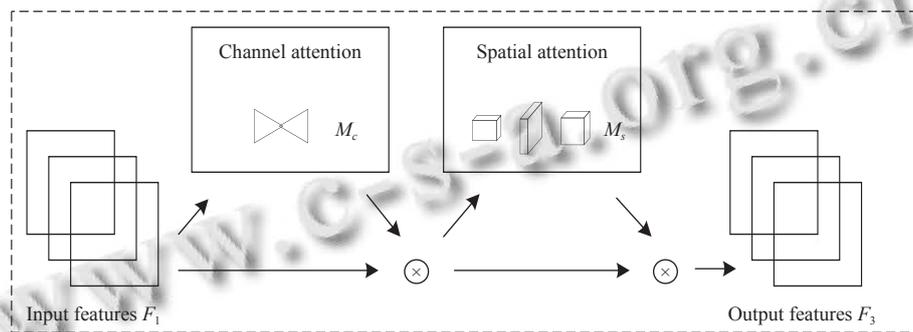


图3 GAM 模块

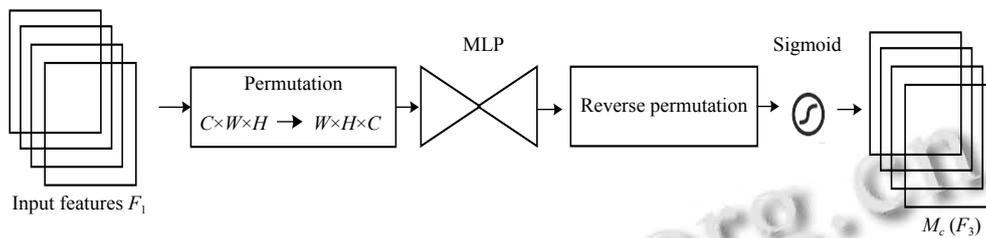


图4 通道注意力子模块

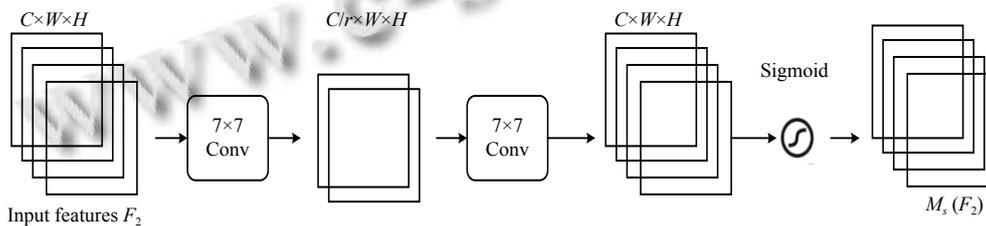


图5 空间注意力子模块

IoU 因其尺度不变性、对称性、同一性、非负性等优点,且输出的损失值介于 $0-1$ 之间,可以很好地表达预测框与真实框的检测效果,但在实际检测中会存在 IoU 损失函数无法使用的情况.

1) 如图 6(a) 所示,当预测框和真实框完全不相交

时, IoU 的输出值为 0 , 损失函数的输出为 1 , 但是从图中可以明显地发现预测框和真实框接近, 损失函数的输出应该更小.

2) 如图 6(b) 所示,当预测框和真实框的交并比相同时,计算出来的损失函数值相同,但不能判断哪种预

测框更加准确。

所以YOLOv5使用的是GIoU损失函数, GIoU在基于IoU特性的基础上引入最小外接框来解决上述IoU损失函数无法使用的情况1), 不仅关注重合区域, 还关注其他非重合区域. 其计算公式如下:

$$L_{GIoU} = 1 - IoU + \frac{|C - (B \cup B^{gt})|}{|C|} \quad (5)$$

其中, B 表示预测框, B^{gt} 表示真实框, C 为 B 和 B^{gt} 的最小包围框.

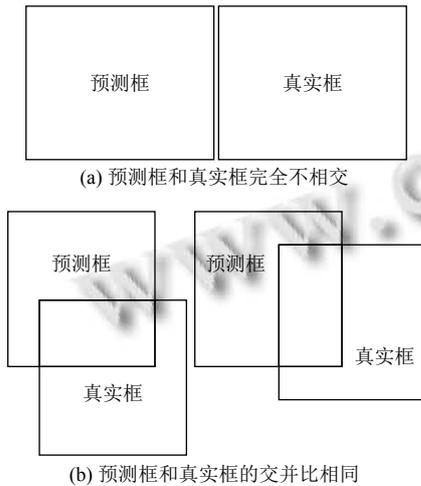


图6 IoU损失函数无法使用的情况

虽然GIoU损失函数有效解决了预测框与真实框不相交的情况, 但是当预测框包含在真实框内时, GIoU就会退化为IoU, 无法区分预测框和真实框的相对位置关系且对每个预测框和真实框而言均要计算最小包围框, 导致计算速度和收敛速度受到限制. 基于GIoU损失, CIoU^[20]损失考虑了边界框的重叠区域、中心点距离以及边界框纵横比的一致性. 所以在本文中, 使用CIoU损失取代GIoU损失, 损失函数定义为:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (6)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

其中, b 为预测框的中心点, b^{gt} 为真实框的中心点, ρ 是两个中心点之间的欧氏距离, c 是最小包围框的对角线长度, α 是权重参数, γ 用来衡量纵横比的一致性, w 和 h 、 w^{gt} 和 h^{gt} 分别表示预测框和真实框的宽与高.

3 实验验证与结果分析

3.1 实验数据集

为了更好地适应我国交通道路场景, 选择使用国内Tsinghua-Tencent 100K (TT100K)数据集^[21]进行训练. TT100K是由清华与腾讯合作制作的大规模交通标志识别数据集, 提供了30 000个交通标志实例, 共128种标志的100 000张高分辨率图像, 部分主要交通标志类型如图7所示. 它包含了光照强度和天气条件巨大变化下的各种图像, 是用于交通标志识别的可靠的基准数据集, 图8所示为其示例.

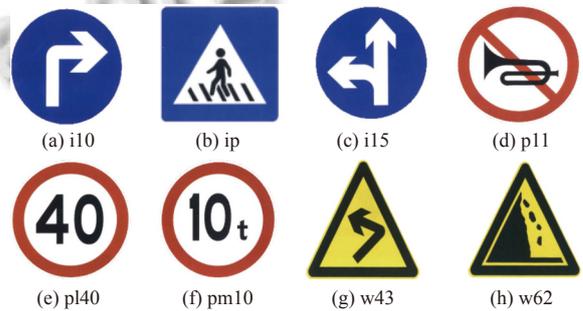


图7 部分主要交通标志类型

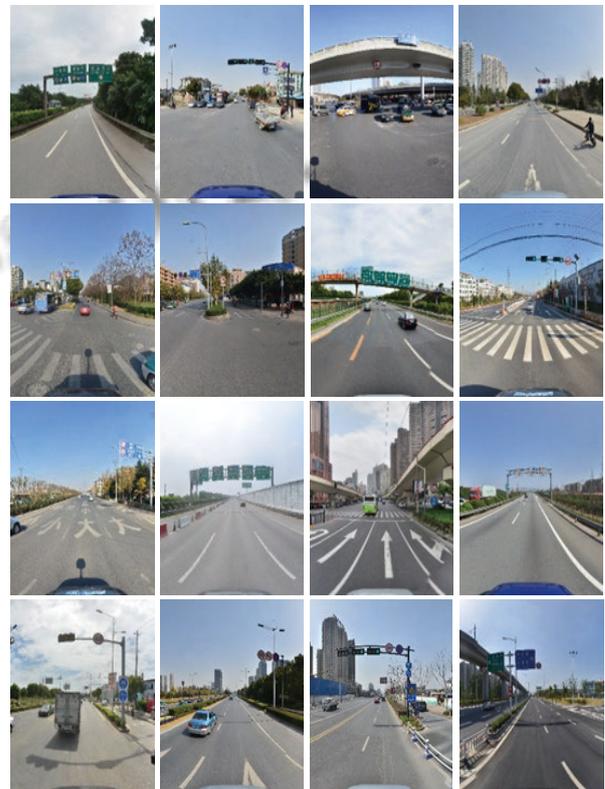


图8 TT100K数据集部分图片

3.2 评价指标

为了客观评价网络模型的性能, 本文实验使用精确率 *Precision*、召回率 *Recall* 和平均精度均值 *mAP* (mean average precision) 作为评价指标, 其中, 精确率是指在模型预测的目标中, 预测正确的比例; 召回率是指在所有真实目标中, 模型预测正确目标的比例; 平均精度均值是所有类别预测精准度的平均值. 具体计算公式如式 (9)–式 (11) 所示:

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (11)$$

其中, *TP* 表示模型预测正确的正样本数量, *FP* 表示模型预测错误的正样本数量, *FN* 表示模型预测错误的负样本数量; *AP* 由集合中某一类别所有精确率之和除以含有该类别目标的图像数量得到.

3.3 实验结果与分析

本文实验环境为 Windows 10 操作系统, PyTorch 深度学习框架, 使用 12 GB 显存的 GPU RTX 3060 训练模型, 训练过程中对网络设定的部分超参数如表 1 所示.

表 1 部分超参数设置

训练参数	参数值
动量 (momentum)	0.937
权重衰减 (weight_decay)	0.0005
初始学习率 (lr0)	0.01
循环学习率 (lrf)	0.2
批次大小 (batch-size)	16
训练代数 (epochs)	300

为了验证上述改进方法的有效性, 本文在 TT100K 数据集上进行了消融对比实验, 结果如表 2 所示, 在模型中加入 GAM 注意力机制和改变损失函数均可以提高网络精度, 两者在 *mAP* 上分别提高了 3.77% 和 2.42%, 而结合两者的改进算法在精确率、召回率和 *mAP* 上均有更好的结果. 图 9 所示为训练所得的精确率-召回率曲线 (*Precision-Recall*), 其能够直观地表现算法性能, 精确率-召回率曲线所围成图形的面积越大表示算法性能越好, 从图 9 可以看出, 训练的模型性能良好, 结合表 2 消融实验结果可知, 本文改进的 YOLOv5

算法与原始的 YOLOv5 算法相比, 在精确率上提升了 7.9%, 召回率提升了 4.43%, *mAP* 提高了 5.72%.

表 2 消融实验结果 (%)

模型	<i>Precision</i>	<i>Recall</i>	<i>mAP</i>
YOLOv5	85.75	85.10	87.28
YOLOv5+GAM	89.40	86.96	91.05
YOLOv5+CIoU	91.33	88.00	89.70
本文算法	93.65	89.53	93.00

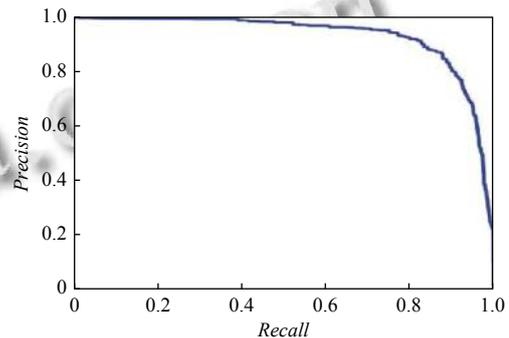


图 9 *Precision-Recall* 曲线

此外, 本文将改进后的 YOLOv5 方法与其他经典的目标检测方法进行对比实验, 实验结果如表 3 所示, 相较于 SSD 和 Faster R-CNN 方法, 本文改进方法的 *mAP* 有明显提升, 提升值分别为 16.51% 和 10.16%, 与 YOLOv4 和 YOLOv5 方法相比 *mAP* 值也有不同程度的提升. 综上所述, 本文改进的 YOLOv5 算法对交通标志识别的综合性能良好.

表 3 对比实验结果 (%)

算法名称	<i>mAP</i>
SSD	76.49
Faster R-CNN	82.84
YOLOv4	85.70
YOLOv5	87.28
本文算法	93.00

最后对训练好的模型进行测试, 测试结果如图 10 所示. 从前两行图中可以看出改进前的 YOLOv5 算法对远处存在的小目标 ph 等存在漏检现象; 第 3、4 行图中存在误检现象, 将第 4 行图中的 p180 误检为 p150; 在第 5 行图中, 即使检测正确但也存在置信度较低的问题. 而改进后的 YOLOv5 算法对远距离的小目标均可以识别到, 漏检和误检问题得到改善, 置信度也相对较高, 相比于原始 YOLOv5 算法在交通标志的识别上具有良好的效果.

- 15 Shi XH, Hu J, Lei XY, *et al.* Detection of flying birds in airport monitoring based on improved YOLOv5. Proceedings of the 6th International Conference on Intelligent Computing and Signal Processing. Xi'an: IEEE, 2021. 1446–1451.
- 16 Yang JF, Fu XY, Hu YW, *et al.* PanNet: A deep network architecture for pan-sharpening. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 1753–1761.
- 17 Liu YC, Shao ZR, Hoffmann N. Global attention mechanism: Retain information to enhance channel-spatial interactions. arXiv:2112.05561, 2021.
- 18 Rezatofighi H, Tsoi N, Gwak J, *et al.* Generalized intersection over union: A metric and a loss for bounding box regression. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 658–666.
- 19 Yu JH, Jiang YN, Wang ZY, *et al.* UnitBox: An advanced object detection network. Proceedings of the 24th ACM International Conference on Multimedia. Amsterdam: ACM, 2016. 516–520.
- 20 Zheng ZH, Wang P, Liu W, *et al.* Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020. 12993–13000.
- 21 Zhu Z, Liang D, Zhang SH, *et al.* Traffic-sign detection and classification in the wild. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2110–2118. [doi: [10.1109/CVPR.2016.232](https://doi.org/10.1109/CVPR.2016.232)]

(校对责编: 孙君艳)