面向长尾分布和特征缺失的交通标志目标检测算法

邓召学[1,2]，刘 灿[1]，郝丙森[1]，彭 冲[3]，隗寒冰[1,2]

（1.重庆交通大学 机电与车辆工程学院，重庆 400074；2.环形山（重庆）科技有限公司，重庆 401100；3.招商局检测车辆技术研究院有限公司，重庆 401122）

摘 要：为解决交通标志检测中的数据长尾分布和小目标特征信息丢失问题，提出了一种交通标志小目标检测网络STF-YOLO。首先，提出了小目标条件扩散模型对TT100K数据集进行数据增强，得到类别均匀的TT100K-P数据集，有效缓解了原数据集长尾分布严重的问题。然后，提出了自适应渐近金字塔网络（AAFPNet），其自适应多尺度通道注意力模块（MC-ACA）在通道层面自适应地融合多层级特征，解决了小目标细节特征丢失问题。最后，设计了K-NCIoU回归损失，结合CIoU与NWD两者优势进一步提升模型检测性能。在三个交通标志检测数据集的测试结果表明，相较于基准模型YOLOv8，STF-YOLO在TT100K-P、TT100K和CCTSDB数据集上mAP@0.5∶0.95分别提升了2.5%、3.3%、2.8%。经实车验证，证明其具备优秀的交通标志实时检测能力。

关键词：交通运输系统工程；小目标检测；长尾分布；小目标条件扩散模型；自适应渐近特征金字塔；多尺度通道注意力

中图分类号：TP242；U491.5 **文献标志码**：A **文章编号**：1671-5497（2026）06-1642-13

DOI：10.13229/j.cnki.jdxbgxb.20241193

Object detection algorithm for traffic signs in long-tail distribution and feature incompleteness

DENG Zhao-xue[1,2], LIU Can[1], HAO Bing-sen[1], PENG Chong[3], WEI Han-bing[1,2]

(1. *College of Electromechanical and Vehicle Engineering, Chongqing Jiaotong University, Chongqing* 400074, *China*; 2. *Crater (Chongqing) Technology Co., Ltd., Chongqing* 401100, *China*; 3. *China Merchants Testing Vehicle Technology Research Institute Co., Ltd., Chongqing* 401122, *China*)

Abstract: To address long-tail distribution and information loss in small traffic sign detection, STF-YOLO was proposed. A small-object conditional diffusion model was employed for augmentation on TT100K. This process created a balanced TT100K-P dataset to alleviate the long-tail issue. An Adaptive Asymptotic Feature Pyramid Network with a Multi-Channel Adaptive Attention (MC-ACA) module was

收稿日期：2024-11-06.

基金项目：国家自然科学基金项目（52172381）；重庆市研究生科研创新项目（CYS260563）；重庆市研究生联合培养基地建设项目（JDLHPYJD2024007）.

作者简介：邓召学（1985-），男，副教授，博士.研究方向：智能车辆自动驾驶行为及协同控制.
E-mail：dengzhaoxue@cqjtu.edu.cn

then introduced. This module fuses multi-level features to recover fine-grained details for small objects. Additionally, the K-NCIoU regression loss was designed, combining the advantages of CIoU and NWD to further enhance the model′s detection performance. The testing results on three traffic sign detection datasets indicate that, compared to the baseline model YOLOv8, STF-YOLO achieves improvements in mAP@0.5:0.95 of 2.5%, 3.3%, and 2.8% on the TT100K-P, TT100K, and CCTSDB datasets, respectively. Real vehicle validation confirms its excellent real-time detection capability for traffic signs.

Key words: engineering of communications and transportation system; small object detection; long-tail distribution; small object conditional diffusion model; adaptive asymptotic feature pyramid; multi-channel attention

交通标志检测是驾驶辅助和自动驾驶系统的重要子模块[1],对于保障驾驶安全、提升驾驶效率和智能化交通管理等方面具有重要作用。然而,交通标志检测在实际应用中面临诸多挑战,如复杂的背景环境[2]、检测数据的长尾分布[3]以及交通标志小目标容易出现信息丢失问题等。解决以上问题能更准确、快速地识别交通标志,对提升自动驾驶系统的可靠性和安全性至关重要。

传统交通标志检测算法依赖于颜色、形状和轮廓等进行特征提取并通过分类器实现交通标志识别[4-6]。但手工提取特征存在人工疲惫和鲁棒性不足的问题,难以适应复杂多变的交通环境。随着卷积神经网络(CNN)的发展[7],众多专家和学者开始采用深度学习方法进行交通标志检测。其中,Wu等[8]结合接收场模块(RFM)和路径聚合网络(PAN)[9]提出了一种基于SSD[10]的交通标志检测模型,该模型对遮挡的交通标志有良好的识别效果。Du等[11]引入空间到深度模块,开发了一种基于YOLOv8的解决多目标漏检问题的交通标志检测方法。类似地,Wang等[12]等利用升级的双向Res2Net对输入特征图进行通道级特征校正,有效增强了交通标志检测精度。然而,虽然现有研究对很多场景(例如多目标与遮挡物)的检测都提出了解决方案,但是,对小目标(特别是较远距离的交通标志小目标)检测仍存在检测与定位精度不足的问题,此问题本质上是检测网络训练时小目标信息丢失所导致的。

另一方面,现实世界中交通标志种类繁多,众多交通标志检测数据集普遍存在样本量不足的问题,数据集的好坏决定了模型能否顺利拟合。为解决此问题,众多学者开始研究有效的数据增强方法。其中,Yun等[13]提出CutMix方法,通过将其他类别信息粘贴到随机抹去的区域,有效提升了区域丢弃策略的效率。Dvornik等[14]通过人工增加训练样本数量,以减少模型的过拟合并提升泛化能力。Georgakis等[15]基于场景的语义信息和上下文合理粘贴物体位置。Yun等[16]在前人研究的基础上结合生成对抗网络优化了Copy-Paste增强算法,生成了更丰富的图像。类似地,Yu等[3]提出多阶段自适应复制粘贴(MSACP)算法,根据每个阶段的训练偏好自适应调整训练集的类信息。总而言之,目前的数据增强方法主要采用数据粘贴技术,即将目标样本裁剪并粘贴到图像的其他区域。然而,这种方法并不适用于种类繁多且大部分类别样本稀少的交通标志。数据的长尾分布问题仍未解决,导致模型对样本较少的类别过拟合,在真实应用中出现部分类别精度不足的问题。

综上,尽管交通标志的检测研究取得了丰硕的成果,但仍存在以下问题:①数据长尾分布,即少数类别样本量极大而大多数类别样本较少,导致模型训练时容易对样本不足的类别过拟合。②小目标在图像中占据的像素较少,容易在下采样和信息融合等过程中丢失关键信息,导致检测精度降低。

针对上述问题,本文提出了一种轻量级实时交通标志检测算法STF-YOLO,有效缓解了检测数据长尾分布与小目标特征缺失问题,实现了对交通标志的实时准确检测,并且避免了计算复杂度的剧增。

1　小目标条件扩散模型

STF-YOLO检测算法具体框架如图1所示。如图2所示,针对交通标志数据长尾分布问题,本文提出利用小目标条件扩散模型增强数据,通过融合边界框、类别信息与纯类别掩码条件精确控制小目标的生成。

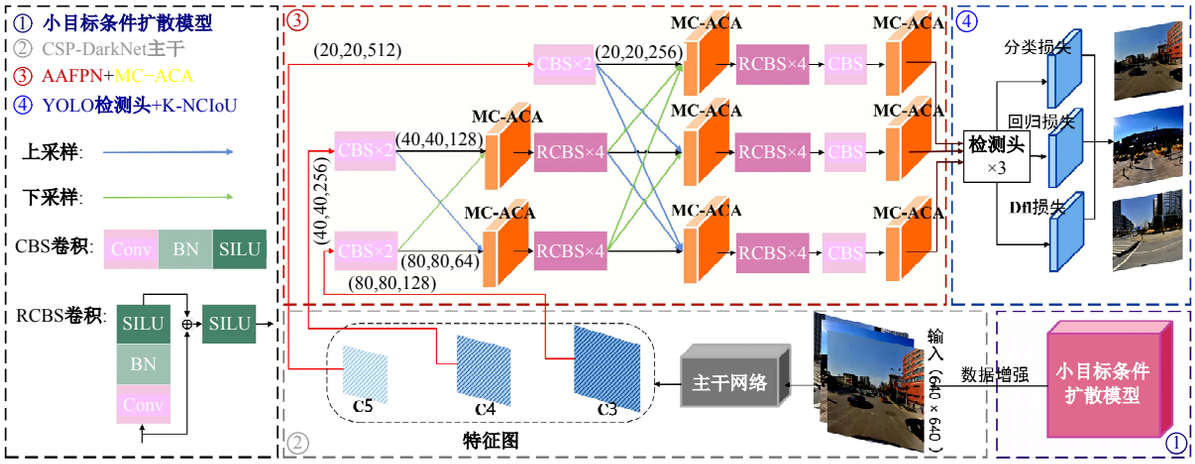

图 1 STF-YOLO 网络结构图

Fig. 1 STF-YOLO network architecture diagram

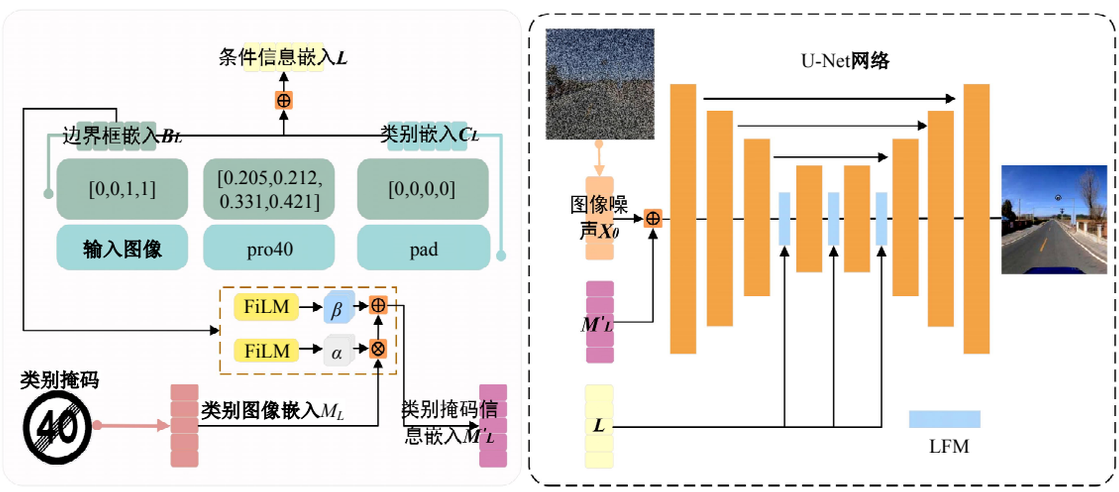

图 2 小目标条件扩散模型结构图

Fig. 2 Structure diagram of the small object conditional diffusion model

1.1 条件信息嵌入

为提高模态间的交互效率,需对条件模态信息进行处理[18]。序列 $l=\{l_1,l_2,\cdots,l_n\}$ 是 n 个对象的集合。每个对象 l_i 表示为 $l_i=\{b_i,c_i\}$,其中 $b_i=(x_{i0},y_{i0},x_{i1}y_{i1})$ 表示一个边界框,$c_i\in[0,C+1]$ 是其类别 ID。为了支持可变长度序列输入,在末尾添加填充 l_o 将序列 l 填充到固定长度 z。具体而言,$b_o=(0,0,0,0)$ 与 $c_o=C+1$ 表示一个没有形状的对象。在填充后,得到一个包含 z 个对象的填充序列 $l=\{l_1,l_2,\cdots,l_z\}$,每个对象具有特定的位置、大小和类别信息。序列 l 通过投影矩阵 W_B 和 W_C 将其转换为条件信息嵌入 $L=\{L_1,L_2,\cdots,L_Z\}\in R^{z\times d_L}$,公式如下:

$$L=B_L+C_L \tag{1}$$

$$B_L=bW_B \tag{2}$$

$$C_L=cW_C \tag{3}$$

式中:B_L、$C_L\in R^{z\times d_L}$ 分别为布局 l 的边界框嵌入与类别嵌入;条件信息嵌入 L 包含了整个输入图像的内容和位置信息;d_L 为嵌入 L 的维度。

1.2 类别掩码信息嵌入

为了确保输出图像的小目标清晰,使用特征线性调制机制(Feature-wise Linear Modulation,FiLM)[17]得到类别掩码嵌入 M'_L 为目标类别和位置提供参考信息。首先,如式(4)所示,类别图像掩码 m 通过投影矩阵 W_M 将其转换为类别图像特征嵌入 M_L。接着,如式(5)(6)所示,边界框嵌入 B_L 通过 FiLM 层生成两个可学习参数 α_i 与 β_i 对类别图像特征嵌入 M_L 进行缩放和偏移,以此得到类别掩码信息嵌入 M'_L。计算过程如下:

$$M_L=mW_M \tag{4}$$

$$\alpha_i、\beta_i=\mathrm{FiLM}(B_L) \tag{5}$$

$$M'_L=\alpha_iM_L+\beta_i \tag{6}$$

式中：$M_L \in R^{z \times d_L}$ 为类别图像嵌入；m 为对应类别的图像掩码信息；α_i 与 β_i 为 FiLM 生成的与 M_L 同样大小的嵌入向量，它们决定了对输入特征的缩放和偏移程度；$M_L' \in R^{z \times d_L}$ 为类别掩码信息嵌入。



图 3　小目标条件扩散模型生成效果图

Fig. 3　Conditional diffusion model results

1.3　特征融合

将图像噪声 X_0 与掩码信息嵌入 M_L' 经拼接后作为 U-Net 的输入。同时，由于条件信息嵌入 L 中对象彼此没有联系，这导致对整个场景的理解较低。使用 Layout Diffusion[18] 中的 LFM 将 L 进行进一步融合后嵌入到 U-Net 中的交叉注意力机制中对其进行融合。

部分生成图像如图 3 所示，受限于计算资源，本工作生成图像的像素设置为 512×512，然后通过 resize 操作将图像扩展到 640×640。图像中小目标清晰，满足模型训练需求，且契合真实交通场景。对比图 4，原始 TT100K-2021(45 类)数据集[19]存在严重的长尾分布问题，近半数的类别样本少于 200。而通过小目标条件扩散模型进行数据增强得到 TT100K-P 数据集中类别稀少的部分都得到了一定程度的扩充，有效地解决了该问题。

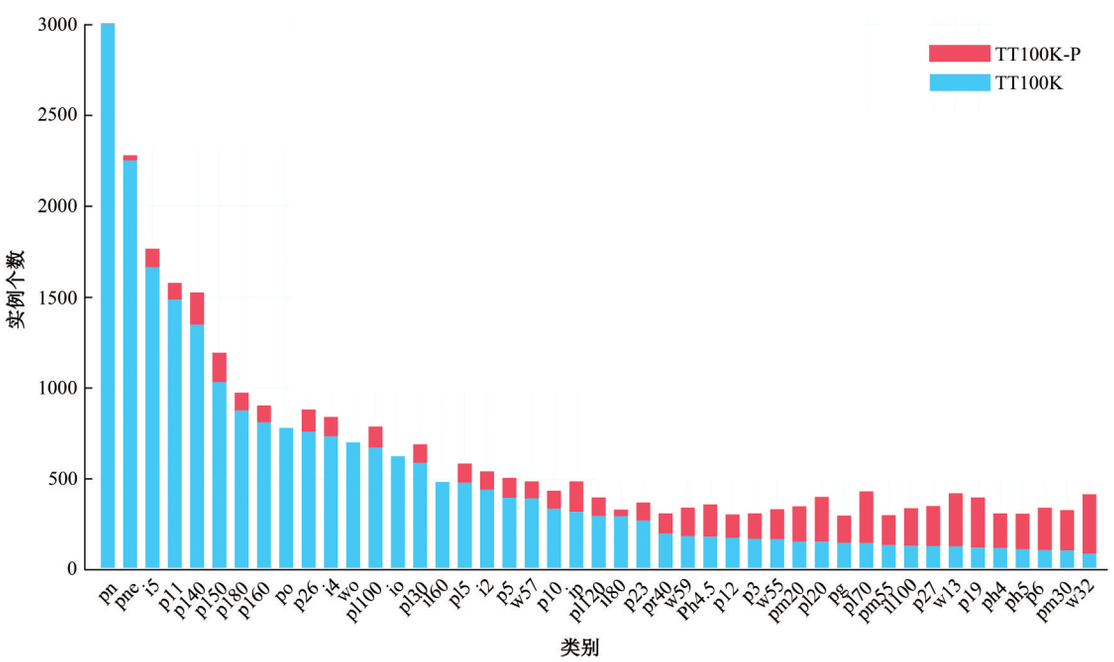

图 4　数据增强前后实例分布对比图

Fig. 4　Instance distribution comparison in augmentation

2　STF-TOLO 检测算法

2.1　多尺度通道注意力模块 MC-ACA

当模型直接融合多层次特征时，尽管输出特征图包含更丰富的信息，但多层级显著信息的融合可能导致微小信息被忽略。在小目标检测中，这些微小信息往往承载着交通标志的关键信息。为了解决这一问题，本文提出了多尺度通道注意力模块(MC-ACA)，其结构如图 5 所示。

MC-ACA 模块整体计算过程如式(7)所示：

$$F'=\sum_{j=1}^{n}\left[w_j\cdot M\left(\sum_{i=1}^{n}F_i\right)\otimes F_j\right] \tag{7}$$

$$w_j=\frac{\exp(a_j)}{\sum_{j=2}^{n}\exp(a_j)} \tag{8}$$

式中:$\sum_{i=1}^{n}F_i$为n个分支总融合特征图;F_j为第j个分支的特征图;n为通过MC-ACA的特征图数目;F'为融合后的特征图;M为通过通道注意力之后的通道注意力权重;w_j为j个分支所对应的权重;a_j为一个可学习三元素一维张量,初始值全为1;exp()为指数函数。

首先,将输入的n个尺寸为$H\times W\times C$的特征图F_i进行逐点相加,获得融合特征图。然后,通过全局平均池化与全局最大池化突出融合特征图的整体信息与最显著特征。为避免过度强调上下文显著信息,引入权重参数λ(设为0.25)调节突出信息比重。接着,受ECA通道注意力[20]启发,MC-ACA通过一维卷积考虑每个通道及其K个邻域的跨通道交互信息,K表示有多少个邻近通道参与该通道进行注意力交互。本文设定,若通道数不超过128,K=3,否则K=5。一维卷积后的特征图经过SiLU激活函数处理将每个加权通道映射到(0,1)范围内。最后,将n个特征图F_i通过各自的权重系数w_i与映射后的特征图逐点相乘后进行融合,得到最终输出特征图F'。

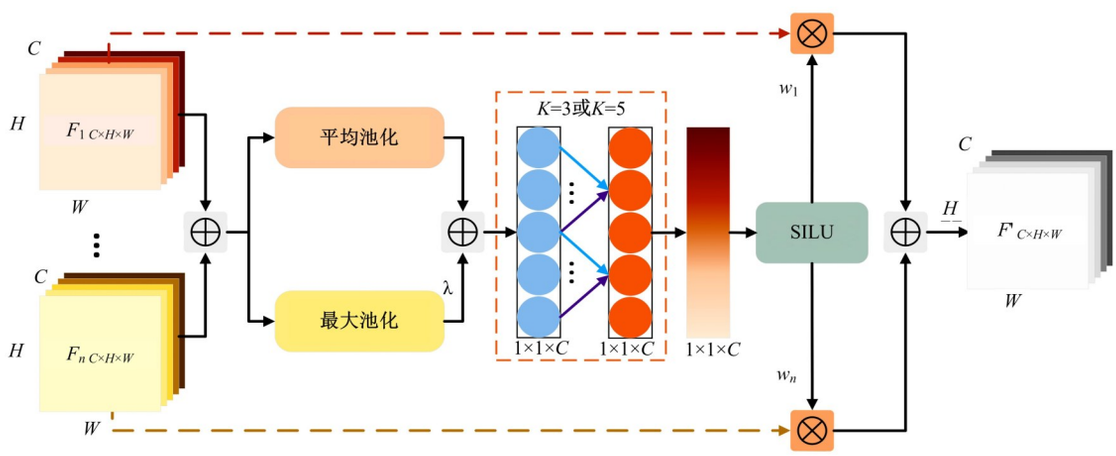

图5 自适应多尺度通道注意力结构图(MC-ACA)

Fig. 5 Structure diagram of adaptive multi-channel attention (MC-ACA)

2.2 自适应渐近金字塔网络AAFPNet

渐近特征金字塔AFPNet[21]通过自适应空间融合模块(ASFF)为不同层级特征分配空间权重,但这会导致对某些特征层的过度关注而忽视其他层的重要信息,从而造成潜在信息的丢失。此外,连续下采样也会导致小目标信息缺失。为此,本文提出自适应渐近金字塔结构AAFPNet,其结构如图6所示。

AAFPNet网络将下采样倍率从8倍调整至2倍,虽增加了颈部参数,但减少了信息缺失,有利于小目标检测。此外,采用MC-ACA替换ASFF,能有效整合不同特征层的信息,充分发挥各层优势。与ASFF相比,MC-ACA不仅考虑空间权重,还通过通道注意力机制动态调整特征贡献,强化关键通道特征,降低信息丢失风险。最后,多个分支融合后,在颈部尾部三个分支均加入单层MC-ACA模块,有效地整合不同分支信息,确保模型关注有利的特征,从而提高检测性能。

2.3 K-NCIoU损失函数

YOLOv8中使用Zheng等[22]提出的CIoU作为回归损失函数。该函数加入中心点距离、重叠面积和长宽比。但在交通标志检测中小目标众多,考虑太多几何因素反而导致其对位置偏差过于敏感,影响准确性。其计算公式如下:

$$\mathrm{CIoU}=1-\mathrm{IoU}+\frac{\rho^2(b_A,b_B)}{c^2}+av \tag{9}$$

$$a=\frac{v}{(1-\mathrm{IoU})+v} \tag{10}$$

$$v=\frac{4}{\pi^2}\left(\arctan\frac{w_B}{h_b}-\arctan\frac{w_A}{h_A}\right)^2 \tag{11}$$

式中:$\rho^2(b_A,b_B)$为预测框与真实框中心点之间的欧几里得距离;c为包围两个框的最小外接框的对角线长度;权重因子用a表示;长宽比一致性用v表示;w_A和h_A分别为预测框的宽度和高度;w_B

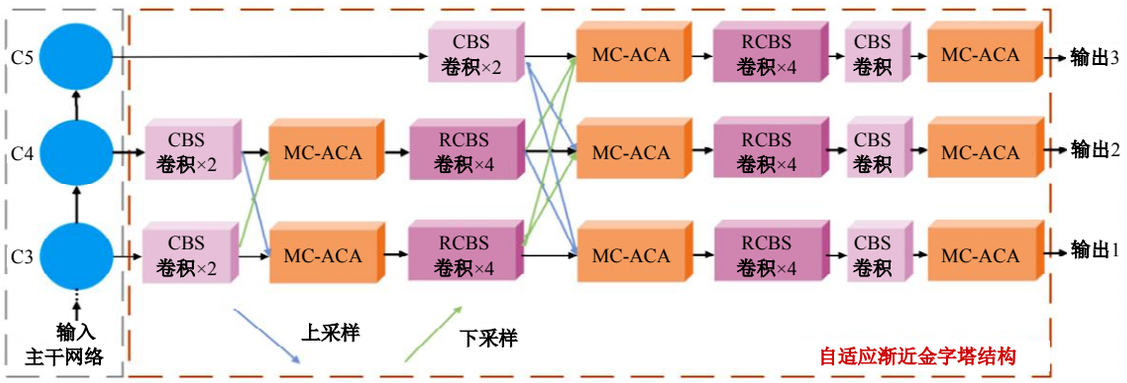

图 6　自适应渐近金字塔结构图

Fig. 6　Structure diagram of adaptive asymptotic feature pyramid (AAFPNet)

和 h_B 分别为真实框的宽度和高度;IoU 为预测框与真实框的交并比。

Wang 等[23]提出的 NWD 则通过将框建模为二维高斯分布,利用归一化瓦瑟斯坦距离对目标尺度不敏感的优点计算相似度,更加稳定于小目标检测。但是,归一化瓦瑟斯坦距离的计算涉及最优传输问题,这是一个非凸优化问题,可能导致收敛速度缓慢。NWD 计算公式如下:

$$W_2^2(NA,NB)=\left\|\left[cx_A,cy_A,\frac{w_A}{2},\frac{h_A}{2}\right]^{\mathrm{T}}-\left[cx_B,cy_B,\frac{w_B}{2},\frac{h_B}{2}\right]^{\mathrm{T}}\right\|^2 \tag{12}$$

$$NWD(NA,NB)=\exp\left(-\frac{w_2^2(NA,NB)}{c}\right) \tag{13}$$

式中:NA 与 NB 是由真实边界框 $A=(cx_A,cy_A,w_A,h_A)$ 和预测边界框 $B=(cx_B,cy_B,w_B,h_B)$ 建模的高斯分布;$w_2^2(NA,NB)$ 为距离度量;c 为与数据集密切相关的常数。

为了解决上述问题,提出了 K-NCIoU 方法,通过自适应参数 k 调整 NWD 损失与 CIoU 损失比重。具体而言,尺寸较大的物体对物体位置的敏感性弱,以 CIoU 作为主导回归损失函数;而对于位置敏感的小目标,则采用 NWD 作为主导损失函数,充分利用其在小目标检测中的优势。同时,保留 CIoU 损失,以加快算法收敛速度。K-NCIoU 计算公式定义如下:

$$L_{\text{K-NCIoU}}=(1-k)\cdot(1-NWD(NA,NB))+k\cdot(1-\text{CIoU}) \tag{14}$$

$$k=\gamma\cdot\frac{w_A\cdot h_A}{w_L\cdot h_L} \tag{15}$$

式中:k 为自适应权重比例系数;w_L 与 h_L 分别为大尺寸样本框的宽度与高度,本文按照小目标定义方式,w_L 与 h_L 取 96;γ 为平衡参数,$\gamma=0.25$。

3　实验结果与分析

3.1　实验数据

本文使用 TT100K[19]、自制 TT100K-P 和 CCTSDB2021[24]数据集作为 STF-YOLO 性能测试的检测基准。

TT100K 交通标志数据集包含 100 000 张高分辨率(2 048×2 048)图像,包含 30 000 个交通标志实例,实例大小分布在 16×20 到 160×160 像素之间。实验选取了公认最常见的 45 个类别,共有 9 659 张完整标注图像,其中包含 6 738 张训练集、1 928 张验证集与 993 张测试集。相应的标志图像和类别名称如图 7 所示,其中 pl* 包括 pl100、pl120、pl20 等,pm* 包括 pm20、pm30 和 pm55,ph* 包括 ph4、ph4.5 和 ph5,il* 包括 il100、il60 和 il80。

TT100K-P 在 TT100K 的基础上,使用所提



图 7　TT100K45 类实例图

Fig. 7　TT100K 45-Class instance diagram

出的小目标条件扩散模型数据增强方法将其扩充至12 025张完整标注信息的图像,包含8 572张训练集、2 460张验证集和993张测试集。

CCTSDB2021数据集共17 856张图像,包含15 886张训练集和1 970张测试集。数据集涵盖城市道路和高速公路等场景,分辨率在600×900到1 024×768之间。交通标志的大小分布在20×20到573×557像素之间,主要分为“警告”、“禁止”和“强制”三类。

3.2 实验环境和参数设置

实验环境为Ubuntu 20.04操作系统,配置64GB RAM、GTX-4090显卡(24 GB显存)。采用PyTorch 1.13.0深度学习框架,编程语言为Python 3.8。模型训练采用随机梯度下降(SGD)优化,初始学习率为0.01,动量为0.937,权重衰减系数为0.000 5。为保证实验公平,不使用预训练权重,训练轮数与YOLOv8一致,均为500轮。TT100K、自制TT100K-P和CCTSDB2021数据集的Batch size均为32。

3.3 实验评估指标

实验采用平均精确率(mAP)、精确率(AP)、浮点运算数(FLOPs)和参数量(Params)作为检测性能评估指标。mAP@0.5表示IoU阈值为0.5时的平均精确率,mAP@0.75表示IoU阈值为0.75时的平均精确率,mAP@0.5:0.95表示IoU从0.5到0.95(步长为0.05)的平均精确率,三者直观反映模型性能。APs、APm和APl分别代表检测小型、中型和大型物体的检测精度。浮点运算次数(Floating point operations,FLOPs)用于衡量模型的复杂程度。参数量(Parms)表示模型的参数数量,衡量对显卡性能的需求。

3.4 消融对比实验

为评估各设计模块对模型性能的贡献,本文将其与YOLOv8s结合,在TT100K-P数据集上进行了5组消融实验,mAP曲线如图8所示彩图参见电子版,下同。结果显示,新增模块的mAP曲线均优于基准模型YOLOv8,表明每个改进模块都为网络带来了一定的性能提升。STF-YOLO的mAP曲线覆盖了单个模块添加时的曲线,表明其改进算法效果达到最佳。

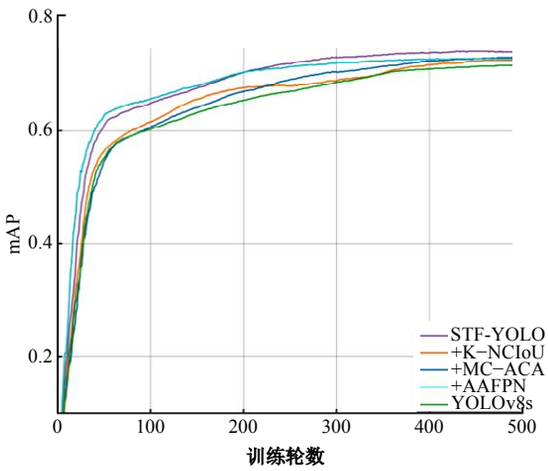

图8 消融实验mAP曲线对比图

Fig. 8 Comparison of mAP curves in ablation experiments

消融对比实验结果如表1所示:引入MC-ACA模块后,mAP@0.5、mAP@0.75和mAP@0.5:0.95分别提升了1.9%、2.0%和1.6%,小、中目标检测精度提升了3.7%和1.1%,由于大目标数据量占比小,且MC-ACA集中关注小目标而忽视了部分大目标全局信息,导致大目标检测精度下降了0.7%。由于基准模型颈部未采用下采样,引入MC-ACA后模型的浮点数运算量和参数量分别增加了30.26 G和23.34 M;以此引入了新的颈部网络AAFPNet,结合下采样、渐近融合与MC-ACA策略,其mAP@0.5、mAP@0.75和mAP@0.5:0.95分别提高了1.2%、2.6%和1.3%,小、中和大目标精度提升了3.5%、0.8%和0.7%。模型的浮点数运算量与参数量仅略有增加;引入K-NCIoU作为回归损失函数后,mAP@0.5、mAP@0.75和mAP@0.5:

表1 消融对比实验表

Table 1 Ablation comparison experiment table

方法	mAP@0.5	mAP@0.75	mAP@0.5:0.95	AP_S	AP_m	AP_l	FLOPs/G	Params/M
YOLOv8s	87.3	81.3	71.4	60.6	78.3	88.7	14.83	11.55
YOLOv8s+MC-ACA	89.2	83.3	73.0	64.3	79.4	88.0	45.09	34.89
YOLOv8s+K-NCIoU	89.1	82.5	72.4	63.2	79.0	88.1	14.83	11.55
YOLOv8s+AAFPNet	88.5	83.9	72.7	64.1	79.1	89.4	18.79	14.89
STF-YOLO	89.6	84.5	73.9	66.0	79.7	88.9	18.79	14.89

0.95分别提升了1.8%、1.2%和1.0%，小、中目标精度提高了2.6%和0.7%，由于自适应参数k引入小部分的NWD损失使模型对大目标的定位不够精细，导致大目标检测精度下降了0.6%。最终，STF-YOLO的mAP@0.5、mAP@0.75和mAP@0.5:0.95分别提高了2.3%、3.2%和2.5%；小目标与中目标检测精度提升了5.4%与1.4%，且大目标检测精度也有小幅度上涨；模型的浮点数运算量与参数量仅略有增加。这表明，本文提出的检测算法在确保轻量化的基础上，显著提高了小目标的检测精度。

3.5　自适应渐近金字塔网络有效性分析

本文提出的自适应渐近金字塔网络AAFPNet通过调整下采样和自适应多尺度通道注意力模块(MC-ACA)，有效解决了小目标细节特征丢失问题。

为了更直观评估AAFPNet对小目标特征的关注，比较了YOLOv8s与应用AAFPNet的特征热力图，如图9所示：(a)组为分散小目标场景，基准模型对边缘“w57”小目标丧失了注意力，AAFPNet增强了对分散小目标特征的关注；(b)组为多目标场景，AAFPNet保持了对多目标的关注能力；(c)组为远距离遮挡场景，基准模型对遮挡目标“pl40”丧失了注意力，AAFPNet提高了对远距离遮挡小目标的关注能力。这些结果表明，AAFPNet显著增强了对小目标区域的关注，有效改善了小目标特征消失的问题。

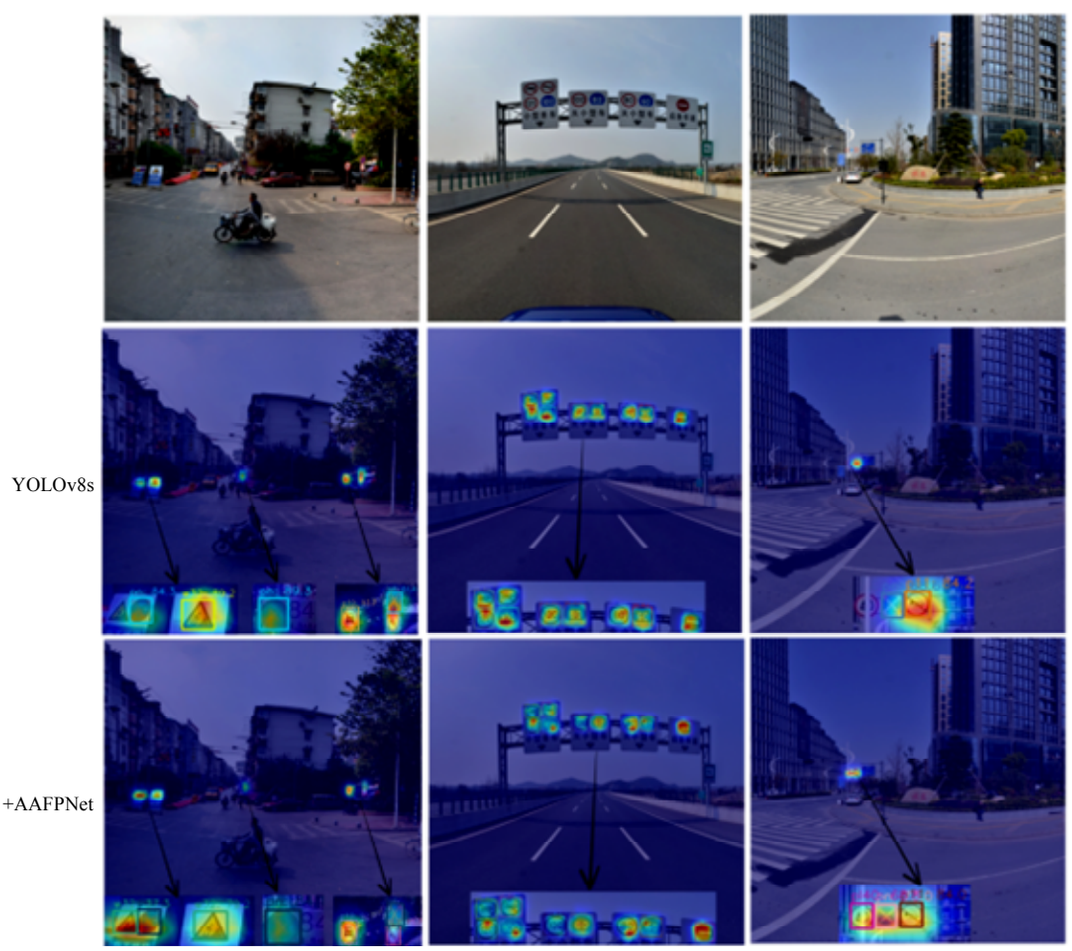

图9　自适应渐近金字塔有效性对比图

Fig. 9　Comparison of effectiveness of adaptive asymptotic feature pyramid

3.6　算法效果对比实验——TT100K数据集

为验证所提出的网络模型的优越性，在TT100K数据集上将STF-YOLO网络与近年针对小目标检测提出的通用框架进行对比，这些算法代表着小目标检测领域的最新进展。

实验结果如表2所示：与大部分算法相比，STF-YOLO的检测平均准确率与小、中和大目标的检测准确率均表现更好，同时浮点数运算量

与模型参数量也远远小于这些算法;与YOLOv8s相比,参数量与浮点数运算量略增,但也满足轻量化的要求。算法的mAP@0.5、mAP@0.75与mAP@0.5:0.95分别比YOLOv8s提高了3%、1.8%与3.3%,其小、中和大目标检测的准确率也超过了YOLOv8s。总的来说,STF-YOLO兼顾了准确率和轻量化,在交通标志检测方面表现优秀。

表2 TT100K数据集上检测模型效果对比实验表

Table 2 Comparison of detection model performance on the TT100K dataset

Model	Year	输入尺寸	mAP0.5	mAP0.75	mAP0.5:0.95	AP_S	AP_m	AP_l	FLOPs/G	Params/M
Faster-RCNN[25]	2015	1 024×10 244	74.1	69.3	58.9	27.7	73.1	81.8	52.5	41.6
RetinaNet[26]	2017	2 048×2 048	61.9	57.9	49.2	—	—	—	—	—
Cascade-RCNN[27]	2018	1 024×1 024	80.1	74.8	63.5	36.2	74.6	83.0	84.5	69.3
FCOS[28]	2019	2 048×2 048	81.3	76.0	63.8	—	—	—	—	—
PSG-YOLOv5[29]	2021	1 024×1 024	69.5	62.8	53.2	35.6	61.2	70.6	131.3	106.2
AIE-YOLO[30]	2022	512×512	78.0	72.9	61.7	—	—	—	—	
YOLOv8s	2023	640×640	81.8	73.5	61.3	37.3	70.5	79.8	14.8	11.6
TF-Fusion[31]	2024	640×640	53.5	49.8	39.5	43.6	43.7	58.6	—	—
STF-YOLO		640×640	84.8	75.3	64.6	44.5	73.8	82.8	19.9	14.9

3.7 算法效果对比实验-TT100K-P数据集

为进一步验证本研究中提出的网络模型的有效性,在自制的TT100K-P数据集上重现了YOLOv5、YOLOX、YOLOv7和YOLOv8算法的检测结果。对比结果见表3,STF-YOLO的mAP@0.5:0.95为73.9%,分别比YOLOv5s、YOLOXs、YOLOv7tiny和YOLOv8s高出14.6%、4.4%、10.9%和2.5%。小目标检测准确率APs较YOLOv5s、YOLOXs、YOLOv7tiny和YOLOv8s分别提升了17.7%、3.8%、14.6%和5.4%。详细的mAP值见表4,STF-YOLO在45个类别中的大多数上均表现出最佳性能。

为验证基于小目标条件扩散模型的数据增强方法所构建的TT100K-P数据集的有效性。通过对比表2与表3的数据,STF-YOLO在TT100K-P数据集相比原始TT100K数据集在mAP及小、中、大目标的检测精度上均有显著提升,分别提高了4.8%、9.4%、9.3%、21.5%、5.9%和6.1%。表4显示,使用TT100K-P数据集训练每个类别拟合良好,且表现出较高的准确性。以上数据都表明,所提出的小目标条件扩散模型数据增强方法有效解决了数据长尾分布的问题。

表3 TT100K-P数据集上检测模型效果对比

Table 3 Comparison of detection model performance on the TT100K-P dataset

Model	Year	mAP@0.5	Map@0.75	mAP@0.5:0.95	AP_S	AP_m	AP_l	FLOPs/G	Params/M
YOLOv5s	2021	77.1	69.4	59.3	49.7	65.6	78.3	8.78	7.24
YOLOXs	2021	87.1	81.0	69.5	62.2	72.9	82.5	13.47	8.97
YOLOv7tiny	2022	79.7	72.9	63.0	51.4	71.0	83.3	7.22	6.54
YOLOv8s	2023	87.3	81.3	71.4	60.6	78.3	88.7	14.83	11.55
STF-YOLO		89.6	84.5	73.9	66.0	79.7	88.9	19.87	14.89

3.8 算法效果对比实验——CCTSDB数据集

为了进一步保证STF-YOLO算法的泛化性和有效性,在CCTSDB2021数据集上进行了第三次对比实验。在CCTSDB2021数据集上STF-YOLO与最新方法的对比如表5所示,其mAP@0.5均优于Faster R-CNN、Cascade-RCNN、YOLOv5s和YOLOv8s等算法。

3.9 检测效果对比分析

为更直观地评价本文改进算法的性能,对YOLOv7tiny和YOLOv8s与本文算法在TT100K-P测试集上的检测结果进行对比。对比结果如图10所示:(a)组为分散小目标场景,YOLOv7tiny表现最差,漏检了“pl70”、“pl50”、“p10”和“i2”类别的小物体,并且“pl40”出现了边界框定位不准问题。YOLOv8s在检测小物体“pl50”、

表 4　45类交通标志检测准确率对比

Table 4　Comparison of detection accuracy for 45 classes of traffic signs

Method	Total	i2	i4	i5	il100	il60	il80	io	ip	p10	p11	p12
YOLOv5s	59.3	51.5	55.5	60.7	58.2	51.5	51.4	61.1	54.1	49.2	50.3	64.6
YOLOXs	69.5	68.3	64.1	65.5	71.2	70.7	65.3	68.7	61.2	64.5	65.2	70.8
YOLOv7tiny	63.0	54.4	57.1	60.8	60.1	54.9	45.2	60.3	55.1	53.0	54.1	69.4
YOLOv8s	71.4	65.5	63.9	64.2	84.3	80.0	80.7	70.6	63.3	59.1	61.1	73.4
STF-YOLO	73.9	66.1	65.1	66.6	85.6	83.2	81.8	75.8	63.9	64.4	65.3	75.8
Method	p19	p23	p27	p3	p5	p6	pg	ph4	ph4.5	ph5	pl100	pl120
YOLOv5s	58.9	67.9	70.1	73.7	71.4	57.5	61.2	47.5	67.8	60.9	63.4	71.1
YOLOXs	70.1	74.9	83.6	78.7	79.1	68.8	69.3	59.6	75.6	72.8	67.1	72.5
YOLOv7tiny	62.7	71.9	80.2	73.8	72.5	61.4	64.6	60.8	69.3	64.1	65.1	75.6
YOLOv8s	70.2	77.0	82.2	86.3	77.6	71.3	74.1	63.1	76.2	76.3	79.0	82.1
STF-YOLO	69.7	80.1	86.5	85.6	77.7	72.3	76.0	70.6	75.8	77.1	82.3	83.6
Method	pl20	pl30	pl40	pl50	pl80	pm20	pm30	pm55	pn	pne	po	pr40
YOLOv5s	45.0	54.6	54.1	44.3	51.7	70.0	71.2	66.6	57.8	58.0	69.6	74.6
YOLOXs	61.3	68.2	68.4	64.6	66.5	74.6	71.1	71.8	70.2	66.5	76.1	75.7
YOLOv7tiny	54.7	62.9	57.0	48.7	58.9	73.7	77.9	75.0	62.1	57.8	72.5	78.4
YOLOv8s	62.1	68.1	64.3	58. 5	68.1	82.0	81.5	79.4	63.6	60.7	79.4	84.8
STF-YOLO	64.1	70.9	67.1	63.1	70.4	81.9	79.8	81.5	66.2	63.4	79.5	85.2
Method	p26	pl5	pl60	pl70	w57	w59	wo	w13	w32	w55		
YOLOv5s	64.7	61.1	56.9	60.1	48.9	64.0	55.5	56.3	46.2	57.2		
YOLOXs	71.6	69.8	66.5	69.9	61.4	70.6	68.5	72.5	58.6	70.4		
YOLOv7tiny	67.8	63.5	63.3	63.3	51.7	65.3	59.2	62.2	49.2	64.9		
YOLOv8s	73.7	69.3	70.2	67.8	60.5	70.5	64.5	70.9	58.5	71.8		
STF-YOLO	76.3	70.7	72.5	71.0	61.5	71.3	64.3	74.5	59.2	73.0		

表 5　CCTSDB数据集上检测模型效果对比实验表

Table 5　Comparison of detection model performance on the CCTSDB dataset

Model	Year	输入尺寸	mAP@0.5	Params /M
Faster R-CNN[25]	2015	1 024×1 024	77.4	41.6
Cascade-RCNN[27]	2018	1 024×1 024	80.3	—
FCOS[28]	2019	2 048×2 048	87.8	35.6
PSG-YOLOv5[29]	2021	1 024×1 024	84.1	106.2
YOLOv5s	2021	640×640	87.5	7.2
YOLOv8s	2023	640×640	93.3	11.6
STF-YOLO		640×640	96.1	14.9

“pn”和“i2”时出现了漏检。而STF-YOLO成功检测了所有标志;(b)组为多目标场景,YOLOv7tiny、YOLOv8s与STF-YOLO均成功检测了所有标志,其中STF-YOLO的检测精度与定位精度均最好;(c)组为遮挡小目标场景,YOLOv7tiny漏检遮挡目标“w55”。YOLOv8s出现了多检问题,即将遮挡物检测为了“w55”与“w53”两种标志。STF-YOLO成功检测出了遮挡物“w55”;(d)组为阴影场景,YOLOv7tiny、YOLOv8s与STF-YOLO均检测了所有标志。以上检测效果表明,STF-YOLO算法在分散小目标、多目标与遮挡小目标等复杂交通场景下依然有很好的性能,进一步证明了检测算法的有效性和实用性。

3.10　实车检测性能测试

为验证本研究的可行性,进行了实车检测性能分析。实验在Ubuntu22.04系统下进行,使用深度学习框架Pytorch1.11,CUDA 11.3,图形处理器(GPU)为NVIDIA RTX3050Ti,通过外接摄像头获取行车实时视频进行检测。实验平均时速为35 km/h,所获取的检测图像如图11所示:STF-YOLO检测算法准确检测到所有标志,通过计算可得在上述实验条件下STF-YOLO模型



图10 复杂场景交通标志检测效果对比

Fig. 10 Comparison of detection results for traffic signs in complex scenes



图11 实车交通标志检测效果

Fig. 11 Detection results of real-world traffic signs

图像帧率(FPS)达到52.04帧/s,证明STF-YOLO检测算法满足实时检测要求。

4 结 论

(1)针对交通标志检测出现的数据长尾分布与小目标特征消失的问题,本文提出STF-YOLO检测算法。相较于基准模型YOLOv8,该算法在自制的TT100K-P、TT100K与CCTSDB数据集上mAP@0.5:0.95分别提升了2.5%、3.3%、2.8%。经实车验证,证明其具备优秀的交通标志实时检测能力。

(2)提出了小目标条件扩散模型数据增强方法对TT100K交通标志数据集进行数据增强,解决了数据集长尾分布问题。

(3)设计了多尺度通道注意力模块(MCACA),用于在通道层面自适应融合多层级特征,并将其引入到设计的自适应渐近金字塔(AAFPNet)中,以解决小目标信息缺失问题,提升小目标检测和定位的准确性。

(4)提出了K-NCIoU回归损失,结合目标样本边界框大小自适应地调整NWD与CIoU损失的权重系数,提高小目标检测回归损失的有效性。

夜间交通标志存在如街道照明和反射光干扰等诸多困难,下一步工作将收集夜间交通标志数据集,重点关注夜景中的交通标志检测。

参考文献:

[1] Liu Y, Peng J, Xue J H, et al. TSingNet: scale-

aware and context-rich feature learning for traffic sign detection and recognition in the wild[J]. Neurocomputing, 2021, 447: 10-22.

[2] 张木易. 复杂背景下交通标志的检测与识别研究[D]. 西安: 西安电子科技大学空间科学与技术学院, 2020.
Zhang Mu-yi. Research on detection and recognition of traffic signs in complex backgrounds[D]. Xi'an: School of Aerospace Science and Technology, Xidian University, 2020.

[3] Yu X, Li F, Liu Y, et al. A multi-stage adaptive copy-paste data augmentation algorithm based on model training preferences[J]. Electronics, 2023, 12(17): 3695-3702.

[4] Zhang T, Zou J, Jia W. Fast and robust road sign detection in driver assistance systems[J]. Applied Intelligence, 2018, 48: 4113-4127.

[5] Yu L, Xia X, Zhou K. Traffic sign detection based on visual co-saliency in complex scenes[J]. Applied Intelligence, 2019, 49: 764-790.

[6] Berkaya S K, Gunduz H, Ozsen O, et al. On circular traffic sign detection and recognition[J]. Expert Systems with Applications, 2016, 48: 67-75.

[7] 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, Las Vegas, NV, USA, 2016: 770-778.

[8] Wu J, Liao S. Traffic sign detection based on SSD combined with receptive field module and path aggregation network[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 4285436.

[9] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018: 8759-8768.

[10] Liu W, Anguelov D, Erhan D, et al. Ssd: single shot multibox detector[C]//Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 2016: 21-37.

[11] Du S, Pan W, Li N, et al. TSD-YOLO: Small traffic sign detection based on improved YOLO v8[J]. IET Image Processing, 2024, 18(11): 2884-2898.

[12] Wang X, Tian Y, Zheng K, et al. C2Net-YOLOv5: A bidirectional Res2Net-based traffic sign detection algorithm[J]. Computers, Materials & Continua, 2023, 77(2): 1949-1965.

[13] Yun S, Han D, Oh S J, et al. CutMix: regularization strategy to train strong classifiers with localizable features[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6023-6032.

[14] Dvornik N, Mairal J, Schmid C. Modeling visual context is key to augmenting object detection datasets [C]//Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 2018: 364-380.

[15] Georgakis G, Mousavian A, Berg A C, et al. Synthesizing training data for object detection in indoor scenes[J]. arXiv preprint arXiv:1702.07836, 2017.

[16] Yun W H, Kim T, Lee J, et al. Cut-and-paste dataset generation for balancing domain gaps in object instance detection[J]. IEEE Access, 2021, 9: 14319-14329.

[17] Perez E, Strub F, de Vries H, et al. FiLm: visual reasoning with a general conditioning layer[C]//Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, 2018: 3942-3951.

[18] Zheng G, Zhou X, Li X, et al. LayoutDiffusion: Controllable diffusion model for layout-to-image generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada, 2023: 22490-22499.

[19] Zhu Z, Liang D, Zhang S, et al. Traffic sign detection and classification in the wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, NV, USA, 2016: 2110-2118.

[20] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, 2020: 11534-11542.

[21] Yang G, Lei J, Zhu Z, et al. AFPN: asymptotic feature pyramid network for object detection[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2023: 2184-2189.

[22] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence, New York, USA, 2020: 12993-13000.

[23] Wang J, Xu C, Yang W, et al. A normalized Gauss-

ian Wasserstein distance for tiny object detection[J]. arXiv preprint arXiv:2110.13389, 2021.

[24] Zhang J, Zou X, Kuang L D, et al. CCTSDB 2021: a more comprehensive traffic sign detection benchmark [J]. Human-centric Computing and Information Sciences, 2022, 12: 23.

[25] Li X, Xie Z, Deng X, et al. Traffic sign detection based on improved faster R-CNN for autonomous driving[J]. The Journal of Supercomputing, 2022, 78: 1-21.

[26] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy, 2017: 2980-2988.

[27] Cai Z, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA, 2018: 6154-6162.

[28] Tian Z, Shen C, Chen H, et al. FCOS: fully convolutional one-stage object detection[J]. arXiv preprint arXiv:1904.01355, 2019.

[29] Hu J, Wang Z, Chang M, et al. Psg-yolov5: a paradigm for traffic sign detection and recognition algorithm based on deep learning[J]. Symmetry, 2022, 14(11): 2262.

[30] Yan B, Li J, Yang Z, et al. AIE-YOLO: auxiliary information enhanced YOLO for small object detection[J]. Sensors, 2022, 22(21): 8221.

[31] Zeng G, Huang W, Wang Y, et al. Transformer fusion and residual learning group classifier loss for long-tailed traffic sign detection[J]. IEEE Sensors Journal, 2024, 24(7): 10551-10560.