文章编号：1003－1251(2026)05－0070－07

显著性引导的 SG-DETR 低光目标检测

魏英姿，张 航，于聚壮

（沈阳理工大学 信息科学与工程学院，沈阳 110159）

摘 要：针对低光图像目标检测中边缘模糊、遮挡及尺度差异等问题，提出一种基于视觉显著性引导的目标检测方法 RT-DETR 的改进模型。首先，利用深度学习模型生成的显著性图做输入级增强，突出目标特征与边缘信息，减少冗余候选框数量。其次，设计门控调制模块，将基于显著性图的注意力权重融入到 RT-DETR 模型的特征图中，确保注意力权重与真实目标区域的精准匹配，以此实现特征级的融合，减少噪声干扰。最后，在损失级利用显著性驱动的动态重加权机制，使训练过程中更关注显著区域的目标，提升模型收敛速度与稳定性。通过自建复杂目标数据集 LlSOD 的测试结果显示，改进模型 mAP@0.5 和 mAP@0.5∶0.95 达到 81.4% 和 60.3%，较原 RT-DETR 模型提升 4.2% 和 1.9%，可有效应用于低光场景目标检测任务。

关 键 词：显著性引导；低光图像；RT-DETR 算法；注意力机制

中图分类号：TP391.4 **文献标志码**：A **DOI**：10.3969/j.issn.1003－1251.2026.05.010

Saliency-Guided SG-DETR Low-Light Scene Detection

WEI Yingzi，ZHANG Hang，YU Juzhuang

（Shenyang Ligong University，Shenyan 110159，China）

Abstract：To address issues such as edge blur，occlusion，and scale differences in low-light image object detection，a RT-DETR object detection method guided by visual saliency was proposed. Firstly，a saliency map generated by a deep learning model was used for input-level enhancement to highlight target features and edge information，reducing the number of redundant candidate boxes. Secondly，a gated modulation module was designed to integrate the attention weights based on the saliency map into the RT-DETR model's feature maps，ensuring precise alignment of attention weights with actual target regions，thereby guaranteeing feature-level fusion and reducing noise interference. Finally，a saliency-driven dynamic reweighting mechanism was applied at the loss level，making the training process focus more on the target instances in salient regions，improving model convergence speed and stability. Tests on a self-built complex object dataset LlSOD show that mAP@0.5 and mAP@0.5∶0.95 of the improved model are 81.4% and 60.3%，respectively，representing improvements of 4.2% and 1.9% over the original RT-DETR model，making it useful for low-light scene object detection tasks.

Key words：saliency guidance；low-light images；RT-DETR algorithm；attention mechanism

目标检测在自动驾驶、智能监控、人机交互等场景中广泛应用[1]。在低光照条件下，图像易受光照强度变化和光线折射的影响，导致对比度降低、颜色失真，使目标与背景的边界难以区分。真

收稿日期：2025－10－16

基金项目：国家重点研发计划课题(2022YFC3302502)

作者简介：魏英姿(1973—)，女，教授，博士。

实场景中目标常因遮挡和重叠导致轮廓模糊,难以完整识别。若目标尺寸较小或特征不明显,检测模型容易将其忽略,进一步增加了检测难度。

目标检测技术分为基于传统的检测方法和基于深度学习的目标检测方法两大类。基于深度学习的目标检测算法主要包括单阶段和双阶段检测算法。Faster-RCNN[2]是两阶段检测算法的代表,首先生成可能包含目标的区域建议,然后对每个建议进行特征提取与分类。与双阶段检测算法不同,单阶段检测算法,如 SSD[3]和 YOLO 系列[4],省略了候选区域生成步骤,直接从图像中回归目标位置与类别,提高了检测效率。

单阶段算法缺乏精细的区域建议机制,在目标密集、遮挡等复杂场景中,其特征表达能力较弱,易出现漏检问题,从而限制了其在复杂任务中的适用性。为此,研究者们针对不同的问题设计了多种改进策略。文献[5]提出基于深度时空残差卷积神经网络的算法,利用时空特征学习实现多目标、遮挡条件下的目标识别。文献[6]通过主干网络的可变形层聚合模块及多尺度注意力并联增强模块提升检测性能,有效解决目标之间相互遮挡造成的漏检。近年来,Transformer 架构逐渐被引入目标检测任务。文献[7]提出了基于 Transformer 架构的目标检测算法 DETR,但由于缺乏有效的区域建议机制,在目标密集的场景下性能有限。文献[8]在 DETR 的基础上构建了 RT-DETR,其在遮挡、密集和小目标场景中检测性能更加良好,同时提升了训练速度和检测精度。

以 DETR 为代表的 Transformer 检测算法在复杂场景中仍存在两方面亟待解决的挑战。一方面是缺乏对关键区域的聚焦能力,另一方面是全局注意力的高计算开销。文献[9]提出跨空间高效多尺度注意力机制,借助并行子网络构建大感受野结构,但其在处理图像时对所有区域一视同仁,忽略了前景与背景在信息价值上的差异。全局自注意力需要对整幅图像进行密集计算,导致显著的计算与存储负担。文献[10]利用高效的部分卷积对 DETR 进行改善,但仍因背景噪声的干扰问题,导致对小目标、遮挡目标或低对比度目标检测性能的下降。视觉显著性作为模拟人类视觉注意力机制的重要技术,为目标检测任务提供了一种有效的先验引导。显著性图能够突出与任务相关的前景区域,同时抑制冗余的背景信息,从而在提升检测精度与效率方面展现出独特优势。文献[11]结合交互式特征增强和跨尺度学习,提出了一种融合 RGB 与深度信息的显著物体检测方法。文献[12]利用显著性图做引导,对图像做增强处理,突出夜间场景中目标边缘轮廓,缓解 Faster-RCNN 检测器特征冗余和目标丢失问题。但在 Transformer 架构下,如何高效、系统性地融合显著性引导信息,仍是一个尚未充分探索的研究方向。

本文提出一种显著性引导优化框架 SG-DETR(saliency guided detection Transformer),主要思路如下:

1)在 RT-DETR 中引入视觉显著性先验信息,将显著性特征与检测网络深层特征融合,以增强语义感知;

2)利用门控调制特征融合模块[13],根据显著性引导对特征级数据进行动态特征筛选与重构,以提高多尺度特征的表达力;

3)采用显著性驱动的动态损失重加权机制,以增强模型训练阶段对显著区域的关注。

1　RT-DETR 目标检测模型

RT-DETR 是基于 DETR 改进的实时目标检测模型,其创新性地融合了 Transformer 全局建模能力与高效检测技术,突破了传统 DETR 模型实时性不足、小目标检测性能有限的瓶颈。RT-DETR 通过架构的深度重构与训练策略的优化,实现了效率与精度的协同提升。骨干网络采用了 ResNet 神经网络与多层次特征融合以增强特征提取能力;基于对象交互关系的高效混合编码器,包括基于注意力的尺度内特征交互模块(AIFI)和基于卷积神经网络的跨尺度特征融合模块(CCFF),提高了特征表达能力与尺度适应性;解码器采用 IoU 感知查询模块,提高目标框的定位精度;采用全连接网络输出头输出预测结果。

2　视觉显著性引导的目标检测

本文提出的基于视觉显著性引导的目标检测框架如图 1 所示。首先,利用显著性检测模型预测图像并生成高置信度的显著性图,该图为目标候选框的生成提供区域约束,有助于减少冗余候选框的数量。该方法基于视觉显著性原理,在通道维度上对输入图像进行增强,通过抑制低响应区域来降低背景复杂纹理的干扰,从而减少计算资源的消耗。这有助于缓解目标特征在提取过程

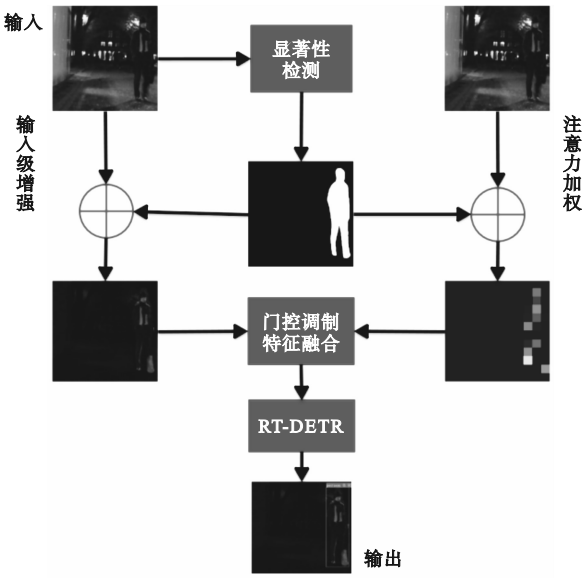


图 1 视觉显著性引导的目标检测架构

Fig. 1 Object Detection Architecture Guided by Visual Saliency

因背景噪声而导致的特征稀释问题,提升模型对关键目标的表征能力。其次,利用显著性图进行注意力加权操作,通过门控调制特征融合模块,利用特征级注意力引导机制将注意力权重融入模型的不同维度特征图中,确保注意力权重与真实目标区域的精准匹配,实现特征级的融合。最后,构建了显著性驱动的动态损失重加权机制,对高显著性区域赋予更高的损失权重,促使网络更准确地拟合显著区域的目标;同时对负样本进行降权,以减轻对背景区域的惩罚。经过显著性引导增强的特征图输入至目标检测的分类与回归分支,实现对目标的精准定位与类别判断。

3 SG-DETR 网络架构

为提升复杂环境下目标检测精度,在 RT-DETR 模型基础上引入视觉显著性先验,构建一个能够有效融合多尺度、多层次信息的整体网络结构。改进后的模型整体结构如图 2 所示。同 RT-DETR 架构,SG-DETR 也由骨干网络、基于对象交互关系的混合编码器、解码器及全连接网络预测头四个部分组成。

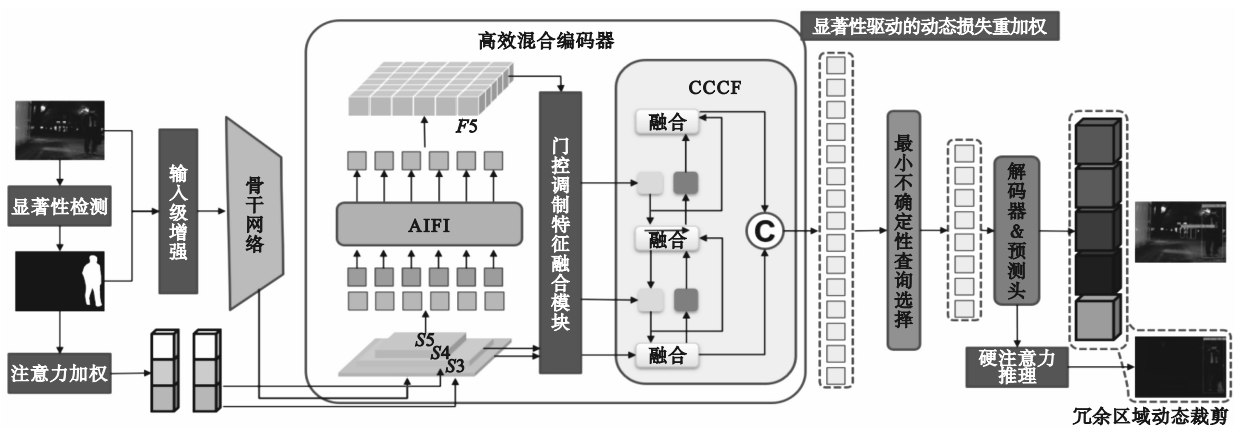


图 2 SG-DETR 网络结构

Fig. 2 Structure of SG-DETR network

3.1 输入级显著性增强

SG-DETR 模型的输入级加权借助 ABiU-Net[14] 模型生成对应输入图像的显著性图。显著性图反映了图像中不同区域对于目标检测任务的重要程度,将生成的显著性图与原始输入图像进行像素级的融合操作。通过对原始输入图像的每个像素值按在显著性图中对应位置的显著值进行加权调整,对原始图像中显著值 $M_{\text{sal}}(x,y)$ 接近 1 的目标区域,通过乘法运算放大其像素亮度与对比度。

$$I_{\text{weighted}}(x,y)=I_{\text{orig}}(x,y)\times M_{\text{sal}}(x,y) \quad (1)$$

式中:$I_{\text{orig}}(x,y)$ 为原始输入图像在像素坐标 (x,y) 处的像素值;$M_{\text{sal}}(x,y)$ 为显著性图在像素坐标 (x,y) 处的像素值;$I_{\text{weighted}}(x,y)$ 为加权后的输入图像在像素坐标 (x,y) 处的像素值。在输入阶段就强化了关键区域的表达,使模型为后续浅层特征 S_3、S_4 的目标定位奠定基础,从而提高对小目标与复杂背景目标的检测能力。输入级处理结果如图 3 所示。

3.2 编码器部分

如图 2 所示,在 SG-DETR 框架的编码器部分,颈部特征融合依旧采用高效混合编码器。骨干网络输出三个不同尺度的特征图 S_3、S_4、S_5。首先,引入视觉显著性先验,利用 ABiU-Net 模型生成显著性图,在输入阶段,基于显著性图对输入



图 3　输入级处理结果

Fig. 3　Input-level processing result

图像进行显著性加权增强,强化关键区域表达,提升小目标与复杂背景目标的检测能力。对于特征图 S_5,运用 AIFI 后得到 F_5。同时,在特征级,通过门控调制特征融合模块,利用显著性引导对 S_3、S_4 以及 F_5 进行动态特征筛选与重构,该模块能依据显著性信息,自适应地调整不同特征图在融合过程中的权重与贡献度,提升模型在多尺度场景下的表达力。其次,将经过门控调制特征融合模块处理后的 S_3、S_4 以及 F_5 三个不同尺度的特征图,通过基于卷积神经网络的 CCCF 进行融合,得到最终输出的查询特征。

3.3　特征级显著性引导

在模型特征处理流程中,骨干网络会输出多尺度特征图 S_3、S_4、S_5,这些特征虽包含目标的层次化语义信息,但在背景干扰密集、小目标占比高的复杂场景下仍存在关键区域特征响应不足的问题。为此,在特征级设计注意力加权与门控调制融合的两级处理机制。首先通过显著性先验增强目标区域特征表达,再通过门控调制特征融合模块平衡显著性特征与原始特征的贡献,充分利用视觉显著性的引导价值,规避显著性预测误差带来的负面影响,为后续混合编码器的特征交互与解码器的精准解码奠定基础。

以骨干网络输出的第 l 个特征图 $\boldsymbol{F}_l \in \mathbf{R}^{C\times H_l\times W_l}$ 为例,其中 C 为特征通道数,H_l,W_l 分别为该尺度特征图的高度与宽度,结合对应尺度下的显著性图 $S_l \in R^{1\times H_l\times W_l}$,注意力加权过程如图 4 所示。

利用空间加权将显著性图作为动态注意力掩码,通过逐元素乘法将显著性信息嵌入特征图,对特征图每个空间位置进行响应调整,使模型更聚焦于目标所在的区域。过程可表达为

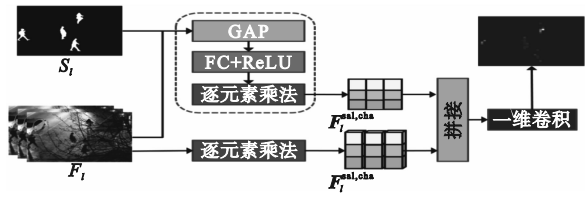


图 4　注意力加权计算过程

Fig. 4　Attention-weighted computation process

$$\boldsymbol{F}_l^{\mathrm{sal,spa}} = \boldsymbol{F}_l \odot (1 + \gamma \cdot \boldsymbol{S}_l) \tag{2}$$

式中:$\odot$ 表示矩阵逐元素乘法;γ 为可学习超参数。操作直接作用于模型的 S_3,S_4 的浅层细节特征与 S_5 的深层语义特征,确保不同尺度下的目标区域均能获得特征响应增益。

不同通道的特征对应目标的不同语义属性,仅通过空间加权难以实现通道维度的针对性增强。为进一步进行全局信息聚合,生成通道权重向量,实现特征的跨通道动态适配。首先,通过全局平局池化(GAP)遍历显著性图 $\boldsymbol{S}_l$ 所有空间位置像素,计算均值以生成单值全局显著性强度指标 $\bar{s}_l = \mathrm{GAP}(\boldsymbol{S}_l)$,将显著性图信息压缩为全局单值,实现从空间维度到单维度的降维,避免单一局部区域的显著性误差对通道权重产生过度干扰;随后,将 $\bar{s}_l$ 输入至含 ReLU 激活函数的全连接层(FC),通过非线性映射生成与特征通道数完全对齐的权重向量 $\boldsymbol{w} \in \mathbf{R}^C$;最后,利用逐通道乘法实现特征重标定,过程可表示为

$$\boldsymbol{F}_l^{\mathrm{sal,cha}}(c) = \boldsymbol{w}(c) \cdot \boldsymbol{F}_l(c) \tag{3}$$

式中:$\boldsymbol{F}_l(c)$ 表示 $\boldsymbol{F}_l$ 的第 c 个通道特征图。从而完成通道维度的动态适配,提升特征图的语义区分度。

综合空间与通道加权的优势,将 $\boldsymbol{F}_l^{\mathrm{sal,spa}}$ 与 $\boldsymbol{F}_l^{\mathrm{sal,cha}}$ 通过通道维度拼接与 1×1 卷积压缩维度,得到最终的加权特征 $\boldsymbol{F}_l^{\mathrm{sal}} \in \mathbf{R}^{C\times H_l\times W_l}$。过程可表示为

$$\boldsymbol{F}_l^{\mathrm{sal}} = C_{1\times1}(\mathrm{Concat}(\boldsymbol{F}_l^{\mathrm{sal,spa}}, \boldsymbol{F}_l^{\mathrm{sal,cha}})) \tag{4}$$

式中:Concat 表示拼接输出特征;$C_{1\times1}$ 表示 1×1 卷积层。不同图像的注意力加权结果如图 5 所示。

考虑到显著性检测模型的预测可能存在漏检、误检的局限性,直接引用可能引入噪声特征,反而影响检测性能。为此在模型中加入了门控调制特征融合模块,通过动态门控权重平衡与原始特征的贡献,实现鲁棒性特征融合。门控模块如图 6 所示。

门控权重以原始特征和显著性加权特征的差异为依据,通过卷积层提取交互信息,经过Sig

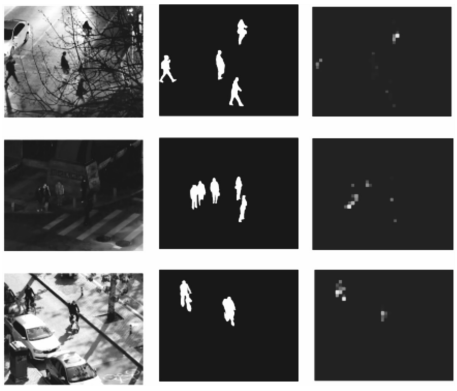


图 5 不同图像的注意力加权图

Fig. 5 Attention-weighted maps of different images

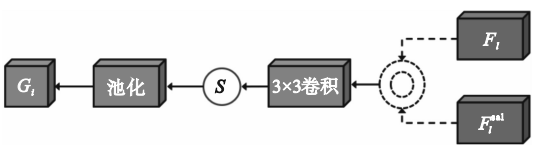


图 6 门控调制特征融合模块

Fig. 6 Gate-controlled Modulation Feature Fusion Module

moid 激活将权重约束至[0,1]区间,过程可表示为

$$\boldsymbol{G}_l = \sigma(C_{3\times3}(\mathrm{Concat}(\boldsymbol{F}_l, \boldsymbol{F}_l^{\mathrm{sal}}))) \tag{5}$$

式中:$\boldsymbol{G}_l \in \mathbf{R}^{C\times H_l\times W_l}$,为门控模块筛选出的特征图;$\sigma(\cdot)$ 为 Sigmoid 激活函数;元素 G_l 的取值接近 1,表示显著性加权有效,若该值接近 0,则表明显著性图可能包含噪声。

基于 G_l,最终的融合特征 F_l^{fused} 通过加权求和实现,过程可表示为

$$\boldsymbol{F}_l^{\mathrm{fused}} = \boldsymbol{G}_l \odot \boldsymbol{F}_l^{\mathrm{sal}} + (1-\boldsymbol{G}_l)\odot \boldsymbol{F}_l \tag{6}$$

融合使模型具备自适应调整能力,在显著性预测可靠的区域,G_l 趋近于 1,主要依赖 F_l^{sal} 的增强特征;反之,G_l 趋近于 0,模型自动保留更多原始特征信息,从而规避显著性误差的负面影响。

3.4 解码器部分

改进后的 SG-DETR 模型中,解码器对最小不确定性查询选择筛选出的特征进行解码。解码器由多层 Transformer 的解码器层组成,对筛选出的查询进行逐级迭代解码,以此逐步修正预测结果。在模型中,融入了显著性相关的机制,在解码过程中,利用显著性先验信息辅助解码,增强对显著区域目标的解码准确性。通过全连接层输出最后的检测框和类别信息。在训练过程中,所有解码器层的分类头和定位头都会输出位置信息和类别信息,并且结合显著性驱动的动态损失重加权机制,这些输出信息会被纳入损失计算,用于约束各个级别的解码器层,增强模型对显著区域目标的关注,提高收敛效率与检测鲁棒性。解码器部分结构如图 7 所示。

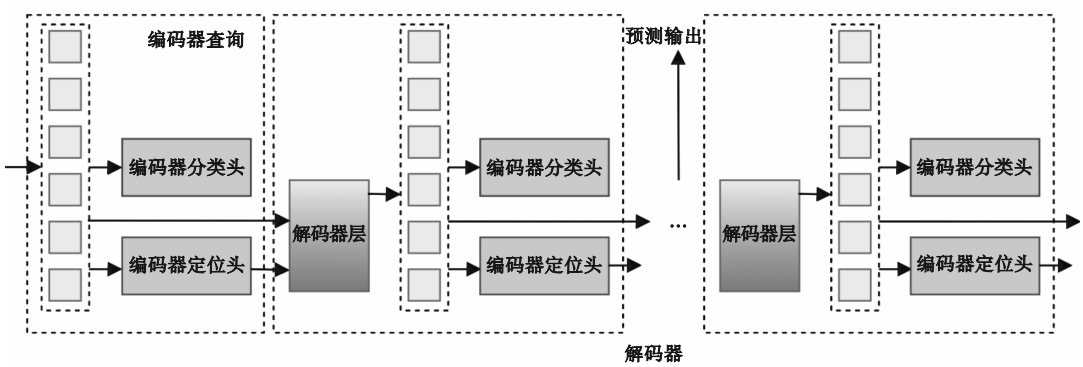


图 7 解码器部分结构

Fig. 7 Decoder partial structure

3.5 动态损失函数

为使模型聚焦于显著区域,定义显著性加权的检测损失,通过动态调整损失权重,增强模型对显著区域目标的关注,提升收敛效率与检测鲁棒性。

对于数据集中的每个标注目标实例 i,设其在图像中的中心坐标为(x_i, y_i),利用显著性图 S,获取该实例中心位置的显著值 $S(x_i, y_i)$。基于显著值,为每个目标实例 i 分配动态损失权重 w_i,过程表示为

$$w_i = 1 + \gamma \cdot S(x_i, y_i) \tag{7}$$

式中:γ 为可调节超参数,用于控制显著性对损失权重的影响。$S(x_i, y_i)$ 较大时,w_i 显著大于 1,该实例的损失在总损失中占比更高;反之损失占比相对正常。

结合上述动态权重,模型的总检测损失 L 被

定义为各目标实例检测损失的加权和,过程表示为

$$L = \sum_i \omega_i \cdot L_{\text{det}}(i) \tag{8}$$

式中:$L_{\text{det}}(i)$为实例 i 的基础检测损失。该损失函数显著区域的损失权重更大,促使模型在反向传播优化过程中,优先对显著区域的检测误差进行修正,强化对显著区域目标的学习。

4　实验与分析

4.1　实验数据集与环境配置

为了验证方法的有效性,构建复杂场景混合数据集。本研究以用于微光视觉的可见 – 红外配对数据集 LLVIP 为基础,LLVIP 包含 15 488 对可见 – 红外图像,使用海康威视 ds-2td8166bjzfy-75h2f/v2 双目相机平台,在晚上 6 ~ 10 点之间从 26 个不同的地点收集,涵盖 24 个不同的夜晚场景和 2 个白天场景。本研究对数据集进行筛选,去除不包含显著目标与相似度高的图像,并采用自然图像质量评估器(NIQE)进行筛选,最终选择 NIQE 大于 6.78 的图片 1 500 张。为保证数据的多样性,采用 YLLSOD 数据集[15]进行扩充。选用领域内通用图像标注工具 Labelme,秉持准确性、样本多样性等原则进行标注,构建包含 3 960 张图片,大于 7 000 个目标实例,涵盖室内、室外、遮挡、重叠等复杂场景与其相应真值标签的数据集 LlSOD,涵盖行人,汽车,路障在内的多类目标。采用 7∶2∶1 比例划分,其中训练集 2 772 张、验证集 792 张、测试集 396 张,确保各子集的场景分布与目标特征保持一致。数据集样本如图 8 所示。

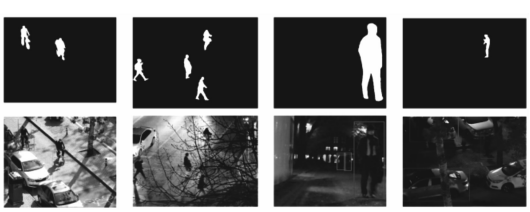


图 8　数据集样本

Fig. 8　Dataset samples

实验硬件配置:NVIDIA GeForce RTX 4090 32 G GPU。软件环境:CUDA 11.8;torch 2.0.1;python 3.8。训练设置大多沿用 RTDETR 的默认训练设置,优化器使用 AdamW,初始学习率为 5×10^{-5},批大小设置为 16,运行 150 轮次,输入图像大小为 640 × 640。

4.2　评价指标

在模型的性能评估中,从精度、效率等维度构建评价体系。精度方面,采用平均精度均值(mAP)作为核心指标,评估模型在不同 IoU 阈值下的综合检测能力;同时计算不同目标尺度下的平均精度 AP_S、AP_M、AP_L 以反映模型对多尺度目标的适应性,并引入边界框偏移误差(BBOX-Err)和背景误检率(FPR)量化显著性引导机制对定位精度与背景抑制的优化效果。效率方面,通过参数量(Params)和浮点运算数(FLOPs)衡量模型计算复杂度,直观展示模型的进一步优化和改进方向。

4.3　消融实验

以 RT-DETR 为基线模型,通过逐一添加核心模块构建 5 组消融实验,分析各模块的功能价值。所有模型均在 LlSOD 训练集上训练。消融实验结果如表 1 所示。第 1 组实验采用原始 RT-DETR 模型;第 2 组实验在原始 RT-DETR 模型的基础上添加输入级显著性增强(IE),使 mAP@0.5∶0.95 提升 1.2%,AP_S 提升 1.7%,表明了显著性在增强目标区域特征响应上的基础作用;第 3 组实验在第 2 组实验的基础上新增特征级显著性引导(CA),由于显著性检测模型自身的预测误差被直接引入特征增强流程,导致噪声干扰抵消了有效特征的增强效果,各项系数提升较少;第 4 组实验在先前实验的基础上增加门控特征融合模块(GM),mAP 提升 1.7%,FPR 从 0.142 降至 0.105,印证了门控机制对特征差异的可靠性判断,平衡特征贡献,提升模型对复杂场景的适配性;第 5 组实验即 SG-DETR,包括动态损失函数,在计算量和参数量仅小幅增加的情况下,实现了更高的检测精度。

为验证 GM 对显著性预测误差的抑制能力,通过引入不同水平的误差,测试 GM 模块在误差干扰下的性能保持能力。以 ABiU-Net 生成的显著性图为基准,通过随机选取背景区域,将其显著值从 0 提升至 0.8 ~ 1.0,模拟误检现象,误检区域占比分别为 0%、10%、20%。在 LlSOD 验证集上,分别采用不同误差水平的显著性图,以 FPR、AP_S、BBOX-Err 为核心指标,对比模型在不同误差下的性能衰减幅度,量化 GM 模块的鲁棒性。性能对比结果如表 2 所示。

表 1 采用不同策略的 DETR 模型消融实验结果

Table 1 Ablation experiment results of DETR models using different strategies

编号	策略				mAP@0.5/%	mAP@0.5:0.95/%	AP_S/%	BBOX-Err/%	FPR/%	Params/10^6	GFLOPs
	IE	CA	GM	DLW							
1					0.762	54.8	0.685	0.182	0.156	19.9	57.0
2	√				0.774	55.2	0.702	0.165	0.142	20.1	57.4
3	√	√			0.788	55.9	0.728	0.138	0.105	20.7	58.6
4	√	√	√		0.805	56.1	0.746	0.121	0.098	20.4	58.2
5	√	√	√	√	0.814	56.7	0.763	0.112	0.087	20.5	58.9

表 2 不同误差水平下 SG-DETR 性能比较

Table 2 Performance comparison of SG-DETR under different error levels

策略	误差/%	FPR/%	BBOX-Err/%	AP_S/%
含 GM	0	0.087	0.112	0.763
	10	0.098	0.125	0.735
	20	0.112	0.140	0.702
无 GM	0	0.105	0.138	0.728
	10	0.153	0.159	0.965
	20	0.179	0.181	0.631

4.4 可视化分析

为了更直观地展示所提方法的优点,算法改进前后的检测结果如图 9 所示。图 9(a)为原始 RT-DETR 模型检测效果,图 9(b)为经显著性增强后的 SG-DETR 检测结果。原始模型在复杂场景下,受边缘纹理等影响,容易出现漏检,误检,改进后的模型能够准确识别出图片中的关键区域信息,表现出更好的检测性能。因此,在各种复杂条件下,本文方法实现了稳定的关键区域定位与识别性能。



图 9 模型检测结果对比图

Fig. 9 Comparison chart of model detection results

5 结论

在 RT-DETR 目标检测任务中,Transformer 的全局自注意力机制需对整幅图像进行密集计算,导致其缺乏对关键区域的聚焦能力。利用视觉显著性引导 RT-DETR 模型数据分析,分别对输入数据、特征级数据以及损失函数实行多维度显著性引导,有效解决了复杂场景下目标定位不准、背景干扰严重的问题。门控调制模块显著降低了模型对显著性预测精度的依赖,增强了模型的场景适配性。在低光、遮挡、小目标等细分复杂场景中展现出较好的性能。

参考文献(References):

[1] 何迪,王静. 公共安全视频监控联网信息安全检测平台的设计与实现[J]. 中国安防,2020(12):17-20.
He D, Wang J. Design and implementation of public security video surveillance network information security detection platform[J]. China Security & Protection, 2020(12):17-20. (in Chinese)

[2] 王君,张德育,康鑫英. 改进 Faster-RCNN 的低空小型无人机检测方法[J]. 沈阳理工大学学报,2021,40(4):23-28.
Wang J, Zhang D Y, Kang X Y. Improved detection method of low altitude small UAV by faster-RCNN[J]. Journal of Shenyang Ligong University, 2021, 40(4):23-28. (in Chinese)

[3] 吴钟仁,周莲英,丁腊春. 基于双向融合和特征增强的 SSD 小目标检测算法[J]. 计算机与数字工程,2025,53(1):115-118,175.
Wu Z R, Zhou L Y, Ding L C. SSD small target detection algorithm based on bidirectional fusion and feature ehancement[J]. Computer & Digital Engineering, 2025, 53(1):115-118,175. (in Chinese)

[4] 冯迎宾,刘艾妮. 基于 YOLOv8n 的夜间车辆检测[J]. 沈阳理工大学学报,2025,44(2):1-6,12.
Feng Y B, Liu A N. Night-time vehicle detection based on YOLOv8n[J]. Journal of Shenyang Ligong University, 2025, 44(2):1-6,12. (in Chinese)

(下转第 84 页)