文章编号:1000-5854(2026)05-0443-10

DNEA-YOLO:一种融合双重归一化高效注意力的遥感目标检测模型

刘 斌[1], 李晓帅[2], 叶 刚[1], 梁 超[1], 胡成玺[2], 邱 磊[2]

(1.国家能源集团宁夏煤业有限责任公司 羊场湾煤矿,宁夏 银川 750411;
2.中国矿业大学(北京) 人工智能学院,北京 100083)

摘 要:物联网数据和遥感的融合极大地扩展了遥感领域的目标检测能力,然而遥感目标检测仍面临复杂背景干扰、密集小目标易漏检、模型难以兼顾精度与边缘部署等问题.为此,提出双重归一化高效注意力 YOLO(dual normalization efficient attention YOLO,DNEA-YOLO)模型,在满足边缘部署需求的同时提升检测精度.在模型设计中,提出双重归一化高效注意力机制,降低计算开销并增强特征选择能力;针对复杂场景特征提取不足问题,将大核选择模块(large selective kernel,LSK)嵌入 C3K2 模块的特征融合分支,形成 C3K2-LSK 集成结构,提升复杂背景下的特征表达能力;针对遥感图像密集目标易漏检问题,采用软非极大值抑制(soft non-maximum suppression,SOFT-NMS)优化后处理策略,进一步提升目标识别准确性.在 SIMD 公开数据集上的实验表明,DNEA-YOLO 参数量为 9.97 M,浮点运算量为 25.2 GFlops,mAP 较基线模型 YOLO11s 提升 2.8%.通过与多个先进模型进行对比实验,结果充分验证了 DNEA-YOLO 在遥感图像目标检测任务中的有效性与优越性.

关键词:目标检测; 遥感图像; 注意力机制; 边缘部署; YOLO

中图分类号:TP 751.1 **文献标志码**:A **doi**:10.13763/j.cnki.jhebnu.nse.202601012

DNEA-YOLO:a remote sensing object detection model integrating dual normalization and efficient attention

Liu Bin[1], Li Xiaoshuai[2], Ye Gang[1], Liang Chao[1], Hu Chengxi[2], Qiu Lei[2]

(1.Yangchangwan Coal Mine,CHN Energy Ningxia Coal Industry Co.,Ltd.,Yinchuan Ningxia 750411,China;
2.School of Artificial Intelligence,China University of Mining and Technology(Beijing),Beijing 100083,China)

Abstract:The integration of IoT data and remote sensing has greatly expanded the object detection capability in the field of remote sensing.However,remote sensing object detection still faces problems such as complex background interference,easy missed detection of dense small targets,and difficulty in balancing accuracy and edge deployment.Therefore,a dual normalization efficient attention YOLO(DNEA-YOLO) model is proposed to improve the detection accuracy while meeting the requirements of edge deployment.In the model design,the dual normalization efficient attention mechanism is proposed to reduce the computational overhead and enhance the feature selection ability.Aiming at the problem of insufficient feature extraction in complex scenes,the large selective kernel(LSK) is embedded into the feature fusion branch of the C3K2 module to form a C3K2-LSK integrated structure to improve the feature expression ability in

收稿日期:2025-11-06; **修回日期**:2026-06-14

基金项目:国家自然科学基金(52274160); 内蒙古自治区重点研发和成果转化计划项目(2025YFHH0001)

作者简介:刘 斌(1984-),男,宁夏银川人,高级工程师,研究方向为煤矿信息处理、矿井安全监控.

通信作者:李晓帅(1998-),男,硕士研究生,研究方向为图像处理.E-mail:1036103149@qq.com

complex backgrounds. Aiming at the problem of missed detection of dense targets in remote sensing images, soft non-maximum suppression (SOFT-NMS) is used to optimize the post-processing strategy to further improve the accuracy of target recognition. Experiments on the SIMD public dataset show that the DNEA-YOLO parameter is 9.97 M, the floating-point operation is 25.2 GFlops, and the mAP is 2.8% higher than the baseline model YOLO11s. The comparison results with several advanced models verify the effectiveness and superiority of DNEA-YOLO in remote sensing image object detection tasks.

Key words: object detection; remote sensing image; attention mechanism; edge deployment; YOLO

近几十年来,随着遥感技术的不断发展和进步,卫星、航空等遥感平台采集的遥感图像数据的分辨率和覆盖面积不断提高,使遥感数据成为地球表面地理空间信息获取的核心手段之一.

诸多研究探索了物联网与遥感技术融合的目标检测方法.2022 年,Ahmed 等[1]结合了物联网与深度学习,提出智能远程监控应用(空中和无人机)以改善智慧城市的监控应用.Xie 等[2]提出一个轻量级模型,旨在将地面端的高性能计算任务迁移至在轨卫星平台,实现在轨解译遥感影像.2023 年,Lu 等[3]提出了一种任务启发式的星地协同目标检测框架,为边缘计算在卫星平台的部署提供了可行的技术方案.

尽管该领域已取得诸多研究进展,但其实际应用仍面临诸多挑战,核心问题在于现有深度学习目标检测模型直接应用于遥感目标检测任务时,检测性能表现不佳.由于图像的观测特性不同,现有的目标检测模型不能直接应用于遥感图像中的目标检测.例如,遥感目标常具有单一的外观特征,如道路、平坦停车场、车辆顶部及建筑物屋顶的纹理信息缺失或稀疏.在车辆等可移动物体的检测任务中,同一幅图像中目标尺度差异显著.同时,遥感图像的分辨率范围很广,从几米到亚米级不等,而现有深度学习目标检测模型大多基于分辨率相对统一的自然图像数据集训练,其对于遥感图像中这种大幅度变化的分辨率适应性较差.此外,复杂背景下的密集目标检测,始终是目标检测算法面临的研究难点,因此遥感图像的目标检测领域仍然面临着诸多挑战:密集小目标和被遮挡的目标;目标尺寸的跨度大;目标背景复杂;数据集分辨率差异大.综上,开展遥感目标检测相关研究,对提升遥感技术的工程应用价值、推动相关领域技术发展具有重要的理论与实践意义.

本文的主要贡献如下:

1) 提出双重归一化高效注意力(dual normalization efficient attention, DNEA)机制,在不增加模型参数量与浮点运算量的前提下,实现目标检测性能的提升;

2) 引入大核选择模块(large selective kernel, LSK),设计了适配 YOLO11s 的 C3K2-LSK 集成部署结构,实现了更优的模块融合方式;

3) 引入软非极大值抑制(soft non-maximum suppression, SOFT-NMS)策略,为双重归一化高效注意力 YOLO(dual normalization efficient attention YOLO, DNEA-YOLO)模型设计了精度优先的后处理方案,进一步提升检测准确性.

1 相关工作

1.1 遥感图像中的目标检测网络

计算机视觉领域的诸多研究均采用深度学习方法应对遥感图像目标检测的各类挑战,并取得了显著成果.国内外学者提出的改进算法对基础模型进行了不同程度的优化,收获了阶段性研究成果,但在平衡检测精度与模型轻量化程度方面仍有优化空间.此外,遥感目标检测场景的背景复杂性,易导致算法的检测性能下降.

2018 年,Deng 等[4]针对快速区域卷积神经网络在遥感图像多类目标检测中检测效果差的问题提出了一种新的多目标即时检测方法,对特征提取器进行了重新设计,增加了接收野大小的多样性.2019 年,涂兵等[5]提出融合递归滤波与 K 近邻算法的高光谱图像分类方法,经主成分分析降维与递归滤波强化空间结构后分类,提升分类精度.2021 年,Cheng 等[6]针对光学遥感图像中目标检测的挑战,提出了端到端跨尺度特征融合框架,有效提高了目标检测效果.2022 年,Liu 等[7]提出自适应平衡网络,通过一种增强的信道注意机制来减轻复杂背景对前景目标的阻碍.2024 年,周莲等[8]针对光学遥感图像小样本目标检测痛点,提出含协同注意力模块与对比学习分支的算法,经数据增强和两阶段微调,在相关数据集上提升平均精度.2025 年,刘

浩然等[9]改进 YOLOv7，融入数字高程模型数据、卷积块注意力模块及最小点距离交并比优化，实现滑坡检测与分割一体化.李愿等[10]构建双重注意力多尺度残差去雾网络，基于大气光幂模型，改善了遥感雾图质量与处理效率.

1.2 注意力机制

注意力机制已在计算机视觉领域得到广泛应用与验证，通过为特征分配不同的权重系数，引导神经网络聚焦输入特征图中的有效特征，从而显著提升目标检测模型的检测性能.

2018 年，Hu 等[11]提出了一种称为“挤压和激励”块的架构单元，通过权重矩阵从通道的角度对图像的不同位置赋予不同的权重.2020 年，Wang 等[12]提出了高效通道注意力机制，该机制使用一维卷积实现局部跨通道交互，避免了在学习通道间依赖关系时压缩和降维操作的影响.2021 年，Yang 等[13]提出简单无参数注意力模块(simple，parameter-free attention module，SIMAM)，可在不增加网络参数的前提下，为层中的特征映射推断出三维注意力权重，有效地提高了卷积神经网络的表示能力.Liu 等[14]提出一种全局注意机制(global attention mechanism，GAM)，通过减少信息损失和放大全局交互表征来提高深度神经网络的性能.针对众多前期研究专注于通过注意力操作捕获显著特征，但缺乏对权重贡献因素的考虑的问题，Liu 等[15]提出归一化注意力机制(normalization based attention module，NAM)，进一步抑制不重要的通道，提高模型精度.2022 年，Xia 等[16]提出可变形注意力机制，克服了常见注意力方法存在的内存计算成本高、受无关区域影响以及数据不可知等问题.2023 年，Shao 等[17]提出多维协作注意力(multidimensional collaborative attention，MCA)，通过构建 2 个并行轴向注意力之间的交互，更有效地利用多尺度特征和全局上下文，获取目标的全局信息.2024 年，Xu 等[18]针对红外小目标检测问题提出分层上下文融合网络，采用多分支特征提取策略能够有效捕获不同尺度和层次的特征信息.

2 方法论

2.1 整体结构图

为减少遥感图像目标检测中的漏检与误检现象，有效提升模型的遥感目标检测性能，以 YOLO11s[19]为基线模型，提出 DNEA-YOLO 模型，其结构如图 1 所示.

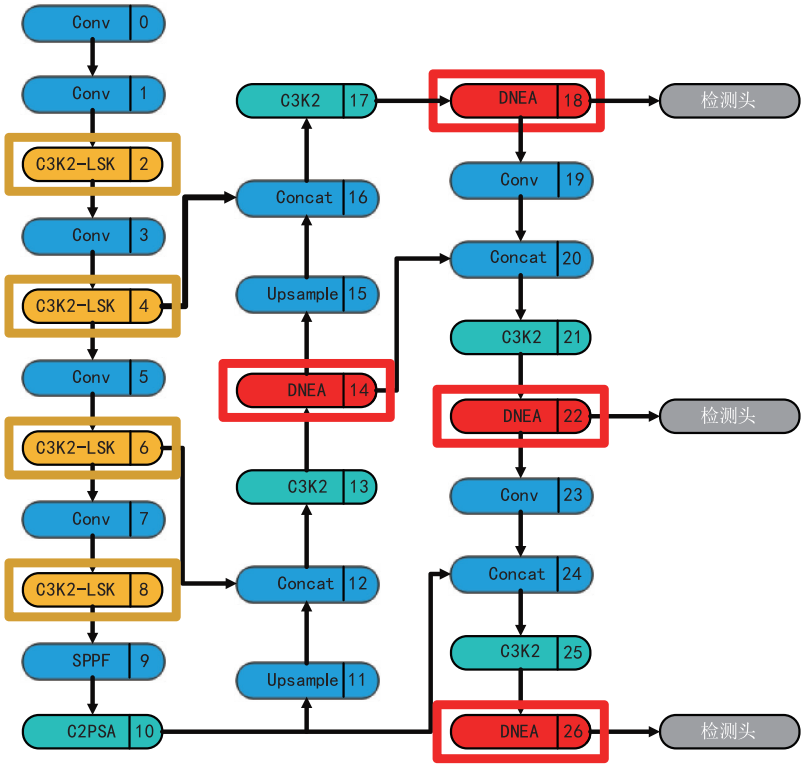


图 1 DNEA-YOLO 整体结构图

Fig.1 Overall structure diagram of DNEA-YOLO

本研究以 YOLO11s 为基线模型构建 DNEA-YOLO，整体结构包含骨干网络、颈部网络、检测头及后处理模块四大部分，相较于基线模型，具体改进及效果为

1）将骨干网络中的 C3K2 模块替换为 C3K2-LSK 集成模块，通过大核卷积捕获更广泛的上下文信息，提升复杂背景下小目标特征提取能力；

2）在颈部网络的特征融合层嵌入 DNEA 双重归一化注意力模块，在零参数量增加的前提下增强模型的特征选择能力，抑制噪声通道和无效空间区域；

3）在检测头的后处理阶段引入 SOFT-NMS 策略，替代传统 NMS，减少密集目标检测中的漏检情况，实现精度优先的检测目标.各改进模块相互协同，在保证模型轻量化的同时，提升整体检测性能.

2.2 DNEA 注意力

针对遥感检测模型难以兼顾低计算开销与高效特征提取能力的问题，本文在现有的轻量化注意力的基础上进行改进，提出了一种轻量级混合注意力 DNEA（图 2），图中 ω 为归一化权值，DA 表示模块的动态调节能力.

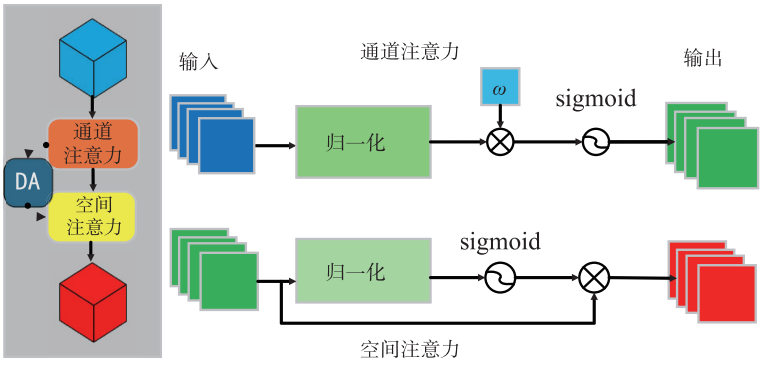


图 2　DNEA 的结构

Fig.2　The structure of DNEA

（1）给出了通道注意力的归一化方法.其中：x_1 为输入张量；μ_{x_1} 为小批量 x_1 的均值；σ_{x_1} 为小批量 x_1 的标准差；γ 和 β 是可训练仿射变换（缩放和平移）的参数.

$$y_1=\gamma\frac{x_1-\mu_{x_1}}{\sqrt{\sigma_{x_1}^2+\varepsilon}}+\beta. \tag{1}$$

（2）中：abs（γ）表示对 γ 参数取绝对值；ω 为基于 γ 计算得到的归一化权重，用于对每个通道的特征图进行加权.

$$\omega=\frac{\mathrm{abs}(\gamma)}{\sum\mathrm{abs}(\gamma)}. \tag{2}$$

（3）对每个通道的特征图进行加权，包括维度替换和恢复操作以便于计算.

$$y_2=\omega\cdot y_1. \tag{3}$$

（4）是通道注意力的残差连接.Sigmoid 用于将归一化特征图的值映射到（0，1）范围内.

$$x_2=\mathrm{sigmoid}(y_2)\times x_1. \tag{4}$$

（5）为空间注意力的归一化方法.其中：x_2 为单个样本值；μ 为数据样本的均值；n 为样本的数量；ε 为防止除零的小数，用以确保数值稳定性.同时该参数用以控制对高对比度区域的关注强度，通过调整 ε 参数可以实现自适应的空间聚焦.较小的 ε 值会增强对这些区域的聚焦，而较大的 ε 值则会导致更平衡的关注分布，从而使 DNEA 能够适应各种输入特征.

$$y_3=\frac{(x_2-\mu)^2}{4\left(\dfrac{\sum(x_2-\mu)^2}{n}+\varepsilon\right)}+0.5. \tag{5}$$

将 DNEA 模块嵌入至模型的颈部网络，以强化多尺度特征的融合效果.通道注意力通过从特征统计中

得出的归一化加权来强调信息丰富的通道并抑制噪声，而空间注意力则聚焦于显著区域.残差连接保留了原始特征.DNEA 注意力机制由通道注意力子模块和空间注意力子模块串行组成，核心通过“无参通道筛选＋自适应空间聚焦”实现特征增强，且无额外参数量和浮点运算量增加.通道注意力子模块复用归一化层的仿射缩放参数(γ)构建通道权重，先对输入特征做批量归一化以消除通道分布差异，再对 γ 参数取绝对值并归一化，生成各通道的重要性权重，通过逐通道加权实现重要通道特征强化、冗余通道特征抑制.空间注意力子模块基于特征的统计特性自适应生成空间权重，计算每个空间位置特征与通道内空间均值的平方差，通过全局归一化和正则化处理后，生成聚焦目标区域的空间权重，抑制背景噪声区域的特征响应.最终，输入特征经通道注意力加权后，再通过空间注意力加权，实现“通道-空间”双重特征增强，提升模型对遥感目标的特征区分能力.

2.3 C3K2-LSK

现有遥感目标检测研究多聚焦于优化目标边界框的表征方式，却忽视了小尺度遥感目标所处的背景上下文信息.针对这一问题，Li 等[20]提出了大选择核网络(large selective kernel network，LSKNet)，其核心 LSK 模块如图 3 所示.LSK 模块的大核卷积可以捕获更广泛的上下文信息，这对于像素信息稀疏的小尺度遥感目标检测尤为关键，进而有效提升模型对小尺度目标的识别与定位能力.

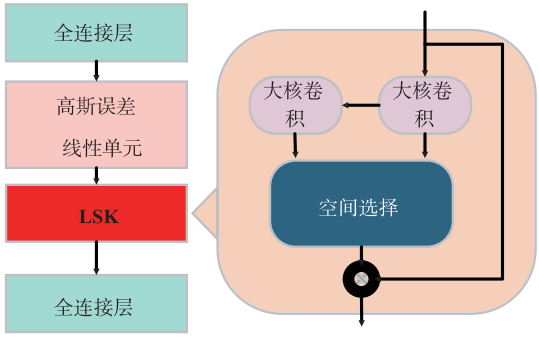


图 3 LSK 模块的基本结构

Fig.3 The basic structure of the LSK module

为降低 LSK 模块在模型训练中的过拟合风险，本文设计了一种更优的集成模块 C3K2-LSK.

跨阶段局部特征融合(cross-stage partial feature fusion，C2F)模块是 YOLOv11 中核心模块跨阶段局部核模块 2(cross-stage partial kernel-2，C3K2)的父类模块，是一种优化的跨阶段部分连接模块，旨在将特征图在不同阶段进行部分连接，以减少计算量并提高特征的复用性.在 C2F 模块的基础上引入了 LSK 模块，增强了网络对小目标特征提取能力.引入 LSK 模块后，C2F-LSK 如图 4 所示.C2F-LSK 能够进一步增强特征提取能力，经由 C3K2 模块继承后，命名为 C3K2-LSK.C3K2-LSK 模块融合大核卷积与动态注意力机制，可有效提升模型的多尺度特征处理能力，能够适配复杂背景下的遥感目标检测任务.

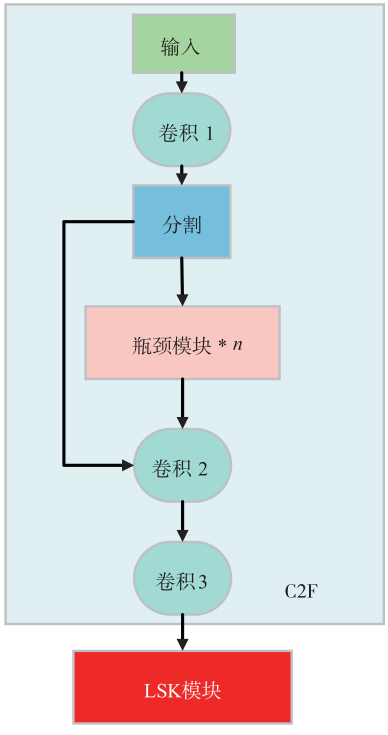


图 4 C2F 模块的流程及 LSK 模块的部署位置

Fig.4 The process of the C2F module and the deployment location of the LSK module

C3K2-LSK 集成模块部署优势体现在 3 个方面：1) 结构适配性，基于 YOLO11s 的核心模块 C3K2 的父类 C2F 模块进行改进，在 C2F 的特征融合层嵌入 LSK 模块，保留原始 C2F 的跨阶段特征复用优势，避免 LSK 模块与基线模型的结构冲突；2) 性能适配性，LSK 大核卷积嵌入在特征融合阶段，可对多尺度特征进行上下文信息补充，针对遥感小目标像素信息少的问题，提升小目标特征的可区分性，通过实验验证，单独引入可使 mAP 提升1.5%；3) 计算量可控性，相较于直接将 LSK 模块替换骨干网络的卷积层，本部署方式仅增加 3.6 GFlops 和 0.53 MParams，保证模型的轻量化特性，满足边缘部署需求.

2.4 SOFT-NMS

传统非极大值抑制算法在目标检测应用中存在固有缺陷：当目标检测框与置信度最高的检测框的交并比超过

预定义阈值时，该检测框的置信度会被置零，易造成目标漏检.针对该缺陷，Navaneeth 等[21]提出了 SOFT-NMS 软非极大值抑制算法.

在目标检测等计算机视觉任务中，模型通常会对同一目标生成多个候选检测框.这些检测框会有重叠，并且会对同一个目标进行多次检测.传统 NMS 的核心思想为，在候选检测结果中筛选出置信度局部最大值的检测框，并抑制其余非极大值检测框.传统 NMS 的检测框筛选策略为

$$s_i=\begin{cases}s_i, \mathrm{IoU}(\mathcal{M},b_i)<N_t,\\ 0, \mathrm{IoU}(\mathcal{M},b_i)\geqslant N_t.\end{cases} \tag{6}$$

其中：b_i 为初始检测框列表；s_i 为对应的检测分数；N_t 为 NMS 阈值；$\mathrm{IoU}(\mathcal{M},b_i)$为重叠值.

SOFT-NMS 改进策略：SOFT-NMS 通过修改传统 NMS 算法的得分衰减方式来解决这个问题.该算法不是将与 $\mathcal{M}$ 有高重叠的检测框的得分直接设置为零，而是根据与 $\mathcal{M}$ 的重叠程度来降低其分类得分.具体来说，提出了 2 种重新评分函数，SOFT-NMS 的框选策略为

$$s_i=\begin{cases}s_i, & \mathrm{IoU}(\mathcal{M},b_i)<N_t,\\ s_i(1-\mathrm{IoU}(\mathcal{M},b_i)), & \mathrm{IoU}(\mathcal{M},b_i)\geqslant N_t.\end{cases} \tag{7}$$

SOFT-NMS 相较于传统 NMS 具有显著优势.该算法通过降低重叠候选检测框的置信度而非直接剔除，有效避免了真实目标检测框的误删，提升了目标检测效果.遥感图像中存在大量密集小目标，如车辆、小型建筑，传统 NMS 在重叠阈值内会直接将非最高得分检测框置零，易造成密集目标的漏检，而 SOFT-NMS 通过根据 IoU 值衰减重叠框的置信度而非直接删除，更适配遥感密集小目标的检测场景.此外，远端遥感监测设备对检测精度的要求高于实时性，SOFT-NMS 的精度优先策略与边缘部署的遥感检测需求高度匹配，可在保证实时性满足需求的前提下，显著减少密集目标、遮挡目标的漏检情况.

3 实　验

3.1 实验设置

本研究采用统一的实验设置.GPU 为 NVIDIA GeForce RTX 2070s，深度学习框架版本为 Pytorch 2.0.0，Python 版本 3.10.15，使用 CUDA 11.8 版本的并行计算框架在 GPU 上进行计算.数据集统一训练 200 个轮次，输入分辨率为 640×640，批大小为 8，不使用预训练权重，优化器使用 SGD，初始学习率为0.01，最终学习率因子为 0.01，开启学习率预热 3 轮，阶段动量 0.8，使用指数移动平均，其余参数采用默认配置.

3.2 数据集及预处理

卫星图像多车辆数据集(satellite imagery multivehicles dataset，SIMD)[22]由巴基斯坦国立科学技术大学提出，具有多类别、中小目标分布丰富、高分辨率等特点，所有图像均为 RGB 格式，支持研究和开发，适用于遥感图像中的目标检测任务，特别是在车辆检测、场景监控和实时目标识别等领域.

由于原始数据集并未划分测试集，为保证模型的泛化性，防止对数据集的过拟合，本文将其按 8∶1∶1 重新划分训练集、验证集和测试集，同时统计了划分后各数据集的类别实例分布，确保按类别随机划分后的数据集符合比例要求，无严重的类别不平衡问题.

实验中所有图像均统一缩放至 640×640 分辨率，保持长宽比不变并以填充方式适配模型输入，同时将标注框坐标归一化至 0～1 区间以符合模型训练格式.训练阶段采用概率为 0.5 的随机水平翻转、色调变化为 ±0.015、饱和度变化为 0.7、亮度变化为 0.4 的 HSV 色彩空间扰动、最大偏移比例为 0.1 的随机平移与尺度范围 0.5 的随机尺度变换的数据增强策略提升样本多样性，并使用概率为 1.0 的 Mosaic 增强进一步丰富训练样本分布，在训练最后 10 个轮次关闭 Mosaic 增强，保证模型收敛稳定.

3.3 评价指标

mAP：指 mAP50-95(%)，目标检测中衡量模型性能的核心指标之一，能综合反映模型在不同重叠程度下的检测能力.

Flops：表示模型在一次前向传播过程中所需的浮点运算次数.

Params：表示模型中可训练参数的总数.参数数量直接影响模型的复杂度和计算量.

FPS：表示每秒处理的帧数，用于衡量模型在给定硬件上的处理速度.

3.4 注意力模块优势性

在 SIMD 数据集上，对比了 DNEA 和其他主流注意力机制的性能（表 1），所有注意力机制均嵌入至 YOLO11s 模型中进行实验.实验结果表明，DNEA 在保持相对较低的计算量（21.5 GFlops）和参数规模（9.43 M）的同时，实现了较高的平均精度（mAP 为 67.1%）.

表 1 DNEA 和其他注意力在 SIMD 验证集的表现

Tab.1 Performance of DNEA and other attentions in SIMD-VAL

模型	mAP/%	Flops/G	Params/M	FPS
YOLO11s	65.5	21.5	9.43	141
NAM[15]	66.2	21.5	9.43	141
SIMAM[13]	65.8	21.5	9.43	152
TripleAttention[23]	65.4	21.6	9.42	154
MCA[17]	65.7	21.3	9.42	244
iRMB[24]	65.3	62.5	10.8	128
MoCAttn[25]	65.2	21.8	9.94	152
GAM[14]	65.1	22.7	11.2	141
MLCA[26]	65.7	21.5	9.42	137
DNEA	67.1	21.5	9.43	135

3.5 消融实验

在 SIMD 数据集上开展消融实验，验证各改进模块的独立有效性及模块间的协同作用，并评估 DNEA-YOLO 模型的综合检测性能，实验结果见表 2. YOLO11s 为基线模型，DNEA-YOLO 同时部署 DNEA、C3K2-LSK 模块以及 SOFT-NMS.结果显示，DNEA-YOLO 在数据集上实现了最高的 mAP 值.尽管受 SOFT-NMS 的计算特性及实验硬件资源的限制，模型的 FPS 有所下降，但其推理速度仍能满足远端遥感监测设备的实时性要求.

表 2 在 SIMD 测试集上的消融实验结果

Tab.2 Results of ablation experiments on SIMD-TEST

模型	DNEA	C3K2-LSK	SOFT-NMS	mAP/%	Flops/G	Params/M	FPS
YOLO11s				63.3	21.5	9.43	141
1	√			65.0	21.5	9.43	135
2		√		64.8	25.1	9.96	128
3			√	64.7	21.5	9.43	85
4	√	√		65.2	25.2	9.97	116
5	√		√	65.7	21.5	9.43	88
6		√	√	65.3	25.1	9.96	79
DNEA-YOLO	√	√	√	66.1	25.2	9.97	71

从消融实验结果可明确各模块独立作用与协同机制：DNEA 注意力模块在不增加模型复杂度的前提下使 mAP 提升 1.7%，核心通过通道与空间双重特征增强抑制复杂背景噪声、强化目标区分性；C3K2-LSK 模块以少量计算量为代价使 mAP 提升 1.5%，依托大核卷积扩大感受野，解决小目标与密集目标的特征提取不足问题；SOFT-NMS 模块在无参数量与计算量增加的情况下使 mAP 提升 1.4%，通过改进后处理策略减少密集目标漏检，仅牺牲一定推理速度.三者协同形成“主干特征提取→颈部特征增强→检测后处理优化”的

全流程适配逻辑，部署于主干的 C3K2-LSK 先完成多尺度特征的基础提取，扩大感受野以捕获遥感小目标与复杂背景下的上下文信息，部署于颈部的 DNEA 再对该特征进行精准筛选与增强，抑制冗余背景噪声、强化有效目标特征，让后续后处理更具针对性，而经主干-颈部两级优化后的清晰特征，又能提升 SOFT-NMS 置信度判断与检测框筛选的精准性，减少有效检测框的误衰减问题.最终三者组合的 DNEA-YOLO 实现了 2.8%的 mAP 最优提升，形成无功能冗余、计算代价可控、性能效果非线性放大的协同优化效果，有效提升了模型对遥感目标的检测能力.

3.6 对比实验

将 DNEA-YOLO 模型与其他主流遥感目标检测模型进行性能对比实验.由表 3 可知，与其他模型相比，DNEA-YOLO 在保持参数量与浮点运算量均为轻量化水平的前提下，取得了最优的检测性能.

表 3 SIMD 测试集上不同模型结果的对比

Tab.3 Comparison of results from different models on the SIMD-TEST

模型	mAP/%	Flops/G	Params/M	FPS
MLCA(2023)[26]	58.3	—	—	—
RS-DETR(2024)[27]	63.1	53.3	17.5	—
YOLO-SSP(2024)[28]	64.9	—	28.4	—
YOLOv8s(2023)	62.4	28.6	11.1	156
YOLOv9s(2024)	64.5	27.0	7.18	143
YOLOv10s(2024)	64.0	24.7	8.06	172
YOLO11s(2024)	63.3	21.5	9.43	141
YOLOv12s(2025)	63.7	19.4	9.12	128
YOLOv13s(2025)	61.8	20.7	9.01	185
DNEA-YOLO	66.1	25.2	9.97	71

图 5 为 DNEA-YOLO 在 SIMD 数据集上的可视化检测结果.可视化结果进一步验证了 DNEA-YOLO 模型的目标检测性能提升效果.

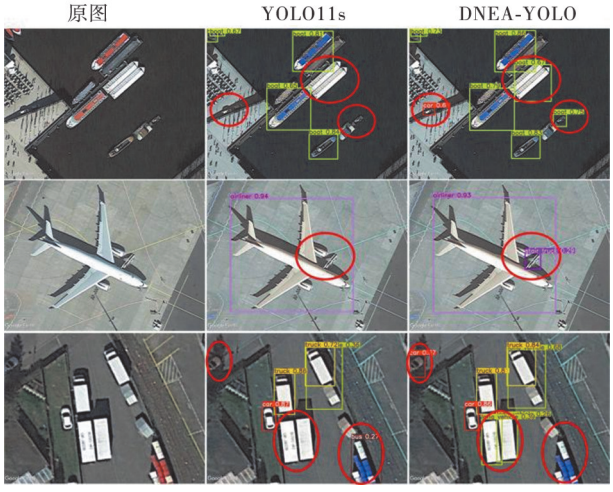


图 5 SIMD 数据集上的可视化结果

Fig.5 Visualization results on the SIMD dataset

图 5 展示了 DNEA-YOLO 与 YOLO11s 在 SIMD 数据集上对舰船、飞机、车辆 3 类典型遥感目标的检测效果对比，检测差异区域以红色圆圈标注.定量与定性分析表明，DNEA-YOLO 有效解决了 YOLO11s 在遥感目标检测中存在的漏检、错检问题.

舰船+车辆检测场景(第一行):总计 11 个目标,YOLO11s 因复杂海面背景干扰及舰船目标尺度差异,出现 5 处漏检,同时因背景纹理与车辆特征相似,发生 1 处复杂背景下车辆目标漏检;而 DNEA-YOLO 通过通道注意力筛选有效特征、空间注意力聚焦目标区域,成功检出 8 个目标,只有 3 处漏检.

飞机+车辆检测场景(第二行):总计 2 个目标,YOLO11s 对像素占比低的小目标特征提取能力不足,出现 1 处小目标漏检;DNEA-YOLO 的 C3K2-LSK 模块增强了小目标上下文信息捕获能力,精准检出该小目标,验证了模型对小尺度遥感目标的检测优势.

车辆目标检测场景(第三行):总计 7 个目标,YOLO11s 受阴影遮挡、目标密集分布影响,出现 2 处卡车目标漏检、1 处阴影区域车辆漏检,同时因背景噪声干扰发生 1 次错检(将非目标区域误判为车辆);DNEA-YOLO 引入的 SOFT-NMS 策略减少了密集目标漏检,且 DNEA 注意力抑制了背景噪声特征,无漏检、错检现象,检测精度显著提升.

4 总 结

本文针对目前遥感目标检测中存在的密集目标检测、复杂背景及环境干扰等问题,提出了 DNEA-YOLO 改进模型.通过提出 DNEA 注意力模块、设计 C3K2-LSK 集成部署结构、引入 SOFT-NMS 后处理策略,DNEA-YOLO 模型有效提升了复杂背景下多类别遥感目标的检测性能;基于 SIMD 数据集上的系列实验,验证了 DNEA-YOLO 模型在遥感图像目标检测任务中的有效性与优越性.

DNEA 模块通过复用归一化参数实现通道与空间双重特征增强,在不增加参数量与计算量的前提下提升模型性能,为轻量化遥感检测模型的设计提供了可行思路.未来可将该轻量化注意力机制与动态卷积、多尺度特征融合、轻量化架构搜索等方法结合,进一步优化模型精度与推理速度,拓展其在在轨实时解译、无人机巡检等边缘部署场景的应用范围.

参考文献:

[1] Ahmed I, Ahmad M, Chehri A, et al. IoT enabled deep learning based framework for multiple object detection in remote sensing images[J]. Remote Sensing, 2022, 14(16): 4107. DOI: 10.3390/rs14164107.

[2] Xie F, Luo H, Li S Q, et al. Using clean energy satellites to interpret imagery: a satellite IoT oriented lightweight object detection framework for SAR ship detection[J]. Sustainability, 2022, 14(15): 9277. DOI: 10.3390/su14159277.

[3] Lu A Q, Cheng Y, Hu Y B, et al. Satellite-terrestrial collaborative object detection via task-inspired framework[J]. IEEE Internet of Things Journal, 2023, 10(23): 20528-20544. DOI: 10.1109/jiot.2023.3287973.

[4] Deng Z P, Sun H, Zhou S L, et al. Multi-scale object detection in remote sensing imagery with convolutional neural networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 145: 3-22. DOI: 10.1016/j.isprsjprs.2018.04.003.

[5] 涂兵,张晓飞,张国云,等.递归滤波与 KNN 的高光谱遥感图像分类方法[J].国土资源遥感,2019,31(1):22-32. DOI:10.6046/gtzyyg.2019.01.04.
Tu Bing, Zhang Xiaofei, Zhang Guoyun, et al. Hyperspectral image classification via recursive filtering and KNN[J]. Remote Sensing for Land & Resources, 2019, 31(1): 22-32.

[6] Cheng G, Si Y J, Hong H L, et al. Cross-scale feature fusion for object detection in optical remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(3): 431-435. DOI: 10.1109/LGRS.2020.2975541.

[7] Liu Y F, Li Q, Yuan Y, et al. ABNet: adaptive balanced network for multiscale object detection in remote sensing imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5614914. DOI: 10.1109/TGRS.2021.3133956.

[8] 周莲,何楚,汪鼎文,等.光学遥感图像的小样本目标检测[J].遥感学报,2024,28(7):1693-1701. DOI:10.11834/jrs.20243209.
Zhou Lian, He Chu, Wang Dingwen, et al. Few-shot object detection in optical remote sensing images[J]. National Remote Sensing Bulletin, 2024, 28(7): 1693-1701.

[9] 刘浩然,严天笑,朱月琴,等.基于一种改进 YOLOv7 算法模型的滑坡识别研究:以四川省白格地区为例[J].自然资源遥感,2025,37(4):48-57. DOI:10.6046/zrzyyg.2024110.
Liu Haoran, Yan Tianxiao, Zhu Yueqin, et al. Landslide identification based on an improved YOLOv7 model: a case study of the Baige area[J]. Remote Sensing for Natural Resources, 2025, 37(4): 48-57.

[10] 李愿,付辉,刘浩志.双重注意力下的多尺度残差遥感图像去雾网络[J].自然资源遥感,2025,37(4):31-39. DOI:10.6046/zrzyyg.2024154.
Li Yuan, Fu Hui, Liu Haozhi. Multi-scale residual dehazing network for remote sensing images based on dual attention

[J].Remote Sensing for Natural Resources,2025,37(4):31-39.

[11] Hu J,Shen L,Albanie S,et al.Squeeze-and-excitation networks[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(8):2011-2023.DOI:10.1109/TPAMI.2019.2913372.

[12] Wang Q L,Wu B G,Zhu P F,et al.ECA-Net:efficient channel attention for deep convolutional neural networks[C]//IEEE.2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR).Seattle:IEEE,2020:11531-11539.DOI:10.1109/cvpr42600.2020.01155.

[13] Yang L X,Zhang R Y,Li L D,et al.SimAM:A simple,parameter-free attention module for convolutional neural networks[C]//Marina M,Tong Z.Proceedings of the 38th International Conference on Machine Learning.New York:PMLR,2021:11863-11874.DOI:10.1109/ICASSP43922.2022.9746294.

[14] Liu Y,Shao Z,Hoffmann N.Global attention mechanism:retain information to enhance channel-spatial interactions[PP/OL].arXiv(2021-12-10)[2025-11-06].https://arxiv.org/abs/2112.05561.

[15] Liu Y C,Shao Z R,Teng Y Y,et al.NAM:normalization-based attention module[PP/OL].arXiv(2021-11-24)[2025-06-07].https://doi.org/10.48550/arXiv.2111.12419.

[16] Xia Z F,Pan X R,Song S J,et al.Vision Transformer with deformable attention[C]//IEEE.2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR).New Orleans:IEEE,2022:4784-4793.DOI:10.1109/cvpr52688.2022.00475.

[17] Shao H,Zeng Q S,Hou Q B,et al.MCANet:medical image segmentation with multi-scale cross-axis attention[J].Machine Intelligence Research,2025,22(3):437-451.DOI:10.1007/s11633-025-1552-6.

[18] Xu S B,Zheng S C,Xu W H,et al.HCF-Net:hierarchical context fusion network for infrared small object detection[C]//IEEE.2024 IEEE International Conference on Multimedia and Expo(ICME).Niagara Falls:IEEE,2024:1-6.DOI:10.1109/icme57554.2024.10687431.

[19] Balakrishna S,Yadao S,Solanki V K.Enhancing YOLOv11 for real-time object detection:advanced architectures and edge-optimized training pipeline[C]//Vijender K S,Tran D T,Pradeep K,et al.Proceedings of the Ninth International Conference on Research in Intelligent Computing in Engineering.Warszawa:Polish Information Processing Society,2025:89-96.DOI:10.15439/2024R115.

[20] Li Y X,Hou Q B,Zheng Z H,et al.Large selective kernel network for remote sensing object detection[C]//IEEE.2023 IEEE/CVF International Conference on Computer Vision(ICCV).Paris:IEEE,2023:16748-16759.DOI:10.1109/ICCV51070.2023.01540.

[21] Bodla N,Singh B,Chellappa R,et al.Soft-NMS:improving object detection with one line of code[C]//IEEE.2017 IEEE International Conference on Computer Vision(ICCV).Venice:IEEE,2017:5562-5570.DOI:10.1109/iccv.2017.593.

[22] Haroon M,Shahzad M,Fraz M M.Multisized object detection using spaceborne optical imagery[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2020,13:3032-3046.DOI:10.1109/JSTARS.2020.3000317.

[23] Misra D,Nalamada T,Arasanipalai A U,et al.Rotate to attend:convolutional triplet attention module[C]//IEEE.2021 IEEE Winter Conference on Applications of Computer Vision(WACV).Waikoloa:IEEE,2021:3138-3147.DOI:10.1109/wacv486 30.2021.00318.

[24] Zhang J N,Li X T,Li J,et al.Rethinking mobile block for efficient attention-based models[C]//IEEE.2023 IEEE/CVF International Conference on Computer Vision(ICCV).Paris:IEEE,2023:1389-1400.DOI:10.1109/iccv51070.2023.00134.

[25] Dai W,Liu R,Wu Z X,et al.Exploiting scale-variant attention for segmenting small medical objects[J].IEEE Transactions on Neural Networks and Learning Systems,2026,37(7):3161-3175.DOI:10.1109/tnnls.2025.3645355.

[26] Wan D H,Lu R S,Shen S Y,et al.Mixed local channel attention for object detection[J].Engineering Applications of Artificial Intelligence,2023,123:106442.DOI:10.1016/j.engappai.2023.106442.

[27] Zhang H,Ma Z,Li X.RS-DETR:an improved remote sensing object detection model based on RT-DETR[J].Applied Sciences,2024,14(22):10331.DOI:10.3390/app142210331.

[28] Liu Y L,Yang D G,Song T T,et al.YOLO-SSP:an object detection model based on pyramid spatial attention and improved downsampling strategy for remote sensing images[J].The Visual Computer,2025,41(3):1467-1484.DOI:10.1007/s00371-024-03434-y.

(责任编辑　赵静慧)