

基于层次化特征融合的输电线路外破检测^①

朱跃飞¹, 程 旭², 刘庆程²

¹(南京信息工程大学 软件学院, 南京 210044)

²(南京信息工程大学 数字取证教育部工程研究中心, 南京 210044)

通信作者: 刘庆程, E-mail: qingcheng_liu@nuist.edu.cn



摘 要: 输电线路通道环境复杂, 各类外破隐患目标在拍摄角度、观测距离等因素影响下尺度差异显著, 进而导致模型在多样化风险目标识别中精度较低, 错检与漏检问题突出. 为解决上述问题, 提出一种基于层次化特征融合的输电线路外破检测方法. 该方法以 RT-DETR 模型为基础, 通过引入轻量化的 C2f_MambaOut 模块, 优化主干网络结构并有效降低模型参数量; 构建融合极性感知注意力与门控机制的 PA_CGLU 模块, 替代原有 AIFI 模块, 以提升查询向量对图像特征的方向感知与显著性建模能力, 增强自适应语义匹配效率; 并设计层次注意力融合块 HAFB, 利用局部与全局注意力分支实现输入特征的多尺度层次化融合与增强, 提升对多类别、多尺度目标的综合识别能力. 同时, 构建了一个涵盖多种真实场景的外破类型且样本分布均衡的输电线路外破检测数据集. 基于该数据集的实验结果表明, 改进后的模型平均精度值 (mAP) 提升了 1.5%, 参数量降低了 20.7%. 实验结果验证所提方法能有效缓解目标尺度差异带来的识别挑战, 增强对多样化外破隐患的综合检测能力, 在模型效率与精度之间取得更优平衡.

关键词: 输电线路巡检; 外破隐患检测; 层次化特征融合; RT-DETR; 轻量化

引用格式: 朱跃飞,程旭,刘庆程.基于层次化特征融合的输电线路外破检测.计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/10096.html>

Transmission Line External Damage Detection Based on Hierarchical Feature Fusion

ZHU Yue-Fei¹, CHENG Xu², LIU Qing-Cheng²

¹(School of Software, Nanjing University of Information Science & Technology, Nanjing 210044, China)

²(Engineering Research Center of Digital Forensics Ministry of Education, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: The transmission line channels feature complex environments, and various hidden external damage targets exhibit significant scale differences due to factors such as shooting angles and observation distances, thereby resulting in the model's low precision and prominent problems of false and missed detection in diversified risk target recognition. To this end, this study proposes a transmission line external damage detection method based on hierarchical feature fusion. The method is based on the RT-DETR model and introduces a lightweight C2f_MambaOut module to optimize the backbone structure and effectively reduce model parameters. Additionally, a PA_CGLU module integrating polarity-aware attention and gating mechanisms is established to replace the original AIFI module, thereby enhancing the query vectors' directional perception of image features and salient modeling capabilities, as well as improving adaptive semantic matching efficiency. Furthermore, a hierarchical attention fusion block (HAFB) is designed to realize multi-scale hierarchical fusion and enhancement of input features by employing local and global attention branches, thus boosting the comprehensive recognition ability of multi-category and multi-scale targets. Additionally, a transmission line external damage detection dataset that covers various real-world scenarios and features a balanced distribution of samples is constructed. Experimental results on this dataset demonstrate that the improved model achieves a 1.5% increase in mean

^① 基金项目: 江苏省研究生科研实践创新计划 (SJCX25_0518); 国家电网公司总部科技项目 (5700-202318309A-1-1-ZN)

收稿时间: 2025-08-18; 修改时间: 2025-09-09, 2025-09-29; 采用时间: 2025-10-14; csa 在线出版时间: 2026-01-15

average precision (mAP) and a 20.7% reduction in the parameter count. The results demonstrate that the effectiveness of the proposed method in mitigating the challenges posed by target scale variation and enhancing the overall detection performance for diverse external damage risks, thereby achieving a better balance between model efficiency and accuracy.

Key words: transmission line inspection; external damage detection; hierarchical feature fusion; RT-DETR; lightweight

电力系统在国家经济发展和社会稳定中具有举足轻重的地位,其中输电线路作为连接发电设施与电力用户之间的重要纽带,其安全稳定运行至关重要^[1].早期输电线路的巡检主要依赖人工巡查和直升机巡检,存在人力物力成本高、作业风险大、巡检效率低等问题.近年来,随着深度学习技术的发展以及无人机与输电监控技术的不断成熟,输电线路巡检正朝着智能化方向稳步推进^[2].

在智能化巡检的发展过程中,研究人员围绕输电线路的多种外破隐患开展了针对性研究.魏贤哲等^[3]针对传统目标检测方法在输电线路外破预警中存在误报率高、漏检率大以及掩模标注成本高的问题,提出将改进的 Mask-RCNN^[4]应用于外破目标检测任务.董卓元等^[5]提出了一种融合残差学习结构的 YOLOv4 输电线路外破检测方法,以同时满足模型在边缘设备上的部署实时性与检测精度需求.高莉莎等^[6]针对模型检测精度不足、推理速度慢等问题,提出了一种基于全局上下文增强的稀疏卷积电网防外力破坏轻量化检测算法 Fast-YOLOv5.赵文清等^[7]针对输电线路智能巡检中对目标检测模型的准确性与检测速度要求高,以及输电线路背景复杂多变的问题,提出一种基于改进 YOLOv8 的输电线路外力破坏检测算法.刘敏等^[8]针对传统外破检测方法感知范围受限的问题,提出了一种基于 YOLO-v10^[9]的全局感知增强检测方法,通过设计全局感知下采样与上采样模块,提升模型对外破目标边界的感知能力.上述现有方法主要面向塔吊、吊车等典型机械类外破目标,且所用数据集存在样本分布不均、过度增强等问题,限制了模型的泛化能力;同时,在检测模型方面,RCNN 系列模型结构复杂、计算开销大,不利于部署;YOLO 系列模型虽检测效率高,但由于召回率偏低,误检与漏检问题较突出,限制了其在复杂输电线路外破场景中的检测精度和鲁棒性.

随着 Transformer^[10]架构在自然语言处理领域的广泛应用,其强大的全局建模能力引起了计算机视觉领域的广泛关注. Facebook 团队于 2020 年提出了基于

Transformer 的端到端目标检测算法 DETR^[11],该方法通过引入序列建模机制,有效简化了传统检测流程,取消了锚框生成与非极大值抑制 NMS 等复杂后处理步骤,在目标检测任务中展现出良好的检测性能.然而,由于 DETR 模型参数量大、训练收敛速度慢等问题,限制了其在实时检测场景中的应用.为提升模型的实用性,百度团队于 2023 年提出了实时检测模型 RT-DETR^[12],在保留 DETR 端到端检测优势的基础上,进一步精简了模型结构,去除了候选框筛选和 NMS 操作,从而显著提升了推理速度与部署效率,实现了检测精度与实时性能的有效平衡,具有较高的工程应用价值.值得关注的是,在输电通道外破检测这一领域,针对 RT-DETR 的应用研究尚处于初步探索阶段.

针对上述问题与研究,本文提出了一种基于 RT-DETR 模型的层次化特征融合方法,用于输电线路外力破坏隐患目标的检测.该方法的主要特色有 4 点.

(1) 设计轻量化的 C2f_MambaOut 模块,优化特征提取主干网络,在显著降低模型参数数量的同时,保持稳定的特征提取性能.

(2) 构建融合极性感知注意力与门控机制的 PA_CGLU 模块,提升查询向量对图像特征的方向感知与显著性建模能力,增强自适应语义匹配效率,尤其适用于复杂场景下的精细目标检测.

(3) 设计层次注意力融合块 HAFB,利用局部与全局注意力分支实现输入特征的多尺度层次化融合与增强,提升多类别、多尺度目标的综合识别能力及模型鲁棒性.

(4) 构建了一个输电通道监控走廊视角,涵盖山火、烟雾、塔吊、吊车、水泥泵车与挖掘机这 6 类目标,且各类样本分布较为均衡的外破隐患图像数据集.

1 方法

在输电线路巡检场景中,大量监控摄像头分布于多条输电通道走廊,并以定时拍摄的方式获取图像数据.这些图像受拍摄角度、观测距离及环境背景等多

因素影响,呈现出复杂多变的视觉特征.为了应对上述场景及问题,本文基于 RT-DETR 模型,提出一种面向输电线路走廊视角的层次化特征融合外破隐患检测方法,旨在提升多视角、多尺度条件下多类别隐患的识别精度与模型鲁棒性.

本文方法结构如图 1 所示.我们将卷积层 (Conv2d)、批归一化层 (batch normalization) 和 SiLU 激活函数组合为一个基本卷积单元,记为 ConvN. 首先,将主干网

络替换为轻量化特征提取网络 CSPNet^[13],并设计 C2f_MambaOut 模块替换原始 C2f 模块.目标图像经主干网络处理后,输出多尺度特征图 {S3, S4, S5}.在改进的高效混合编码器模块中,采用 PA_CGLU 模块替代 AIFI 模块,对深层特征 S5 进行序列建模与上下文交互,生成包含丰富全局语义信息的特征表示 F5;随后, F5 与浅层特征 S3、中层特征 S4 输入至跨尺度特征融合模块,实现多尺度特征交互与语义对齐.

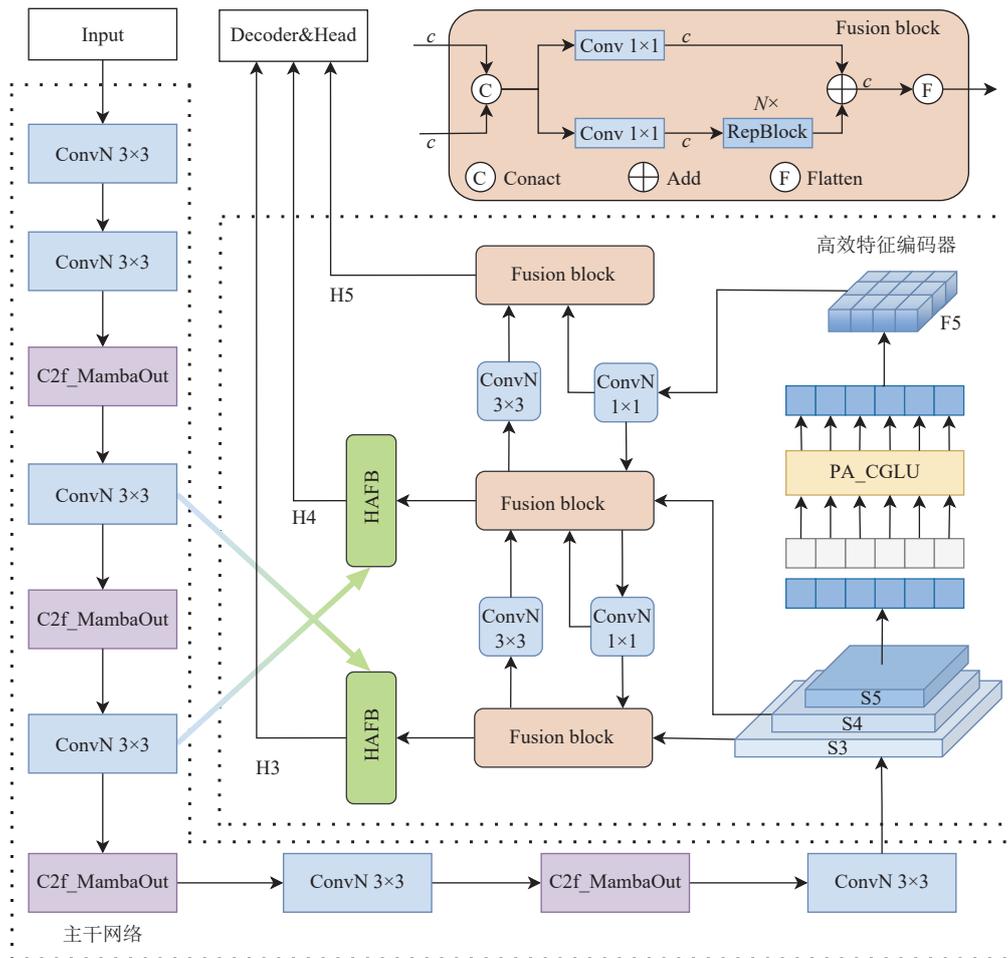


图 1 改进 RT-DETR 算法结构图

为进一步增强不同层级特征间的互补性,层次注意力融合块 HAFB 作为精融合核心,对主干输出的原始浅层和中层特征 (S3, S4) 与跨尺度融合后对应的特征进行二次融合,充分结合空间细节与语义信息.最终,生成优化后的特征映射 {H3, H4, H5}, 输入解码器与检测头以实现输电通道外破隐患目标的精准检测.

1.1 主干网络优化

RT-DETR 模型的参数量较大,在主干网络特征提

取模块存在较多冗余,为了便于模型部署及后续优化,本文引入 CSPNet 对主干网络部分进行轻量化改进. CSPNet 通过通道划分策略将输入特征图一分为二,其中一路直接进行跨层连接以促进梯度流动,另一路则通过残差结构进行深度特征学习,最后在输出端进行融合,有效平衡了模型性能与计算复杂度.然而,在输电线路监控走廊这一复杂背景场景中, CSPNet 中的 C2f 模块由于特征划分策略和卷积结构的限制,可能导致

局部细节特征的提取能力不足,影响对小目标或细粒度缺陷的检测性能。

如图2所示,本文在 CSPNet 网络结构基础上,借鉴 MambaOut^[14]模型中的 Gated CNN block,构建了新的 C2f_MambaOut 模块,对 C2f 模块中的 Bottleneck 残差结构进行替换,旨在提升特征提取的效率与表达能力。该模块将输入特征划分后并行送入多个 Gated CNN block 进行建模,再通过特征拼接与卷积融合,构建出兼具细粒度细节与全局语义的表达特征,从而有效增强模型在复杂输电场景中对外破隐患目标的感知能力和识别精度。

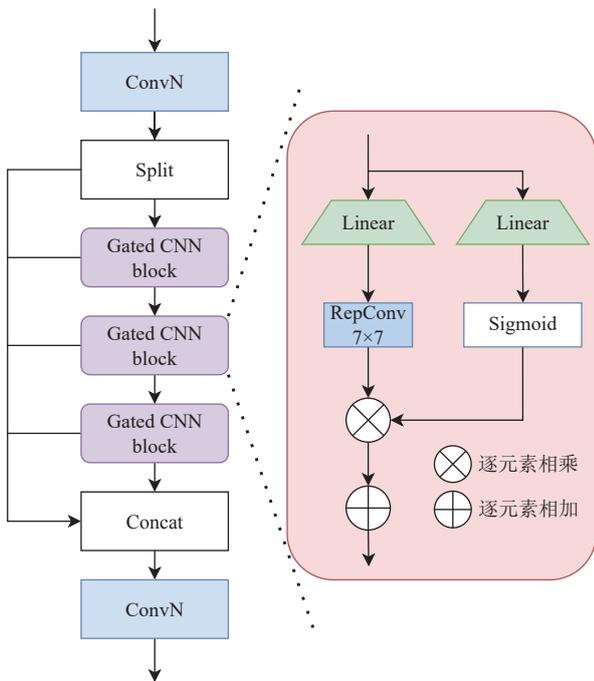


图2 C2f_MambaOut 结构图

Gated CNN block 结构如图2右侧所示。该设计移除了 Mamba^[15]模型提出的结构化状态空间模型 SSM (structured state space model), 使用门控机制与较大卷积核的局部建模策略以提升模型对复杂纹理与边缘细节的提取能力。具体而言,该模块首先对输入特征进行通道扩展,并划分为两个功能分支,分别用于生成门控权重和进行信息保留与深度卷积建模。卷积分支采用 7×7 深度卷积以增强局部上下文感知,同时兼顾计算效率;门控机制则通过调节两条分支特征的融合强度,引导模型聚焦于更具判别性的关键区域。最终,融合后的特征通过残差连接与原始输入相加,提升了特征表示的完整性与训练过程的稳定性。

1.2 融合极性感知注意力与门控机制

RT-DETR 模型中的注意力增强特征集成模块 AIFI 采用全局自注意力机制,有效建模特征间的长程依赖,增强了跨区域特征的交互能力。然而在输电通道场景中,外破隐患目标往往分布于自然环境与工业设施的交界区域,既包括山火、烟雾等自然类干扰,也涵盖塔吊、挖掘机等大型工程机械。整体背景呈现出强干扰、高速挡和结构复杂等典型特征。在此类复杂背景下,传统自注意力机制在局部几何结构建模方面存在局限,难以有效捕捉空间结构信息。此外,AIFI 所基于的标准自注意力结构对细粒度目标的动态筛选能力不足,难以抑制冗余或无关特征。

为增强局部结构感知能力,并提升模型对复杂背景下关键特征的动态响应效率,本文设计了极性感知门控卷积注意力模块 PA_CGLU,其核心结构如图3所示。该模块摒弃了 AIFI 中标准的 Transformer 编码器结构,融合了极性感知线性注意力与门控卷积单元,构建了一个异构的双路径特征处理框架:一方面,极性感知线性注意力 PA 增强了模型对空间极性方向与纹理流动趋势等结构性信息的建模能力,提升了在弱纹理或模糊区域中目标的判别能力;另一方面,门控卷积单元 CGLU^[16]可动态调节局部通道响应,有效筛选对任务敏感的细粒度特征,从而增强模型在多尺度目标下的语义表达能力。

极性感知线性注意力 PA^[17]结构如图3右侧所示,其设计核心在于解决传统线性注意力机制在建模负向分量时的能力不足问题。具体实现过程主要包括3项关键步骤:查询向量与键向量的极性分解、极性注意力权重计算以及极性感知混合策略。首先,如式(1)所示将查询向量 Q 和键向量 K 分解为正分量和负分量:

$$\begin{cases} Q = Q^+ - Q^- \\ K = K^+ - K^- \end{cases} \quad (1)$$

其中, $Q = \{Q_i\}_{i \in [d]} \in \mathbb{R}^d$, $K = \{K_i\}_{i \in [d]} \in \mathbb{R}^d$, 初始化采用 Xavier 初始化, $Q_i^+ = \max(Q_i, 0)$ 和 $Q_i^- = \max(-Q_i, 0)$ 分别代表 Q 的正分量和负分量, K 的分解方式相同;随后,如式(2)所示:

$$\begin{aligned} \langle Q, K \rangle = & \langle Q^+, K^+ \rangle + \langle Q^-, K^- \rangle \\ & - \langle Q^+, K^- \rangle - \langle Q^-, K^+ \rangle \end{aligned} \quad (2)$$

其中, $\langle \cdot, \cdot \rangle$ 表示向量内积计算,查询向量 Q 和键向量 K 定义与分解方式与式(1)相同。通过计算分解后的查

询和键向量之间的加权内积, 构建注意力权重, 该过程同时考虑了正分量相似度与负分量相似度, 分别用于建模符号一致分量与符号相反分量之间的响应关系; 为克服直接相减可能带来的信息损失, 引入了一个可学习的混合机制, 该机制通过两个可学习的系数矩阵 G^S 和 G^O 来分别处理相同符号和相反符号的响应. 随后, 将 G^S 和 G^O 分别应用于值向量 V 的两个子分量 V^S 和 V^O , 结果经拼接后作为最终的注意力表示.

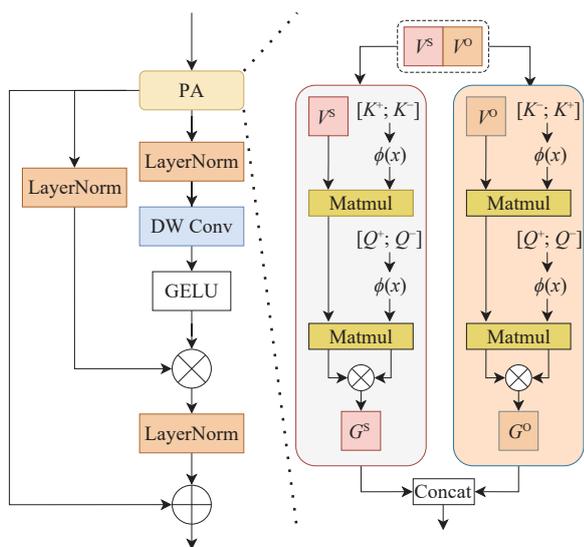


图3 PA_CGLU 结构图

1.3 层级注意力融合块

在深度学习, 尤其是计算机视觉任务中, 多尺度特征融合与注意力机制已被广泛应用, 并在目标检测、图像分割等复杂场景中展现出显著的性能优势. 随着 RT-DETR 在目标检测领域的逐步应用, 各领域的学者也进行了多尺度特征融合与注意力机制方面的改进研究. 张筱等^[18]基于原始 CCFM 模块设计了多尺度特征增强融合的金字塔结构, 通过增强特征层与细粒度信息融合, 有效提升了模型对不同尺度水下目标的定位与识别能力. 刘思元等^[19]引入双向特征金字塔网络 BiFPN 优化特征融合路径, 并结合 CARAFE 上采样算子, 提升了多尺度特征的融合效率与表达能力. 秦伦明等^[20]采用高层筛选特征金字塔网络 HS-FPN 替换原模型中的跨尺度上下文特征混合器 CCFM, 通过高层特征筛选并融合低层特征信息, 提升模型在极端天气下对低对比度和模糊目标的检测精度. 刘泽等^[21]通过引入自适应边界聚合思想, 构建了 BABF-FPN 多尺度特征融合模块, 利用双向机制自适应聚合不同分辨率的

特征表示, 增强了小目标模糊边界的细粒度表达能力.

然而, 在融合多源特征时, 现有方法普遍存在局部结构信息与全局语义之间协同不足的问题, 难以实现高效的信息耦合与统一表达, 从而限制了特征表征的鲁棒性与任务适配能力. 为解决上述问题, 本文设计了层次注意力融合块 HAFB, 通过引入结构分层的注意力机制, 自适应协调局部细节与全局语义的特征表达, 实现信息层级间的深度融合, 从而显著增强模型表征的判别性与泛化能力. HAFB 的核心改进在于融合局部与全局注意力分支, 并构建层次化特征路径以实现输入特征的精细融合. 这种设计既保留了输入的空间细节, 又借助自适应加权机制强化了全局语义的建模能力.

如图 4 所示, HAFB 首先对输入的两路特征分别通过 1×1 卷积降维, 统一通道数并降低计算开销; 随后, 采用 3×3 卷积对处理后的特征进行初步融合, 形成基线特征. 同时, 降维后的特征被送入两侧层次注意力分支. 分支通过局部不重叠分块和前馈注意力机制提取空间特征, 结合任务相关特征筛选与通道重构, 最终输出语义增强的特征; 最后, 将两路经层次注意力分支增强后的特征与基线特征进行拼接, 并依次通过 1×1 卷积、 3×3 可重参数化卷积以及另一个 1×1 卷积进行特征整合. 其中, 首个 1×1 卷积用于压缩通道以降低计算开销, 可重参数化卷积通过结构重参数化实现高效特征重组, 显著提升参数利用率和多分支融合能力, 最终的 1×1 卷积进一步精炼输出特征, 增强表达一致性.

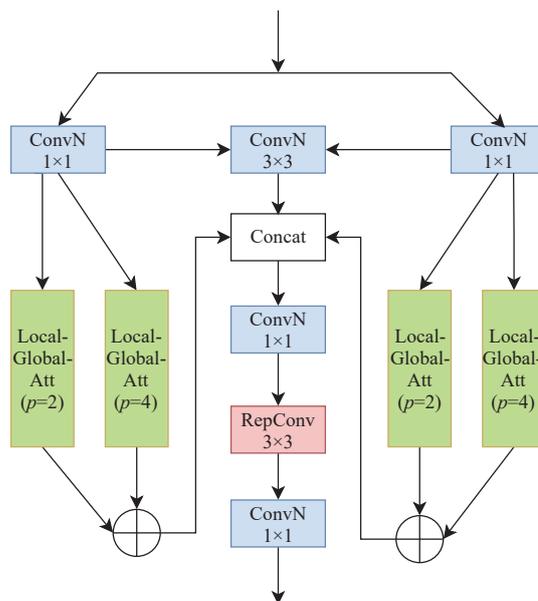


图4 HAFB 结构图

层次注意力分支 (LocalGlobalAtt) 结构如图 5 所示, 局部分支和全局分支的区别通过控制分块大小参数 p 实现, 这依赖于空间维度上不重叠分块的聚合与位移. 此外, 计算不重叠分块之间的注意力矩阵, 以实现局部和全局特征的提取与交互. 具体而言, 首先使用高效计算操作 Unfold 将输入特征向量 $F \in \mathbb{R}^{H' \times W' \times C}$ 划分为一组连续分块 $(p \times p, H'/p, W'/p, C)$. 然后进行通道平均得到 $(p \times p, H'/p, W'/p)$, 并通过 FFN 进行线性计算,

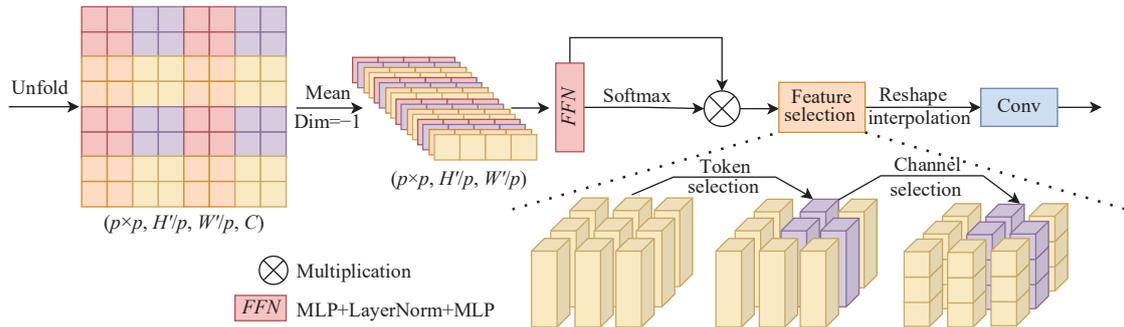


图 5 层次注意力分支结构图

接下来对线性计算结果应用 Softmax 激活函数以获得空间维度上的概率分布, 并相应调整其权重. 在加权结果中, 利用特征选择模块从 token 和通道中选择与任务相关的特征. 具体来说, 另 $d = H' \times W' / (p \times p)$, 加权结果表示为 $(t_i)_{i=1}^C$, 其中 $t_i \in \mathbb{R}^d$ 表示第 i 个输出 token. 特征选择操作每个 token, 输出如式 (4) 所示:

$$\hat{t}_i = P \cdot \sin(t_i, \xi) \cdot t_i \quad (4)$$

其中, $\xi \in \mathbb{R}^C$ 和 $P \in \mathbb{R}^{C \times C}$ 是任务相关参数为可学习参数, 初始化采用均匀分布. $\sin(\cdot, \cdot)$ 是一个值范围为 $[0, 1]$ 的余弦相似度函数. 然后对每个 token 通过 P 的线性变换进行通道选择, 随后进行 reshape 和插值操作, 最终生成特征 $F_{\text{local}} \in \mathbb{R}^{H' \times W' \times C'}$ 和 $F_{\text{global}} \in \mathbb{R}^{H' \times W' \times C'}$.

2 实验设计与结果分析

2.1 数据集构建

本文中使用的原始输电通道外破隐患图像数据来自某电力科学研究院, 数据采集自多地多条输电走廊多个监控摄像头, 主要涉及平原及丘陵地貌. 数据包含了不同时间段、气象条件和环境背景下的实际应用场景, 涵盖晴天、阴天等多种天气类型及晨昏、强光、逆光等不同光照条件. 场景背景类型丰富, 既包括靠近城镇、道路、工地等人类活动密集区域, 也包括农田、林地、河道等自然或半自然区域.

FFN 计算如式 (3) 所示:

$$FFN(X) = \max(0, XW_1 + b_1)W_2 + b_2 \quad (3)$$

其中, $W_1 \in \mathbb{R}^{C \times 4C}$ 为第 1 层的权重矩阵, 通常将维度拓展 4 倍; $b_1 \in \mathbb{R}^{4C}$ 为第 1 层的偏置项; $W_2 \in \mathbb{R}^{4C \times C}$ 为第 2 层的权重矩阵, 将维度降回原来的大小; $b_2 \in \mathbb{R}^C$ 为第 2 层的偏置项, 权重矩阵初始化采用 He 正态分布, 偏置项初始化为 0.

由于初始数据集中存在一定程度的冗余与质量差异, 本文在数据预处理阶段对原始图像进行了筛选与清洗处理. 具体而言, 对于同一摄像头在短时间内连续采集的图像, 本文去除了其中内容高度相似的样本, 以减少冗余性; 此外, 本文还剔除了部分画质低劣、目标轮廓模糊甚至难以通过人眼识别缺陷信息的图像. 经过严格筛选, 最终保留了 4520 张具有代表性且图像质量较高的原始样本图像, 涵盖典型场景与目标类型. 随后, 依据江苏省某地区《输电通道隐患标注文件》的相关要求, 采用 LabelImg 工具对原始图像进行了人工标注.

在完成细致的数据整理与标注工作后, 本文构建了一个面向输电线路外破隐患目标检测的高质量数据集, 共包含 4520 张图像、8409 个标注框, 涵盖山火、烟雾、塔吊、吊车、水泥泵车和挖掘机这 6 类典型目标. 图 6 展示了各类别目标的代表性样本, 可以看出, 输电通道监控场景背景复杂、光照条件多变, 外破隐患目标分布位置不固定, 且在形态、尺度上差异显著. 因此, 开展多类别典型外破隐患智能识别的相关研究, 对于提升输电通道运行的安全水平与巡检智能化能力具有重要意义.

在综合考虑图像数量与标签分布的基础上, 本文按照 8:1:1 的比例将数据集划分为训练集、验证集与测试集, 具体划分结果如表 1 所示.



图6 输电线路外破隐患数据集示例

表1 自建数据集划分表

类别	训练集 标签数	验证集 标签数	测试集 标签数	总标签数
山火	660	83	84	827
烟雾	1324	165	166	1655
塔吊	1534	191	193	1918
吊车	1188	148	149	1485
水泥泵车	733	91	91	915
挖掘机	1286	161	162	1609

2.2 实验环境和评估指标

本文实验基于 Python 3.8.0 版本的 Torch 1.10.0 深度学习框架的软件运行环境. 服务器系统为 CentOS7, 搭载 NVIDIA GeForce RTX3090 显卡. 实验超参数设置如下, 输入图像设置为 640×640, 训练轮数为 150 轮, 初始学习率为 0.0001, 批量大小为 4, 优化器为 AdamW, 训练所采用的损失函数为 GIoU 损失函数, 其余设置均与原 RT-DETR 算法相同.

本文采用 4 项轻量化指标和两项准确度指标进行评估. 轻量化指标包括计算量 GFLOPs、帧率 FPS、参数量 Params 和模型大小 model size; 准确度指标包括各类别平均精度 AP 和平均检测精度 mAP. 其中, Params 和 model size 分别表示模型结构中的参数数量和整体权重大小, 模型越小, 所需存储和计算资源越少, 训练与推理的效率也越高; FPS 表示模型在推理阶段的处理速度, 用于衡量其实时性能; GFLOPs 表示模型在一次前向推理中所需的浮点运算量, 用于衡量模型的计

算复杂度.

2.3 消融实验

本文基于 RT-DETR 模型进行了一系列的改进, 通过消融实验验证所提算法模块的有效性. 实验在自建的输电线路外破检测数据集上进行.

从表 2 中各类别平均精度的消融实验结果可以看出, 引入轻量化特征提取模块 C2f_MambaOut 后, 模型整体精度略有下降, 其中对水泥泵车类别的检测精度影响最为显著. 这表明该类别在特征表达上对网络结构压缩较为敏感, 可能由于其形态特征复杂、纹理变化不明显, 在轻量化条件下难以有效建模其判别特征. 在此基础上引入的 PA_CGLU 模块在塔吊、吊车等结构清晰、轮廓特征明显的机械类目标中实现了显著的性能提升, 进一步验证了该模块在增强多尺度结构性特征表达方面的有效性, 尤其适用于具有复杂构型的刚性目标识别任务. 此外, 融合了局部与全局注意力机制的 HAFB 模块在多数类别上均提升了检测精度, 表明其在特征聚合与跨尺度建模方面具有良好适应性. 然而, 在烟雾类别上, 其检测精度出现一定幅度的下降, 可能是由于烟雾目标本身边缘模糊、透明度高、与背景差异弱等因素, 导致模型难以聚焦于有效特征区域, 加之类内尺度和形态变化大, 对注意力机制的泛化能力构成了更高要求. 综上所述, 尽管各模块在不同类别上的性能表现存在差异, 但当本文方法将 C2f_MambaOut、

PA_CGLU 与 HAFB 模块进行融合后,模型的整体识别能力得到显著增强.在保持较低模型复杂度的前提下,实现了对多类别外破隐患目标的精准识别,尤其在复杂背景、尺度变化大和结构特征弱的场景下均表现出良好的鲁棒性.融合模型在各类别上均取得了较高的检测精度,充分验证了所提方法在输电线路外破隐患检测任务中的有效性与实用价值.

表 2 各类别平均精度 AP@0.5 消融实验

类别	RT-DETR	+C2f_MambaOut	+PA_CGLU	+HAFB	Ours
山火	0.715	0.753	0.746	0.777	0.767
烟雾	0.72	0.71	0.7	0.679	0.735
塔吊	0.838	0.853	0.874	0.85	0.844
吊车	0.849	0.855	0.872	0.852	0.855
水泥泵车	0.746	0.687	0.748	0.753	0.752
挖掘机	0.877	0.868	0.859	0.881	0.881
mAP@0.5	0.791	0.788	0.8	0.799	0.806

表 3 展示了各模块在不同指标下的消融实验结果.引入轻量化模块 C2f_MambaOut 后,尽管整体检测

表 3 不同改进策略的消融实验

RT-DETR	C2f_MambaOut	PA_CGLU	HAFB	mAP@0.5	GFLOPs	Params (MB)	Model size (MB)
√	—	—	—	0.791	57.0	19.9	77.0
√	√	—	—	0.788	46.6	13.8	53.7
√	—	√	—	0.800	57.2	20.0	77.6
√	—	—	√	0.799	58.4	21.6	84.0
√	√	√	√	0.806	50.1	15.8	61.6

图 7 展示了两者的热力图对比结果.

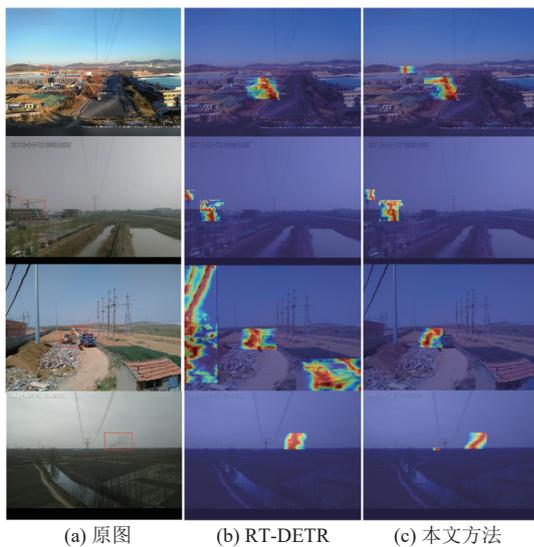


图 7 热力图对比

可以观察到,原始 RT-DETR 模型的关注区域分布较为分散,与目标实际轮廓的契合度较低,且存在一定

精度略有下降,但其参数量与模型大小分别减少了 30.6% 和 30.3%,计算量也下降 18.2%,充分验证了主干网络轻量化设计在压缩模型规模方面的有效性.PA_CGLU 模块在几乎不增加复杂度的情况下提升了检测精度,增强了模型的特征表达与语义建模能力.HAFB 模块由于集成了多层次的注意力分支机制,导致模型参数量与计算量有所增加.然而结合表 2 中的检测结果可见,该模块在多数目标类别上均带来了精度提升,在一定程度上弥补了轻量化带来的性能损失,体现了其在上下文建模与结构感知方面的优势.综合来看,本文方法融合上述 3 个模块,在检测性能与模型轻量化之间实现了较优的平衡.在参数量和模型大小分别减少 20.7% 和 20%、计算量减少了 12.1% 的同时,平均检测精度提升了 1.5 个百分点.该结果表明,所提方法不仅有效压缩了模型规模,还显著增强了其在多类别外破隐患目标检测中的综合表现,为实际场景中实现高效、精确的智能巡检提供了有力支撑.

程度的漏检与误检现象.相比之下,本文方法所提结构在关注区域上更加聚焦于目标本体,热力图显著贴合目标真实形状,显示出更强的判别能力.整体来看,本文方法在鲁棒性与检测精度方面均优于原始模型,展现出其在输电线路外破隐患检测任务中的更强适应性与实用性.

2.4 对比实验

2.4.1 不同注意力机制对比实验

为了验证本文所提出的 PA_CGLU 模块的有效性,本文选取 ELA^[22]、CA^[23]、SE^[24]等不同的注意力机制进行对比实验,实验结果如表 4 所示.

由表 4 可见,不同注意力机制在提升模型精度和控制计算复杂度方面表现存在差异.ELA 机制在 mAP@0.5 指标上取得了最高的 0.802,但其参数规模与计算量相对较大;CA 与 SE 机制虽然在参数量上更具优势,但在信息建模过程中主要依赖通道加权,空间上下文建模能力不足,因而在精度上存在一定损失.相

比之下,本文提出的 PA_CGLU 模块在保持较低计算开销的同时,通过极性感知线性注意力与门控卷积单元的协同建模,实现了全局结构信息与局部细粒度特征的兼顾表达.与 ELA 强调全局长程依赖、CA 与 SE 偏重通道加权不同,PA_CGLU 能够在捕捉空间极性方向与纹理取向信息的同时,动态调节通道响应并筛选关键信息,从而在轻量化条件下依然维持较高的检测精度.

表 4 注意力机制对比实验

注意力机制	mAP@0.5	GFLOPs	Params (MB)
ELA	0.802	58.9	24.9
CA	0.796	58.4	20.2
SE	0.793	57.1	19.9
PA_CGLU	0.800	57.2	20.0

2.4.2 不同多尺度融合方法对比实验

为了验证本文所提出的层次化多尺度特征融合模块的有效性,将特征融合部分分别替换为 BiFPN^[25]、Slimneck^[26]、HSFPN^[27]进行对比实验,实验结果如表 5 所示.

表 5 多尺度融合对比实验

特征融合网络	mAP@0.5	GFLOPs	Params (MB)
BiFPN	0.796	64.3	20.3
Slimneck	0.785	53.3	19.3
HSFPN	0.788	53.4	18.1
HAFB	0.799	58.4	21.6

由表 5 可见,不同多尺度特征融合方法在检测精度与复杂度之间表现存在明显差异. BiFPN 在跨层特征交互方面表现较好,但由于其采用多层可学习权重分配与多次迭代融合,计算复杂度显著提升,不利于轻量化部署. Slimneck 与 HSFPN 在计算量和参数规模上具有一定优势,但融合过程中对层次信息的建模能力不足,导致检测精度相对较低.相比之下,本文提出的 HAFB 模块在保持适中计算开销的同时,实现了较高的检测精度,优于其他 3 种对比方法.其原因在于, HAFB 通过层次化注意力引导的多尺度特征融合,能够在充分利用浅层细节特征与深层语义信息的同时,动态调整不同尺度特征在融合过程中的权重分配,既避免了 BiFPN 的冗余开销,又弥补了 Slimneck 与 HSFPN 在跨层交互建模方面的不足.因此, HAFB 在检测性能与计算效率之间实现了更优的平衡,展现出较强的实用性与可扩展性.

2.4.3 不同检测模型对比实验

为验证本文方法在输电线路外破检测任务中的性

能表现,选取多个主流目标检测模型进行对比实验,涵盖了 YOLOv5、YOLOv8、YOLOv10、Deformable DETR^[28]、DINO^[29]等代表性方法,以及文献[6]和文献[18]两种改进方法,实验均在相同的环境下使用相同的数据集,并在相同测试集上进行比较,对比结果如表 6 所示.表 6 的结果显示,本文方法在检测性能方面表现优异,mAP@0.5 和 mAP@0.5:0.95 分别达到 80.6% 和 45.2%,整体精度优于其余所有对比模型.相较于基线模型 RT-DETR,本文通过一系列有效改进,在参数量减少超过 20% 的前提下,使 mAP@0.5 和 mAP@0.5:0.95 分别提升了 1.5 个和 0.8 个百分点.与文献[18]的改进方法相比,本文方法的检测精度和计算复杂度指标要更好,说明本文改进更适合于输电线路外破检测.文献[6]的改进是对 YOLOv5s 版本模型进行较大幅度的轻量化改进,其检测速度得到了较大幅度的提升,但是其检测精度过低,在实际运用中难以达到很好的效果.

表 6 对比实验

模型	mAP@0.5	mAP@0.5:0.95	GFLOPs	FPS (f/s)	Params (MB)
YOLOv5m	0.726	0.402	64	71	25.0
YOLOv8m	0.745	0.415	78.7	70	25.8
YOLOv10m	0.686	0.377	58.9	71	15.3
Deformable DETR	0.735	0.421	78.4	29	40.0
DINO	0.782	0.440	166.8	24	47.1
文献[6]	0.612	0.308	7.8	126	3.6
文献[18]	0.788	0.436	66.0	32	14.8
RT-DETR	0.791	0.444	57.0	63	19.9
本文方法	0.806	0.452	50.1	43	15.8

此外,与 Deformable DETR 和 DINO 等方法相比,本文方法在参数量和计算量显著降低的基础上实现了更优的检测性能,进一步验证了所提结构在保持效率的同时具备良好的检测能力与部署价值.与 YOLO 系列方法相比,本方法检测精度远高于 YOLOv5、YOLOv8 等方法,尽管推理速度相对较低,但仍满足实时检测 FPS 指标大于 30 的基本需求,且本文方法的计算量和模型复杂度更低.综上所述,本文方法与主流算法相比,在保证输电线路外破隐患目标实时检测的同时,实现更优的检测精度和鲁棒性,具有良好的实用价值.

2.5 部署实验

为了验证模型在计算能力相对受限设备上的性能表现,本文将在 RTX3090 设备上训练好的模型分别移植到 RTX2080Ti 和 CPU 设备中进行部署实验对比测试.测试过程中使用相同的测试集,其中 CPU 设备型

号为 Intel(R) Core(TM) i7-14650HX.

实验结果如表 7 所示. 可以看出, 在不同设备上, 模型的检测精度基本保持不变, 具备良好的跨平台一致性, 而模型的推理速度在资源受限的设备上不可避免地出现了下降. 总体而言, 本文提出的模型在计算资源受限环境下仍能保持较好的性能平衡, 具备在实际嵌入式设备中部署的可行性, 在输电外破实际应用场景中基本可以满足实际检测需求.

表 7 部署实验

模型	设备	mAP@0.5	FPS (f/s)	模型内存占用 (MB)
RT-DETR	RTX3090	0.791	63	19.9
	RTX2080Ti	0.789	54	19.9
	CPU	0.790	11	19.9
本文方法	RTX3090	0.806	43	15.8
	RTX2080Ti	0.805	35	15.8
	CPU	0.804	9	15.8

此外, 模型在轻量化部署方面也具备进一步优化的潜力. 通过量化可将模型参数从高精度浮点数压缩为低比特表示, 减小存储占用; 通过剪枝可去除冗余通道和层结构, 降低计算开销. 在保证检测精度基本不变的前提下, 这些策略能够显著提升模型在低算力设备上的推理速度与能效, 为实际嵌入式部署提供更多可行性和灵活性.

2.6 检测结果可视化分析

为了更直观地展示本文方法的有效性, 本文对不同场景下的目标检测效果进行了可视化分析, 结果如图 8 所示.

图 8 展示了 4 组典型场景中输电线路外破隐患目标的检测结果. 图 8(a) 为基线模型 RT-DETR 的检测结果, 图 8(b) 为本文方法的检测结果, 按从上到下顺序依次排列. 第 1 组场景为城镇的河边, 图中包含塔吊和水泥泵车目标, 基线模型将道路转弯区域误检为吊车目标. 第 2 组场景为城郊工厂, 基线模型将吊车错检为挖掘机, 而本文方法能够正确识别目标, 且置信度更高、边界框定位更精确. 第 3 组场景为暗光条件林地, 本文方法在山火与烟雾目标的检测中表现出更高的准确性和置信度, 展现出良好的鲁棒性. 第 4 组场景为农田, 基线模型漏检了一处挖掘机目标. 综合以上分析, 本文方法在多场景下表现出更高的检测精度与鲁棒性, 能够有效减少误检与漏检现象, 具备更强的实用性与泛化能力.



图 8 基线模型与本文方法在不同场景下检测结果

3 结论与展望

本文围绕输电线路外破隐患目标检测中的关键点展开研究, 聚焦于复杂环境下多尺度、多类别目标的识别精度与模型效率之间的矛盾问题. 在深入分析现有检测模型在外破隐患识别中存在如精度不足、鲁棒性较差等问题的基础上, 本文提出了一种基于层次化特征融合的输电线路外破检测方法. 该方法在保持推理效率的同时, 显著提升了对目标尺度差异与结构复杂性的适应能力. 通过模块化设计与整体结构优化, 模型不仅在检测精度方面实现了提升, 同时在参数压缩与特征表达能力方面也展现出良好的综合性能.

此外, 本文还构建了一个具备实际代表性和类间平衡性的数据集, 为后续相关研究提供了可靠的实验基础. 实验结果进一步验证了所提方法在检测精度与模型效率之间实现了良好平衡, 在典型输电场景下具备较强的工程应用潜力. 后续将结合输电线路巡检的实际需求, 进一步提升模型的综合检测性能, 探索更多类别、更复杂场景及恶劣天气下的检测方法.

参考文献

- 胡毅, 刘凯, 吴田, 等. 输电线路运行安全影响因素分析及

- 防治措施. 高电压技术, 2014, 40(11): 3491–3499.
- 2 伍钰婷, 宋先琴, 伍继红, 等. 基于深度学习的输电线路检测综述. 2024年电力行业技术监督专业技术交流研讨会优秀论文集. 南宁: 中国电力技术市场协会, 2024. 751–756.
 - 3 魏贤哲, 卢武, 赵文彬, 等. 基于改进 Mask R-CNN 的输电线路防外破目标检测方法研究. 电力系统保护与控制, 2021, 49(23): 155–162.
 - 4 He KM, Gkioxari G, Dollár P, *et al.* Mask R-CNN. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2980–2988.
 - 5 董卓元, 高永亮, 袁斌, 等. 一种改进型 YOLOv4 输电线路防外破检测方法. 电网与清洁能源, 2023, 39(6): 17–25.
 - 6 高莉莎, 郭乐乐, 韩硕, 等. 全局上下文增强的稀疏卷积电网防外力破坏检测. 计算机系统应用, 2024, 33(6): 81–90. [doi: 10.15888/j.cnki.csa.009514]
 - 7 赵文清, 蔡建颖, 马玉梅, 等. 基于改进 YOLOv8 的输电线路外力破坏检测算法. 山东电力技术, 2025, 52(5): 1–8.
 - 8 刘敏, 陈明, 武明虎, 等. GAE-YOLO: 全局感知增强的输电线路外破隐患目标检测方法. 仪器仪表学报, 2025, 46(2): 267–278.
 - 9 Wang A, Chen H, Liu LH, *et al.* YOLOv10: Real-time end-to-end object detection. Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2025. 3429.
 - 10 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
 - 11 Carion N, Massa F, Synnaeve G, *et al.* End-to-end object detection with Transformers. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 213–229.
 - 12 Zhao YA, Lv WY, Xu SL, *et al.* DETRs beat YOLOs on real-time object detection. Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024. 16965–16974.
 - 13 Zhang DH, Hao XY, Liang LL, *et al.* A novel deep convolutional neural network algorithm for surface defect detection. Journal of Computational Design and Engineering, 2022, 9(5): 1616–1632. [doi: 10.1093/jcde/qwac071]
 - 14 Yu WH, Wang XC. MambaOut: Do we really need Mamba for vision? Proceedings of the 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2025. 4484–4496.
 - 15 Gu A, Dao T. Mamba: Linear-time sequence modeling with selective state spaces. arXiv:2312.00752, 2023.
 - 16 Shi D. TransNeXt: Robust foveal visual perception for vision Transformers. Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024. 17773–17783.
 - 17 Meng WK, Luo YD, Li X, *et al.* PolaFormer: Polarity-aware linear attention for vision Transformers. arXiv:2501.15061, 2025.
 - 18 张筱, 刘孙俊, 陈懿, 等. 多尺度特征增强融合的 ME-RTDETR 水下目标检测. 微电子学与计算机, <https://link.cnki.net/urlid/61.1123.TN.20250508.1234.002>. (2025-05-08).
 - 19 刘思元, 高凯, 雍龙泉. 改进 RT-DETR 的航拍小目标检测算法. 计算机工程与应用, 2025, 61(4): 272–281.
 - 20 秦伦明, 张云起, 崔昊杨, 等. 基于改进 RT-DETR 的极端天气下交通标志检测方法. 电子测量技术, 2025, 48(9): 56–64.
 - 21 刘泽, 宋廷伦, 石先让, 等. 基于改进 RT-DETR 的路面异常检测技术研究. 计算机工程, <https://doi.org/10.19678/j.issn.1000-3428.00EC0070182>. (2024-11-21).
 - 22 Xu W, Wan Y, Zhao WN. ELA: Efficient location attention for deep convolution neural networks. Journal of Real-time Image Processing, 2025, 22(4): 140. [doi: 10.1007/s11554-025-01719-6]
 - 23 Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13708–13717.
 - 24 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
 - 25 Tan MX, Pang RM, Le QV. EfficientDet: Scalable and efficient object detection. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 10778–10787.
 - 26 Oguntola I, Olubeko S, Sweeney C. SlimNets: An exploration of deep model compression and acceleration. Proceedings of the 2018 IEEE High Performance Extreme Computing Conference (HPEC). Waltham: IEEE, 2018. 1–6.
 - 27 Chen YF, Zhang CY, Chen B, *et al.* Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases. Computers in Biology and Medicine, 2024, 170: 107917. [doi: 10.1016/j.combiomed.2024.107917]
 - 28 Zhu XZ, Su WJ, Lu LW, *et al.* Deformable DETR: Deformable Transformers for end-to-end object detection. arXiv:2010.04159, 2021.
 - 29 Zhang H, Li F, Liu SL, *et al.* DINO: DETR with improved denoising anchor boxes for end-to-end object detection. arXiv:2203.03605, 2022.

(校对责编: 张重毅)