

基于轻量化边缘增强的电力作业穿戴检测

柏帆¹, 郭鹏程², 佟鑫², 王荣历², 刘佳杰¹, 金彬¹

(1. 宁海县雁苍山电力建设有限公司, 浙江 宁波 315600;

2. 国网浙江省电力有限公司宁海县供电公司, 浙江 宁波 315600)

摘要: 为解决电力作业环境中背景复杂和目标尺寸多样导致的穿戴检测难题, 提出一种轻量化的边缘感知检测算法; 采用多尺度边缘信息增强方法在浅层高效融合边缘特征以抑制背景噪声; 设计 ShiftRepC3 模块于混合编码器中高效提取并融合局部特征, 并利用 DySample 改进上采样增强重建能力; 引入基于注意力机制的剪枝方法降低模型参数量与计算量, 满足边缘设备部署需求; 实验结果表明, 和基线模型 RT-DETR 相比, 所提方法在两个数据集上的 $mAP50$ 分别提升了 2.8% 和 2.1%, 在 $mAP50-95$ 分别提升了 2.6% 和 1.5%, 并且保持了较低计算开销。

关键词: 边缘特征增强; RT-DETR; 电力作业安全; 注意力机制; 轻量化

Edge-enhanced Power Operation Wearable Detection Based on Lightweight

BAI Fan¹, GUO Pengcheng², TONG Xin², WANG Rongli², LIU Jiajie¹, JIN Bin¹

(1. Ninghai Yancangshan Electric Power Construction Co., Ltd., Ningbo 315600, China;

2. Ninghai County Power Supply Company, State Grid Zhejiang Electric Power Co., Ltd., Ningbo 315600, China)

Abstract: To solve the wearable detection of complex backgrounds and diverse target sizes in power operation environments, a lightweight edge-aware detection algorithm is proposed. A multi-scale edge information enhancement method is employed to efficiently fuse edge features at shallow layers, suppressing background noise. A ShiftRepC3 module is designed within the hybrid encoder to effectively extract and fuse local features, and a DySample method is utilized to improve upsampling for the capability of enhanced reconstruction. An attention mechanism-based pruning method is introduced to reduce the parameters and computational costs of the model, meeting the deployment requirement of edge devices. Experimental results demonstrate that, compared with the baseline real-time detection transformer (RT-DETR) model on two datasets, the proposed method improves the $mAP50$ by 2.8% and 2.1%, and the $mAP50-95$ by 2.6% and 1.5%, respectively, and maintains a low computational cost.

Keywords: edge feature enhancement; RT-DETR; electrical operation safety; attention mechanism; lightweight

0 引言

电力作业环境复杂且危险性高, 作业人员的安全防护显得至关重要。在电力作业中, 违规穿戴行为是导致事故发生的主要因素之一^[1]。然而, 传统的人工巡查方法效率低, 且容易受到人为因素的干扰, 难以实现实时、全面的安全监控^[2]。

目前基于深度学习的电力现场作业人员违规穿戴检测方法可以分为两类: 两阶段检测方法和单阶段检测方法

法^[3]。经典的两阶段检测方法代表是快速区域卷积神经网络^[4]和加速区域卷积神经网络^[5]。这些方法先生成候选框, 再对候选框内的物体进行识别。单阶段检测方法的代表是单次多框检测器^[6]和 YOLO^[7]系列算法, 它们无需生成候选框, 能够直接输出预测结果。随着 Transformer 在计算机视觉领域的广泛应用, 其强大的全局建模能力和对长程依赖关系的捕捉能力逐渐成为研究热点。文献 [8] 提出了一种基于 Transformer 的全局目标检测模型检 (DETR, detection transformer), 简化了现有

收稿日期:2025-05-30; 修回日期:2025-06-19。

基金项目:宁波永耀电力投资集团科技项目(CF058211002024001)。

作者简介:柏帆(1982-),男,大学本科,高级工程师。

通讯作者:郭鹏程(1983-),男,大学本科,工程师。

引用格式:柏帆,郭鹏程,佟鑫,等. 基于轻量化边缘增强的电力作业穿戴检测[J]. 计算机测量与控制, 2025, 33(8):112-119, 128.

目标检测方法中复杂的后处理步骤, 为目标检测的发展开辟了新的方向。然而, DETR 模型存在训练收敛速度慢和对小目标检测能力不足的问题。因此, 文献 [9] 提出了一种实时端到端目标检测模型 (RT-DETR, real-time detection transformer), 通过设计高效的混合编码器来加速多尺度特征处理, 并采用最小不确定性查询选择策略, 提高了模型的检测速度和精度。轻量化方面, 针对电力作业场景中设备种类复杂、检测任务对效率和部署资源要求较高的特点, 近年来轻量化检测算法成为研究热点。例如, 文献 [10] 提出的轻量级 YOLOv5 在保持较高检测精度的同时显著减少了模型参数量, 适用于边缘设备部署; 文献 [11] 设计了进 YOLOv4 模型, 在电力设备识别任务中取得了良好的实时性与泛化能力; 文献 [12] 进一步结合注意力机制与 MobileNetV3, 提升了穿戴规范检测的准确性和鲁棒性。

电力作业环境中的背景复杂和目标尺寸多样等因素, 严重影响了模型对作业人员安全装备的识别与检测精度。同时, 现场部署的边缘设备通常面临计算资源受限和存储容量有限等问题。针对电力作业场景中背景复杂且干扰因素多的问题, 文献 [13] 通过嵌入非对称卷积模块和 Transformer 模块, 增强了网络对不规则物体特征和上下文信息的建模能力。文献 [14] 提出一种安全穿戴检测模型, 通过结合多种数据增强策略, 提升了模型在复杂环境下的鲁棒性。文献 [15] 通过引入高效金字塔分离注意力网络 (EPSANet, efficient pyramid split attention network) 注意力模块和改进的空间金字塔池化结构, 提高了对安全带的检测精度。文献 [16] 提出一种少样本学习的高空作业安全带检测方法, 通过两阶段检测流程, 先检测出高空作业人员, 再利用少样本度量学习方法识别是否佩戴安全带。针对电力作业场景中的

边缘设备资源受限的问题, 文献 [17] 通过引入深度可分离卷积网络和优化的空间金字塔池化结构, 有效降低了模型参数量, 并采用 Soft-NMS 算法减少漏检率。刘昶成等人 [18] 通过集成轻量化卷积模块和自适应空间特征融合网络, 降低了计算复杂度并提高了检测精度。

现有研究尽管在提高检测精度和模型轻量化方面取得了一定进展, 但仍然存在一些需要解决的问题。首先, 背景噪声干扰问题依然存在, 尤其是在复杂环境中, 目标与背景的对比度较低, 导致检测效果不稳定。此外, 现有方法虽然能有效提高检测精度, 但由于参数量和计算开销过大, 无法满足实际电力作业场景中对实时性和计算效率的要求。针对以上挑战, 本文基于 RT-DETR [9] 提出了一种轻量化的边缘特征增强的检测算法 (LERT-DETR, lightweight edge-enhanced RT-DETR), 以提升检测精度和效率。

1 LERT-DETR 算法

1.1 算法介绍

LERT-DETR 由优化后的主干网络、混合编码器以及解码器 3 个部分组成, 网络结构如图 1 所示。主干网络采用 ResNet18 [16] 提取特征, 并通过多尺度边缘信息提取器从浅层中提取边缘信息, 随后, 将最后 3 个阶段经过 EFF 融合边缘信息后输出的特征作为混合编码器的输入。混合编码器对最高层级的特征进行编码, 同时实现多尺度的特征融合。通过设计 ShiftRepC3 高效提取局部特征信息, 移位操作 [17] 提升特征表示的丰富性。通过 DySample [18] 对上采样进行优化实现更精细的特征重建, 从而有效提升模型对多尺度特征的捕获能力。在解码器部分, 首先通过 IOU-感知查询模块对初始查询进行优化, 利用候选框的 IoU 信息动态调整查询嵌入,

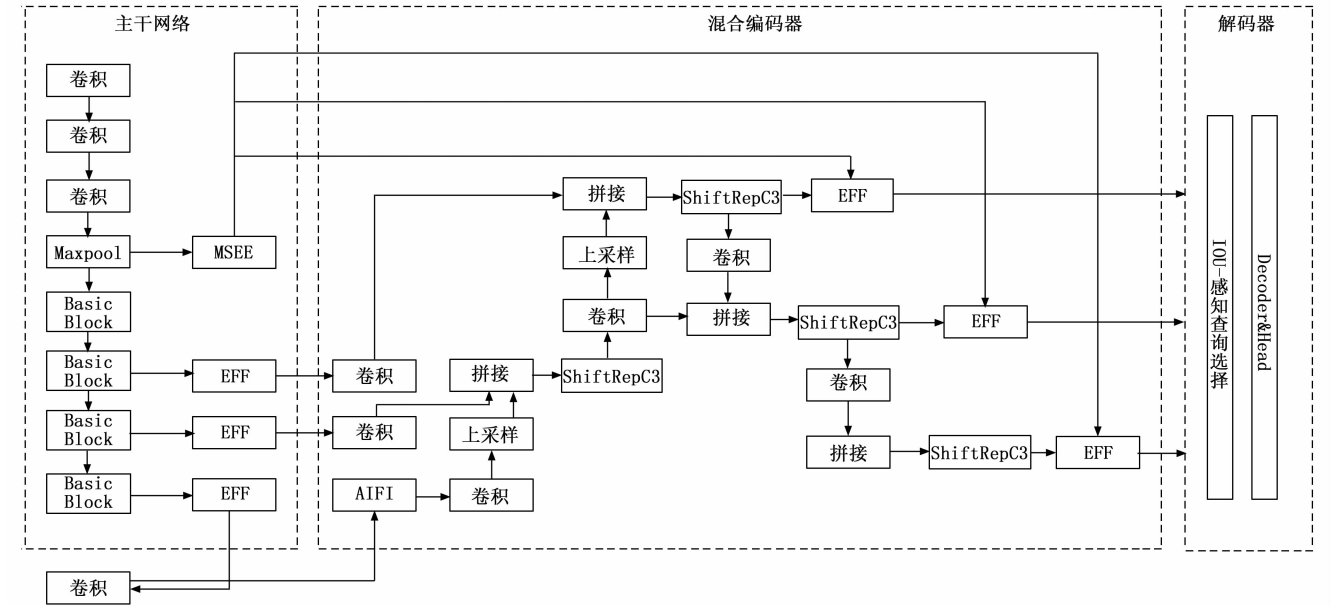


图 1 LERT-DETR 网络结构

从而增强查询对目标位置和形状的感知能力。优化后的查询向量再与混合编码器输出的多尺度特征进行交互,生成更精确的目标特征表示。

整个目标检测流程如下:输入图像首先经过主干网络提取多尺度特征,浅层特征进入边缘信息增强模块以提取细粒度边缘信息,深层特征则与边缘信息融合后传入混合编码器;在编码器阶段,特征先经 ShiftRepC3 模块进行重构,随后通过动态上采样模块提升空间分辨率;剪枝后的 Transformer 编码器进一步进行全局建模与多尺度特征整合。最终,在解码器部分,优化后的目标查询与融合特征交互,实现边界框与类别的精确预测,输出目标检测结果。

1.2 多尺度边缘信息增强方法

在目标检测任务中,边缘信息的准确提取对于提升检测精度至关重要。边缘信息作为目标物体轮廓及细节的关键表征,能够为检测模型提供丰富且具有区分度的特征,帮助模型在复杂场景中有效区分目标与背景。然而,现有主流目标检测网络往往侧重于对目标整体特征的学习,未能充分提取和有效利用边缘信息。特别是检测小目标时,由于其尺寸微小且特征相对不明显,检测网络易受到背景噪声的干扰,从而影响整体性能。为解决这一问题,本文提出了一种多尺度边缘信息增强(MEIE, multi-scale edge information enhancement)方法,该方法包括两个核心模块:多尺度边缘信息提取器(MSEE, multi-scale edge extractor)和边缘特征融合(EFF, edge feature fusion)模块。MEIE 方法通过在浅层中提取边缘信息传递到主干网络中,并与不同尺度的特征信息融合,为后续检测任务提供边缘增强的特征表示。为了缓解原始图像中边缘特征中包含的噪声信息,本文利用浅层的卷积层有效过滤不重要的背景信息,在浅层加入了 MSEE 模块。该模块生成多个不同尺度的边缘特征,能够优化主干网络的特征提取过程,并增强后续的特征融合效果,从而提高模型对复杂背景目标的检测精度。MSEE 模块结构如图 2(a)所示,对于输入的图像,首先利用 Sobel 卷积提取图像中的边缘特征,其利用 Sobel 算子分别提取输入特征图在水平方向和垂直方向的边缘梯度信息,通过深度卷积实现对每个输入通道的独立梯度计算,并输出高质量的边缘特征。Sobel 算子定义的水平梯度核 G_x 和垂直梯度核 G_y 为:

$$G_x = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (1)$$

对输入特征图 $X^{C \times H \times W}$ 的每个通道 C 独立卷积计算:

$$E_x^c = X^c \otimes G_x, E_y^c = X^c \otimes G_y \quad (2)$$

其中: $\otimes$ 表示卷积操作。

此时,边缘特征图可以表示为:

$$E^c = E_x^c + E_y^c \quad (3)$$

由于平均池化操作可能导致边缘特征被平滑处理,从而削弱边缘的显著性和对背景噪声的抑制能力,而最大池化操作可以保留局部区域中最关键的特征信息。为了能有效保留边缘特征信息并进一步增强的同时进行下采样操作,本文选择多次最大池化操作逐步降低特征图的空间分辨率。这种方式能够更加集中地捕捉边缘的关键信息,从而生成一系列具有不同尺度的特征图:

$$E_i = \text{MaxPool}(E_{i-1}), i = 1, 2, \dots, n \quad (4)$$

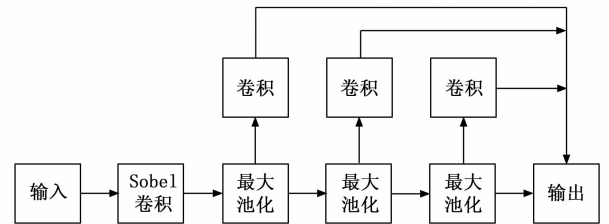
其中: n 为特征图的尺度层数。

接着,对每个尺度 E_i 的特征图应用 1×1 卷积层进行通道压缩,优化计算效率:

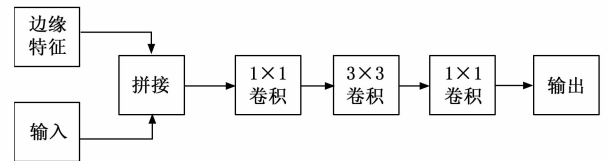
$$\hat{E}_i = \text{Conv}_{1 \times 1}(E_i), \hat{E}_i \in \mathbb{R}^{C' \times H_i \times W_i} \quad (5)$$

最终输出为多尺度的特征图集合 $\{\hat{E}_1, \hat{E}_2, \dots, \hat{E}_n\}$, 为后续特征融合提供丰富的边缘信息。

此外,由于特征融合的过程中边缘信息可能出现一定程度的丢失,因此在输出到检测头之前,本文设计了 EFF 模块对边缘信息进行进一步融合和增强。该模块将多尺度边缘特征融合为统一表示,同时提取更具表达能力的高级特征,有效整合了细粒度和粗粒度的边缘信息。如图 2(b)所示,首先对输入的两个尺度边缘特征在通道维度上拼接为一个复合张量。接着通过 1×1 卷积对拼接后的特征进行通道压缩,消除冗余信息。随后使用 3×3 卷积对融合后的特征进行局部提取和高级表示生成。最后使用 1×1 卷积调整输出通道数,确保特征的维度与后续网络结构兼容,并输出融合后的特征图。



(a) 多尺度边缘提取器



(b) 边缘特征融合模块

图 2 多尺度边缘信息提取器和边缘特征融合模块结构

1.3 ShiftRepC3 模块

特征融合阶段的性能直接影响模型对多尺度目标的感知能力和上下文信息的捕获能力。在 RT-DETR 中,特征融合阶段采用了 RepC3 结构,该结构利用标准卷积层来提取局部特征。然而 RepC3 的这种设计感受野有限,导致在处理复杂背景或多尺度目标时,模型可能

无法充分整合全局信息。此外,随着网络深度的增加,模型的计算开销显著上升,尤其是在高分辨率特征图上进行操作时,计算负担进一步加重。因此,本文设计了ShiftRepC3模块以克服了传统卷积的局限性,提供了更高效的特征融合能力,在检测精度和计算效率之间实现了更好的平衡,网络结构如图3所示。

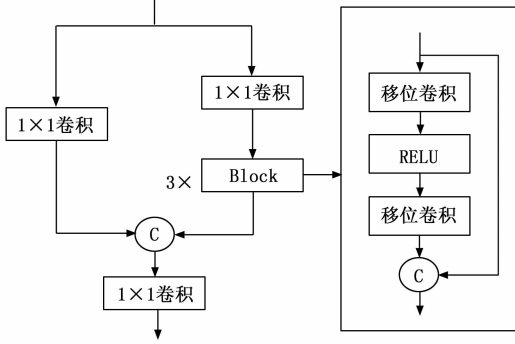


图3 ShiftRepC3 模块结构

ShiftRepC3通过移位卷积替代了传统卷积的操作,通过组合移位操作和 1×1 卷积来实现高效地提取局部特征信息,在不同偏移方向的特征组合提高特征表示的丰富性。此外,shift操作不需要额外的参数或计算,因此计算复杂度接近于 1×1 卷积,使得ShiftRepC3模块在保持较低计算复杂度的前提下,扩大了感受野,从而有效捕获更大的局部上下文信息。其网络结构如图4所示。

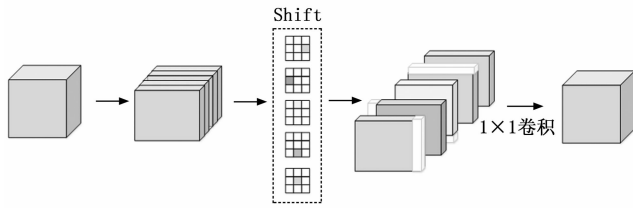


图4 移位卷积

假设输入特征图为 $X \in \mathbb{R}^{C \times H \times W}$,其中 C 、 H 、 W 分别是通道数、高度和宽度。将输入特征 X 分为 $G=5$ 个通道组:

$$\mathbf{X} = [X_1, X_2, \dots, X_5], X_i \in \mathbb{R}^{G/5 \times H \times W} \quad (6)$$

其中: $i = 1, 2, \dots, 5$ 。

对每个组执行不同方向的偏移操作,其中 X_1 向左移动 s 个像素、 X_2 向右移动 s 个像素、 X_3 向上移动 s 个像素、 X_4 向下移动 s 个像素、 X_5 不移动:

$$\text{Shift}(X_1) = \text{Roll}(X_1, -s, \dim = W),$$

$$\text{Shift}(X_2) = \text{Roll}(X_2, +s, \dim = W),$$

$$\text{Shift}(X_3) = \text{Roll}(X_3, -s, \dim = H),$$

$$\text{Shift}(X_4) = \text{Roll}(X_4, +s, \dim = H),$$

$$\text{Shift}(X_5) = X_5 \quad (7)$$

其中: $\text{Roll}(X, s, \dim)$ 表示沿指定维度 $\dim$ 滚

动 s 个像素。

将偏移后的特征重新组合为:

$$\mathbf{X}_{\text{shifted}} = [\text{Shift}(X_1), \text{Shift}(X_2), \dots, \text{Shift}(X_5)] \quad (8)$$

之后,对 $\mathbf{X}_{\text{shifted}}$ 施加一个标准的 1×1 卷积操作 W_{conv} ,用以融合通道间的信息:

$$\mathbf{X}_{\text{conv}} = W_{\text{conv}} * \mathbf{X}_{\text{shifted}} \quad (9)$$

其中: $*$ 表示卷积操作。

在两个移位卷积之间插入ReLU激活函数,使得特征的表达能力进一步提升。最终的输出可以表示为:

$$\mathbf{X}_{\text{out}} = W_{\text{conv2}} * \text{Shift}\{\text{ReLU}[W_{\text{conv1}} * \text{Shift}(\mathbf{X})]\} \quad (10)$$

其中: W_{conv1} 和 W_{conv2} 表示两次移位卷积的卷积核参数,Shift($\cdot$)表示偏移操作。

1.4 动态上采样方法

在RT-DETR中,默认采用固定规则的最邻近插值方法。然而,该方法难以处理复杂的几何变换或非线性结构,并且由于输入特征图通常是均匀缩放的,它在面对多尺度特征或不规则特征图时表现出明显的局限性。为解决这一挑战,本文采用动态上采样方法DySample,从点采样的角度对上采样的过程进行重构,通过生成内容感知的采样点来对特征图进行重新采样,实现高效的特征上采样。相比传统的动态卷积方法,它避免了复杂计算,同时能够更有效地保留输入特征的细节信息。其过程包括初始采样位置、计算偏移量、像素重排以及重新采样,DySample结构如图5所示。

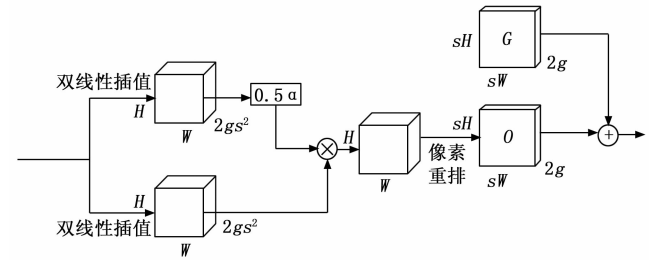


图5 DySample 结构

对于输入特征图 X ,假设维度为 $C \times H \times W$,首先生成一个初始采样网格 G ,通过采用双线性插值法,使其均匀分布于特征图中。同时,特征图 X 经过两层线性变换,每层将 X 投影到不同的维度,生成大小为 $2gs^2 \times H \times W$ 的偏移量 O :

$$O = \text{linear}(X) \quad (11)$$

接着,通过通道重排操作,将偏移量重新排列为 $2 \times sH \times sW$ 。其中,动态范围因子通过对输入特征图进行线性投影并结合Sigmoid函数生成,其取值范围被限制在 $0 \sim 0.5$ 之间:

$$O = 0.5 \text{sigmoid}[\text{linear}_1(X) \cdot \text{linear}_2(X)] \quad (12)$$

对初始采样网格 G 与偏移量相加,得到采样集 S 。利用该采样集对输入的特征图进行重新采样,生成上采样后的特征图。

1.5 基于注意力机制的剪枝方法

现有模型参数量大且计算复杂度较高,在资源受限的设备上难以部署。因此,本文引入 SAL (Shift Attention Layer)^[16]方法对 LERT-DETR 进行轻量化剪枝。其核心原理是利用注意力机制动态学习卷积核中最重要的连接,通过在训练过程中自动对卷积层进行结构化剪枝,并将其转换为更加高效的移位层,同时通过学习选择最优的移位操作来进一步优化网络性能。具体来说,假设输入张量 $X \in \mathbb{R}^{C \times L}$ 和卷积权重 $W \in \mathbb{R}^{D \times C \times S}$, 其中 C 为通道数, L 为每个特征图的长度, D 为输出通道数, S 为卷积核大小。

在训练阶段,使用注意力张量 $A \in \mathbb{R}^{D \times C \times S}$ 表示每个权重连接的重要性,动态学习哪些连接需要保留。首先,对每个第 d 输出通道和第 c 输入通道的切片 $A_{d,c,\cdot}$ 进行标准差归一化,使其分布范围一致:

$$A_{d,c,\cdot} \leftarrow \frac{A_{d,c,\cdot}}{\text{std}(A_{d,c,\cdot})} \quad (13)$$

之后,对每个通道对 (d, c) 的卷积核位置维度通过温度 T 调节 Softmax 函数,生成归一化概率分布:

$$A_{d,c,\cdot} \leftarrow \text{Softmax}(T \cdot A_{d,c,\cdot}) \quad (14)$$

温度 T 的大小控制 A 的分布是否更加接近于二值化。较大的 T 产生均匀分布,而较小的 T 会使注意力集中在某些权重上。每次迭代后,温度 T 按照衰减因子 α 更新:

$$T \leftarrow \alpha \cdot T \quad (15)$$

在剪枝阶段,使用注意力张量 A 对原始卷积权重 W 进行加权,仅保留最重要的连接。首先通过逐元素乘积更新权重张量:

$$W^A = W \odot A \quad (16)$$

其中: $\odot$ 表示逐元素相乘, W^A 中大部分权重会被削弱,仅保留 A 中重要性高的权重。

之后,使用剪枝后的权重 W^A 进行标准卷积操作:

$$Y_{d,l} = \sum_{c=1}^C \sum_{s=1}^S x_{c,l+s-[S/2]} W_{d,c,s} \quad (17)$$

其中, l 表示特征图的位置索引, s 表示卷积核的移位索引。

然后,通过反向传播同时更新 W 和 A ,使 A 学习到对目标任务最优的剪枝策略。

在推理阶段,注意力张量 A 被二值化,完成最终剪枝。首先,对 A 的每个切片 $A_{d,c,\cdot}$ 应用硬最大化操作,保留其中值最大的一个位置,其余值置为 0:

$$A_{d,c,s} = \begin{cases} 1 & \text{if } s' = \arg\max_s A_{d,c,s'} \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

根据二值化的 A ,每个切片 $W_{d,c,\cdot}$ 中仅保留一个权重,对应最优移位索引 $s^* = \arg\max_s A_{d,c,s'}$ 。此时,卷积运算转化为:

$$Y_{d,l} = \sum_{c=1}^C x_{c,l+s^*-[S/2]} W_{d,c,s^*} \quad (19)$$

2 实验与结果分析

2.1 数据集与评估指标

本文数据源自天池大数据竞赛平台,分别选取广东电网智慧作业挑战赛中的绝缘手套检测和规范着装检测训练集 2 519 张和 2 940 张。两类数据均采集自电力作业场景,涵盖了户外变电站、输电线路、配电检修等典型工作环境。图像数据包含多角度拍摄、不同背景、不同人员姿态和穿戴状态,并在一定程度上反映了不同光照条件下的作业情况,如晴天、阴天和遮挡光照。图像分辨率统一,来源设备包括移动设备和现场摄像头,能较好反映电力作业中的实际视觉感知条件,具备一定的多样性和应用普适性。

所有图像数据均经 Labelimg 工具标注处理。鉴于官方仅开放训练集数据,本文对该数据集进行随机拆分,按 7:2:1 比例随机划分训练集、验证集和测试集,此外,进行五倍数据增强以扩充数据集,包括旋转、水平翻转、随机裁剪、亮度增强与仿射变换。

绝缘手套和工作服数据集样本分布如表 1 和表 2 所示。

表 1 绝缘手套数据集介绍

标签名称	具体含义	目标数量
badge	监督员袖章	789
person	作业人员	4 793
glove	佩戴绝缘手套	7 025
wrongglove	未佩戴绝缘手套	5 179
operatingbar	手持操作杆	918
powerchecker	验电笔	794

表 2 工作服手套数据集介绍

标签名称	具体含义	目标数量
badge	监督员袖章	1 339
person	作业人员	6 837
clothes	合规工作服	4 888
wrongclothes	不合规工作服	1 718

2.2 实验设置

网络训练平台配置如下:实验环境为 ubuntu20.04 操作系统, GPU 使用 NVIDIA GeForce RTX 4090 (24 GB), CPU 使用英特尔 (R) 至强 (R) 8336C@2.30 GHz, CUDA 版本为 12.1, PyTorch 版本为 2.2.2, Python 版本为 3.8.19。

实验参数设置如下:训练迭代次数为 200, 批次大小为 8, 输入图像大小为 640×640。通过经验调整选择初始学习率为 0.000 1, 权重衰减系数为 0.001, 选择 AdamW 作为优化器。

2.3 评价指标

为了全面评估不同的模型的检测性能, 实验选择模型的参数量、计算量、 mAP_{50} 和 mAP_{50-95} 作为主要的评价指标, 具体计算公式如下:

$$P = \frac{TP}{TP + FP} \tag{20}$$

$$AP = \int_0^1 P(R) dR \tag{21}$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_c \tag{22}$$

TP (真阳性) 表示正确预测的正样本数量, FP (假阳性) 表示错误预测的正样本数量, $P(R)$ 是 PR 曲线上的精确率和召回率的函数, N 是类别数量。 mAP_{50} 代表在 IoU (Intersection over Union) 阈值为 0.5 的条件下计算的平均精度, mAP_{50-95} 代表步长 0.05、 IoU 阈值从 0.5~0.95 范围内计算平均精度的综合指标, 综合反映了模型的总体检测性能。

2.4 对比实验

在表 3 中, 对比了所提方法与几种主流目标检测方法的性能^[19-22]。本文将 RT-DETR-R18^[9] 模型作为基线模型通过实验验证了所提方法在多个方面的优势。与主流方法相比, 所提方法不仅在检测精度上表现优异, 同时在模型大小和计算复杂度方面也具有显著优势。具体而言, 与基线模型相比, 所提方法在绝缘手套数据集和工作服数据集上的 mAP_{50} 分别提升了 2.8% 和 2.1%, mAP_{50-95} 分别提升了 2.6% 和 1.5%, 表明其在高置

信度区域和更严格 IoU 阈值范围内的检测能力均得到了显著增强。此外, 所提方法通过引入 MEIE 模块、ShiftRepC3 模块、DySample 以及基于注意力机制的剪枝方法, 大幅优化了模型的参数量和计算效率, 参数量从 19.9 M 降至 7.1 M, $FLOPs$ 从 57.0 G 降至 31.1 G, 分别减少了 12.8 M 和 25.9 G。综上所述, 所提方法在精度、效率和轻量化方面实现了良好的平衡, 展现出较强的综合性能和应用潜力。

表 4 展示了不同剪枝方法在电力作业场景违规穿戴检测任务中的性能表现, 以评估不同剪枝策略对模型 mAP_{50} 、 mAP_{50-95} 、参数量和计算量的影响。实验结果显示, 与未剪枝模型相比, 所提 SAL 方法在两个数据集上的 mAP_{50} 仅下降 0.3%, mAP_{50-95} 下降 0.2%~0.4%, 但参数量和计算量分别减少了 67.3% 和 50.0%, 显著优于其他剪枝方法。例如, L_1 方法^[23]虽然压缩率较高, 但精度损失较大; Taylor-FO-BN^[24]和 Slimming^[25]在轻量化效果上不如 SAL, mAP_{50} 和 mAP_{50-95} 总体也略低于本文方法; 而 DepGraph^[26]尽管参数量较低, 但检测性能有较大的降低。综上所述, 本文的 SAL 方法在精度和效率之间实现了更优的平衡, 证明了其在保持高精度的同时有效降低计算复杂度的能力, 更适合资源受限的边缘设备部署。

2.5 消融实验

表 5 展示了在绝缘手套数据集上模型选择不同配置的消融实验结果, 旨在评估各个模块对整体模型性能的

表 3 与主流目标检测模型的性能对比

Model	$mAP_{50}/\%$		$mAP_{50-95}/\%$		$Parameters/M$	$FLOPs/G$
	insulation glove	workwear	insulation glove	workwear		
RT-DETR-R18 ^[19]	84.8	96.1	59.6	82.4	19.9	57.0
YOLOv8m	80.8	94.3	59.0	81.6	25.8	78.7
YOLOv10m ^[20]	80.2	94.0	58.3	81.2	16.5	63.4
YOLOv11m	82.4	94.9	59.3	82.0	20.0	67.7
Faster R-CNN ^[5]	81.5	94.7	58.9	81.4	41.4	206.7
Mask R-CNN ^[21]	82.1	94.6	58.3	81.1	44.0	236.2
RT-DETR-L ^[9]	85.9	96.5	60.5	82.7	28.0	100.6
SSD	78.7	93.5	57.6	80.6	12.3	63.2
RetinaNet ^[22]	82.2	94.4	59.6	81.5	35.4	204.0
Ours	87.6	98.2	62.2	83.9	7.1	31.1

表 4 不同剪枝方法对比

Methods	$mAP_{50}/\%$		$mAP_{50-95}/\%$		$Parameters/M$	$FLOPs/G$
	insulation glove	workwear	insulation glove	workwear		
未剪枝模型	87.9	98.5	62.4	84.3	21.7	62.2
L_1 ^[20]	85.5	97.3	60.8	83.7	7.5	31.1
Taylor-FO-BN ^[21]	86.3	98.0	62.0	83.8	12.8	31.1
Slimming ^[22]	86.1	97.4	61.4	83.4	8.2	31.1
DepGraph ^[23]	87.0	97.5	61.1	82.9	6.8	31.1
Ours	87.6	98.2	62.2	83.9	7.1	31.1

表 5 在绝缘手套数据集上的消融实验结果

Baseline	MEIE	ShiftRepC3	DySample	SAL	<i>mAP</i> 50/%	<i>mAP</i> 50-95/%	<i>Parameters</i> /M	<i>FLOPs</i> /G
✓					84.8	59.6	19.9	57.0
✓	✓				86.1	61.0	21.8	63.2
✓		✓			85.6	60.5	19.8	56.0
✓			✓		85.3	59.9	19.9	57.0
✓	✓	✓			86.4	61.7	21.7	62.2
✓	✓		✓		86.3	61.2	21.8	63.2
✓		✓	✓		85.7	60.9	19.7	56.0
✓	✓	✓	✓		87.9	62.4	21.7	62.2
✓	✓	✓	✓	✓	87.6	62.2	7.1	31.1

贡献。实验中使用 RT-DETR-R18^[9] 模型作为基线模型。从实验结果中可以看出,随着 MEIE、ShiftRepC3、DySample 模块逐步引入,模型的检测性能不断得到提升,证明了各模块的有效性。

在绝缘手套数据集上基线模型的 *mAP*50 和 *mAP*50-95 分别为 84.8% 和 59.6%,但计算量为 57.0 G,参数量为 19.9 M。单独引入 MEIE 方法后, *mAP*50 和 *mAP*50-95 分别提升至 86.1% 和 61.0%。这表明 MEIE 方法利用 MSEE 模块与 EFF 模块精准捕捉目标边缘,降低背景噪声干扰,提高模型的检测精度。单独采用 ShiftRepC3 模块时, *mAP*50 和 *mAP*50-95 分别为 85.6% 和 60.5%,参数量和计算量和基线模型相当,说明其以移位卷积替代传统卷积,在不增加资源消耗前提下,能有效拓展感受野,提升特征提取能力。单独引入 DySample 模块, *mAP*50 和 *mAP*50-95 分别提升至 85.3% 和 59.9%。这表明 DySample 模块通过优化上采样过程,能有效地保留输入特征的细节信息。

结合上述 3 个模块后,模型的 *mAP*50 和 *mAP*50-95 分别提升至 87.9% 和 62.4%,模型的检测性能达到最优。进一步,引入 SAL 方法后,模型的参数量和计算量分别降至 7.1 M 和 31.1 G,同时检测性能仍保持的较高水准。这表明 SAL 方法通过动态学习关键连接,将传统卷积转换为高效移位层,实现了模型轻量化,同时最大程度保留了检测精度。

2.6 可视化分析

在本节中,通过对基线模型与 LERT-DETR 模型的全面可视化比较,展示了它们在电力作业场景下的违规穿戴检测任务上的性能差异。图 6 展示了原始图片、基线模型以及 LERT-DETR 模型在绝缘手套数据集和工作服数据集上的识别结果。由图 6 (a) 可见,在绝缘手套数据集上(第 1~3 列)基线模型对于作业人员的手套出现了冗余检测框问题,即对同一目标生成多个重叠预测框,导致检测置信度分散与定位精度下降。由图 6 (b) 可见,在工作服数据集上,基线模型在第 1 列中存在误检情况,将背景中的红旗误判为

作业人员;同时,在第 2 列中,基线模型存在漏检问题,部分目标因遮挡和尺度变化未被有效识别。相比之下,改进的 LERT-DETR 模型通过引入 MEIE 方法、ShiftRepC3 模块和动态上采样方法,减少了误检和漏检现象并优化了检测结果,在复杂电力场景下具备更强的鲁棒性,为安全装备规范穿戴的智能化监管提供了可靠的技术支撑^[27-29]。



图 6 本文方法与 RT-DETR 方法的检测结果的可视化对比

3 结束语

本文提出了一种轻量化的边缘特征增强检测算法 (LERT-DETR, lightweight edge-enhanced RT-DETR) 用于检测电力作业场景下作业人员的穿戴情况,旨在解决电力作业场景中目标尺寸多样、背景复杂等挑战。LERT-DETR 通过优化主干网络、混合编码器和解码器结构,提升了模型对不同尺度目标和复杂背景的识别能力。主干网络采用 ResNet18,并提出一种多尺度边缘信息增强方法,充分利用边缘特征信息,以减少背景噪声的干扰。混合编码器部分设计了 ShiftRepC3 模块,替代传统卷积层并扩展感受野,有效提升特征提取效率;利用动态上采样方法实现高效的特征上采样,增强了对不规则特征图的处理能力。此外, LERT-DETR 模

型还通过引入基于注意力机制的SAL剪枝方法,减小了参数量和计算量。实验结果表明,LERT-DETR在电力作业场景中的检测精度和效率均优于现有方法,尤其在复杂背景和多样化目标尺寸的情况下,表现出了较强的鲁棒性和实时性。未来,LERT-DETR算法在电力作业安全监管系统中具有良好的应用潜力。结合多任务学习框架,该方法可进一步扩展至作业行为识别、安全违规检测等任务,实现多维度作业状态感知与综合分析。此外,针对不同地区、不同作业场景的视觉差异,跨领域迁移能力的提升将是算法实用化的重要方向。后续研究将重点探索轻量化部署、领域自适应与多模态融合等策略,以增强模型在实际复杂环境中的泛化能力与鲁棒性。

参考文献:

- [1] ZHANG W K, PAN L Z, GUO Z B, The visual detection method of the abnormal state of insulating gloves based on RetinaNet in power scenarios [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2022, 37 (1): 85 - 91.
- [2] COLLINS S, DEANE J P, PONCELET K, et al. Integrating short term variations of the power system into integrated energy system models; a methodological review [J]. Renewable and Sustainable Energy Reviews, 2017, 76 (1): 839 - 856.
- [3] WANG B, MA F, GE L, et al. Icing-EdgeNet: a pruning lightweight edge intelligent method of discriminative driving channel for ice thickness of transmission lines [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1 - 12.
- [4] GAVRILESCU R, ZET C, FOSALAU C, et al. Faster R-CNN: an approach to real-time object detection [C] // 2018 International Conference and Exposition on Electrical and Power Engineering, IEEE, 2018: 0165 - 0168.
- [5] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6): 1137 - 1149.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // Proceedings of the 2016 European Conference on Computer Vision, Cham: Springer, 2016: 21 - 37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2016: 779 - 788.
- [8] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C] // Proceedings of the 2020 European Conference on Computer Vision, Cham: Springer, 2020: 213 - 229.
- [9] ZHAO Y, LV W, XU S, et al. Detrs beat yolos on real-time object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965 - 16974.
- [10] SUN Y, LI X, LI J, et al. Yolov5-lite-based real-time UAV detection for high-security zones; ensuring airspace protection [C] // 2024 9th International Conference on Control, Robotics and Cybernetics, IEEE, 2024: 12 - 17.
- [11] LI J, XU Y, NIE K, et al. PEDNet: A lightweight detection network of power equipment in infrared image based on YOLOv4-tiny [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1 - 12.
- [12] CHEN J, DENG S, WANG P, et al. Lightweight helmet detection algorithm using an improved YOLOv4 [J]. Sensors, 2023, 23 (3): 1256.
- [13] 杨乐, 郭一鸣, 霍勇博, 等. 改进YOLOv5在电力生产违规穿戴检测中的应用 [J]. 电力系统保护与控制, 2023, 51 (14): 160 - 168.
- [14] 曹捷, 郭志彬, 潘立志, 等. 高空作业场景下的安全带穿戴检测 [J]. 湖南科技大学学报 (自然科学版), 2022, 37 (1): 92 - 99.
- [15] 李永福, 陈立斌, 惠君伟, 等. 基于EPSA-YOLOv5电力高空作业安全带佩戴检测 [J]. 西安工程大学学报, 2024, 38 (2): 18 - 25.
- [16] 石彦鹏, 潘作为, 成浩天. 基于YOLOv5和少样本学习的高空作业安全带检测 [J]. 电子器件, 2024, 47 (1): 215 - 220.
- [17] 顿伟超, 王淑青, 张鹏飞, 等. 基于改进YOLOv4的电力高空作业安全带检测 [J]. 湖北工业大学学报, 2022, 37 (5): 6 - 11.
- [18] 刘昶成, 邵文权, 李玲陶. 基于改进YOLOv5s的电力作业人员安全帽检测算法研究 [J]. 国外电子测量技术, 2024, 43 (2): 34 - 42.
- [19] 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, 2016: 770 - 778.
- [20] ZHANG X, ZENG H, GUO S, et al. Efficient long-range attention network for image super-resolution [C] // European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 649 - 667.
- [21] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6027 - 6037.

(下转第128页)