

# 空间域与频域特征融合的电力安全装备检测方法

周龙伟, 郭鹏程, 张仕勇, 彭家从, 田斌, 袁天霖

(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:** 电力安全装备检测对保障作业人员人身安全、降低事故风险和经济损失至关重要; 针对电力作业场景背景复杂和样本不均衡导致检测精度不高的问题, 提出一种融合空间域与频域特征的双域门控融合检测方法; 该方法设计了3种核心技术模块: 双域特征增强模块, 融合空间域与频域信息, 提升细节与边缘感知能力; 聚焦融合模块, 结合深度可分离卷积与门控机制, 聚焦关键区域, 优化多尺度特征融合效果; 时序平滑滑动损失函数, 引入动态加权策略, 提升困难样本学习表现并增强检测稳定性; 采用 RT-DETR 作为基线框架, 在绝缘手套和工作服数据集上进行了对比实验; 实验结果表明, 所提方法在绝缘手套与工作服数据集上的  $mAP_{50}$  分别提升了 3.1 和 2.4 个百分点, 在  $mAP_{50-95}$  分别提升了 2.8 和 1.8 个百分点; 所提出的 D2GF-DETR 在两个数据集上的检测精度较现有主流方法有显著提升, 同时保持了较低计算开销。

**关键词:** 电力安全装备检测; 目标检测算法; RT-DETR; 双域特征增强; 多尺度特征融合; 损失函数

## Power Safety Equipment Detection Methods Based on Spatial and Frequency Domain Feature Fusion

ZHOU Longwei, GUO Pengcheng, ZHANG Shiyong, PENG Jiacong, TIAN Bin, YUAN Tianlin  
(Ninghai County Power Supply Company, State Grid Zhejiang Electric Power Co., Ltd.,  
Ningbo 315600, China)

**Abstract:** It is of great significance for power safety equipment detection methods to ensure the personal safety of workers and reduce the risk of accidents and economic losses. In order to solve the low detection accuracy caused by complex backgrounds and imbalanced samples in power operation scenarios, a dual-domain gate fusion detection method is proposed, which integrates spatial and frequency domain features. This method includes three key technical modules: (1) a dual-domain feature enhancement module, which integrates spatial and frequency domain information to improve sensitivity to edge and detail features; (2) a focused fusion module, which combines depthwise separable convolutions with gating mechanisms to emphasize critical regions and optimize multi-scale feature integration; (3) a temporally smoothed slide loss function, which introduces that dynamic weighting strategy to enhance the learning of hard examples and improve detection stability. By using a real-time detection transformer (RT-DETR) as the baseline framework, a comparison of experiments is made on the datasets of insulated gloves and working clothes. Experimental results demonstrate that the proposed method increases by 3.1% and 2.4% in  $mAP_{50}$ , and 2.8% and 1.8% in  $mAP_{50-95}$  on the above datasets, respectively, and the proposed dual-domain gate fusion detection transformer (D2GF-DETR) algorithm is superior in detection accuracy and computational costs to mainstream algorithms.

**Keywords:** electric power safety equipment detection; target detection algorithms; RT-DETR; dual-domain feature enhancement; multi-scale feature fusion; loss function

收稿日期:2025-05-07; 修回日期:2025-06-11。

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

作者简介:周龙伟(1994-),男,工程师。

通讯作者:郭鹏程(1983-),男,工程师。

引用格式:周龙伟,郭鹏程,张仕勇,等.空间域与频域特征融合的电力安全装备检测方法[J].计算机测量与控制,2025,33(7):

139-145.

### 0 引言

电力行业作业环境复杂、任务强度高，作业人员所使用的安全装备（如绝缘手套、防护服、监督袖章等）是否正确佩戴，直接关系到其人身安全和电力作业的可靠性。传统的人工巡检与视频监控手段效率低、误判率高，难以满足大规模、高频次巡检任务的智能化要求。借助深度学习实现的自动化图像检测技，已成为提升作业安全水平的重要技术路径。

近年来，以 YOLO 系列<sup>[1]</sup>为代表单阶段目标检测方法广泛应用于电力安全装备检测中，该方法通过端到端方式直接预测目标的位置和类别，实时性较高。此外，Zhao 等人<sup>[2]</sup>提出的 RT-DETR（Real-Time DETection TRansformer）模型，通过高效混合编码器和查询选择策略，简化了单阶段方法的后处理步骤，提高模型的检测精度。然而，电力作业图像往往存在背景干扰强、目标尺度多变、样本不均衡等问题，导致模型检测精度难以进一步提升。

针对电力场景的特殊需求，研究者们对目标检测算法进行改进。例如，李建路等<sup>[3]</sup>通过自适应中值滤波法、简化传统脉冲耦合神经网络（PCNN, pulse coupled neural network）以及卷积神经网络（CNN, convolutional neural network）与门控循环单元（GRU, gated recurrent unit）网络结构融合，提升了模型的检测效率与精度。范宇恒等<sup>[4]</sup>通过引入 SE 注意力机制和加权双向特征金字塔网络增强了模型的特征提取能力，并优化损失函数提高了模型的检测精度。花磊等<sup>[5]</sup>优化网络结构，引入小目标检测分支和 WIOU Loss，提升了模型在复杂场景下的检测性能。

尽管深度学习在电力安全装备检测领域取得了进展，但现有方法依然存在一些问题。首先，现有方法主

要依赖单一的空间特征进行检测，容易受到复杂背景中噪声和冗余信息的干扰，降低检测精度。其次，在多尺度特征融合时，存在特征间信息冗余和信息丢失的问题，影响模型的鲁棒性。此外，实际场景中缺陷类样本稀缺，导致训练中出现难易度不平衡问题。因此，本文基于 RT-DETR 提出一种安全装备检测算法 D2GF-DETR（Dual-Domain Gate Fusion DETection TRansformer, D2GF-DETR），并设计了 3 种核心技术模块：1）双域特征增强模块，融合 Scharr 空间梯度与频域高频分量，增强目标边缘与细节建模能力；2）聚焦融合模块，结合深度可分离卷积与门控机制，有效过滤冗余特征、突出关键区域；3）时序平滑滑动损失函数，引入动态加权机制，提升对困难样本的学习稳定性。本文在广东电网绝缘手套与工作服数据集上进行了系统实验验证，结果表明该方法提升了模型的特征表达能力和检测精度，在复杂场景下具有更强的适应性和鲁棒性。

### 1 方法

目标检测领域中检测精度与计算复杂度的平衡至关重要。传统方法往往通过增加模型的计算量和参数量来提高精度，但这不仅加大了对计算资源的需求，还使得模型难以在边缘计算设备上高效部署。针对这一问题，本文提出一种 D2GF-DETR 检测模型，在保持降低参数量与计算量的情况下，有效提升对电力安全装备检测的精度，网络结构如图 1 所示。主干网络用于提取图像的多尺度特征，为了有效地抑制背景噪声，设计双域增强（DDFE, dual-domain feature enhancement）模块来结合图像的空间域和频率域特征，实现更精细的特征建模。为了消除特征融合时冗余信息引入导致关键信息丢失的问题，在混合编码器部分，设计了聚焦融合模块以增强特征融合能力，提升模型对复杂场景的适应性并保持模

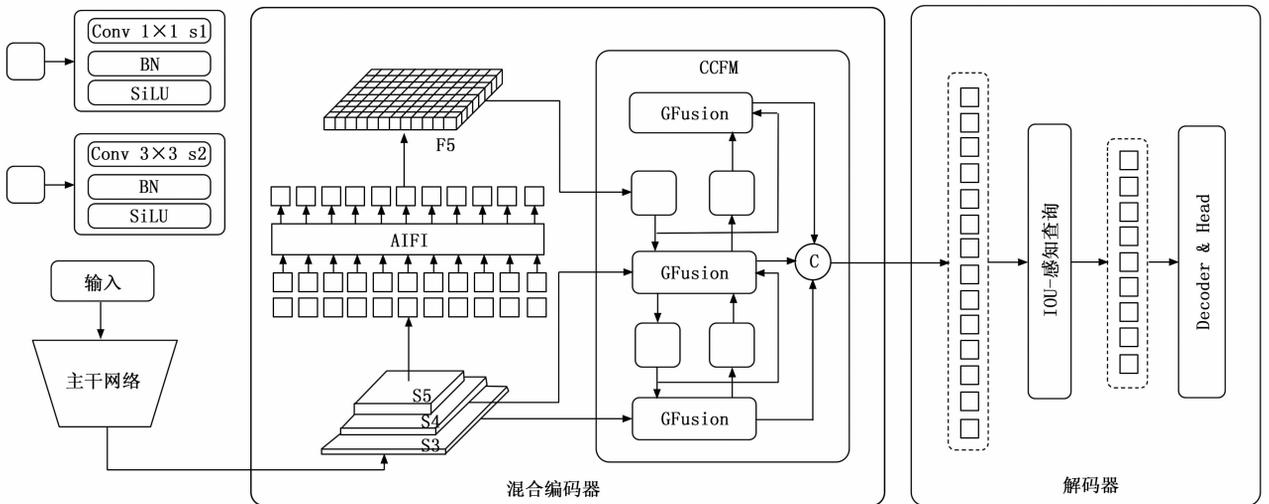


图 1 D2GF-DETR 网络结构

型的轻量化。在解码器部分, 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, 从而增强查询对目标位置和形状的感知能力。

### 1.1 双域特征增强模块

在目标检测任务中, 模型通常依赖多尺度特征提取来平衡计算效率与检测精度。然而, 现有的特征提取方法主要集中在空间域, 忽视了频域特征的重要性。空间域特征主要捕捉图像的整体轮廓和结构等低频信息, 但对细节、边缘等高频信息的敏感度较低。而频域特征中则包含了图像的细节、边缘以及局部纹理等高频信息, 有助于模型区分不同类别物体、提升小物体和边界模糊物体的检测精度, 并能够有效地抑制背景噪声, 减少冗余信息。因此, 本文提出了 DDFE 模块, 通过结合空间域和频率域特征, 实现更精细的特征建模。DDFE 模块首先在空间域中通过 Scharr 算子提取梯度特征, 以增强局部边缘信息。Scharr 算子用于计算图像在水平和垂直方向的梯度信息, 提取边缘特征。其在每个通道上分别进行卷积操作, 输出空间域梯度图。在频域处理方面, 图像块先被裁剪为固定大小 (64×64), 然后通过 2D 快速傅里叶变换在频率域中捕获图像的高频与低频成分, 获取全局特征, 最后将空间域与频域信息融合, 从而形成增强特征图。DDFE 模块结构如图 2 所示, 整个过程可以表示为:

$$F_{\text{final}} = H\{S(X), F^{-1}\{G[F(X)]\}\} \quad (1)$$

其中:  $S(\cdot)$  表示空间域处理,  $F(\cdot)$  和  $F^{-1}(\cdot)$  分别表示 FFT 及其逆变换,  $G(\cdot)$  表示在频域上进行的卷积操作,  $H(\cdot)$  表示双域特征增强模块中的特征融合操作, 指的是对由 Scharr 算子提取的空间域梯度图与频域中的高频特征图在通道维度上的融合。

在空间域部分, DDFE 模块采用 Scharr 算子进行梯度检测。Scharr 算子是一种改进的差分近似算子, 相较于 Sobel 算子, 其梯度近似误差更小, 并能更好地保留边缘、细节等高频信息。设 Scharr 算子的卷积核为  $K_x$  和  $K_y$ , 对于输入图像  $K$ , 其梯度计算方式为:

$$\left. \begin{aligned} G_x &= X * K_x \\ G_y &= X * K_y \end{aligned} \right\} \quad (2)$$

经过上述处理后, 得到的边缘特征图  $G=0.5G_x+0.5G_y$  经过后续卷积层融合原始输入信息, 形成空间增强特征  $F_s$ 。具体而言, 对于每一个输入图像块, 首先裁剪为大小为 64×64 的窗口, 然后在每个通道上独立执行二维快速傅里叶变换, 输出复数形式的频谱图。所得复数张量包含实部与虚部两个部分, 分别表示图像在水平与垂直方向上各频率成分的响应强度。随后计算幅度谱, 以完整反映频率能量分布。其幅度通过沿通道拼

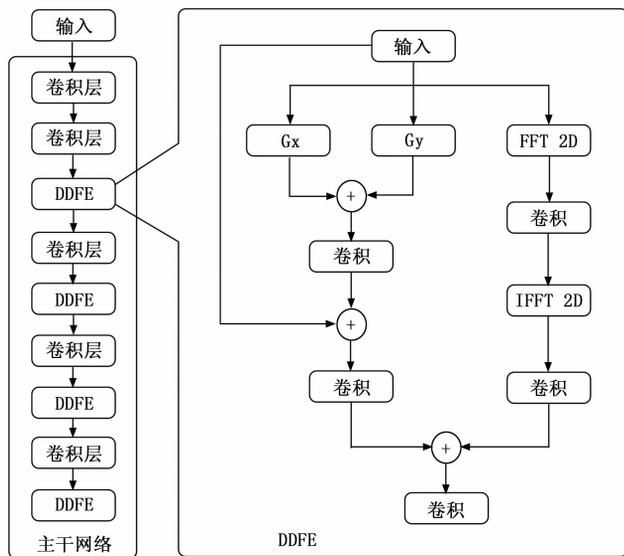


图 2 双域特征增强模块结构

接, 将两部分合并:

在频域部分, 输入  $X$  通过 2D 快速傅里叶变换 (FFT) 后, 得到频域表示  $F_y$ , 其包括实部和虚部, 通过沿通道拼接, 将两部分合并:

$$F_f^{\text{cat}} = \text{Concat}(F_f^{\text{real}}, F_f^{\text{imag}}) \quad (3)$$

然后, 对  $F_f^{\text{cat}}$  进行卷积操作, 以增强频域特征, 并通过逆 FFT 将增强特征投影回空间域:

$$F_f^{\text{spatial}} = F^{-1}[\text{Conv}(F_f^{\text{cat}})] \quad (4)$$

最后, 将空间域特征  $F_s$  与频域特征  $F_f^{\text{spatial}}$  逐元素相加, 并通过 1×1 卷积层进行通道整合, 形成增强后的特征图。

### 1.2 聚焦融合模块

特征融合模块的设计直接影响模型的检测性能。RT-DETR<sup>[2]</sup> 通过 Fusion 模块来提取多尺度特征并进行融合, 但在复杂的背景环境下对细节特征的处理过程容易引入冗余信息, 导致部分重要信息丢失。针对这一问题, 本文设计了聚焦融合 (FF, focused fusion) 模块以增强特征融合能力, 提升模型对复杂场景的适应性。通过深度可分离卷积和门控卷积对原有的融合模块改进以更好的抑制融合过程中噪声或冗余信息并保持模型的轻量化。网络结构如图 3 所示, 深度可分离卷积采用 3×3 标准内核, 包含深度可分离卷积和点卷积两个阶段, 分别用于提取空间特征与通道整合。整个结构由 3 个串联的 1×1 s 深度可分离卷积块和 ReLU 激活组成。在每层卷积后嵌入门控机制, 用于生成通道注意力掩码。输出与主干路径相乘实现显式特征筛选, 最后经 1×1 卷积整合输出特征。

对输入门控卷积的特征  $x$  首先进行归一化处理得到

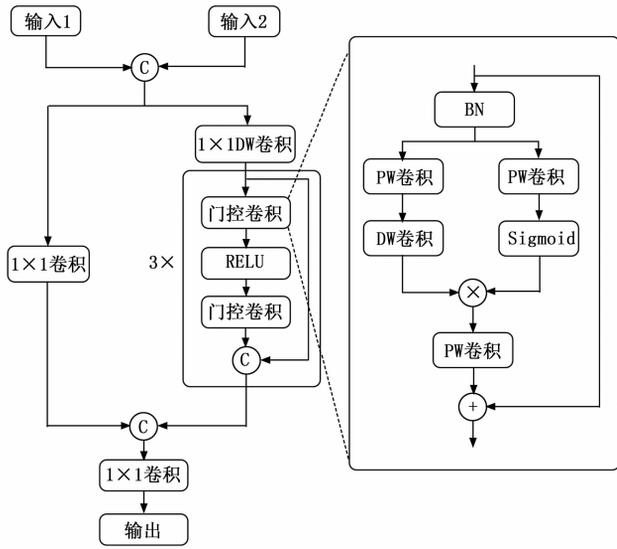


图 3 聚焦融合模块结构

$\hat{x}$ ，之后对  $\hat{x}$  经过门控路径和特征变换路径，门控路径采用  $1 \times 1$  逐点卷积生成门控权重，然后通过 Sigmoid 激活函数进行激活得到  $x_1$ 。特征变换路径首先通过  $1 \times 1$  逐点卷积提取特征，随后通过  $3 \times 3$  深度可分离卷积进行局部特征提取得到  $x_2$ ：

$$\begin{aligned} x_1 &= \sigma[PWConv_1(\hat{x})] \\ x_2 &= DWConv[PWConv_2(\hat{x})] \end{aligned} \quad (5)$$

然后，通过逐元素相乘，将门控路径产生的  $x_1$  作为门控信号调节特征变换路径输出的  $x_2$ ，从而实现信息流的有效筛选：

$$y' = x_1 \cdot x_2 \quad (6)$$

最后，再次经过  $1 \times 1$  逐点卷积，将融合后的特征投影回原始维度，并与输入特征进行残差连接，形成最终输出。

### 1.3 时序平滑滑动损失

在电力安全装备检测任务中，样本的难易不平衡问题较为突出。电力作业场景中绝缘手套、监督员袖章、验电笔等工具的尺度变化范围大，复杂背景中相似形状物体混杂增加了困难样本的识别难度。这种样本复杂度分布失衡现象可能导致模型在训练过程中对简单样本产生过拟合现象，同时在困难样本的学习上表现出表征能力不足的问题。Slide Loss<sup>[6]</sup>通过对难以判别的困难样本赋予更高权重，从而使它们在训练时能够获得更多关注。具体方法是将所有边界框的 *IoU* 值平均值作为阈值，低于该阈值的视为负样本，高于阈值的视为正样本。对 *IoU* 值接近阈值的样本，分配更高的权重，有助于模型更好地优化并处理这些模糊样本。

然而，在实际训练时，Slide Loss 在调整较难区分样本的权重时，可能会因损失值波动而导致训练不稳

定。为了解决这个问题，本文引入了指数移动平均的思想对其进行优化，提出了时序平滑滑动损失函数（TSS Loss, temporally smoothed slide Loss）。通过构建动态权重平滑函数，将历史梯度信息融入当前权重计算过程，有效抑制训练过程中的梯度震荡现象。具体而言，每次更新步骤  $t$  可表示为：

$$IOU'_t = d_t \cdot (IOU'_t) + (1 - d_t) \cdot IOU_t \quad (7)$$

其中： $d_t$  为步骤  $t$  的衰减因子，定义为：

$$d_t = decay \cdot (1 - e^{-t/\tau}) \quad (8)$$

其中： $decay$  为 0.999， $\tau$  为 200。用于调节样本加权强度随训练时间变化的平滑程度。

此时，加权函数  $f(x)$  可以表示为：

$$f(x) = \begin{cases} 1 & x \leq IOU'_t - 0.1 \\ e^x - x & IOU'_t - 1 < x < IOU'_t \\ e^x - 1 & x \geq IOU'_t \end{cases} \quad (9)$$

最终，TSS Loss 计算公式如下：

$$L = loss \times f(x) \quad (10)$$

其中： $loss$  为基础损失函数输出值。

通过指数移动平均的思想，TSS Loss 能够在平滑及动态调整 *IoU* 值的同时，关注难以处理的困难样本，并有效平衡当前观测数据与历史数据的影响，使损失值更加平稳，从而提升模型的鲁棒性。

## 2 实验

### 2.1 数据集与评价指标

本文数据来源于天池大数据竞赛中广东电网智慧现场作业挑战赛，包括绝缘手套数据集和工作服数据集，分别包含 2 519 张和 2 940 张关于电力作业现场的图片，采用 Labelimg 工具标注。绝缘手套数据集图像分辨率集中于  $1\ 280 \times 720$  和  $1\ 920 \times 1\ 080$ ，工作服数据集分辨率分布更广，存在部分高动态范围图像。目标物体尺寸变化大，存在大尺度人物与小尺度装备穿戴部件。由于竞赛只公开了训练集，本文将训练集重新按 7 : 2 : 1 比例随机划分为训练集、验证集和测试集，并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。

为了比较不同的模型对电力安全装备的检测性能，实验将采用的参数指标有模型参数量、计算量、 $mAP_{50}$  和  $mAP_{50-95}$ 。参数量和计算量用于衡量模型的复杂度和检测速度。参数量越小，模型越轻量化；计算量越小，对设备硬件要求越低，检测速度越快。同时， $mAP$  用于衡量模型在多类别检测任务中的整体性能，值越大代表模型识别精度越高。 $mAP_{50}$  代表在 *IoU* 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 到 0.95 范围内计算平均精度的

综合指标, 综合反映了模型在该阈值下的检测性能。

## 2.2 实验设置

本实验环境基于 ubuntu20.04 操作系统, GPU 为 NVIDIA GeForce RTX 4090 (24 GB), CPU 为英特尔 8336C@2.30 GHz, CUDA 版本为 12.3, 采用 Pytorch 2.2.2, Python 版本为 3.8.17。在超参数配置方面, 训练的总迭代次数为 200, 批次大小为 8, 输入图像的尺寸为  $640 \times 640$ 。初始学习率设置为 0.000 1, 权重衰减系数为 0.001, 优化器选择 AdamW。

## 2.3 对比实验

表 1 对比了所提方法与几种主流目标检测方法的性能, 以 RT-DETR-R18<sup>[2]</sup> 模型作为基线模型。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将  $mAP_{50}$  分别提升了 3.1 和 2.4 个百分点, 在绝缘手套数据集和工作服数据集上将  $mAP_{50-95}$  分别提升了 2.8 和 1.8 个百分点, 证明所提方法具有更高的检测精度。在计算效率方面, 所提方法的计算量和参数量分别为 54.8 G 与 19.3 M, 相较于 Faster R-CNN<sup>[8]</sup>、Mask R-CNN<sup>[9]</sup> 等主流的目标检测模型保持了相对较低的开销, 并且具有更高的检测精度。此外, 与 YOLO 系列的方法如: YOLOv8m、YOLOv10m<sup>[7]</sup>、YOLOv11m 等模型相比, 在绝缘手套数据集和工作服数据集上取得了更高的检测精度。与计算量和参数量相近的 YOLOv10m<sup>[7]</sup> 模型相比, 所提方法在绝缘手套数据集和工作服数据集上将  $mAP_{50}$  分别提高了 7.7 和 4.5 个百分点, 在绝缘手套数据集和工作服数据集上将  $mAP_{50-95}$  分别提升了 4.1 和 3.2 个百分点。这表明, D2GF-DETR 采用了更加高效的特征提取和定位策略, 从而在复杂场景下能够更好地捕捉细节信息。综上所述, 所提方法在保持较低参数量和计算量的同时, 能够实现较高的检测精度, 具有更好的性能-效率平衡。

表 1 与主流目标检测模型的性能对比

模型	$mAP_{50} / \%$		$mAP_{50-95} / \%$		参数量 / M	计算量 / G
	绝缘手套数据集	工作服数据集	绝缘手套数据集	工作服数据集		
RT-DETR-R18 <sup>[2]</sup>	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 <sup>[7]</sup>	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 <sup>[8]</sup>	81.5	94.7	58.9	81.4	41.4	206.7
Mask R-CNN <sup>[9]</sup>	82.1	94.6	58.3	81.1	44.0	236.2
Deformable-DETR <sup>[10]</sup>	85.2	95.7	60.0	82.5	40.1	196.0
RT-DETR-L <sup>[2]</sup>	85.9	96.5	60.5	82.7	28.0	100.6
SSD <sup>[11]</sup>	78.7	93.5	57.6	80.6	<b>12.3</b>	63.2
RetinaNet <sup>[12]</sup>	82.2	94.4	59.6	81.5	35.4	204.0
D2GF-DETR	<b>87.9</b>	<b>98.5</b>	<b>62.4</b>	<b>84.2</b>	19.3	<b>54.8</b>

表 2 对比了不同损失函数在电力安全装备检测任务中的表现。本实验将使用 GIoU 的 RT-DETR-R18<sup>[2]</sup> 模型作为基线模型, 比较了不同的损失函数。实验结果显示, 与基线模型相比, TSS Loss 在绝缘手套数据集和工作服数据集上  $mAP_{50}$  均提升了 1.2 个百分点, 在绝缘手套数据集和工作服数据集上将  $mAP_{50-95}$  分别提升了 1.2 和 0.7 个百分点。TSS Loss 在两个数据集上的提升幅度不同, 在工作服数据集上提升幅度更大, 主要由于其样本遮挡情况严重, TSS 机制可在目标时序位置略偏移时稳定优化边界预测效果; 而在绝缘手套数据集中, TSS Loss 则更注重特征学习的平滑连续性。此外, 表 2 中结果可以看出, Inner-IoU<sup>[13]</sup> 和 MPDIoU<sup>[14]</sup> 等损失函数相较于基线模型性能略有提升。Focal Loss 和 Focaler-IoU 在数据集上检测性能均不佳, 未能有效解决背景复杂和难分样本的问题。与 Slide Loss<sup>[6]</sup> 相比, TSS Loss 通过引入时序平滑机制, 减少了由于权重波动引起的训练不稳定性, 使得模型在面临复杂背景和小尺度目标时表现得更加鲁棒。

表 2 与不同损失函数性能对比

方法	$mAP_{50} / \%$		$mAP_{50-95} / \%$	
	绝缘手套数据集	工作服数据集	绝缘手套数据集	工作服数据集
基线模型	84.8	96.1	59.6	82.4
Inner-IoU <sup>[13]</sup>	84.9	95.7	59.9	82.6
MPDIoU <sup>[14]</sup>	85.7	96.3	60.2	82.9
Slide Loss <sup>[6]</sup>	85.3	96.5	60.4	82.8
Focal Loss <sup>[15]</sup>	84.6	95.8	59.7	82.2
Focaler-IoU <sup>[16]</sup>	83.6	95.7	59.5	82.1
TSS Loss	<b>86.0</b>	<b>97.3</b>	<b>60.8</b>	<b>83.1</b>

表 3 显示了不同特征融合方法的对比实验结果, 使用 RT-DETR-R18<sup>[2]</sup> 模型作为基线模型, 与 RMT<sup>[17]</sup>、Conv3XC<sup>[18]</sup> 等五种方法进行对比。实验结果表明, FFusion 在绝缘手套和工作服数据集上的  $mAP_{50}$  和  $mAP_{50-95}$  等指标均优于其他方法, 同时保持了较低的计算开销。相较其他方法, FFusion 有效提升了多尺度特征的融合效果, 提升了模型的检测精度。

表 3 不同方法对 Fusion 改进对比

模型	$mAP_{50} / \%$		$mAP_{50-95} / \%$		参数量 / M	计算量 / G
	绝缘手套数据集	工作服数据集	绝缘手套数据集	工作服数据集		
基线模型	84.8	96.1	59.6	82.4	19.9	57.0
RMT <sup>[17]</sup>	85.4	96.5	59.8	82.6	18.5	50.2
Conv3XC <sup>[18]</sup>	84.3	95.7	59.5	82.1	19.9	57.0
DGCST <sup>[19]</sup>	84.5	95.8	59.2	82.0	18.5	50.1
DBB <sup>[20]</sup>	85.2	96.4	<b>60.5</b>	82.7	<b>17.7</b>	57.0
DRB <sup>[21]</sup>	85.0	96.3	59.8	82.5	18.1	<b>48.3</b>
FFusion	<b>85.8</b>	<b>96.9</b>	60.4	<b>82.9</b>	18.7	51.1

表 4 在绝缘手套数据集上消融实验结果

基线模型	DDFE	FFusion	TSS Loss	$mAP_{50}$ /%	$mAP_{50-95}$ /%	参数量/M	计算量/G
✓				84.8	59.6	19.9	57.0
✓	✓			86.2	61.0	20.5	60.6
✓		✓		85.8	60.4	18.7	51.1
✓			✓	86.0	60.8	19.9	57.0
✓	✓	✓		87.0	61.7	19.3	54.8
✓	✓		✓	87.3	61.3	20.5	60.6
✓		✓	✓	86.7	60.8	<b>18.7</b>	<b>51.1</b>
✓	✓	✓	✓	<b>87.9</b>	<b>62.4</b>	19.3	54.8

## 2.4 消融实验

表 4 展示了在绝缘手套数据集上的消融实验结果, 使用 RT-DETR-R18<sup>[2]</sup> 模型作为基线模型。模型的  $mAP_{50}$  和  $mAP_{50-95}$  相较于基线模型均提高了 1.4 个百分点。引入 DDFE 模块后, 这表明 DDFE 模块通过融合空间域与频域的特征, 增强了对细节和边缘信息的捕捉。引入 FFusion 模块后, 模型的  $mAP_{50}$  和  $mAP_{50-95}$  相较于基线模型分别提升了 1.0 和 0.8 个百分点, 并且计算量和参数量分别减少了 1.2 G 和 6.9 M, 表明 FFusion 模块有效抑制冗余信息, 优化多尺度特征融合效果并减轻了计算开销。引入 TSS Loss 后, 模型的  $mAP_{50}$  和  $mAP_{50-95}$  相较于基线模型分别提升 1.2 和 0.8 个百分点, 说明 TSS Loss 通过动态调整样本权重, 提升了模型在困难样本上的学习效果, 提高了模型的检测精度和鲁棒性。当同时加入上述 3 个模块后, 模型的  $mAP_{50}$  和  $mAP_{50-95}$  达到了最优, 并在计算效率上保持了较好的平衡。

## 2.5 可视化分析

在本节中, 通过对基线模型与 D2GF-DETR 模型的全面可视化比较, 展示了在电力作业场景安全装备检测任务中的性能差异。图 4 (a) ~ (b) 中的每一列分别展示了原始图片、基线模型以及 D2GF-DETR 模型在绝缘手套数据集和工作服数据集上的识别结果, 其中预测框中的类别及其对应的置信度得分均显示在框内。在图 4 中, D2GF-DETR 相较于基线模型在处理小尺度目标和遮挡目标方面展现出更强的检测能力。以图 4 (a) 第 1 列为例, 基线模型未能识别出未佩戴绝缘手套的作业人员, 存在明显漏检, 而 D2GF-DETR 成功定位并识别该目标, 体现出其对小目标的识别优势; 图 4 (a) 第 2 列中, 基线模型对绝缘手套产生冗余检测框, 而 D2GF-DETR 抑制了该类误检。在图 4 (b) 第 1 列中, 基线模型未能检测出监督员袖章, 而 D2GF-DETR 通过频域增强抑制了背景干扰, 准确识别出该目标, 说明其在复

杂场景中具有更强的鲁棒性与可靠性。进一步分析发现, 典型误检往往发生在背景纹理与装备图案相似的位置, 尤其在绿植背景和塔架结构下容易将非目标区域误判为手套或袖章; 而漏检则多集中于装备被遮挡、光照强反差、穿戴形态变形等情况, 导致边界模糊或特征退化。D2GF-DETR 通过频域增强提取边缘及纹理主导的高频成分, 有效提升了对遮挡区域中局部显著特征的响应能力, 同时, FFusion 模块中的门控机制可动态抑制冗余背景响应, 提升关键区域特征表达。此外, TSS Loss 对困难样本设置动态加权, 增强了模型在样本特征模糊或位置偏移时的稳定性。因此, 整体模型对误检和漏检的抵抗能力显著优于基线方法。

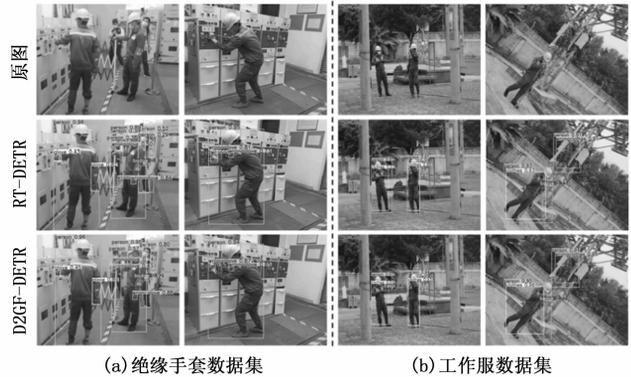


图 4 绝缘手套与工作服数据集可视化对比

## 2.6 应用流程说明

本节介绍所提模型应用于电力安全装备检测的具体实现过程, 包括现场图像采集、模型推理流程、嵌入式部署、坐标还原与穿戴状态判断等关键步骤。具体实现流程如下: 首先, 部署于作业现场的工业相机定时采集高分辨率图像, 通过边缘计算终端进行图像归一化、分辨率缩放等预处理。随后, 图像输入已部署 D2GF-DETR 模型的嵌入式推理设备 (如 Jetson Xavier NX), 模型提取空间与频域特征, 完成特征融合后输出检测结果。模型输出的类别与边界框被实时解析, 通过坐标映射将装备位置还原至原图上, 并结合已设定规范 (如绝缘手套是否佩戴) 进行装备穿戴状态判定。判定结果用于驱动后端告警系统或生成检测报告。整个流程具备低延迟、高鲁棒性和良好的现场适应性, 可支撑全天候巡检需求。

## 3 结束语

论文提出了一种 D2GF-DETR 模型, 通过提出的双域特征增强模块、聚焦融合模块和时序平滑滑动损失函数, 提升了电力作业环境下的目标检测精度和鲁棒性, 以提高电力行业安全监管的效率和准确性。DDFE 模块

结合空间域与频域特征,通过提取梯度特征和频域信息,增强了对局部边缘信息及细节的捕捉能力。FFusion 特征融合模块通过引入深度可分离卷积与门控卷积,优化了多尺度特征的融合过程,有效抑制了冗余信息的干扰,保持了特征的表达能力。TSS Loss 利用指数移动平均机制平滑损失函数的训练过程,动态调整样本的权重,确保困难样本得到更多关注,并避免训练过程中的过拟合现象。实验结果表明,所提出的 D2GF-DETR 模型在电力安全装备检测任务中,取得了更高的检测精度和较低的计算复杂度。尽管模型的检测精度表现优异,但计算复杂度仍然较高,未来将进一步优化计算效率,以更好地满足电力作业现场的实时性要求。

#### 参考文献:

- [1] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2016: 779-788.
- [2] 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, Piscataway: IEEE, 2024: 16965-16974.
- [3] 李建路,朱 珠,李 柯,等. 基于智能感知与 AR 的电力设备故障检测识别技术 [J]. 信息技术, 2024, (12): 136-140.
- [4] 范宇恒,焦良葆,郑良成,等. 基于改进 YOLOv8 的电力作业人员安全带检测 [J]. 计算机测量与控制, 2024, 32 (11): 140-145.
- [5] 花 磊,米奕莹. 基于 Protect-YOLO 的变电站电力作业人员佩戴安全防护用具检测 [J]. 机电信息, 2024, (18): 9-12.
- [6] YU Z, HUANG H, CHEN W, et al. Yolo-facev2: A scale and occlusion aware face detector [J]. Pattern Recognition, 2024, 155: 110714.
- [7] WANG A, CHEN H, LIU L, et al. Yolov10: Real-time end-to-end object detection [J]. ArXiv Preprint ArXiv: 2405.14458, 2024.
- [8] 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.
- [9] HE K, GKIOXARI G, DOLLAR P, et al. Mask r-cnn [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2017: 2961-2969.
- [10] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable transformers for End-to-End object detection [C] // International Conference on Learning Representations, 2021: 1-16.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C] // Proceedings of the European Conference on Computer Vision, Cham: Springer International Publishing, 2016: 21-37.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2017: 2980-2988.
- [13] ZHANG H, XU C, ZHANG S. Inner-IoU: more effective intersection over union loss with auxiliary bounding box [J]. ArXiv Preprint ArXiv: 2311.02877, 2023.
- [14] MA S, XU Y. Mpdious: a loss for efficient and accurate bounding box regression [J]. ArXiv Preprint ArXiv: 2307.07662, 2023.
- [15] ROSS T Y, DOLLAR G. Focal loss for dense object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2017: 2980-2988.
- [16] ZHANG H, ZHANG S. Focaler-IoU: More focused intersection over union loss [J]. ArXiv Preprint ArXiv: 2401.10525, 2024.
- [17] FAN Q, HUANG H, CHEN M, et al. Rmt: Retentive networks meet vision transformers [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2024: 5641-5651.
- [18] WAN C, YU H, LI Z, et al. Swift parameter-free attention network for efficient super-resolution [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2024: 6246-6256.
- [19] GONG W. Lightweight object detection: A study based on YOLOv7 integrated with ShuffleNetv2 and vision transformer [J]. ArXiv Preprint ArXiv: 2403.01736, 2024.
- [20] DING X, ZHANG X, HAN J, et al. Diverse branch block: Building a convolution as an inception-like unit [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2021: 10886-10895.
- [21] DING X, ZHANG Y, GE Y, et al. UniRepLKNNet: A universal perception large-kernel ConvNet for audio video point cloud Time-Series and image recognition [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2024: 5513-5524.