光电技术及应用

DOI：10.16818/j.issn1001－5868.20260412001

# 融合双域特征的绝缘子缺陷检测算法

党宏社[1*]，张代言[1]，张兰馨[1]，魏赵恺[1]，张选德[2]
（陕西科技大学1. 电气与控制工程学院；2. 电子信息与人工智能学院，西安 710021）

**摘　要：** 在无人机航拍输电线路巡检场景中，针对复杂环境干扰导致的绝缘子缺陷检测精度不高、误检、漏检等问题，文章基于实时端到端目标检测模型（RT-DETR）算法，提出一种融合双域特征的绝缘子缺陷检测算法。在主干网络中设计空频特征融合（FreqSF）模块，通过融合空间域与频率域信息，增强了绝缘子缺陷细节纹理特征的表达能力；引入高低频注意力（HiLo）机制替换多头自注意力（MHSA）机制，将特征图中的高频与低频特征分开处理，以强化对局部细节特征的提取；采用三重特征编码器（TFE）与尺度序列特征融合（SSFF）模块，将深层语义信息与浅层细节信息进行深度融合。实验结果表明，所提算法相较于主流算法取得了较好的检测效果，其mAP@0.5达到0.974，相较于基线算法提高1.5%，精确率和召回率分别提高了1.2%和1.1%。

**关键词：** 绝缘子缺陷；RT-DETR；目标检测；特征融合
**中图分类号：** TP391.41　**文章编号：** 1001－5868(2026)05－0953－08

## Insulator Defect Detection Algorithm Based on Dual-Domain Feature Fusion

DANG Hongshe[1], ZHANG Daiyan[1], ZHANG Lanxin[1], WEI Zhaokai[1], ZHANG Xuande[2]
(1. School of Electrical and Control Engineering; 2. School of Electrical Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, CHN)

**Abstract:** To address the issues of low detection accuracy and frequent false positives and false negatives caused by complex environmental interferences during unmanned aerial vehicle inspections of power transmission lines, this study proposes a dual-domain feature fusion algorithm for insulator defect detection based on Real-Time Detection Transformer (RT-DETR). First, a Spatial-Frequency Feature Fusion module is designed within the backbone network. By integrating spatial- and frequency-domain information, this module significantly enhances the feature-representation capability for fine textures of insulator defects. Second, a High-Low-Frequency Attention mechanism is introduced to replace the conventional Multi-Head Self-Attention mechanism. Processing the high- and low-frequency features separately strengthens the extraction of local-detail features. Finally, a Triple Feature Encoder and a Scale Sequence Feature Fusion module are employed to achieve the deep fusion of high-level semantic information and low-level detail information. Experimental results demonstrate that the proposed algorithm outperforms mainstream detection models. Specifically, it achieves a mAP@0.5 of 0.974, which is 1.5% higher than that of the baseline algorithm, as well as improved precision and recall by 1.2% and 1.1%, respectively.

**Key words:** insulator defect; RT-DETR; object detection; feature fusion

**收稿日期：**2026－04－12.
**基金项目：**国家自然科学基金项目（61871260）.
*通信作者：党宏社

## 0 引言

随着我国输电网络规模的不断扩大,线路中绝缘子的数量持续增加。为了保障电力系统的安全稳定运行,对输电线路绝缘子进行定期巡检十分重要。近年来,随着无人机技术的发展,基于无人机图像的非接触式巡检方式因其效率高、覆盖范围广,已逐渐成为输电线路巡检的主要手段[1]。面对海量巡检图像数据,基于深度学习的检测方法凭借其优越的特征提取与泛化能力,相较传统视觉方法显著提升了检测性能,已成为输电线路视觉检测领域的重要研究方向[2]。

针对绝缘子缺陷检测精度不高等问题,研究者从多尺度特征提取、注意力机制以及特征融合结构等方面开展了大量研究。在多尺度特征提取方面,孙虹等[3]在YOLOv7中引入多尺度卷积块注意力模块并改进空间金字塔池化结构,有效提升了多尺度特征表达能力;王艳等[4]设计了多尺度特征提取模块并结合自适应交互融合机制,提高了多尺度目标的检测精度;曾业战等[5]利用密集连接结构与多尺度特征提取模块增强特征复用能力,从而提高了检测精度。在注意力机制方面,刘宏志等[6]通过在YOLOv11n算法中结合通道注意力与特征聚合机制,以增强通道间特征依赖关系,提高小目标特征提取能力;彭晏飞等[7]在YOLOv5s中引入CA注意力机制并结合Transformer结构,以增强复杂背景下的特征表征能力。在特征融合方面,Tao等[8]设计了动态卷积颈部结构与多尺度注意力模块,强化了复杂背景下小目标细节特征提取能力;Wu等[9]通过引入自适应卷积结构并结合加权特征融合策略,进一步优化了复杂背景下的检测效果。

随着Transformer在目标检测领域的广泛应用,基于DETR框架的检测方法逐渐成为研究热点[10]。Chen等[11]设计双向跨尺度特征融合网络并提出动态直方图自注意力机制,以增强低对比度缺陷的特征响应能力。Cheng等[12]在检测框架中引入适配器模块对预训练权重进行自适应调节,并结合高斯显著性引导机制与跨层信息交互,提高了小规模数据条件下的绝缘子缺陷检测性能。Chen等[13]通过设计双域选择机制模块,并结合自适应稀疏自注意力结构,提升了复杂背景下输电线路缺陷检测性能。

综上,现有研究虽取得一定进展,但在复杂场景下仍存在两点局限:一是模型对绝缘子细小纹理与模糊边缘特征的表达能力不足;二是复杂背景易引发特征混淆,导致误检率偏高。为此,本文基于实时端到端目标检测模型(Real-Time Detection Transformer,RT-DETR)[14]算法,提出了一种融合双域特征的绝缘子缺陷检测算法。首先,设计空频特征融合(Frequency-Spatial Fusion,FreqSF)模块,通过融合空间域与频率域的特征,提高对纹理信息及模糊缺陷边缘特征的表达能力。其次,采用高低频注意力(High-Low Frequency Attention,HiLo)机制,替换尺度内特征交互模块中的多头自注意力(Multi-Head Self-Attention,MHSA)机制,提高网络捕捉绝缘子图像局部细节特征的能力。最后,引入三重特征编码器(Triple Feature Encoder,TFE)并结合尺度序列特征融合(Scale Sequence Feature Fusion,SSFF)模块,来融合不同尺度特征图中的深层语义信息与浅层细节信息。

## 1 融合双域特征的绝缘子缺陷检测算法

### 1.1 算法整体架构

本文算法的整体架构如图1所示,主要由主干网络、编码器和解码器三个部分组成。首先,输入图像经主干网络进行特征提取,该阶段引入FreqSF模块,通过融合空域与频域信息补充原始特征,对细小纹理与缺陷边缘的特征进行增强。随后,对提取出的特征输入编码器进行处理:在尺度内特征交互阶段,引入HiLo机制替换MHSA机制,以强化对局部细节特征的捕捉能力;在多尺度特征融合阶段,采用TFE与SSFF模块,将提取到的深层语义信息与浅层细节信息进行深度融合。最后,融合后的特征输入解码器,完成绝缘子缺陷的最终分类与定位。

### 1.2 FreqSF

RT-DETR主干网络主要依赖卷积操作进行特征提取,在复杂背景下对绝缘子缺陷边缘细节的表达能力不足,同时对全局上下文信息的建模能力有限,难以准确捕捉缺陷的细节特征。为此,本文设计了FreqSF模块,该模块通过两条并行支路,将空域的局部梯度响应与频域的全局频谱分布相结合。利用空域捕捉图像的细节与边缘信息,同时利用频域映射补充卷积在全局信息建模能力上的不足,从而提升模型对绝缘子缺陷边界的判别能力。

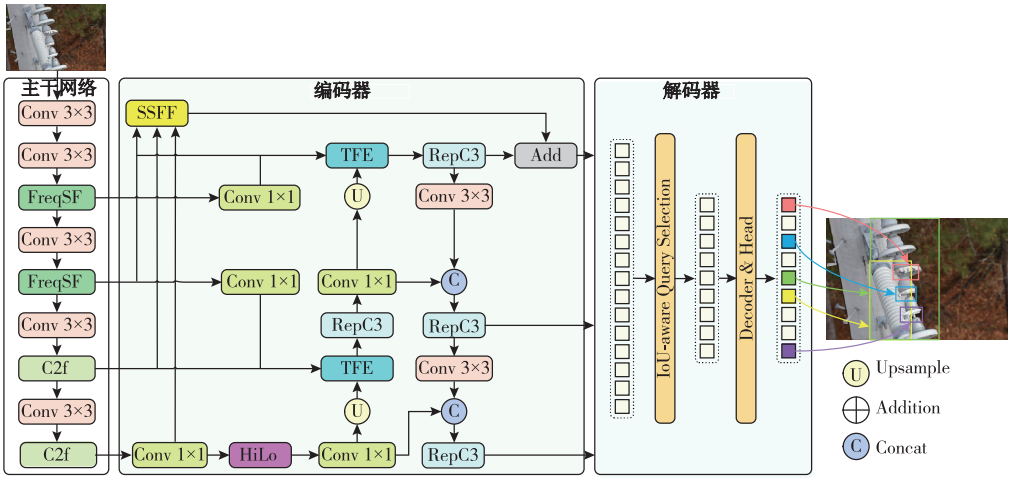



图 1　算法整体架构

FreqSF 模块基于 CSPNet[15] 架构思想设计，在每个阶段中将输入特征沿通道维度划分为两部分：一部分通过跨阶段连接传递至阶段末端，另一部分进入 FreqSF 进行特征提取，最终在阶段末端进行拼接融合，从而在控制参数量的同时提升了特征传递效率。由于浅层网络包含较多的高频细节特征，将该模块部署于主干网络前两阶段的高分辨率特征层，以强化细节信息提取，而在深层阶段采用轻量化 C2f 结构进行特征聚合，从而在控制计算复杂度的同时提升深层语义表达能力。如图 2 所示，FreqSF 主要由空域分支与频域分支构成。

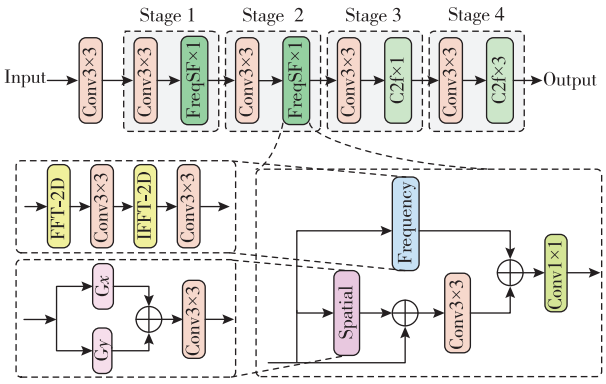



图 2　FreqSF 结构

在空域分支中，首先利用 Scharr 算子分别计算图像在水平与垂直方向的梯度，以提取图像中主要的结构特征，随后将水平与垂直方向的梯度响应图进行线性加权合并，生成初始的边缘特征图。该特征图经过标准卷积层增强后，通过残差连接与原始输入特征直接相加，可以防止细小的边缘特征在处理过程中丢失，从而有效突出绝缘子缺陷的局部空间特征。在频域分支中，首先对输入特征进行快速傅里叶变换（FFT-2D），将其从物理坐标系映射至复数频率空间，其离散正变换原理如式（1）所示。

$$F(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\mathrm{e}^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{1}$$

式中，$M$ 与 $N$ 分别表示特征图的高与宽，$(x,y)$ 为空间坐标，$(u,v)$ 为对应的频率坐标。经过频域映射后，将包含振幅与相位信息的实部与虚部沿通道维度拼接，并引入卷积算子对其进行全局频谱重构，从而跨越空间距离的限制，在全局视野下完成上下文特征的提取。随后，通过快速傅里叶逆变换（IFFT-2D）将处理后的频域分量转换至空间域，并使用卷积层来消除频域处理过程中可能产生的边缘噪点，其逆变换公式为

$$f(x,y)=\frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}F(u,v)\mathrm{e}^{j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{2}$$

最终，将频域分支输出与空域分支输出进行加权融合，并经过卷积层整合，生成模块的最终输出特征图，从而提升对绝缘子缺陷细节纹理特征的表达能力。

### 1.3　HiLo

绝缘子缺陷图像通常包含局部破损与裂纹等高频信息，而大范围的背景区域主要包含低频信息。然而，RT-DETR 中的尺度内特征交互（AIFI）模块采用标准多头自注意力机制，对所有图像块应用统一的全局注意力，未能区分特征中所包含的不同频率信息，从而在处理高分辨率图像时带来较大的计算开销。为此，本文在编码器中引入 HiLo[16] 对 AIFI 中 MHSA 进行替换，其结构如图 3

所示。设$X$为输入特征图，$N_h$表示当前层注意力头总数，$\alpha\in(0,1)$表示高低频注意力头的比例，其中$(1-\alpha)N_h$个头用于高频注意力分支，$\alpha N_h$个头用于低频注意力分支。

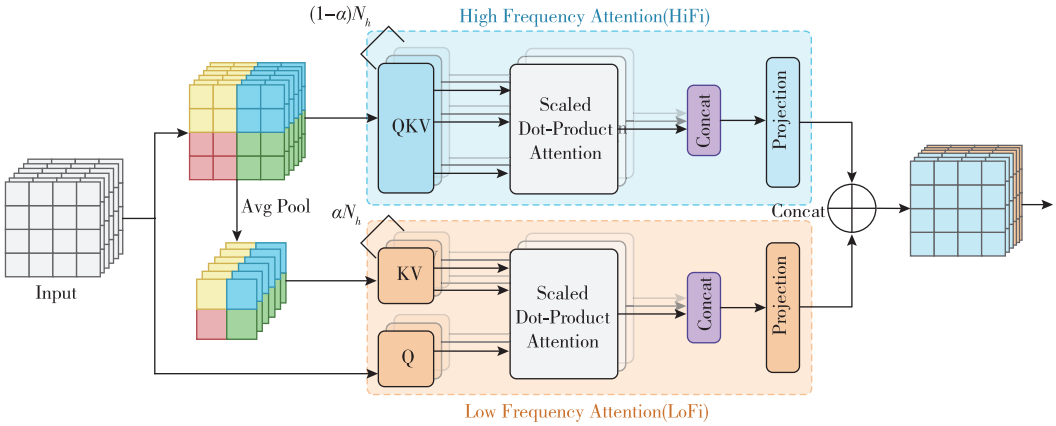



图3 HiLo注意力机制结构

高频注意力分支用于捕获特征图中的局部细节信息，采用局部窗口自注意力机制，将特征图划分为一系列非重叠的$2\times2$窗口，并在每个局部窗口内部执行标准自注意力运算，得到高频输出特征HiFi($X$)，从而在保持细节表达能力的同时提高计算效率。低频注意力分支旨在捕获输入特征图的整体结构信息，首先对输入特征$X$采用大小为$s\times s$的窗口进行平均池化，生成低分辨率的特征图$X_{\text{pool}}$，以滤除高频噪声并保留低频结构信息。随后将低频特征映射为键($K$)和值($V$)，而查询($Q$)仍由原始高分辨率特征$X$生成。其线性映射过程如式(3)与式(4)所示。

$$K=X_{\text{pool}}W_k,\ V=X_{\text{pool}}W_v \tag{3}$$

$$Q=XW_q \tag{4}$$

式中，$W_q$, $W_k$, $W_v$为可学习的投影权重矩阵。在此基础上，通过标准注意力机制完成全局信息交互，计算低频输出特征LoFi($X$)：

$$\text{LoFi}(X)=\text{Softmax}\left(\frac{QK^T}{\sqrt{D_h}}\right)V \tag{5}$$

式中，$D_h$为单个注意力头的特征维度。最后，将两分支提取到的特征在通道维度进行串联拼接，并通过线性投影层进行融合，其输出形式如式(6)所示。

$$\text{HiLo}(X)=\left[\text{HiFi}(X);\ \text{LoFi}(X)\right] \tag{6}$$

该机制将特征图中的高频与低频特征分开处理，在降低复杂度和计算成本的同时，提高了网络捕捉图像全局依赖关系和局部细节特征的能力。

## 1.4 多尺度特征融合结构改进

### 1.4.1 TFE

在输电线路巡检场景下，为提升复杂环境下细小绝缘子缺陷的识别能力，可以对比分析不同尺度下目标的形状与外观变化。由于主干网络不同特征层的尺度存在差异，传统的融合机制通常仅对小尺寸特征图进行上采样，并与更浅层的特征进行拼接或相加，从而忽略了大尺寸特征图中丰富的细节信息。为此，本文引入了TFE模块，对大、中、小三种尺寸的特征图进行并行处理，以有效地捕捉不同尺度下的缺陷特征。其结构如图4所示。

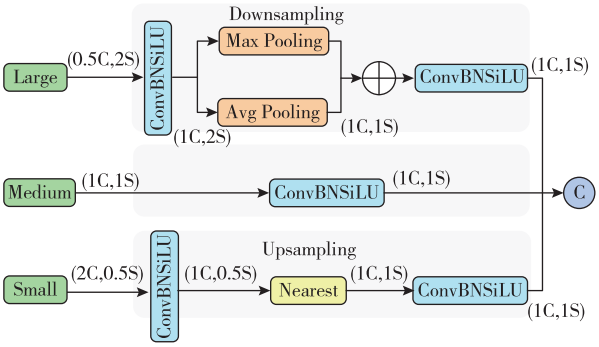



图4 TFE模块结构

在特征编码前，模块首先调整各输入特征的通道数，使其与中等尺寸特征图保持一致。大尺寸特征图在经过卷积模块处理后将其通道数调整为1C，然后采用最大池化与平均池化相结合的方式进行下采样，以降低特征的空间维度并保留显著的局部细节。小尺寸特征图在调整通道数后使用最近邻插值进行上采样，有助于保留局部特征

并防止小目标特征信息的丢失。最后，将对齐处理后的大、中、小三个尺寸的特征图进行卷积操作，然后在通道维度上进行拼接。其计算过程如式(7)所示。

$$F_{\mathrm{TFE}}=\mathrm{Concat}(F_{\mathrm{l}},F_{\mathrm{m}},F_{\mathrm{s}}) \tag{7}$$

式中，$F_{\mathrm{l}}$，$F_{\mathrm{m}}$，$F_{\mathrm{s}}$分别为经过处理后的大、中、小尺寸特征图，$F_{\mathrm{TFE}}$为 TFE模块的输出特征图。

1.4.2　SSFF

对于复杂背景下绝缘子图像的多尺度问题，现有研究通常采用特征金字塔结构进行特征融合，但其融合方式多局限于简单的元素相加或通道拼接，难以有效利用所有金字塔特征图之间的相关性。为此，本文引入了 SSFF 模块[17]，该模块可以将深层特征图的语义信息与浅层特征图的细节信息进行深度融合。SSFF 模块的输入计算公式为：

$$F_{\sigma}(w,h)=G_{\sigma}(w,h)\times f(w,h) \tag{8}$$

$$G_{\sigma}(w,h)=\frac{1}{2\pi\sigma^{2}}\mathrm{e}^{-(w^{2}+h^{2})/2\sigma^{2}} \tag{9}$$

式中，$f(w,h)$代表输入宽度为 $w$、高度为 $h$ 的二维特征图，$G_{\sigma}(w,h)$为二维高斯滤波器，$F_{\sigma}(w,h)$是使用二维高斯滤波器进行一系列卷积平滑生成的尺度空间。

SSFF 模块结构如图 5 所示。首先，通过两个独立的 1×1 卷积将 P4 和 P5 特征图的通道数统一调整为 256，并对其进行上采样，使其空间尺寸与 P3 特征图保持一致。随后，采用 Unsqueeze 操作为各层特征增加一个维度，并沿深度维度进行拼接。接着，通过 3D 卷积提取跨尺度特征之间的关联信息以增强特征表达，并结合批量归一化和 SiLU 激活函数进行非线性变换，然后引入 3D 最大池化压缩不同尺度的特征。最后，通过 Squeeze 操作去除冗余维度并输出融合特征。

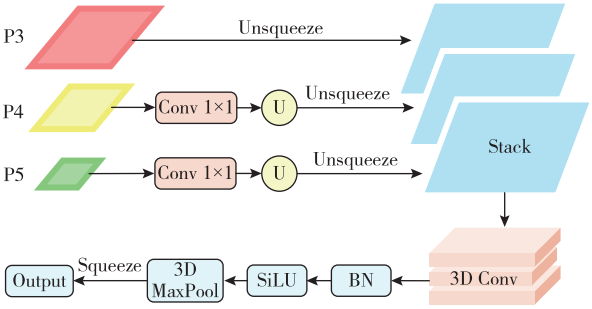



图 5　SSFF 模块结构

## 2　实验与结果分析

### 2.1　数据集

数据集来源于 IEEEDataPort 公开的绝缘子缺陷图像数据集(Insulator Defect Image Dataset，IDID)，共 1 600 张图片。由于原始数据集未提供具体缺陷位置标注信息，本文利用 LabelImg 工具对数据集中破损与闪络缺陷进行重新标注。为增强模型的泛化能力并防止过拟合，通过高斯模糊、随机旋转翻转、改变亮度与对比度等方式对数据集进行扩充，最终得到 3 840 张图像。按照 8∶1∶1 将数据集划分为训练集、验证集和测试集。破损与闪络缺陷数据集图片示例见图 6。



(a) 闪络缺陷　　(b) 破损缺陷

图 6　缺陷数据集图片示例

### 2.2　实验环境

实验在 Ubuntu22.04 操作系统下完成，实验平台配置为 Intel Xeon Silver 4214R CPU 和 NVIDIA RTX 3080 Ti GPU，显存为 12 GB，深度学习框架采用 PyTorch2.1.2，CUDA 版本为 11.8。训练过程中迭代轮数为 200，批量大小为 8，优化器采用 AdamW，学习率设置为 0.000 1，动量大小为 0.9，权重衰减系数为 0.000 1。

### 2.3　评价指标

为了评估模型的检测性能，采用精确率(Precision，P)、召回率(Recall，R)、平均精度均值(Mean Average Precision，mAP)、模型参数量(Params)和计算量(GFLOPs)作为模型性能的评价指标，mAP@0.5 表示当 IoU 阈值为 0.5 时的平均精度。计算公式分别为：

$$P=\frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}}+N_{\mathrm{FP}}} \tag{10}$$

$$R=\frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}}+N_{\mathrm{FN}}} \tag{11}$$

$$AP=\int_{0}^{1}P(R)\mathrm{d}R \tag{12}$$

$$mAP=\frac{1}{n}\sum_{i=1}^{n}AP \tag{13}$$

式中，$N_{TP}$为正样本被正确预测为正样本的数量，$N_{FP}$为负样本被错误预测为正样本的数量，$N_{FN}$为正样本被错误预测为负样本的数量，$AP$用于评估单个类别上的检测准确性，$mAP$为所有类别$AP$的平均值。

### 2.4 消融实验

由表1消融实验结果可知，在主干网络中引入FreqSF模块后，模型的mAP@0.5提升了0.8%，破损与闪络缺陷的AP值分别提升了2.7%与1.0%，同时参数量下降19.6%，这说明FreqSF模块通过融合频域与空域信息，有效增强了模型对缺陷细节特征的表达能力。在采用HiLo后，mAP@0.5达到0.962，破损及闪络缺陷的$AP$值分别提升了1.2%与0.9%。引入TFE+SSFF模块使mAP@0.5提高了0.7%，两类缺陷的$AP$值分别提升2.5%与1.1%，表明该模块有效增强了深层语义与浅层细节之间的信息交互能力。当同时使用三种改进策略时，模型综合性能达到最优，相较于基线RT-DETR，破损与闪络缺陷的AP值分别提升了5.0%与2.0%，mAP@0.5整体提升了1.5%。实验结果表明，所提算法能够克服复杂环境干扰，实现对绝缘子缺陷的高精度检测。

**表1 消融实验结果**

| 实验 | FreqSF | HiLo | TFE+SSFF | *AP* | | | mAP@0.5 | *Params*/M | *GFLOPs* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | insulator | broken | flashover | | | |
| 1 | | | | 0.992 | 0.911 | 0.958 | 0.959 | 19.9 | 56.9 |
| 2 | √ | | | 0.992 | 0.938 | 0.968 | 0.967 | 16.0 | 57.8 |
| 3 | | √ | | 0.992 | 0.923 | 0.967 | 0.962 | 19.7 | 57.1 |
| 4 | | | √ | 0.994 | 0.936 | 0.969 | 0.966 | 20.2 | 61.5 |
| 5 | √ | √ | | 0.994 | 0.942 | 0.974 | 0.971 | 16.2 | 59.4 |
| 6 | √ | | √ | 0.993 | 0.944 | 0.977 | 0.972 | 16.6 | 64.8 |
| 7 | √ | √ | √ | 0.995 | 0.961 | 0.978 | 0.974 | 16.4 | 63.5 |

### 2.5 对比实验

为确定FreqSF模块替换主干网络中C2f模块的最佳数量，在基线算法的主干网络中分别替换不同数量的C2f进行对比实验。从表2可知，仅替换1个C2f时参数量最低，但精度提升较为有限；相比于替换3个或4个模块，替换2个C2f在没有增加多少参数量和计算量的情况下检测精度提升最多。因此，选择替换主干网络中2个C2f模块，以在检测精度与计算效率之间取得较好平衡。

**表2 替换不同数量FreqSF对比实验结果**

| 模型 | *P* | mAP@0.5 | *Params*/M | *GFLOPs* |
|---|---|---|---|---|
| RT-DETR | 0.962 | 0.959 | 19.9 | 56.9 |
| RT-DETR+1FreqSF | 0.963 | 0.962 | **15.8** | **56.6** |
| RT-DETR+2FreqSF | **0.968** | **0.967** | 16.0 | 57.8 |
| RT-DETR+3FreqSF | 0.965 | 0.965 | 17.4 | 60.8 |
| RT-DETR+4FreqSF | 0.967 | 0.964 | 20.5 | 60.9 |

表3列出了本文算法与当前主流目标检测算法SSD[18]、Faster R-CNN[19]与YOLO系列[20]以及近年最新改进算法IDD-DETR[11]、AdIn-DETR[12]和PLD-DETR[13]在测试集上的对比试验结果。从表中可知，与经典检测算法SSD与Faster R-CNN相比，所提算法的mAP@0.5分别提升了7.3%和4.9%；相较于YOLO系列，在保持高精确率与召回率的同时，mAP@0.5仍有约1.3%的性能优势。在Transformer同类框架对比中，所提算法较Deformable-DETR的mAP@0.5与精确率分别提升3.0%与2.6%；与基线算法RT-DETR相比，mAP@0.5提高至0.974，精确率与召回率分别提升1.2%与1.1%。相较于近年提出的代表性改进算法，所提算法在模型参数量更低的同时，检测精度更高。综上所述，所提算法较好地平衡了检测精度与模型参数量，在复杂背景的缺陷检测任务中展现出更高的稳定性和鲁棒性。

**表3 对比实验结果**

| 模型 | *P* | *R* | mAP@0.5 | *Params*/M | *GFLOPs* |
|---|---|---|---|---|---|
| SSD | 0.913 | 0.884 | 0.901 | 26.3 | 62.0 |
| Faster R-CNN | 0.932 | 0.903 | 0.925 | 137.1 | 370.1 |
| YOLOv8m | 0.962 | 0.939 | 0.961 | 25.8 | 78.7 |
| YOLOv11m | 0.958 | 0.932 | 0.963 | 20.0 | 67.7 |
| Deformable-DETR | 0.948 | 0.923 | 0.944 | 40.1 | 123.0 |
| PLD-DETR | 0.965 | 0.935 | 0.967 | 21.1 | 59.4 |
| IDD-DETR | 0.970 | 0.938 | 0.969 | 21.6 | 65.7 |
| AdIn-DETR | 0.972 | 0.939 | 0.971 | 44.0 | 144.0 |
| RT-DETR | 0.962 | 0.930 | 0.959 | 19.9 | **56.9** |
| 本文 | **0.974** | **0.941** | **0.974** | **16.4** | 63.5 |

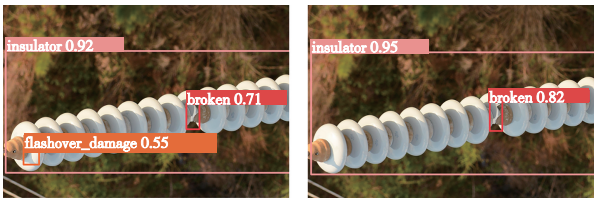
### 2.6 可视化分析

为进一步验证所提算法的实际检测效果，从测试集中选取了四组不同场景的图像进行可视化对比，结果如图7所示。其中图(a)～(d)分别展示了

RT-DETR（左侧）与所提算法（右侧）在冰雪覆盖、复杂植被、运动模糊和结构遮挡四种场景下的检测结果。从图中可以发现，RT-DETR容易受到复杂环境干扰，在冰雪覆盖、复杂植被与结构遮挡场景下均产生了误检，并在运动模糊场景下发生了漏检。相比之下，所提算法能够有效克服上述环境干扰，准确检测出绝缘子缺陷，且整体检测置信度更高。




（a）冰雪覆盖场景




（b）复杂植被场景




（c）运动模糊场景




（d）结构遮挡场景

图 7　检测结果可视化对比

## 3　结论

本文针对无人机航拍输电线路巡检中复杂环境干扰导致绝缘子缺陷检测精度不高以及误检漏检的问题，提出了一种基于RT-DETR且融合双域特征的缺陷检测算法。该算法通过引入FreqSF模块、HiLo注意力机制，并结合多尺度特征融合，提高了局部细节与边缘纹理的表征能力。实验结果表明，所提算法较基线RT-DETR在mAP@0.5、精确率及召回率上分别提高了1.5%、1.2%和1.1%，改善了复杂场景下的误检与漏检问题，具备较高的工程应用潜力。未来将进一步丰富数据集，并探索所提算法在无人机边缘计算平台上的轻量化部署，从而实现高效的智能在线巡检。

**参考文献：**

[1] 刘传洋，吴一全．基于深度学习的输电线路视觉检测方法研究进展[J]．中国电机工程学报，2023，43(19)：7423-7445.
Liu Chuanyang，Wu Yiquan．Research progress of vision detection methods based on deep learning for transmission lines [J]．Proceedings of the CSEE，2023，43(19)：7423-7445.

[2] 刘传洋，吴一全，刘景景．无人机航拍图像中绝缘子缺陷检测的深度学习方法研究进展[J]．电工技术学报，2025，40(9)：2897-2916.
Liu Chuanyang，Wu Yiquan，Liu Jingjing．Research progress of deep learning methods for insulator defect detection in UAV based aerial images[J]．Transactions of China Electrotechnical Society，2025，40(9)：2897-2916.

[3] 孙　虹，郭义菡，雷经发，等．基于改进YOLOv7的多尺度特征提取绝缘子缺陷检测算法[J]．高压电器，2025，61(6)：102-112.
Sun Hong，Guo Yihan，Lei Jingfa，et al．Insulator defect detection algorithm for multi-scale feature extraction based on improved YOLOv7 [J]．High Voltage Apparatus，2025，61(6)：102-112.

[4] 王　艳，钟昕圻，王子剑，等．多尺度特征提取与自适应交互融合的绝缘子及其缺陷检测算法[J]．电力系统及其自动化学报，2026，38(4)：147-158.
Wang Yan，Zhong Xinqi，Wang Zijian，et al．Insulator and defect detection algorithm based on multi-scale feature extraction and adaptive interactive fusion [J]．Proceedings of the CSU-EPSA，2026，38(4)：147-158.

[5] 曾业战，陈天航，邓　倩，等．基于多尺度特征提取与密集连接的绝缘子缺陷检测算法[J]．计算机工程与应用，2026，62(16)：368-378.
Zeng Yezhan，Chen Tianhang，Deng Qian，et al．Insulator defect detection based on multi-scale feature extraction and dense connection[J]．Computer Engineering and Applications，2026，62(16)：368-378.

[6] 刘宏志，马　跃，邱　彬，等．改进YOLOv11n在输电线路绝缘子主要缺陷检测中的应用研究[J]．高压电器，2025，61(10)：149-158.
Liu Hongzhi，Ma Yue，Qiu Bin，et al．Research on application of improved YOLOv11n in detection of main defects of insulators on transmission lines[J]．High Voltage Apparatus，2025，61(10)：149-158.

[7] 彭晏飞，袁晓龙，赵　涛，等．基于改进YOLOv5s的绝缘子缺陷检测方法[J]．高压电器，2026，62(1)：134-142.
Peng Yanfei，Yuan Xiaolong，Zhao Tao，et al．Detection

method for insulator defect based on improved YOLOv5s[J]. High Voltage Apparatus, 2026, 62(1): 134-142.

[8] Tao Z, He Y, Lin S, et al. SnakeNet: an adaptive network for small object and complex background for insulator surface defect detection[J]. Computers and Electrical Engineering, 2024, 117: 109259.

[9] Wu Y, Mu X, Shi H, et al. An object detection model AAPW-YOLO for UAV remote sensing images based on adaptive convolution and reconstructed feature fusion[J]. Scientific Reports, 2025, 15: 16214.

[10] 张靖雯, 孙 坚, 徐红伟, 等. 基于改进RT-DETR的玻璃绝缘子缺陷检测算法[J]. 电子测量技术, 2025, 48(14): 96-105.
Zhang Jingwen, Sun Jian, Xu Hongwei, et al. Defect detection algorithm for glass insulators based on improved RT-DETR[J]. Electronic Measurement Technology, 2025, 48(14): 96-105.

[11] Chen W, Li S, Han X. IDD-DETR: insulator defect detection model and low-carbon operation and maintenance application based on bidirectional cross-scale fusion and dynamic histogram attention[J]. Sensors, 2025, 25(18): 5848.

[12] Cheng Y, Liu D. AdIn-DETR: adapting detection transformer for end-to-end real-time power line insulator defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 3528511.

[13] Chen J, Zhang X, Feng D, et al. PLD-DETR: a method for defect inspection of power transmission lines[J]. Electronics, 2025, 14(20): 4107.

[14] Zhao Y, Lü W, Xu S, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

[15] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 1571-1580.

[16] Pan Z, Cai J, Zhuang B. Fast vision transformers with HiLo attention[C]//Advances in Neural Information Processing Systems 35, 2022: 14541-14554.

[17] Kang M, Ting C M, Ting F F, et al. ASF-YOLO: a novel YOLO model with attentional scale sequence fusion for cell instance segmentation[J]. Image and Vision Computing, 2024, 147: 105057.

[18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//Computer Vision – ECCV 2016, 2016: 21-37.

[19] 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 and Machine Intelligence, 2017, 39(6): 1137-1149.

[20] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

**作者简介：**

**党宏社**(1962—)，男，陕西省咸阳市人，博士，教授，主要研究方向为数字图像处理、模式识别；

**张代言**(2002—)，男，陕西省宝鸡市人，硕士研究生，主要研究方向为深度学习、图像处理。