

# 基于多尺度特征增强的无人机多模态电线检测\*

田洪坤<sup>1</sup> 姜 因<sup>1</sup> 李溯源<sup>1</sup> 任 涛<sup>2</sup>

(1. 中国刑事警察学院公安信息技术与情报学院 沈阳 110854; 2. 东北大学软件学院 沈阳 110819)

**摘要:**针对无人机电线检测任务中面临的复杂背景干扰、电线尺度随机分布且差异大的挑战,提出一种基于多尺度特征增强的算法框架。该框架以多模态特征融合输入为基础,集成了3个核心模块:多尺度特征增强模块(multi-scale feature enhance, MSFE)用于捕捉不同尺度的电线特征;双路径交互注意力模块(dual-path interactive attention, DPIA)通过通道与空间双路径机制聚焦电线关键区域;自适应特征融合模块(adaptive feature fusion, AFF)动态平衡编码器语义信息与解码器边缘信息,旨在提升复杂场景下的检测鲁棒性。实验结果表明,所提方法在雾天、雪天、夜间等多种复杂场景中均表现优异。相较于现有方法,各项指标均取得了显著提升。消融实验充分验证了各模块的有效性;在基线模型基础上, IoU、Robj、Recall、Precision 和 F1-Score 分别提升了 12.56%、5.53%、4.54%、3.03% 和 3.78%, 表明其对增强模型检测能力至关重要。定性分析结果进一步证实了该算法在复杂场景下对电线细长结构的精准定位能力及其对背景噪声的有效抑制。因此,该方法对无人机电线检测任务具有实际应用价值。

**关键词:**多尺度特征;无人机;电线检测;注意力机制;特征融合

**中图分类号:** TP391.4; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 470

## UAV multi-modal transmission line detection based on multi-scale feature enhancement

Tian Hongkun<sup>1</sup> Jiang Nan<sup>1</sup> Li Suyuan<sup>1</sup> Ren Tao<sup>2</sup>

(1. School of Public Security Information Technology and Intelligence, Criminal Investigation Police University of China, Shenyang 110854, China; 2. School of Software, Northeastern University, Shenyang 110819, China)

**Abstract:** Aiming at the challenges of complex background interference, random distribution and large differences in transmission line scales in UAV transmission line detection tasks, this paper proposes an algorithm framework based on multi-scale feature enhancement. The framework is based on multi-modal feature fusion input and integrates three core modules: the multi-scale feature enhancement module (multi-scale feature enhance, MSFE) is used to capture transmission line features of different scales; the dual-path interactive attention module (dual-path interactive attention, DPIA) focuses on key transmission line regions through a dual-path mechanism of channel and space; the adaptive feature fusion module (adaptive feature fusion, AFF) dynamically balances the semantic information of the encoder and the edge information of the decoder, aiming to improve the detection robustness in complex scenarios. Experimental results show that the proposed method performs excellently in various complex scenarios such as foggy days, snowy days, and low light. Compared with existing methods, all indicators have achieved significant improvements. Ablation experiments fully verify the effectiveness of each module: on the basis of the baseline model, IoU, Robj, Recall, Precision, and F1-Score are increased by 12.56%, 5.53%, 4.54%, 3.03%, and 3.78% respectively, indicating that they are crucial for enhancing the detection capability of the model. Qualitative analysis results further confirm that the algorithm can accurately locate the slender structure of transmission lines in complex scenarios and effectively suppress background noise. Therefore, the method in this paper has practical application value for UAV transmission line detection tasks.

**Keywords:** multi-scale features; UAV; transmission line detection; attention mechanism; feature fusion

## 0 引言

无人机对电线的精准感知是多领域任务的关键支撑,一方面,在电力系统运维中,可实现对电线老化、腐蚀等缺陷的快速巡检,显著降低人工成本与作业风险<sup>[1]</sup>;另一方面,在自主导航与避障中,实时电线检测是保障无人机在复杂空域安全作业的前提<sup>[2-4]</sup>,尤其在军警用侦察、巡逻等场景,可有效规避电线碰撞风险,保护无人机和电力设施<sup>[5]</sup>。然而,光照突变、雾气干扰、复杂背景干扰等环境的复杂性,电线细长形态、尺度差异、随机分布等自身特性,使无人机在各类场景下的电线检测仍面临精度不足、鲁棒性欠缺等挑战,亟需更有效的检测方法。

早期无人机电线检测研究以单模态视觉数据为核心,聚焦于可见光数据。Song 等<sup>[6]</sup>利用改进的 Radon 变换提取电线线性特征,Liao 等<sup>[7]</sup>结合局部特征与空间顺序实现电线定位,但这类传统的特征提取策略均受限于复杂未知背景的干扰,在树木、建筑等复杂环境中鲁棒性不足。随着深度学习技术发展<sup>[8]</sup>,基于卷积神经网络(CNN)的方法逐步成为主流;Gao 等<sup>[9]</sup>提出并行分支网络提取多级特征;Song 等<sup>[10]</sup>通过优化特征提取流程增强检测鲁棒性。然而,这些方法仍未能突破可见光成像固有的光照敏感瓶颈,即在夜间、大雾等复杂光照场景下,其纹理细节模糊导致漏检率显著上升。

红外模态数据<sup>[11]</sup>凭借对光照变化的稳定性展现出独特优势。Li 等<sup>[12]</sup>利用红外图像的温度分布特征定位电线,有效规避了光照干扰;Nguyen 等<sup>[13]</sup>对红外纹理分析,实现低光照环境下目标检测。然而,红外图像的分辨率通常较低,且语义信息匮乏,难以捕捉电线的细微形态,仅依靠单一红外数据仍无法应对复杂场景下的检测需求。

将可见光与红外等多模态数据融合使用,成为突破单模态局限的重要方向。Choi 等<sup>[14]</sup>基于可见光与红外多模态数据,构建了涵盖昼夜、雾、雪等场景的无人机电线检测数据集,并设计了基于多分支特征提取与通道注意力的融合网络,系统验证了多模态数据在电线检测中的优势,为后续研究提供了数据支撑与方法参考。尽管如此,该方法在面对如铁路轨道、山地地形等复杂背景干扰及电线细长形态、尺度差异等自身特性挑战时,仍存在鲁棒性不足的问题。现有研究中,虽有依托多模态数据实现绝缘子、输电塔等线路部件巡检的成功案例<sup>[15-16]</sup>,但这些方法针对特定部件设计,难以直接迁移用于电线检测这一更精细的任务。此外,一些目标检测、图像分割方法<sup>[17-19]</sup>也无法直接适用于电线检测任务,仍难以解决多尺度特征融合不充分、注意力机制低效等问题。

为了解决这些问题,本文提出一种基于多尺度特征

增强的无人机多模态电线检测方法,创新性提出多尺度特征增强策略,适配电线目标在复杂背景下从局部边缘到全局形态的多尺度特性,增强检测的鲁棒性;提出通道空间双路径交互注意力机制,精准聚焦电线特征,抑制复杂背景,为无人机电线检测任务提供关键的语义与空间指导;提出自适应调整编码特征和解码特征的权重策略,有效平衡两者之间的关系,提高特征语义与边缘的交互效果。

## 1 基于多尺度特征增强的无人机多模态电线检测算法

为有效应对无人机电线检测时的复杂背景干扰、电线分布随机且存在尺度差异等导致检测精度不足,而现有方法对多尺度特征融合不充分、注意力机制针对性弱等问题,提出基于多尺度特征增强的无人机电线检测框架,通过强化多尺度特征表达与目标聚焦能力,提升在各类场景下的电线检测效果。

### 1.1 算法框架

基于多尺度特征增强的无人机电线检测算法,融合了多模态输入、编码阶段的多尺度特征增强、注意力机制以及解码阶段的特征自适应融合,实现复杂场景下电线目标的精准检测。算法整体结构分为编码器、瓶颈层及解码器 3 个阶段,框架图如图 1 所示。

编码阶段,输入层采用多模态数据融合策略,将无人机采集的可见光图像与红外图像拼接为 4 通道输入,通过残差双卷积模块(residual double conv, RDC)提取特征,为后续特征学习提供基础。基于此,编码器由三级下采样模块构成,每级通过最大池化与 RDC 组合实现特征提取与降维。其中,RDC 模块通过两层  $3 \times 3$  卷积、批归一化(BN)、ReLU 激活及残差连接,有效缓解深层网络梯度消失问题。编码器输出不同尺度特征,为解码阶段提供多分辨率跳跃连接基础。为了强化多尺度特征表征能力,在编码阶段针对不同层级特征设计多尺度特征增强模块(multi-scale feature enhance, MSFE),形成多尺度互补的编码特征集。

瓶颈阶段作为全局特征枢纽,位于编码与解码的衔接处,聚焦特征优化,集成了 MSFE 与双路径交互注意力机制(dual-path interactive attention, DPIA)。MSFE 融合多尺度上下文信息;DPIA 通过通道与空间双维度注意力机制对全局语义特征进行筛选,聚焦电线目标区域特征,确保解码阶段优先传递目标相关信息。

解码阶段,以自适应特征融合(adaptive feature fusion, AFF)为核心,通过三级上采样与精细特征整合逐步恢复空间分辨率。每级上采样模块先通过双线性插值提升上层特征分辨率,再通过 AFF 动态调整编码特征和

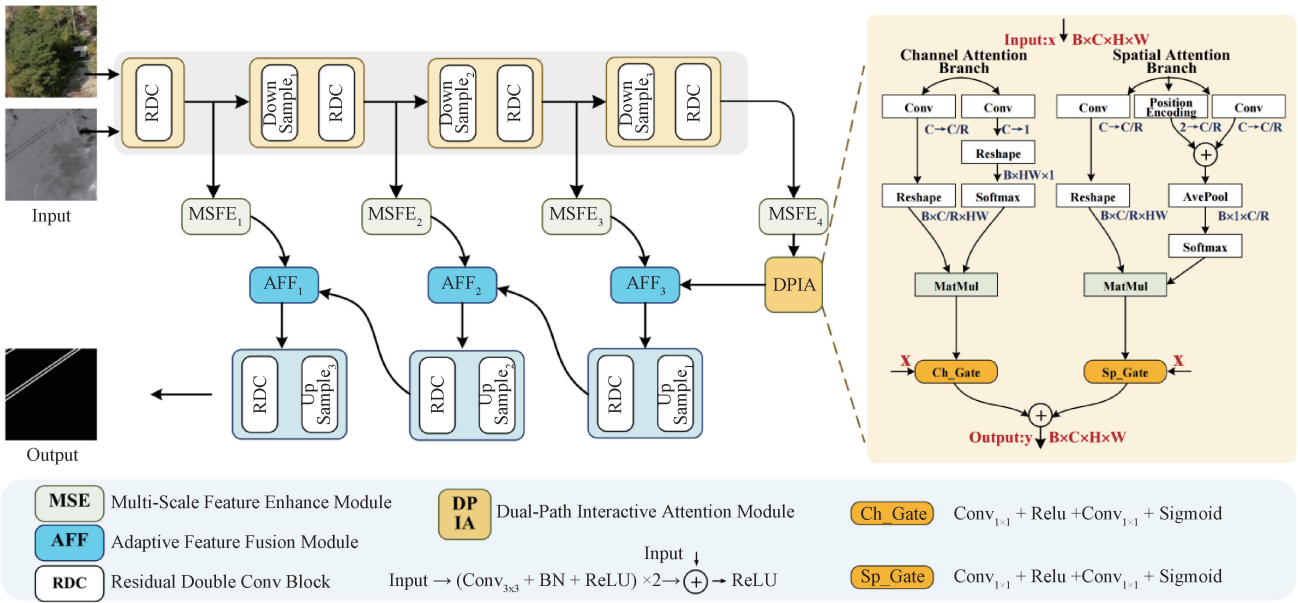


图 1 多尺度特征增强无人机电线检测网络

Fig. 1 Framework of multi-scale feature enhancement UAV power line detection network

解码特征两者权重,解决电线检测中“语义模糊”与“边缘丢失”的矛盾。三级解码过程形成递进式融合链,每级融合后均通过 RDC 进一步提取特征,抑制融合过程中可能引入的噪声。最后,经输出卷积层将特征图映射为二值化检测结果,完成电线目标像素级精准定位。

算法通过二元交叉熵损失 (binary cross entropy loss, BCE) 和 Dice 损失加权优化网络参数。

BCE 损失用于衡量每个像素预测和真实标签间差异,如式(1)所示。

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (1)$$

式中:  $y_i$  是第  $i$  个像素的真实标签;  $p_i$  是第  $i$  个像素的预测概率;  $N$  是总像素数。

Dice 损失用于预测结果和真实标签之间的区域相似度度量,如式(2)所示。

$$L_{Dice} = 1 - \frac{2 \cdot \sum_{i=1}^N p_i y_i + \epsilon}{\sum_{i=1}^N p_i + \sum_{i=1}^N y_i + \epsilon} \quad (2)$$

总损失是 BCE 损失和 Dice 损失的加权和,  $\alpha$  在本算法中设置为 0.5。

$$L_{total} = \alpha \cdot L_{BCE} + (1 - \alpha) \cdot L_{Dice} \quad (3)$$

### 1.2 MSFE 模块

MSFE 的设计受空洞空间金字塔池化 (atrous spatial pyramid pooling, ASPP)<sup>[20]</sup> 经典多尺度结构启发,核心围绕无人机电线检测的任务特性进行优化。与 ASPP 不同的是,MSFE 在残差连接、层级适配、特征融合 3 个核心环

节进行了针对性创新,这些设计使其不仅具备多尺度特征提取能力,还能针对性匹配电线细长形态、随机尺度分布的特性,更能精准解决电线检测中的“细节丢失”“全局关联性不足”等问题,适配无人机视角下的精准感知需求,其结构如图 2 所示。

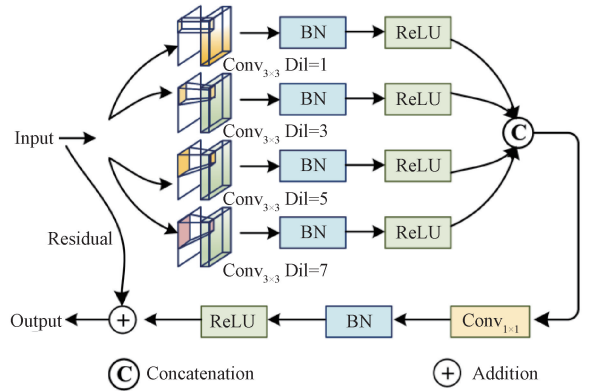


图 2 多尺度特征增强模块

Fig. 2 Multi-scale feature enhance block

MSFE 具体由 4 条并行空洞卷积分支、特征融合单元及残差连接构成,核心设计围绕电线检测任务特性优化:每条分支采用 3×3 卷积,空洞率设置为 1、3、5、7 的连续递增配置,既覆盖从局部到全局的感受野范围,又避免了传统结构空洞率间隔过大导致的尺度适配盲区,配合 BN 与 ReLU 激活提升特征表达能力。特征融合单元先将多分支结果沿通道维度拼接,再通过 1×1 卷积统一通道数并经 BN+ReLU 激活,有效缓解多模态数据的分布差异;

残差连接则通过  $1 \times 1$  卷积实现输入特征与融合特征的通道对齐后相加,避免直接融合导致的特征冲突。该结构采用分层动态调整策略:浅层启用小空洞率分支保留电线边缘、断点等细节,中层用中等空洞率平衡局部特征与上下文关联,深层启用全部分支捕捉长距离电线的全局连续性,实现多尺度特征的精准捕获与增强,如式(4)所示。

$$\begin{cases} F_{MSFE} = \text{ReLU}[BN(\text{Conv}_{1 \times 1}(C)) + \text{Residual}(X)] \\ C = [F_1, F_2, \dots, F_N] \\ F_k = \text{ReLU}(BN(\text{Conv}_{3 \times 3, d_k}(X))) \end{cases} \quad (4)$$

式中:  $F_{MSFE}$  为 MSFE 模块的输出特征;  $C$  为多分支特征经通道拼接后的聚合特征;  $F_k$  为第  $k$  条并行分支的输出特征;  $d_k$  为该分支的空洞率;  $\text{Residual}(X)$  为残差连接。

### 1.3 DPIA 模块

DPIA 是针对无人机电线检测任务设计的注意力模块,由通道注意力路径与空间注意力路径并行构成。其中,通道路径源自 Polarized Self-Attention<sup>[21]</sup>,负责捕捉通道间的特征依赖;空间路径则在其基础上改进,通过引入位置编码与精细化权重生成机制,解决原有空间注意力对细长目标定位模糊的问题,DPIA 结构如图 1 所示。

通道路径直接沿用 Polarized Self-Attention 通道注意力结构,通过均值、归一化等消除多模态特征的分布偏差,再通过 Softmax 与矩阵乘法学习通道间的依赖,为解码阶段提供可靠语义依据。其核心式为:

$$\alpha_c = \text{Sigmoid}\left(\text{Conv}_{1 \times 1}\left(\text{ReLU}\left(\text{Conv}_{1 \times 1}\left(\frac{F_{cv} - \mu}{\sigma + \epsilon} \cdot \text{Softmax}(F_{cq})\right)\right)\right)\right) \quad (5)$$

其中,  $F_{cv} = \text{Conv}_{1 \times 1}(X)$  为降维后的通道价值特征 ( $C \rightarrow C/R, R$  为缩减率),  $F_{cq} = \text{Conv}_{1 \times 1}(X)$  为通道查询特征 ( $C \rightarrow 1$ ),  $\mu$  和  $\sigma$  分别为  $F_{cv}$  的均值和标准差,  $\cdot$  表示矩阵乘法。

通道注意力路径通过 4 步操作实现对特征通道的选择性增强:首先用两个  $1 \times 1$  卷积将输入特征降维为通道价值特征  $F_{cv}$  和查询特征  $F_{cq}$ ,压缩通道以减少计算量;接着对  $F_{cv}$  进行通道维度的均值与标准差归一化,消除多模态数据的通道分布差异;再将归一化后的  $F_{cv}$  与 Softmax 函数激活的  $F_{cq}$  矩阵相乘,计算通道间相关性,突出与电线检测相关的特征通道;最后通过  $1 \times 1$  卷积升维  $\rightarrow$  ReLU 激活  $\rightarrow 1 \times 1$  卷积降维  $\rightarrow$  Sigmoid 的非线性映射生成最终通道权重。

对于无人机多模态电线检测任务, Polarized Self-Attention 的空间路径存在两点不足,一是位置编码不足,难以定位电线的细长连续结构,容易将背景纹理误判为

电线;二是权重生成较简单,直接通过矩阵乘法输出,对复杂场景及天气的干扰抑制能力弱。为此,本算法设计了改进空间注意力路径,核心式为:

$$\alpha_s = \text{Sigmoid}\left(\text{Conv}_{1 \times 1}\left(\text{ReLU}\left(\text{Conv}_{1 \times 1}(F_{sv} \cdot \text{Softmax}(F_{sq} + P))\right)\right)\right) \quad (6)$$

其中,  $F_{sv} = \text{Conv}_{1 \times 1}(X)$  为空间价值特征 ( $C \rightarrow C/R, R$  为缩减率),  $F_{sq} = \text{Conv}_{1 \times 1}(X)$  为空间查询特征 ( $C \rightarrow C/R, R$  为缩减率),  $P$  为通过拼接高、宽方向线性坐标向量并卷积生成的位置编码,  $P = \text{Conv}_{1 \times 1}(\text{cat}(h_{pos}, w_{pos}))$ 。

相比 Polarized Self-Attention 的空间注意力路径,本文提出的 DPIA 空间注意力路径,引入了线性位置编码,并精细化特征权重的生成,具体通过 4 步协同操作精准定位电线空间分布:首先,用两个  $1 \times 1$  卷积将输入特征降维为空间价值特征  $F_{sv}$  和查询特征  $F_{sq}$ ;其次,引入沿高、宽维度的线性位置编码,经卷积转换后与  $F_{sq}$  融合,注入坐标信息锁定电线连续细长结构,解决传统注意力位置感知缺失的问题,关键的是,融合位置信息的  $F_{sq}$  会经自适应平均池化压缩为全局向量,通过聚合空间全局信息强化与  $F_{sv}$  的相关性计算,既过滤局部噪声,又突出电线整体形态的一致性;随后将展平空间维度的  $F_{sv}$  与池化后的  $F_{sq}$  做矩阵乘法,经 Softmax 函数得到初步权重;最后通过  $1 \times 1$  卷积升维  $\rightarrow$  ReLU  $\rightarrow 1 \times 1$  卷积降维  $\rightarrow$  Sigmoid 构成的权重计算网络生成空间特征权重,实现对复杂背景抑制。

通道路径、空间路径双路径特征,通过逐元素相加并残差连接实现交互整合,输出最终增强特征,其核心公式为:

$$F_{\text{DPIA}} = X + \alpha_c \cdot X + \alpha_s \cdot X \quad (7)$$

将 DPIA 部署于整个网络瓶颈层的核心优势在于:瓶颈层作为编码器的最高语义抽象层,特征图语义信息最丰富,在此处部署应用 DPIA,可通过通道、空间双路径协同交互,精准聚焦电线特征,并抑制复杂背景,为无人机电线检测任务提供关键的语义与空间指导。

### 1.4 AFF 模块

在电线检测任务中,“语义模糊”与“边缘丢失”是两个关键且相互矛盾的问题。编码特征包含丰富的语义信息,但在定位细节边缘时能力较弱;解码特征在边缘细节定位上表现出色,却可能缺乏足够的语义信息。本文构建的 AFF 模块旨在动态调整编码特征和解码特征的权重,有效平衡两者之间的关系,从而解决上述矛盾,提升电线检测的准确性和可靠性,如图 3 所示。

AFF 模块通过引入可学习的权重,实现对编码特征和解码特征的自适应融合。首先,使用  $1 \times 1$  卷积分别调整编码特征  $f_1'$  和解码特征  $f_2'$  的通道数,使其达到相同的输出通道数。然后,将解码特征  $f_2'$  上采样至与编码特

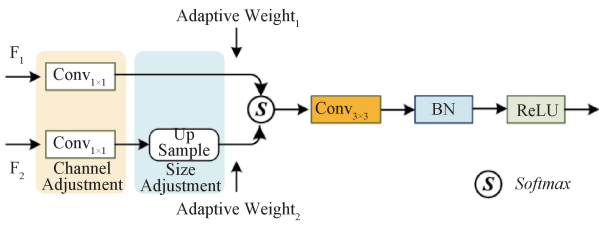


图3 自适应特征融合模块

Fig. 3 Adaptive feature fusion block

征  $f_1'$  相同的尺寸。接下来,利用 Softmax 函数对可学习的权重  $w_1$  和  $w_2$  进行归一化处理,确保两者之和为 1。最后,根据归一化后的权重对调整后特征进行加权求和,得到融合后的特征  $F_{\text{AFF}}$ 。核心公式如下:

$$F_{\text{AFF}} = \frac{w_1}{\sum_{i=1}^2 w_i} \cdot f_1' + \frac{w_2}{\sum_{i=1}^2 w_i} \cdot f_2' \quad (8)$$

式中:  $f_1'$  和解码特征  $f_2'$  分别是调整通道数后的编码特征和解码特征。通过这种方式, AFF 能根据不同的输入动态调整特征的权重,充分发挥编码和解码特征的优势。

## 2 实验与结果分析

### 2.1 实验环境

本文实验环境配置如表 1 所示。

表 1 实验环境配置

Table 1 Experiment environment

Experiment environment	Version
Operating System	Ubuntu 18.04
CPU	Intel Core i7-10700
GPU	NVIDIA GeForce RTX 3080
Deep learning framework	Pytorch 1.7.1, Cuda 11.0
Programming language	Python 3.7

### 2.2 数据集

实验数据集主要采用 VITLD 数据集<sup>[14]</sup>,包含 420 对可见光、红外图像及对应的像素级标注真值图像,分辨率为  $256 \times 256$ ,用于无人机输电线路检测。数据集初始划分为训练集(280 对)、验证集(60 对)、测试集(80 对)。为更好验证算法在复杂场景中的检测能力,参照前期相关研究设定,在原始测试集的可见光图像基础上人工扩充了白天、雾天、夜间及雪天等多种场景,形成共 400 对测试数据。训练阶段对数据进行丰富的增强,包括对输入图像旋转、随机翻转、边界裁剪,以及对图像的颜色、雾面、雪景等进行调整,以适配不同场景条件。

### 2.3 网络参数设置及评价指标

网络训练过程中使用 RAdam 优化器,初始学习率为 0.001,权重衰减 0.0005,搭配 StepLR 学习率调度器,每

20 个 epoch 衰减至 0.5 倍。批量大小设为 4,总训练轮次 200 轮。

参照前期相关研究的设定,本研究采用的评价指标主要包括交并比(intersection over union, IoU):用于衡量预测结果与真实标签的重叠程度,通过计算预测区域与目标区域交集和并集之比评估检测精度,直接反映目标定位的准确性;目标识别率(recognition rate of objects, Robj)通过正确检测到的目标数量与真实目标总数之比计算得出,用于评估模型对目标的整体识别能力,电线检测任务中,Robj 侧重衡量模型对电线这类细长目标的完整识别率,其值越高,表明目标漏检率越低,模型能更全面地捕捉目标存在性。为更全面地评估算法的性能,实验中还补充了召回率(recall)、精确率(precision)及 F1-Score 等指标,从多维度量化模型的检测效果。

### 2.4 比较实验

本文对比了多项相关算法,包括 Choi 等<sup>[14]</sup>提出的基于 UNet 的无人机电线检测算法、基于 UMFNet 多模态特征融合的无人机电线检测算法、Zhou 等<sup>[25]</sup>提出的基于知识蒸馏的无人机电线检测算法 BGRNet、Zhang 等<sup>[26]</sup>提出的基于多模态特征增强的无人机电线检测算法 HMMEN、Qin 等<sup>[17]</sup>提出的显著目标检测代表性算法 U2-Net、Gao 等<sup>[18]</sup>和 Zhou 等<sup>[19]</sup>提出的图像分割代表性算法 UTNet、UNet++等;此外,为了更全面的比较,实验中,还在 Choi 等算法基础上,部署了基于 Woo 等<sup>[22]</sup>提出的 CABM、Mehta 等<sup>[23]</sup>提出的 MobileViT 及 Jannu 等<sup>[24]</sup>提出的 Shuffle Attention 的改进算法。

表 2 为比较实验结果,本文方法总体表现出较强的竞争力,优于其他对比方法。在关键指标 IoU 上,本文方法达到 66.30%,仅略低于最优方法 HMMEN 的 66.60% (差值 0.3%);相较于 BGRNet 的 62.30%、Shuffle Attention 的 60.60%等主流方法同样具备显著优势,充分体现其在电线目标区域定位精度上的突出表现。在反映目标区域整体捕捉能力的 Robj 指标上,本文方法取得 80.99% 的成绩:既与当前该指标表现最优的 CBAM 的 80.41% 接近且小幅领先,更显著高于 Shuffle Attention 的 76.99% 等其他对比方法,表明其对目标区域的整体覆盖能力处于领先水平。在 Recall 指标上,本文方法的 75.95% 较 CBAM 的 75.54% 提升 0.41%,较 UNet++ 的 73.52% 提升 2.43%,这意味着其对真实电线目标的识别更全面,漏检率更低;在 Precision 指标上,本文方法的 74.88% 略低于 UNet++ 的 75.53%,但略高于 Shuffle Attention 的 74.85%,整体保持了较高的识别精准度。F1-Score 是 Recall 与 Precision 的综合衡量指标,本文方法的 F1-Score 为 75.41%,较 Shuffle Attention 的 74.13% 提升 1.28%、较 CBAM 的 73.79% 提升 1.62%,说明其在“识别全面性”与“识别准确性”之间实现了更优的平衡。

表 2 对比实验结果  
Table 2 Comparative experimental results

Methods	IoU/%	Robj/%	Recall/%	Precision/%	F1-Score/%	FPS
UNet <sup>[14]</sup>	52.92	74.31	73.22	70.96	72.07	68.03
UMFNet <sup>[14]</sup>	57.07	73.98	72.12	73.01	72.56	52.63
U2-Net <sup>[17]</sup>	53.89	73.24	71.74	72.52	72.13	51.06
UTNet <sup>[18]</sup>	56.66	74.39	70.22	71.32	70.76	65.63
UNet++ <sup>[19]</sup>	58.59	74.72	73.52	<b>75.53</b>	74.51	46.08
CBAM <sup>[22]</sup>	59.96	80.41	75.54	72.11	73.79	65.64
MobileViT <sup>[23]</sup>	56.85	72.33	67.14	68.37	67.75	60.76
Shuffle Attention <sup>[24]</sup>	60.60	76.99	73.42	74.85	74.13	64.01
BGRNet <sup>[25]</sup>	62.30	-	68.10	71.30	69.40	11.80
HMMEN <sup>[26]</sup>	<b>66.60</b>	-	72.30	73.00	72.60	15.15
本文	66.30	<b>80.99</b>	<b>75.95</b>	74.88	<b>75.41</b>	<b>60.17</b>

注:粗体表示最优,下同

FPS 是衡量模型推理效率的核心指标,数值越高代表推理速度越快。本文方法的 FPS 为 60.17,虽略低于 UNet(68.03)、UTNet(65.63)等轻量方法,但显著高于 BGRNet(11.80)、HMMEN(15.15)等 IoU 表现较优的方法;同时与 MobileViT(60.76)的推理速度接近,结合前文的精度优势,说明本文方法在检测精度与推理效率间实现了较好平衡,具备更强的实际部署适用性。

图 4 为本文方法在各类场景下的检测结果,在多场景下展现出优异鲁棒性。从实验结果图可见,原始、白天、雾天、夜晚、雪天等复杂场景中,可见光(VL)与红外(IR)输入经处理后,结果与真值(GT)高度契合。在面对电线尺度不一、分布密度随机、背景复杂等挑战,无论常规光照、恶劣天气,还是昼夜环境差异等,均能精准检测目标,有效验证方法对复杂场景的强适应性。

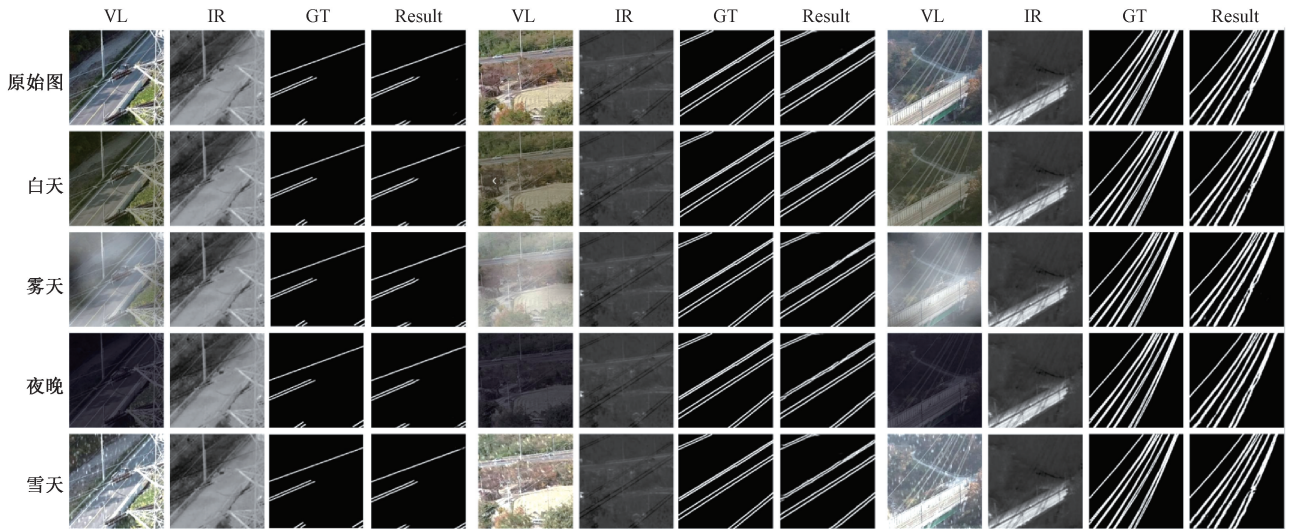


图 4 本文模型在各类场景下无人机电线检测效果

Fig. 4 The detection effect of the model in this paper on UAV power line detection in various scenarios

综上所述,本文方法在电线检测的定位精度、区域覆盖率等方面均展现出显著优势,验证了所提方法在复杂场景下的有效性和鲁棒性。

### 2.5 模块级消融实验

为验证 MSFE、AFF、DPIA 模块的有效性,开展消融实验,结果如表 3 所示。

基线模型(Baseline) IoU 为 53.74%, Robj 为 75.46%, Recall、Precision、F1-Score 分别为 71.41%、

表 3 消融实验结果

Table 3 Ablation experiments results (%)

MSFE	AFF	DPIA	IoU	Robj	Recall	Precision
-	-	-	53.74	75.46	71.41	71.85
✓	-	-	56.64	78.60	74.93	73.52
✓	✓	-	59.58	79.92	75.06	74.39
✓	-	✓	63.67	78.85	74.42	<b>75.33</b>
✓	✓	✓	<b>66.30</b>	<b>80.99</b>	<b>75.95</b>	74.88

71.85%、71.63%。引入 MSFE 后, IoU 提升至 56.64%, Robj 达 78.60%, 各指标均有增长。MSFE 可增强不同尺度特征, 丰富特征表达, 助力更精准的目标定位与识别。在 MSFE 基础上加入 AFF, IoU 进一步升至 59.58%, Robj、Recall 等指标同步实现优化。AFF 依据特征相关性自适应融合, 强化有效特征传递, 使特征互补更充分。在 MSFE 基础上加入 DPIA 时, IoU 达到了 63.67%。DPIA 构建双路径交互, 与 MSFE 协同增强特征交互与依赖挖掘, 精准捕捉细节与全局关联, 助力性能跃升。当 3 个模块全引入时, IoU 达 66.30%, Robj、Recall、Precision 及 F1-Score 等指标达到最优。三模块协同, 全方位提升模型对复杂场景检测能力, 也进一步验证了各模块必要性与协同优势。

如图 5 所示, 对比本文算法与基线算法在夜间、雪天等复杂场景下对多尺度电线的检测效果。图中红色标记处, 基线算法存在漏识别、误识别问题; 而本文算法凭借多尺度特征增强、自适应特征融合、双路径交互注意力机制等模块的协同作用, 可精准、完整检测电线, 直观验证了三模块融合增强的性能优势, 展现算法对复杂场景的适配性。

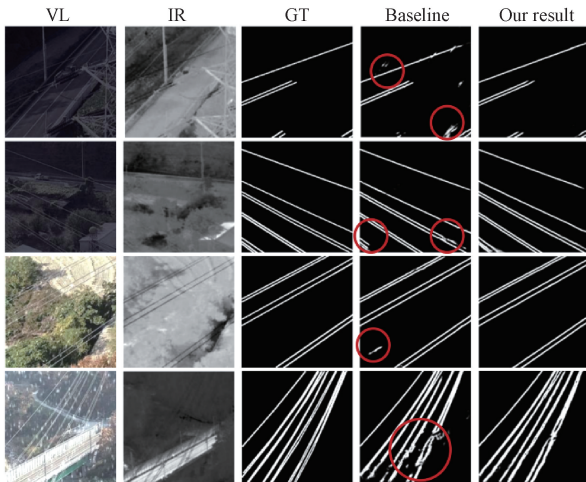


图 5 本文模型与基线模型效果对比

Fig. 5 Comparison of the Effectiveness Between the Model in This Paper and the Baseline

## 2.6 细粒度消融实验

### 1) MSFE 细粒度消融实验结果

为验证所设计多尺度特征增强模块 MSFE 的有效性, 以经典多尺度结构 ASPP 为对照开展细粒度消融实验。如表 4 所示: 采用 ASPP 时 IoU 为 65.88%, 采用 MSFE 时 IoU 提升至 66.30%; 同时, MSFE 将反映目标整体捕捉能力的 Robj 指标从 79.71% 提升至 80.99%, 综合性能指标 F1-Score 也从 74.01% 提升至 75.41%。这一结果归因于 MSFE 针对无人机电线长形态、随机尺度的检

测特性设计, 相比通用型 ASPP 更适配任务需求。

表 4 MSFE 细粒度消融实验结果

Table 4 MSFE fine-grained ablation experiment results (%)

Settings	IoU	Robj	F1-Score
With ASPP	65.88	79.71	74.01
With MSFE	<b>66.30</b>	<b>80.99</b>	<b>75.41</b>

### 2) DPIA 细粒度消融实验结果

为验证所设计双路径交互注意力模块 DPIA 的有效性, 以 PolarizedSelfAttention 为对照开展细粒度消融实验。如表 5 所示: 采用 PolarizedSelfAttention 时 IoU 为 65.91%, 采用 DPIA 时 IoU 提升至 66.30%; 同时, DPIA 将反映目标整体捕捉能力的 Robj 指标从 79.84% 提升至 80.99%, 综合性能指标 F1-Score 也从 75.15% 提升至 75.41%。这一结果归因于 DPIA 针对无人机电线细长连续结构、复杂场景的检测特性设计, 相比 PolarizedSelfAttention 更适配任务需求。

表 5 DPIA 细粒度消融实验结果

Table 5 DPIA fine-grained ablation experiment results (%)

Settings	IoU	Robj	F1-Score
With PolarizedSelfAttention	65.91	79.84	75.15
With DPIA	<b>66.30</b>	<b>80.99</b>	<b>75.41</b>

## 2.7 典型失败示例分析

图 6 为本文模型在两类高挑战巡检场景下的典型失败案例。

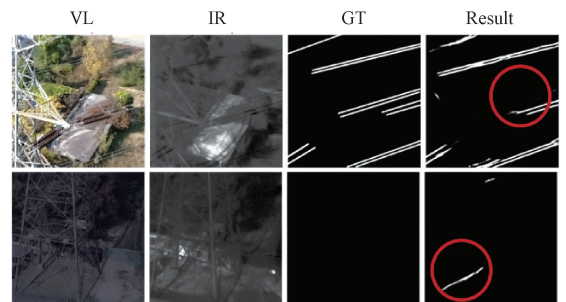


图 6 本文模型典型失败示例分析

Fig. 6 Analysis of typical failure examples of the model

尽管模型在多数场景中已实现电线目标的有效识别, 但面对复杂干扰与极端环境, 仍存在客观优化空间。第 1 个场景表现为漏检: 可见光图像中电线与铁塔视觉信息交织, 目标特征被干扰稀释; 红外图像受铁塔热辐射影响, 目标与背景对比度降低。由于模型未引入多模态融合机制, 无法互补两类图像的特征优势, 最终导致部分

目标漏检。第2个场景表现为误检:极端夜间低光照下,可见光图像模糊、目标细节丢失,红外图像则存在大量类目标噪声。缺乏多模态融合导致模型无法通过跨模态校验区分虚假特征与真实目标,引发误判。后续可将多尺度特征增强与多模态融合技术相结合,进一步提升模型对复杂背景干扰及极端环境的适应能力。

### 3 结 论

本文聚焦无人机多模态电线检测任务,提出一种基于多尺度特征增强的算法框架。该框架集成 MSFE、DPIA 和 AFF,旨在提升复杂背景干扰、电线尺度随机分布且差异大等场景下的检测效果。实验验证表明,本文方法在多种复杂场景中表现优异,相较于其他方法,在 IoU、Robj、Recall、Precision 及 F1-Score 等指标上均有显著提升。尤其在消融实验中,各模块有效性得到充分验证:在基线模型基础上,IoU、Robj、Recall、Precision 及 F1-Score 分别提升 12.56%、5.53%、4.54%、3.03% 和 3.78%,提升效果显著,表明它们对增强模型检测能力至关重要。此外,定性分析结果进一步证实了该算法在复杂场景下的强适应性与鲁棒性。因此,本文所提方法对无人机电线检测任务具有实际应用价值,有望为相关领域研究提供新的思路与技术支持。

### 参考文献

- [ 1 ] 黎淦钧, 黄英, 杨劲涛, 等. 基于先验特征的配电线检测[J]. 电子设计工程, 2020, 28(21): 173-177.  
LI H J, HUANG Y, YANG J T, et al. Distribution line detection based on prior features[J]. Electronic Design Engineering, 2020, 28(21): 173-177.
- [ 2 ] 何奇晗, 李中旭, 任凌飞, 等. 细节与全局多级融合的红外无人机目标检测[J]. 红外技术, 2025, 47(12): 1483-1490.  
HE Q H, LI ZH X, REN L F, et al. Infrared UAV target detection with multi-level fusion of details and global information [J]. Infrared Technology, 2025, 47(12): 1483-1490.
- [ 3 ] 苏雨蕾, 黄丹丹, 刘智, 等. 无人机对地小目标检测方法研究[J]. 电子测量与仪器学报, 2024, 38(9): 144-154.  
SU Y L, HUANG D D, LIU ZH, et al. Research on UAV ground small target detection methods[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 144-154.
- [ 4 ] 孙冰男, 于思恺, 张宇, 等. 融合动态多尺度轻量化无人机小目标检测算法[J]. 电子测量技术, 2025, 48(21): 199-206.  
SUN B N, YU S K, ZHANG Y, et al. Dynamic multi-scale fusion lightweight UAV small target detection algorithm [J]. Electronic Measurement Technology, 2025, 48(21): 199-206.
- [ 5 ] 张彼德, 王泽林, 廖其龙, 等. 基于轻量化改进 YOLOv8 的光伏阵列表面缺陷检测[J]. 电子测量与仪器学报, 2025, 39(6): 100-111.  
ZHANG B D, WANG Z L, LIAO Q L, et al. Surface defect detection of photovoltaic array based on lightweight improved YOLOv8 [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(6): 100-111.
- [ 6 ] SONG B, LI X. Power line detection from optical images[J]. Neurocomputing, 2014, 129: 350-361.
- [ 7 ] LIAO S, AN J. A robust insulator detection algorithm based on local features and spatial orders for aerial images [J]. IEEE Geoscience and Remote Sensing Letters, 2014, 12(5): 963-967.
- [ 8 ] 张彪, 荀荣科, 许家忠. 基于深度学习的工业轴承缺陷检测算法研究[J]. 仪器仪表学报, 2025, 46(4): 136-149.  
ZHANG B, XUN R K, XU J ZH. Research on industrial bearing defect detection algorithm based on deep learning[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 136-149.
- [ 9 ] GAO Z, YANG G, LI E, et al. Efficient parallel branch network with multi-scale feature fusion for real-time overhead power line segmentation [J]. IEEE Sensors Journal, 2021, 21(10): 12220-12227.
- [ 10 ] SONG J, QIN X, LEI J, et al. A fault detection method for transmission line components based on synthetic dataset and improved YOLOv5[J]. International Journal of Electrical Power & Energy Systems, 2024, 157: 109852.
- [ 11 ] 马滔, 熊英灼, 潘庆娜, 等. 基于 VSA-YOLO 11n 的无人机红外小目标检测[J]. 红外技术, 2025, 47(12): 1510-1517.  
MA T, XIONG Y ZH, PAN Q N, et al. UAV infrared small target detection based on VSA-YOLO 11n [J]. Infrared Technology, 2025, 47(12): 1510-1517.
- [ 12 ] LI B, WU D, CONG Y, et al. A method of insulator detection from video sequence [C]. 2012 Fourth International Symposium on Information Science and Engineering. Piscataway. IEEE, 2012: 386-389.
- [ 13 ] NGUYEN V N, JENSSEN R, ROVERSO D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning[J]. International Journal of Electrical Power & Energy Systems, 2018, 99: 107-120.

- [14] CHOI H, YUN J P, KIM B J, et al. Attention-based multimodal image feature fusion module for transmission line detection [J]. IEEE Transactions on Industrial Informatics, 2022, 18(11): 7686-7695.
- [15] YOU X, ZHAO X. An insulator defect detection network based on improved YOLOv7 for UAV aerial images[J]. Measurement, 2025, 253: 117410.
- [16] BAO W, DU X, WANG N, et al. A defect detection method based on BC-YOLO for transmission line components in UAV remote sensing images[J]. Remote Sensing, 2022, 14(20): 5176.
- [17] QIN X, ZHANG Z, HUANG C, et al. U2-Net: Going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404.
- [18] GAO Y, ZHOU M, METAXAS D N. Utnet: A hybrid transformer architecture for medical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2021: 61-71.
- [19] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, et al. Unet ++: A nested U-Net architecture for medical image segmentation[C]. International Workshop on Deep Learning in Medical Image Analysis. Cham: Springer International Publishing, 2018: 3-11.
- [20] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [21] LIU H, LIU F, FAN X, et al. Polarized self-attention: Towards high-quality pixel-wise mapping [J]. Neurocomputing, 2022, 506: 158-167.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [23] MEHTA S, RASTEGARI M. Mobilevit: Lightweight, general-purpose, and mobile-friendly vision transformer[J]. ArXiv preprint arXiv:211002178, 2021.
- [24] JANNU C, VANAMBATHINA S D. Shuffle attention U-Net for speech enhancement in time domain [J]. International Journal of Image and Graphics, 2024,

24(4): 2450043.

- [25] ZHOU W, JI C, FANG M. Transmission line detection through bi-directional guided registration with knowledge distillation [J]. IEEE Transactions on Industrial Informatics, 2023, 20(4): 5671-5682.
- [26] ZHANG S, ZHANG X, REN W. Bringing RGB and IR together: Hierarchical multi-modal enhancement for robust transmission line detection [J]. ArXiv preprint arXiv:250115099, 2025.

## 作者简介



**田洪坤**, 博士, 现为中国刑事警察学院讲师, 主要研究方向为人工智能和低空安全防控。

E-mail: thkcipuc@163.com

**Tian Hongkun**, Ph. D., now he is a lecturer at Criminal Investigation Police University of China. His main research interests include artificial intelligence, low-altitude security and prevention.



**姜囡**(通信作者), 博士, 现为中国刑事警察学院教授, 主要研究方向为人工智能和低空安全防控。

E-mail: zgxx\_jiangnan@126.com

**Jiang Nan** (Corresponding author), Ph. D., now she is a professor at Criminal Investigation Police University of China. Her main research interests include artificial intelligence, low-altitude security and prevention.



**李溯源**, 博士, 现为中国刑事警察学院讲师, 主要研究方向为人工智能和低空安全防控。

E-mail: 965643937@qq.com

**Li Suyuan**, Ph. D., now he is a lecturer at Criminal Investigation Police University of China. His main research interests include artificial intelligence, low-altitude security and prevention.



**任涛**, 博士, 现为东北大学教授, 主要研究方向为通信工程和人工智能。

E-mail: chinarentao@163.com

**Ren Tao**, Ph. D., now he is a professor at Northeastern University. His main research interests include communication engineering and artificial intelligence.