

控制与决策

Control and Decision

基于跨模态交互融合与全局特征校准的无人机电线检测

田洪坤, 姜囡, 任涛

引用本文:

田洪坤, 姜囡, 任涛. 基于跨模态交互融合与全局特征校准的无人机电线检测[J]. *控制与决策*, 2026, 41(3): 639–650.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2025.0764>

您可能感兴趣的其他文章

Articles you may be interested in

面向多目标侦察任务的无人机航线规划

UAV trajectory planning for multi-target reconnaissance missions

控制与决策. 2021, 36(5): 1191–1198 <https://doi.org/10.13195/j.kzyjc.2019.1284>

多无人机协同直播场景下自适应任务卸载决策

Adaptive task offloading decision of multi-UAVs cooperation in live broadcasting scenario

控制与决策. 2021, 36(4): 974–982 <https://doi.org/10.13195/j.kzyjc.2019.1104>

基于凸面体圆弧航路的无人机自主避障算法

Autonomous obstacle avoidance algorithm designed for UAV based on convex circular trajectory

控制与决策. 2021, 36(3): 653–660 <https://doi.org/10.13195/j.kzyjc.2019.0768>

天临空协同对地观测任务规划模型与并行竞争模因算法

Planning model and parallel competing memetic algorithm for space–near space–air based cooperative earth observation missions

控制与决策. 2021, 36(3): 523–533 <https://doi.org/10.13195/j.kzyjc.2020.0732>

分布式无人机的时变编队非线性控制设计

Time-varying formation nonlinear control of distributed multiple UAVs

控制与决策. 2021, 36(10): 2490–2496 <https://doi.org/10.13195/j.kzyjc.2020.0136>

基于跨模态交互融合与全局特征校准的无人机电线检测

田洪坤¹, 姜 囡^{1†}, 任 涛²

(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: 无人机电线的精确感知是电力系统运维与无人机自主避障的核心基础, 但复杂环境下的光照变化、背景干扰及电线自身特性, 给精准识别带来极大挑战. 现有单一模态检测方法依赖可见光或红外数据, 因复杂背景适应性差, 在恶劣条件下表现不佳, 存在明显局限性; 多模态检测虽通过融合可见光与红外数据提升了鲁棒性, 但在复杂环境适应性及任务特异性挖掘上仍有不足. 为此, 提出可见光与红外数据跨模态交互融合与全局特征校准检测方案: 通过跨模态交互引导融合模块 (CIGF) 实现双模态特征深度交互与优势互补, 通过全局特征重要性校准器 (GFSM) 精准校准枢纽特征并增强关键信息, 通过多感受野增强解码器 (MRED) 高效重建电线目标精细空间结构并实现像素级定位; 三大核心模块协同, 形成从特征提取、交互融合到全局校准再到精细解码的完整技术链路. 在无人机电线检测权威数据集 VITLD 上的实验显示, 该算法可满足检测精度与实时性的双重需求, 尤其在夜间低光照、雾天模糊、雪天遮挡等复杂极端环境中仍保持高精度, 突破传统方法应用瓶颈. 该方案可为解决无人机电线检测问题提供有效思路, 具有重要理论意义与实际应用价值.

关键词: 无人机; 电线检测; 自主巡检; 多模态融合; 目标识别; 计算机视觉

中图分类号: TP391.4 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2025.0764

引用格式: 田洪坤, 姜囡, 任涛. 基于跨模态交互融合与全局特征校准的无人机电线检测 [J]. 控制与决策, 2026, 41(3): 639-650.

UAV transmission line detection based on cross-modal interaction fusion and global feature modulation

TIAN Hong-kun¹, JIANG Nan^{1†}, REN Tao²

(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: Accurate perception of transmission lines by unmanned aerial vehicles (UAVs) is a core foundation for power system operation and maintenance as well as UAV autonomous obstacle avoidance. However, in complex environments, factors such as light changes, background interference, and the inherent characteristics of transmission lines themselves pose significant challenges to accurate identification. Existing single-modal detection methods rely on visual or infrared data. Restricted by their poor adaptability to complex backgrounds, these methods perform poorly under harsh conditions and have obvious limitations. Although multi-modal detection, which fuses visual and infrared data, has improved robustness compared with single-modal methods, existing multi-modal approaches still have shortcomings in adapting to complex environments and mining task-specific characteristics. To address these issues, a transmission line detection scheme based on cross-modal interaction fusion and global feature calibration for visual and infrared data is proposed. Specifically, the cross-modal interaction guided fusion (CIGF) module enables deep interaction and complementary advantages of dual-modal features; the global feature significance modulator (GFSM) accurately calibrates pivotal features and enhances key information; and the multi-receptive enhanced decoder (MRED) efficiently reconstructs the fine spatial structure of transmission line targets and achieves pixel-level positioning. These three core modules work together to form a complete technical chain covering feature extraction, interaction fusion, global calibration, and fine decoding. Experiments on VITLD, the authoritative dataset for UAV transmission line detection, show that the proposed algorithm meets the dual requirements of detection accuracy and real-time

收稿日期: 2025-07-21; 录用日期: 2025-11-28.

基金项目: 辽宁省自然科学基金项目 (2025-BS-0254, 2025-BS-0248); 国家重点研发计划项目 (2017YFC0821005); 辽宁省科技厅联合开放基金机器人学国家重点实验室开放基金项目 (2020-KF-12-11); 辽宁省研究生教育教学改革研究项目 (LNYJG2024310); 辽宁省教育科学“十四五”规划课题项目 (JG25DB472).

†通信作者. E-mail: zgxj_jiangnan@126.com.

performance. Notably, it maintains high accuracy even in complex and extreme environments such as low light at night, foggy blur, and snowy occlusion, breaking through the application bottleneck of traditional methods. This scheme provides an effective approach to solving the problem of UAV transmission line detection and holds important theoretical significance and practical application value.

Keywords: unmanned aerial vehicle (UAV); transmission line detection; autonomous inspection; multimodal fusion; target recognition; computer vision

0 引言

无人机对电线的精确感知是支撑多个领域任务开展的核心基础:其一,在电力系统的日常运维工作中,借助这一能力能够快速排查电线的老化、锈蚀等各类故障,削减人工投入成本的同时,降低现场作业时可能出现的安全隐患^[1];其二,在无人机自主导航与障碍物规避过程中,对电线进行实时检测是确保其在复杂空中环境下安全运行的必要条件^[2-3].由于自然环境中存在光照骤变、雾气笼罩、复杂背景干扰等不确定因素,以及电线本身具有尺度不一、分布无规律等特点,无人机在不同场景下进行电线检测时,依旧面临着识别精度不高、适应能力较弱等难题.因此,迫切需要更高效的检测手段来解决这些问题.

基于单一模态的无人机电线检测多依赖可见光或者红外数据,主要分为基于传统方法和深度学习方法两类.在可见光数据方面,传统方法主要通过边缘检测算子、霍夫变换、灰度变换等提取线路的几何特征,较具代表性的研究如 Song 等^[4]利用边缘检测算法从光学图像中提取输电线; Baker 等^[5]基于霍夫变换提取电线特征;黎溟钧等^[6]提出改进的分段灰度变换增强方法,通过抑制非输电线灰度区间突出目标.然而,这类方法对未知复杂背景敏感,可迁移性不强.随着深度学习技术的兴起,基于卷积神经网络

实现端到端检测逐渐成为主流.刘嘉玮等^[7]提出全卷积网络用于电力线识别,提升了复杂背景下的分割效果; Li 等^[8]采用金字塔多尺度特征分类实现输电线路区域定位,但其像素级精度受网格划分粒度的限制,难以达到理想的效果;李树挺等^[9]通过改进 YOLOv5 算法,引入注意力机制与分层特征融合结构,并优化训练权重提升了检测准确率; Zhang 等^[10]通过约束中间特征图的形态特征,并结合多尺度和结构化的先验信息提升电线检测效果.然而,受限于可见光图像的成像特性,上述研究在低光照或恶劣天气等条件下难以适用.红外模态凭借其对照光变化的鲁棒性,可用于补充可见光模态的不足.较有代表性的研究如 Jalil 等^[11]提出了利用红外图像检测电力设备故障;黎新等^[12]针对架空高压电线巡检难度大、检测效率低等问题,提出了基于红外图像的架空高压导线分割与检测的方法.然而,仅依托单模态红外数据,受限于红外数据的低分辨率和低对比度,在诸多复杂任务场景中也存在一定的局限性.

基于多模态数据的无人机电线检测通过不同的框架策略整合可见光与红外数据,以提升检测鲁棒性.从融合架构来看,可分为早期融合、后期融合和分阶段融合 3 种类型.早期融合也称为简单拼接融合,如图 1(a) 所示,将多模态图像直接拼接为多通道

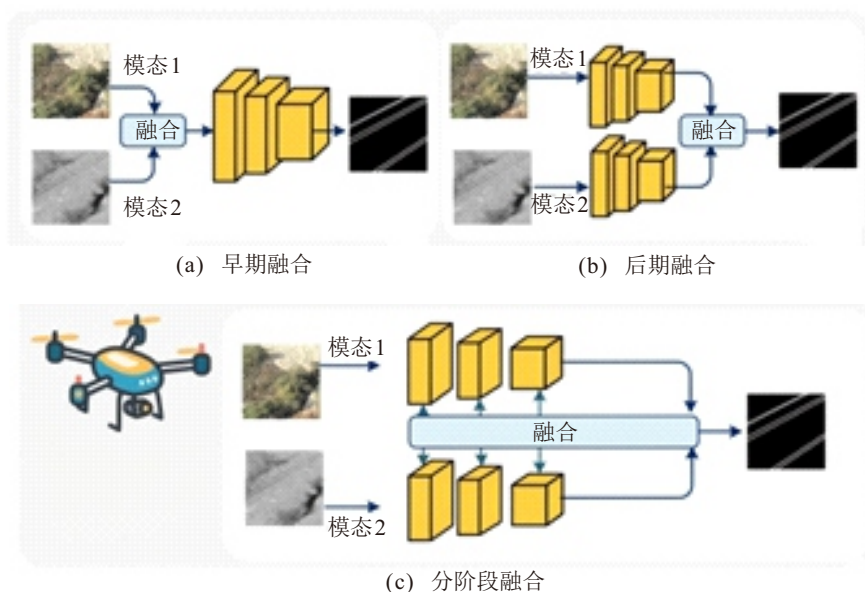


图1 常用的跨模态融合架构

输入, 如 Choi 等^[13] 将可见光与红外图像拼接为 4 通道输入. 相比于单模态数据, 该融合方式可以有效提升效果, 但这类结构未考虑模态差异, 在一些复杂场景下性能甚至低于单一模态. 后期融合如图 1(b) 所示, 是指采用双分支网络分别编码可见光和红外特征, 解码阶段通过特征相加或拼接融合. Hazirbas 等^[14] 提出的 FuseNet 在编码阶段保持模态独立, 解码时融合中间特征. 然而, 后期融合由于缺乏跨模态关联建模, 易出现特征错位. 分阶段融合如图 1(c) 所示, 是指在不同模态独立编码的同时进行分阶段的交互融合. 当前主流的方式是引入注意力机制动态学习模态权重, 如 Hu 等^[15] 构建的 SE 模块通过通道注意力校准特征; Choi 等^[13] 则提出多分支特征提取结构, 并结合注意力模块, 实现分阶段的特征建模与融合. 相比早期融合和后期融合, 分阶段融合可以更好地捕捉跨模态数据的内在联系, 提高特征利用率. 然而, 当前跨模态分阶段融合方法仍存在一定的局限性: 一方面是对复杂环境的适应性较差, 未能充分挖掘不同模态数据间的信息互补性、克服模态差异性, 在复杂条件下的表现缺乏鲁棒性; 另一方面是对任务特异性的挖掘不够, 未能针对电线分布随机、尺度差异大等特点优化跨模态交互和特征校准.

多模态(可见光、红外、深度)融合技术已在无人机及机器人目标感知领域开展大量探索, 相关研究针对不同任务的特性设计专用融合策略, 以图 1(c) 所示的分阶段融合为主, 且已取得显著成效. 针对无人人群计数任务, Zhou 等^[16] 提出可见光与红外双模态融合网络 DEFNet. 该网络通过密集数据增强模块融合同尺寸跨模态特征, 再借助高效空洞融合模块优化背景抑制效果, 最终利用融合解码模块结合高低层特征生成密度图, 有效提升了人群计数精度. 然而, 该方法聚焦于密集人群的统计性检测, 对双模态特征的交互深度仍存在优化空间. 在显著目标检测任务中, Fu 等^[17] 提出 JL-DCF 框架, 通过联合学习模块依托共享骨干网络提取可见光与深度双模态协同特征, 再经跨模态融合模块完成特征压缩与显式融合. 尽管该方法有效提升了复杂场景下显著目标的定位与识别效果, 但双模态交互的动态适配性较弱, 对小尺度细节的捕捉能力不足, 难以应对电线的尺度差异问题. Peng 等^[18] 针对复杂场景下可见光与热红外显著目标检测问题, 提出双模态融合网络 TCAINet, 该网络通过多维度数据预处理提升模型抗干扰能力, 在并行提取双模态多尺度特征后, 进一步借助通道注意力筛选关键信息, 并通过跨模态注意力实现协同融合, 有效提升了中大型显著目标的检

测精度. 但该方法聚焦于中大型目标的整体轮廓与位置检测, 对电线这类细长目标的局部细节关注不足, 无法满足电线检测对精细结构还原的需求. 针对无人机智能交通全天候检测场景, Ji 等^[19] 提出 FRCNet, 该网络采用双流骨干网络并行提取可见光与红外双模态多尺度特征, 通过监督级联可变形配准与多维度相似性匹配消除模态空间错位, 再借助增强模态自适应选择融合策略实现特征互补, 显著提升了车辆、行人等目标的检测鲁棒性. 然而, 该架构侧重局部上下文信息的利用, 无法同时兼顾电线这类线性目标的整体延伸方向(长距离连续性)与局部细枝细节. 面向机器人非结构化环境精准感知需求, Tong 等^[20] 提出基于可见光与深度双模态的 HFNet, 该网络通过分层特征融合与精细位姿对齐策略, 缓解了双模态特征层间语义与空间属性差异的问题, 进而显著提升了机器人对目标的空间定位能力. 但此类方法的核心目标是刚性物体的位姿感知, 侧重全局空间匹配而非像素级细节还原, 无法挖掘电线这类细长目标的专属特征, 难以满足无人机电线像素级检测的需求.

综上, 现有多模态融合方法均针对特定任务的目标特性设计, 虽在各自领域表现优异, 但因未适配电线的细长、随机分布等形态特征, 以及像素级定位、多尺度兼顾的检测需求, 难以直接应用于无人机电线检测任务. 因此, 亟需设计面向电线特性的专用跨模态融合与特征校准机制.

为了解决这些问题, 本文提出一种基于跨模态交互融合与全局特征校准的无人机多模态电线检测方法, 包括以下几个方面的创新:

- 1) 提出跨模态交互引导融合策略, 通过可见光与红外数据的深度交互与优势互补, 适配复杂环境下双模态信息的协同表征需求, 提升特征融合的有效性;
- 2) 提出全局特征重要性校准策略, 依托多尺度扩展与长距离上下文建模, 实现全局特征优化与校准, 为无人机电线检测任务提供全局层面的特征指导;
- 3) 提出多感受野增强解码策略, 采用渐进式上采样架构与多尺度上下文信息捕获机制, 有效地结合编码阶段高层语义与解码阶段边缘细节的关系, 提升电线目标精细结构的重建精度.

1 算法设计

为有效解决无人机电线检测中复杂背景干扰严重、电线尺度差异大、空间分布随机等导致的检测精

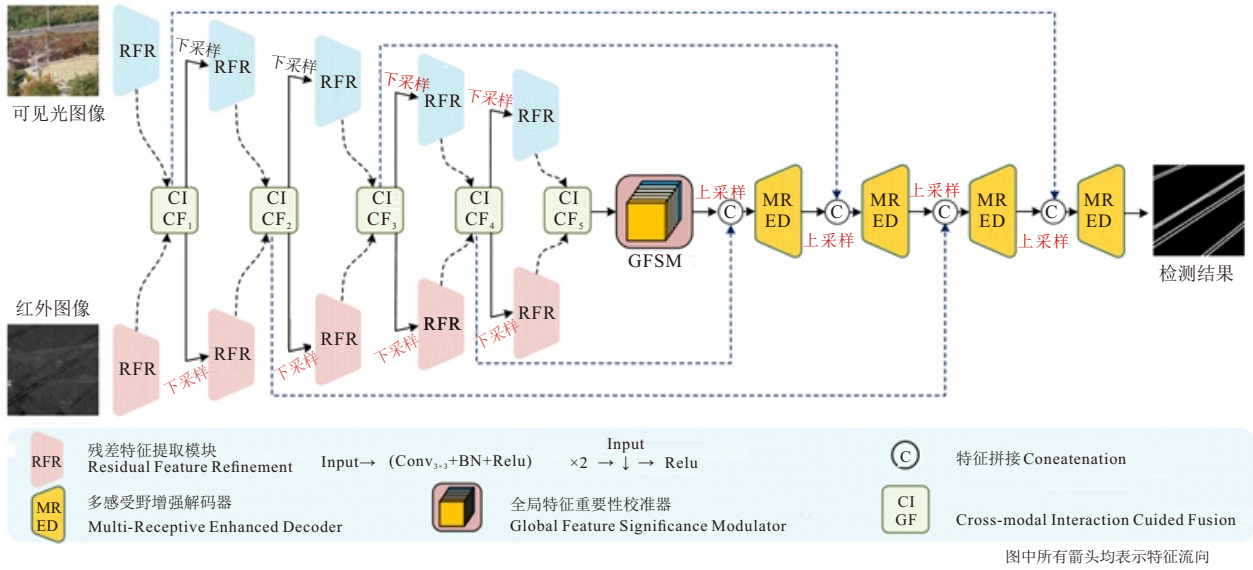


图2 基于跨模态交互融合与全局特征校准的无人机电线检测算法

度不足问题,同时克服现有方法在多模态图像融合不充分、特征校准能力弱等缺陷,本文提出基于跨模态交互融合与全局特征校准的创新算法框架,涵盖跨模态交互引导融合、全局特征重要性校准器及多感受野增强解码器等关键模块。

1.1 算法总框架

算法框架主要由跨模态交互引导融合 (cross-modal interaction guided fusion, CIGF) 编码阶段、全局特征重要性校准阶段 (global feature significance modulator, GFSM) 和多感受野增强解码阶段 (multi-receptive enhanced decoder, MRED) 三个关键阶段组成,旨在实现高效的特征提取与融合,提升无人机电线检测性能。总框架如图2所示。

在 CIGF 编码阶段,首先采用残差特征提取模块 (residual feature refinement, RFR) 对可见光与红外模态分别进行特征提取。该模块通过双层 3×3 卷积架构结合批归一化与 ReLU 激活函数,引入残差连接构建特征优化单元,有效缓解深层网络梯度消失问题。基于提取的双模态特征, CIGF 通过双向特征重建机制实现模态间信息互补,输出模态级特征图与融合特征图。模态级特征经最大池化下采样逐步提取高层语义特征,而融合特征则通过跳跃连接传递至解码阶段,为解码阶段提供跨模态协同信息。

在 GFSM 校准阶段,主要任务是对全局特征的重要性进行校准,以增强对重要特征的关注。GFSM 位于连接编码阶段与解码阶段的关键节点,承担着特征优化与转换的全局特征枢纽任务。通过多尺度扩展、长距离上下文建模以及特征校准,为后续解码阶段提供更具判别力的特征表示。

在 MRED 解码阶段,采用渐进式上采样架构,通过双线性插值恢复空间分辨率并与编码阶段跳跃特征拼接融合。在此基础上,通过 MRED 模块感受野扩展捕获多粒度上下文信息,经过多级特征融合与增强操作,逐步重建电线目标的精细空间结构,最终通过输出卷积层将解码特征映射为二值化电线检测图,实现像素级电线定位。

本算法采用组合损失函数以协同优化像素级分类和目标形状匹配。损失函数结合二元交叉熵 (BCE) 损失和 Dice 损失: BCE 损失确保像素级预测准确性,通过惩罚错误分类像素来优化分类边界; Dice 损失则专门针对细长电线结构优化,通过最大化预测与真实掩模的重叠面积来提升目标完整性检测能力。具体如下所示:

$$\mathcal{L}_{\text{total}} = \underbrace{\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]}_{\text{BCELoss}} + \underbrace{\lambda \left(1 - \frac{2 \sum y_i \hat{y}_i + \epsilon}{\sum y_i + \sum \hat{y}_i + \epsilon} \right)}_{\text{DiceLoss}}. \quad (1)$$

其中: y_i 为真实标签; \hat{y}_i 为预测概率; N 为像素总数; ϵ 为平滑系数,设置为 1; λ 为平衡权重,设置为 0.5。

1.2 跨模态交互引导融合模块 (CIGF)

CIGF 通过双向特征重建机制实现模态间信息互补,通过 3 个阶段实现,分别是模态通道交互增强阶段、全局模态权重评估阶段及多粒度特征融合阶段,输出模态级特征图与融合特征图。CIGF 整体框架如图3所示。

在模态通道交互增强阶段,依托双向通道交互

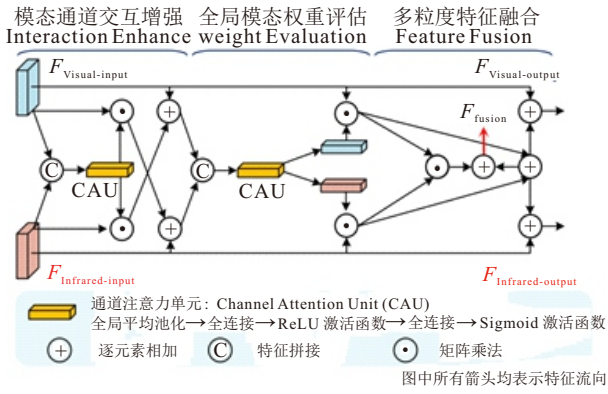


图3 跨模态交互引导融合模块

机制, 精准强化双模态特征判别能力. 此阶段调用轻量级通道注意力单元, 构建非对称特征引导逻辑: 先将红外特征 (I) 与可见光特征 (V) 沿通道维度拼接, 经全局平均池化 (GAP) 层与两层全连接层完成非线性变换, 再通过 Sigmoid 激活生成通道权重. 权重以残差形式对互补模态的特征响应实施调制, 达成模态间通道交互的增强目标. 该过程数学表达如下:

$$\begin{cases} \mathcal{W}_I = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \text{GAP}(\text{concat}(V, I)))) \\ \mathcal{W}_V = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \text{GAP}(\text{concat}(I, V)))) \\ I' = I + \mathcal{W}_I \odot V \\ V' = V + \mathcal{W}_V \odot I \end{cases} \quad (2)$$

其中: δ 为 ReLU 激活函数, 引入非线性变换能力; σ 为 Sigmoid 函数, 生成归一化通道权重; \mathbf{W}_n 为全连接计算矩阵; \odot 表示矩阵乘法, 通过各环节协同实现模态通道交互的有效增强.

在全局模态权重评估阶段, 基于上一阶段的特征增强成果, 模块借助轻量级注意力单元开展双模态全局重要性的评估. 具体而言, 先将经交互增强后的特征 (V' , I') 沿通道维度拼接, 以此捕获跨模态依赖关系; 再输入压缩激励单元, 生成通道注意力向量 A , 并将其均等分为可见光模态专属权重 A_V 与红外模态专属权重 A_I 这两个分量. 通过双模态特征质量动态调整, 形成关键特征强化响应, 该过程数学描述如下:

$$\begin{cases} A = \sigma(\mathbf{W}_4 \delta(\mathbf{W}_3(\text{GAP}(\text{concat}(V', I'))))) \\ [A_V, A_I] = \text{split}(A) \end{cases} \quad (3)$$

其中: δ 为 ReLU 激活函数, 引入非线性变换能力; σ 为 Sigmoid 函数, 生成归一化通道权重; \mathbf{W}_n 为全连接计算矩阵.

在多粒度特征融合阶段, 聚焦双模态信息的深度整合, 采用对齐路径、交互路径分层融合策略, 精准平衡共性特征提取与特性特征保留的需求, 构建多维度、多尺度的特征融合逻辑: 首先对齐路径, 基于上阶段输出的模态专属权重 A_V 、 A_I , 对原始双模

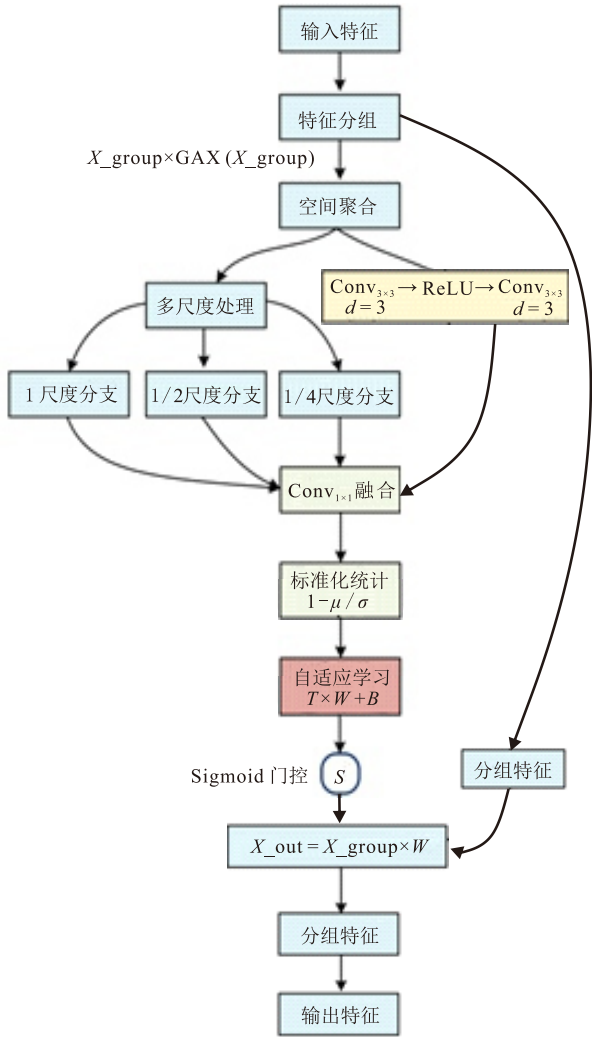
态特征分别执行加权对齐操作, 通过 $V \odot A_V$ 与 $I \odot A_I$ 的计算将双模态特征映射至统一特征空间, 消除模态间基础特征分布差异, 实现跨模态特征的初步校准; 然后交互路径, 为突破简单加权的局限, 引入跨模态交互项 $(V \odot A_V) \odot (I \odot A_I)$, 通过特征矩阵乘法深度挖掘双模态特征间的协同关联, 强化跨模态特征的非线性耦合, 适配多粒度特征需求; 最后为充分利用原始特征与对齐特征的互补性, 对两类特征执行元素级相加, 形成包含多尺度关联信息的融合特征图融合特征图 F_{fusion} . 同时, 为保留模态专属特性, 采取残差保留与跨模态融入策略, 在维持视觉模态专属特性的同时引入跨模态互补信息, 形成模态级特征图 V_{out} 与 I_{out} . 该过程数学表达如下:

$$\begin{cases} F_{\text{fusion}} = \underbrace{V \odot A_V + I \odot A_I}_{\text{对齐路径}} + \underbrace{(V \odot A_V) \odot (I \odot A_I)}_{\text{交互路径}} \\ V_{\text{out}} = \frac{1}{2}(V + V \odot A_V + I \odot A_I) \\ I_{\text{out}} = \frac{1}{2}(I + V \odot A_V + I \odot A_I) \end{cases} \quad (4)$$

其中 $1/2$ 为融合缩放系数, 平衡模态内特征与跨模态交互特征的贡献权重, 既让模态专属特征融入跨模态对齐信息, 又避免原始特征被过度地覆盖, 实现“融合增效”与“特性保留”的多粒度特征输出.

1.3 全局特征重要性校准器 (GFSM)

GFSM 作为衔接编码与解码流程的关键功能模块, 承担精准校准全局特征重要性的作用. 其运作依托特征分组策略, 将高维输入特征沿通道维度拆解为若干子特征空间, 以此为基础依次执行多模态特征提取、自适应特征评估与选择性特征校准 3 个递进阶段, 通过分组分层处理机制, 在有效降低计算复杂度的同时实现对关键特征的增强与冗余信息的抑制, 为下游任务输送具备更高判别力的特征表征, 适配复杂场景下多样化的特征表达需求, 整体框架如图 4 所示. 在多模态特征提取阶段, 借助多尺度操作和长距离上下文建模达成. 针对输入特征 \mathbf{X}_g , 构建多分支, 实现多尺度特征挖掘. 利用卷积与全局平均池化获取基础特征, 通过不同下采样系数 (2 倍、4 倍) 的下采样配合卷积, 挖掘多尺度上下文. 同时, 用带空洞卷积的链式操作拓宽感受野, 突破局部的限制, 捕获长距离依赖. 多尺度和长距离特征经通道拼接与 1×1 卷积融合, 生成多维度上下文表征 \mathbf{T} , 该过程数学表达如下:



图中所有箭头均表示特征流向

图4 全局特征重要性校准器

$$T = \text{Conv}_{1 \times 1} \left(\text{concat} \begin{bmatrix} \text{Conv}(\text{GAP}(\mathbf{X}_g)) \\ \text{Downsample}(\text{Conv}(\mathbf{X}_g), 2) \\ \text{Downsample}(\text{Conv}(\mathbf{X}_g), 4) \\ \text{Conv}_{d=5}(\text{ReLU}(\text{Conv}_{d=3}(\mathbf{X}_g))) \end{bmatrix} \right) \quad (5)$$

其中: GAP为全局平均池化, Downsample为下采样操作, $\text{Conv}_{d=n}$ 代表空洞率为 n 的空洞卷积, concat为通道拼接。

在自适应特征评估阶段, 核心是依据统计特性

动态生成特征重要性权重. 首先, 对于上阶段融合所得特征 T 计算其全局统计量, 遍历特征的所有空间位置, 得到特征均值和标准差; 然后, 引入标准化与可学习调整机制, 使特征重要性评估贴合数据内在分布; 最后, 通过 Sigmoid 函数映射输出归一化的特征重要性权重 W . 该过程数学表达如下:

$$W = \sigma \left(\gamma \cdot \frac{T - \mu}{\varphi} + \beta \right). \quad (6)$$

其中: μ 为均值, φ 为标准差, γ 、 β 为可学习参数, σ 为 Sigmoid 函数。

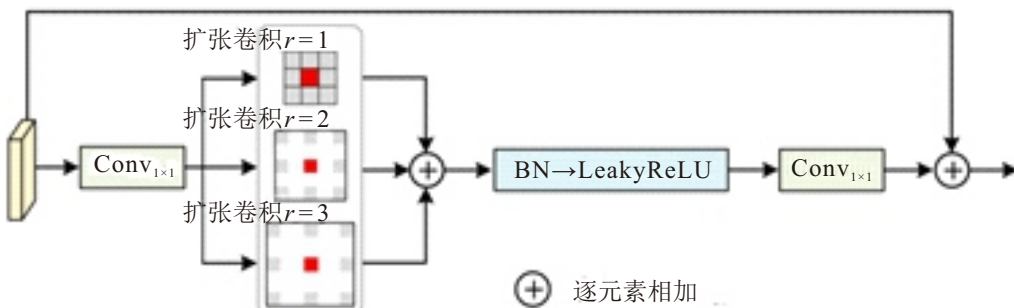
在选择性特征校准阶段, 通过重要性权重 W 精细调控原始特征 X . 先分组处理特征, 对每组特征 X_g , 复用多模态特征提取阶段输出 T_g , 计算其均值、标准差, 结合可学习参数生成组内权重; 再将权重与原始特征矩阵相乘, 增强关键特征、抑制非重要特征. 所有分组校准后的特征经通道拼接 ($\text{Concat}_{g=1}^G$, $G=8$), 输出最终校准特征 Y . 该过程数学表达如下:

$$Y = \text{GFMSM}(X) = \begin{cases} \text{Concat}_{g=1}^G \left(X_g \odot \sigma \left(\gamma_g \cdot \frac{T_g - \mu_g}{\varphi_g} + \beta_g \right) \right), \\ T_g = f_{\text{multi-scale}}(X_g) + f_{\text{long-range}}(X_g). \end{cases} \quad (7)$$

1.4 多感受野增强解码器 (MRED)

MRED 作为解码阶段核心组件, 以渐进式上采样为基础架构, 融合编码阶段跳跃特征, 借助多感受野扩展捕获多粒度上下文, 经多级特征操作重建电线精细结构, 最终输出二值化检测图, 实现像素级定位, 承担从抽象特征到精细检测结果的转换任务, 整体框架如图5所示。

在 MRED 模块中, 先经 1×1 卷积降维与初步特征变换, 再利用空洞率分别为 1、2、3 的空洞卷积, 构建多尺度感受野, 覆盖电线目标从局部细节到全局轮廓的多粒度上下文. 多分支卷积结果经逐元素相加、批归一化与 LeakyReLU 激活, 强化特征表达的多样性与区分度. 在此基础上, 经过 1×1 卷积调整通道后, 与残差输入 (原始上采样拼接特征 X) 相



图中所有箭头均表示特征流向

图5 多感受野增强解码器

加, 通过残差机制保留原始特征信息, 进一步融合多级特征. 多次重复该过程, 逐步优化特征, 重建目标的精细结构. 该过程数学表达如下:

$$\text{MRED}(X) = X + \text{Conv}_{1 \times 1} \left(\delta_{\text{leaky}} \left(\text{BN} \left(\sum_{i=1}^3 \text{Conv}_{d_i} (\text{Conv}_{1 \times 1}(X)) \right) \right) \right). \quad (8)$$

其中: $\sum_i^3 \text{Conv}_{d_i}$ 为带空洞率 d_i 的空洞卷积, 实现多感受野上下文捕获, 多分支卷积结果逐元素相加, 融合不同感受野特征; BN 为批归一化, 标准化特征分布, 加速训练与提升稳定性; δ_{leaky} 为 LeakyReLU 激活函数, 引入非线性, 增强特征表达能力.

2 实验结果与分析

2.1 实验环境

本文实验环境搭建如下: 操作系统采用 Ubuntu 18.04, 硬件配备 Intel Core i7-10700 CPU 与 NVIDIA GeForce RTX 3080 GPU, 算法以 Python 3.7 为编程语言, 基于 PyTorch 1.7.1 深度学习框架开展算法构建与实验.

2.2 数据集

实验选用 VITLD 数据集^[13] 用于无人机输电线路检测任务, 该数据集涵盖 420 组可见光与红外图像对, 配套像素级标注真值图. 数据划分遵循常规实验流程, 280 组用于模型训练、60 组用于模型验证, 80 组用于测试. 为深度检验算法在复杂天气、复杂光照等环境下的鲁棒性, 实验中对测试集中可见光图像, 补充白天常规光照、浓雾遮蔽、夜间低照度以及雪花覆盖等多样化样本, 经过扩充后的测试数据规模共达到 400 组.

训练阶段通过多维度数据增强策略提升模型泛化能力: 对输入图像执行旋转、随机翻转、边界裁剪等几何变换, 同时模拟颜色失真、雾效遮挡、覆雪覆盖等环境干扰, 使模型在训练过程中可充分学习不同场景下的特征模式, 适配实际无人机输电线路感知的复杂条件.

2.3 网络参数及评价指标

网络训练时, 优化器采用 RAdam, 初始学习率设为 0.001, 权重衰减取 0.0005, 搭配 StepLR 调度器, 每 20 个轮次学习率衰减为原有的 1/2; 批量大小为 4, 总训练 200 轮次.

参考相关研究实践, 涵盖以下实验评价指标:

1) 交并比 (intersection over union, IoU), 用于衡量检测结果与真实目标的定位重叠度, 有

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|}. \quad (9)$$

其中: 集合 A 为预测目标区域, 集合 B 为真实目标区域, $|A \cap B|$ 为交集的像素数量, $|A \cup B|$ 为并集的像素数量.

2) 目标识别率 (object recognition rate, Robj), 用于评估算法对电线等细长目标的完整识别能力, 有

$$\text{Robj} = \frac{\sum \text{TP}_{\text{obj}}}{\sum \text{GT}_{\text{obj}}}. \quad (10)$$

其中: TP_{obj} 为单张图像中被正确识别的目标数量, GT_{obj} 为单张图像中真实目标的总数.

3) 召回率 (Recall), 用于衡量算法对真实电线目标的检出能力, 有

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (11)$$

其中: TP 为预测为目标且真实为目标像素数量, FN 为真实为目标但预测为非目标的像素数量.

4) 精确率 (Precision), 用于衡量算法预测结果中真实电线像素的占比, 有

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (12)$$

其中: TP 为预测为目标且真实为目标像素数量, FN 为真实为目标但预测为非目标的像素数量, FP 为预测为目标但真实为非目标的像素数量.

5) F1 分数 (F1-Score) 用于综合评估召回率与精确率的平衡性能, 有

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (13)$$

6) 帧率 (frames per second, FPS), 用于衡量算法的实时处理能力, 反映单位时间内算法可处理的图像帧数, 直接关系到电线检测任务中实时监测需求的满足程度, 有

$$\text{FPS} = \frac{1}{\text{avg_infer_time}}. \quad (14)$$

其中 avg_infer_time 为单张图像的平均推理时间 (单位: s), 由测试集总推理时间除以图像总数得到.

通过上述多维度评价指标, 全面量化各算法的检测性能.

2.4 比较实验

在相同配置环境下对比多项相关算法, 包括无人机电线检测算法 UNet^[13] (算法 1)、基于多模态特征融合的无人机电线检测算法 UMFNet (算法 2)、DEFNet^[16] (算法 3)、JL-DCF^[17] (算法 4)、HFNet^[20] (算法 5)、FRCPNet^[19] (算法 6) 以及 TCAINet^[18] 等多模

态图像融合算法. 在对比 DEFNet、JL-DCF、HFNet、FRCPNet、TCAINet 算法时, 为更精准地衡量不同算法在跨模态数据融合方面的效果, 采用特定的对比策略: 将本文算法中的 CIGF 融合模块替换为对比算法中的核心融合模块, 以确保对比的针对性和有效性. 各对比方法选用的模块分别为算法 3 的密集型数据增强模块 (intensive data-enhancement module, IDEM)、算法 4 的跨模态融合模块 (cross-modal fusion module, CAM)、算法 5 的高层跨模态特征融合模块 (high-level cross-modal feature fusion, HCFF)、算法 6 的增强型模态自适应选择融合模块 (enhanced modal adaptive choice fusion module, EMACFM) 以及算法 7 的跨模态融合模块 (cross-modal fusion module, CAF).

表 1 所示数据为各算法在原始、白天、雾天、夜间和雪天等多场景的测试平均值. 从核心评价指标看, 本文算法在多数关键指标上表现最优, 实现了检测精度与运行效率的较好平衡. 在 IoU 指标上, 本文算法达到 67.79%, 相比次优算法 5 (65.14%) 高出 2.65%, 表明本文算法在检测定位的精准度上更具优势; 在 Robj 指标上, 本文算法达到 80.62%, 相比次优的算法 4 (79.51%) 高出 1.11%, 体现出本文算法对电线目标的检测完整度更优; 在 Recall 指标上, 本文算法达到 76.57%, 相比次优算法 4 (76.03%) 高出 0.54%, 表明本文算法对目标的覆盖能力更全面, 能捕捉到更多真实存在的目标; 在 Precision 指标上, 算法 7 以 76.29% 为最优, 本文算法达 75.95%, 相比其低 0.34%, 但结合 IoU、Robj、Recall 等指标看, 本文算法均优于算法 7, 整体检测性能更均衡; 在 F1-Score 指标上, 本文算法达到 76.26%, 相比次优的算法 4 (75.93%) 高出 0.33%, 这进一步验证了本文算法在检测性能的均衡性上更突出, 能够在保证召回率的同时兼顾精确率; 在 FPS 指标上, 本文算法为 47.62, 虽并非最高 (如算法 1 的 FPS 为 68.03), 但在 IoU、Robj、Recall、F1-Score 等核心精度指标均表现

表1 对比实验结果

比较算法	IoU (%)	Robj (%)	Recall (%)	Precision (%)	F1-Score (%)	FPS
算法1 ^[13]	52.92	74.31	73.22	70.96	72.07	68.03
算法2 ^[13]	57.07	73.98	72.12	73.01	72.56	52.63
算法3 ^[16]	61.22	74.68	74.31	75.89	75.09	46.08
算法4 ^[7]	61.38	79.51	76.03	75.83	75.93	51.02
算法5 ^[20]	65.14	78.27	76.01	75.16	75.58	38.61
算法6 ^[19]	65.45	74.97	72.24	75.72	73.94	36.63
算法7 ^[18]	65.74	79.01	75.07	76.29	75.58	52.91
本文算法	67.79	80.62	76.57	75.95	76.26	47.62

优异的情况下, 仍保持了较为合理的运行效率, 充分体现了精度与效率的良好平衡.

图 6 展示了本文算法在各类场景下的检测结果, 涵盖无电线、有少量电线、有大量电线以及有大量电线且尺度差异较大等多种场景, 充分体现了该算法在多场景下的鲁棒性. 从实验结果能够清晰地看到, 在原始、白天、雾天、夜晚以及雪天等多样化复杂场景中, 经过处理的可见光与红外数据其结果与真值高度契合. 即使面临电线尺度不一、分布密度随机、背景复杂等挑战, 无论是在常规光照、恶劣天气, 还是昼夜环境差异等情况下, 对于无电线场景能准确判断无目标, 对于有电线场景也能清晰分辨不同尺度电线, 均可以精准检测目标, 有效验证了本文算法对复杂场景的适应性.

为了更全面地比较各算法在复杂场景下的性能表现, 采用气泡图可视化方法分别在夜间 (图 7)、雪天 (图 8) 和雾天 (图 9) 场景下综合评估算法性能. 气泡图横轴为不同算法, 纵轴代表 IoU 指标; 橙色参考线标注对比算法的最佳 IoU 值, 气泡大小与 Robj 指标正相关, 通过双指标结合可全面衡量算法在复杂环境下的鲁棒性. 在夜间场景 (图 7) 中, 光照不足导致目标与背景对比度骤降, 对算法检测性能形成严峻挑战. 对比算法的 IoU 值分布在 51.54% ~ 65.75% 之间, Robj 值处于 71.34% ~ 78.97% 的范围. 本文算法展现出显著优势, IoU 达到 69.74%, Robj 为 81.86%, 两项指标均高于所有对比算法, 表明本文算法在低光照条件下仍能精准检测电线目标. 在雪天场景 (图 8) 中, 雪花对目标的遮挡给算法检测提出了更高要求. 对比算法的 IoU 在 53.75% ~ 66.24% 之间, Robj 在 73.20% ~ 78.14% 范围内. 本文算法的 IoU 达到 67.92%, Robj 为 80.41%, 两项指标均位居首位. 在雾天场景 (图 9) 中, 气溶胶散射效应造成图像模糊, 增加了目标识别难度. 对比算法的 IoU 范围为 51.95% ~ 64.49%, Robj 为 71.55% ~ 78.76%. 本文算法再次表现最优, IoU 为 69.48%, Robj 为 81.03%, 体现出本文算法有效克服了雾气对图像质量的不良影响. 综上所述, 本文算法在夜间、雾天、雪天等复杂环境下均具备较强的鲁棒性与适应性.

2.5 消融实验

1) 模块级消融实验.

为验证所提出各模块在无人机电线检测任务中的有效性进行消融实验, 结果如表 2 所示. 当不使用任何模块时, 模型 IoU 为 53.74%, Robj 为 75.46%,

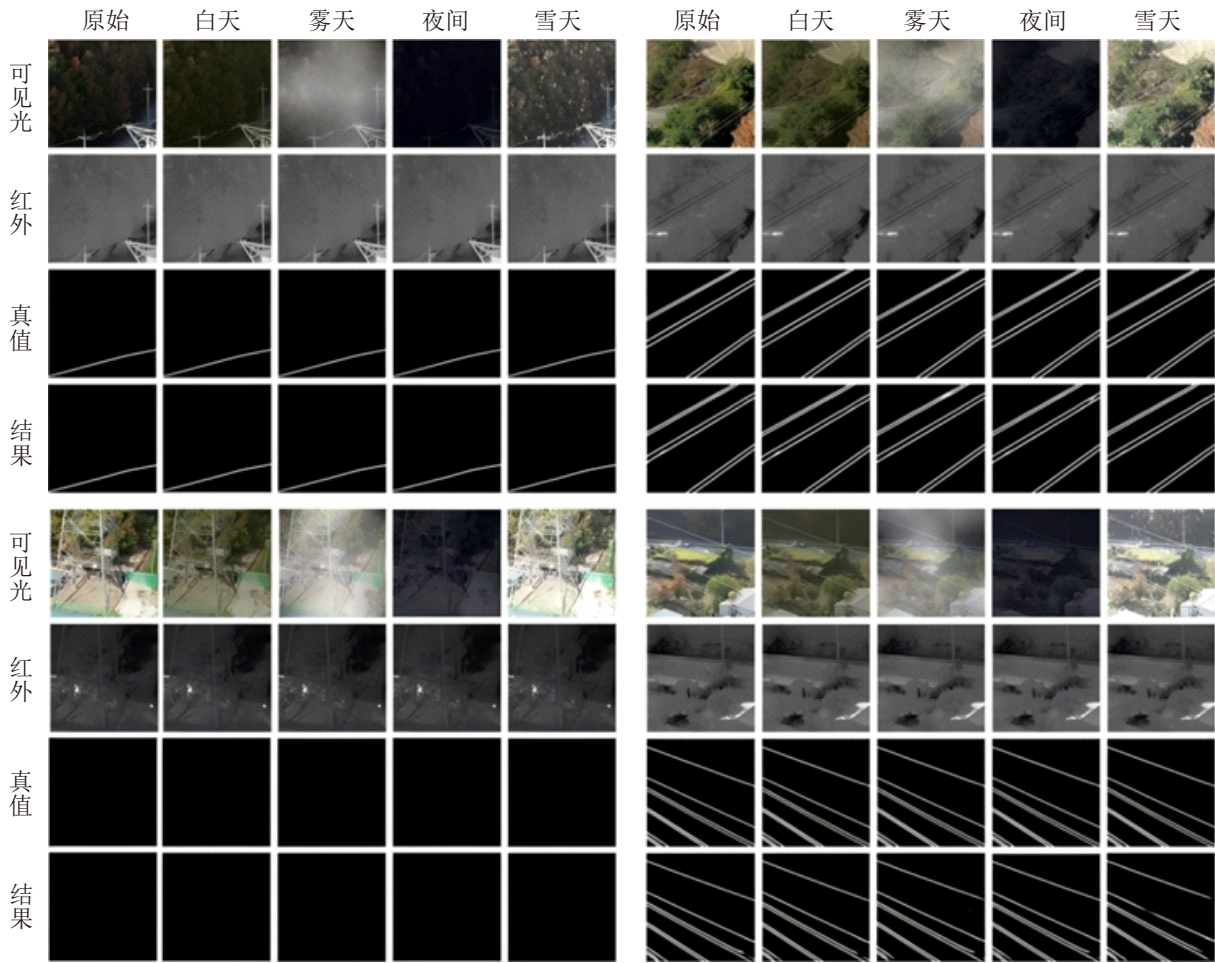


图6 本文算法在各类场景下无人机电线检测效果

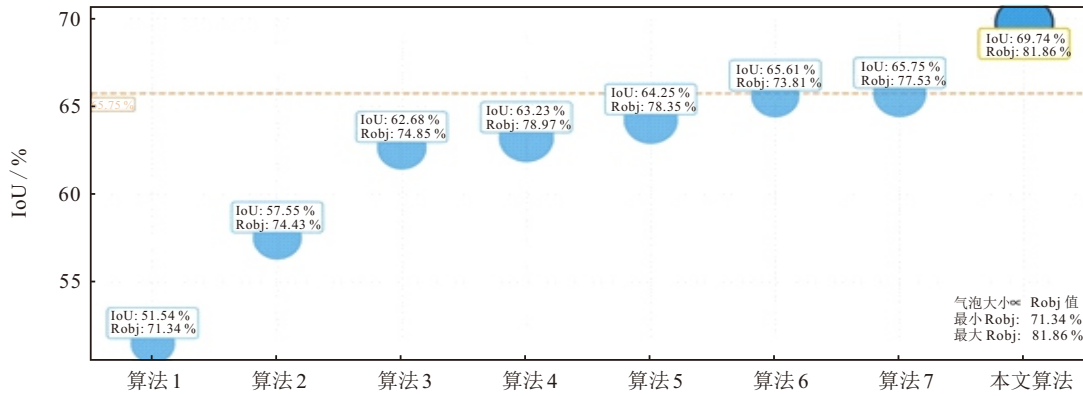


图7 各算法在夜间场景下表现对比

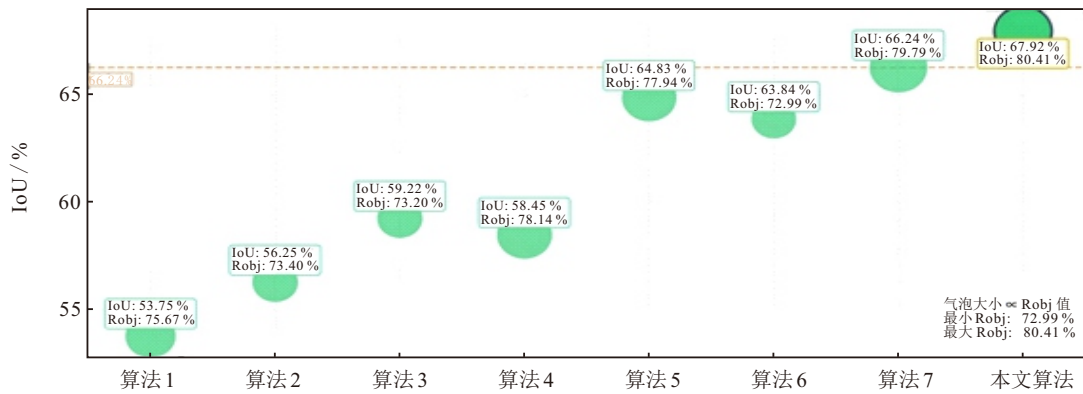


图8 各算法在雪天场景下表现对比

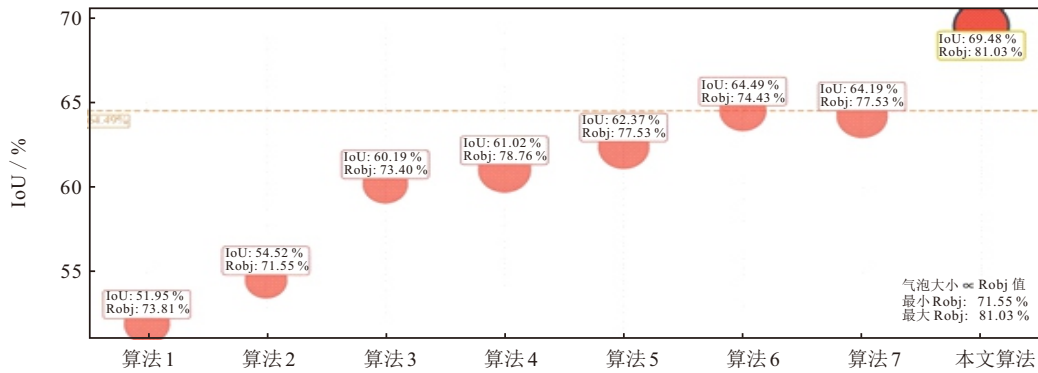


图9 各算法在雾天场景下表现对比

Recall 为 71.41%, Precision 为 71.85%, F1-Score 为 71.63%。加入跨模态交互引导融合模块 (CIGF) 后, 各项指标均有明显提升, IoU 达到 62.51%, Robj 为 77.69%, Recall、Precision 和 F1-Score 分别提高至 75.06%、73.36% 和 74.20%, 表明 CIGF 模块能够有效促进不同模态信息的融合, 提升检测性能。在此基础上引入全局特征重要性校准器 (GFSM), 模型性能得到进一步改善, IoU 升至 66.61%, Robj 为 78.56%, Recall、Precision 和 F1-Score 分别为 75.87%、74.28% 和 75.07%, 表明 GFSM 可对全局特征进行有效校准, 为下游任务输送具备更高判别力的特征表征。当同时使用跨模态交互引导融合模块 (CIGF)、全局特征重要性校准器 (GFSM) 和多感受野增强解码器 (MRED) 时, 模型性能达到最佳, IoU 为 67.79%, Robj 为 80.62%, Recall、Precision 和 F1-Score 分别为 76.57%、75.95% 和 76.26%, 体现了 MRED 在丰富特征感受野、提升检测精度方面的积极作用。综上所述, 本文所提出的 3 个模块在无人机电线检测中均能发挥正向作用, 且协同使用时效果更佳。

表2 模块级消融实验结果 %

GIGF	GFSM	MRED	IoU	Robj	Recall	Precision	F1-Score
—	—	—	53.74	75.46	71.41	71.85	71.63
√	—	—	62.51	77.69	75.09	73.36	74.20
√	√	—	66.61	78.56	75.87	74.28	75.07
√	√	√	67.79	80.62	76.57	75.95	76.26

2) 细粒度消融实验。

为进一步验证各模块在不同场景下的细粒度有效性, 基于雾天、夜间、雪天等典型场景开展消融实验, 结果以热力图形式呈现, 如图 10 所示。以“无模块”为基准, 模型在夜间 IoU 仅为 51.71%, 雾天 52.29%、雪天 54.35%, 各场景 Robj 与 F1-Score 也处于较低水平, 反映出无模块时算法在夜间场景的性能相对薄弱, 且雾天 (能见度低)、雪天 (背景干扰大) 等复杂场景下, 特征表征与多模态融合能力不足。引入跨模态交互引导融合模块 (CIGF) 后, 夜间性

能提升最显著, IoU、Robj、F1-Score 均明显改善; 雾天 IoU 升至 61.32%, 雪天达 63.09%, 验证了 CIGF 对多模态信息的融合能力, 尤其适配夜间、雾天等场景。加入全局特征重要性校准器 (GFSM, 模块组合为 CIGF + GFSM) 后, 夜间 IoU 进一步提升至 67.55%, 雪天 IoU 达 64.73%, 各场景 Robj 与 F1-Score 同步优化, 表明 GFSM 能为复杂场景输送更具判别力的特征。再引入多感受野增强解码器 (MRED, 模块组合为 CIGF + GFSM + MRED), 各场景性能达到最优: 夜间 IoU 达 67.66%、雾天达 66.03%、雪天达 67.08%, Robj 接近 80%, F1-Score 为各场景中最佳, 体现出 3 个模块的协同效应。综上, 所提出 3 个模块在雾天、夜间、雪天等不同场景下均能发挥正向作用, 且协同使用时可更好地适应无人机电线检测的复杂场景需求。

3) CIGF 细粒度消融实验。

为进一步验证跨模态交互引导融合 (CIGF) 模块的有效性, 及不同融合策略对检测性能的影响, 开展 CIGF 模块细粒度消融实验, 结果如表 3 所示。当 CIGF 采用“简单叠加融合”策略时, IoU 为 62.28%, Robj 为 78.43%, Recall 为 74.37%, Precision 为 75.48%, F1-Score 为 74.92%; 采用“双向重建融合”策略时, 各项指标均得到进一步提升, IoU 达 67.79%, Robj 为 80.62%, Recall、Precision 和 F1-Score 分别提高至 76.57%、75.95% 和 76.26%, 表明使用“双向重建融合”策略相较于“简单叠加融合”更能深度挖掘模态间的互补性, 进一步强化特征融合效果, 为后续检测任务提供更优质的特征支撑。

4) GFSM 细粒度消融实验。

为验证全局特征重要性校准器 (GFSM) 中多尺度建模、上下文建模分支的有效性, 及协同作用, 开展细粒度消融实验, 结果如表 4 所示。当“移除多尺度建模分支”时, 相较于“使用多尺度 + 上下文建模”的情况 (IoU 为 67.79%), IoU 下降至 64.90%, 变化较为显著, 表明多尺度建模分支可丰富特征的尺度表

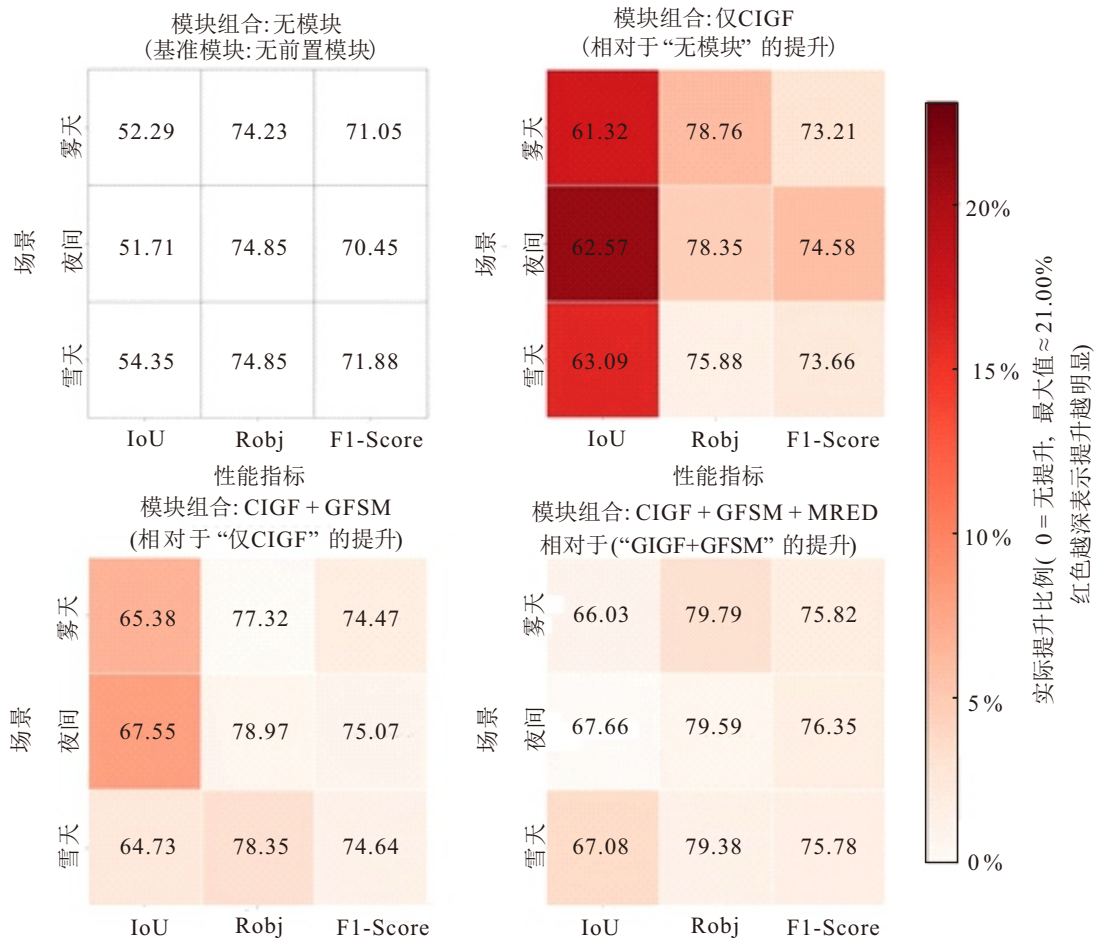


图10 模块级消融实验细粒度热力图

表3 CIGF 细粒度消融实验结果 %

CIGF设定	IoU	Robj	Recall	Precision	F1-Score
使用简单叠加融合	62.28	78.43	74.37	75.48	74.92
使用双向重建融合	67.79	80.62	76.57	75.95	76.26

表4 GFISM 细粒度消融实验结果 %

GFISM设定	IoU	Robj	Recall	Precision	F1-Score
移除多尺度建模分支	64.90	76.70	74.62	73.91	74.26
移除上下文建模分支	66.94	77.73	75.55	74.69	75.12
使用多尺度 + 上下文建模	67.79	80.62	76.57	75.95	76.26

征能力, 对 IoU 的提升起到关键作用, 有助于更全面地捕捉不同尺度下的特征, 优化检测区域的交叠精度. 当“移除上下文建模分支”时, 相较于“使用多尺度 + 上下文建模”的情况 (Robj 为 80.62%), Robj 下降至 77.73%, 变化较为显著, 表明上下文建模分支能强化长距离特征的关联与传递, 对 Robj 的提升效果突出, 可以更好地覆盖目标区域、减少漏检.

若使用多尺度与上下文建模两分支协同 (即 GFISM 同时发挥双分支作用的设定), 则模型的 IoU、Robj、Recall、Precision、F1-Score 均达到最优 (分别

为 67.79%、80.62%、76.57%、75.95%、76.26%), 明显优于“单独移除某一支”的表现, 体现出多尺度建模与上下文建模存在良好的协同效应, 二者结合可以从尺度多样性与长距离关联两个维度共同优化特征表征, 为下游检测任务输送更具判别力的全局特征, 进而全面提升无人机电线检测性能.

3 结论

本文针对无人机电线检测这一技术难题, 提出了一种融合可见光与红外数据的跨模态检测方案. 通过构建跨模态交互引导融合模块 (CIGF)、全局特征重要性校准器 (GFISM) 和多感受野增强解码器 (MRED), 形成基于跨模态交互融合与全局特征校准的算法框架, 显著提升了复杂场景下的检测鲁棒性. 实验结果表明, 所提出算法帧率 (FPS) 达到 47.62, 结合其在 IoU、Robj、Recall、Precision 及 F1-Score 等核心指标上的优势, 可满足检测精度与实时性的双重需求, 尤其在夜间低光照、雾天图像模糊及雪天目标遮挡等极端环境中, 仍能保持高精度检测性能, 有效突破了传统方法在复杂气象条件下的应用瓶颈. 此外, 通过消融实验进一步验证了所提出各模块的有效性, 相比于基线算法, 本文算法在 IoU、Robj、

Recall、Precision 及 F1-Score 指标上分别提高了 14.05%、5.16%、5.16%、4.10%、4.63%，表明 CIGF、GFSM 和 MRED 在无人机电线检测中能够发挥正向作用，且协同使用时效果更佳。综上所述，所提出方法为解决无人机电线检测问题提供了一种有效的思路，具有重要的理论意义和实际应用价值。

参考文献 (References)

- [1] 张贵峰, 吴新桥, 张巍, 等. 一种基于视频序列的输电线跟踪技术研究[J]. 电气时代, 2017(2): 54-57.
(Zhang G F, Wu X Q, Zhang W, et al. Research on power line tracking technology based on video sequence[J]. Electric Age, 2017(2): 54-57.)
- [2] 丁浩晗, 贺万程, 万俊, 等. DI-YOLO: 一种面向无人机航拍图像的高效小目标检测框架[J]. 控制与决策, 2025, 40(10): 3106-3116.
(Ding H H, He W C, Wan J, et al. DI-YOLO: An efficient small object detection framework for UAV aerial imagery[J]. Control and Decision, 2025, 40(10): 3106-3116.)
- [3] 陈志旺, 肖迪创, 吕昌昊, 等. 基于多尺度融合和高分辨特征增强的无人机航拍目标检测[J]. 控制与决策, 2025, 40(7): 2290-2299.
(Chen Z W, Xiao D C, Lv C H, et al. UAV aerial target detection based on multi-scale fusion and high-resolution feature enhancement[J]. Control and Decision, 2025, 40(7): 2290-2299.)
- [4] Song B Q, Li X L. Power line detection from optical images[J]. Neurocomputing, 2014, 129: 350-361.
- [5] Baker L, Mills S, Langlotz T, et al. Power line detection using Hough transform and line tracing techniques[C]. 2016 International Conference on Image and Vision Computing New Zealand. Palmerston North, 2016: 1-6.
- [6] 黎洪钧, 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. Prior characteristics-based distribution line identification[J]. Electronic Design Engineering, 2020, 28(21): 173-177.)
- [7] 刘嘉玮, 李元祥, 龚政, 等. 全卷积网络电线识别方法[J]. 中国图象图形学报, 2020(5): 956-966.
(Liu J W, Li Y X, Gong Z, et al. Power line recognition method via fully convolutional network[J]. Journal of Image and Graphics, 2020(5): 956-966.)
- [8] Li Y, Pan C F, Cao X B, et al. Power line detection by pyramidal patch classification[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2019, 3(6): 416-426.
- [9] 李树挺, 陈奇, 徐一鸣, 等. 基于改进 YOLOv5 的低空空域电线检测[J]. 中国科技信息, 2025(5): 124-127.
(Li S T, Chen Q, Xu Y M, et al. Wire detection in low altitude airspace based on improved YOLOv5[J]. China Science and Technology Information, 2025(5): 124-127.)
- [10] Zhang H, Yang W, Yu H, et al. Detecting power lines in UAV images with convolutional features and structured constraints[J]. Remote Sensing, 2019, 11(11): 1342.
- [11] Jalil B, Leone G R, Martinelli M, et al. Fault detection in power equipment via an unmanned aerial system using multi modal data[J]. Sensors, 2019, 19(13): 3014.
- [12] 黎新, 曾祥进, 洪俐, 等. 红外架空导线图像的分割与检测方法[J]. 红外技术, 2024, 46(12): 1390-1398.
(Li X, Zeng X J, Hong L, et al. Wire segmentation and detection method for infrared overhead wire images[J]. Infrared Technology, 2024, 46(12): 1390-1398.)
- [13] 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.
- [14] Hazirbas C, Ma L N, Domokos C, et al. FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture[C]. Computer Vision – ACCV 2016. Cham: Springer International Publishing, 2017: 213-228.
- [15] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7132-7141.
- [16] Zhou W J, Pan Y, Lei J S, et al. DEFNet: Dual-branch enhanced feature fusion network for RGB-T crowd counting[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24540-24549.
- [17] Fu K R, Fan D P, Ji G P, et al. JL-DCF: Joint learning and densely-cooperative fusion framework for RGB-D salient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 12345-12354.
- [18] Peng H, Hu Y F, Yu B C, et al. TCAINet an RGB T salient object detection model with cross modal fusion and adaptive decoding[J]. Scientific Reports, 2025, 15: 14266.
- [19] Ji Y Y, Song K C, Wen H W, et al. UAV applications in intelligent traffic: RGBT image feature registration and complementary perception[J]. Advanced Engineering Informatics, 2025, 63: 102953.
- [20] Tong L, Qian K. HFNet: High-precision robotic grasp detection in unstructured environments using hierarchical RGB-D feature fusion and fine-grained pose alignment[J]. Measurement, 2025, 253: 117775.

作者简介

田洪坤 (1996–), 讲师, 博士, 主要研究方向为智慧警务、多模态数据融合、低空安全防控, E-mail: thkcipuc@163.com;

姜因 (1979–), 教授, 博士, 主要研究方向为智慧警务、智慧侦查, E-mail: zgjx_jiangnan@126.com;

任涛 (1980–), 教授, 博士, 主要研究方向为无人系统智能控制, E-mail: chinarentao@163.com.