

文章编号 1004-924X(2025)08-1274-15

电力巡检中的偏振图像特征融合

倪梦瑶¹, 彭元龙², 胡 尚¹, 闫龙川², 郑锦坤³, 曹丹华^{1*}

(1. 华中科技大学 光学与电子信息学院, 湖北 武汉 430074;

2. 国家电网有限公司信息通信分公司, 北京 100761;

3. 国网江西省电力有限公司, 江西 南昌 330096)

摘要:为提升智能电网监测系统在复杂光照条件下的隐患目标识别精度, 本文提出了一种基于偏振度与偏振光强双模态信息融合的检测方法。首先, 设计双路特征融合网络, 通过密集卷积模块分别提取偏振强度图像和偏振度图像的特征, 以增强浅层信息的保留能力; 同时, 引入空间-通道协同注意力机制, 在空间与通道维度上建立特征依赖关系, 从而提升模型的特征选择性聚合能力。最后, 引入多尺度自适应结构相似性损失函数并设计加权算法优化参考图像的生成质量, 优化融合图像的结构保真度与目标显著性, 进一步提升其质量。实验结果表明, 与当前最先进的图像融合算法相比, 本研究提出的方法在多种评价指标上均展现出了显著的性能提升。同时, 通过消融实验进一步验证了网络模块和损失函数的有效性与实用性。在自建的目标检测数据集中, 采用本研究方法生成的融合图像在目标识别精度上达到了 91.5%, 同时 mAP@0.5 指标也达到了 0.916, 在客观评价方面优于对比方法, 提升了后续目标检测网络的识别准确性。

关键词: 电力巡检; 机器视觉; 偏振成像; 图像融合; 目标检测

中图分类号: TP391.4 **文献标识码:** A

doi: 10.37188/OPE.20253308.1274

CSTR: 32169.14.OPE.20253308.1274

Polarized image feature fusion in power inspection

NI Mengyao¹, PENG Yuanlong², HU Shang¹, YAN Longchuan², ZHENG Jinkun³, CAO Danhua^{1*}

(1. School of Optical and Electronic Information, Huazhong University of Science and Technology,
Wuhan 430074, China;

2. State Grid Information & Telecommunication Branch, Beijing 100761, China;

3. State Grid JiangXi Electric Power Supply Co., Nanchang 330096, China)

* Corresponding author, E-mail: dhcao@hust.edu.cn

Abstract: To enhance target detection accuracy in smart grid monitoring, especially under the challenging conditions posed by complex outdoor lighting, a comprehensive framework for fusing polarized and intensity images was proposed. To improve the accuracy of detecting potential hazards in smart grid monitoring systems under complex lighting conditions, this paper proposed a detection method based on the fusion of polarization and light intensity dual-modal information. This framework addressed the inherent difficulties of image analysis in diverse lighting scenarios, ensuring robust and accurate monitoring. Firstly, a dual-path feature fusion network was designed, which used dense convolutional modules to extract features

收稿日期: 2024-11-08; 修订日期: 2025-01-13.

基金项目: 国家电网公司科技项目 (No. 5700-202325308A-1-1-ZN)

from polarized intensity images and polarization degree images separately, thereby enhancing the retention capability of shallow information. Simultaneously, by constructing feature dependencies in both spatial and channel dimensions, new feature maps were selectively generated, solving the feature aggregation problem in feature fusion. Finally, a multi-scale adaptive structural similarity loss function was introduced, and a weighted algorithm was designed to optimize the quality of reference image generation, enhancing the structural fidelity and target saliency of the fused images, and further improving their quality. Experimental results demonstrate that, compared to state-of-the-art image fusion algorithms, the proposed method shows significant performance improvements across multiple evaluation metrics, compared to intensity images(S_0). These improvements are not only statistically significant but also visually apparent, as the fused images produced by our method are clearer, more detailed, and more informative. Ablation experiments further validate the effectiveness and practicality of the network modules and loss function. In a custom target detection dataset, the fused images generated by this method achieve a recognition accuracy of 91.5%, with an mAP@0.5 score of 0.916. These results showcase the superior performance of our method in objective evaluations and highlight its significant contribution to enhancing the detection accuracy of subsequent target detection networks in smart grid monitoring.

Key words: power inspection; machine vision; polarization imaging; image fusion; object detection

1 引言

电力设备巡检对保障电网安全运行至关重要。传统基于光强图像的巡检方法^[1-4]易受光照条件和天气影响,导致图像质量下降。相比之下,偏振成像技术通过解析光波与物体表面的相互作用,能克服雾霾等复杂气象条件的影响,显著提升目标纹理特征的恢复能力。偏振成像技术的应用价值与其物理特性密切相关,在不同领域展现出独特优势。在电力巡检领域,通过分析铁塔、导线等设备表面反射光的偏振态,可有效识别设备损伤、腐蚀等异常状态^[5];在水面船舶成像领域^[6],该技术能够抑制水面镜面反射干扰,显著提升目标-背景对比度,为船舶分类识别提供可靠依据;在军事图像领域,利用目标与背景的偏振特性差异,可突破传统伪装识别瓶颈,显著提升战场态势感知能力^[7]。基于上述特性,多模态目标检测研究主要沿两个技术路线展开:一是通过特征级融合实现不同物理特性的优势互补,二是基于决策级融合提升目标识别的可靠性。

互补特征融合通过整合多源图像特征提升检测性能,主要分为传统方法和深度学习方法两类。传统融合方法涵盖多尺度变换^[8-10]、稀疏表示^[11-12]、子空间分析^[13]及基于显著性的方法^[14],其中多尺度变换最为常用,包括离散小波变换(Dis-

crete Wavelet Transform, DWT)^[8]、曲线波变换(Curvelet Transform, CVT)^[9]及非下采样轮廓波变换(Non-Subsampled Contourlet Transform, NSCT)^[10]等。Su等人^[15]基于NSCT的融合算法,结合高效的降噪方法,有效保留了视觉真实性与偏振信息。

深度学习方法中Jung等人^[16]则设计了深度图像融合网络(DIFNet),通过CNN实现特征提取与融合,在多曝光图像融合中表现优异。陈永等人^[17]采用改进CBAM模块增强特征提取,结合L1范数优化特征融合。张等人^[18]提出的偏振图像融合网络(PFNet),引入了多尺度结构相似性(Structural Similarity Index Measure, SSIM)损失增强融合图像的细节与对比度。

Ma等人将生成对抗网络(Generative Adversarial Network, GAN)引入图像融合,先后提出FusionGAN^[19]和DDcGAN^[20],通过单/双判别器机制解决融合图像与源图像相似度过高的问题。

相比之下,决策级融合则通过整合不同物理特征的输出结果优化识别性能。Zhang等人^[21]提出的引导注意力特征融合(Guided Attention Feature Fusion, GAFF),利用模态内注意力模块增强行人区域特征,提高了行人检测系统的鲁棒性与可靠性。2022年,Chen等人^[22]提出的基于概率集成的决策信息融合方法,将先验分布与单模

态分类器结合,解决了特定模态中的漏检问题。

针对复杂场景下电网巡检目标识别准确率不足的问题,本研究提出了一种基于偏振-光强双模态融合的检测方案,主要工作有:(1)建立了端到端的双模态融合检测框架,通过无监督双路特征融合网络有效整合光强图像的结构信息与偏振图像的纹理特征,显著提升目标检测性能;(2)空间-通道协同的特征融合网络,在空间维度实现特征自适应选择,在通道维度完成特征动态聚合,充分挖掘双模态数据的互补特性;(3)建立了面向电力巡检的偏振-光强配准图像数据集(IPF),包含 3000 对高质量样本,涵盖输电线路、

城市道路等典型场景,为算法研究提供了可靠的基准数据。

2 方 法

2.1 电力巡检中的目标检测框架

为精准提取目标边缘与纹理特征,同时有效抑制复杂背景干扰,本研究创新性地引入偏振成像技术,构建了如图 1 所示的偏振-光强双模态融合检测框架。该框架通过融合偏振图像的细节分辨能力与光强图像的结构信息,显著提升了复杂环境下电力设备的检测性能。

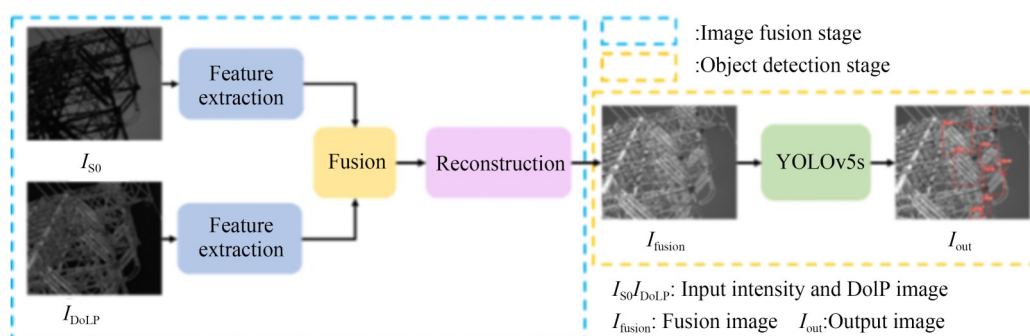


图 1 偏振图像融合与目标检测框架

Fig. 1 Overall framework of images fusion detection

在电力设施监测中,偏振信息主要采集自自然光照环境。尽管自然光本身具有随机偏振特性,但其在经物体表面反射或散射后会产生偏振态变化。这种偏振反射特性与表面材质的折射率及入射角密切相关,可通过 Fresnel 方程定量描述。具体而言,具体地,对于 s 偏振光(电场垂直于入射面)和 p 偏振光(电场平行于入射面),偏振反射率分别为:

$$R_s = \left(\frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} \right)^2, \quad (1)$$

$$R_p = \left(\frac{n_1 \cos \theta_t - n_2 \cos \theta_i}{n_1 \cos \theta_t + n_2 \cos \theta_i} \right)^2, \quad (2)$$

其中: n_1 和 n_2 分别为入射介质和目标材质的折射率, θ_i 表示入射角, θ_t 表示折射角。

不同材质表面的反射特性存在显著差异:玻璃(折射率约 1.5)相较于塑料(折射率约 1.3)具有更强的偏振反射特性。根据 Fresnel 方程,较高折射率导致更大的偏振反射率,这使得玻璃绝缘子表面的反射光表现出更显著的偏振特性^[23],

在偏振强度上明显优于其他材质表面。

为满足自然光条件下的实时成像需求,本研究采用分焦平面偏振相机进行图像采集。如图 2 所示,该相机在每个光强传感器前集成特定方向的偏振片,可同步捕获场景在 0° 、 45° 、 90° 和 135° 四个方向上的线偏振光强度并记录为 I_0 、 I_{45} 、 I_{90} 、 I_{135} ,通过这四个分量可使用 Stokes 矢量描述光的

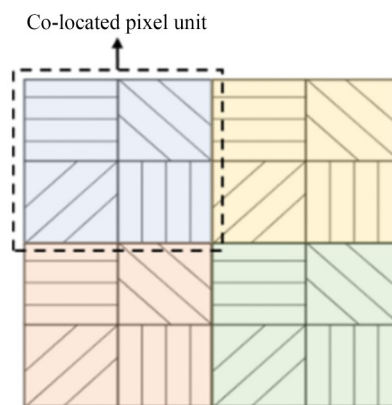


图 2 分焦平面偏振成像像素单元

Fig. 2 Focal plane polarization imaging pixel unit

偏振态,其表示如式(3)所示^[24-25]:

$$S = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ I_R - I_L \end{bmatrix}, \quad (3)$$

其中, I_R 和 I_L 分别表示右旋圆偏振光和左旋圆偏振光。

本文所研究的融合对象分别为 Stokes 参数计算得到的偏振强度图像和线偏振度图像。强度图像来源于 Stokes 参量的 S_0 分量:

$$S_0 = I_0 + I_{90}. \quad (4)$$

线偏振度(DoLP)图像是部分偏振光的强度与整个光的强度之比,表示如式(5):

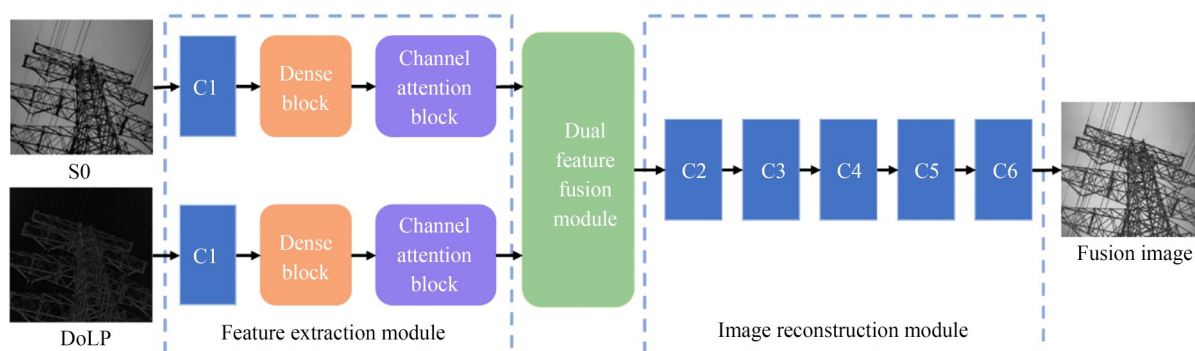
$$DoLP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \times 100\% = \frac{\sqrt{(I_0 - I_{90})^2 + (I_{45} - I_{135})^2}}{I_0 + I_{90}} \times 100\%. \quad (5)$$

本研究采用的线偏振度(DoLP)计算方法在提升图像对比度的同时保持了较低的计算复杂度,同时该方法能够有效提取传统光强图

像难以捕获的材质特征信息,为电力设备的精确识别提供关键判别依据。相比之下,偏振角图像易受复杂背景干扰,且对观测距离变化敏感,导致检测结果稳定性不足。基于线偏振度在金属目标检测中的显著优势^[26],本研究设计了一种级联检测框架:首先通过双路融合网络整合光强图像(S_0)与线偏振度图像(DoLP)的特征信息,随后将融合结果输入改进的 YOLOv5 模型^[27]进行目标检测。该方案充分利用了线偏振度的材质分辨能力,同时保持了检测算法的实时性。

2.2 偏振图像融合网络

在图像融合阶段,核心挑战在于有效提取并整合光强图像与偏振图像的互补特征,以实现最优融合效果。为此,本研究提出了一种无监督双分支融合网络(Polarization Dual-Branch Fusion Network, PDBFNet)。该网络在 PFNet 的基础上,通过引入通道注意力块和双路特征融合(Dual Feature Fusion, DFF)模块来实现。PDBFNet 的架构如图 3 所示,它主要由特征提取模块、特征



(a) 整个算法的结构
(a) Structure of the proposed PDBFNet

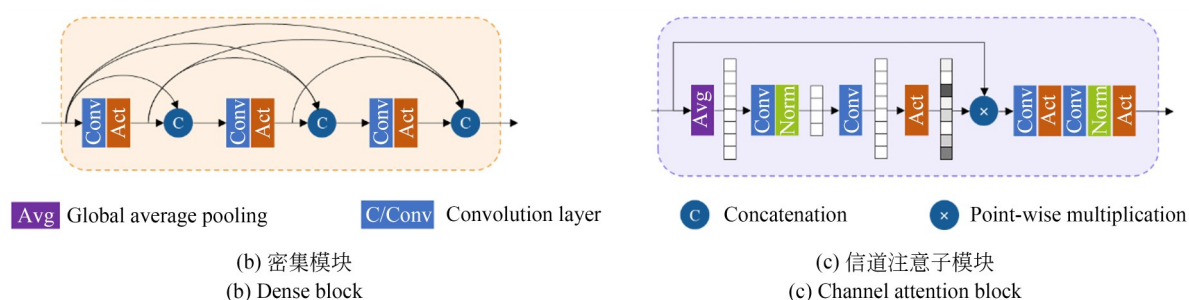


图3 PDBFNet 架构 (a) 显示了整个算法的结构;(b)和(c)给出了密集模块和信道注意力模块的细节

Fig. 3 Architecture of proposed PDBFNet. (a) shows the structure of the whole algorithm. (b) and (c) give the details of the dense block and the channel attention block

融合模块和图像重建模块组成。

(1)特征提取模块:采用双分支对称结构分别处理 S_0 和 $DoLP$ 图像。输入图像首先通过 16 通道卷积层(C1)进行初步特征提取,随后经过密集连接块^[27-28](包含三个 16 通道卷积层)提取高维非线性特征,通过特征复用机制有效缓解过拟合问题。最后引入通道注意力机制(SE 块^[29])动态优化特征表示,增强关键特征的表达能力。

(2)特征融合模块:采用全局特征聚合策略,通过空间-通道协同注意力机制建立特征依赖关系。与 PFNet 的简单拼接策略不同,DFE 模块通过自适应权重计算实现特征动态融合,显著提升重要特征的表达能力,其有效性将在消融实验中得到验证。

(3)图像重建模块:从融合特征图中生成最终融合图像,由五个逐渐减少通道的卷积层(C2-C6)串联组成。

2.3 双路特征融合模块

在本团队研究工作的基础上^[30],在融合网络中构建了双路特征融合(DFE)模块,如图 4 所示。DFE 模块采用特征分组-重组策略,将输入特征 F_{S_0} 和 F_{DoLP} 划分为两部分,形成新特征组 F_C 和 F_S :

$$F_C = \text{Concat}(F_{S1}, F_{D1}), \quad (6)$$

其中, F_{S1} 和 F_{D1} 分别通过分离函数得到。具体而言,对于 F_{S_0} 特征,其奇数通道构成子特征 F_{S1} ,偶数通道构成子特征 F_{S2} ,如式(7)所示:

$$\begin{cases} F_{S1} = \sum F_{S_0}(B, C_i, H, W), i \in 2n+1 (n=0, 1, 2, \dots) \\ F_{S2} = \sum F_{S_0}(B, C_i, H, W), i \in 2n (n=0, 1, 2, \dots) \end{cases} \quad (7)$$

其中, B, C, H, W 分别指代特征块的 *batchSize*, *channelSize*, *height*, *width* 数量。

DFE 模块的核心创新在于空间-通道并行融合机制:在空间维度执行特征选择,在通道维度实现特征聚合,生成增强特征 O_1 和 O_2 输出。最终通过特征拼接获得融合特征表示 F_{fusion} ,如公式(8)所示:

$$F_{\text{fusion}} = \text{Concat}(O_1, O_2) = \text{Concat}(\text{CFB}(F_C), \text{SFB}(F_S)). \quad (8)$$

通道融合分支通过加权融合策略突出关键特征:首先对强度信息特征和偏振度信息特征 $\{F_{S1}, F_{D1}\} \in R^{H \times W \times C_1}$ 逐点相加,随后通过全局平均池化提取空间维度全局信息 $F_{\text{Avg}} \in R^{1 \times 1 \times C_1}$,如式(9):

$$F_{\text{Avg}} = \text{GAP}(F_{\text{add}}) = \text{GAP}(F_{S1}, F_{D1}). \quad (9)$$

DFE 模块通过 Squeeze 操作实现特征通道压缩,如式(10)所示。该操作在保留全局上下文信息的同时显著降低特征冗余,有效提升模型泛化能力。

$$C_{\text{squeeze}} = F_{\text{sq}}(F_{\text{Avg}}) = \text{CBA}(F_{\text{Avg}}), \quad (10)$$

其中: F_{sq} 采用降采样卷积实现通道压缩,参照 SE-Net 设计,本研究中令 $C_{\text{squeeze}} \in R^{1 \times 1 \times \frac{C_1}{4}}$ 。随后,通过 Weighted 模块重建通道间依赖关系,生成通道融合权重,如公式(11)所示:

$$C_{\text{mid}} = F_{\text{ex}}(C_{\text{squeeze}}) = \text{Softmax}(\text{Conv}(C_{\text{squeeze}})). \quad (11)$$

F_{ex} 采用上采样卷积将通道数扩展至 $C_{\text{mid}} \in R^{1 \times 1 \times 2C_1}$,并分离出通道权重参数 $\{C_S, C_D\} \in R^{1 \times 1 \times C_1}$ 。DFE 使用 softmax 函数对 C_S 和 C_D 进行通道级归一化,指导输入特征的加权融合,得到通道融合分支的最终输出 O_1 :

$$\begin{cases} O_1 = F_{S1} \otimes C_S + F_{D1} \otimes C_D \\ C_S + C_D = 1 \end{cases} \quad (12)$$

空间融合分支旨在强化特征图的空间依赖关系,突出关键空间特征。与通道融合分支类似,首先对输入特征 $\{F_{S2}, F_{D2}\} \in R^{H \times W \times C_1}$,随后沿通道维度执行全局平均池化和最大池化,获得空间特征 $\{S_{\text{Avg}}, S_{\text{Max}}\} \in R^{H \times W \times 1}$ 。随后空间特征拼接生成中间向量 $S_{\text{mid}} \in R^{H \times W \times 2}$,再通过一个 Weighted 卷积模块

生成空间融合权重 $S_{\text{weight}} \in R^{H \times W \times 1}$:

$$\begin{cases} S_{\text{mid}} = \text{Concat}(\text{GAP}(S_{\text{add}}), \text{GMP}(S_{\text{add}})) \\ S_{\text{weight}} = F_{\text{ex}}(S_{\text{mid}}) = \text{Softmax}(\text{Conv}(S_{\text{mid}})) \end{cases} \quad (13)$$

最终对 S_{weight} 使用 softmax 函数归一化得到 S_S 和 S_D ,并完成输入特征子块执行空间维度的加权融合:

$$\begin{cases} O_2 = F_{S2} \otimes S_S + F_{D2} \otimes S_D \\ S_S + S_D = 1 \end{cases} \quad (14)$$

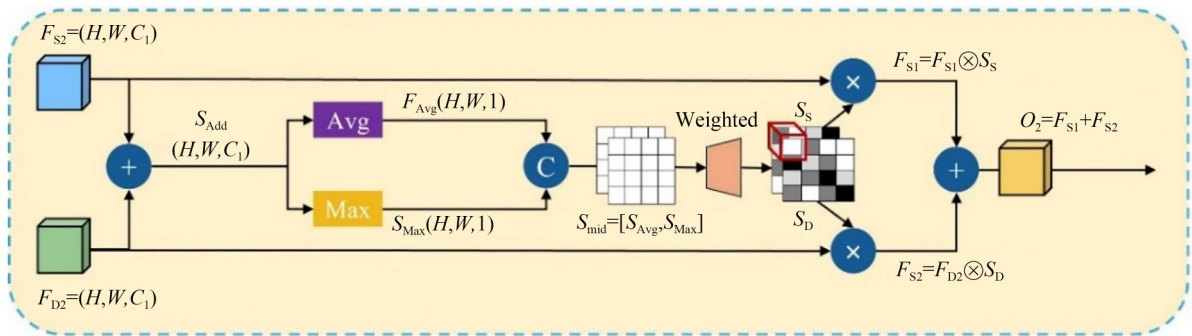
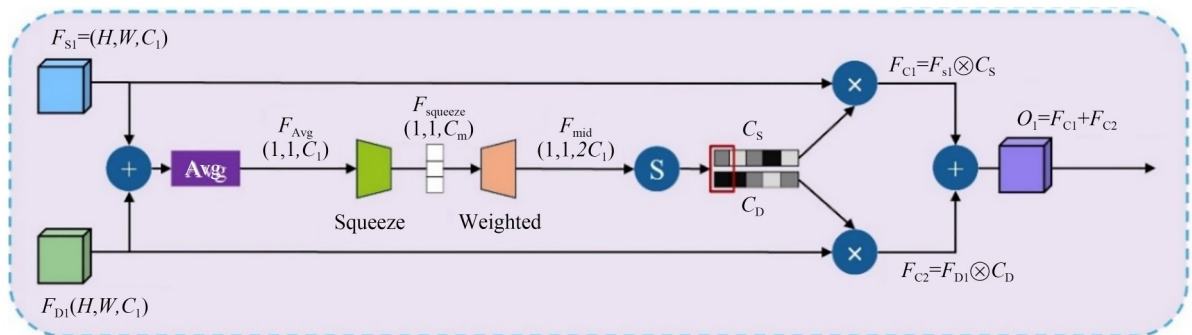
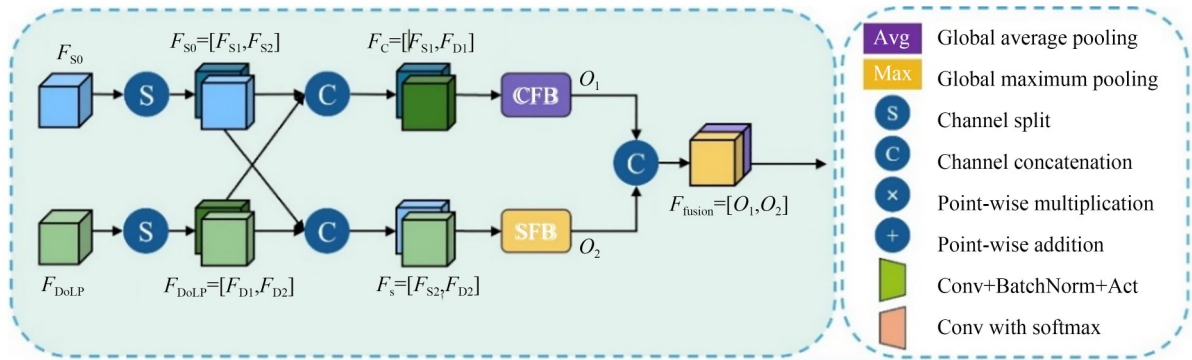


图 4 双特征融合模块示意图

Fig. 4 Illustration of dual feature fusion module

实验表明,该双分支特征融合模块在仅引入少量参数的情况下,显著提升了PDBFNet模型的细节保留能力。

2.4 损失函数

本文所使用的损失函数由多尺度自适应SSIM(MASSIM)损失函数和内容损失函数 $Loss_{\text{con}}$ 组成。对于内容损失函数,本文设计融合方法结合强度图像 I_{S_0} 和偏振度图像 I_{DoLP} ,得到新

的参考图像 I_G (Ground True Image)。

对于MASSIM,采用多尺度自适应结构相似度(MASSIM)作为损失函数,通过多组不同尺度的卷积窗口并行计算融合图像与参考图像的结构相似度(SSIM),并对各尺度结果进行加权平均。其中, $[W_n]$ 指卷积窗口的大小列表,卷积窗口根据输入图像的真实大小按比例计算。

$$Loss_{\text{MASSIM}} = 1 - \frac{1}{2n} \sum_{w \in [W_n]} (loss_{\text{SSIM}}(I_{S_0}, I_f; w) +$$

$$loss_{\text{SSIM}}(I_{\text{DoLP}}, I_f; w)),$$

$$w_i = \max(H, W) * \alpha_i \in [scale_n]. \quad (16)$$

$loss_{\text{SSIM}}(I^i, I^j; w)$ 指图像 I 和图像 I 与窗口 w

之间的局部结构相似度, 可以表示为:

$$loss_{\text{SSIM}}(I^i, I^j; w) = \frac{(2\overline{I_w^i} \overline{I_w^j} + C_1)(2\delta_{I_w^i I_w^j} + C_2)}{(\overline{I_w^i}^2 + \overline{I_w^j}^2 + C_1)(\delta_{I_w^i}^2 + \delta_{I_w^j}^2 + C_2)}, \quad (17)$$

其中: C_1, C_2 为稳定计算设置的常数, 分别设为 1×10^{-4} 和 9×10^{-4} 。 I_w^i 表示图像 I 在窗口 w 内的区域, $\overline{I_w^i}$ 表示区域 I_w^i 的平均值。 $\delta_{I_w^i}^2$ 和 $\delta_{I_w^j}^2$ 分别表示区域 I_w^i 的方差与 I_w^i 和 I_w^j 的协方差。

平均绝对误差(MAE)的计算公式如式(18)所示:

$$Loss_{\text{mae}} = \frac{1}{MN} \|I_G - I_f\|. \quad (18)$$

设图像分辨率为 $M \times N$, I_G 是 I_{S_0} 和 I_{DoLP} 生成的参考图像, I_f 为网络输出的融合图像。基于 S_0 富含对比度与细节信息、DoLP 包含边缘梯度信息的特性, 本章设计了一种频域分解融合算法: 通过对输入图像进行多尺度分解, 提取不同频段的特征信息, 并基于显著性权重 μ_i 自适应确定各输入图像的贡献度, 最终生成细节丰富、对比度优化的参考图像 I_G , 其数学表达为:

$$I_{\text{Base}}^i = I_i * \mu(x, y), \quad (19)$$

$$I_{\text{Mid}}^i = I_i * \sigma(x, y), \quad (20)$$

$$I_{\text{Detail}}^i = I_i - I_{\text{Base}}^i, \quad (21)$$

其中: $I_{\text{Base}}^i, I_{\text{Detail}}^i$ 和 I_{Mid}^i 分别为 I_i 的均值滤波图像、细节图像和中值滤波图像, $\mu(x, y)$ 和 $\sigma(x, y)$ 为均值滤波和中值滤波。生成图像 I_G 由细节部分和背景部分相加构成, μ_i 通过计算两幅图像的 I_{Base} 和 I_{Mid} 的差异来得到:

$$I_G = (0.5 \cdot I_{\text{Base}}^i + 0.5 \cdot I_{\text{Base}}^j) + (\mu_i \cdot I_{\text{Detail}}^i + (1 - \mu_i) \cdot I_{\text{Detail}}^j), \quad (22)$$

$$\mu_i = \frac{(|I_{\text{Base}}^i - I_{\text{Mid}}^i|)}{(|I_{\text{Base}}^i - I_{\text{Mid}}^i|) + (|I_{\text{Base}}^j - I_{\text{Mid}}^j|)}. \quad (23)$$

相较于直接选取两幅图像的平均值, 采用上述方法能够更好地保留和增强图像的细节与边缘信息, 通过自适应显著性权重确保重要区域的

清晰度和对比度。此外设置加权参数 γ 调节两个损失函数的比例, 保证 $Loss_{\text{MASSIM}}$ 和 $Loss_{\text{mae}}$ 在同一个量级。最终的损失函数可以表示为:

$$Loss = Loss_{\text{MASSIM}} + \gamma Loss_{\text{mae}}. \quad (24)$$

3 试验结果与分析

3.1 实验环境和数据集

本研究在 TensorFlow 框架下实现了所提算法, 采用 Adam 优化器进行模型训练。具体训练参数设置为: batch size=32, epoch=10, 初始学习率=1e-3, 损失函数加权参数 $\gamma=0.2$ 。所有实验均在 NVIDIA 3090 GPU 和 Intel Core i9-10900K CPU 平台上完成。

实验评估基于自建 IPF 数据集和公开 PolarLITIS 数据集^[31]。其中, PolarLITIS 数据集包含道路场景的 RGB-偏振多模态数据, 本研究选取 1 206 张图像作为测试集。IPF 数据集聚焦电力巡检场景, 采集了高压输电线路和城市道路的多距离、多角度图像, 以绝缘子作为主要识别目标。如图 5 所示, 数据集涵盖绝缘子、铁塔、施工机械、吊车等多种目标, 图像分辨率为 $1\,224 \times 1\,024$, 总样本量达 3 000 张, 其中 435 张用于 PDBFNet 网络训练。通过 64×64 的滑窗裁剪, 共获得 182, 528 组训练样本。网络训练时, 将 S_0 和 DoLP 图像作为独立输入, 并通过融合模块实现特征整合。



图 5 IPF 数据集的光强图像与偏振度图像

Fig. 5 Intensity image, DoLP image in IPF dataset

如表 1 所示, 与主流数据集^[17,31,35]相比, IPF 数据集在样本规模、图像分辨率等方面具有优势, 为算法性能评估提供了可靠的数据支撑。

表 1 相似数据集对比
Tab. 1 Comparison of similar datasets

Image	Quantity	Resolution	Type	Content
TNO	261	768×576	IR+RGB	Military Scene
M3FD	4 500	1 024×768	IR+RGB	Road Scene
RoadScene	221	irregular	IR+RGB	Road Scene
PolarLITIS	2 569	500×500	Polar+RGB	Road Scene
IPF (ours)	3 000	1224×1024	Polar+RGB	Electricity Scene

3.2 融合网络对比实验

本研究首先在 IPF 验证集上开展了损失函数优化实验。通过对比单一窗口尺寸的 SSIM 和 MAE 损失函数性能。其中,SSIM 采用了不同的 w 大小进行实验评估,选择信息熵(Information Entropy, $EN^{[32]}$),平均梯度(Mean Gradient, MG),标准差(Standard Deviation, $SD^{[33]}$)以及结构相似度(SSIM)作为评价指标,结果如表 2 和图 6 所示。

表 2 卷积窗口 w 对图像融合质量的影响
Tab. 2 Effect of w on image fusion quality in SSIM loss functions

Window Size	EN	MG	SD	SSIM	
				DoLP	S_0
5×5	7.045	9.202	50.573	0.382	0.419
7×7	6.685	11.069	35.712	0.521	0.330
9×9	6.993	9.299	49.130	0.373	0.413
11×11	7.104	8.370	57.981	0.292	0.409
13×13	7.018	<u>10.324</u>	48.249	<u>0.429</u>	0.394
15×15	7.007	9.267	<u>51.329</u>	0.355	<u>0.422</u>
17×17	<u>7.049</u>	9.139	49.762	0.367	0.428

注:*加粗字体表示最优结果,斜体加下划线表示次优

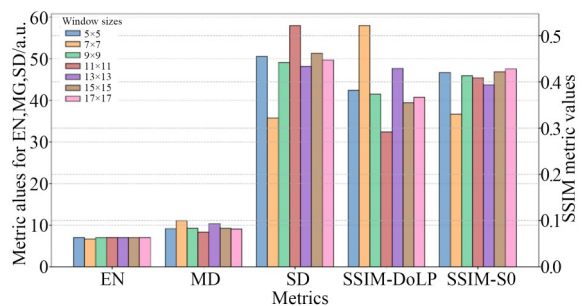


图 6 不同卷积窗口 w 下的图像融合质量

Fig. 6 w on image fusion quality in SSIM loss functions

不同窗口大小对评价指标的影响存在差异:11×11 窗口在信息熵(EN)和标准差(SD)指标上表现最优,而7×7窗口则在平均梯度(MG)和结构相似度(SSIM)上取得最佳效果。这表明单一窗口尺寸难以实现全局最优,验证了引入多尺度 SSIM 的必要性。

对于 w 的选择,本研究设计了三种方案:最优窗口集合,即能使测评指标获得最优结果的窗口大小集合 set1=[7,11,17];最优和次优窗口集合 set2=[7,11,13,15,17];以及全部窗口大小集合 set3=[5~17],分别对应 MASSIM-I, MASSIM-II 和 MASSIM-III。

表 3 多尺度卷积窗口 w 对图像融合质量的影响
Tab. 3 Effect of multi-scale w on image fusion quality in MASSIM

Method	EN	MG	SD	SSIM-DoLP	SSIM- S_0
MASSIM-I	7.038	8.569	50.223	0.345	0.426
MASSIM-II	7.155	9.547	55.968	0.355	0.411
MASSIM-III	6.902	9.108	46.199	0.381	0.432

注:*加粗字体表示最优结果

表 3 展示了 MASSIM 不同卷积窗口组合的测评效果,MASSIM-II 在各项指标上均达到最优或次优水平,其中信息熵指标甚至超越单一窗口最优结果,而 MASSIM-III 因引入过多噪声导致性能下降。因此,最终选择 MASSIM-II 作为模型训练的损失函数。

在融合网络对比实验中,PDBFNet 与现有先进方法进行了全面对比,包括传统方法(NSCT,CVT)和深度学习方法(DIFNet,DDcGAN,PFNet)。为更全面评估融合效果,在常规指标基础上引入互信息(MI)和 $Q^{AB/F[34]}$ 指标,并

优化 SSIM 计算方式:分别以 DoLP 和 S0 作为参考图像计算 SSIM 值后取平均。实验结果如表 4 和表 5 所示:表 4 展示了自建数据集未融合的 S₀ 和 DoLP 图像的客观质量评价结果,作为性能评估基准;表 5 展示了各算法在自建数据集上的融合图像客观质量评价结果,组 1 和组 2 分别对应传统方法和深度学习方法。图 7 给出了全部客观

表 4 自建数据集原始图像的客观质量评价

Tab. 4 Objective quality evaluation of raw images from IPF

Image	EN	SD	MG
S0	6.143	59.397	25.162
DoLP	5.771	32.919	72.400

质量评价结果的柱状对比。

表 5 融合模型在自建数据集上的客观质量评价

Tab. 5 Objective quality evaluation of fusion model on IPF

NO	Fusion method	EN	SD	MG	MI	Q ^{AB/F}	SSIM
1	NSCT	6.920	<u>67.307</u>	82.080	1.826	0.401	0.513
	CVT	6.863	58.376	<u>76.557</u>	<u>1.830</u>	0.392	0.358
2	DIFNet	5.760	20.828	46.199	1.167	0.383	0.466
	DDcGAN	5.630	23.311	33.996	1.098	0.388	0.438
	PFNet	<u>7.116</u>	59.395	74.625	1.769	0.398	<u>0.492</u>
	PDBFNet	7.206	76.917	66.170	2.744	0.401	0.469

注:*加粗字体表示最优结果,斜体+下划线表示次最优,全部指标均为越高越好

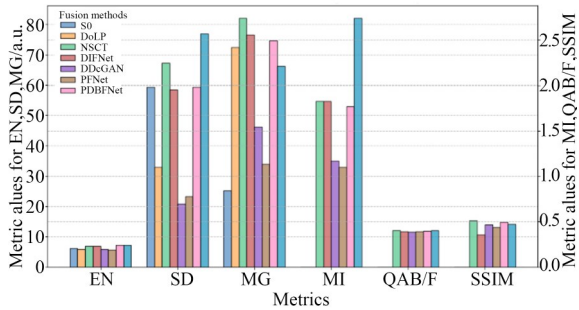


图 7 自建数据集的客观质量评价对比

Fig. 7 Comparison of objective quality assessment of IPF

实验结果显示,PDBFNet 在 0.28 s 内迅速完成处理,并在众多性能指标上占据领先地位,尤其在边缘细节的高保真保留与噪声的有效抑制方面表现卓越。相比之下,尽管 NSCT 在某些指标上展现出优越性,其长达 156.70 s 的处理时间无法满足实时性要求。而 PFNet 则因采用了较为简化的特征拼接方法,导致边缘信息的保留不够充分,且对比度有所降低。

此外,图 8 展示了各算法的融合图像。主视觉视觉上,NSCT 突出边缘,但引入噪声且丢失纹理细节;CVT 生成的伪边缘信息导致图像严重退化;DIFNet 因稀疏残差块难以捕捉 DoLP 图像

中的边缘细节,使图像模糊。尽管 DDcGAN 能融合两图像,但对背景过于敏感,产生伪影。PFNet 缺乏加权合并模块,导致 DoLP 特征利用不足,整体对比度降低。在最后一组图中,PDBFNet 的融合图像不仅保留了吊车倾斜支柱的纹理与边缘细节,且在细节图右侧边缘的施工人员的比 PFNet 的效果更加清晰,这说明融合图像在抑制噪声方面表现明显。

表 6、表 7 和图 9 展示了在公开数据集上的测试结果,表 6 为 PolarLITIS 数据集原始图像的客观质量评价,表 7 为各融合模型在 PolarLITIS 数据集上融合图像的客观质量评价,图 9 给出了表格的数据对比。实验得出,PDBFNet 在 EN 和 SSIM 指标上达到最优,其他指标也接近最优水平,在融合效果与计算效率之间实现了良好平衡。

表 6 PolarLITIS 数据集原始图像的客观质量评价

Tab. 6 Objective quality evaluation of the PolarLITIS

Image	EN	SD	MG
S0	6.562	62.292	25.827
DoLP	3.363	40.420	34.619

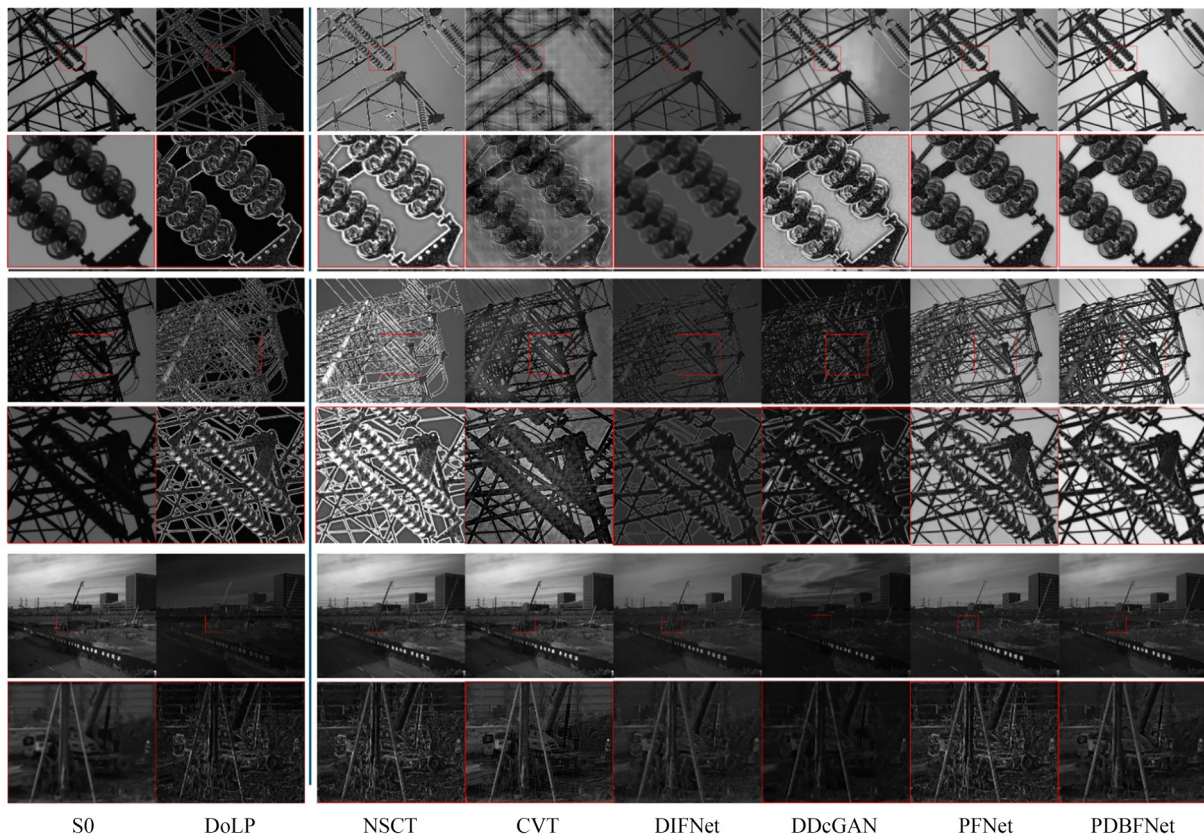


图 8 不同算法的主观融合结果对比,每组图都设有放大区域以显示融合细节

Fig. 8 Fusion samples of comparison algorithms. The area in the box region is enlarged to display the details

表 7 融合模型在 PolarLITIS 数据集的客观质量评价

Tab. 7 Objective quality evaluation of fusion model on PolarLITIS

NO	Fusion method	EN	SD	MG	MI	$Q^{AB/F}$	SSIM
1	NSCT	<u>6.765</u>	54.187	44.853	4.055	0.393	0.518
	DIFNet	5.673	20.743	16.435	3.214	0.381	0.472
2	DDcGAN	5.751	22.706	23.424	1.831	0.371	0.366
	PFNet	6.707	41.557	<u>39.686</u>	3.572	0.382	<u>0.534</u>
	PDBFNet	6.787	<u>52.328</u>	39.024	<u>3.780</u>	<u>0.383</u>	0.537

注:*加粗字体表示最优结果,斜体+下划线表示次最优,全部指标均为越高越好

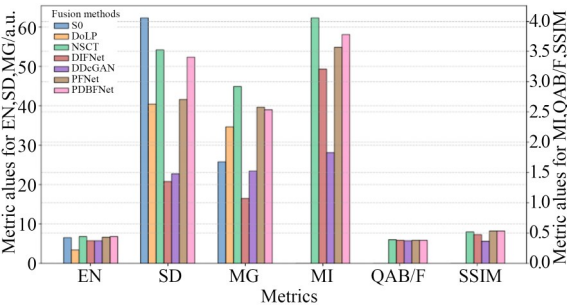


图 9 PolarLITIS 数据集的客观质量评价对比

Fig. 9 Comparison of objective quality assessment of PolarLITIS

3.3 消融实验

3.3.1 模型结构消融实验

为验证 DFF 模块和通道注意力机制的有效性,本文设计了两组消融实验,具体配置与结果如表 8 所示。

A 组实验重点评估了不同融合策略的性能差异。实验表明,相较于简单的点加或拼接融合方法,DFF 模块通过自适应调整通道和空间维度的融合权重,实现了强度图像细节与偏振图像边缘特征的优势互补,显著提升了融合性能。单独

表 8 网络结构消融实验结果
Tab. 8 Results of structure ablation experiment

Group	Description	EN	MG	SD	SSIM	
					DoLP	S ₀
A	Point Plus	7.127	9.359	53.189	0.337	0.219
	Concat	7.114	8.837	51.813	0.334	0.225
	Only-Spatial Fusion	7.082	9.170	51.684	0.350	0.226
	Only-Channel Fusion	7.005	9.212	49.755	0.347	0.223
B	No SE	7.151	9.307	54.208	0.342	0.225
	PDBFNet(ours)	7.155	9.547	55.968	0.341	0.231

注:*加粗字体表示最优结果

使用通道或空间融合模块均导致性能下降,证实了双维度融合策略的必要性。DFE 模块通过特征分组机制,实现了空间与通道权重的独立优化,从而获得更优的融合效果。B 组实验验证了通道注意力模块的有效性。实验结果显示,引入通道注意力机制后,PDBFNet 的特征提取能力显著提升。该模块通过强化密集块输出特征中的关键通道信息,进一步改善了融合质量。

3.3.2 损失函数消融实验

此外,本研究对损失函数进行了深入评估。在 $Loss_{SSIM}$ 对比了单一窗口 Single(11×11)与组合窗口(Win-sets)的性能差异; $Loss_{mea}$ 对比了 I_G 的两种生成方法: S_0 和 DoLP 的均值融合(average)、通过式(22)得到的加权融合(weight)。实验基于优化后的网络模型,使用 2 923 张测试图像进行评估。

表 9 结果显示,同时优化 $Loss_{MASSIM}$ 和 $Loss_{mea}$ 显著提升了信息熵和标准差指标,同时保持其他

质量指标处于次优水平。这种双损失优化策略有效平衡了细节保留与整体质量的需求,证实了所提方法的有效性。

3.4 目标检测对比实验

为确定最适合本研究任务的目标检测网络,对主流 YOLO 系列网络进行了系统性评估,包括 YOLOv5-n/s/m, YOLOv8-n/s 及 YOLOv11-s。评估指标涵盖网络参数量、推理时间、识别准确率、召回率以及 mAP 0.5。实验在 IPF 数据集上进行,采用 2 133 张训练图像和 432 张验证图像,所有网络均训练 100 个 epoch。推理阶段设置置信度阈值为 0.05,IOU 阈值为 0.2,并在 NVIDIA GeForce RTX 3090 显卡上测试推理时间,结果如表 10 所示。实验表明,YOLOv5-s 在检测精度和推理速度之间实现了最佳平衡,虽然部分模型在精度上有小幅提升,但其计算效率和参数量难以满足实时性和边缘部署需求。因此,本研究

表 10 使用偏振原始图像及融合图像的目标检测结果

Tab. 10 Object detection results of comparative periment

Network	Params/ M	Time /ms	Precision	Recall	mAP 0.5
Yolov5-n	1.9	10.5	0.838	0.832	0.874
Yolov5-s	7.2	22.3	<u>0.891</u>	0.854	0.902
Yolov5-m	21.2	50.8	0.915	<u>0.861</u>	0.911
Yolov8-n	3.2	8.7	0.873	0.826	0.877
Yolov8-s	11.2	26.2	0.878	0.85	0.882
Yolov11-s	9.4	26	0.886	0.874	<u>0.903</u>

注:*加粗字体表示最优结果

表 9 损失函数消融实验结果

Tab. 9 Results of Loss function ablation experiment

$Loss_{SSIM}$	$Loss_{mae}$	EN	MG	SD	SSIM	
					DoLP	S ₀
Single	Average	7.048	9.025	52.652	0.396	0.331
Win-sets	Average	7.227	9.590	57.770	0.396	0.360
Single	Weight	7.196	10.061	57.942	0.354	0.344
Win-sets	Weight	7.248	10.118	55.158	0.442	0.388

选择 YOLOv5-s 作为后续融合网络的检测算法。

表 11 和表 12 展示了各融合网络的检测性能,表 11 为使用偏振原始图像的目标检测结果,作为评估的基准。表 12 展示了融合图像的目标检测效果,其中 PDBFNet 在所有指标上均取得最优结果,证实了其生成的融合图像能有效降低误检率和漏检率。表中前两行分别展示了未融合的 S0 和 DoLP 图像的检测结果作为基准。

图 10 展示了典型场景下的检测效果:(a)对于严重遮挡的绝缘子目标,PDBFNet 显著提升

表 11 使用偏振原始图像的目标检测对比结果

Tab. 11 Target detection results using raw images

Image	mAP 0.5	Precision	Recall
S ₀	0.891	0.854	0.902
DoLP	0.852	0.843	0.873

表 12 使用融合图像的目标检测对比结果

Tab. 12 Target detection results using fused images

Fusion method	mAP 0.5	Precision	Recall
NSCT	0.861	0.866	0.898
CVT	0.851	0.871	0.887
DIFNet	0.858	0.848	0.885
DDcGAN	0.874	0.852	0.889
PFNet	0.852	0.812	0.863
PDBFNet	0.915	0.873	0.916

注:*加粗字体表示最优结果

了目标召回率;(b)在面对镜面反射干扰时,PDBFNet 有效抑制了误报;(c)在远距离铁塔检测中,PDBFNet 保留了更多有效信息,实现了 100% 的目标召回。这些结果充分证明了 PDBFNet 在复杂场景下的优越性能。

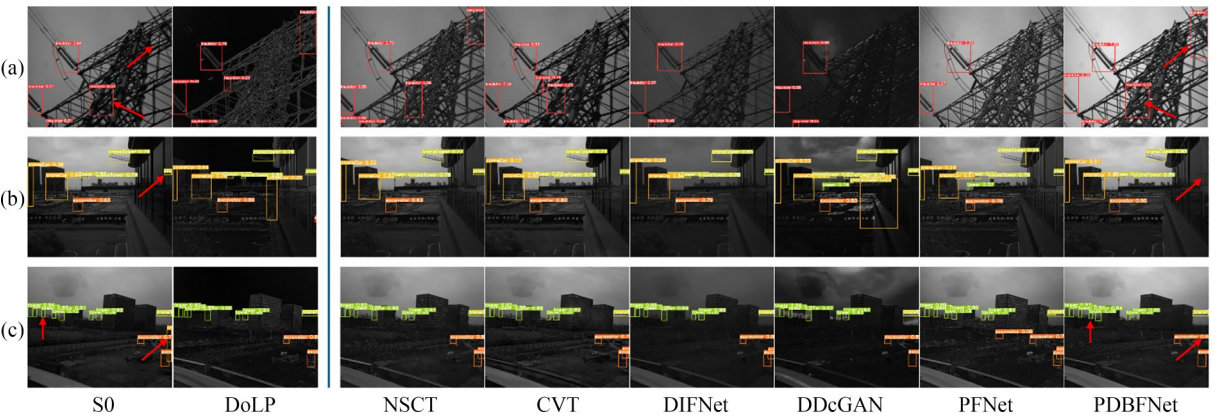


图 10 使用不同融合算法生成图的目标检测结果示例

Fig. 10 Example of target detection results using different fusion algorithms to generate graphs

4 结 论

针对输电线路巡检中复杂场景下的目标识别难题,本研究提出了一种基于光强-偏振双模态融合的目标检测框架,并创新性地设计了偏振图像融合网络 PDBFNet。

PDBFNet 网络采用双路融合模块(DFE)增强的密集卷积结构进行特征提取,有效提升了复杂背景和光照变化下的特征表达能力。通过优化多约束损失函数,进一步强化了网络对巡检目标的识别能力。为支撑算法研究,构建了多维光

学图像数据集 IPF,该数据集涵盖丰富的真实电网场景,为算法训练与评估提供了可靠的数据基础。

实验结果表明,PDBFNet 在多项评价指标上均优于现有先进算法,其生成的融合图像在保持纹理细节和边缘信息的同时,显著抑制了噪声干扰。在目标检测任务中,PDBFNet 实现了 91.5% 的识别精度和 0.916 的 mAP@0.5,在误检率和漏检率之间取得了良好平衡。

未来研究将引入红外等多模态信息,通过多维光学特征融合,进一步提升复杂场景下的目标

检测性能,推动输电线路智能巡检技术的发展。

作者贡献声明:

倪梦瑶:测量方法的提出,论文构思和撰写;

彭元龙:为了项目研究成果能够发表而去争取并获得资助;

胡尚:使用图像分析技术合成研究数据;

闫龙川:为研究活动的策划和执行进行管理和协调;

郑锦坤:为了项目研究成果能够发表而去争取并获得资助;

曹丹华:论文审核与编辑写作。

参考文献:

- [1] CHEN J L, FU Z J, CHENG X, *et al.* An method for power lines insulator defect detection with attention feedback and double spatial pyramid[J]. *Electric Power Systems Research*, 2023, 218: 109175.
- [2] TAO X, ZHANG D P, WANG Z H, *et al.* Detection of power line insulator defects using aerial images analyzed with convolutional neural networks[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, 50(4): 1486-1498.
- [3] WANG S S, ZOU X Y, ZHU W, *et al.* Insulator defects detection for aerial photography of the power grid using you only look once algorithm[J]. *Journal of Electrical Engineering & Technology*, 2023, 18(4): 3287-3300.
- [4] ZHENG W J, LI C Q, CUI X Y, *et al.* A multi-objective parallel detection algorithm for images of power transmission line corridors[C]. 2020 *IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*. June 12-14, 2020. Chongqing, China. IEEE, 2020.
- [5] HUANG Y Y, WANG J, SONG Y P, *et al.* A novel defect detection method with eliminating dust for specular surfaces based on structured-light modulation analysis technique[J]. *Optics & Laser Technology*, 2021, 141: 107089.
- [6] 李林, 张国清, 周战荣, 等. 偏振探测成像技术研究进展及其军事应用[J]. 火箭军工程大学学报, 2024, 38(5): 109-118.
LI L, ZHANG G, ZHOU Z, *et al.*. Progress and military applications of polarization detection imaging technology [J]. *Journal of Rocket Force University of Engineering*, 2024, 38(5): 109-118. (in Chinese)
- [7] REN Q M, XIANG Y F, WANG G C, *et al.* The underwater polarization dehazing imaging with a lightweight convolutional neural network [J]. *Optik*, 2022, 251: 168381.
- [8] SHENSA M J. The discrete wavelet transform: wedding the a trous and Mallat algorithms [J]. *IEEE Transactions on Signal Processing*, 1992, 40(10): 2464-2482.
- [9] ZHANG Q, GUO B L. Multifocus image fusion using the nonsubsampling contourlet transform[J]. *Signal Processing*, 2009, 89(7): 1334-1346.
- [10] HEDIFA D, FELLA C, ABDERRAZAK B. Multimodal medical image fusion using guided filter and curvelet transform[C]. 2022 *2nd International Conference on New Technologies of Information and Communication (NTIC)*. December 21-22, 2022, Mila, Algeria. IEEE, 2022: 1-5.
- [11] JI X X. An improved image fusion method of infrared image and sar image based on contourlet and sparse representation [C]. 2015 *7th International Conference on Intelligent Human-Machine Systems and Cybernetics*. August 26-27, 2015, Hangzhou, China. IEEE, 2015: 282-285.
- [12] LI Y H, LIU G, BAVIRISETTI D P, *et al.* Infrared-visible image fusion method based on sparse and prior joint saliency detection and LatLRR-FPDE [J]. *Digital Signal Processing*, 2023, 134: 103910.
- [13] KANNAN K. Application of partial differential equations in multi focused image fusion[J]. *International Journal of Advanced Networking and Applications*, 2022, 14(1): 5266-5270.
- [14] YANG Y, GAO C R, MING Z Q, *et al.* LatLRR-CNN: an infrared and visible image fusion method combining latent low-rank representation and CNN [J]. *Multimedia Tools and Applications*, 2023, 82(23): 36303-36323.
- [15] SU Y X, LIANG X, CAO D H, *et al.* Research on a multi-dimensional image information fusion al-

- gorithm based on NSCT transform [J]. *Frontiers of Optoelectronics*, 2024, 17(1): 4.
- [16] JUNG H, KIM Y, JANG H, *et al.* Unsupervised deep image fusion with structure tensor representations [J]. *IEEE Transactions on Image Processing*, 2020.
- [17] 陈永, 张娇娇, 王镇. 多尺度密集连接注意力的红外与可见光图像融合[J]. *光学精密工程*, 2022, 30(18): 2253-2266.
- CHEN Y, ZHANG J J, WANG Z. Infrared and visible image fusion based on multi-scale dense attention connection network [J]. *Opt. Precision Eng.*, 2022, 30(18): 2253-2266. (in Chinese)
- [18] ZHANG J C, SHAO J B, CHEN J L, *et al.* PF-Net: an unsupervised deep network for polarization image fusion [J]. *Optics Letters*, 2020, 45(6): 1507-1510.
- [19] MA J Y, YU W, LIANG P W, *et al.* Fusion-GAN: a generative adversarial network for infrared and visible image fusion [J]. *Information Fusion*, 2019, 48: 11-26.
- [20] MA J Y, XU H, JIANG J J, *et al.* DDcGAN: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4980-4995.
- [21] ZHANG H, FROMONT E, LEFEVRE S, *et al.* Guided attentive feature fusion for multispectral pedestrian detection [C]. 2021 *IEEE Winter Conference on Applications of Computer Vision (WACV)*. January 3-8, 2021, Waikoloa, HI, USA. IEEE, 2021: 72-80.
- [22] CHEN Y T, SHI J H, YE Z L, *et al.* *Multimodal Object Detection Via Probabilistic Ensembling* [M]. Computer Vision-ECCV 2022. Cham: Springer Nature Switzerland, 2022: 139-158.
- [23] 李校博. Stokes 矢量及 Mueller 矩阵的优化测量方法研究 [D]. 天津: 天津大学, 2020.
- LI X B. *Research on Optimal Measurement Method of Stokes Vector and Mueller Matrix* [D]. Tianjin: Tianjin University, 2020. (in Chinese)
- [24] RAMAN C V. Christiaan Huyghens and the wave theory of light [J]. *Proceedings of the Indian Academy of Sciences-Section A*, 1959, 49(4): 185-192.
- [25] GOLDSTEIN D H. *Polarized Light* [M]. Boca Raton: CRC Press, 2017.
- [26] SHEN Y, LIN W F, WANG Z F, *et al.* Rapid detection of camouflaged artificial target based on polarization imaging and deep learning [J]. *IEEE Photonics Journal*, 2021, 13(4): 7800309.
- [27] JOCHER G, STOKEN A, CHAURASIA A, *et al.* Ultralytics/yolov5: v6.0-YOLOv5n 'Nano' models, Roboflow integration, TensorFlow export, OpenCV DNN support [R]. *Zenodo*, 2021.
- [28] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 2261-2269.
- [29] HU J, SHEN L, SUN G. Squeeze-and-excitation Networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [30] SUN Y X, NI M Y, ZHAO M, *et al.* Low-light enhancement method with dual branch feature fusion and learnable regularized attention [J]. *Frontiers of Optoelectronics*, 2024, 17(1): 28.
- [31] BLIN R, AINOUS S, CANU S, *et al.* The PolarLITIS dataset: road scenes under fog [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 10753-10762.
- [32] VAN AARDT J. Assessment of image fusion procedures using entropy, image quality, and multispectral classification [J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 023522.
- [33] MOULDEN B, KINGDOM F, GATLEY L F. The standard deviation of luminance as a metric for contrast in random-dot images [J]. *Perception*, 1990, 19(1): 79-101.
- [34] PIELLA G, HEIJMANS H. A new quality metric for image fusion [C]. *Proceedings 2003 International Conference on Image Processing*. September 14-17, 2003, Barcelona, Spain. IEEE, 2003: III-173.
- [35] LIU J Y, FAN X, HUANG Z B, *et al.* Target-aware dual adversarial learning and a multi-scenario multi-modality benchmark to fuse infrared and visi-

ble for object detection [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recog-*

nition (CVPR). June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 5792-5801.

作者简介:



倪梦瑶(2000—),女,2022年于华中科技大学获得学士学位,现于华中科技大学攻读硕士学位,主要研究方向为目标检测、图像增强算法。
E-mail: m202272709@hust.edu.cn

通讯作者:



曹丹华(1966—),女,博士,教授,博士生导师,现就职于华中科技大学光学与电子信息学院,主要研究方向为光电传感和信号处理、机器视觉算法。
Email: dhcao@hust.edu.cn