doi:10.16112/j.cnki.53-1223/n.202512030001
引用格式:项导,施汉琴,鲍蓉,等. 变化注意力引导的变电站缺陷检测算法[J]. 昆明理工大学学报(自然科学版),2026,51(3):109-119.
Citation:XIANG Dao,SHI Hanqin,BAO Rong,et al. Change Attention Guided Defect Detection Algorithm for Substations[J]. Journal of Kunming University of Science and Technology(Natural Science),2026,51(3):109-119.

变化注意力引导的变电站缺陷检测算法

项 导[1],施汉琴[1],鲍 蓉[1],刘 羽[2]
(1. 徐州工程学院 信息工程学院,江苏 徐州 221018; 2. 合肥工业大学 仪器科学与光电工程学院,安徽 合肥 230009)

摘要: 变电站缺陷及时检测和处理是保障电力系统稳定运行的基础.针对现有变电站缺陷检测算法存在的背景干扰抑制不足、多尺度目标识别能力薄弱及双时相信息利用不充分等问题,本文提出一种变化注意力引导的变电站缺陷检测算法.该方法充分挖掘变电站双时相巡检图像中的时序差异信息,通过构建孪生骨干网络提取多尺度特征,并设计变化注意力引导模块,以双时相特征差异为注意力权重,动态增强缺陷区域的显著性,抑制背景干扰.同时,引入多尺度空洞特征提取模块与自适应特征融合结构,提升模型对不同尺度缺陷的检测能力.在自建的变电站双时相缺陷数据集上的实验结果表明,所提方法在平均精度均值、精确率、召回率等指标上显著优于 Faster RCNN 等主流检测模型.消融实验与热力图可视化分析进一步验证了变化注意力机制与特征融合策略的有效性,证明了该方法在复杂环境下具备良好的鲁棒性与工程应用潜力.

关键词: 双时相图像;变电站巡检;变化注意力;变电站缺陷检测;自适应特征融合
中图分类号:TP391.41;TM63 **文献标识码**:A **文章编号**:1007-855X(2026)03-0109-11

Change Attention Guided Defect Detection Algorithm for Substations

XIANG Dao[1],SHI Hanqin[1],BAO Rong[1],LIU Yu[2]
(1. School of Information Engineering,Xuzhou University of Technology,Xuzhou 221018,Jiangsu,China;
2. School of Instrument Science and Opto-electronics Engineering,Hefei University of Technology,Hefei 230009,China)

Abstract: Timely detection and handling of substation defects are fundamental to ensuring the stable operation of power systems. To address the issues in current substation defect detection algorithms,such as insufficient suppression of background interference,weak capability in recognizing multi-scale targets,and inadequate utilization of dual-temporal information,this paper proposes a change attention-guided substation defect detection algorithm. This method fully explores the temporal difference information in dual-temporal substation inspection images,extracts multi-scale features by constructing a Siamese backbone network,and designs a change attention guidance module. Taking the dual-temporal feature difference as attention weights,the module dynamically enhances the saliency of defect regions and suppresses background interference. Meanwhile,a multi-scale atrous feature extraction module and an adaptive feature fusion structure are introduced to improve the model' s capability in detecting defects of varying scales. Experimental results on a self-built bi-temporal substation defect dataset demonstrate that the proposed method is significantly superior to mainstream detection models such as Faster

收稿日期:2025-12-03. **基金项目**:国家自然科学基金项目(62576132).
作者简介:项导(1989-),男,博士,讲师. 主要研究方向:缺陷检测,图像显著性分析,深度学习.
E-mail:xiangdao@xzit.edu.cn

RCNN in terms of mean Average Precision, Precision, Recall and other indicators. Ablation experiments and heatmap visualization analysis further validate the effectiveness of the change attention mechanism and feature fusion strategy, proving that this method has good robustness and engineering application potential in complex environments.

Keywords: dual - temporal images; substation inspection; change attention; substation defect detection; adaptive feature fusion

0 引 言

作为电力系统“发、输、变、配、用”的核心枢纽,变电站的稳定运行关系到电网安全和国计民生. 受自然环境、高压工作负荷和人为破坏等因素影响,变电站容易出现表计破损、箱柜门打开等缺陷,影响设备正常工作,甚至造成停电[1]. 传统人工巡视机制存在工作效率低、漏检误检多等问题,难以满足变电站规模快速扩张下的运维需求[2]. 为提升巡视效率和运维质量,国家电网公司从“十二五”开始就将智能巡检机器人作为电网智能化的核心技术装备进行推广应用,构建“集控站 + 无人值守”的新型智能运维模式,推动远程智能巡视替代现场人工巡视[3].

目前,智能机器人、监控摄像机等巡检装备虽然已初步替代人工实现了远程巡视,但受限于缺乏高精度缺陷检测算法,现阶段缺陷分析仍大量依赖人工核查,成为阻碍全流程智能运维体系落地的关键瓶颈[4]. 中国计算机学会(CCF)在发布“电力视觉”术语时指出,现有的变电站缺陷检测研究大多基于通用目标检测框架,主要面向自然场景中常规目标检测问题进行设计和优化,未充分考虑变电场景的特殊性,存在模型泛化能力不足、可解释性欠缺及过拟合风险高等问题[5]. 变电站巡检装备在相同点位、不同时间采集的样本蕴含丰富的时相信息[6],缺陷在双时相样本中表现为区域特征的变化. 为充分利用巡检样本中的时相信息,本文在通用目标检测框架中嵌入变化注意力机制,以输入图像的特征差异为注意力权重,对巡检图像的特征图进行引导,增强存在特征变化的区域,并抑制变化较小的背景.

1 相关方法

1.1 变电站缺陷检测

作为工业缺陷检测的一个细分研究领域,变电站缺陷检测方法可以分为传统方法和基于深度学习的方法. 传统方法主要利用底层图像处理或传统机器学习来提取缺陷,仅适用于成像条件稳定、背景简单的场景. 基于深度学习的方法可以直接从原始图像中学习缺陷的特征表示,是当前变电站缺陷检测研究的主流方法,其从模型结构上可以分为二阶段(Two - stage)模型和一阶段(One - stage)模型.

二阶段模型将目标检测分为两步:首先生成区域候选框,再对每个候选框进行位置回归和类别分类. Faster RCNN[7]是采用此类框架的经典算法. Ying 等[8]在 Faster R - CNN 中引入类间采样、中心引导非极大值抑制和类别自适应阈值等改进策略,缓解了样本长尾分布和边界框冗余导致的漏检或误报问题. 张铭泉等[9]在 Faster R - CNN 网络中添加空间金字塔池化模块并改进特征融合方式,增强了对小目标物体和设备渗漏油的检测性能. 二阶段模型的优势在于检测准确率高,但计算复杂度高、检测速度较慢,难以满足实时性要求很高的应用场景.

一阶段模型不依赖候选框,直接通过骨干网络提取的特征预测缺陷的位置和类别. YOLO 系列模型[10]是该类框架的典型代表. Dong 等[11]在 YOLOv5 中嵌入并行混合注意力机制,引导模型在复杂背景中聚焦关键设备区域. Yuan 等[12]基于 YOLOv7 提出了一种新的多分支骨干和颈部网络融合框架,提高了对多尺度目标的检测性能. 一阶段模型具有较高的检测效率,但在复杂场景下的多尺度目标检测效果仍有待进一步提升.

上述一阶段和二阶段模型主要基于通用目标检测框架,通过嵌入注意力机制、优化损失函数等方式降低背景干扰、增强多尺度缺陷特征的提取能力,在特定类型的变电站缺陷检测上取得了一定进展,但随着变电站缺陷类型的多样化及检测场景的复杂化,现有方法仍存在以下局限:1) 在复杂背景下骨干网络特

征捕获能力不足,对相似目标的辨识能力有限,且周围环境的特征干扰容易影响缺陷目标的识别;2）在面对尺度差异较大、形状变化复杂的目标时,从颈部到检测头的输入特征表达能力明显不足,最终导致模型的检测性能不佳. 总体而言,当前研究侧重于直观感知类问题,未充分挖掘和利用变电领域知识,模型的性能难以满足实用化需求,亟需进行“专用化”定制.

1.2 变化检测

变化检测是指对同一地理位置、不同时间采集的图像序列进行像素级或特征级分析,以识别实体或现象状态差异的技术. 早期的变化检测主要通过对人工选取的特征进行对比分析得到变化检测结果,存在抗噪能力差、大量伪变化等问题. 随着深度学习技术的兴起,基于孪生网络的端到端变化检测方法已成为研究的主流.

孪生网络模型通过共享权重的子网络分别提取双时相输入的特征,然后利用特征的差值进行变化检测. Chen 等[13]提出时空关系建模机制,提高了变化信息提取的准确性. Feng 等[14]通过自注意力和交叉注意力机制增强特征之间的信息耦合,提升了对前景和背景类别分布不均衡数据集的变化检测性能.

变电站巡检装备在不同时间采集的样本之间具有时相关系,通过检测不同时相巡检样本的区域变化,可以辅助检测缺陷. 近年来,已经有研究开始关注时相信息在变电站缺陷检测领域的应用[15-16],但总体而言,相关的研究仍然偏少,时相特征的价值亟待深入挖掘与工程化利用.

2 本文方法

本节详细介绍所提出的变化注意力引导的变电站缺陷检测算法 CAGNet. 如图 1 所示,CAGNet 由变化编码器、变化解码器和缺陷检测头构成. 首先,将变电站巡检装备在相同点位、不同时间采集的双时相巡检图像,即不含缺陷的参考图像 $T_1 \in R^{H\times W\times 3}$ 与可能含有缺陷的待检测图像 $T_2 \in R^{H\times W\times 3}$,经像素级配准[17]后输入变化编码器,提取多尺度全局上下文特征信息并强化缺陷变化区域的特征表达,随后的变化解码器使用类 FPN(Feature Pyramid Network)结构[18],通过上采样操作融合高层和底层特征信息,生成预测特征图. 最后将预测特征图输入到缺陷检测头中,进行缺陷的定位和分类.

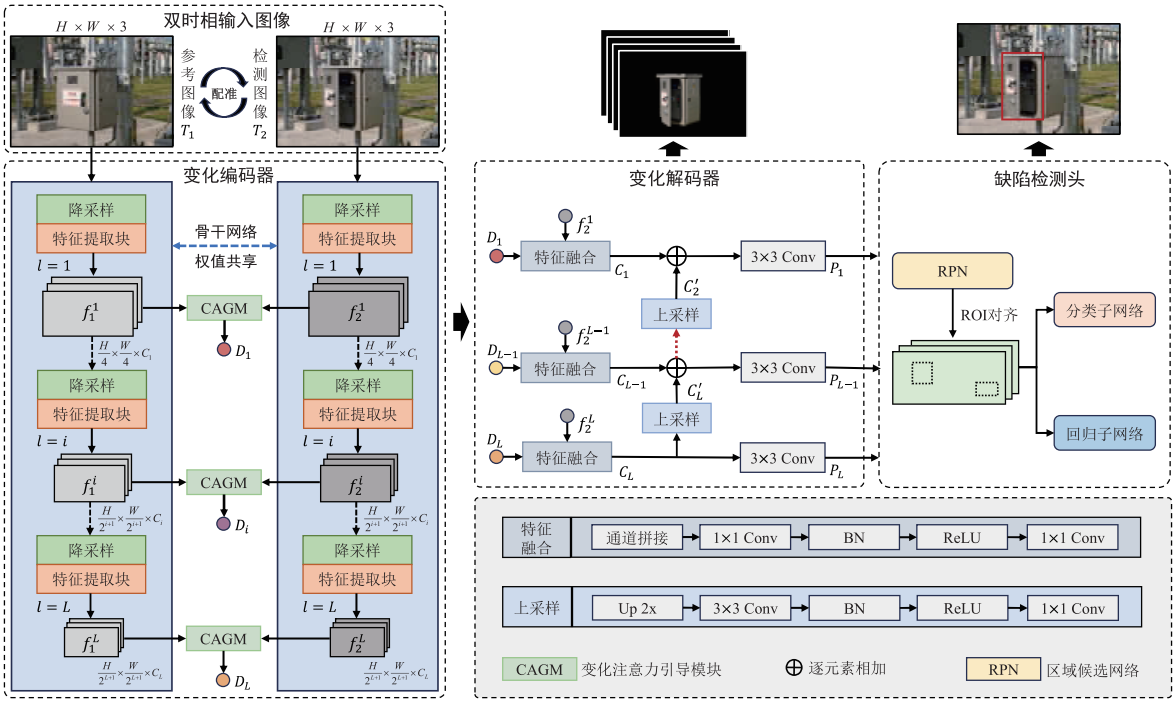


图 1　CAGNet 网络整体结构图

Fig. 1　The framework of CAGNet

2.1 变化编码器

变化编码器由共享权重的孪生骨干网络和变化注意力引导模块(CAGM)组成. 孪生骨干网络由 L 层降采样和特征提取块构成,负责从双时相巡检图像中提取多尺度缺陷特征. CAGM 模块负责对孪生骨干网络提取的各层特征进行融合和映射.

2.1.1 孪生骨干网络

鉴于变电站缺陷图像具有背景复杂、特征多样和尺度多变的特点,为了更好地捕捉缺陷的细节和整体结构,本文采用 Noman 等[19]提出的高效局部 - 全局上下文聚合(ELGCA)模块来构建骨干网络.

结合变电站缺陷图像的尺度先验信息,本文采用 $L=4$ 层骨干网络来提取多尺度缺陷特征,各层特征图相对于输入图像的缩小倍数分别为 1/4、1/8、1/16 和 1/32,相应的通道数分别为 64、96、128 和 256.

2.1.2 变化注意力引导模块(CAGM)

在变电站场景中,背景干扰是造成缺陷漏检或误检的重要因素[6]. 在孪生骨干网络提取的多尺度特征中,背景区域的特征相似性很强,变化区域的特征相似性则较低,而缺陷正处于变化区域中. 为了抑制背景干扰,本文设计了变化注意力引导模块(Change Attention Guidance Module, CAGM),如图 2 所示. CAGM 模块以孪生骨干网络提取的第 i 层参考图像特征图 f_1^i 和巡检图像特征图 f_2^i 为输入,首先对输入特征图执行绝对差操作,以突出相邻区域的差异性、强调变化区域. 该过程可以用公式(1)表示:

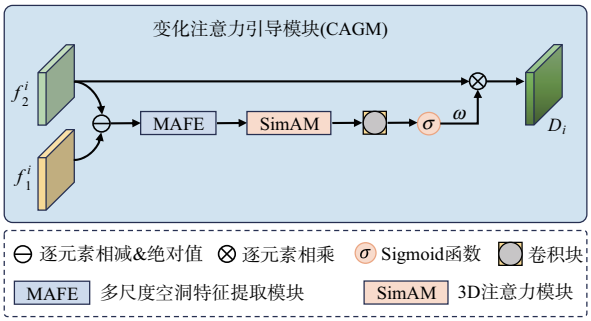


图 2 变化注意力引导模块结构图

Fig. 2 Change attention guidance module structure diagram

$$f_{\text{diff}}^i = |f_1^i - f_2^i| \tag{1}$$

其中:$|\cdot|$ 代表绝对值,f_{diff}^i 表示输出的初始差异特征. 接下来,将 f_{diff}^i 分别经过多尺度空洞特征提取模块(Multiscale Atrous Feature Extraction Module, MAFE)、3D 注意力模块 SimAM[20] 和卷积块处理,进一步增强差异特征信息.

如图 3 所示,MAFE 模块由负责提取多尺度特征的多分支结构以及用于融合多分支特征的 1×1 卷积构成. 结合变电站缺陷的尺度先验及孪生骨干网络输出的多层级特征图尺寸,在多分支结构中设计 4 个并行空洞卷积分支来扩展感受野,各分支由不同空洞率的空洞卷积混合构成,有助于减轻网格效应(gridding effect)[21]. MAFE 提取的特征经通道拼接后,送入 1×1 卷积进行多尺度特征聚合. 该设计基于以下考虑:1) 变电站缺陷目标在图像中呈现显著尺度差异,如小尺度的绝缘子破损至大尺度的箱门闭合异常. 多分支感受野设计旨在覆盖从局部细节到全局结构的特征提取需求;2) 骨干网络输出的特征图尺寸逐层递减,其有效感受野逐层增大. MAFE 模块通过递增的感受野与不同层级的特征图形成尺度对齐,增强多尺度缺陷特征的提取能力. MAFE 的处理过程可以用公式(2)表示为:

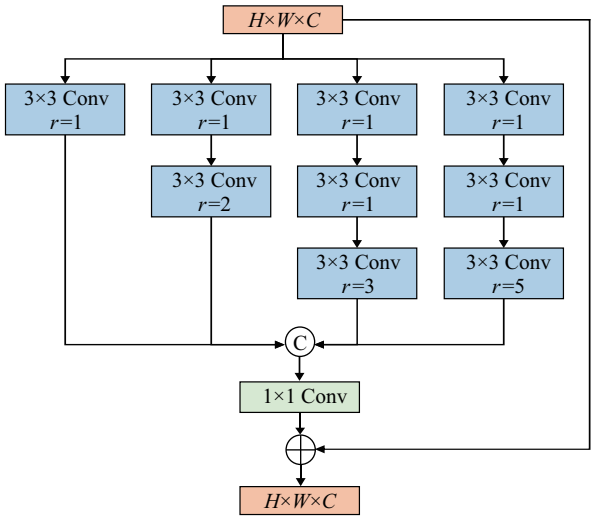


图 3 多尺度空洞特征提取模块结构图

Fig. 3 Structure diagram of the multiscale atrous feature extraction module

$$F^i_{\mathrm{MAFE}} = \mathrm{Conv}_1\left(\mathop{\mathrm{Cat}}_{k=1}^{4}(\mathrm{Branch}_k(f^i_{\mathrm{diff}}))\right) + f^i_{\mathrm{diff}} \tag{2}$$

其中:Conv_1 表示 1×1 卷积,Cat 表示通道拼接操作;$\mathrm{Branch}_{k\in\{1,2,3,4\}}$ 表示第 k 个空洞卷积分支,有助于提高模型在细节识别和局部特征表达上的能力.

差异特征经 MAFE 扩大感受野后,再送入 SimAM 建模每个像素点在局部邻域中的相对重要性. SimAM 是一种轻量级、无参数化的 3D 注意力模块,能够动态建模特征间的空间相关性和对比性,并自适应地调整特征响应,从而识别出对任务最为重要的区域和细节特征,提高变化区域的检测准确性. 最后,经卷积块进一步增强差异信息后,送入 Sigmoid 函数生成注意力权重 ω 并应用到原始特征图 f^i_2 上,使模型更加精准地聚焦于缺陷信息,提高特征表达的效率与准确性. 该过程可以表示为:

$$D_i = \sigma(\mathrm{Conv}_3(\mathrm{SimAM}(F_{\mathrm{MAFE}}))) \otimes f^i_2 \tag{3}$$

其中:Conv_3 表示由 3×3 卷积、批量归一化(Batch Normalization)和 ReLU 激活函数构成的卷积块,σ 代表 Sigmoid 激活函数,$\otimes$ 代表逐元素相乘操作. 最终,经 CAGM 模块增强缺陷语义特征、抑制无关背景干扰后,变化编码器输出多尺度缺陷特征图 $D_{i\in\{1,2,3,4\}}$,作为变化解码器模块的输入.

2.2 变化解码器

变化解码器的核心任务是对不同层级的缺陷信息进行自适应融合,实现信息互补与特征增强,采用类似 FPN 的金字塔结构,以自底向上 + 横向连接的方式将高层级的语义信息和低层级的细节信息进行深度融合,促进缺陷特征的高效解码,为后续的缺陷检测提供更丰富的特征支持.

以第 i 层和第 $i+1$ 层的特征融合为例,首先将编码器输出的特征 D_i 和 D_{i+1} 分别送入特征融合模块处理得到 C_i 和 C_{i+1},然后将 C_i 和经上采样模块处理的 C_{i+1} 逐像素相加,实现相邻层级特征的融合,最后再经 3×3 卷积强化特征的表达能力. 该过程可以表示为:

$$C_i = \mathrm{Fuse}(D_i, f^i_2) \tag{4}$$

$$C_{i+1} = \mathrm{Fuse}(D_{i+1}, f^{i+1}_2) \tag{5}$$

$$P_i = \mathrm{Conv}_3(C_i + \mathrm{Up}(C_{i+1})) \tag{6}$$

其中,Fuse(·)代表特征融合模块,其主要由通道拼接和 1×1 卷积构成. 考虑到 CAGM 模块在双时相特征交互过程中可能会损失一部分缺陷细节信息,本文设计特征融合模块,将编码器输出的蕴含双时相差异信息的特征 D_i 和骨干网络提取的保留原始细节信息的特征 f^i_2 进行融合. 具体地,特征融合模块首先通过通道拼接操作将 D_i 和 f^i_2 在通道维度上拼接,形成一个综合特征图 $[D_i; f^i_2]$,随后,应用多个 1×1 卷积对拼接后的特征进行通道整合和特征变换,通过自动学习各通道的权重,实现特征的自适应融合. 该机制旨在补充缺陷细节信息,提高缺陷局部定位的准确性及模型在复杂场景下的鲁棒性. Up(·)代表上采样操作,主要由双线性插值、3×3 卷积和 1×1 卷积等构成,实现高级特征和低级特征在空间尺度上的对齐. 需要指出的是,对于最底层特征 D_4 而言,经过特征融合模块和 3×3 卷积处理后,直接输出增强特征 P_4. 最终,变化解码器输出多尺度增强特征 $P_{i\in\{1,2,3,4\}}$,作为缺陷检测头的输入.

3 实验结果和分析

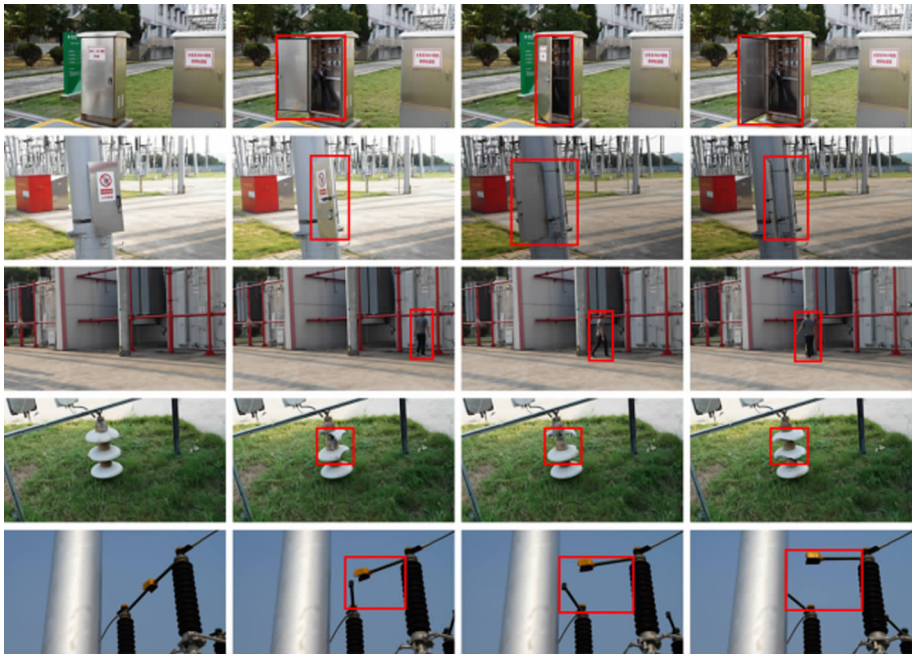
3.1 实验环境

本文实验在配备 2 块 NVIDIA RTX 4090 24 GB GPU 和 1 颗 Intel Xeon W5 - 2455x 3.2 GHz CPU 的工作站上开展,操作系统为 64 位 Ubuntu 22.04 LTS,代码基于 PyTorch 2.2.1 深度学习框架. 模型训练选择 AdamW 优化器,设置 $\beta_1 = 0.5$,$\beta_2 = 0.999$,权重衰减系数设置为 5e - 4,批量大小设置为 8,迭代次数设置为 200 轮,编码器网络通过加载预训练权重[19]进行初始化,其余网络采用 Kaiming 初始化[22]. 学习率由余弦退火调度函数调整,初始学习率设置为 0.000 1. 双时相输入图像尺寸统一调整为 640×640×3.

3.2 数据集说明

本文采用变电站巡检机器人以及固定式监控摄像机,拍摄了大量巡检图像,经人工筛选和图像配准[17]处理后,整理得到变电站双时相缺陷样本集 SDDD(Substation Dual - phase Defect Dataset),总共包含

1 915 组巡检图像,如图 4 所示. 每组图像由相同点位、不同时间采集的 1 张正常图像和 3 张缺陷图像构成. 数据集按 9∶1 划分为训练集和测试集,划分过程中保持各类别占比一致. 样本集覆盖箱门闭合异常(1 284 对)、异物入侵(912 对)、绝缘子破损(480 对)、表计破损(639 对)、硅胶桶破损(546 对)、开关分合变化(1 056 对)和设备装置位置变化(828 对)共 7 种常见类型.



注:第一列为正常图像,其余三列为缺陷图像

图 4 变电站双时相缺陷样本集示例图

Fig. 4 Examples of the dual – temporal substation defect dataset(The first column shows normal images, and the remaining three columns show defect images)

3.3 性能对比实验

为了验证本文方法的有效性,本文使用精确率(Precision)、召回率(Recall)、平均精度(Average Precision,AP),以及平均精度均值(mean Average Precision,mAP)作为模型性能的评价指标[16]. 在变电站双时相缺陷样本集 SDDD 上与 5 种主流的目标检测算法进行性能对比,包括一阶段模型 YOLOX[10]、YOLOv8m[23]、YOLOv10m[24]和 Gold – YOLO[25],二阶段模型 Faster RCNN[7].

为了适配变电站缺陷检测任务,对上述算法进行两种改进:1)单输入模式(Conventional Mode):将双时相输入图像在通道维度拼接,作为模型输入,用“ – c”表示;2)双输入模式(Dual – temporal Mode):基于公式(1),利用孪生骨干网络计算绝对差异特征,作为后续缺陷检测网络的输入,用“ – d”表示. 模型训练的实验环境和超参数设置与本文方法一致.

如表 1 所示,给出了本文所提出的 CAGNet 与其他对比方法的实验结果. 易见,基于孪生网络的方法普遍优于基于通道拼接的方法,表明采用孪生网络提取的缺陷特征具有更强的表达能力,验证了利用双时相信息抑制背景干扰、提升检测性能的有效性. 在上述 5 种对比算法中,Faster RCNN – d 的 *mAP* 指标最高,相较于 YOLOX – d 领先幅度最高达 3.72%,且 Faster RCNN 算法在从单输入切换到双输入的性能提升幅度取得最大值 1.63%,表明二阶段模型能更加有效地应对背景复杂、特征和尺度多样的变电站巡检场景,这也是本文采用二阶段框架构建网络模型的重要原因.

本文方法在整体性能指标上达到了最高精度,*mAP* 为 80.62%,与 Faster RCNN – d 相比提升 2.65%,不仅在绝缘子破损、硅胶桶破损等变化特征相对微弱的类别上提升显著,在箱门闭合异常、异物入侵等传统方法较易识别的类别上,本文方法也均达到最高水平. 本文提出的变化注意力机制和特征融合机制,通过显式建模特征差异并引导特征强化,显著超越了其他算法在双时相输入模式下的性能.

为全面评估所提方法在变电站缺陷检测任务中的综合性能,本文进一步对比了不同模型的精确率(Precision)、召回率(Recall)及其衍生指标漏检率(1 - Recall)和误检率(1 - Precision),如表 2 所示. 实验结果表明,本文方法在精确率和召回率上均显著优于对比算法. 其中,本文方法的精确率较 YOLOX - d、Faster RCNN - d 分别提升 1.24% 和 2.95%,表明本文方法在正确识别缺陷、减少背景误判方面表现更优,可有效降低人工复核过程中对非缺陷区域的无效核查成本;本文方法的召回率较次优的 Faster RCNN - d 提升 5.86%,体现了模型对真实缺陷的高覆盖性,这对变电站运维中避免缺陷漏判引发安全风险具有重要意义.

表 1　与先进检测算法性能对比

Tab. 1　Performance comparison with state - of - the - art detection algorithms

模型	Backbone	*AP*/%							*mAP*/%	*mAP* 提升幅度/% d - c
		箱门闭合异常	异物入侵	绝缘子破损	表计破损	硅胶桶破损	开关分合变化	设备装置位置变化		
YOLOX - c	DarkNet53	76.57	73.81	74.04	64.82	66.39	76.15	80.72	73.21	1.04
YOLOX - d		78.12	75.20	74.83	65.75	67.81	76.87	81.14	74.25	
YOLOv8m - c	CSPDarkNet	77.65	75.24	70.51	68.98	66.32	81.24	83.05	74.86	1.31
YOLOv8m - d		79.80	76.12	71.27	71.63	69.25	81.61	83.48	76.17	
YOLOv10m - c	CSPDarkNet	80.32	74.61	77.75	71.45	72.13	78.62	82.37	76.75	0.71
YOLOv10m - d		81.76	74.83	78.05	72.19	73.04	79.22	83.16	77.46	
Gold - YOLOm - c	EfficientRep	81.14	78.92	76.31	67.25	72.24	79.83	81.49	76.74	1.06
Gold - YOLOm - d		81.51	79.54	77.12	70.30	72.84	80.72	82.58	77.80	
Faster RCNN - c	Resnet50	80.06	74.21	76.14	72.24	70.35	80.12	81.27	76.34	**1.63**
Faster RCNN - d		81.32	78.64	76.56	**73.17**	71.48	81.93	82.71	77.97	
本文方法	ELGCNet	**84.96**	**81.71**	**79.25**	73.08	**75.82**	**83.91**	**85.63**	**80.62**	-

为了更全面地评估本文方法,将本文方法与上述 5 种主流缺陷检测算法进行了模型复杂度对比分析,包括模型浮点运算数(*FLOPs*)、参数规模(*Params*)和帧率(*FPS*),结果如表 3 所示. 从结果可见,YOLO 系列模型因其端到端的设计,在浮点运算数和帧率等指标上展现出显著优势,这种特性使得 YOLO 系列模型在算力资源有限、对实时性要求高的边缘端设备部署场景中,具有更强的适用性.

表 2　不同模型的精确率和召回率对比

Tab. 2　Comparison of precision and recall among different models　%

方法	精确率	召回率	漏检率	误检率
YOLOX - d	74.86	72.54	27.46	25.14
YOLOv8m - d	72.42	77.28	22.72	27.58
YOLOv10m - d	71.01	76.03	23.97	28.99
Gold - YOLOm - d	72.05	76.57	23.43	27.95
Faster RCNN - d	73.15	77.87	22.13	26.85
本文方法	**76.10**	**83.73**	**16.27**	**23.90**

本文方法在两阶段检测框架中引入了 ELGCA 骨干网络以及变化注意力机制,计算复杂度高于其他对比模型,帧率仅为 21.2FPS,在所有参与对比的算法中处于最低水平. 但需要指出的是,尽管 YOLO 系列一阶段模型在计算效率和实时性方面表现出色,但在算法性能上与本文方法仍存在较大差距. 在实际的变电站远程智能巡视系统应用场景中,更看重算法的高性能表现,以确保能够准确地检测出变电站设备缺陷,降低漏检、误检引发的运维风险. 因此,尽管本文方法在算法效率方面不占优势,但从系统整体性能和实际应用需求的角度来看,其更适合部署在智能分析主机上.

表 3　计算复杂度对比结果

Tab. 3　Comparison of computational complexity

模型	*FLOPs*/G	*Params*/M	*FPS*
YOLOX - d	114.6	25.3	**96.8**
YOLOv8m - d	104.2	25.8	89.3
YOLOv10m - d	**84.5**	**15.8**	95.9
Gold - YOLOm - d	115.8	41.2	94.9
Faster RCNN - d	218.8	41.3	23.8
本文方法	227.2	31.9	21.2

3.4 消融实验

为了验证本文方法中不同改进的效果,设计了消融实验,针对基线模型 Faster RCNN – d(骨干网络为 Resnet50,采用绝对差异特征,解码器采用标准 FPN 设计,即将特征融合模块替换为 1×1 卷积)逐步引入特定的改进,以分析每个改进的独立贡献,以及不同改进之间的相互作用.

实验结果如表 4 所示,改进 1 方法将基线模型的 ResNet50 骨干网络替换为 ELGCNet,*mAP* 提升 0.37%,表明 ELGCNet 相比传统的 ResNet50 具有更强的全局特征提取能力. 改进 2 方法在改进 1 的基础上将绝对差异特征替换为变化注意力机制,*mAP* 提升明显,其增益显著超过骨干网络改进,表明变化注意力机制是性能提升的关键驱动力,可以有效放大缺陷区域的特征差异,同时抑制背景干扰. 改进 3 方法在改进 1 的基础上引入特征融合模块,实现 *mAP* 提升 0.52%,表明本文设计的特征融合模块与标准 FPN 相比能提取更丰富的缺陷细节特征,从而增强模型在复杂场景下的缺陷定位准确性与鲁棒性. 改进 4 方法在改进 3 的基础上,将骨干网络恢复为 ResNet50,同时差异特征采用变化注意力机制,*mAP* 提升 1.12%,进一步印证了注意力机制在模型中的不可替代性.

表 4 消融实验结果

Tab. 4 Results of ablation experiment

方法	骨干网络		差异特征		解码器		*mAP*/%
	Resnet50	ELGCNet	绝对差异特征	变化注意力机制	1×1 卷积	特征融合模块	
基线模型	√		√		√		77.97
改进 1		√	√		√		78.34
改进 2		√		√	√		79.93
改进 3		√	√			√	78.86
改进 4	√			√		√	79.98
本文方法		√		√		√	**80.62**

为进一步验证变化注意力引导模块(CAGM)中各组件的独立贡献,本文针对多尺度空洞特征提取模块(MAFE)、SimAM 注意力模块等设计了专项消融实验,结果如表 5 所示. 实验数据显示,当移除 SimAM 模块(对比本文方法与改进 6)时,模型 *mAP* 从 80.62% 降至 80.01%,性能下降 0.61%,表明 SimAM 通过动态建模特征间的空间相关性与对比性,能够有效增强缺陷区域的特征响应,其轻量级无参数化的特性使其在复杂场景下仍能保持高效计算. 对于 MAFE 模块,移除该结构(对比本文方法与改进 5)导致 *mAP* 下降 0.98%,验证了多尺度空洞卷积分支通过感受野扩展对捕捉不同尺度缺陷特征的关键作用,其多分支感受野的设计能有效提升模型对多尺度缺陷的检测鲁棒性. 卷积块的设计有助于进一步强化差异特征的语义表达,移除该模块(对比本文方法与改进 7)将导致 *mAP* 下降 0.25%. 如果将 MAFE、SimAM 和卷积块均移除(对比本文方法与基线模型),*mAP* 则大幅下降 2.61%,这验证了多组件联合优化的必要性及模块设计的合理性. MAFE、SimAM 和卷积块通过尺度适配、局部聚焦和特征增强的互补机制,共同实现了差异特征提取与强化的核心功能.

表 5 CAGM 模块消融实验结果

Tab. 5 Results of CAGM ablation experiments

方法	MAFE	SimAM	卷积块	*mAP*/%
基线模型				78.01
改进 5		√	√	79.64
改进 6	√		√	80.01
改进 7	√	√		80.37
本文方法	√	√	√	80.62

为进一步验证 MAFE 模块 4 分支设计的合理性,本文设计了对照消融实验. 各分支感受野按等差递增方式设计,实验结果如表 6 所示. 实验结果表明,4 分支设计在检测精度与计算效率之间达到了最佳平衡. 3 分支方案因缺少感受野为 15 的大尺度分支,无法有效覆盖箱门闭合异常等大尺度缺陷的特征提取需求,*mAP* 下降 0.47%. 5 分支和 6 分支方案新增的更大感受野分支超出了变电站缺陷的实际尺度分布范围,不仅未带来检测精度的提升,反而因参数冗余增加了模型过拟合风险,同时带来推理效率的下降,验证了 MAFE 模块 4 分支设计的合理性.

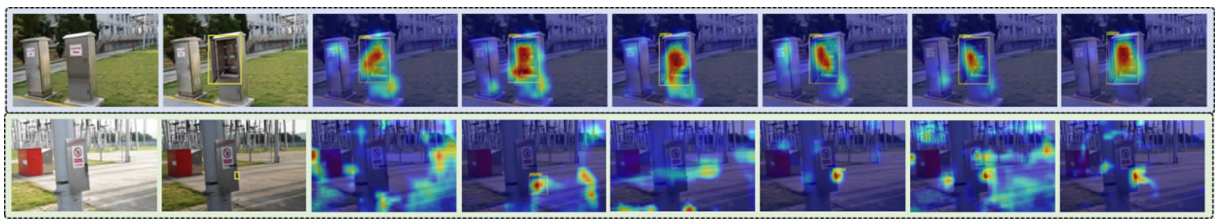
表 6　MAFE 模块消融实验结果

Tab. 6　Results of MAFE ablation experiments

分支数量	感受野组合	*mAP*/%	FPS
3 分支	3,7,11	80.15	21.9
4 分支	3,7,11,15	80.62	21.2
5 分支	3,7,11,15,19	80.54	20.7
6 分支	3,7,11,15,19,23	80.51	20.1

3.5 检测结果可视化

为进一步验证本文方法在变电站缺陷检测任务中的优越性,选取典型场景下的代表性样本图像,对本文方法及对比模型的缺陷检测结果及激活热力图[26]进行可视化,结果如图 5 所示.



(a)　(b)　(c)　(d)　(e)　(f)　(g)　(h)

(a)参考图;(b)标注真实值的检测图;(c)YOLOX - d;(d)YOLOv8m - d;(e)YOLOv10m - d;
(f)Gold - YOLOm - d;(g)Faster RCNN - d;(h)本文方法

图 5　不同方法测试结果可视化

Fig. 5　Visualization of test results of different models: (a) reference image; (b) inspection image with Ground Truth annotation; (c) YOLOX - d; (d) YOLOv8m - d; (e) YOLOv10m - d; (f) Gold - YOLOm - d; (g) Faster RCNN - d; (h) ours

在大尺度缺陷目标检测场景中(如第一行样本所示),本文方法与对比模型均能有效识别缺陷目标. 然而,YOLOX 等对比模型在背景区域也存在较高的响应值,表明模型对背景干扰的抑制能力较弱,而本文方法的热力图则集中分布于缺陷目标区域,且定位框与真实值更为接近,体现出更优的目标定位精度.

在小尺度缺陷目标检测场景中(如第二行样本所示),受光线变化等因素干扰,YOLOX 及 YOLOv10m 模型均出现漏检现象. 通过激活热力图分析可见,上述对比模型的响应区域主要集中于图像亮度变化剧烈的区域,未能准确聚焦小缺陷. 相比之下,本文方法通过变化注意力引导策略和多尺度特征融合机制,其热力图在小缺陷区域形成明显峰值响应,验证了本文方法在小目标缺陷检测中的鲁棒性.

总体而言,本文方法在复杂变电站场景下不仅能够有效抑制背景干扰,提升大尺度缺陷目标的定位精度,还能通过增强对微小缺陷的特征感知能力,显著改善小目标检测性能,展现出更全面的缺陷检测能力.

4 结　论

针对变电站缺陷检测中背景干扰强、小目标漏检率高、时相信息利用不充分等挑战,本文提出了一种变化注意力引导的缺陷检测模型 CAGNet. 通过引入变化注意力机制、优化特征提取与融合网络,该模型实

现了缺陷区域响应的有效放大,提升了缺陷检测的精度与鲁棒性. 实验结果表明,CGANet 在 SDDD 数据集上 *mAP* 达到 80.62%,显著优于 Faster RCNN 等主流检测模型,在复杂场景下展现出优异的检测性能.

本文的核心创新点体现在两个方面:一是设计变化注意力引导模块,将双时相特征差异显式建模为注意力权重,通过多尺度空洞特征提取与 SimAM 注意力协同增强缺陷显著性;二是构建 ELGCA 孪生骨干网络与自适应特征融合结构,在实现长程依赖建模的同时弥补细节信息损失.

本研究推动了变电站缺陷检测从单时相感知向双时相变化推理转变,适用于变电站后端智能分析主机,具备良好工程应用潜力. 未来将开展基于知识蒸馏的模型轻量化研究,实现算法在边缘设备的高效部署.

参考文献:

[1] 罗悦冬,杜思超,李猷,等. 基于深度特征感知的变电站设备典型缺陷检测方法[J/OL]. 南方电网技术,2025 (2025-03-28). https://link.cnki.net/urlid/44.1643.TK.20250328.1131.003.
LUO Y D,DU S C,LI Y,et al. Typical defect detection of substation equipment based on deep feature perception[J/OL]. Southern Power System Technology,2025 (2025-03-28). https://link.cnki.net/urlid/44.1643.TK.20250328.1131.003.

[2] 白艳峰,王立彪,高卫东,等. 面向电力设备缺陷检测的多模态层次化分类[J]. 中国图象图形学报,2024,29(7):2011-2023.
BAI Y F,WANG L B,GAO W D,et al. Multi-modal hierarchical classification for power equipment defect detection[J]. Journal of Image and Graphics,2024,29(7):2011-2023.

[3] 秦佳鸣,梁颖,金继开. 变电站智能巡检系统实用化应用分析[J]. 电力与能源,2024,45(3):389-391.
QIN J M,LIANG Y,JIN J K. Practical application analysis of intelligent substation inspection system[J]. Power & Energy,2024,45(3):389-391.

[4] 齐冬莲,韩译锋,周自强,等. 基于视频图像的输变电设备外部缺陷检测技术及其应用现状[J]. 电子与信息学报,2022,44(11):3709-3720.
QI D L,HAN Y F,ZHOU Z Q,et al. Review of defect detection technology of power equipment based on video images[J]. Journal of Electronics & Information Technology,2022,44(11):3709-3720.

[5] 赵振兵,冯烁,席悦,等. 大模型时代:电力视觉技术新起点[J]. 高电压技术,2024,50(5):1813-1825.
ZHAO Z B,FENG S,XI Y,et al. The era of large models:A new starting point for electric power vision technology[J]. High Voltage Engineering,2024,50(5):1813-1825.

[6] 陈鹏辉. 变电站场景下的图像变化检测技术研究[D]. 杭州:浙江大学,2021.
CHEN P H. Research on image change detection technology in substation scene[D]. Hangzhou:Zhejiang University,2021.

[7] REN S Q,HE K M,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.

[8] YING Y,WANG Y Z,YAN Y F,et al. An improved defect detection method for substation equipment[C]//2020 39th Chinese Control Conference (CCC),27-29 July 2020,Shenyang,China. IEEE,2020:6318-6323.

[9] 张铭泉,邢福德,刘冬. 基于改进 Faster R-CNN 的变电站设备外部缺陷检测[J]. 智能系统学报,2024,19(2):290-298.
ZHANG M Q,XING F D,LIU D. External defect detection of transformer substation equipment based on improved Faster R-CNN[J]. CAAI Transactions on Intelligent Systems,2024,19(2):290-298.

[10] GE Z,LIU S T,WANG F,et al. YOLOX:Exceeding YOLO series in 2021[EB/OL]. 2021:arXiv:2107.08430. https://arxiv.org/abs/2107.08430.

[11] DONG H,YUAN M,WANG S,et al. PHAM-YOLO:A parallel hybrid attention mechanism network for defect detection of meter in substation[J]. Sensors,2023,23(13):6052.

[12] YUAN J H,WANG F L,HU Q L,et al. The substation defect detection method based on FEFPN-YOLOv7[C]//2024 International Conference on Image Processing,Computer Vision and Machine Learning (ICICML),22-24 November 2024,Shenzhen,China. IEEE,2024:419-423.

[13] CHEN H R X,SONG J,HAN C X,et al. ChangeMamba:Remote sensing change detection with spatiotemporal state space model[J]. IEEE Transactions on Geoscience and Remote Sensing,2024,62:1-20.

[14] FENG Y C,JIANG J W,XU H H,et al. Change detection on remote sensing images using dual - branch multilevel intertemporal network[J]. IEEE Transactions on Geoscience and Remote Sensing,2023,61:4401015.

[15] 王鹏,赵春晖,周华良,等. 基于多时相巡检图像的变电设备抗干扰缺陷检测[J]. 控制与决策,2024,39(3):885 - 892.
WANG P,ZHAO C H,ZHOU H L,et al. Anti - interference defect detection of substation equipment based on multi - temporal inspection images[J]. Control and Decision,2024,39(3):885 - 892.

[16] 赵振兵,马迪雅,石颖,等. 基于改进 YOLOX 的变电站仪表外观缺陷检测算法[J]. 图学学报,2023,44(5):937 - 946.
ZHAO Z B,MA D Y,SHI Y,et al. Appearance defect detection algorithm of substation instrument based on improved YOLOX[J]. Journal of Graphics,2023,44(5):937 - 946.

[17] LINDENBERGER P,SARLIN P E,POLLEFEYS M. LightGlue:Local feature matching at light speed[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV),1 - 6 October 2023,Paris,France. IEEE,2023:17581 - 17592.

[18] HU M,LI Y L,FANG L,et al. A^2 - FPN:Attention aggregation based feature pyramid network for instance segmentation [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),20 - 25 June 2021,Nashville,TN,USA. IEEE,2021:15338 - 15347.

[19] NOMAN M,FIAZ M,CHOLAKKAL H,et al. ELGC - net:Efficient local - global context aggregation for remote sensing change detection[J]. IEEE Transactions on Geoscience and Remote Sensing,2024,62:1 - 11.

[20] YANG L X,ZHANG R Y,LI L D,et al. Simam:A simple,parameter - free attention module for convolutional neural networks [C]//Proceedings of the 38th International conference on machine learning. PMLR,2021:11863 - 11874.

[21] 李志财,龙军,钱育蓉,等. 基于差异感知增强的遥感图像耕地变化检测网络[J]. 计算机工程与应用,2026,62(4):224 - 237.
LI Z C,LONG J,QIAN Y R,et al. Difference - aware enhancement network for cropland change detection in remote sensing images[J]. Computer Engineering and Applications,2026,62(4):224 - 237.

[22] HE K M,ZHANG X Y,REN S Q,et al. Delving deep into rectifiers:Surpassing human - level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision (ICCV),07 - 13 December 2015,Santiago,Chile. IEEE,2015:1026 - 1034.

[23] SOHAN M,THOTAKURA S R,REDDY V R. A review on YOLOv8 and its advancements[C]//International Conference on Data Intelligence and Cognitive Informatics. SingaporeSpringer Nature Singapore2024:529 - 545.

[24] WANG A,CHEN H,LIU L H,et al. YOLOv10:Real - time end - to - end object detection[C]//Advances in Neural Information Processing Systems 37,10 - 15 December 2024,Vancouver,BC,Canada. Neural Information Processing Systems Foundation,Inc. (NeurIPS),2024:107984 - 108011.

[25] WANG C C,HE W,NIE Y,et al. Gold - YOLO:Efficient object detector via gather - and - distribute mechanism[C]//Advances in Neural Information Processing Systems 36,10 - 16 December 2023,New Orleans,Louisiana,USA. Neural Information Processing Systems Foundation,Inc. (NeurIPS),2023:51094 - 51112.

[26] Anugerah Galang Persada,Igi Ardiyanto,Muhammad Bayu Sasongko,et al. A review of CAM - based visual explanation on diabetic retinopathy[J]. IEEE Access,2026,14:13737 - 13767.

(编　辑:朱银周)