



解耦与引导协同的输电线路螺栓缺陷检测

赵文清, 张乐

引用本文:

赵文清, 张乐. 解耦与引导协同的输电线路螺栓缺陷检测[J]. 智能系统学报, 2026, 21(5): 1142–1153.

ZHAO Wenqing, ZHANG Le. Synergistic decoupling and guidance for transmission line bolt defect detection[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1142–1153.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202509002>

您可能感兴趣的其他文章

双向特征融合与注意力机制结合的目标检测

Target detection based on bidirectional feature fusion and an attention mechanism

智能系统学报. 2021, 16(6): 1098–1105 <https://dx.doi.org/10.11992/tis.202012029>

融合共现推理的Faster R-CNN输电线路金具检测

Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection

智能系统学报. 2021, 16(2): 237–246 <https://dx.doi.org/10.11992/tis.202012023>

隔级融合特征金字塔与CornerNet相结合的小目标检测

Small target detection based on a combination of feature pyramid and CornerNet

智能系统学报. 2021, 16(1): 108–116 <https://dx.doi.org/10.11992/tis.202004033>

嵌入遮挡关系模块的SSD模型的输电线路图像金具检测

Fittings detection in transmission line images with SSD model embedded occlusion relation module

智能系统学报. 2020, 15(4): 656–662 <https://dx.doi.org/10.11992/tis.202001008>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

基于改进的Faster R-CNN高压线缆目标检测方法

Object detection of high-voltage cable based on improved Faster R-CNN

智能系统学报. 2019, 14(4): 627–634 <https://dx.doi.org/10.11992/tis.201905026>

DOI: 10.11992/tis.202509002

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20260311.0850.002>

解耦与引导协同的输电线路螺栓缺陷检测

赵文清^{1,2}, 张乐¹

(1. 华北电力大学 控制与计算机工程学院, 河北 保定 071003; 2. 华北电力大学 河北省能源电力知识计算重点实验室, 河北 保定 071003)

摘要: 针对输电线路巡检图像中螺栓等紧固件目标微小、细节易丢失、上下文信息利用不充分以及多尺度特征融合存在语义冲突的问题, 本文提出一种融合解耦与引导思想的螺栓缺陷检测方法。该方法主要由以下 3 部分构成: 1) 高频细节引导的特征增强模块利用高频分量驱动门控机制增强微弱目标的细节表达, 2) 解耦上下文与显著性金字塔池化模块通过并行分支分别捕获峰值响应与区域统计特征以实现信息互补, 3) 非对称引导融合金字塔网络以非对称路径分流和细节引导融合策略缓解多尺度特征的语义冲突。在自建螺栓缺陷数据集上的实验结果表明, 所提方法性能优越, 在 mAP@50 和 mAP@.5:.95 指标上达到 93.9% 和 63.4%, 较基线模型 YOLOv11s 分别提升了 3.1 和 2.9 个百分点, 并在公开数据集 RSOD 与 DIOR 上表现出良好的泛化能力。

关键词: 螺栓; 缺陷识别; 输电线路巡检; 小目标检测; 特征融合; 解耦金字塔池化; 深度学习; 一阶段检测器
中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2026)05-1142-12

中文引用格式: 赵文清, 张乐. 解耦与引导协同的输电线路螺栓缺陷检测 [J]. 智能系统学报, 2026, 21(5): 1142-1153.

英文引用格式: ZHAO Wenqing, ZHANG Le. Synergistic decoupling and guidance for transmission line bolt defect detection[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1142-1153.

Synergistic decoupling and guidance for transmission line bolt defect detection

ZHAO Wenqing^{1,2}, ZHANG Le¹

(1. School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China; 2. Hebei Key Laboratory of Knowledge Computing for Energy & Power, North China Electric Power University, Baoding 071003, China)

Abstract: This paper proposes a novel detection method integrating decoupling and guidance strategies to address the challenges of detecting small fasteners in UAV imagery, including detail loss, insufficient contextual information, and conflicts in multi-scale feature fusion. The method mainly consists of the following three modules: a high-frequency detail-guided feature enhancement (HFDG) module that leverages high-frequency components to drive a gating mechanism, thereby enhancing the representation of faint targets; a decoupled context and salient pyramid pooling (DCSPP) module that captures peak responses and regional statistics in parallel to achieve complementary information; and an asymmetric guidance fusion pyramid network (AGF-PN) that mitigates semantic conflicts via asymmetric path shunting and a detail-guided fusion strategy. Experiments on a self-built bolt defect dataset demonstrate superior performance, achieving 93.9% mAP@50 and 63.4% mAP@.5:.95, surpassing the YOLOv11s baseline by 3.1 and 2.9 percentage points, respectively. The proposed method also demonstrates strong generalization on the public RSOD and DIOR datasets.

Keywords: bolt; defect detection; power line inspection; small object detection; feature fusion; decoupled pyramid pooling; deep learning; one-stage detector

输电线路作为电力系统的动脉, 其安全稳定运行至关重要^[1]。近年来, 无人机巡检已成为保

障线路健康、替代高风险人工作业的主流技术^[2-3]。在巡检任务中, 对螺栓、销钉等关键紧固件的缺陷检测是一项核心环节。这些部件虽小, 但其松动、缺损等状态异常可能引发导线脱落等严重事故, 对电网安全构成重大威胁。

收稿日期: 2025-09-01. 网络出版日期: 2026-03-11.

基金项目: 国家自然科学基金项目 (62371188).

通信作者: 赵文清. E-mail: zhaowenqing@ncepu.edu.cn.

©《智能系统学报》编辑部版权所有

然而,实现对此类目标的精确检测难度较大。首先,螺栓等紧固件属于典型的小目标,其在图像中有效像素占比极低,其轮廓、纹理等高频特征在深度卷积网络(convolutional neural network, CNN)的层级特征提取中,极易因空间分辨率的不断降低而被削弱或丢失^[4],这种现象可被视为一种网络固有的低通滤波效应。此外,最大池化(max pooling)等降采样操作在提取显著特征的同时,会舍弃大量对小目标识别同样重要的上下文统计信息,如均值、方差等特征被压缩或忽略^[5],从而使模型在高空复杂背景下难以有效定位目标。另外,现有的特征金字塔网络,如PANet(path aggregation network)^[6],在融合不同层级特征时,其自顶向下与自底向上的信息流路径存在“紧耦合”,这使得中间特征需同时适应来自不同语义层次的监督信号^[7],从而引发语义冲突与梯度混淆。最后,螺栓状态的判定本身具有很强的上下文依赖性,例如,在悬垂线夹上,无销钉螺栓为竖直方向的被认为是正常状态,而水平方向则为异常状态,这种视觉歧义对模型的场景理解能力提出了更高要求。

为应对上述问题,本文提出了一种用于输电线路螺栓缺陷检测的方法。其核心思想为“解耦与引导(decoupling and guidance)”,即通过结构性解耦分离原本冲突或冗余的信息流,并利用一种信息如高频细节去自适应地引导和增强另一种信息如语义特征的表达。本文的贡献如下。

1) 针对细节信息丢失问题,本文提出了高频细节引导的特征增强模块(high-frequency detail-guided feature enhancement, HFDG),通过并行结构与高频分量驱动的门控机制,以增强小目标区域中微弱特征的表达能力。

2) 为弥补上下文信息的损失,本文设计了解耦上下文与显著性金字塔池化模块(decoupled context and salient pyramid pooling, DCSPP),将提取峰值响应的显著特征分支与捕获区域统计特征的上下文分支进行解耦并行处理,并通过融合实现信息互补。

3) 为缓解特征金字塔中的语义冲突,本文构建了非对称引导融合金字塔网络(asymmetric guidance fusion pyramid network, AGF-PN),采用非对称路径分流为不同方向的信息流提供专属的梯度回传路径,并利用细节引导融合机制实现更高效的自适应多尺度信息整合。

1 相关工作

1.1 目标检测方法

近年来,基于深度学习的目标检测算法取得

了显著进展,以YOLO(you only look once)为代表的单阶段检测器通过不断引入如C2f、SPPF、PANet等高效模块实现了速度与精度的平衡,使其成为工业界的主流架构^[8]。此外,基于Transformer的检测器开创了新的端到端检测范式,Deformable DETR(deformable transformers for end-to-end object detection)^[9]通过多尺度可变形注意力缓解了原始DETR收敛慢和小目标检测困难的问题;DAB-DETR(dynamic anchor boxes are better queries for DETR)^[10]与Conditional DETR(conditional DETR for fast training convergence)^[11]分别在查询设计与收敛速度上进行改进;DINO(DETR with improved denoising anchor boxes)^[12]与DDQ(dense distinct queries for end-to-end object detection)^[13]则通过更加致密、去噪的查询机制,进一步提升了端到端检测的精度与稳定性。

然而,这些目标检测算法多基于PASCAL VOC、MS COCO等自然场景目标检测数据集设计,其网络结构在直接应用于输电巡检时,难以适应此类存在极端小目标的场景。小目标检测是计算机视觉领域的关键难题,目前主要有多尺度特征融合方法、利用上下文信息融合的方法和特征增强的方法。

为同时捕获高层语义信息与底层高分辨率细节,针对多尺度特征融合的特征金字塔网络(feature pyramid network, FPN)及其变体被广泛采用^[14-16]。然而,诸如PANet的对称耦合结构会导致信息冗余与语义冲突。因此,本文提出AGF-PN,通过非对称路径分流在结构上解耦两条不同方向的路径,并通过细节引导融合策略,利用底层细节去指导高层语义的融合,实现了更具针对性的信息交互。

小目标周围上下文信息的有效利用是提升检测性能的有效策略。现有方法通过引入全局或区域语义建模来扩展感知范围^[17-18],或者采用高效的池化压缩策略以较低的计算代价实现多尺度上下文聚合。然而,最大池化在强化峰值响应的同时存在明显信息压缩问题^[19],容易造成上下文丢失。为此,本文提出了DCSPP,将提取峰值响应的“显著特征分支”与提取区域均值的“并行上下文增强分支”进行解耦,协同增强实现信息互补,在继承SPPF高效性的同时,缓解其固有的信息损失。

利用特征表示增强的方法关注增强小目标自身的特征表达^[20-21]。不同于上述已有方法依赖特征激活强度作为门控信号,本文提出的HFDG采

用一种门控机制,通过显式提取的高频分量驱动门控选择,引导网络聚焦于对小目标检测中至关重要的高频细节。

1.2 深度学习在电力巡检中的应用

深度学习已被广泛应用于输电线路巡检^[22]。关于螺栓、销钉等紧固件的检测,现有研究主要有基于主流目标检测框架的方法、通过优化训练策略来提升模型性能的方法和利用大模型与多模态的方法。

基于主流目标检测框架的方法多基于通用框架进行改良。He 等^[23]设计了集成全局与局部感知的 Res-PANet 结构,并引入 MPPD(minimum peripheral point distance)边界框约束以解决样本不均衡问题。张珂等^[2]构建了基于自监督学习的 E-Swin(efficient shifted windows)模型,利用未标记数据挖掘特征以降低标注依赖。Luo 等^[24]则提出了 UBDDM,采用“先粗后精”的两阶段策略应对超小目标检测。近期,Mamba-YOLO-SPDC^[25]引入状态空间模型(state space model, SSM),通过集成 VSS(visual state space)模块与 SPD-Conv(space-to-depth convolution)构建两阶段框架,旨在增强全局上下文感知并缓解细粒度特征丢失。然而,由于这些方法多是在于自然场景的架构上进行设计,在应对输电巡检场景中仍存在结构适配性不足的问题。通过优化训练策略的方法也受到关注。阎光伟等^[26]通过融合先验信息与特征约束,应对螺栓缺陷的类间相似与类内多样性挑战。Zhang 等^[27]则针对样本稀缺问题,设计了结合双判别器与伪增强策略的生成对抗网络以扩充训练数据。这些

方法主要在训练层面提供优化,但并未对检测器核心的网络结构进行显著改进。近期有研究者将基础大模型与多模态技术应用于电力巡检。PowerGPT^[28]作为电力领域通用基础模型,利用海量时序数据预训练以提升异常检测能力。针对视觉样本不均衡问题,一种基于 VLM(vision-language model)的方法^[29]通过将绝缘子物理规范转化为先验知识嵌入特征空间,实现了缺陷样本的生成式扩充。Zhao 等^[30]则聚焦螺栓检测的可解释性,通过构建螺栓知识图谱并利用大语言模型进行跨模态推理,赋予了检测结果明确的语义解释。但是,这些方法需要巨大的计算资源,且大模型并不适用于所有应用场景。

基于以上分析,本文提出一种融合解耦与引导思想的螺栓缺陷检测方法。具体而言,HF-DG 解耦了主干与高频特征提取,并利用高频信息引导主干特征的增强;DCSPP 解耦了显著性与上下文信息的提取,并利用上下文信息引导显著特征的增强;AGF-PN 则解耦了特征金字塔的上下行路径,并通过细节信息引导语义特征的融合。

2 本文方法

本文所提方法的整体架构如图 1 所示。输入的巡检图像首先由 2 个 3×3 卷积组成的下采样模块(Stem)进行 4 倍下采样并送入层次化引导网络(hierarchical guided network, HGN),完成特征提取与增强。由 HGN 得到的多尺度特征通过 AGF-PN 进行跨尺度特征融合,融合后特征经由检测头(Detect)得到最终的检测结果。

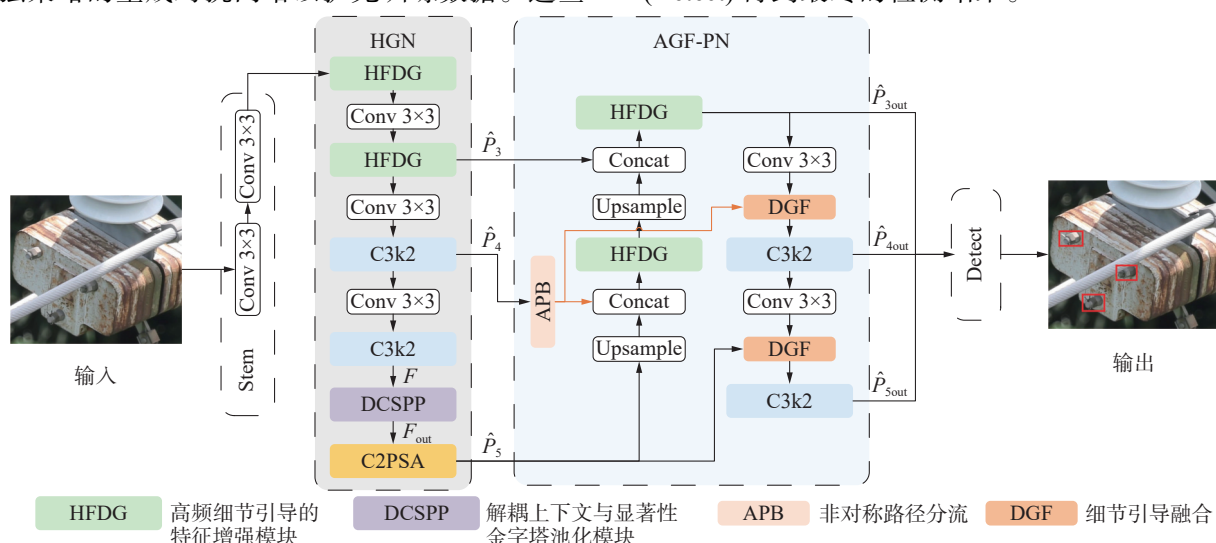


图 1 解耦与引导协同的输电线路螺栓缺陷检测整体架构

Fig. 1 Overall architecture of synergistic decoupling and guidance for transmission line bolt defect detection

2.1 层次化引导网络 (HGN)

如图 1 所示, HGN 主要由 HFDG、C3k2、负责

下采样的 3×3 卷积组成,且在末端通过 DCSPP 和 C2PSA 进行特征预融合增强,从而为送入 AGF-PN

做好准备。其中, C3k2 和 C2PSA 是 YOLOv11 的标准模块。本节将重点介绍 HFDG 和 DCSPP。

2.1.1 高频细节引导的特征增强模块 (HFDG)

为缓解深度网络在处理螺栓、销钉等小目标的特征时高频细节信息易丢失的问题, 本文提出了 HFDG, 如图 2 所示。该模块通过并行分支设计、高频信号驱动的门控机制以及分段式的信息融合保留并增强了对缺陷识别至关重要的细节特征。

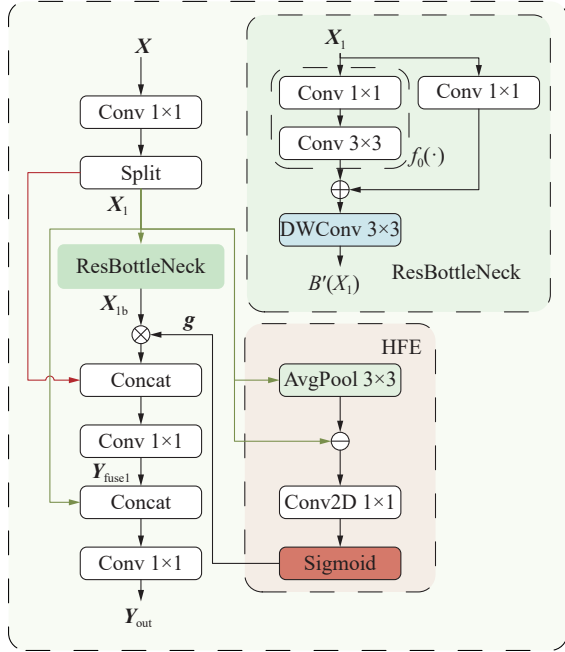


图 2 HFDG 模块结构

Fig. 2 Structure of the HFDG module

记输入特征图为 $X \in \mathbb{R}^{C_{in} \times H \times W}$, 模块的最终输出为 $Y_{out} \in \mathbb{R}^{C_{out} \times H \times W}$, HFDG 的过程为

$$[X_0, X_1] = \text{Split}(\text{Conv}_{1 \times 1}(X)) \quad (1)$$

1) 初始通路初始化。为保留原始的浅层信息, 输入特征 X 经过 1×1 卷积处理后, 被划分为主干处理分支 X_1 和浅层短路分支 X_0 , 如式 (1) 所示。

2) 高频引导门控。为利用特征图中的细节信息引导特征提取, 高频细节提取器 (high-frequency detail extractor, HFE) 首先从主干分支中提取高频分量。该操作通过从原始特征中减去平均池化 (一种低通滤波) 的结果来实现。然后, 提取的高频特征将被送入由一个 1×1 卷积构成的门控生成器, 并经 Sigmoid 激活函数 $\sigma(\cdot)$ 生成各个特征点的通道门控向量 g , 该门控向量将调控特征的重要性, 公式为

$$g = \sigma(\text{Conv}_{2D 1 \times 1}(X_1 - \text{AvgPool}_{3 \times 3}(X_1)))$$

3) 并行局部残差增强。为缓解串行的特征提取带来的信息衰减, 主干分支 X_1 被送入一个并行的局部残差增强块 ResBottleNeck 进行处理, 记为 $B'(\cdot)$, 从而增强对中间层特征的保持能力, 公式表

示为

$$f_0(\cdot) = \text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(\cdot))$$

$$B'(X_1) = \text{DWConv}_{3 \times 3}(f_0(X_1) + \text{Conv}_{1 \times 1}(X_1))$$

并行分支的输出 $B'(X_1)$ 与前述的门控向量 g 进行逐通道相乘得到 X_{1b} , 实现对特征的自适应增强。

4) 两阶段分段式融合。最后, 为整合模块内的多个信息流, 设计了一种两阶段融合策略以保留特征的完整性。

第一阶段, 将浅层短路分支 X_0 与门控增强后的主干分支 X_{1b} 进行拼接融合, 得到初步融合结果 Y_{fuse1} , 公式表示为

$$Y_{fuse1} = \text{Conv}_{1 \times 1}(\text{Concat}(X_0, X_{1b}))$$

第二阶段, 为构建直接的信息补偿通路来保留原始语义, 将 Y_{fuse1} 与未经处理的原始主干分支 X_1 进行拼接融合, 得到最终结果 Y_{out} , 公式为

$$Y_{out} = \text{Conv}_{1 \times 1}(\text{Concat}(Y_{fuse1}, X_1))$$

2.1.2 解耦上下文与显著性金字塔池化 (DCSPP)

为缓解最大池化带来的上下文信息损失并增强网络对小目标及其复杂背景的理解能力, 本文提出了 DCSPP。该模块通过解耦显著特征分支与上下文增强分支, 实现两种异构特征间的互补增强, 其整体结构如图 3 所示。

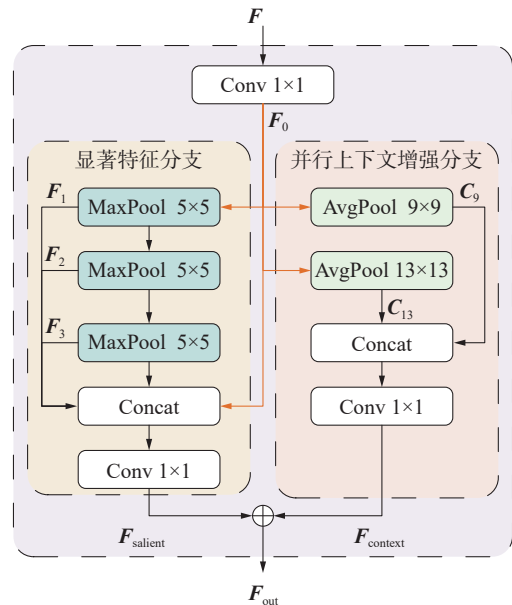


图 3 DCSPP 模块结构

Fig. 3 Structure of the DCSPP module

1) 显著特征分支。该分支用于提取多尺度的局部峰值响应。首先, 输入特征 F 经过一个瓶颈卷积得到中间特征 F_0 。随后, F_0 与其分别经过 3 次串联的 5×5 最大池化的输出进行拼接, 以聚合不同尺度的局部最强响应, 最终由一个融合卷积输出显著特征 $F_{salient}$, 公式表示为

$$F_0 = \text{Conv}_{1 \times 1}(F)$$

$$F_i = \text{MaxPool}_{5 \times 5}(F_{i-1}), i = 1, 2, 3$$

$$F_{\text{salient}} = \text{Conv}_{1 \times 1}(\text{Concat}(F_0, F_1, F_2, F_3))$$

2) 并行上下文增强分支。该分支用于捕获与显著特征互补的多尺度区域统计信息, 以中间特征 F_0 为输入, 采用多个并行的大核平均池化提取多尺度特征的一阶统计信息 (均值) 来获得平滑的区域上下文信息, 从而补充显著响应中的局部偏置。并由融合卷积进行自适应加权, 得到上下文特征 F_{context} , 公式表示为

$$C_9 = \text{AvgPool}_{9 \times 9}(F_0)$$

$$C_{13} = \text{AvgPool}_{13 \times 13}(F_0)$$

$$F_{\text{context}} = \text{Conv}_{1 \times 1}(\text{Concat}(C_9, C_{13}))$$

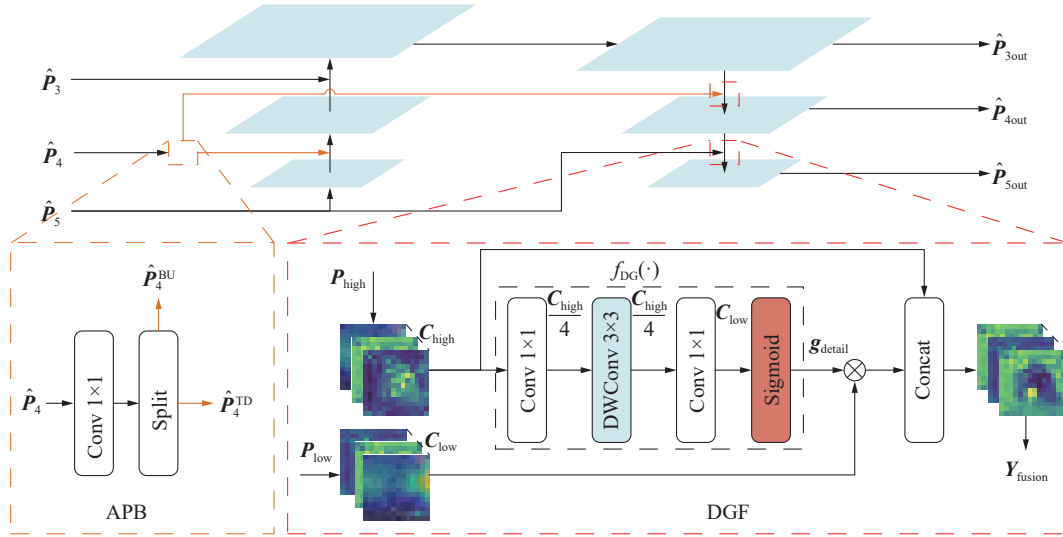


图 4 AGF-PN 结构

Fig. 4 Structure of the AGF-PN

AGF-PN 接收来自 HGN 输出的多尺度特征 $\hat{P}_3$ 、 $\hat{P}_4$ 、 $\hat{P}_5$, 通过 APB 和 DGF 进行处理。

2.2.1 非对称路径分流 (APB)

为解耦上下行路径, AGF-PN 首先对 $\hat{P}_4$ 使用 1 个 1×1 卷积实现通道级的自适应特征重标定。随后在通道维度上划分为用于自顶向下路径的 $\hat{P}_4^{\text{TD}}$ 和自底向上路径的 $\hat{P}_4^{\text{BU}}$, 公式表示为

$$\hat{P}_4^{\text{TD}}, \hat{P}_4^{\text{BU}} = \text{Split}(\text{Conv}_{1 \times 1}(\hat{P}_4))$$

式中: $\hat{P}_4^{\text{TD}}$ 专用于与上采样后的 $\hat{P}_5$ 融合, 来执行自顶向下路径的语义增强任务; $\hat{P}_4^{\text{BU}}$ 则保留了未经高层语义调制的部分原始 $\hat{P}_4$ 特征, 用于后续自底向上路径的融合, 以进行细节保持。

2.2.2 细节引导融合 (DGF)

在自底向上的融合阶段, 为实现更具选择性的信息整合, 引入 DGF 来取代简单的特征拼接。当融合来自高分辨率特征图的 P_{high} 与来自低分辨率

3) 特征融合。最后, 将两个分支的输出进行逐元素相加, 得到模块的最终输出 F_{out} , 公式为

$$F_{\text{out}} = F_{\text{salient}} + F_{\text{context}}$$

该模块为显著特征引入了互补的上下文信息作为调制与语义补充, 从而避免了最大池化导致的信息损失。

2.2 非对称引导融合金字塔网络 (AGF-PN)

针对传统 PANet 结构中因上下行路径耦合导致的语义冲突与信息冗余, 本文提出了 AGF-PN。该网络通过非对称路径分流 (asymmetric path bifurcation, APB) 和细节引导融合 (detail-guided fusion, DGF) 两项机制, 实现了结构解耦和细节引导的跨尺度信息整合。其整体结构如图 4 所示。

率特征图的 P_{low} 时, DGF 首先利用保留了更精确定位细节的 P_{high} 来生成空间门控信号 g_{detail} 。随后, 将 g_{detail} 乘性地施加于低分辨率但语义更丰富的特征 P_{low} 上, 并与原始的 P_{high} 进行拼接得到融合特征 Y_{fusion} , 公式表示为

$$g_{\text{detail}} = f_{\text{DG}}(P_{\text{high}})$$

$$Y_{\text{fusion}} = \text{Concat}((P_{\text{low}} \otimes g_{\text{detail}}), P_{\text{high}})$$

式中 $f_{\text{DG}}(\cdot)$ 是一个轻量级门控生成网络, 由以下 4 层组成: 首先通过 1×1 卷积将输入通道数压缩为原来的 $1/4$, 接着使用 3×3 的深度卷积提取局部上下文, 然后通过 1×1 卷积对通道信息整合并扩张至低分辨率特征 P_{low} 的通道大小, 最后经 Sigmoid 激活函数生成门控向量。

DGF 使网络能够基于高分辨率特征中更可靠的空间细节, 有选择性地增强低分辨率特征中的关键区域, 实现定位精度和语义信息的协同优化。

最终, 这些经过增强后的多尺度特征 $\hat{P}_{3out}$ 、 $\hat{P}_{4out}$ 、 $\hat{P}_{5out}$ 被送往检测头输出检测结果。

3 实验结果与分析

基于自建输电线路螺栓缺陷数据集和公开数据集, 本文分别进行了实验验证。考虑到该任务中螺栓状态的判定与其所处的金具类型和空间方向紧密相关 (例如悬垂线夹上的无销钉螺栓垂直方向时为正常, 水平则为异常), 本文在进行数据标注时采取了细致的策略来提供更多的监督信息。最终的数据集共包含 1 343 张图像, 总计 5 355 个标注实例。其中, 核心螺栓标签采用了包含状态、方向 and 有无销钉的复合类别, 具体分为 5 类: 正常垂直无销钉螺栓 (2 092 个)、正常水平有销钉螺栓 (850 个)、异常水平无销钉螺栓 (531 个)、正常无方向特征有销钉螺栓 (460 个), 以及异常无方向特征无销钉螺栓 (255 个)。为提供丰富的上下文信息, 还额外标注了 6 种螺栓所处的金具部件类型作为辅助标签, 仅作为训练时的信息补充, 共计 1 167 个, 包括悬垂线夹 XG 型 (283 个)、直角挂板 (259 个)、悬垂线夹 XGU 型 (229 个)、重锤 (202 个)、悬垂线夹 XT 型 (130 个) 和悬垂线夹 XTS 型 (64 个)。

所有实验均在 Ubuntu 22.04 操作系统上进行, 硬件包括 Intel Core i7-14700K 处理器和 NVIDIA

GeForce RTX 5090D 显卡, 模型训练在 GPU 上进行, 所有推理实验均在 CPU 上进行。深度学习框架为 PyTorch 2.6.0, CUDA 版本为 12.8。

本文采用平均精度 (average precision, AP) 和平均精度均值 (mean average precision, mAP) 作为主要评价指标, 其中 mAP@50 表示 IoU(intersection over union) 为 0.5 时的 mAP, mAP@.5:.95 表示 IoU 从 0.5 到 0.95 范围内计算得到的 mAP。同时报告模型参数量 (parameters, Params)、计算量 (GFLOPs) 以及推理速度 (每秒帧数, frames per second, FPS) 用于评估模型复杂度与检测效率。

3.1 对比实验

为验证本文所提方法的有效性, 将本文方法与一系列先进的目标检测模型进行了综合比较, 包括不同版本的 YOLO 系列以及主流的 DETR 变体。此外, 附录中表 A1 对本文采用的代表性 YOLO 版本的核心骨干、检测头与匹配策略及关键特性进行了归纳整理。

所有模型均在自建的输电线路螺栓缺陷数据集上进行训练和评估以保证公平性, 详细的对比结果如表 1 所示。其中“A_H_N”代表“异常水平无销钉螺栓”, “N_X”代表“正常无方向特征有销钉螺栓”, “A_X”代表“异常无方向特征无销钉螺栓”, “N_V_N”代表“正常垂直无销钉螺栓”, “N_H_P”代表“正常水平有销钉螺栓”。

表 1 与其他模型的对比
Table 1 Comparison with other models

模型	A_H_N/%	N_X/%	A_X/%	N_V_N/%	N_H_P/%	mAP@50/%	mAP@.5:.95/%	Params/ 10^6	GFLOPs	FPS/(帧/s)
DDQ ^[13]	87.8	84.0	62.6	94.1	86.5	83.0	56.3	48.30	228.0	0.5
DAB-DETR ^[10]	93.8	90.8	84.4	95.4	89.3	90.8	58.7	43.70	84.4	2.3
Cond-DETR ^[11]	96.6	90.9	85.7	95.3	91.0	91.9	59.8	43.40	83.1	2.7
DINO ^[12]	86.7	78.3	60.8	94.8	86.4	81.4	55.2	47.50	228.0	0.6
Def-DETR ^[9]	95.8	89.8	80.7	95.6	90.3	90.4	59.4	41.00	163.0	0.8
YOLOv5s	90.6	87.4	81.3	97.6	91.2	89.6	58.1	9.13	24.1	21.3
YOLOv6s	94.7	89.4	82.2	96.8	92.9	91.2	59.8	16.31	44.2	26.6
YOLOv8s	94.9	88.7	79.3	97.8	92.0	90.5	59.3	11.14	28.7	20.0
YOLOv9s ^[31]	93.6	87.9	83.9	97.4	88.7	90.3	57.8	7.29	27.4	23.9
YOLOv9m	95.3	88.3	86.5	97.0	91.2	91.7	60.6	20.17	77.6	12.2
YOLOv10s ^[32]	92.4	87.2	83.1	96.3	89.9	89.8	60.3	8.07	24.8	33.9
YOLOv10m	92.0	92.3	85.2	95.4	92.8	91.6	61.7	16.50	64.0	15.7
YOLOv11s	93.5	90.9	81.1	97.8	90.8	90.8	60.5	9.44	21.5	32.0
YOLOv12s ^[33]	92.7	90.8	83.2	97.5	92.3	91.3	59.2	9.24	21.2	26.7
本文模型	95.3	93.9	89.8	97.5	93.2	93.9	63.4	10.12	24.7	28.6

注: 加粗表示本列最优结果。

表 1 的详细对比结果表明了本文所提方法的综合优势。与基线模型 YOLOv11s 相比, 本文方法在代表整体性能的 $mAP@50$ 和 $mAP@.5:.95$ 上分别取得了 3.1 和 2.9 个百分点的提升, 这主要得益于对关键缺陷类别识别能力的增强。本文方法对于视觉辨识难度大的“异常无方向特征无销钉螺栓”类别, AP 值实现了从 81.1% 到 89.8% 的增长 (+8.7 百分点), 这证明了本文所提模块在处理细节、上下文及语义信息上的有效性。

值得注意的是, 本文所提模型的性能提升是

在保持模型轻量化的前提下实现的。相较于基线模型, 本文模型的参数量仅微增 7.2%, 推理速度则保持在同一水平。若与 YOLOv9m 等更复杂的模型相比, 本文所提模型能以约 1/2 的参数量和 1/3 的计算量达到更优的检测精度, 验证了本文方法的先进性。

3.2 消融实验

为评估本文所提出的各模块的有效性, 设计了消融实验。所有实验均在基线模型 YOLOv11s 上进行增量式对比, 实验结果如表 2 所示。

表 2 消融实验结果
Table 2 Results of ablation study

%

方法	A_H_N	N_X	A_X	N_V_N	N_H_P	$mAP@50$	$mAP@.5:.95$
Baseline(YOLOv11s)	93.5	90.9	81.1	97.8	90.8	90.8	60.5
Baseline+HFDG	94.3	92.4	85.3	97.6	93.2	92.6	60.7
Baseline+AGF-PN	91.9	90.7	88.3	96.8	93.1	92.2	62.1
Baseline+DCSPP	94.2	90.3	86.2	97.3	90.7	91.7	62.0
Baseline+AGF-PN+HFDG	94.4	91.8	89.1	97.7	91.4	92.9	61.7
Baseline+DCSPP+HFDG	96.6	91.1	89.9	97.6	91.9	93.4	62.9
Baseline+AGF-PN+DCSPP	96.2	94.6	87.1	97.6	92.6	93.6	62.3
Baseline+HFDG+AGF-PN+DCSPP	95.3	93.9	89.8	97.5	93.2	93.9	63.4

注: 加粗表示本列最优结果。

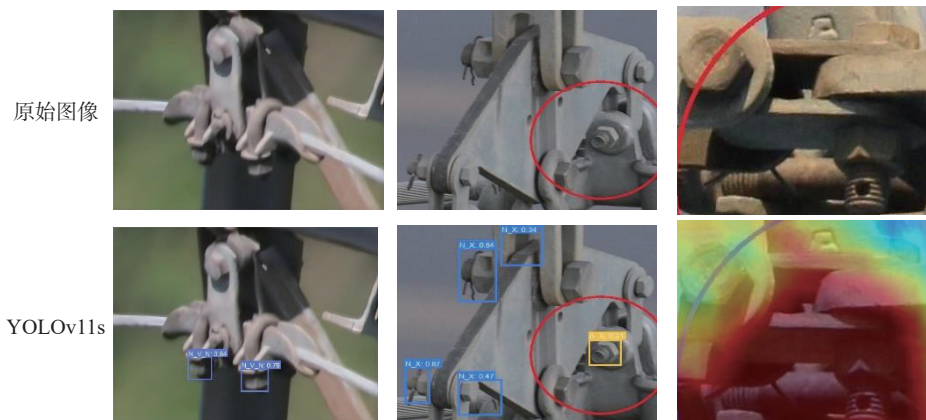
1) 各模块的性能分析。将 3 个模块分别引入基线模型后进行测试, 结果表明每个模块均能带来明显的性能提升。以 HFDG 为例, 其将 $mAP@50$ 从 90.8% 提升至 92.6%(+1.8 百分点), 在“异常无方向特征无销钉螺栓”这一难点类别上提升了 4.2 百分点, 验证了 HFDG 在增强困难语义小目标细节特征上的有效性。AGF-PN 和 DCSPP 则在更为严格的 $mAP@.5:.95$ 指标上提升了 1.6 和 1.5 百分点, 表明本文方法通过优化特征融合和增强上下文信息改善了模型的定位精度。

2) 模块间的协同效应分析。将模块进行两两组合测试, 结果表明模块间的组合取得了比单个模块更好的性能。例如, Baseline+DCSPP+HFDG 的组合在 $mAP@.5:.95$ 上达到了 62.9% 的检测

精度, Baseline+AGF-PN+DCSPP 的组合则在 $mAP@50$ 上达到了 93.6% 的高点。当同时集成 3 个模块时, 本文方法达到了 93.9% 的 $mAP@50$ 和 63.4% 的 $mAP@.5:.95$, 均为所有配置中的最优值。

3.3 实验结果可视化

图 5 给出了本文方法与基线模型在 2 个典型场景下的检测效果。在图 5(a) 中, 基线模型 YOLOv11s 由于对微弱细节不敏感而漏检了 2 个“正常水平有销钉螺栓”, 而本文方法借助 HFDG 对局部细节特征的强化, 避免了漏检。在图 5(b) 的复杂背景下, 基线模型将两处无关的金属结构识别为螺栓, 导致了误检, 且将“异常水平无销钉螺栓”检测为“异常无方向特征无销钉螺栓”。本文方法通过引入 DCSPP 与 AGF-PN 获得的上下文信息避免了误检。



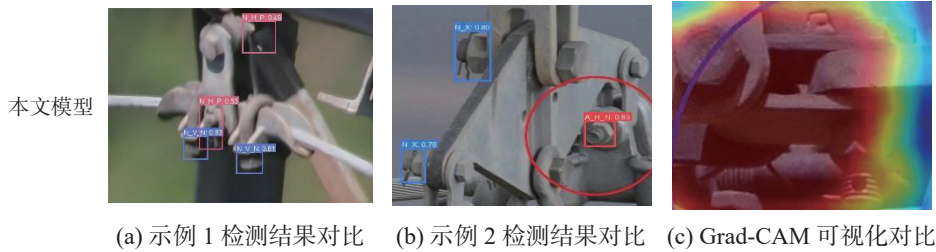


图5 检测结果的可视化

Fig. 5 Visualization of detection results

图5(c)给出了“异常无方向特征无销钉螺栓”类别的一个典型检测场景及其对应的 Grad-CAM 可视化结果。本文模型不仅关注螺栓的局部纹理,还成功捕获了背景金具的全局结构语义。这种上下文感知能力使得模型能够利用金具的几何形态来辅助推断微小螺栓的存在及其语义,从而提高检测精度,这一点也与该类别 AP 值由 81.1%

提升至 89.8% 的结果相吻合。

图6给出了基线模型与本文方法在各类别上的 PR(precision-recall) 曲线对比,其中图6(a)为正常类别,图6(b)为异常类别。可以看出,本文方法在大部分类别上的 PR 曲线整体高于基线模型,尤其在中高召回率区间优势更为明显。这表明所提方法在保证精度的同时,对缺陷目标具有更好的检出能力。

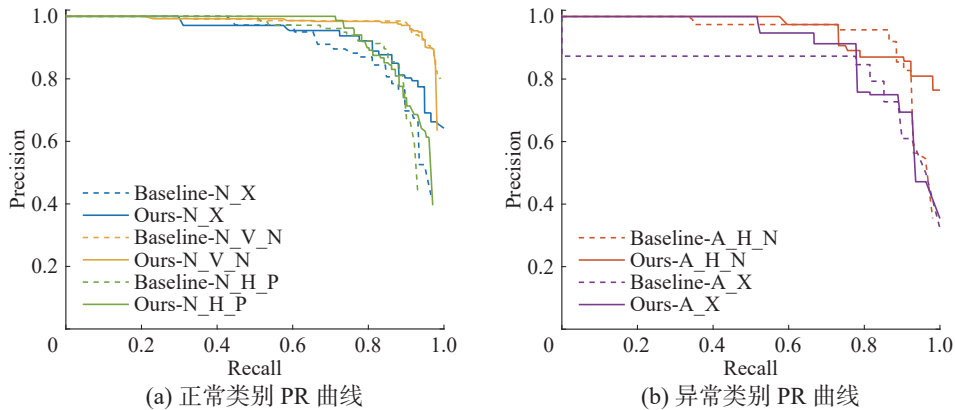


图6 各类别 PR 曲线

Fig. 6 Precision-Recall curves for each class

3.4 在其他数据集上的实验

为进一步验证本文所提模型的泛化性,本文在两个公开数据集 RSOD(remote sensing object detection)^[34] 和 DIOR(dataset for object detection in optical remote sensing images)^[35] 上进行了评估。与基于螺栓缺陷数据集的对比实验不同,RSOD 和 DIOR 在相关研究中已有常用的参考方法,因此本节在保留通用检测基线及本文方法的同时,增补了若干在相关工作中常用的代表性方法,用以更全面地展示本文方法在公开基准上的性能表现。结果分别如表3和表4所示。

RSOD 由武汉大学发布,包含飞机、操场、立交桥和油桶 4 类目标,共 976 张图像、6 950 个标注实例,并指出油桶等目标往往密集分布且背景复杂,对模型的小目标检测能力具有较强考验。对比实验结果如表3所示。本文提出的方法在 mAP@50 指标上取得了 96.5% 的精度,较基线模型 YOLOv11 提升了 3.5 百分点,且性能在所有对

比方法中位居榜首,验证了本文所提方法的泛化能力。

表3 不同模型在 RSOD 上的检测结果

Table 3 Detection results of different models on RSOD

方法	输入尺寸	mAP@50/%
SSD	300×300	76.4
Faster R-CNN	600×1 000	80.5
CenterNet	640×640	77.3
RetinaNet	640×640	81.7
YOLOv3	640×640	81.6
YOLOv4	640×640	87.8
YOLOv5	640×640	83.6
YOLOX	640×640	89.4
YOLOv8	640×640	89.4
IRSAM ^[36]	640×640	94.7
YOLOv11	640×640	93.0
YOLOv12	640×640	93.5
本文模型	640×640	96.5

注:加粗表示本列最优结果。

表 4 不同模型在 DIOR 上的检测结果
Table 4 Detection results of different models on DIOR

方法	输入尺寸	mAP@50/%
SSD	300×300	55.7
Faster R-CNN	600×1 000	52.1
CenterNet	640×640	55.2
RetinaNet	640×640	61.6
YOLOv3	640×640	58.0
YOLOv4	640×640	65.0
YOLOv5	640×640	69.6
YOLOX	640×640	72.2
YOLOv8	640×640	77.2
AAFNet ^[37]	640×640	75.3
SCFNet ^[38]	640×640	69.9
StrMCsDet ^[39]	608×608	65.6
YOLOv11	640×640	77.6
YOLOv12	640×640	77.6
本文模型	640×640	78.6

注: 加粗表示本列最优结果。

为突出 HFDG 与典型频域增强方法及经典非频域通道注意力模块之间的差异与优势, 本文在 RSOD 上以 YOLOv11 作为统一基线, 在与 HFDG 相同的特征层位置分别引入 FcaNet(frequency channel attention network)^[40] 与 ECA(efficient channel attention)^[41], 并保持其余配置不变进行对比实验。各方法在 RSOD 上的结果如表 5 所示。对比结果表明所提出的 HFDG 优于典型的通道注意力增强模块或传统频域重加权策略。

表 5 HFDG 与 FcaNet、ECA 模块的对比
Table 5 Comparison between HFDG, FcaNet and ECA modules

方法	输入尺寸	mAP@50/%
Baseline	640×640	93.0
Baseline+ECA	640×640	93.6
Baseline+FcaNet	640×640	93.1
Baseline+HFDG	640×640	94.7

注: 加粗表示本列最优结果。

DIOR 由西北工业大学发布, 为当前遥感目标检测领域应用广泛的大规模基准数据集, 包含 23 463 张图像和 192 472 个实例, 覆盖 20 类常见地物目标 (如储罐、风力发电机、港口等), 该数据集类间相似性高、类内差异大且目标尺度变化显著, 对检测器的整体性能提出了更高要求。如表 4 所示, 本文提出的方法在 mAP@50 的指标上较基线模型 YOLOv11 提升了 1.0 百分点。RSOD 与 DIOR

的实验结果表明, 本文所提方法在不同检测场景下仍能保持较好的检测性能, 体现出较强的泛化能力。

3.5 部署与量化性能

为进一步评估所提方法在资源受限的工业巡检场景中的部署潜力, 本文基于螺栓缺陷数据集, 在 Intel Core i7-14700K 平台上进行了量化推理测试。实验设置了两种数值精度配置进行对比: 以原生 PyTorch 框架进行的 FP32 全精度推理作为性能基准, 利用 OpenVINO 推理引擎实现的 INT8 量化推理以模拟实际部署环境。具体性能对比结果如表 6 所示。

表 6 不同数值精度下的检测性能
Table 6 Detection performance under different numerical precision settings

方法	数值精度	mAP@50/%	FPS/(帧/s)
Baseline	FP32	90.8	32.0
	INT8	90.6	84.3
本文模型	FP32	93.9	28.6
	INT8	92.4	79.4

实验结果表明, INT8 量化显著提升了推理效率。本文方法的推理速度跃升至 79.4 帧/s, 满足边缘侧实时巡检需求。在精度方面, 本文方法在量化后 mAP@50 由 93.9% 下降至 92.4%, 高于基线模型 0.2 百分点的精度损失。这表明, 本文所提方法在一定程度上更依赖于螺纹纹理、销钉边缘等高频细节特征, 因此对量化过程中的截断误差更为敏感; 相比之下, 基线模型在量化后精度变化较小, 但在微小缺陷的识别能力上仍明显弱于本文方法。实验证明本文方法可在边缘计算场景下实现精度与速度平衡。

4 结束语

本文提出了一种融合解耦与引导思想的螺栓缺陷检测方法。设计了利用高频分量强化微弱目标细节的高频细节引导的特征增强模块 (HFDG)、通过并行分支捕获互补上下文信息的解耦上下文与显著性金字塔池化模块 (DCSPP) 以及缓解多尺度特征语义冲突的非对称引导融合金字塔网络 (AGF-PN)。在自建的输电线路螺栓缺陷数据集上的实验结果表明, 本文方法性能优越, mAP@50 和 mAP@.5:.95 分别达到了 93.9% 和 63.4%, 相较于基线模型 YOLOv11s 分别提升了 3.1 和 2.9 百分点。此外, 本文方法在公开数据集 RSOD 与 DIOR 的表现证明了其良好的泛化能力。

未来的研究工作主要可从以下 3 方面展开: 一是面向算力更受限的嵌入式边缘设备, 在现有 INT8 量化的基础上尝试网络剪枝、知识蒸馏等模型压缩策略, 进一步提升推理效率; 二是针对输电线路缺陷样本稀缺、标注成本高的问题, 探索自监督学习与小样本学习范式, 利用未标注巡检数据减轻对精细标注的依赖, 改善长尾缺陷类别的检测性能; 三是结合多模态大模型的发展, 尝

试将本文框架与视觉-语言模型相结合, 引入文本先验辅助复杂缺陷场景的判别。

附录

本附录对本文实验中涉及的代表性 YOLO 系列模型进行了简要归纳。表 A1 总结了各版本在核心骨干、检测头与匹配策略及关键特性等方面的差异。

表 A1 不同 YOLO 版本的结构差异概览
Table A1 Structural comparison of different YOLO versions

版本	核心骨干	检测头与匹配策略	关键特性
YOLOv5	CSPDarknet主干, 包含Focus、C3等结构, SPPF+PANet颈部。	一阶段anchor-based检测头, 自适应锚框。	引入Mosaic数据增强和自适应锚框计算。
YOLOX	CSPDarknet主干, PANet颈部。	解耦检测头, anchor-free, SimOTA动态匹配。	构建统一的anchor-free解耦检测框架, 被后续多代YOLO模型沿用。
YOLOv6	EfficientRep主干, Rep-PAN颈部。	解耦检测头, SimOTA样本分配。	使用重参数化结构, 训练与推理结构分离以提升部署效率。
YOLOv8	C2f主干, PAN-like颈部。	解耦检测头, anchor-free, TAL分配。	引入C2f模块增强梯度流, 采用Task-Aligned分配策略。
YOLOv9	GELAN主干与颈部。	一阶段anchor-free检测头, TAL分配。	提出PGI机制(可编程梯度信息)与GELAN结构, 缓解深层信息瓶颈。
YOLOv10	改进C2f主干, FPN-PAN颈部。	NMS-free一阶段检测头, 双重一致性分配。	在保持精度的前提下移除NMS, 降低端到端推理延迟。
YOLOv11	C3k2作为主干核心, 配合SPPF与FPN-PAN颈部。	解耦anchor-free检测头(含C2PSA模块)。	通过C3k2与C2PSA加强局部细节与上下文信息建模能力。
YOLOv12	R-ELAN主干, 结合区域注意力的颈部。	解耦anchor-free检测头。	Attention-centric设计, 引入区域注意力提升全局感知能力。

参考文献:

[1] 李文娟, 杨生婧, 于昊, 等. 基于光电转换与多光谱成像的输电线路红外图像智能处理研究[J]. 光学与光电技术, 2025, 23(4): 16–27.
LI Wenjuan, YANG Shengjing, YU Hao, et al. Intelligent processing of infrared images for transmission lines based on photoelectric conversion and multispectral imaging[J]. Optics & optoelectronic technology, 2025, 23(4): 16–27.

[2] 张珂, 周睿恒, 石超君, 等. 自监督 E-Swin 的输电线路金具检测[J]. 中国图象图形学报, 2023, 28(10): 3064–3076.
ZHANG Ke, ZHOU Ruiheng, SHI Chaojun, et al. Self-supervised E-Swin based transmission line fittings detection[J]. Journal of image and graphics, 2023, 28(10): 3064–3076.

[3] 闫建红, 冉同霄. 基于 YOLOv8 的轻量化无人机图像目标检测算法[J]. 图学学报, 2024, 45(6): 1328–1337.
YAN Jianhong, RAN Tongxiao. Lightweight UAV image target detection algorithm based on YOLOv8[J]. Journal of graphics, 2024, 45(6): 1328–1337.

[4] XIAO Yao, XU Tingfa, YU Xin, et al. A lightweight fusion strategy with enhanced interlayer feature correlation for small object detection[J]. IEEE transactions on geoscience and remote sensing, 2024, 62: 4708011.

[5] ZAFAR A, SABA N, ARSHAD A, et al. Convolutional neural networks: a comprehensive evaluation and benchmarking of pooling layer variants[J]. Symmetry, 2024, 16(11): 1516.

[6] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759–8768.

[7] ZHAO Yuhao, LIU Qing, SU Hu, et al. Attention-based multiscale feature fusion for efficient surface defect detection[J]. IEEE transactions on instrumentation and measurement, 2024, 73: 5013310.

[8] 李珍峰, 符世琛, 徐乐, 等. 基于 MBI-YOLOv8 的煤矸石目标检测算法研究[J]. 图学学报, 2024, 45(6): 1301–1312.
LI Zhenfeng, FU Shichen, XU Le, et al. Research on gangue target detection algorithm based on MBI-YOLOv8

- [J]. *Journal of graphics*, 2024, 45(6): 1301–1312.
- [9] ZHU Xizhou, SU Weijie, LU Lewei, et al. Deformable DETR: deformable Transformers for end-to-end object detection[EB/OL]. (2021–01–13)[2025–09–01]. <https://openreview.net/forum?id=gZ9hCDWe6ke>.
- [10] LIU Shilong, LI Feng, ZHANG Hao, et al. DAB-DETR: dynamic anchor boxes are better queries for DETR [EB/OL]. (2022–01–29)[2025–09–01]. <https://openreview.net/forum?id=oMI9PjOb9Jl>.
- [11] MENG Depu, CHEN Xiaokang, FAN Zejia, et al. Conditional DETR for fast training convergence[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 3631–3640.
- [12] ZHANG Hao, LI Feng, LIU Shilong, et al. DINO: DETR with improved denoising anchor boxes for end-to-end object detection [EB/OL]. (2023–02–02)[2025–09–01]. <https://openreview.net/forum?id=3mRwyG5one>.
- [13] ZHANG Shilong, WANG Xinjiang, WANG Jiaqi, et al. Dense distinct query for end-to-end object detection[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7329–7338.
- [14] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10778–10787.
- [15] LIU Songtao, HUANG Di, WANG Yunhong. Learning spatial fusion for single-shot object detection[EB/OL]. (2019–11–21)[2025–09–01]. <https://arxiv.org/abs/1911.09516>.
- [16] ZHANG Yin, YE Mu, ZHU Guiyi, et al. FFCA-YOLO for small object detection in remote sensing images[J]. *IEEE transactions on geoscience and remote sensing*, 2024, 62: 5611215.
- [17] LIU Ze, LIN Yutong, CAO Yue, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2022: 9992–10002.
- [18] OUYANG Daliang, HE Su, ZHANG Guozhong, et al. Efficient multi-scale attention module with cross-spatial learning[C]//2023 IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes: IEEE, 2023: 1–5.
- [19] LU Yong, SUN Minghao. Lightweight multidimensional feature enhancement algorithm LPS-YOLO for UAV remote sensing target detection[J]. *Scientific reports*, 2025, 15: 1340.
- [20] DING Shijiao, XIONG Maomao, WANG Xinghua, et al. Dynamic feature and context enhancement network for faster detection of small objects[J]. *Expert systems with applications*, 2025, 265: 125732.
- [21] HE Luhao, ZHOU Yongzhang, LIU Lei, et al. Research on object detection and recognition in remote sensing images based on YOLOv11[J]. *Scientific reports*, 2025, 15: 14032.
- [22] 张辉, 杜瑞, 钟杭, 等. 电力设施多模态精细化机器人巡检关键技术及应用[J]. *自动化学报*, 2025, 51(1): 20–42.
- ZHANG Hui, DU Rui, ZHONG Hang, et al. The key technology and application of multi-modal fine robot inspection for power facilities[J]. *Acta automatica sinica*, 2025, 51(1): 20–42.
- [23] HE Min, QIN Liang, DENG Xinlan, et al. Bolt-YOLO: research on an algorithm framework for detecting bolt defects in transmission lines[J]. *IEEE transactions on power delivery*, 2025, 40(3): 1718–1729.
- [24] LUO Peng, WANG Bo, WANG Hongxia, et al. An ultrasmall bolt defect detection method for transmission line inspection[J]. *IEEE transactions on instrumentation and measurement*, 2023, 72: 5006512.
- [25] WANG Zeyu, LI Chen, XU Huiying, et al. Mamba YOLO: a simple baseline for object detection with state space model[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2025, 39(8): 8205–8213.
- [26] 阎光伟, 周香君, 焦润海, 等. 融合先验信息和特征约束的杆塔螺栓缺陷检测[J]. *中国图象图形学报*, 2023, 28(11): 3497–3508.
- YAN Guangwei, ZHOU Xiangjun, JIAO Runhai, et al. Defect detection of tower bolts by fusion of priori information and feature constraints[J]. *Journal of image and graphics*, 2023, 28(11): 3497–3508.
- [27] ZHANG Ke, XIAO Yangjie, WANG Jiacun, et al. DP-GAN: a transmission line bolt defects generation network based on dual discriminator architecture and pseudo-enhancement strategy[J]. *IEEE transactions on power delivery*, 2024, 39(3): 1622–1633.
- [28] TU Shihao, ZHANG Yupeng, HUAJIANG Yan, et al. PowerGPT: foundation model for power systems [EB/OL]. (2024–02–11)[2025–09–01]. <https://openreview.net/forum?id=ntSP0bzr8Y>.
- [29] LIU Shanfeng, ZHANG Weijian, YUAN Shaoguang, et al. A lightweight model for insulator defect detection based on vision–language modeling and prior knowledge in power systems[J]. *Processes*, 2025, 13(11): 3714.
- [30] ZHAO Wenqing, DING Yingxue, ZHANG Le, et al. CMKR-PBDM: a transmission line pin-missing bolts detection method based on cross-media and knowledge reasoning[J]. *IEEE transactions on power delivery*, 2025, 40(2): 1030–1039.
- [31] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9:

- learning what you want to learn using programmable gradient information[C]//Computer Vision–ECCV 2024. Cham: Springer Nature Switzerland, 2024: 1–21.
- [32] CHEN Hui, CHEN Kai, DING Guiguang, et al. YOLOv10: real-time end-to-end object detection[C]//Advances in Neural Information Processing Systems 37. Vancouver: NeurIPS, 2024: 107984–108011.
- [33] TIAN Yunjie, YE Qixiang, DOERMANN David. YOLOv12: attention-centric real-time object detectors [EB/OL]. (2025–09–19)[2025–12–02]. <https://openreview.net/forum?id=gCvByDI4FN>.
- [34] LONG Yang, GONG Yiping, XIAO Zhifeng, et al. Accurate object localization in remote sensing images based on convolutional neural networks[J]. *IEEE transactions on geoscience and remote sensing*, 2017, 55(5): 2486–2498.
- [35] LI Ke, WAN Gang, CHENG Gong, et al. Object detection in optical remote sensing images: a survey and a new benchmark[J]. *ISPRS journal of photogrammetry and remote sensing*, 2020, 159: 296–307.
- [36] 赵文清, 赵振寰, 巩佳潇. 结合倒残差自注意力机制的遥感图像目标检测[J]. *智能系统学报*, 2025, 20(1): 64–72.
ZHAO Wenqing, ZHAO Zhenhuan, GONG Jiaxiao. Remote sensing image object detection based on inverted residual self-attention mechanism[J]. *CAAI transactions on intelligent systems*, 2025, 20(1): 64–72.
- [37] ZHAO Wenqing, KANG Yijin, CHEN Hao, et al. Adaptively attentional feature fusion oriented to multiscale object detection in remote sensing images[J]. *IEEE transactions on instrumentation and measurement*, 2023, 72: 5008111.
- [38] YUE Chenke, YAN Junhua, ZHANG Yin, et al. SCFNet: Semantic correction and focus network for remote sensing image object detection[J]. *Expert systems with applications*, 2023, 224: 119980.
- [39] REN Shougang, FANG Zhiruo, GU Xingjian. A cross stage partial network with strengthen matching detector for remote sensing object detection[J]. *Remote sensing*, 2023, 15(6): 1574.
- [40] QIN Zequn, ZHANG Pengyi, WU Fei, et al. FcaNet: frequency channel attention networks[C]//2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2022: 763–772.
- [41] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531–11539.

作者简介:



赵文清, 教授, 博士, 主要研究方向为人工智能与图像处理。获河北省科技进步二等奖、三等奖各 1 项。发表学术论文 50 余篇。E-mail: zhao-wenqing@ncepu.edu.cn。



张乐, 硕士研究生, 主要研究方向为人工智能与图像处理。E-mail: 220232221016@ncepu.edu.cn。