

# 基于线性全局注意力与专家几何知识的防振锤 缺陷级联检测

黄志威<sup>1,2</sup>, 邹国锋<sup>1,2†</sup>, 傅桂霞<sup>1,2</sup>, 陈玉山<sup>1,2</sup>, 王伟健<sup>1,2</sup>, 高明亮<sup>1,2</sup>

(1. 山东理工大学 山东省新型配用电技术与装备重点实验室, 山东 淄博 255000;

2. 山东理工大学 电气与电子工程学院, 山东 淄博 255000)

**摘要:** 防振锤是输配电线路无人机航拍巡检中重点关注的金具之一。受制于无人机巡检安全距离与拍摄视角, 巡线图像中防振锤目标具有像素占比少、缺陷特征差异不明显的特点, 给防振锤的细粒度缺陷检测带来了严峻挑战。为此, 提出基于线性全局注意力与专家几何知识的防振锤缺陷级联检测方法, 将防振锤细粒度缺陷检测任务解耦为定位和分类两个阶段。在定位阶段, 提出基于线性全局注意力的轻量化定位网络, 在保证实时性的同时增强网络的全局上下文感知能力。在缺陷分类阶段, 提出专家几何知识嵌入分类网络, 将防振锤缺陷分类中的专家几何知识转化为显式特征表示, 并嵌入深度学习网络, 增强网络对缺陷类别的语义理解深度。最后, 在自建数据集上实验表明: 第一级定位网络实现 95% 召回率下 84.1% 的定位精确率, 第二级缺陷分类准确率达 93.42%, 验证所提出方法在细粒度缺陷检测中的有效性。

**关键词:** 输配电线路; 防振锤; 级联框架; 缺陷检测; 线性全局注意力; 专家知识

中图分类号: TP391 文献标志码: A

DOI: 10.13195/j.kzyjc.2026.0092

引用格式: 黄志威, 邹国锋, 傅桂霞, 等. 基于线性全局注意力与专家几何知识的防振锤缺陷级联检测 [J]. 控制与决策.

## A cascaded defect detection method for vibration dampers based on linear global attention and expert geometric knowledge

HUANG Zhi-wei<sup>1,2</sup>, ZOU Guo-feng<sup>1,2†</sup>, FU Gui-xia<sup>1,2</sup>, CHEN Yu-shan<sup>1,2</sup>, WANG Wei-jian<sup>1,2</sup>, GAO Ming-liang<sup>1,2</sup>

(1. Shandong Provincial Key Laboratory of New Power Distribution & Utilization Technology and Equipment, Shandong University of Technology, Zibo 255000, China; 2. School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255000, China)

**Abstract:** Vibration dampers are critical hardware components and high-priority inspection targets in UAV-based power line surveys. Constrained by safety distances and camera perspectives, vibration dampers in inspection images often exhibit small pixel occupancy and subtle defect features, posing significant challenges for fine-grained defect detection. To address this, we propose a cascaded defect detection method for vibration dampers based on linear global attention and expert geometric knowledge, decoupling the fine-grained detection task into localization and classification stages. In the localization stage, a lightweight linear global attention-based network is developed to enhance global contextual awareness while maintaining real-time performance. In the defect classification stage, an expert geometric knowledge embedded network is presented to formalize geometric priors into explicit feature representations, augmenting the semantic comprehension of defect categories. Experimental results on a custom dataset show that the localization network achieves 84.1% precision at a 95% recall rate, and the classification accuracy reaches 93.42%, validating the efficacy of the proposed method for fine-grained defect detection.

**Keywords:** power transmission and distribution lines; vibration dampers; cascaded framework; defect detection; linear global attention; expert knowledge

收稿日期: 2026-01-26; 录用日期: 2026-04-30.

基金项目: 国家自然科学基金项目 (52377110); 山东省自然科学基金项目 (ZR2022QE100).

责任编辑: 张丹.

†通信作者. E-mail: zgf841122@163.com.

## 0 引言

防振锤作为抑制导线微风振动的核心金具<sup>[1]</sup>,长期暴露于野外环境易发生形变或脱落,威胁输电线路安全,因此常态化巡检至关重要<sup>[2]</sup>.目前,无人机智能化巡检因具有受地形条件影响小、巡检范围广、作业效率高等特点已成为运维的主流手段<sup>[3,4]</sup>.但面对海量巡检图像,人工判读不仅效率低下,且因主观疲劳易导致漏检<sup>[5]</sup>.因此,需要引入智能化识别方法实现巡检图像的自动、高效解译,提升电网智能化运维的可靠性与效率<sup>[6]</sup>.

面向巡线图像的防振锤识别技术历经了基于传统图像处理到深度学习的演进过程<sup>[7]</sup>.传统方法多依赖手工特征工程,利用阈值分割<sup>[8]</sup>、边缘检测<sup>[9]</sup>等方法锁定防振锤位置后,再提取相应区域的颜色<sup>[10]</sup>、纹理<sup>[11]</sup>等特征,最后利用逻辑规则和 SVM<sup>[12]</sup>等方法进行缺陷判别.这种基于专家经验的硬编码虽然具有较强的可解释性,但却存在泛化能力不足的局限<sup>[13]</sup>.

深度学习通过构建深层神经网络实现特征的自动提取,已成为现代目标检测任务的核心驱动力.在防振锤识别任务中,根据是否生成候选框,主流检测架构可分为双阶段(Two-stage)与单阶段(One-stage)两类.在 Two-stage 研究中,文献[14]基于 Cascade R-CNN 架构,利用递归特征金字塔优化特征表示,实现了防振锤的高精度检测,但受制于候选框生成机制,推理速度难以满足电力巡检的实时性要求. One-stage 网络则通过同步回归目标位置与类别信息,显著提升了推理效率,在防振锤实时检测领域得到广泛应用.文献[15]通过在 EfficientDet 中引入高分辨率特征层,增强了对小尺度防振锤的感知能力.文献[16]在 YOLOv5 中集成具有自注意力机制的 AFF-Transformer 与空间通道注意力模块,有效提升了密集场景下防振锤的检测性能,但自注意力机制的高计算复杂度限制了模型的运行速率.文献[17]提出的 YOLOv8-SPH 结合多尺度特征提取与自适应稀疏剪枝技术,在保持检测精度的同时,显著降低了参数量与计算负担.尽管上述框架在防振锤定位与缺失检测任务中表现优异,但在应对更精细的细粒度缺陷时,模型面临判别性特征提取不足的挑战,难以在统一特征空间内兼顾稳健的定位信息与支撑精准分类的微观细节.

综上所述,为解决目标占比小、缺陷类间差异不显著导致的防振锤细粒度缺陷检测难题,本文提出了基于线性全局注意力与专家几何知识的防振锤缺陷级联检测方法.首先,构建防振锤定位与缺陷分类

数据集,为模型的训练与评估提供数据基础.其次,为克服级联架构中首级网络对系统整体推理速度的影响,提出基于线性全局注意力的轻量化定位网络(Linear Global Attention YOLO, LGA-YOLO),通过线性全局注意力模块与局部特征提取模块,在确保目标精准定位的同时,降低计算开销,满足实时检测需求.最后,针对纯数据驱动模型在复杂环境下难以建立稳定的判别边界问题,提出专家几何知识嵌入分类网络(Geometric Expert Knowledge Classification Network, GEK-Net),将专家几何知识转化为显式特征表示,并通过输入端、特征层及决策层嵌入深度学习网络,增强网络对缺陷类别的语义理解深度.

## 1 防振锤数据集构建与特征分析

本文所用原始图像数据采集于 2020 年 3 月至 2021 年 12 月之间,使用 DJI Mavic 2 Pro 无人机,搭载 1 英寸 2000 万像素传感器,采集范围覆盖 35 kV 与 110 kV 架空输电线路,共 33 条、1672 个基塔,时间跨度确保样本涵盖不同季节与光照条件.采集场景包括乡村、河岸、农田、城区、住宅及工业区等典型场景,包含了桶型、音叉型、扭转型等主流防振锤型号<sup>[18]</sup>.基于无人机采集的图像,标注构建了防振锤定位数据集与缺陷分类数据集.

### 1.1 防振锤定位数据集

防振锤定位数据集包含分辨率为 5472×3648 像素的巡检图像 6025 张,标注了 28544 个防振锤目标.定位数据集不区分具体缺陷类别,统一标注为“Damper”,数据集的部分图像如图 1 所示.

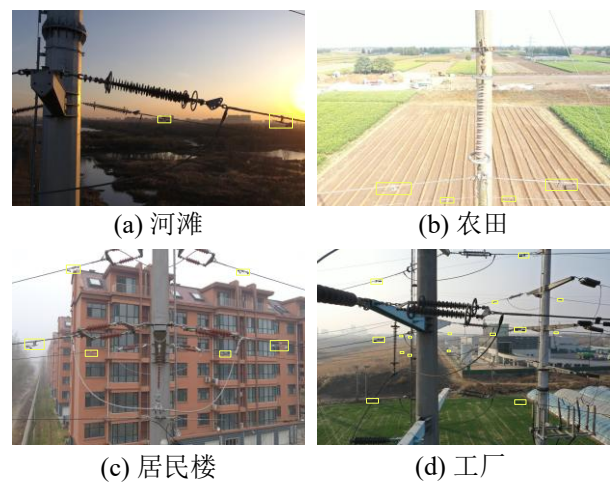


图1 防振锤定位数据集不同场景

图 2 展示了防振锤定位数据集中防振锤目标在整幅图像中的像素占比分布.其中,99.2%的防振锤目标像素占比小于 2%,57.39%的防振锤目标像素占比不足 0.2%,呈现小目标特点,这导致防振锤特征难以得到有效表达,增加了缺陷类别细分时的错

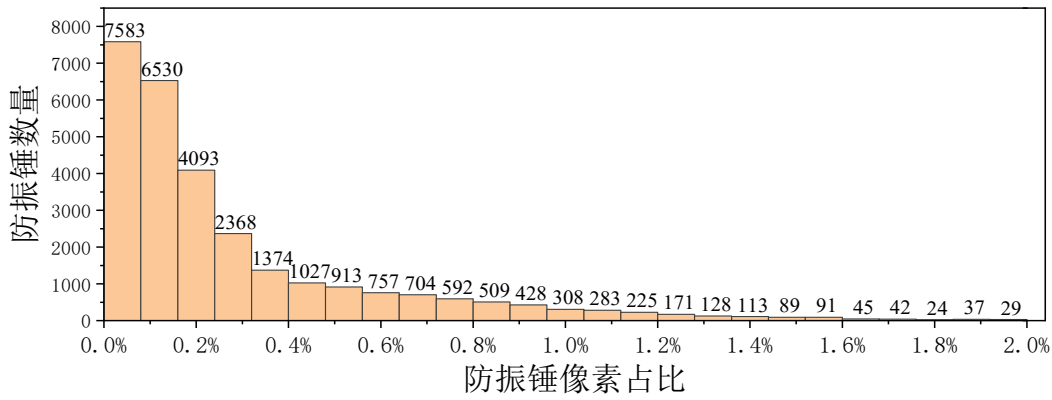


图2 防振锤像素占比

检风险.

### 1.2 缺陷分类数据集

为实现对防振锤状态的精细化表征, 本文基于其物理结构与失效机理定义了缺陷类别. 图3展示了防振锤的组成部件, 包括线夹、钢绞线、锤头及紧固件. 依据各组件的典型失效状态, 本文定义了正常类别和4种缺陷类别.

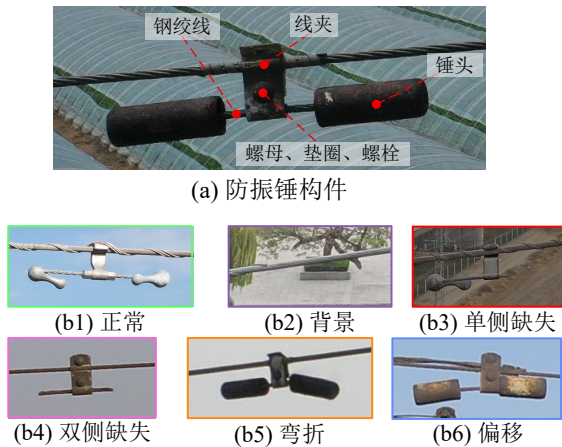


图3 防振锤构件与缺陷类别展示

此外, 考虑到级联检测框架在定位阶段产生背景误检, 缺陷分类数据集增设了背景类. 背景类样本提取自定位网络在训练集上的高置信度误检窗口, 包含塔材交错结构、导线和线夹等易与防振锤混淆的背景区域. 通过引入背景类, 分类网络能够区分防振锤目标与干扰背景, 进而抑制级联误差传播. 缺陷分类数据集类别定义和数量如表1所示.

## 2 防振锤细粒度缺陷级联检测方法

### 2.1 防振锤定位与分类任务解耦的必要性分析

现有防振锤的缺陷检测方法多采用单级的检测框架, 试图利用单一网络同时实现防振锤定位与缺陷分类, 然而, 此类架构在处理细粒度缺陷检测任务时, 检测精度往往难以达到预期.

表1 缺陷分类数据集各类别定义和数量

类别	类别定义	样本数
正常	连接完好、无松动、无缺陷	3978
单侧缺失	一侧锤头脱落	3664
双侧缺失	两侧锤头均脱落	2758
弯折	钢绞线强度下降导致整体弯曲	2019
偏移	线夹松动造成锤头相对于线夹位移	2141
背景	非防振锤区域(误检窗口)	4877

为探究单级的检测框架在细粒度缺陷检测任务中定位与分类任务的难度差异, 本文在 640×640、960×960 和 1280×1280 三种分辨率下, 对比了经典的 Ones-tage 目标检测方法 YOLOv8 在以下两种任务模式中的表现:

多类别检测: 防振锤标注了具体的缺陷类别, 模型在定位防振锤的同时也需对防振锤缺陷类别进行分类, 即标准的端到端检测任务.

单类别检测: 将所有状态防振锤标注为同一类别, 模型仅区分防振锤与背景, 不涉及具体的缺陷类别细分.

图4展示了 YOLOv8 在两种任务模式的 P-R 曲线, 并对比了召回率为 95% 时对应的精确率. 由图4可知, 在召回率为 95% 时, 进行单类别检测时模型在 640×640 的较低分辨率下就能有 83.2% 的较高精确率. 而进行多类别检测任务时, 在 1280×1280

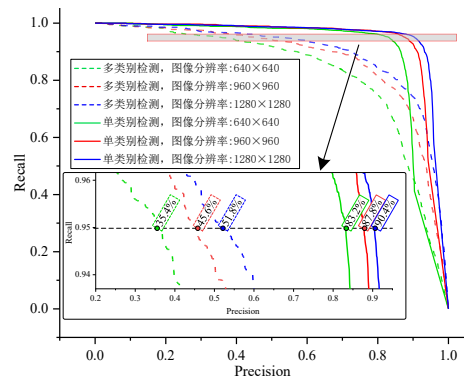


图4 YOLOv8 在两种任务模式下性能对比

的较高分辨率下精确率也仅达 51.8%，性能出现显著落后。

上述结果表明，单级架构的失效根源并非单纯的分辨率不足，而在于任务耦合导致的特征表征冲突，定位任务依赖大感受野的全局上下文感知，而细粒度缺陷判别则依赖局部细节特征。在共享特征空间中，不同状态防振锤的共性结构特征往往占据主导，而缺陷判别性特征易被淹没。即便将分辨率提升至  $1280 \times 1280$ ，单级网络仍因同时承担双重任务而难以有效提取细粒度判别特征。因此，将防振锤细粒度缺陷检测任务进行解耦，分别聚焦定位与分类任务是提升缺陷检测性能的关键。

## 2.2 级联检测方法

基于上述解耦思路，本文提出防振锤缺陷级联检测方法，该级联方法的设计初衷在于平衡检测效率与分类精度。首先，定位网络作为第一级任务，其计算开销直接制约了整体算法的实时性表现。考虑到防振锤定位高度依赖大感受野下的全局上下文信息，本文设计了 LGA-YOLO。通过引入线性全局注意力在控制模型复杂度的前提下，增强网络全局感

知能力。

其次，针对电力巡检的行业标准 DL/T 2691—2023 中 85% 的召回率要求，LGA-YOLO 采用较低的置信度阈值以减少漏检。然而，这不可避免地引入了塔材交错、导线等易与防振锤混淆的背景误检，加之细粒度缺陷类别间特征差异细微，纯数据驱动模型难以建立稳定的判别边界，因此本文提出 GEK-Net 完成分类任务，利用防振锤缺陷分类中的专家几何知识增强网络对缺陷语义的理解深度，并通过增设的背景类过滤第一级产生的背景误检。

最终，基于 LGA-YOLO 与 GEK-Net 构建了如图 5 所示的级联检测方法。对于原始高分辨率巡线图像，首先将其缩放至低分辨率输入进 LGA-YOLO 获得防振锤候选框，并设置较低置信度阈值以保留潜在防振锤区域。随后，依据第一级输出的边界框坐标在原始的高分辨率图像上裁剪出防振锤预测子图。最后将子图输入进 GEK-Net 执行细粒度分类，通过背景类别判别滤除第一级引入的误检干扰，输出防振锤缺陷细粒度检测结果。

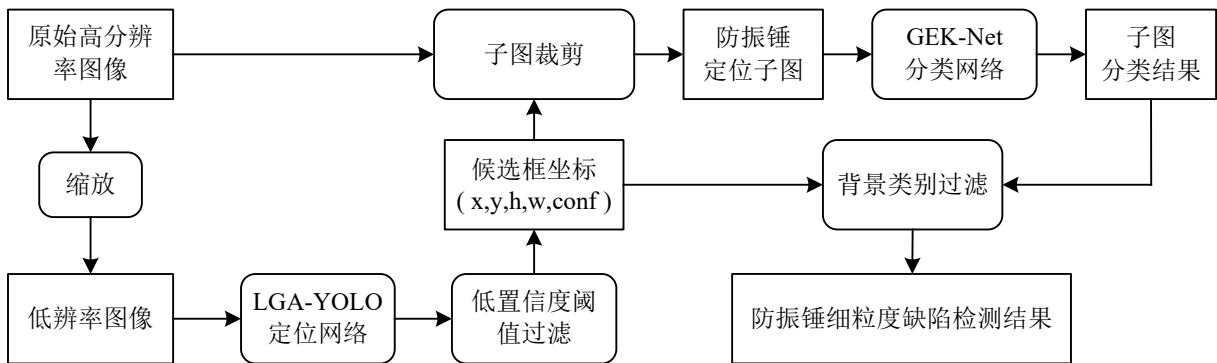


图5 级联检测方法整体流程

## 2.3 基于线性全局注意力的轻量化定位网络

本文提出的 LGA-YOLO 主要由主干网络、颈部网络和检测头三部分组成，结构如图 6 所示。

主干网络由 LGA-Block 和 LFE-Block 组合而

成，在浅层阶段主要堆叠 LFE-Block，利用深度可分离卷积快速提取图像的通用纹理特征。在网络的深层阶段，交错插入 LGA-Block，利用其全局感受野捕捉防振锤结构信息与长距离语义依赖。

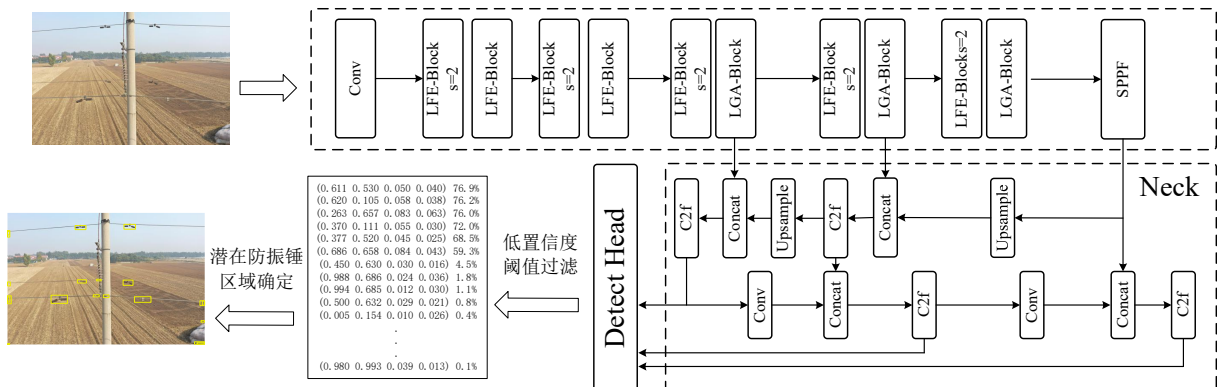


图6 LGA-YOLO 结构图

颈部网络采用 PANet 结构进行特征融合, 通过上采样与 C2f 模块, 将深层的语义信息与浅层的空间信息充分交互, 增强对多尺度防振锤的适应性。

最终, 将融合特征图经检测头处理, 并行输出边界框坐标预测与防振锤置信度。

### 2.3.1 线性全局注意力模块

在防振锤定位中, 网络高度依赖大感受野下的全局上下文信息以区分防振锤目标与背景。卷积神

经网络受限于局部感受野, 难以提取全局上下文信息。而标准 Transformer 虽具备全局建模能力, 但其庞大的计算开销限制了实时性能。

为此, 本文设计了线性全局注意力模块 (LGA-Block), 如图 7 所示。该模块借鉴深度可分离卷积的通道独立假设, 将注意力计算限制在各通道内部独立进行。通过取消通道间映射并仅在空间维度建立像素级关联, 将计算复杂度降至线性。

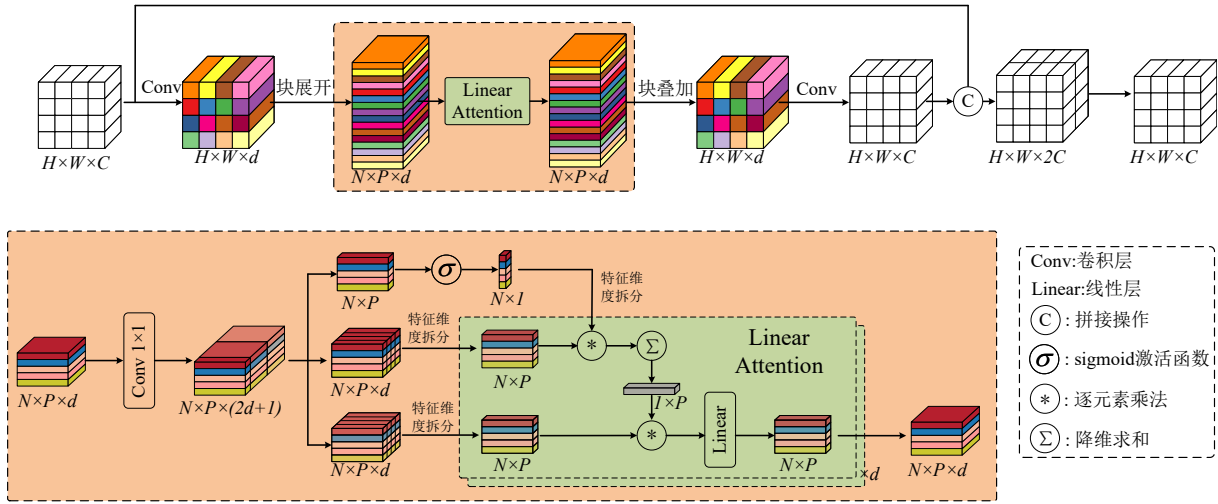


图7 LGA-Block 模块

对于输入特征图, 首先通过块展开将其重塑为序列信号  $X \in \mathbb{R}^{P \times N \times d}$ , 其中  $P$  为 Patch 数量,  $N$  为通道数,  $d$  为嵌入维度。LGA-Block 的推理过程包含以下三个步骤:

首先生成异构特征, 利用共享权重的线性变换一次性生成查询 ( $Q$ )、键 ( $K$ ) 和值 ( $V$ ):

$$[Q, K, V] = WX. \quad (1)$$

其中  $Q \in \mathbb{R}^{P \times N \times 1}$ , 在通道维度压缩以提取空间显著性,  $K, V \in \mathbb{R}^{P \times N \times d}$  保留完整维度以承载语义信息。

然后进行通道解耦线性交互, 在  $d$  个特征维度内独立并行注意力计算。对于第  $c$  个通道的特征分量  $Q_c, K_c, V_c \in \mathbb{R}^{P \times N}$ , 执行以下操作:

利用 Sigmoid 处理  $Q$  生成空间权重分布  $\alpha \in \mathbb{R}^{N \times 1}$ ; 将  $\alpha$  与  $K_c$  进行加权聚合, 提取该通道的全局上下文向量  $s_c \in \mathbb{R}^{1 \times P}$ :

$$s_c = \alpha^\top K_c. \quad (2)$$

利用  $s_c$  对  $V_c$  进行特征增强, 并通过线性层映射回原特征空间:

$$Y_c = \text{Linear}(s_c \otimes V_c) \in \mathbb{R}^{N \times P}. \quad (3)$$

最后经过并行操作获取的全局上下文特征经线性层融合后, 通过块叠加操作还原为原始张量形状  $\mathbb{R}^{H \times W \times C}$ , 并利用残差与卷积组合将信息融合回原

始分辨率。

### 2.3.2 局部特征提取模块

图 8 展示了倒残差结构的 LFE-Block, LFE-Block 采用升维—卷积—降维的处理流程, 首先利用  $1 \times 1$  卷积扩展通道维度, 然后采用  $3 \times 3$  深度可分离卷积提取局部空间特征, 并根据网络层级设置步长以实现特征保持或下采样, 最后经  $1 \times 1$  卷积投影回低维空间。特别地, 仅在步长为 1 且输入输出维度一致时引入残差连接, 确保深层网络的信息传递。

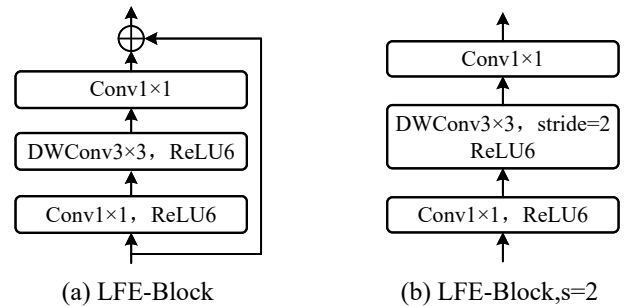


图8 LFE-Block 模块

## 2.4 专家几何知识嵌入分类网络

在防振锤缺陷的人工判读中, 领域专家并非依赖像素级纹理统计, 而是遵循三层几何认知逻辑: 关注边缘轮廓以判断结构性缺陷; 聚焦防振锤本体区域以排除背景干扰; 基于对称性分析与轴向分布规律以

判定缺陷状态.

为将专家几何知识转化为深度学习可学习的特征约束, 弥补纯数据驱动模型在细粒度分类中的不足, 本文提出专家几何知识嵌入分类网络 (GEK-Net). 如图9所示, 对于输入的防振锤定位子图, 该网络首先通过专家几何知识提取模块输入图像进行处理, 依次提取边缘特征、空间注意力掩码和几何投影特征, 以映射人工判读中的三层认知逻辑, 并通过输入端、特征层和决策层三级嵌入深度分类网络, 实现背景误检过滤与缺陷精确判别.

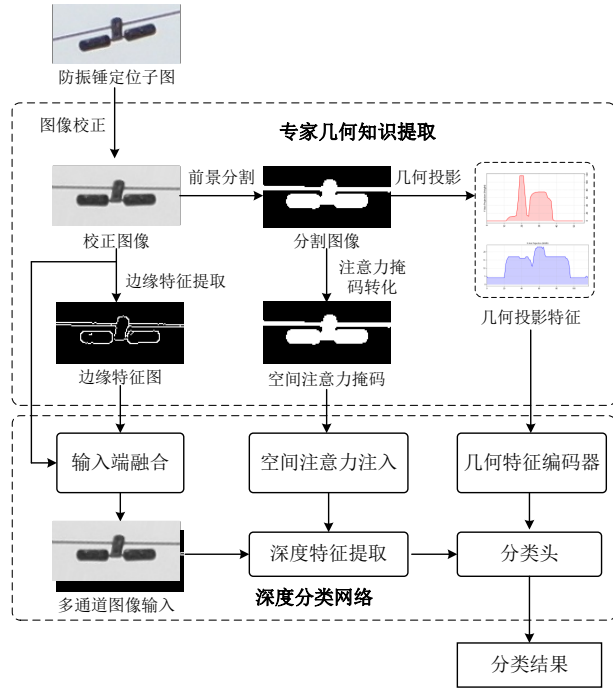


图9 GEK-Net 网络结构图

### 2.4.1 专家几何知识提取模块

为有效提取专家几何知识, 本文将人工判读流程建模为可计算的算法流程. 对于输入的防振锤定位子图, 首先执行几何校正以实现视角的校正, 然后通过前景分割排除环境背景以聚焦目标前景, 最后利用几何投影将空间质量分布转化为特征. 上述步骤依次生成校正后的图像与边缘特征、空间注意力掩码和几何投影特征, 共同构成多维度的专家几何知识表征. 具体操作如下:

(1) 几何校正, 针对无人机视角多变导致的子图旋转偏差的问题, 本文基于边缘梯度进行几何校正.

首先, 利用一阶微分算子提取梯度, 经非极大值抑制与双阈值滞后处理获得边缘点集. 随后, 采用改进的RANSAC算法进行导线估计, 为抑制航拍图像的轻微扰动与防振锤目标的尺度变化引入两项约束:

自适应距离阈值: 使用与图像高度 $H$ 线性相关的动态阈值 $\tau_{dist}$ 以适配不同拍摄尺度.

$$\tau_{dist} = \max(\gamma \cdot H, \tau_{min}). \quad (4)$$

其中,  $\gamma$ 为比例系数,  $\tau_{min}$ 为最小容差.

斜率约束: 基于导线水平分布先验, 限制拟合斜率 $|k| \leq \tan(60^\circ)$ , 剔除垂直方向的干扰, 缩小RANSAC解空间.

最后, 计算导线偏角 $\alpha = \arctan(\hat{k})$ 并绕图像几何中心 $(c_x, c_y)$ 进行旋转校正. 对于产生的三角空白区域, 采用边界复制填充以保持纹理连续, 获得对齐后的图像 $I_{rot}$ . 各类别图像几何矫正结果如图10所示.



图10 各类别几何矫正结果

(2) 前景分割, 因为无人机图像采集时间不同与环境, 巡线航拍图像存在多种成像条件, 为增强场景自适应能力本文采用多策略互补的投票机制确定前景区域, 分割阈值的三种算法如下:

全局统计分割: 计算整幅ROI灰度直方图, 基于最大类间方差准则自动确定全局阈值 $T_{global}$ :

$$T_{global} = \arg \max_t [\omega_0(t)\omega_1(t)(\mu_0(t) - \mu_1(t))^2]. \quad (5)$$

其中 $\omega_0$ 和 $\omega_1$ 为两类像素概率,  $\mu_0, \mu_1$ 为两类类内灰度均值. 该策略利用整体统计特性抑制全局噪声, 适合光照均匀、对比度显著的场景.

几何直方图分割: 对直方图进行高斯平滑后, 连接峰值点与最远端点构成三角形, 取最大垂距点对应灰度作为阈值 $T_{geo}$ :

$$T_{geo} = \arg \max_t d(t, \text{line}_{\text{peak-far}}). \quad (6)$$

其中 $d(\cdot)$ 表示点到直线的垂距. 该策略利用几何极值定位, 适合阴天等弱对比度场景.

局部自适应分割: 以每个像素为中心取 $m \times m$ 邻域, 计算局部均值 $\mu_{local}$ 作为此像素的阈值:

$$T_{local}(x, y) = \mu_{local}(x, y) - C. \quad (7)$$

其中 $C$ 为微调常数. 该策略通过局部对比度保留细

节结构, 适用于局部光照不均的场景。

最后通过逐像素多数投票, 仅保留被至少两路算法确认为前景的区域. 该互补机制生成的空间注意力掩码  $M_{prior}$  精准表征防振锤前景区域。

(3) 几何投影, 在专家判读经验中, 类别分类依赖于对防振锤结构对称性与轴向分布的观察. 为把几何结构知识转化为可计算特征, 本文将空间注意力掩码  $M_{prior}$  分别沿 X 轴和 Y 轴对二值掩码进行积分投影. 其中, X 轴投影  $P_x$  刻画横向对称性与结构完整性, Y 轴投影  $P_y$  则表征防振锤的下垂程度及线夹纵向位置, 各类别的分割与投影结果如图 11 所示。

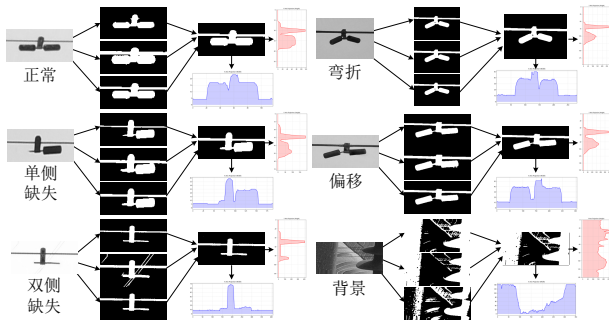


图11 各类别分割与几何投影结果

综上, 几何校正后图像的边缘特征、空间注意力掩码与几何投影特征共同构成了专家几何知识表征, 实现了特征从视觉感官向几何结构属性的转化。

#### 2.4.2 深度分类网络

为充分利用专家几何知识, 本文构建了深度分类网络, 如图 12 所示. 该网络以 LGA-Block 与 LFE-Block 的堆叠为主干提取深层特征, 并通过输入端、特征层、决策层的三级融合机制, 将多维度的专家几何知识表征嵌入深度分类网络:

在输入端进行模态融合, 将几何校正后的图像  $I_{rot}$  与边缘特征  $I_{edge}$  在通道维度拼接, 构建复合输入张量  $X_{input} \in \mathbb{R}^{H \times W \times 4}$ . 该策略在输入阶段即引入显式结构先验, 引导底层卷积核在提取纹理特征的同时, 增强对防振锤边缘的感知能力。

在深层特征提取阶段, 将空间注意力掩码  $M_{prior}$  作为硬注意力门控以避免背景噪声. 将高分辨率掩码  $M_{prior}$  下采样至与网络第4层相同尺寸, 得到  $M'$ . 随后, 通过元素级乘法将空间注意力注入特征图:

$$\hat{F}_4 = F_4 \otimes (1 + \eta \cdot M') \quad (8)$$

式中,  $\eta$  表示增强系数,  $\otimes$  表示哈达玛积. 该机制通过防振锤位置边界的硬约束抑制了掩码外部的背景响应, 迫使网络聚焦于防振锤本体区域。

在决策层通过独立的几何特征编码器处理一维投影信号  $P_x, P_y$ . 该分支利用多层全连接网络将几

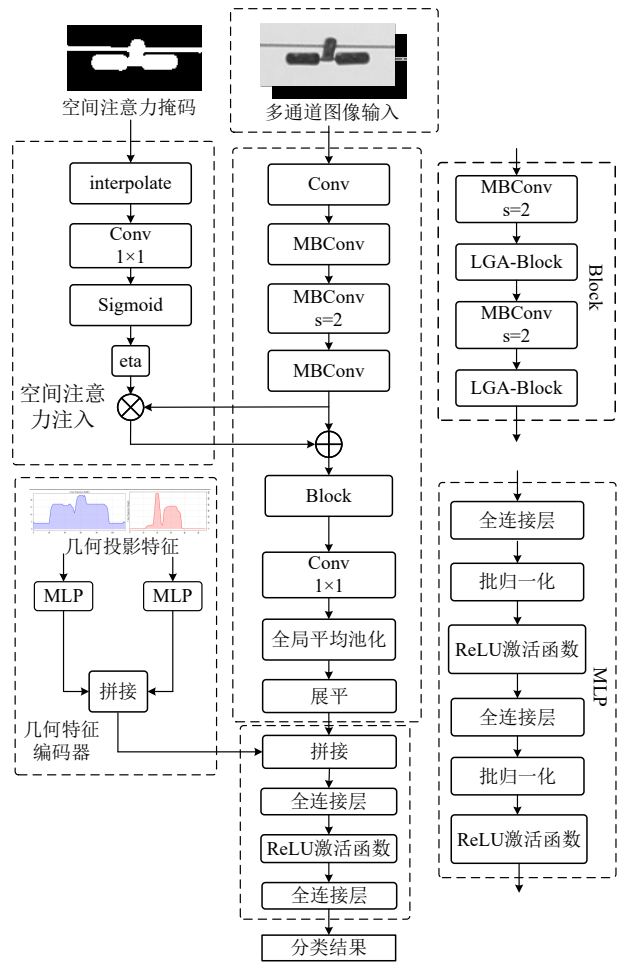


图12 深度分类网络结构图

何特征映射为高维语义向量  $V_{geo}$ , 以提取防振锤的跨度分布与垂度特征. 最终, 在决策层将视觉特征  $V_{vis}$  与几何特征  $V_{geo}$  拼接, 实现类别预测:

$$H = ReLU(W_1 Concat(V_{vis}, V_{geo}) + b_1), \quad (9)$$

$$\hat{y} = Softmax(W_2 H + b_2). \quad (10)$$

综上, 该架构实现了视觉表现特征与专家几何知识的深度耦合. 通过多层级嵌入, 网络在利用空间拓扑逻辑有效剔除背景干扰的同时, 显著增强了对形变、缺失等细粒度缺陷的判别鲁棒性。

### 3 实验结果与分析

#### 3.1 实验设置

##### (1) 数据集设置

本文采用第 1 章构建的两套数据集分别评估定位与分类性能, 其中防振锤定位数据集包含 6025 张巡检图像, 分辨率采用保持长宽比的 Letterbox Resizing 策略缩放至  $640 \times 640$  像素. 按照 8:2 的比例划分为训练集与测试集. 为验证模型在复杂工况下的鲁棒性, 从测试集中筛选出典型困难场景子集, 具体定义如表 2 所示。

缺陷分类数据集涵盖正常、单侧缺失、双侧缺

表2 典型困难场景子集定义与数量

子集类型	定义	图像数量
多目标	目标数量 $\geq 4$	423
小目标	目标像素占比 $< 0.2\%$	434
多尺度	最大最小目标面积之比 $> 10$ 且目标面积变异系数 $> 0.5$	99

失、弯折、偏移及背景 6 类样本, 总计 19437 张, 并重采样至  $32 \times 80$  像素, 按 8:2 比例划分训练集与测试集.

## (2) 软硬件环境与参数设置

实验在 Windows10 系统上进行, GPU 为英伟达 RTX 4080 SUPER 显卡和 32GB 内存. 软件环境包括 PyTorch 2.0.0、CUDA 11.8 和 Python 3.10.

在定位模型训练中, Batch Size 设为 16, 总训练轮数为 1000, 并引入早停机制, 即连续 100 个 Epoch 性能无提升时自动终止训练. 采用 SGD 优化器, 初始学习率设定为 0.05, 最终通过 0.5 的衰减因子调整至 0.025.

在图像分类模型训练中, Batch Size 设置为 48, 总训练轮数为 120, 其中包括 5 个冻结轮次. 采用 SGD 优化器, 初始学习率设为 0.01, 并遵循余弦退火调度规律进行衰减, 权重衰减系数为  $5 \times 10^{-4}$ .

## 3.2 评价指标

防振锤定位采用平均精度 (Average Precision, AP) 作为核心评价标准<sup>[19]</sup>. AP 定义为精确率-召回率 (Precision-Recall) 曲线下的面积, 能够反映模型整体检测性能, 其计算公式为:

$$AP = \int_0^1 P(R) dR. \quad (11)$$

其中, P 代表精确率, R 代表召回率.

缺陷分类实验采用准确率 (Acc)、精确率 (P) 和召回率 (R) 进行评估<sup>[20]</sup>.

本文采用参数量 (Param)、GFLOPs 和 FPS 评估模型的复杂度与检测速度. 其中, 参数量 (Parameters) 代表模型中全部可训练权重与偏置的总数, 以  $10^6$  (M) 为单位; GFLOPs 表示单张图像完成一次前向传播所需的浮点运算量, 以  $10^9$  为单位; FPS 表示模型每秒预测的图片数量, 用于衡量模型的实时性能.

## 3.3 防振锤目标定位实验

本节通过基准性能对比、高召回率下性能分析、挑战子集测试与可视化对比分析, 全面评估 LGA-YOLO 作为级联架构首级定位的有效性.

### 3.3.1 对比实验

为对比 LGA-YOLO 有效性, 本节在自建防振锤定位数据集上, 与轻量化配置下的 ResNet<sup>[21]</sup> 网络和

YOLO 网络<sup>[22]</sup>、轻量化网络 (MobileNet<sup>[23]</sup>、GhostNetV2<sup>[24]</sup>、StarNet<sup>[25]</sup>、FasterNet<sup>[26]</sup>、MobileMamba<sup>[27]</sup>) 以及基于 Transformer 的网络 (SwinTransformer<sup>[28]</sup>、EfficientViTv2<sup>[29]</sup>) 进行模型复杂度、检测精度与检测速度的全面对比, 结果如图 13 所示. 图中气泡的横、纵坐标分别代表检测 FPS 与 AP, 气泡直径大小对应模型的 GFLOPs.

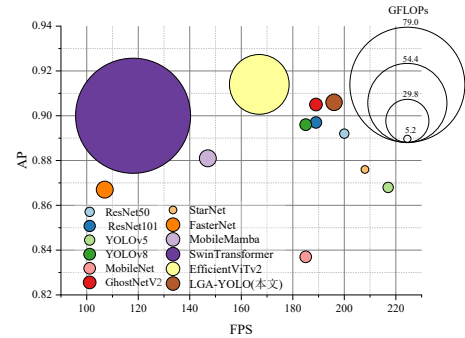


图13 不同模型性能对比

图 13 表明, 在检测精度方面, 本文提出的 LGA-YOLO 表现优异, 其 AP 值仅次于高计算复杂度的 EfficientViTv2, 而计算量仅约为其 1/4. 在检测速度方面, LGA-YOLO 的 FPS 达 196, 虽较 YOLOv5 略低 21 帧, 但检测精度提升了 3.8 个百分点, 能够满足实时性检测需求. 综上所述, LGA-YOLO 在性能表现与资源开销之间取得了平衡, 其高效的推理速度与优秀的精度表现, 使其能够契合级联检测架构中第一级的性能要求, 展现出卓越的整体性能.

### 3.3.2 高召回率下精确率对比

定位模型作为级联架构的第一级, 需在保障高召回率的前提下抑制背景误检. 因此, 高召回率下的精确率更能直观反映第一级的工程可用性. 基于 DL/T 2691-2023 标准要求, 本文以 95% 召回率为基准, 并增设 97% 召回率验证更严苛安全等级下的性能边界. 各算法的精确率及 P-R 曲线如图 14, 直观展示了第一级定位模型在不同召回率要求下的精确率.

图 14 表明, LGA-YOLO 在 95% 召回率下的精确率达到 84.10%, 与最优的 EfficientViTv2 相比落后 0.5 个百分点. 在更严苛的 97% 召回率下, LGA-YOLO 的召回率达到最高的 79.55% 精确率, 仍保持较高性能.

### 3.3.3 挑战子集对比

为进一步评估模型在复杂工况下的可靠性, 本节在测试集中筛选三类挑战子集, 以对比各方法在面对目标密集、尺度剧变及微小目标时的性能, 对比结果如表 3 所示.

实验结果表明, LGA-YOLO 在三大挑战子集上

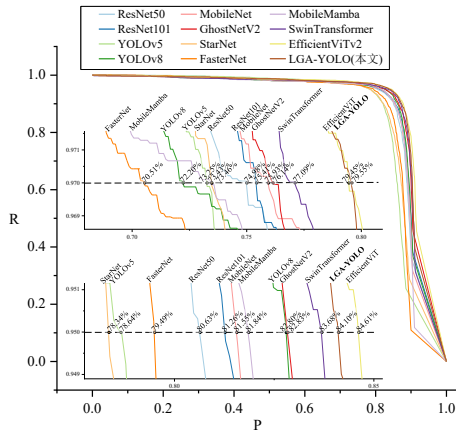


图14 防振锤目标定位方法 P-R 曲线对比

的表现与整体测试集上保持一致, 均取得了次优性能, 且在多目标与多尺度场景下表现出极强的竞争力, 其 AP 值与 EfficientViTv2 仅有微小差距. 结果表明 LGA-YOLO 能够在复杂环境下保持稳健的检测性能, 为无人机电力巡检中的防振锤定位任务提供可靠保障.

表3 三类挑战子集结果对比 (AP)

Backbone	多目标	多尺度	小目标
ResNet50	0.835	0.600	0.767
ResNet101	0.844	0.605	0.779
YOLOv5	0.807	0.573	0.737
YOLOv8	0.841	0.595	0.775
MobileNet	0.840	0.598	0.773
GhostNetV2	0.853	0.618	<u>0.790</u>
StarNet	0.822	0.583	0.748
FasterNet	0.844	0.607	0.773
MobileMamba	0.827	0.586	0.754
SwinTransformer	0.845	0.613	0.787
EfficientViTv2	<b>0.864</b>	<b>0.625</b>	<b>0.808</b>
LGA-YOLO(本文)	<u>0.854</u>	<u>0.622</u>	0.787

### 3.3.4 可视化对比

为直观对比 LGA-YOLO 在实际场景下的定位性能, 图 15 展示了在置信度阈值设为 0.1 时本文与其他方法的对比结果. 图中数字为检测框置信度, 绿色框为正确检测, 红色框为背景误检, 粉色框为目标漏检.

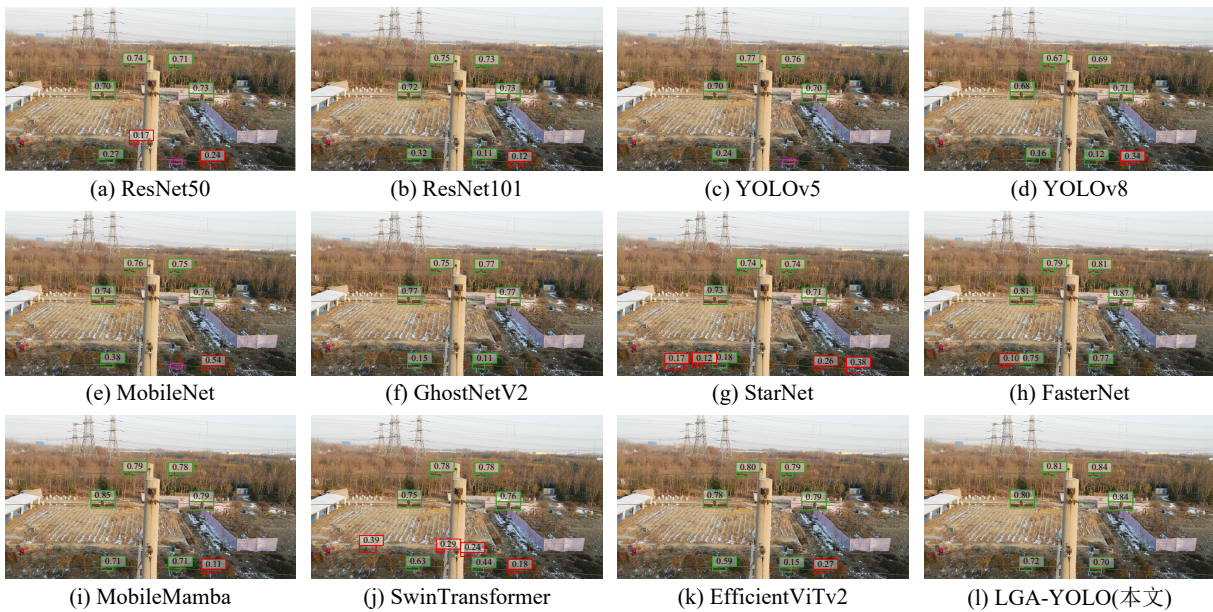


图15 可视化对比

图 15 表明, 本文方法能以显著高于其他方法的置信度精确锁定防振锤. 在检测完整性上, LGA-YOLO 成功捕获了图像中所有的目标, 有效规避了 YOLOv5、MobileNet 存在的漏检问题; 在背景鉴别力上, 相比于 YOLOv8 和 StarNet 等方法引入的背景误检, 本文方法展现出背景抑制能力, 未引入无关噪声的干扰.

### 3.4 缺陷分类实验

本节通过消融实验、基准模型对比及混淆矩阵可视化分析, 系统验证 GEK-Net 在防振锤缺陷分类任务中的有效性. 系统评估各专家几何知识对性能

的贡献, 并全面分析模型对各类精细化缺陷的辨识能力与分类鲁棒性.

#### 3.4.1 消融实验

为验证各专家几何知识对缺陷分类性能的贡献, 本节以由 LGA-Block 与 LFE-Block 的堆叠提取防振锤深度特征提取网络为 Baseline, 通过逐步引入边缘特征 ( $\mathcal{F}_{in}$ )、空间注意力注入 ( $\mathcal{M}_{prior}$ ) 及几何投影特征 ( $\mathcal{G}_{enc}$ ) 开展消融实验, 结果如表 4 所示.

由表 4 可知, 各模块均能有效提升缺陷分类性能. 其中, 引入边缘特征对性能提升最显著, 准确率较 Baseline 增长 3.21%, 验证了输入端边缘先验对细

表4 PPA-Net 分类网络消融实验结果

Baseline	$\mathcal{F}_{in}$	$\mathcal{M}_{prior}$	$\mathcal{G}_{enc}$	评价指标 (%)		
				Acc	P	R
√	-	-	-	89.10	89.13	88.99
√	√	-	-	92.31	92.21	91.61
√	-	√	-	90.67	90.65	90.23
√	-	-	√	90.85	90.67	90.51
√	√	√	√	<b>93.42</b>	<b>93.01</b>	<b>93.28</b>

微缺陷特征的强化作用. 几何投影特征与空间注意力也分别带来了 1.75% 与 1.57% 的准确率提升. 最终, 完整模型在三者协同下达到 93.14% 的最高准确率, 较 Baseline 提升 4.32%, 证明不同维度的专家几何知识在增强模型判别力方面具有良好的互补性.

### 3.4.2 对比实验

为验证所提分类网络的先进性, 本节与经典残差网络 (VGG11<sup>[30]</sup>、ResNet50)、轻量化架构网络 (MobileNet、StarNet) 及轻量化设置的 Vision Transformer<sup>[31]</sup> 模型进行对比, 结果如表 5 所示.

表5 缺陷分类方法对比结果

方法	Acc (%)	P (%)	R (%)	Param	GFLOPs	FPS
VGG11	90.75	90.43	90.80	128.79	0.500	102.94
ResNet50	76.32	76.51	74.54	23.52	0.218	304.88
MobileNet	<u>91.36</u>	<u>90.74</u>	<u>91.09</u>	2.23	<b>0.017</b>	<u>353.24</u>
StarNet	82.83	81.69	81.64	2.68	<u>0.023</u>	312.95
ViT	73.08	71.88	71.17	<b>0.40</b>	0.065	<b>492.52</b>
GEK-Net(本文)	<b>93.42</b>	<b>93.01</b>	<b>93.28</b>	<u>0.69</u>	0.033	232.72

由表 5 可知, 本文方法在保持 232.72 FPS 的实时性前提下, 准确率达到 93.42%, 全面超越了其他对比方法. 在轻量化水平上, 本文方法的参数量仅为 0.69M, 约占 MobileNet 的 31%, 且准确率领先其 2.06 个百分点, 实现了更小的模型体积与更高的识别精度. 相较于传统的网络 VGG11, 本文方法在参数量大幅压缩的同时, 准确率提升了 2.67%. 而 ViT 由于缺乏卷积神经网络的局部感性归纳偏置, 导致其辨识精度远低于本文方法. 综上所述, 所提算法在分类精度上表现出显著优势, 且具备实时运行效率, 能够有效完成缺陷分类任务.

### 3.4.3 混淆矩阵对比

为进一步分析模型对各细分缺陷类别的辨识精度及误判分布, 本节给出本文方法与其他模型的混淆矩阵可视化对比, 如图 16 所示.

图 16 表明, 本文方法的混淆矩阵对角线最为集中, 展现了最优的分类准确率. 实验同时也揭露了防振锤细粒度缺陷检测时的共性难题: 正常类别、弯折及偏移缺陷之间存在一定程度的相互混淆, 究其原因

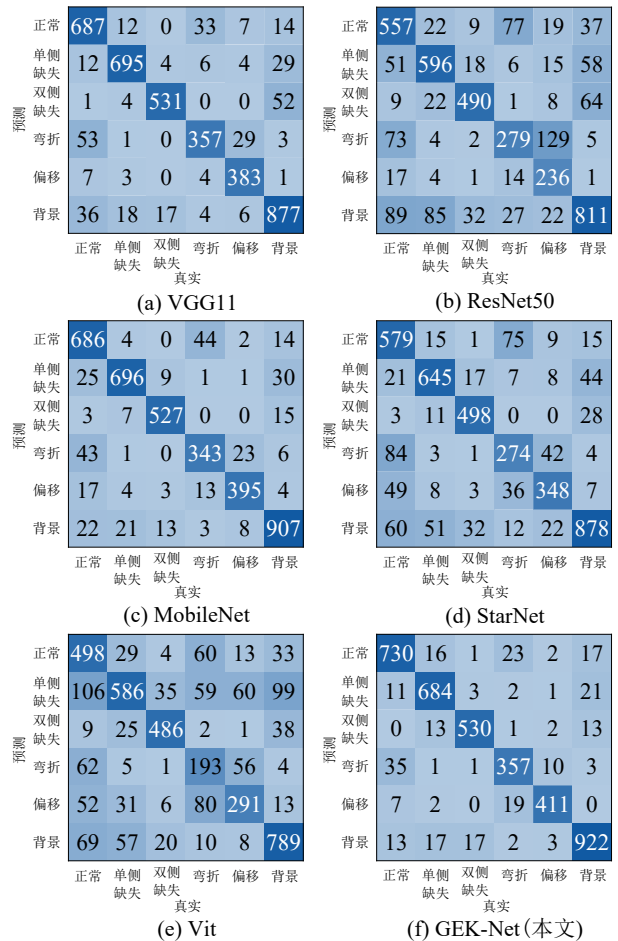


图16 混淆矩阵对比

因, 此类缺陷在视觉表征上具有极高的形态相似性, 其差异特征往往处于特征空间的判别边际, 增加了辨识难度. 即便如此, 本文方法凭借专家几何知识对结构拓扑的增强, 相较于基准模型显著降低了非对角线区域的离散分布, 证明了专家几何特征在处理边际模糊样本时的作用.

## 4 结论

本文提出了基于线性全局注意力与专家几何知识的防振锤缺陷级联检测方法, 有效解决无人机巡检图像中防振锤缺陷检测的表征局限与算力瓶颈. 通过线性全局注意力机制与专家几何知识, 在实现轻量化边缘部署的同时, 有效解决了纯数据驱动模型在复杂巡检背景下的特征漂移与辨识精度瓶颈. 在自建数据集上的实验结果验证了该算法在精度、速度与模型体量间的平衡, 能够满足巡检要求. 未来工作将聚焦增量学习与自进化机制, 减少对人工标注数据的依赖, 构建具备在线学习能力的智能电力巡检体系.

### 参考文献 (References)

[1] 王少华, 蒋兴良, 孙才新. 输电线路导线舞动的国内外研究现状[J]. 高压技术, 2005(10): 14-17.

- (Wang S H, Jiang X L, Sun C X. Study status of conductor galloping on transmission line[J]. *High Voltage Engineering*, 2005(10): 14-17.)
- [2] 毛建旭, 贺振宇, 王耀南, 等. 电力巡检机器人路径规划技术及应用综述[J]. *控制与决策*, 2023, 38(11): 3009-3024.  
(Mao J X, He Z Y, Wang Y N, et al. Review of research and applications on path planning technology for power inspection robots[J]. *Control and Decision*, 2023, 38(11): 3009-3024.)
- [3] 苏怡萱, 李智, 盘书宝. 基于改进 YOLOv8 的输电线路绝缘子缺陷检测方法[J]. *电子测量与仪器学报*, 2025, 39(1): 14-23.  
(Su Y X, Li Z, Pan S B. Improved YOLOv8-based insulator defect detection method for transmission lines[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(1): 14-23.)
- [4] 李辉, 董燕, 刘祥, 等. 基于两阶段深度网络的输电线路异常目标检测方法[J]. *控制与决策*, 2022, 37(7): 1873-1882.  
(Li H, Dong Y, Liu X, et al. Transmission line abnormal object detection method based on deep network of two-stage[J]. *Control and Decision*, 2022, 37(7): 1873-1882.)
- [5] 刘传洋, 吴一全. 基于深度学习的输电线路视觉检测方法研究进展[J]. *中国电机工程学报*, 2023, 43(19): 7423-7445.  
(Liu C Y, Wu Y Q. Research progress of vision detection methods based on deep learning for transmission lines[J]. *Proceedings of the CSEE*, 2023, 43(19): 7423-7445.)
- [6] Shao N, Zou G F, Wei L Y, et al. Weak feature defect detection of insulators considering class boundaries and interclass dependencies[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 1-13.
- [7] 翟永杰, 杨旭, 赵振兵, 等. 融合共现推理的 Faster R-CNN 输电线路金具检测[J]. *智能系统学报*, 2021, 16(2): 237-246.  
(Zhai Y J, Yang X, Zhao Z B, et al. Integrating co-occurrence reasoning for Faster R-CNN transmission line fitting detection[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(2): 237-246.)
- [8] Huang X B, Zhang X L, Zhang Y, et al. A method of identifying rust status of dampers based on image processing[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(8): 5407-5417.
- [9] 肖志云, 王海强. 图像双分割与小波域多特征融合的高压输电线路典型小目标故障识别[J]. *电网技术*, 2021, 45(11): 4461-4469.  
(Xiao Z Y, Wang H Q. Typical small target fault identification of high-voltage transmission lines based on image double segmentation and fusion of multi-features in wavelet domain[J]. *Power System Technology*, 2021, 45(11): 4461-4469.)
- [10] 宋伟, 左丹, 邓邦飞, 等. 高压输电线防震锤锈蚀缺陷检测[J]. *仪器仪表学报*, 2016, 37(S1): 113-117.  
(Song W, Zuo D, Deng B F, et al. Corrosion defect detection of earthquake hammer for high voltage transmission line[J]. *Chinese Journal of Scientific Instrument*, 2016, 37(S1): 113-117.)
- [11] Shen H K, Chen P H, Chang L M. Automated steel bridge coating rust defect recognition method based on color and texture feature[J]. *Automation in Construction*, 2013, 31: 338-356.
- [12] 吴志成, 林秀贵, 许家浩, 等. 一种采用航拍图像识别高压输电线防震锤缺失的检测算法[J]. *应用科技*, 2021, 48(5): 35-41.  
(Wu Z C, Lin X G, Xu J H, et al. An algorithm for missing detection of anti-vibration hammer on high voltage transmission line based on aerial image[J]. *Applied Science and Technology*, 2021, 48(5): 35-41.)
- [13] Nguyen V N, Jenssen R, Roverso D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning[J]. *International Journal of Electrical Power & Energy Systems*, 2018, 99: 107-120.
- [14] 赵振兵, 熊静, 李冰, 等. 基于改进 Cascade R-CNN 的典型金具及其部分缺陷检测方法[J]. *高电压技术*, 2022, 48(03): 1060-1067.  
(Zhao Z B, Xiong J, Li B, et al. Typical fittings and its partial defect detection method based on improved Cascade R-CNN[J]. *High Voltage Engineering*, 2022, 48(03): 1060-1067.)
- [15] 宋立业, 刘帅, 王凯, 等. 基于改进 EfficientDet 的电网元件及缺陷识别方法[J]. *电工技术学报*, 2022, 37(9): 2241-2251.  
(Song L Y, Liu S, Wang K, et al. Identification method of power grid components and defects based on improved EfficientDet[J]. *Transactions of China Electrotechnical Society*, 2022, 37(9): 2241-2251.)
- [16] 赵振兵, 王帆帆, 刘良帅, 等. 基于注意力特征融合 YOLOv5 模型的无人机输电线路航拍图像金具检测方法[J]. *电测与仪表*, 2023, 60(3): 145-152.  
(Zhao Z B, Wang F F, Liu L S, et al. Transmission line image fitting detection method based on attention feature fusion YOLOv5 model[J]. *Electrical Measurement & Instrumentation*, 2023, 60(3): 145-152.)
- [17] 张吉祥, 方东. 基于 YOLOv8-SPH 的防震锤缺失检测[J]. *电子测量与仪器学报*, 2025, 39(2): 251-261.  
(Zhang J X, Fang D. Detection of missing shockproof hammers based on YOLOv8-SPH[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(2): 251-261.)
- [18] 余晓东. 电力金具设计[M]. 北京: 中国电力出版社, 2019.  
(Yu X D. Design of electric power fittings[M]. Beijing: China Electric Power Press, 2019.)
- [19] Everingham M, Van Gool L, Williams C K I, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [20] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context[C]. *European Conference on*

- Computer Vision. Zurich, 2014: 740-755.
- [21] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 770-778.
- [22] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 779-788.
- [23] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[J/OL]. 2017, arXiv: 1704.04861.
- [24] Tang Y, Han K, Guo J, et al. GhostNetv2: Enhance cheap operation with long-range attention[C]. Advances in Neural Information Processing Systems. New Orleans, 2022: 9969-9982.
- [25] Ma X, Dai X Y, Bai Y, et al. Rewrite the stars[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2024: 5694-5703.
- [26] Chen J R, Kao S H, He H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, 2023: 12021-12031.
- [27] He H Y, Zhang J N, Cai Y X, et al. MobileMamba: Lightweight multi-receptive visual mamba network[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, 2025: 4497-4507.
- [28] Liu Z, Lin Y T, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]. IEEE/CVF International Conference on Computer Vision. Montreal, 2021: 9992-10002.
- [29] Li Y Y, Hu J, Wen Y, et al. Rethinking vision transformers for MobileNet size and speed[C]. IEEE/CVF International Conference on Computer Vision. Paris, 2023: 16843-16854.
- [30] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J/OL]. 2014, arXiv: 1409.1556.
- [31] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J/OL]. 2020, arXiv: 2010.11929.

### 作者简介

黄志威 (1999-), 男, 硕士生, 主要研究方向为输电线路缺陷检测, E-mail: [ishuangzhiwei@163.com](mailto:ishuangzhiwei@163.com);

邹国锋 (1984-), 男, 副教授, 博士, 主要研究方向为电力故障诊断, E-mail: [zgf841122@163.com](mailto:zgf841122@163.com);

傅桂霞 (1985-), 女, 副教授, 博士, 主要研究方向为智能信息处理, E-mail: [fgx45101@163.com](mailto:fgx45101@163.com);

陈玉山 (2000-), 男, 硕士生, 主要研究方向为行人重识别, E-mail: [cys001223@163.com](mailto:cys001223@163.com);

王伟健 (2004-), 男, 硕士生, 主要研究方向为输电线路缺陷检测, E-mail: [19862511191@163.com](mailto:19862511191@163.com);

高明亮 (1985-), 男, 副教授, 博士, 主要研究方向为计算机视觉、教育大模型, E-mail: [sdut\\_mlgaao@163.com](mailto:sdut_mlgaao@163.com).