

# 基于无人机图像的建筑起重机械表面缺陷 视觉智能诊断方法<sup>\*</sup>

常潇丹<sup>1</sup> 冯 浩<sup>1</sup> 殷晨波<sup>2</sup> 陈鸣隽<sup>2</sup> 王 军<sup>3</sup>

(1. 南京信息工程大学人工智能学院 南京 210044; 2. 南京工业大学机械与动力工程学院 南京 211816;  
3. 江苏天宙检测科技有限公司 南京 210035)

**摘要:**建筑起重机械是现代工程的核心装备,其高空作业的高风险性易引发重大事故及经济损失,严重威胁安全。为了提升缺陷识别的效率和精度,降低操作人员登高巡查的风险,提出了一种基于无人机图像的表面缺陷智能检测方法 FRE。建筑起重机械表面缺陷种类多样、尺度微小且背景复杂,传统 YOLOv8 网络因多尺度特征融合能力不足及环境适应性局限,难以实现高精度缺陷检测。利用无人机巡检施工设备,建立了钢丝绳缺陷、金属结构锈蚀两个典型的起重机械缺陷图像数据集。将 YOLOv8 骨干网络中的 C2F 模块替换为 RepViT Block 模块,提升模型在图像理解和处理中的性能和效率,显著降低了计算复杂度和延迟,训练速度分别提高了 46.4%、2.6%;将 FasterNet Block 模块替换颈部网络的 C2F 模块,提高对缺陷的定位性能,提高了检测小目标的能力;将高效多尺度注意力(EMA)模块嵌入到骨干网络中,抑制背景信息的干扰,使模型更加关注缺陷特征。与现有的缺陷检测相比,该模型的检测精度分别达到了 88.0%、94.1%。同时,模型参数量相较于 YOLOv8 模型下降了 23.26%。结果表明,该方法可以快速、准确的检测出建筑起重机械表面缺陷,具有一定的社会应用价值。

**关键词:** 表面缺陷检测;无人机图像;改进 YOLOv8;建筑起重机械;EMA 模块

中图分类号: TH218; TN98 文献标识码: A 国家标准学科分类代码: 460.50

## Visual intelligent diagnosis method for surface defects of construction hoisting machinery based on UAV images

Chang Xiaodan<sup>1</sup> Feng Hao<sup>1</sup> Yin Chenbo<sup>2</sup> Chen Mingjun<sup>2</sup> Wang Jun<sup>3</sup>

(1. School of Artificial Intelligence, Nanjing University of Information Science & Technology, Nanjing 210044, China;  
2. School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211816, China;  
3. Jiangsu Tianzhou Testing Co., Ltd., Nanjing 210035, China)

**Abstract:** Construction cranes are the core equipment of modern engineering, and their high-risk operation at height is prone to cause major accidents and economic losses, seriously threatening safety. In order to improve the efficiency and accuracy of defect recognition and reduce the risk of operators climbing up to inspect, a surface defect intelligent detection method FRE based on UAV images is proposed. The surface defects of construction cranes are diverse, tiny in scale and complex in background, and the traditional YOLOv8 network is difficult to realize high-precision defect detection due to the lack of multi-scale feature fusion capability and the limitation of environmental adaptability. Utilizing the UAV inspection construction equipment, two typical lifting machine defect image datasets of wire rope defects and metal structure corrosion are established. The C2F module in the YOLOv8 backbone network is replaced with the RepViT Block module to improve the performance and efficiency of the model in image understanding and processing, which significantly reduces the computational complexity and latency, and the training speed is increased by 46.4% and 2.6%, respectively; the C2F module in the neck network is replaced by the FasterNet Block module, which improves the performance of the localization of defects and improves the ability of detecting small targets; the EMA module is embedded into the backbone network to suppress the interference of background information and make the model more focused on defect features. Compared with the existing defect detection, the detection

accuracy of the model reaches 88.0% and 94.1%, respectively. Meanwhile, the number of model parameters decreased by 23.26% compared with the YOLOv8 model. The results show that the method can quickly and accurately detect the surface defects of construction cranes, which has certain social application value.

**Keywords:** surface defect detection; drone images; improved YOLOv8; construction crane machinery; EMA module

## 0 引言

建筑起重机械是现代建筑施工的重要设备,涵盖了塔式起重机、施工升降机、吊篮、桥式起重机、移动式起重机等设备,具有专业技术含量高、危险性较大等特点<sup>[1]</sup>。由于过去重建轻养,随着时间的推移,建筑机械也面临着各种损伤和老化问题,很多建筑起重机械已进入病害高发期,常见锈蚀、开裂等缺陷<sup>[2]</sup>。若建筑起重机械“带病上岗”,极易出现倒塌、断臂等情形,甚至出现群死群伤的情况。

但是,人工巡检受限于建筑机械的高度和外形,容易遗漏死角和盲点,且高空作业安全风险高、工作效率低,传统的检测方法难以全面、准确地评估设备的状态。现有的一些视觉手段也只是“人工巡检为主,视频监测为辅”<sup>[3-4]</sup>。因此,亟需一种安全、智能、易操作的方式实现建筑机械缺陷的全面安全检测,克服人工巡检“费、慢、难、险”的弊端。

近年来,深度学习凭借自动化特征提取和高精度模式识别在缺陷检测领域发挥了重要作用。针对风力涡轮机、输电线路、钢材表面等工业场景的检测需求,国内外学者基于YOLO系列框架提出了一系列改进模型。张银胜等<sup>[5]</sup>提出了一种改进YOLOv5s的风力涡轮机表面缺陷检测模型以解决传统方式检测风力涡轮机表面缺陷时出现的精度不足、泛化性较差问题;Liu等<sup>[6]</sup>提出了一种基于深度学习目标检测网络的高压输电线路关键目标和缺陷的检测方法;Tie等<sup>[7]</sup>提出了一种基于YOLOv8n目标检测框架的轻型钢表面缺陷测试模型LSKAYOLOv8;赵佰亭等<sup>[8]</sup>设计的ECC-YOLO模型在钢表面缺陷检测中综合性能突出。然而深度学习模型在边缘设备部署时计算开销过大,难以完全满足实际工业场景的实时性与低成本需求。

与此同时,在复杂缺陷的三维识别与动态场景检测中,计算机视觉与摄影测量技术的结合成为重要突破口。Lobanov等<sup>[9]</sup>利用无人机高分辨率影像与摄影测量技术构建塔体三维模型,实现碎片穿透、裂缝等缺陷的立体化定位;刘金海等<sup>[10]</sup>采用主动小样本学习策略,减少管道焊缝缺陷检测的数据依赖;李可等<sup>[11]</sup>提出改进辅助分类生成对抗网络,增强了钢表面缺陷检测效果;马燕婷等<sup>[12]</sup>提出了一种基于YOLOv5网络改进的算法模型MT-YOLOv5,更好地均衡了检测速度与检测精度;Yuan等<sup>[13]</sup>为了实现输电线路典型缺陷的自动识别,提出了一种基

于YOLOv5算法的典型输电线路缺陷识别方法;范先友等<sup>[14]</sup>提出了一种基于改进YOLOv7的液晶面板电极缺陷视觉检测方法;Duan等<sup>[15]</sup>提出了一种基于无人机航空摄影场景的目标检测模型M-YOLOv8s;He等<sup>[16]</sup>提出了一种自适应多尺度检测方法,实现了风机叶片表面缺陷的准确分类和定位;Feng等<sup>[17]</sup>提出基于无人机遥感图像的路面损伤混合模型,通过图像拼接与深度学习完成大范围损伤检测。此类技术突破了传统二维检测的局限性,为缺陷的空间分析与动态监测提供了新思路。然而三维建模与图像拼接的计算复杂度高,实时性较差,且对复杂背景的鲁棒性仍需提升,难以满足高速工业检测需求。面对无人机巡检、生产线高速检测等动态场景,实时性与轻量化成为关键研究方向。杨宇龙等<sup>[18]</sup>设计风力发电机表面缺陷检测模型,优化推理流程以满足实时性要求;Yang等<sup>[19]</sup>改进的YOLOv6s-GRE均通过模型压缩与加速技术降低计算负载。这些研究显著提升了模型在边缘设备上的部署能力,但如何在轻量化过程中避免精度损失仍是难点。

针对上述挑战,本文提出了一种基于改进YOLOv8n的表面缺陷视觉检测算法。根据实际工业场景搭建无人机图像采集装置,创建了钢丝绳缺陷、金属结构锈蚀数据集。首先,改进骨干网络(backbone),将RepViT Block模块引入Backbone以轻量化模型,提高训练速度;其次,基于FasterNet Block模块重新设计颈部网络(neck)的C2F模块,提高对缺陷的定位性能,提高了检测小目标的能力;最后,将高效多尺度注意力模块(efficient multi-scale attention, EMA)嵌入到Backbone中,抑制背景信息的干扰,使模型更加关注缺陷特征。

## 1 建筑起重机械表面缺陷视觉诊断方法

### 1.1 YOLOv8 网络模型

YOLOv8n是YOLOv8的轻量级变体,专为资源受限的嵌入式设备设计。在整体设计上,YOLOv8n模型包括输入端、Backbone、Neck和头部(head)4部分。对于输入端,会通过数据预处理和增强的操作被转换为模型可以处理的格式。Backbone部分是YOLOv8的主干网络,主要负责提取图像的特征。Neck是连接Backbone和Head的部分,主要作用是特征融合和处理,以提高检测的准确性和效率。Head部分是模型最后一层,其结构会根据不同的任务而有所不同。

表 1 为各检测模型在典型工业缺陷检测数据集上的对比结果,由表 1 可见,YOLOv8 在缺陷检测中展现出显著优势,其通过优化的网络架构与动态标签分配策略,在

保持轻量化设计的同时,实现了更高的检测精度与实时性,相较于 YOLOv5、Faster R-CNN 等模型,更适用于复杂工业场景下的实时检测需求。

表 1 缺陷检测方法性能对比

Table 1 Performance comparison of defect detection methods

模型	测试数据集	mAP@ 0.5/%	检测速度/fps	Params/( $\times 10^6$ )
YOLOv8n	MS COCO	78.5	120	3.2
YOLOv5n	MS COCO	72.3	145	1.9
YOLOv7-tiny	MS COCO	75.6	110	6.0
YOLOv9	MS COCO	82.1	80	8.7
YOLOv10n	MS COCO	79.8	115	3.5
YOLOv11s	MS COCO	80.2	100	4.8
Mask R-CNN	VOC2007	80.5	15	44.3
Faster R-CNN	VOC2007	77.2	20	38.6

虽然原始 YOLOv8n 模型在性能方面具有较高的准确性和速度,但也存在一些缺点。首先,YOLOv8 的参数量相较于其前身 YOLOv4 有所增加,这可能会导致更大的存储空间需求和计算负担。其次,YOLOv8n 在处理工业缺陷检测等需要高精度和复杂场景的任务时,可能会受到其轻量化设计和资源限制的影响,导致一定的检测精度和处理能力上的限制。因此,选择 YOLOv8n 作为基准模型,并进行了改进。

## 1.2 FRE(FasterNet-RepViT-EMA)网络模型

提出了一种基于 YOLOv8n 的轻量级建筑起重机械表面缺陷检测模型 FRE。其核心改进包括 Backbone 的 RepViT Block 模块替换、Neck 的 FasterNet Block 模块替

换,以及 EMA 多尺度注意力机制的嵌入。图 1 所示为 FRE 模型结构示意图。由图 1 可见,RepViT Block 模块替换骨干网络的 C2F 模块,融合 Transformer 的全局感知与卷积神经网络(convolutional neural networks, CNN)的局部特征提取能力,结合重参数化技术提升训练效率;FasterNet Block 模块嵌入颈部网络,利用部分卷积(PConv)降低计算冗余,显著增强小目标定位精度;EMA 通过跨空间特征聚合抑制复杂背景干扰。三者协同作用在轻量化基础上实现高精度,同时保持实时性,有效解决了建筑起重机械缺陷检测中多尺度特征融合不足、动态背景噪声敏感及边缘设备资源受限的难题。

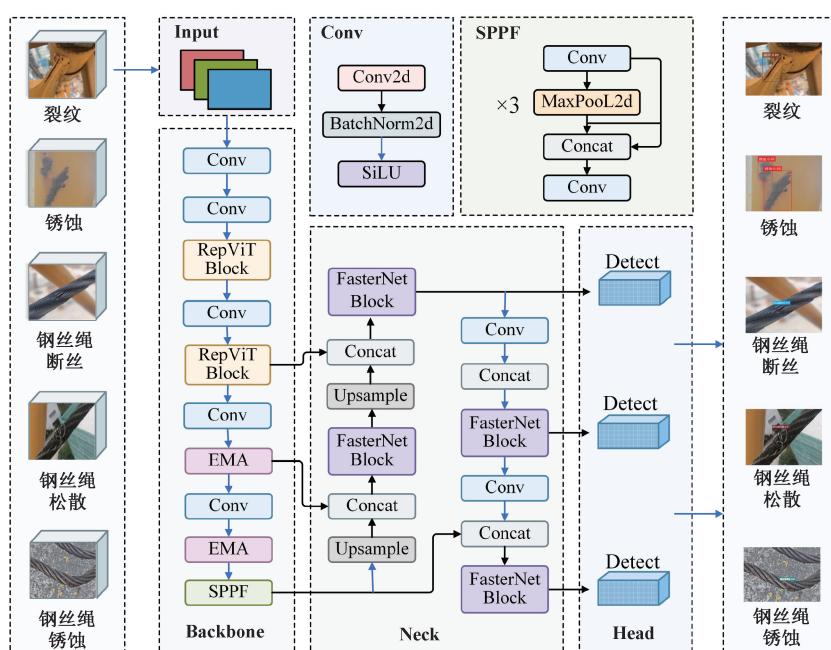


图 1 FRE 模型架构

Fig. 1 FRE model architecture

### 1) RepViT Block 模块

RepViT Block 模块<sup>[20]</sup>被用来替换主干中的 C2F 模块以增强模型特征提取能力。RepViT Block 是 RepViT 模型中的创新模块,而 RepViT 模块是受 Transformer 的启发的视觉处理模型,主要包括模块(stem)、下采样

层(downsampling)、全局平均池化层(pooling)、全连接层(fully connected, FC)以及多个阶段的 RepViT Block 组成,旨在提升模型在图像理解和处理中的性能和效率。RepViT 结构如图 2 所示。

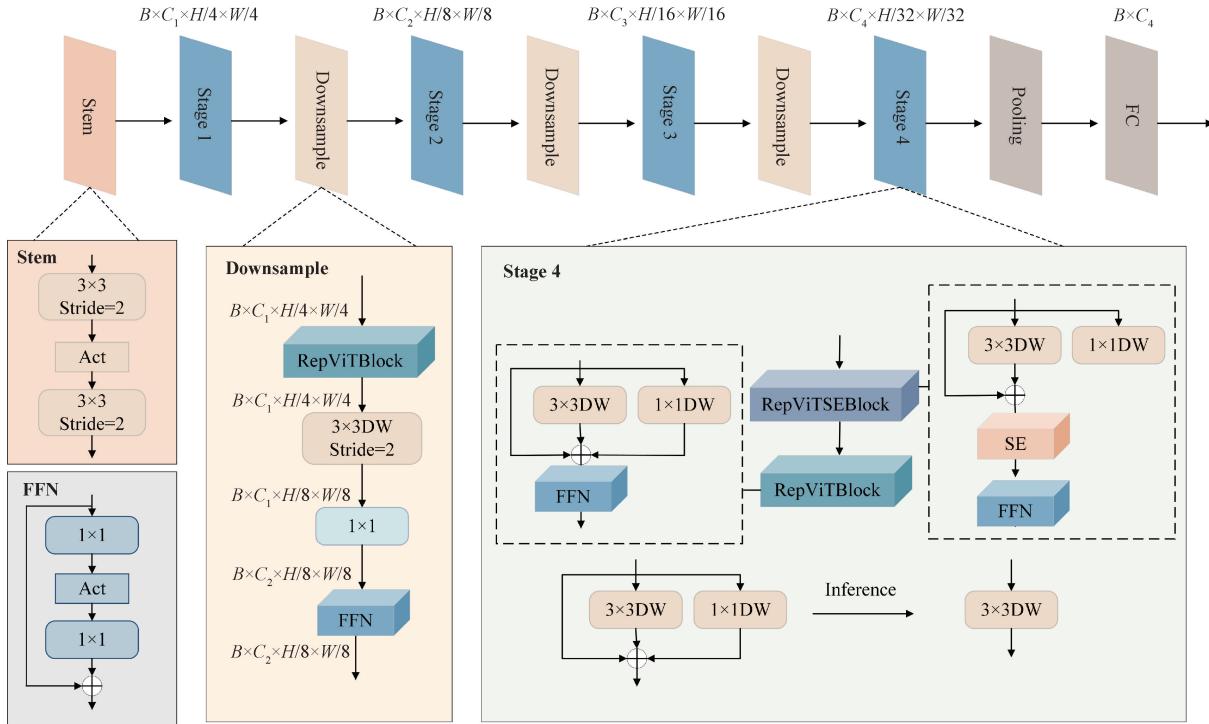


图 2 RepViT 结构

Fig. 2 RepViT structure diagram

首先,RepViT 模型在初期引入两个步长等于 2 的  $3 \times 3$  卷积层,用于对输入图像进行处理。这种设计简化了初始的计算负担,同时提升了模型的优化稳定性和表现力。其次,RepViT 在每个阶段后的下采样层中加入了深度卷积和点卷积相结合的方式,这不仅实现了空间分辨率的降低,还进一步加强了通道维度的信息表达能力。此外,RepViT 还在这些下采样层中增加了 RepViT Block,使网络在保持高效下采样的同时拥有更好的特征提取能力。其中,RepViT 有 4 个阶段,每个阶段由多个 RepViTBlock 组成,以及一个可选择的 RepViTSEBlock,包含深度可分离卷积( $3 \times 3$ DW), $1 \times 1$  卷积,压缩激励模块(squeeze-and-excitation, SE)和前馈网络(feed-forward network, FFN)。最后通过全局平均池化层和全连接层,用于最终的类别预测。

RepViT 模块中的核心单元是 RepViTBlock,RepViTBlock 结构如图 3 所示。RepViT Block 模块使用  $3 \times 3$  深度卷积进行空间信息融合,然后使用  $1 \times 1$  卷积进行通道间交互,并结合可选择的 SE 层进一步优化。进一步采用了一种广泛应用于 DW 层的结构再参数化技术,

以增强训练过程中的模型学习。

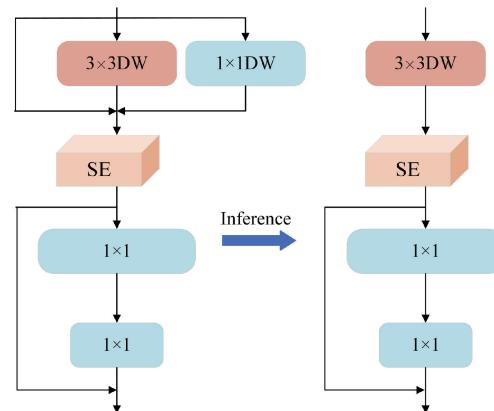


图 3 RepViT Block 结构

Fig. 3 RepViT Block structure diagram

受 Transformer 启发,RepViT 作为一种新的轻量级 CNN 架构,在保持移动设备友好性的同时提高了性能和效率。RepViT Block 模块的引入显著提高了模型在移动设备上的推理速度和性能表现,使得 FRE 能够在保持较

高精度的同时,显著降低计算复杂度和延迟,从而适合资源受限的边缘设备部署。

### 2) FasterNet Block 模块

为进一步减少模型的参数量和计算量,将颈部中的 C2F 模块替换为 FasterNet Block 模块<sup>[21]</sup>。FasterNet Block 是 FasterNet 网络的基本构建单元,负责处理特征的提取和转换。FasterNet Block 集成了 PConv 和点线卷积(PWConv),以减少冗余计算。PConv 的浮点操作运算(FLOPs)记为:

$$h \times w \times k^2 \times C_p^2 \quad (1)$$

式中: $h$ 是输出特征映射的高度; $w$ 是输出特征映射的宽

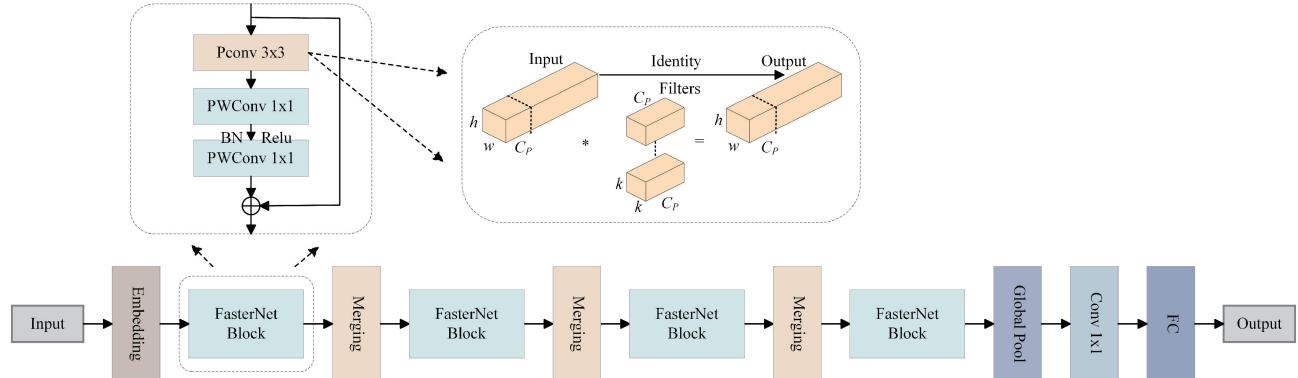


图 4 FasterNet Block 网络结构

Fig. 4 FasterNet Block network structure

PConv 采用标准卷积法对部分输入信道进行空间特征提取,同时保持剩余信道不变。通过将第一个或最后一个连续的  $C_p$  通道作为整个特征图的表示,保证了其通用性,同时保持了输入和输出特征图通道数量的一致性。

FasterNet Block 通过使用一种新颖的部分卷积,减少了冗余计算和内存访问。

### 3) EMA 机制

在复杂场景下,由于不同尺度的缺陷特征难以完全捕捉,受到注意力机制的启发,引入了 EMA<sup>[22]</sup>方法,通过采用跨尺度的特征建模以及精确空间信息融合,能够有效解决复杂场景下的微小缺陷的检测问题。EMA 方法的流程如图 5 所示。

高效多尺度注意力机制 EMA 通过将输入特征  $X$  按照跨信道维度方向划分为子特征  $X'$ ,以学习不同的语义特征表达。其中, $X$  和  $X'$  分别为:

$$X \in R^{C \times H \times W} \quad (3)$$

$$X' = [X_0, X_1, \dots, X_{G-1}] \in R^{C/G \times H \times W} \quad (4)$$

令  $G \ll C$  并且假设利用学到的注意权重描述符来增强每个子特征中感兴趣区域的特征表示。为了收集多尺度的空间信息,EMA 利用 3 条平行路径来提取分组特征图的注意权重描述。其中两条平行路由是  $1 \times 1$  分支,第 3 条路由是  $3 \times 3$  分支。更具体的来说,在  $1 \times 1$  分支中,

度; $k$  是卷积核的大小; $C_p$  是每个卷积核属于输出的信道数。由于部分卷积模式的部分比  $r$  为:

$$r = \frac{C_p}{c} = \frac{1}{4} \quad (2)$$

所以 PConv 的浮点操作运算只有标准卷积的  $1/16$ ,构建了更强大和高效的空间特征提取的轻量级网络。

每个 FasterNet Block 都有一个 PConv 层和两个 PWConv(或 Conv  $1 \times 1$ )层。它们共同构成反向残差块,中间层有一个扩展的通道数量,并放置一个快捷连接以重用输入特性。PConv 的机理如图 4 所示。

有两个 1D 全局平均池化操作分别沿两个空间方向对通道进行编码,在  $3 \times 3$  分支中只有一个  $3 \times 3$  卷积核堆叠,以捕获多尺度特征表示。

接着,两个由  $1 \times 1$  分支编码的特征按照图像高度  $H$  的方向进行拼接,使它共享相同的  $1 \times 1$  卷积,而不在  $1 \times 1$  分支中降维。将  $1 \times 1$  卷积的输出分解为两个向量后,采用两个非线性 Sigmoid 函数拟合线性卷积上的二维二项分布。为了在  $1 \times 1$  分支中实现两条平行路由之间不同的跨通道交互特征,通过一个简单的乘法来聚合每一组内的两个通道级注意映射。另一方面, $3 \times 3$  分支通过  $3 \times 3$  卷积捕获局部跨通道交互,以扩大特征空间。这样,EMA 不仅对信道间信息进行编码,以调整不同信道的重要性,而且还将精确的空间结构信息保存到信道中。

为了在不同空间维度方向上聚合跨空间信息,以实现更丰富的特征聚合,引入了两个张量,其中一个是  $1 \times 1$  分支的输出,另一个是  $3 \times 3$  分支的输出。然后,利用二维全局平均池化对  $1 \times 1$  分支输出中的全局空间信息进行编码,在通道特征的联合激活机制之前,将最小分支的输出直接转换为相应的维度形状( $R^{1 \times C//G} \times R^{C//G \times HW}$ )。2D 全局池化操作可以表示为:

$$z_e = \frac{1}{HW} \sum_j^H \sum_i^W x_e(i, j) \quad (5)$$

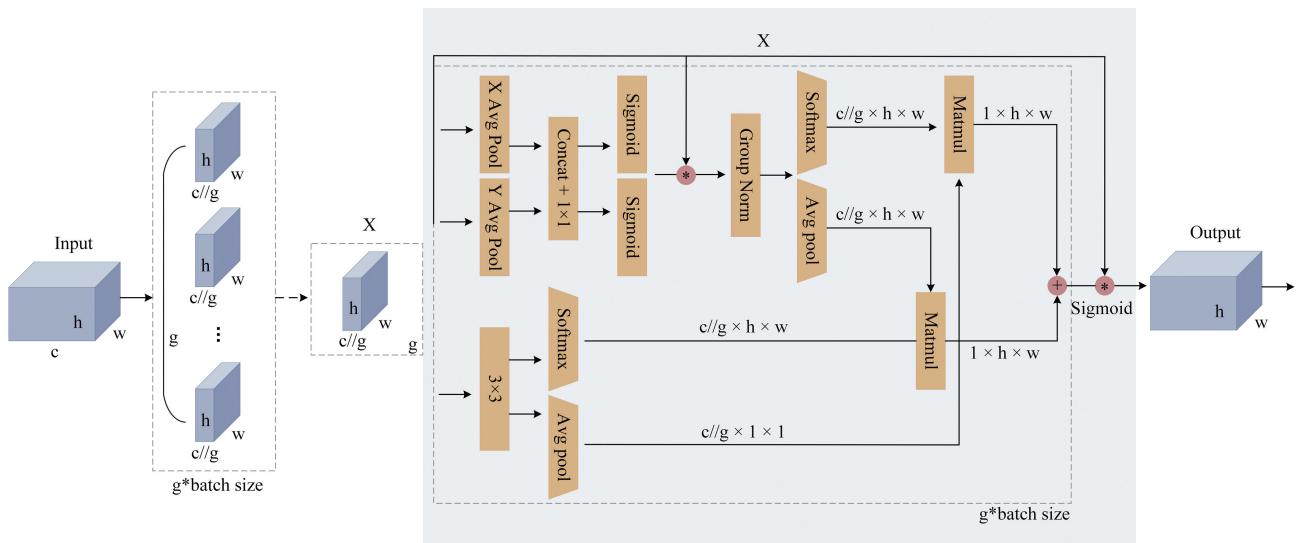


图 5 EMA 原理图

Fig. 5 EMA schematic

为了提高计算效率,2D 全局平均池化使用二维高斯映射的自然非线性函数 Softmax 来拟合其上的线性变化。上述并行操作处理的输出通过点乘得到了第 1 个空间注意力图,描述了收集的不同尺度的空间信息。与此同时,利用 2D 全局平均池化编码的  $3 \times 3$  分支的全局空间特征与维度变换一致的  $1 \times 1$  分支特征进行联合激活以获得第 2 个保留整个精确的空间位置信息的空间注意力图。最后,将每一组内的输出特征图计算为所生成的两个空间注意权重值的聚合并经过 Sigmoid 函数计算。它捕获像素级的成对关系,并突出显示所有像素的全局上下文。

综上所述,EMA 采用跨空间信息聚合方法,对远程依赖关系进行建模,并将精确的位置信息嵌入之中。融合了不同上下文信息,使得改进后的模型能够获得更高级别的像素级关注,以提升复杂场景下微小缺陷的检测能力。

## 2 无人巡检系统

### 2.1 无人巡检系统

基于 AI 的建筑起重机械无人巡检系统主要由行业级多旋翼无人机、图像自动采集模块、缺陷智能识别模块、可视化安全智能评估模块等关键部分组成。采用无人机搭载专用高清相机对建筑机械进行全覆盖高清图像采集,利用 FRC 智能算法自动完成对建筑起重机械结构表面锈蚀及钢丝绳锈蚀、断丝、松散等缺陷进行智能识别、定性判断。

### 2.2 缺陷数据集

建筑起重机械的缺陷图像通过大疆 Mavic 3E 无人机获得,无人机采集塔式起重机表面缺陷图像的场景如

图 6 所示。



图 6 无人机正在巡检塔式起重机

Fig. 6 Drones inspecting tower cranes

塔式起重机的常见故障缺陷中,金属结构的锈蚀和钢丝绳缺陷是两类关键问题,因此将其作为重点研究对象。金属结构的锈蚀甚至开裂直接影响整体稳定性,极易引发倒塌,重要性居于首位;钢丝绳缺陷(如断丝、锈蚀、变形)则涉及承载安全,可能导致重物坠落,紧随其后。两者均为塔式起重机安全运行的关键检测重点。建筑起重机械损坏的类型和数量如表 2 所示。为了保证图像数据集的多样性,采集了多种类型的缺陷图像。所采集的图像的示例如图 7 所示。

表 2 塔式起重机典型缺陷数据集的参数

Table 2 Parameters of typical defect

dataset for tower cranes

类型	数量/张	大小	迭代次数
钢丝绳缺陷	1 463	640×640×3	200
金属结构锈蚀	806	1 280×1 280×3	300



图 7 多种类型的缺陷数据集

Fig. 7 Various types of defect datasets

### 3 实验与结果分析

#### 3.1 实验环境

该模型的训练过程在 Windows 10 操作系统和 PyTorch 2.1 框架下运行。软件环境配置为 CUDA 11.1 和 Python 3.8。采用双显卡进行模型训练, 其型号均为 NVIDIA GeForce RTX 3090 24 G。在模型训练中采用随机梯度下降算法对网络模型的权重进行更新和优化。具体参数设置如下: 批量大小 (Batch Size) 为 16, 初始学习率 (Initial Learning Rate) 为 0.01, 动量参数 (Momentum) 为 0.937, 权重衰减因子 (Weight\_decay) 为 0.000 5。

#### 3.2 评价指标

为了验证所提出的 FRE 模型的性能, 使用平均精度均值 (mAP)、精度 (precision)、召回率 (recall)、参数量 (parameters)、浮点计算量以及平均模型检测处理时间 Speed-GPU 作为测量指标。Precision 计算为正确预测的阳性样本数与预测为阳性样本的样本数的比例, 定义如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

Recall 计算为所有被预测的正确目标的比例, 定义如下:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

式中:  $TP$  表示正确分类的正例子的数量;  $FP$  表示错误分类的负例子的数量;  $FN$  表示错误分类的正例子的数量;  $TN$  表示正确分类为负例子的负例子的数量。

平均精度 (AP) 计算公式如下:

$$AP = \int_0^1 P(R) dR \quad (8)$$

mAP 计算公式如下:

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (9)$$

$mAP@0.5$  为 IoU 设置为 0.5 时所有类别的 mAP。 $mAP@0.5 : 0.95$  为不同 IoU 阈值下的 mAP, 其中 IoU 的变化范围为 0.5~0.95, 步长为 0.05。

#### 3.3 典型缺陷的实验分析

##### 1) 钢丝绳缺陷

通过消融实验验证网络结构变化引起的性能变化。开展了 YOLOv8n、YOLOv8n\_RVB、YOLOv8n\_Faster、YOLOv8n\_Faster\_RVB 和 FRE 共 5 个实验。图 8 所示为不同模型在数据集上的精度曲线, 结果如表 3 所示。

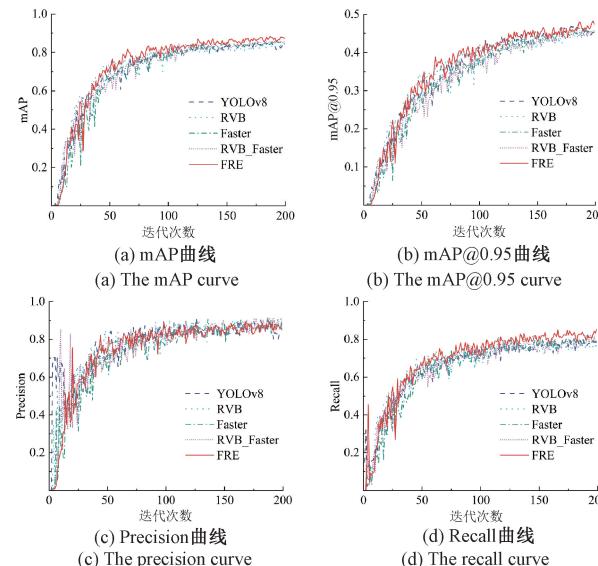


图 8 模型在钢丝绳缺陷数据集上的实验结果对比

Fig. 8 Comparison of experimental results of the model on the wire rope defect dataset

用 RepViT Block 模块替换 YOLOv8n 骨干网络中的 C2F 后, 改进后的模型的平均检测处理时间减少了 46.43%, 与 YOLOv8n 相比, 计算量和模型参数数量分别减少了 22.22% 和 24.25%, 同时模型精度提升了 0.5%。

用 FasterNet Block 模块替换 YOLOv8n 颈部网络中的 C2F 后, 改进后的模型的平均检测处理时间同样减少显著, 减少了 50%, 且模型参数数量减少了 23.59%, 同时模型精度提升了 1.0%。将 YOLOv8n 中所有 C2F 用 RepViT Block 模块与 FasterNet Block 模块分别替换后, 模型精度提升了 1.9%, 平均检测处理时间降低了 64.29%, 参数量减少了 23.92%。此外, EMA 模块添加在 YOLOv8n 骨干网络后, 计算量减少了 20.99%, 平均模型检测处理时间减少了 50%, 同时模型精度提高了 3.7%。

表 4 为该方法与基准模型在钢丝绳缺陷数据集测试集上的单类比较结果。在 Break 和 Loose 类别中, 该方法的单类 AP 值均高于基准模型。综上所述, 改进的 FRE 模型具有更好的识别效果。

表3 不同模型在钢丝绳缺陷数据集上的结果

Table 3 Results of different models on the wire rope defect dataset

模型	mAP@0.5/%	mAP@0.5;0.95/%	Precision/%	Recall/%	Params/( $\times 10^6$ )	计算量/GFLOPs	Speed-GPU/ms
YOLOv8n	84.3	46.8	79.0	90.0	3.01	8.1	2.8
YOLOv8n_RVB	84.8	44.9	81.1	91.0	2.28	6.3	1.5
YOLOv8n_Faster	85.3	45.9	75.7	92.0	2.30	6.3	1.4
YOLOv8n_Faster_RVB	86.2	46.7	78.1	91.0	2.29	6.3	1.0
FRE	88.0	49.0	81.9	94.0	2.31	6.4	1.4

表4 测试集单类型AP结果

Table 4 AP results of a single type on the test set(%)

类别	YOLOv8n	FRE
断丝(break)	73.8	76.9
锈蚀(rust)	96.9	93.9
松散(loose)	82.3	93.3

图9为所提出的方法与YOLOv8n在不同场景中的比较。图9(a1)~(c1)为YOLOv8n的检测结果,图9(a2)~(c2)为本文方法的检测结果。

对于密集分布的模糊目标缺陷,由图9(a)可见,FRE可以准确检测密集分布的模糊缺陷,而YOLOv8n没有完全检测出所有明显的断丝缺陷。在图9(b)中,当同时出现两种缺陷时,该方法准确地识别到了两种缺陷,而YOLOv8n只识别到了锈蚀缺陷。在图9(c)中,该方法相较于YOLOv8n识别到了更多严重遮挡的松动缺陷。结果表明,该方法优于YOLOv8n。

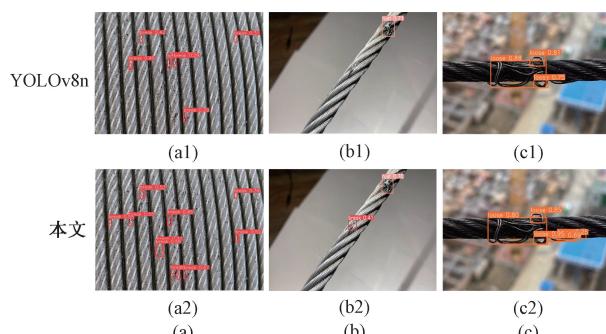


图9 模型在钢丝绳缺陷数据集上的检测结果对比

Fig. 9 Comparison of detection results of the model on the wire rope defect dataset

为了进一步验证该方法的性能,将其与Faster R-CNN、YOLOv5s、YOLOv7-tiny-SiLU等流行检测网络进行了比较,并保持参数设置一致,结果如图10和表5所示。图10为缺陷数据集在不同算法下的检测结果比较。

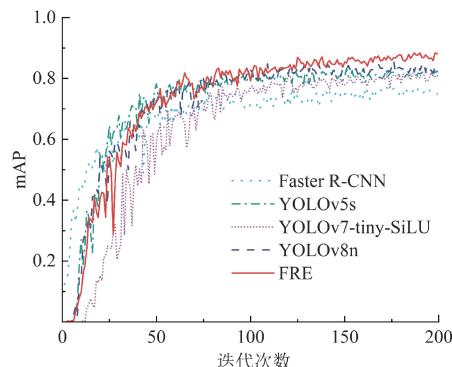


图10 钢丝绳数据集在不同算法下的检测结果比较

Fig. 10 Comparison of detection results of wire rope dataset under different algorithms

从表5可以看出,与Faster R-CNN、YOLOv5s、YOLOv7-tiny-SiLU相比,该方法在保证运行速度的前提下,大大降低了计算量和参数数量,该方法的精度分别提高了10.8%、5.6%和7.0%。尤其是与YOLOv5s比较,计算量降低了73.11%,参数量降低了74.64%。结果表明,该方法比其他主流轻量级网络具有优势。

## 2) 金属结构锈蚀

在标准节缺陷数据集上的实验结果如图11和表6所示。从表6可以看出,用RepViTBlock模块替换YOLOv8n骨干网络中的C2F后,计算量和模型参数量的变化趋势与钢丝绳缺陷数据集变化一致,同时模型精度提升了2.0%,平均检测处理时间小幅减少了2.6%。

表5 不同轻量级网络性能的比较

Table 5 Comparison of performance of different lightweight networks

模型	mAP@0.5/%	mAP@0.5;0.95/%	Precision/%	Recall/%	Params/( $\times 10^6$ )	计算量/GFLOPs	Speed-GPU/ms
Faster R-CNN	77.2	34.2	76.5	83.2	28.28	—	—
YOLOv5s	82.4	41.1	88.3	87.0	9.11	23.8	2.9
YOLOv7-tiny-SiLU	81.0	40.5	80.2	92.0	6.20	13.8	2.8
YOLOv8n	84.3	46.8	79.0	90.0	3.01	8.1	2.8
FRE	88.0	49.0	81.9	94.0	2.31	6.4	1.4

表 6 不同模型在金属结构锈蚀数据集上的结果

Table 6 Results of different models on the metal structure corrosion dataset

模型	mAP@0.5/%	mAP@0.5;0.95/%	Precision/%	Recall/%	Params/( $\times 10^6$ )	计算量/GFLOPs	Speed-GPU/ms
YOLOv8n	86.2	71.3	92.3	85.0	3.01	8.1	3.9
YOLOv8n_RVB	88.0	75.6	87.0	88.0	2.28	6.3	3.8
YOLOv8n_Faster	87.0	73.0	88.0	88.0	2.30	6.3	3.6
YOLOv8n_Faster_RVB	90.6	79.3	57.7	88.0	2.29	6.3	3.3
FRE	94.1	85.1	70.0	94.0	2.30	6.4	3.6

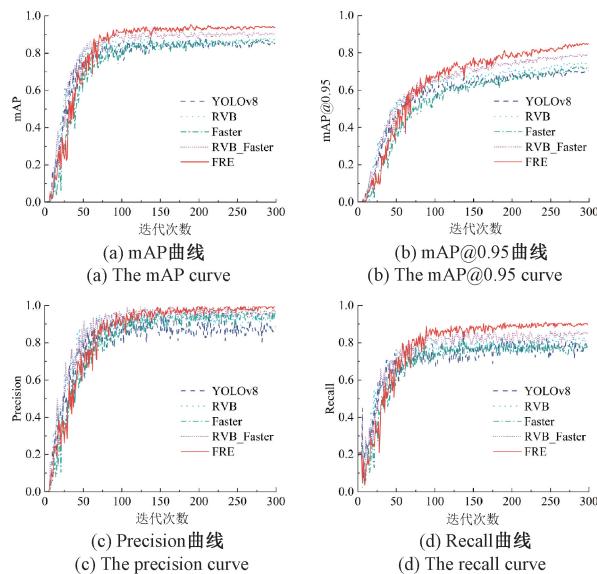


图 11 模型在金属结构锈蚀数据集上的实验结果对比

Fig. 11 Comparison of experimental results of the model on the metal structure rust dataset

用 FasterNetBlock 模块替换 YOLOv8n 颈部网络中的 C2F 后, 模型精度提升了 0.9%, 平均检测处理时间减少了 8.3%。将 YOLOv8n 中所有 C2F 用 RepViTBlock 模块与 FasterNetBlock 模块分别替换后, 模型精度相比于 RepViTBlock 和 FasterNetBlock 分别提升了 3.0% 和 4.1%, 平均检测处理时间分别降低了 15.15% 和 9.09%。

此外, 加入 EMA 模块后, 与 YOLOv8n 相比, 模型精度提高了 9.2%, 平均检测处理时间减少了 8.3%。

因此, 验证了该方法的通用性, 通过重新训练该模型的结构, 该模型可以适应不同的数据集或场景。

图 12 所示为该方法与 YOLOv8n 在标准节缺陷数据

集上的比较。图 12(a1)~(c1) 是 YOLOv8n 的检测结果, 图 12(a2)~(c2) 是 FRE 的检测结果。对于密集分布的模糊目标缺陷, 由图 12(a) 可见, FRE 可以分别检测出密集分布的缺陷, 而 YOLOv8n 则是识别为一个缺陷目标。在图 12(b) 中, 该方法相较于 YOLOv8n 识别到了更多的锈蚀缺陷。在图 12(c) 中, 当出现与目标缺陷有部分相同特征的目标时, 该方法可以更加准确地区分目标缺陷与相似目标。结果表明, 该方法优于 YOLOv8n。

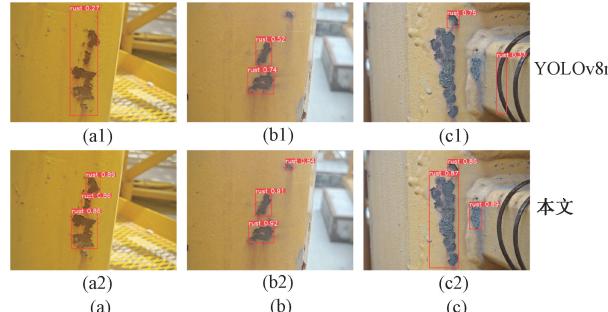


图 12 模型在金属结构锈蚀数据集上的检测结果对比

Fig. 12 Comparison of detection results of the model on the metal structure rust dataset

为了验证该方法的优越性, 同样与 Faster R-CNN、YOLOv5s、YOLOv7-tiny-SiLU 和 YOLOv8n 进行对比。图 13 所示为锈蚀缺陷数据集在不同检测算法下的 mAP 曲线对比。比较结果如表 7 所示。由表 7 可知, 与其他检测算法相比, 该方法在保证运行速度的前提下, 大大降低了计算量和参数数量, 该方法的精度分别提高了 10.3%、9.3%、11.0% 和 9.2%。且相较于 YOLOv5s 与 YOLOv7-tiny-SiLU, 参数量分别减少了 74.8% 和 62.9%。因此, 基于该模型的总体检测性能指标, 提出的算法在识别精度和轻量级两方面都具有很大优势。

表 7 不同轻量级网络性能的比较

Table 7 Comparison of performance of different lightweight networks

模型	mAP@0.5/%	mAP@0.5;0.95/%	Precision/%	Recall/%	Params/( $\times 10^6$ )	计算量/GFLOPs	Speed-GPU/ms
Faster R-CNN	85.3	56.6	82.2	92.0	28.28	—	—
YOLOv5s	86.1	70.7	85.0	86.0	9.11	23.8	3.6
YOLOv7-tiny-SiLU	84.8	68.4	79.6	93.0	6.20	13.8	3.8
YOLOv8n	86.2	71.3	88.3	92.0	3.01	8.1	3.9
FRE	94.1	85.1	70.0	94.0	2.30	6.4	3.6

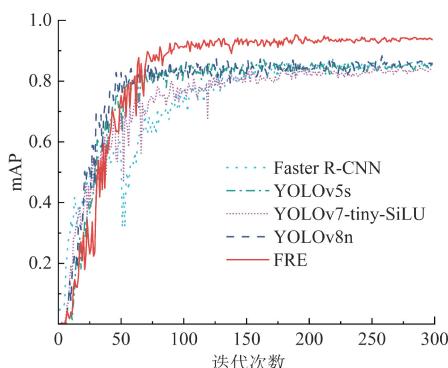


图 13 金属结构锈蚀数据集在不同算法下的检测结果比较

Fig. 13 Comparison of detection results of metal structure rust dataset under different algorithms

## 4 结 论

针对建筑起重机械领域巡检存在的问题,本文提出了一种基于无人机图像的表面缺陷智能检测方法 FRE。该方法以 YOLOv8 为基础模型,在 Backbone 中引入 RepViT Block 模块以轻量化模型,从而提高了模型的训练速度;基于 FasterNet Block 模块改进 Neck 部分,提高对缺陷的定位性能,提高了检测小目标的能力;将 EMA 注意力模块引入 Backbone 中,增强多尺度特征融合能力,抑制背景信息的干扰,使模型更加关注缺陷特征。在自定义的钢丝绳缺陷、金属结构锈蚀两个典型缺陷的起重机械缺陷图像数据集上分别进行测试,实验结果表明,与现有的目标检测方法相比,提出的算法模型在自制的两种数据集上的平均精度分别达 88.0% 和 94.1%,模型大小为  $2.3 \times 10^6$ ,可以快速、准确的检测出建筑起重机械表面缺陷,避免了传统检测方式下人员登高存在的安全隐患问题。在后期的工作中,将进一步优化模型结构,丰富现有的缺陷数据集,引入更多的缺陷类型,验证该方法的适用性和可靠性。

## 参考文献

- [ 1 ] ZHOU Y, GUO H, MA L, et al. Image-based onsite object recognition for automatic crane lifting tasks [ J ]. Automation in Construction, 2021, 123: 103527.
- [ 2 ] ZHANG D L, LI M C, TIAN D, et al. Intelligent text recognition based on multi-feature channels network for construction quality control [ J ]. Advanced Engineering Informatics, 2022, 53: 101669.
- [ 3 ] YU L, HUANG M M, JIANG S, et al. Unmanned aircraft path planning for construction safety inspections [ J ]. Automation in Construction, 2023, 154: 105005.
- [ 4 ] SUN Z, ZHU Z, XIONG R, et al. Dynamic human systems risk prognosis and control of lifting operations during prefabricated building construction [ J ]. Developments in the Built Environment, 2023, 14: 100143.
- [ 5 ] 张银胜, 杨宇龙, 吉茹, 等. 改进 YOLOv5s 的风力涡轮机表面缺陷检测 [ J ]. 电子测量与仪器学报, 2023, 37(1): 40-49.
- ZHANG Y SH, YANG Y L, JI R, et al. Surface defect detection of wind turbine based on YOLOv5s [ J ]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 40-49.
- [ 6 ] LIU Z, WU G, HE W, et al. Key target and defect detection of high-voltage power transmission lines with deep learning [ J ]. International Journal of Electrical Power & Energy Systems, 2022, 142: 108277.
- [ 7 ] TIE J, ZHU C, ZHENG L, et al. LSKA-YOLOv8: A lightweight steel surface defect detection algorithm based on YOLOv8 improvement [ J ]. Alexandria Engineering Journal, 2024, 109: 201-212.
- [ 8 ] 赵佰亭, 张晨, 贾晓芬, 等. ECC-YOLO:一种改进的钢材表面缺陷检测方法 [ J ]. 电子测量与仪器学报, 2024, 38(4): 108-116.
- ZHAO B T, ZHANG CH, JIA X F, et al. ECC-YOLO: An improved method for detecting surface defects in steel [ J ]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [ 9 ] LOBANOV L, STELMAKH D, SAVITSKY V, et al. Damage detection and analysis using unmanned aerial vehicles ( UAVs ) and photogrammetry method [ J ]. Procedia Structural Integrity, 2024, 59: 43-49.
- [ 10 ] 刘金海, 赵真, 付明芮, 等. 基于主动小样本学习的管道焊缝缺陷检测方法 [ J ]. 仪器仪表学报, 2022, 43(11): 252-261.
- LIU J H, ZHAO ZH, FU M R, et al. Active small sample learning based the pipe weld defect detection method [ J ]. Chinese Journal of Scientific Instrument, 2022, 43(11): 252-261.
- [ 11 ] 李可, 祁阳, 宿磊, 等. 基于改进 ACGAN 的钢表面缺陷视觉检测方法 [ J ]. 机械工程学报, 2022, 58(24): 32-40.
- LI K, QI Y, SU L, et al. Visual inspection of steel surface defects based on improved auxiliary classification generation adversarial network [ J ]. Journal of Mechanical Engineering, 2022, 58(24): 32-40.
- [ 12 ] 马燕婷, 赵红东, 阎超, 等. 改进 YOLOv5 网络的带钢表面缺陷检测方法 [ J ]. 电子测量与仪器学报, 2022, 36(8): 150-157.
- MA Y T, ZHAO H D, YAN CH, et al. Strip steel surface defect detection method by improved YOLOv5

- network [ J ]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8) : 150-157.
- [13] YUAN J, ZHENG X, PENG L, et al. Identification method of typical defects in transmission lines based on YOLOv5 object detection algorithm [ J ]. Energy Reports, 2023, 9: 323-332.
- [14] 范先友,过峰,俞建峰,等.基于改进YOLOv7的液晶面板电极缺陷视觉检测技术研究 [ J ]. 电子测量与仪器学报, 2023, 37(9) : 225-233.
- FAN X Y, GUO F, YU J F, et al. Research on visual detection technology for liquid crystal panel electrode defect by improved YOLOv7 [ J ]. Journal of Electronic Measurement and Instrumentation, 2023, 37 ( 9 ) : 225-233.
- [15] DUAN S, WANG T, LI T, et al. M-YOLOv8s: An improved small target detection algorithm for UAV aerial photography [ J ]. Journal of Visual Communication and Image Representation, 2024, 104: 104289.
- [16] HE Y, NIU X, HAO C, et al. An adaptive detection approach for multi-scale defects on wind turbine blade surface [ J ]. Mechanical Systems and Signal Processing, 2024, 219: 111592.
- [17] FENG S, GAO M, JIN X, et al. Fine-grained damage detection of cement concrete pavement based on UAV remote sensing image segmentation and stitching [ J ]. Measurement, 2024, 226: 113844.
- 杨宇龙,张银胜,段修贤,等. 基于轻量级卷积网络的风力发电机表面缺陷检测研究 [ J ]. 电子测量与仪器学报, 2024,38(8) :36-45.
- YANG Y L, ZHANG Y SH, DUAN X X, et al. Research on surface defect detection of wind turbine based on lightweight convolutional network [ J ]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8):36-45.
- [19] YANG X, DEL REY CASTILLO E, ZOU Y, et al. UAV-deployed deep learning network for real-time multi-class damage detection using model quantization techniques [ J ]. Automation in Construction, 2024, 159: 105254.
- [20] WANG AO, CHEN H, LIN Z J, et al. Repvit: Revisiting mobile CNN from vit perspective [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 15909-15920.
- [21] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [22] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning [ C ]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing ( ICASSP ). IEEE, 2023; 1-5.

### 作者简介



常潇丹,现为南京信息工程大学硕士研究生,主要研究方向为无人机智能视觉检测。

E-mail: 202312620011@nuist.edu.cn



冯浩(通信作者),2020 年于南京工业大学获得博士学位,现为南京信息工程大学副教授、硕士生导师、系主任,主要研究方向为人工智能技术。

E-mail: fenghao@nuist.edu.cn

**Feng Hao** ( Corresponding author ) received his Ph. D. degree in 2020 from Nanjing Tech University. He is now an associate professor, master's supervisor, and department head of Nanjing University of Information Science and Technology. His main research interest includes artificial intelligence technology.