

基于无人机图像和机器学习技术的钢结构表面腐蚀 智能识别方法研究

左吉寅¹, 周学军^{1*}, 王 振², 刘 哲¹, 孙继祥¹

(1. 山东建筑大学 土木工程学院, 济南 250101; 2. 山东交通学院 交通土建工程学院, 济南 250153)

摘 要: 随着以机器学习为核心的人工智能技术与无人机(unmanned aerial vehicle, UAV)图像技术不断发展, 为高空在役钢结构表面腐蚀的快速识别提供了新的思路。文中利用 UAV 采集既有钢结构建筑的腐蚀图像并建立数据集; 在现有机器学习方法的基础上, 提出了一种基于分层注意力的 FasterViT 与 YOLOv8 网络结构相结合的机器学习新算法, 实现了钢结构表面腐蚀区域的精准定位与腐蚀程度的快速智能识别。该算法识别准确率高于 YOLO 系列、FasterViT 单独使用情况下的准确率, 准确率达到 92.6%。文中所提方法可为在役钢结构服役性能的量化评估提供技术支撑。

关键词: 无人机图像技术; 机器学习; 分层注意力; 钢结构表面腐蚀; 智能识别

中图分类号: TP391.4

文献标志码: A

文章编号: 1671-9379(2026)09-0109-10

DOI: 10.13969/j.jzgjgz.20250206001

Research on Intelligent Identification Method for Steel Structure Surface Corrosion Based on UAV Images and Machine Learning Technology

ZUO Jiyin¹, ZHOU Xuejun^{1*}, WANG Zhen², LIU Zhe¹, SUN Jixiang¹

(1. School of Civil Engineering, Shandong Jianzhu University, Jinan 250101, China; 2. School of Transportation Civil Engineering, Shandong Jiaotong University, Jinan 250153, China)

*Corresponding author: xuejunzhou@sdjzu.edu.cn

Abstract: With the development of artificial intelligence technology centered on machine learning and unmanned aerial vehicle (UAV) imaging technology, new approaches are provided for the rapid identification of surface corrosion in high-altitude in-service steel structures. In this paper, UAV imaging technology was used to photograph existing steel structure buildings to obtain corrosion images and establish a dataset, and on the basis of existing machine learning methods, a new machine learning algorithm combining hierarchical-attention-based FasterViT and YOLOv8 network structure was proposed, which enabled the rapid and intelligent identification of corrosion severity and corrosion locations on the surface of steel structures. The recognition accuracy of this method was higher than those of the YOLO series and FasterViT alone, with an accuracy of 92.6%. The proposed method can provide technical support for the quantitative assessment of the service performance of in-service steel structures.

收稿日期: 2025-02-06

基金项目: 国家自然科学基金(52278507), 山东省重点研发计划(重大科技创新工程)项目(2024CXGC10321), 国家重点研发计划(SQ2020YFF0426523)

作者简介:

左吉寅(1998—), 男, 硕士研究生, 主要从事钢结构基本理论与空间钢结构方面的研究。E-mail: zuojiyin006@163.com。

通信作者:

周学军(1965—), 男, 博士, 教授, 主要从事钢结构基本理论与空间钢结构方面的研究。E-mail: xuejunzhou@sdjzu.edu.cn。

Keywords: unmanned aerial vehicle (UAV) imaging technology; machine learning; hierarchical attention; steel structure surface corrosion; intelligent identification

钢结构被广泛应用于大型体育场馆、高铁站房、大跨度桥梁等重大基础设施工程中。受自身构造特点以及风吹、雨淋、日晒等长期环境因素共同作用,钢结构表面会发生不同程度锈蚀。锈蚀不仅会降低钢结构的承载力与耐久性,严重时还会引发结构坍塌等重大安全事故。因此,及时、准确地检测识别钢结构表面腐蚀状况,做好维护管养工作,对保障工程安全具有重要意义。

针对高空在役钢结构,传统人工腐蚀识别方法存在效率低、识别精度有限、作业风险高等不足。随着无人机(unmanned aerial vehicle, UAV)图像技术的发展,UAV能够便捷抵达结构各类待检部位;UAV搭载三维高清扫描激光相机可采集钢结构表面图像数据。同时,以机器学习为核心的人工智能技术,已在工业制造、医疗诊断、安防监控等诸多领域展现出良好应用潜力。将两类技术有机结合,能够实现高空在役钢结构腐蚀的自动化识别分析,可大幅提升腐蚀识别的效率与准确度。

常规的钢结构腐蚀识别技术主要包括超声波检测^[1]、电磁检测^[2]、声发射检测^[3]等,但上述方法用于高空在役钢结构的腐蚀检测时,存在识别局限性较大、效率偏低、现场灵活性不足的问题。随着UAV与机器学习技术的快速发展,众多学者转向以卷积神经网络^[4](convolutional neural network, CNN)为基础的机器学习图像识别技术开展研究。CHEN等^[5]采用旋翼UAV采集钢结构构件图像,所提出的腐蚀智能识别算法可自动识别图像内的腐蚀区域,识别精度较高。HAN等^[6]基于UAV图像技术提出了一种腐蚀图像检测方法,建立腐蚀分割数据集,结合特征金字塔网络(FPN)与路径聚合网络(PANet)算法,得到识别精度较高的检测模型。JIANG等^[7]提出了一种融合深度学习与UAV视觉的桥梁螺栓状态自动分析方法,该方法能有效消除UAV采集数据的固有缺陷,实现螺栓缺陷的准确分类。JIANG等^[8]研发了一种爬壁无人机(UAS)系统,可获取无失真的混凝土构件裂缝细节图像;采用CNN对图像集开展训练与验证,证实了该系统与方法具备工程实用性。逯鹏等^[9]提出了一种基于Vision Transformer网络的钢结构表面锈蚀程度识别方法,该方法对锈蚀程度的分类准确率可达90%。

但目前公开的钢结构腐蚀开源数据集数量较少,而神经网络训练需要大量样本数据,造成钢结构腐蚀识别模型精度受限。为此,本文提出一种基于分层注意力的FasterViT^[10]与YOLOv8^[11]网络相结合的方法,开展钢结构表面腐蚀智能识别研究。利用UAV拍摄采集钢结构表面腐蚀图像,采用LabelImg工具完成腐蚀分类标

注,构建钢结构表面腐蚀数据集;基于自建数据集完成模型训练与测试,验证该算法对腐蚀识别检测的有效性。

1 UAV图像技术

1.1 图像采集

山东体育学院训练馆的外部采用了螺栓球节点空间网格结构(图1),但由于长期暴露于自然环境,在水分、氧气、污染物以及风霜雨雪、日照等因素作用下,钢构件易产生腐蚀。网架处于高空位置,采用人工识别腐蚀状态难度大。本文采用大疆 MATRICE 300 SERIES 四旋翼UAV,搭载ZENMUSE P1相机采集钢结构腐蚀高清图像,UAV设备参数见表1。依托大疆平台实现UAV精准定位、智能巡航与定点拍摄;该平台以多传感器信息融合系统为基础,集成视觉传感、超声传感、激光传感等多类传感器模块。针对UAV设计对应的飞行策略,对传感器采集信息进行补充与处理,提升整套系统智能化水平^[12]。

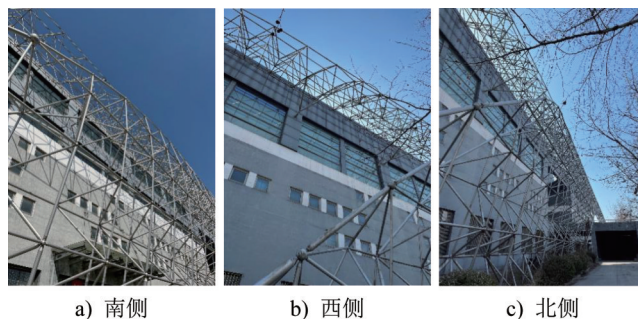


图1 外侧网架部分图片

Fig. 1 Partial picture of the outer mesh frame

表1 UAV设备参数

Table 1 UAV equipment parameters

设备	参数	取值
	尺寸/mm	810×670×430
	叶片直径/mm	15
	最大俯仰角度/(°)	30
	云台最大负重/g	930
	定位精度	1 cm + 1 ppm(水平) 1.5 cm + 1 ppm(垂直)
	飞行时间/min	30

1.2 实景模型建立

将UAV实拍照片导入到3D建模软件中,利用软件

中的空中三角测量功能,对导入的照片进行处理和分析,生成目标物体的三维坐标与纹理信息;手动添加像控点以提升模型精度,添加完成后重新进行空中三角测量,优化模型细节与精度^[13]。基于处理后的数据生成初始3D模型,对模型进行检查、优化、裁剪、缩放、旋转等操作,将优化后的3D模型导出为fbx格式,如图2所示。



图2 训练馆实景模型

Fig. 2 Training hall live model

依托该实景模型对训练馆螺栓球节点(图3)开展初步分析对比,对照模型与实拍影像,筛选腐蚀状态匹配的

螺栓球节点、桁架图像,作为数据集原始素材。

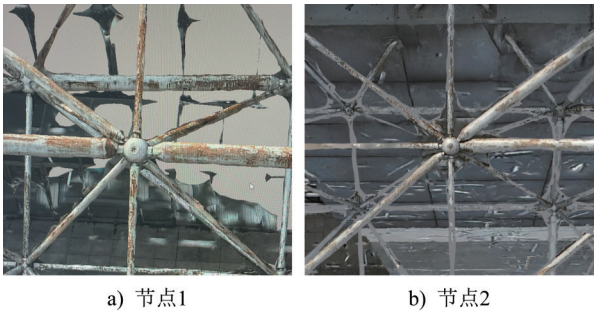


图3 螺栓球节点

Fig. 3 Bolt ball joints

1.3 数据集的分类及标注

本文以UAV采集的钢结构腐蚀高清图像作为数据源并完成标注,构建适配本文网络模型的数据集用于算法训练,数据集制作流程见图4。数据集划分为训练集、测试集、验证集三类:训练集用于模型迭代训练;测试集用于检验模型腐蚀识别功能是否正常;验证集用于评估训练完成模型对腐蚀程度的识别准确度。

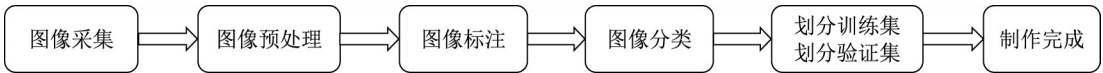


图4 数据集制作流程图

Fig. 4 Dataset production flowchart

对UAV原始图像做预处理,统一缩放分辨率至360像素×360像素;该分辨率既可实现数据集轻量化、提升识别推理速度,也可匹配后续腐蚀识别系统交互界面的输入要求。通过人工筛选剔除画面模糊、与腐蚀目标无

关的图像。最终构建的数据集共包含500张图像(图5),涵盖不同拍摄角度,同时包含点蚀、氧化皮局部剥落、大面积锈蚀、构件本体缺陷等多种环境干扰工况;其中训练集400张,验证集50张,测试集50张,部分样本见图5。

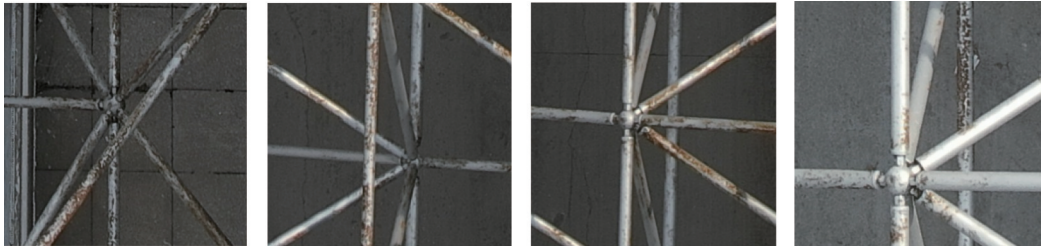


图5 腐蚀图像数据集部分样本

Fig. 5 Some samples from corrosion image dataset

数据集划分完成后,依据我国现行标准^[14]根据图像中钢结构腐蚀颜色深浅、腐蚀面积占比,将腐蚀划分为轻度腐蚀、中度腐蚀、重度腐蚀、极度腐蚀4个等级。当图像内存在多等级腐蚀区域重叠时,按照由重到轻的顺序极度、重度、中度、轻度标注最高腐蚀等级;轻度与中度腐蚀重叠时,若中度腐蚀区域占比大于50%,整体标记为中度腐蚀;重度与极度腐蚀重叠时,直接标记为极度腐蚀;通过统计区域内各等级腐蚀点密度,综合判定腐蚀等级,

分类标准见表2。

确定分类标准后开展数据标注,生成4类腐蚀标签;采用LabelImg工具完成数据集标注,标注过程如图6所示。在标注过程中对于区分阴影与锈蚀,通过视觉特征对比:(1)阴影颜色为单一灰色或黑色;纹理平滑渐变,构件表面无损伤;边缘与遮挡物形状一致;(2)锈蚀颜色为多色混杂包括橙红、黄褐、棕黑等颜色;表面存在颗粒、起泡、剥落或裂纹;边缘不规则扩散,边界模糊。对于灰

表2 腐蚀程度分类标准

Table 2 Classification criteria of corrosion degree

腐蚀程度	腐蚀程度表现状态
轻度腐蚀(mild_corrosion)	腐蚀颜色较浅,面积较小,表面涂层有轻微氧化皮剥落,存在少量点蚀。
中度腐蚀(moderate_corrosion)	腐蚀颜色较深但面积小或腐蚀面积大但腐蚀颜色较浅,腐蚀比较均匀,腐蚀颜色为浅棕色。
重度腐蚀(severe_corrosion)	表面已存在面积较大的腐蚀,表面涂层已有氧化皮剥落,腐蚀颜色较深且为棕色。
极度腐蚀(extreme_corrosion)	腐蚀颜色深且面积大,表面涂层已大部分剥落,腐蚀部分表现颜色呈黑色与深红色。

尘等附着物:(1) 颗粒分布均匀,无表面结构破坏;(2) 颜色与环境污染物相关(如灰色、土黄色)。标注时借助高通滤波器强化锈蚀纹理细节,进一步区分灰尘干扰;锈蚀多集中在焊缝、螺栓球节点、网架上侧等易积水高湿位置,灰尘更易堆积于水平构件表面。标注完成后输出目标框坐标,与图像一一对应;该标注方式可降低复杂边缘、密集点蚀区域的标注难度,得到标注精度较高的完整腐蚀数据集。

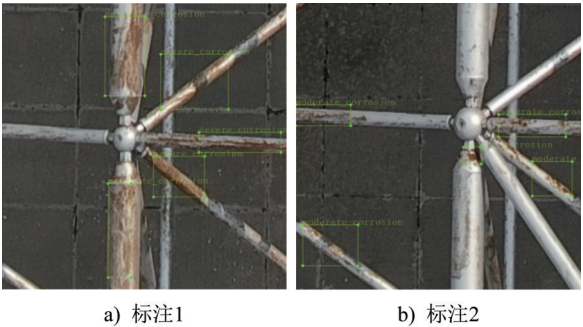


图6 腐蚀图像标注过程

Fig. 6 Corrosion image labeling process

2 机器学习原理

机器学习的核心是从数据中学习特征模式,实现对未知样本的预测与判别。机器学习^[15]的发展分为浅层学习、深度学习两个阶段。浅层学习模型结构简单,隐含层数量少甚至无隐含层,具备较强的自学习能力、非线性映射能力与鲁棒性^[16]。深度学习采用多层隐含层网络结构,特征提取学习能力更强;通过逐层初始化策略,有效降低神经网络训练难度。浅层学习与深度学习的原理对比见图7。

机器学习主要分为数据准备、模型训练、模型评估三个环节。首先采集足量任务相关数据,结合任务与数据特征选取适配模型,使用预处理后的数据开展训练;将数据集划分为训练集、验证集,通过多轮训练与验证循环,获取模型在多组子集上的性能指标,客观评价模型泛化能力。分类任务常用评价指标包含准确率、精确率、召回率、 F_1 分数;回归任务常用均方误差(MSE)、平均绝对误差(MAE)、决定系数(R^2)作为评价指标。

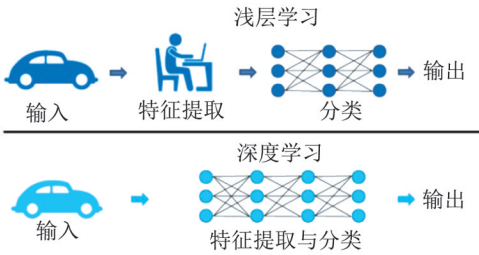


图7 浅层学习与深度学习^[16]

Fig. 7 Shallow learning and deep learning

2.1 YOLOv8 网络模型

YOLOv8是在YOLO系列算法基础上整合优化得到的网络结构,见图8。从图中分析可知,YOLOv8依靠多层卷积堆叠提取图像多尺度特征。同时融入残差连接(residual connections)^[17]缓解深层网络训练的梯度消失、梯度爆炸问题,在降低计算开销的同时提升训练效率与模型性能。YOLOv8采用了全新的主干网络(backbone)和检测头(head),兼顾检测精度与推理速度,适合实时检测场景;集成深度可分离卷积、特征金字塔网络(FPN)^[18]、路径聚合网络(PAN)等技术,以减少计算量并提高计算效率,且在各种复杂场景下表现出较高的鲁棒性和稳定性,相较于该系列中的旧版本,YOLOv8有较高的目标检测能力。为进一步提升识别精度与推理速度,本文对YOLOv8网络做改进优化处理。

2.2 FasterViT 网络模型

FasterViT (fast vision transformers with hierarchical attention)是NVIDIA团队提出的一种高效神经网络结构。该模型融合CNN局部特征学习与ViT全局建模优势,引入分层注意力(HAT)^[19]模块,在降低计算量的前提下增加窗口间的交互,可用作即插即用的增强模块,用于分类、对象检测、图像分割等任务,兼顾识别精度与图像推理吞吐量。FasterViT由4个不同阶段组成,在高分辨率阶段使用残差卷积块,后续阶段采用Transformer块;阶段之间通过步长卷积层降低输入图像分辨率,并利用HAT模块提取长、短距离空间关系,进行有效的跨窗口交互。FasterViT网络结构如图9所示,本文结合该网络优势对YOLOv8网络模型进行改进。

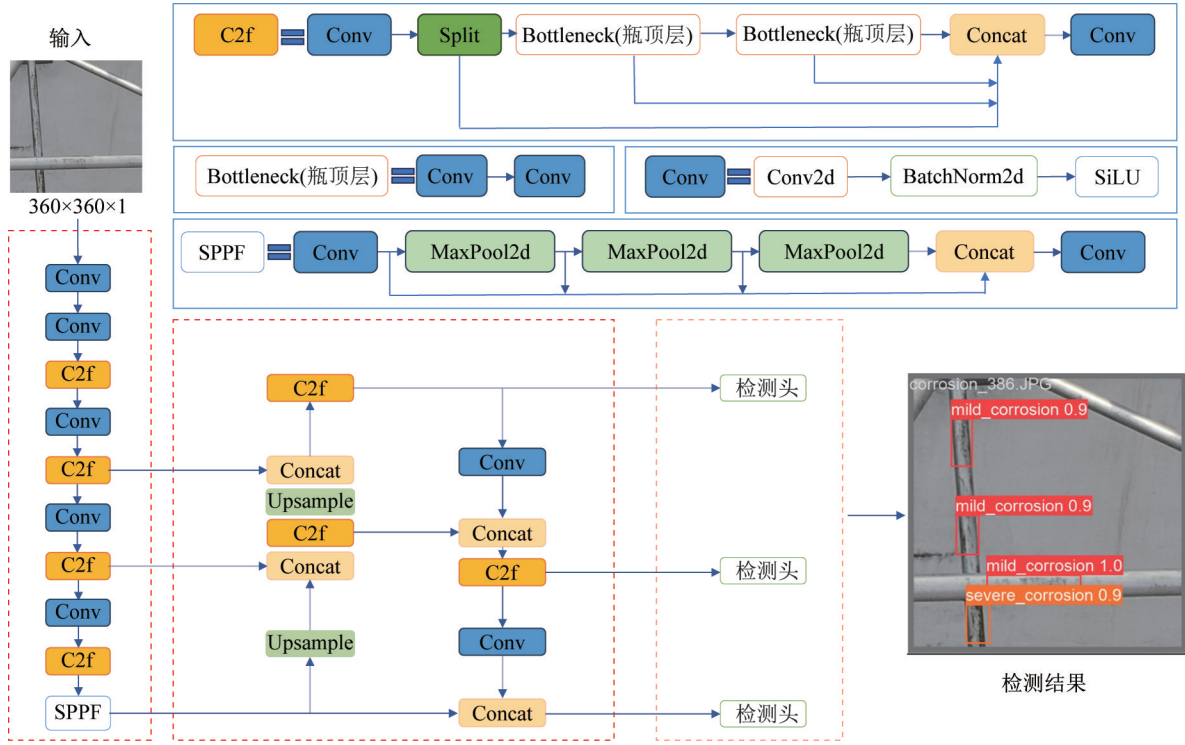


图8 YOLOv8网络结构

Fig. 8 YOLOv8 network structure

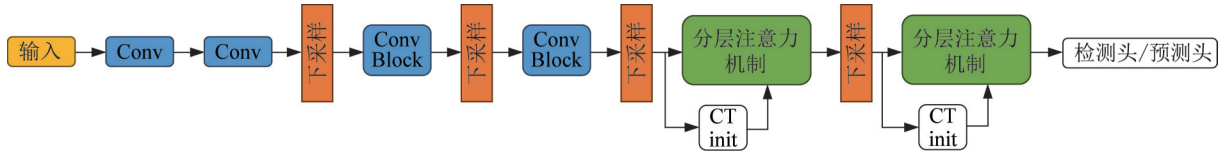


图9 FasterViT网络结构

Fig. 9 FasterViT network structure

2.3 改进 YOLOv8算法

为提升模型检测精度、优化计算资源分配,本文采用 FasterViT 模块与 YOLOv8 模块结合的方法,将 FasterViT 模块集成到 YOLOv8 颈部网络,替换颈部中原有的 C2f 模块,强化网络对腐蚀图像特征的提取能力,获得 FasterViT-YOLOv8 算法,如图 10 所示。由图可知, YOLOv8 的骨干网络 (backbone) 通过多次卷积输出多分辨率特征图,包含细粒度到全局不同层级视觉特征,送入颈部网络完成多尺度特征融合处理。

输入图像 $x \in R^{H \times W \times 3}$ 先经过两次 3×3 的卷积层,每个卷积层步长为 2,每次卷积后接入 ReLU 激活函数; FasterViT 对空间特征应用层归一化,再使用内核为 3×3 且步长为 2 的卷积层,将空间分辨率降低一半。

$$\hat{x} = \text{GELU}(\text{BN}(\text{Conv}_{3 \times 3}(x))) \quad (1)$$

$$x = \text{BN}(\text{Conv}_{3 \times 3}(\hat{x})) + x \quad (2)$$

式中:BN 表示批归一化; $\hat{x}$ 表示中间输出特征,经过 3×3 卷积、批归一化、GELU 激活运算后得到的中间特征; x 表

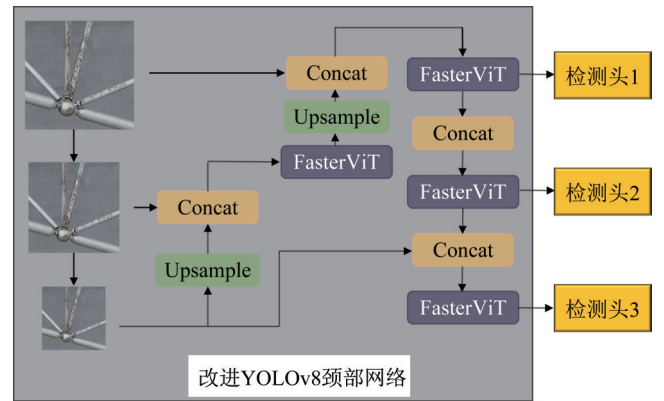


图10 FasterViT-YOLOv8颈部网络

Fig. 10 FasterViT-YOLOv8 neck network

示输入特征图,来自 FasterViT 模块的输入,为网络原始输入经过前期卷积处理后的特征张量;GELU 表示高斯误差线性单元非线性激活函数。

在 Swin-Transformer 局部窗口基础上, FasterViT 通过其窗口注意力模块引入 carrier tokens (CTs),用于汇总

窗口内特征信息,完成窗口之间交互。当给出一个输出特征图 $x \in R^{H \times W \times d}$, 其中 H 、 W 、 d 分别表示特征映射的高度、宽度、数量,设置 $H=W$, 以 $n=H^2/k^2$ 将输入特征图划分为 $n \times n$ 个局部窗口, 其中 k 为窗口大小, 计算公式为:

$$\hat{x} = \text{Split}_{k \times k}(x) \quad (3)$$

通过构件 CTs 可实现成本降低, 并获得较本地窗口更大的注意力覆盖范围。首先, 对各窗口的 L2c 特征执行池化操作, 完成 CT 初始化, 计算公式为:

$$\hat{x}_{ct} = \text{Conv}_{3 \times 3}(x) \quad (4)$$

$$\hat{x}_{ct} = \text{AvgPool}_{H^2 \rightarrow n^2 L}(\hat{x}_c) \quad (5)$$

式中: $\text{Conv}_{3 \times 3}$ 为 Twins 中使用的高效位置编码; $\hat{x}_{ct}$ 和 AvgPool 分别表示 carrier token 和特征池化操作。

在每个 HAT 块中, CT 都会经过以下处理:

$$\hat{x}_{ct} = \hat{x}_{ct} + \gamma_1 \cdot \text{MHSA}(\text{LN}(\hat{x}_{ct})) \quad (6)$$

$$\hat{x}_{ct} = \hat{x}_{ct} + \gamma_2 \cdot \text{MLP}_{d \rightarrow 4d \rightarrow d}(\text{LN}(\hat{x}_{ct})) \quad (7)$$

式中: LN 表示层归一化; MHSA 表示多头自注意力; γ 为可学习各通道特定的缩放因子; $\text{MLP}_{d \rightarrow 4d \rightarrow d}$ 为带有 GELU 激活函数的两层 MLP 结构; $\hat{x}_{ct}$ 为 FasterViT 提出的载体令牌特征; γ_2 为可训练参数。

最后, token 被进一步拆分为局部特征和 CT, 用于后续的分层注意力层:

$$\hat{x}_1, \hat{x}_{ct,1} = \text{Split}(\hat{x}_w) \quad (8)$$

式中: $\hat{x}_1$ 为局部窗口特征; $\hat{x}_{ct,1}$ 为拆分之后得到的新一代 carrier-token 载体令牌; $\hat{x}_w$ 为输入至 Split 操作的组合 token 特征。

改进后的模型通过 FasterViT 实现空间特征层归一化, 实现精度与吞吐量的平衡。在目标尺寸、形态等特征显著变化的情况下, 注意力机制可自适应调整不同特征权重, 确保模型训练过程中聚焦多样化的特征。适配 UAV 拍摄视角下的复杂检测场景, 提升模型鲁棒性与检测性能。

2.4 训练参数

本文采用计算机编程语言 Python3.8.10 构建 FasterViT-YOLOv8 模型, 使用自建钢结构腐蚀数据集开展训练, 本文所采用的设备和参数见表 3。训练初始学习率 (Learning Rate) 设置为 0.001, 用于控制参数更新速率, 阈值设置为 0.4, 训练循环数 epochs 设置为 500, 每轮训练包含 100 次迭代 (Steps)。为防止过拟合, 启用早停策略: 持续监控验证集指标, 当验证集性能不再提升甚至下降时终止训练, 保证训练损失与验证损失同步下降; 采用较小的 batch size 提高梯度更新频率。最终在独立测试集完成模型评估, 减小验证集参与超参数调优带来的结果偏差。

表 3 实验设备及参数

Table 3 Experimental equipment and parameters

名称	版本或型号
操作系统	windows10(64 bit)
GPU	NVIDIA GeForce RTX 3080
运行内存	32 GB
Cuda	Cuda-10.0
编程语言	Python3.8.10
编程工具	PyCharm
深度学习框架	PyTorch

3 智能识别结果分析

3.1 评价指标

利用不同算法识别出的结果分为四类, 分别为正确预测为正类的样本数 (T_P), 模型预测为正; 正确预测为负类的样本数 (F_P), 模型预测为负类; 错误预测为正类的样本数 (F_N), 该样本实际为负类, 但模型误判为正类; 错误预测为负类的样本数 (T_N), 该样本实际为正类, 但模型误判为负类。

本文采用准确率 A_{cc} (Accuracy)、精确率 P (Precision)、召回率 R (Recall)、平均准确率均值 m_{AP} (mean Average Precision)、 F_1 值及混淆矩阵综合评估模型性能。召回率和精确率是二元分类任务的关键评估指标, 其有助于评估模型的预测准确性。对模型进行分析, 各指标计算见式 (9)~(13):

(1) 准确率 (Accuracy)

准确率 A_{cc} 是正确预测为正的样本 (T_P) 占预测为正的总样本 ($T_P + F_P$) 的比例, 其计算方法如下:

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_N + F_P} \quad (9)$$

(2) 精确率 (Precision)

精确率 P 为所有被预测为正例的样本中, 真正正例所占比例, 该指标用于衡量模型预测正例的准确程度, 其计算方法如下:

$$P = \frac{T_P}{T_P + F_P} \quad (10)$$

(3) 召回率 (Recall)

召回率 R 表示所有正例样本中, 被正确识别为正例的样本占例, 反映模型对真正正例的检出覆盖水平, 其计算方法如下:

$$R = \frac{T_P}{T_P + F_N} \quad (11)$$

(4) F_1 值

F_1 值是精确率和召回率的调和平均值, 其计算方法如下:

$$F_1=\frac{2T_p}{2T_p+F_p+F_N}$$

(12)

(5) 平均准确率均值(mean Average Precision)

平均准确率均值 m_{AP} 则是用来度量模型在所有类别上的平均精确率,其计算方法如下:

$$m_{AP}=\frac{1}{N}\sum_{i=1}^NAP,P=\int_0^1P(R)dR$$

(13)

式中: N 为数据集的目标检测类别总数量。

3.2 模型训练

通过搭建 YOLOv8 改进模型网络环境,完成改进模

型网络、数据集配置文件、模型配置文件编写。在配置完毕后,依据电脑的具体性能参数(如 CPU、GPU 型号及内存大小等),合理设置训练过程中的各项参数开始训练。训练结果如图 11 所示,box_loss(边界框损失)、dfl_loss(目标损失)曲线呈持续下降并逐步趋于平稳,损失稳态值越低,代表边界框定位效果越好^[20]。同时,cls_loss(分类损失均值)曲线则反映分类的准确性,其值越小代表分类结果越准确。此外 P (精确度)曲线、 R (召回率)曲线及 m_{AP} (平均精度均值)随训练轮次增加逐步上升并趋近 0.9,说明模型训练收敛效果良好,检测精度较高。

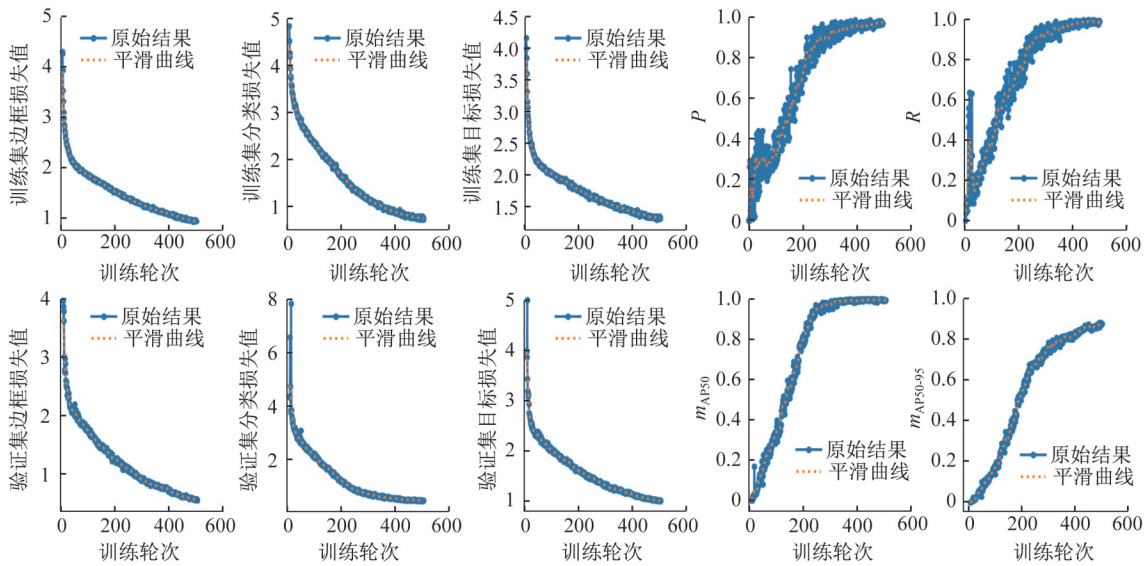


图 11 训练参数变化曲线

Fig. 11 Variation curves of training parameters

训练结束自动生成权重文件;将测试代码中的权重路径更新为生成的 best 权重文件,保持参数与训练阶段一致并运行测试代码,系统会输出测试结果。对比检测结果可知:原始 YOLOv8 模型存在部分识别缺陷;当采用改进后的算法进行检测时,上述问题得到显著改善,腐蚀图像检测置信度最高可达 92%,检测效果良好。

3.3 对比分析

为充分验证本文所提改进算法的优势与有效性,使用自建数据集对 YOLOv5、YOLOv6、YOLOv7、YOLOv8、FasterViT 及 FasterViT-YOLOv8 开展对比实验,测试结果见表 4。

表 4 不同的算法模型性能对比

Table 4 Performance comparison of different algorithmic models

算法	模型参数量/ $\times 10^4$	$R/\%$	$A_{cc}/\%$	$P/\%$	$m_{AP}@0.5/\%$	F_{PS}
YOLOv5	785	80.7	81.1	80.2	81.7	105
YOLOv6	973	81.2	81.7	81.6	81.9	113
YOLOv7	695	82.1	82.4	82.9	87.6	119
YOLOv8	562	82.6	85.3	83.8	88.2	136
FasterViT	634	80.3	83.6	83.9	84.7	115
FasterViT-YOLOv8	696	88.7	90.1	89.6	92.6	169

注: $m_{AP}@0.5$ 表示模型每秒钟可以处理完成的图像帧数,用来衡量目标检测模型的推理速度。

对比实验结果表明:本文提出FasterViT-YOLOv8综合指标最优;准确率为90.1%, $m_{AP}@0.5$ 为92.6%,优于其他YOLO系列以及原始FasterViT模型;相比FasterViT,FasterViT-YOLOv8的准确率提升了6.5%, $m_{AP}@0.5$ 提升了7.9%。综上所述,本文所提FasterViT-YOLOv8算法兼顾检测精度与推理速度。

表5为在自建数据集上各类别 $m_{AP}@0.5$ 的对比结果,FasterViT-YOLOv8算法在轻度、中度、重度、极度腐蚀四类上检测精度整体最优,其中中度腐蚀类别精度达到87.6%;仅极度腐蚀精度略低于YOLOv7精度,造成这一现象的原因在于网络结构、模型复杂度与调优程度存在差异。

表5 数据集各类别精度对比

Table 5 Accuracy comparison of dataset categories

%

算法	轻度腐蚀精度	中度腐蚀精度	重度腐蚀精度	极度腐蚀精度
YOLOv5	69.2	70.4	64.2	62.5
YOLOv6	73.8	71.9	69.7	68.4
YOLOv7	75.9	76.7	74.2	78.5
YOLOv8	80.3	82.7	80.6	74.5
FasterViT	76.5	77.1	71.5	69.8
FasterViT-YOLOv8	84.5	87.6	83.1	77.3

综上所述,FasterViT-YOLOv8的各项指标相比YOLOv8的各项指标提升明显,在全部对比模型中综合性能占优,验证了该改进算法在钢结构腐蚀目标检测任务中的有效性。

3.4 可视化分析

从自建数据集随机选取多组不同角度、场景复杂的

图像,对比YOLOv8与FasterViT-YOLOv8可视化检测结果,见图12。通过图12目标检测可视化结果对比可以看出,原始YOLOv8存在目标漏检、置信度偏低现象;改进的FasterViT-YOLOv8可以检出漏检腐蚀目标,检测置信度更高,有效缓解漏检、误检问题,整体检测性能得到提升。

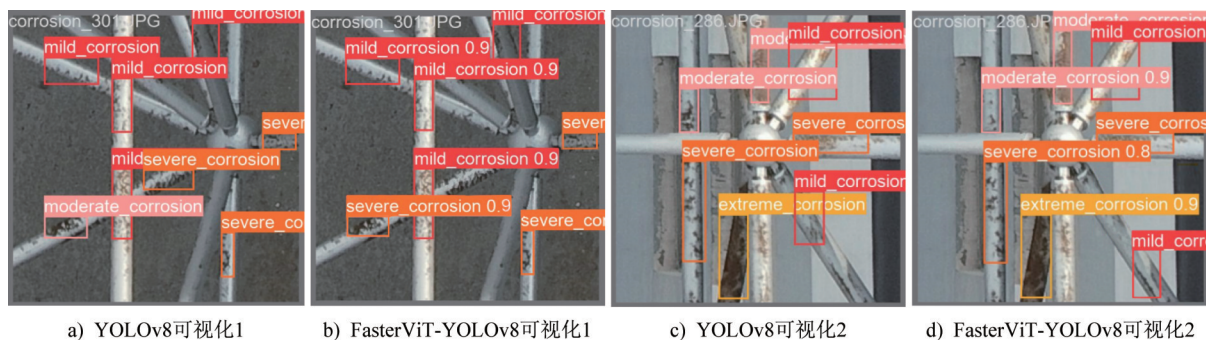


图12 可视化图像对比

Fig. 12 Visual image comparison

4 结论

本文以UAV采集的500张钢结构表面腐蚀图像作为数据源,对6种机器学习模型开展钢结构腐蚀识别对比实验,得到主要结论如下:

(1) 采用搭载高清相机的多旋翼UAV,可采集高空在役钢结构的腐蚀高清图像,以此构建适配模型训练的数据集;采集多视角、多干扰条件下的样本,是提升复杂现场下腐蚀识别性能的重要基础。

(2) 本文采用FasterViT模型对YOLOv8目标检测算法进行优化,优化后的YOLOv8算法在精确率上相对

于YOLOv5、YOLOv6及YOLOv7分别提高了9.0%、8.4%、7.7%,同时在 $m_{PA}@0.5$ 指标上分别提高了10.9%、10.7%、5.0%,目标检测能力显著增强。

(3) 本文构建的FasterViT-YOLOv8智能识别模型,通过对网络颈部模块优化,提升了钢结构腐蚀识别精度。将该算法应用于山东体育学院训练馆网架的腐蚀识别,钢管腐蚀识别准确率可达92.6%,具备工程应用价值。

参考文献:

[1] 李霄,梁中红,曹廷义.超声波C扫描技术在涉硫设备及管

- 道腐蚀检测中的应用[J].当代化工研究,2021(9):136-137. DOI:10.3969/j.issn.1672-8114.2021.09.063.
- LI Xiao, LIANG Zhonghong, CAO Tingyi. Application of ultrasonic C scanning technology in corrosion detection of sulfur-related equipment and pipeline [J]. Modern Chemical Research, 2021(9): 136-137. DOI: 10.3969/j.issn.1672-8114. 2021.09.063. (in Chinese)
- [2] 乔宁宁.大型钢结构钢板腐蚀的电磁检测方法[J].工业技术创新,2021,8(5):89-93. DOI: 10.14103/j.issn.2095-8412. 2021.10.016.
- QIAO Ningning. Electromagnetic detection method for steel plate corrosion in large steel structures [J]. Industrial Technology Innovation, 2021, 8(5): 89-93. DOI: 10.14103/j. issn.2095-8412.2021.10.016. (in Chinese)
- [3] 鞠晓臣,梁永奇,赵欣欣,等.采用声发射技术检测钢桥疲劳裂纹[J].哈尔滨工程大学学报,2023,44(4):649-656. DOI: 10.11990/jheu.202202028.
- JU Xiaochen, LIANG Yongqi, ZHAO Xinxin, et al. Fatigue crack detection of steel bridges by the acoustic emission technique[J]. Journal of Harbin Engineering University, 2023, 44(4): 649-656. DOI: 10.11990/jheu.202202028. (in Chinese)
- [4] 逯鹏,赵天淞,王剑,等.基于计算机视觉的钢结构表面损伤识别与健康监测综述[J].工业建筑,2022,52(10):22-27. DOI:10.13204/j.gyzG22071401.
- LU Peng, ZHAO Tiansong, WANG Jian, et al. Review on damage identification and health monitoring of steel structures based on computer vision [J]. Industrial Construction, 2022, 52(10): 22-27. DOI: 10.13204/j.gyzG22071401. (in Chinese)
- [5] CHEN Q C, WEN X, LU S, et al. Corrosion detection for large steel structure base on UAV integrated with image processing system [C]//Proceedings of 2019 International Conference on Aeronautical Materials and Aerospace Engineering (AMAE 2019). Shanghai: Shanghai Institute of Special Equipment Inspection & Technical Research, 2019: 6.
- [6] HAN Q H, ZHAO N, XU J. Recognition and location of steel structure surface corrosion based on unmanned aerial vehicle images [J]. Journal of Civil Structural Health Monitoring, 2021, 11(5): 1375-1392. DOI: 10.1007/s13349-021-00515-7.
- [7] JIANG S, ZHANG J. Real-time crack assessment using deep neural networks with wall-climbing unmanned aerial system [J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(6): 549-564. DOI: 10.1111/mice.12519.
- [8] JIANG S, ZHANG J, WANG W G, et al. Automatic inspection of bridge bolts using unmanned aerial vision and adaptive scale unification-based deep learning [J]. Remote Sensing, 2023, 15(2): 328. DOI: 10.3390/rs15020328.
- [9] 逯鹏,赵天淞,王剑,等.基于计算机视觉的钢结构表面锈蚀程度检测方法[J].工业建筑,2024,54(8):133-139. DOI: 10. 3724/j.gyzG23062009.
- LU Peng, ZHAO Tiansong, WANG Jian, et al. A method for detecting surface corrosion degree of steel structures based on computer vision [J]. Industrial Construction, 2024, 54(8): 133-139. DOI: 10.3724/j.gyzG23062009. (in Chinese)
- [10] KHAYATAZAD M, DE PUE L, DE WAELE W. Detection of corrosion on steel structures using automated image processing [J]. Developments in the Built Environment, 2020, 3: 100022. DOI: 10.1016/j.dibe.2020.100022.
- [11] 梁礼明,龙鹏威,李俞霖.改进轻量高效FMG-YOLOv8s的钢材表面缺陷检测算法[J].计算机工程与应用,2025, 61(3): 84-93. DOI: 10.3778/j.issn.1002-8331.2406-0358.
- LIANG Liming, LONG Pengwei, LI Yulin. Improved lightweight and efficient FMG-YOLOv8s algorithm for steel surface defect detection [J]. Computer Engineering and Applications, 2025, 61(3): 84-93. DOI: 10.3778/j.issn.1002-8331.2406-0358. (in Chinese)
- [12] 唐永,么学春,王晨,等.基于无人机图像技术和支持向量机(SVM)的桥梁裂缝自动识别系统[J].公路工程,2024, 49(6): 49-56. DOI: 10.19782/j.cnki.1674-0610.2024.06.008.
- TANG Yong, YAO Xuechun, WANG Chen, et al. Automatic bridge crack detection system based on UAV image technology and support vector machine (SVM) [J]. Highway Engineering, 2024, 49(6): 49-56. DOI: 10.19782/j.cnki.1674-0610.2024.06. 008. (in Chinese)
- [13] 王爽.基于无人机摄像的桥梁结构位移检测研究[D].重庆:重庆交通大学,2022.
- WANG Shuang. Research on bridge structure displacement detection of unmanned aerial vehicle camera [D]. Chongqing: Chongqing Jiaotong University, 2022. (in Chinese)
- [14] GB/T 8923.1—2011 涂覆涂料前钢材表面处理 表面清洁度的目视评定 第1部分:未涂覆过的钢材表面和全面清除原有涂层后的钢材表面的锈蚀等级和处理等级[S].
- GB/T 8923.1—2011 Preparation of steel substrates before application of paints and related products—Visual assessment of surface cleanliness: Part 1: Rust grades and preparation grades of uncoated steel substrates and of steel substrates after overall removal of previous coatings[S]. (in Chinese)
- [15] 肖乾浩.基于机器学习理论的机械故障诊断方法综述[J].现代制造工程,2021(7):148-161. DOI: 10.16731/j.cnki.1671-3133.2021.07.022.
- XIAO Qianhao. Review on mechanical fault diagnosis methods based on machine learning theories [J]. Modern Manufacturing Engineering, 2021(7): 148-161. DOI: 10.16731/j.cnki.1671-3133.2021.07.022. (in Chinese)
- [16] 郝善行.基于计算机视觉的钢结构局部腐蚀检测及特征提取[D].哈尔滨:哈尔滨工业大学,2022.
- HAO Shanxing. Local corrosion detection and feature extraction of steel structure based on computer vision [D]. Harbin: Harbin Institute of Technology, 2022. (in Chinese)
- [17] HATAMIZADEH A, HEINRICH G, YIN H X, et al. FasterViT: Fast vision transformers with hierarchical attention [EB/OL]. [2025-02-06]. 2023: arXiv: 2306.06189. <https://arxiv.org/abs/2306.06189>.

