

基于SPDConv-EMA与自适应加权融合的改进YOLO11n风机叶片缺陷检测方法

朱 向¹, 张伶俐¹, 卢成武^{1,2*}

¹重庆文理学院群与图的理论及应用重庆市高校重点实验室, 重庆

²重庆第二师范学院数学与大数据学院, 重庆

收稿日期: 2026年7月3日; 录用日期: 2026年8月3日; 发布日期: 2026年8月11日

摘 要

针对无人机航拍风机叶片缺陷检测中小目标细节易丢失、缺陷边缘纹理特征不明显、复杂背景干扰明显及多尺度特征融合不足等问题, 本文提出一种融合SPDConv、C3k2_EMA与AWFF的改进YOLO11n检测方法。该方法在Backbone关键下采样位置采用SPDConv替换普通跨步卷积, 以减少裂纹、边缘磨损等细长小目标在降采样过程中的空间信息损失; 在深层特征提取单元中嵌入EMA注意力机制, 构建C3k2_EMA模块, 以增强缺陷区域响应并抑制阴影、反光和叶片纹理等背景干扰; 在Neck关键融合节点引入AWFF模块, 通过可学习权重自适应平衡浅层细节与深层语义特征。实验结果表明, 改进模型的Precision、Recall、mAP@0.5和mAP@0.5:0.95分别达到84.8%、84.3%、86.7%和57.6%, 较原始YOLO11n分别提高1.2、3.9、2.5和2.4个百分点, 验证了本文方法在风机叶片表面缺陷检测任务中的有效性。

关键词

风机叶片, 表面缺陷检测, YOLO11n, SPDConv, C3k2_EMA, AWFF

An Improved YOLO11n Method for Wind Turbine Blade Defect Detection Based on SPDConv-EMA and Adaptive Weighted Fusion YOLO11n

Xiang Zhu¹, Lingli Zhang¹, Chengwu Lu^{1,2*}

¹Chongqing Key Laboratory of Group & Graph Theories and Applications, Chongqing University of Arts and Sciences, Chongqing

²School of Mathematics and Big Data, Chongqing University of Education, Chongqing

*通讯作者。

文章引用: 朱向, 张伶俐, 卢成武. 基于 SPDConv-EMA 与自适应加权融合的改进 YOLO11n 风机叶片缺陷检测方法[J]. 计算机科学与应用, 2026, 16(8): 65-79. DOI: 10.12677/csa.2026.168263

Abstract

To address the loss of small-defect details, weak edge-texture representation, complex background interference, and insufficient multi-scale feature fusion in UAV-based wind turbine blade inspection, this paper proposes an improved YOLO11n method integrating SPDConv, C3k2_EMA, and AWFF. SPDConv is used to replace ordinary strided convolutions at key down sampling positions in the backbone, reducing spatial information loss for slender cracks and edge defects. The EMA attention mechanism is embedded into deep feature extraction units to construct C3k2_EMA, which enhances responses to defective regions while suppressing shadows, highlights, and blade texture interference. In the neck, AWFF is introduced at key fusion nodes to adaptively balance shallow detail features and deep semantic features through learnable weights. Experimental results show that the improved model achieves Precision, Recall, mAP@0.5, and mAP@0.5:0.95 of 84.8%, 84.3%, 86.7%, and 57.6%, respectively, improving the original YOLO11n by 1.2, 3.9, 2.5, and 2.4 percentage points and demonstrating its effectiveness for wind turbine blade surface defect detection.

Keywords

Wind Turbine Blade, Surface Defect Detection, YOLO11n, SPDConv, C3k2_EMA, AWFF

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着全球能源结构加快变革、“双碳”行动持续深入推进,“以风代煤”的风力发电作为清洁能源之一,已成为构建绿色低碳能源体系的关键支撑。《风能北京宣言 2.0》提出,“十五五”期间我国风电年度新增装机容量不低于 1.2 亿千瓦;到 2030 年,力争实现我国风电总装机量达到 13 亿千瓦,到 2035 年达到 20 亿千瓦,到 2060 年达到 50 亿千瓦[1]。近年来,风电装机容量逐渐增大,对风电机组安全高效运行以及智能化运维提出了更高的要求。但风电机组在恶劣环境条件下常年运转,其主要部件的状态好坏直接关系到机组的出力性能、稳定可靠性及使用年限等问题,因此对风电机组设备进行状态监测和故障识别的研究具有重要的现实意义。

风机叶片是作为风电机组进行风能捕获并完成能量转换的关键装置之一,在运行过程中始终处于高风速、雨雪、沙尘、紫外线、温差、海上盐雾等多种环境下,易出现裂纹、损伤、腐蚀、污染、磨损、表皮脱落等表面缺陷,并可能导致其在长期交变载荷的作用下继续发展。造成叶片气动性能恶化及结构强度减弱,甚至是严重的安全事故。传统的风机叶片检测手段有目视、望远镜观测、吊篮及绳索作业等,也有利用超声检测、红外热成像以及振动信号分析进行无损检测的研究工作,但人工检测具有效率低、费用高、主观性强、安全性差等缺点,而一些无损检测的方法又存在着检测设备复杂、检测环境受限以及专业技术人员要求高等缺陷,不能适应大容量风电场的快速巡视工作。

随着无人机平台、高清视觉传感器和深度学习技术的发展,基于无人机航拍图像的风机叶片表面缺陷检测逐渐成为智能巡检的重要方式。Shihavuddin 等将深度学习方法与无人机图像相结合,用于风机叶片表面检测,验证了视觉智能检测方法在提升叶片巡检自动化水平方面的可行性[2]。Xu 等利用无人机拍

摄图像进行风机叶片表面检测研究,进一步说明无人机航拍图像与深度学习方法相结合能提高叶片表面缺陷检测的自动化程度[3]。Yang 等将迁移学习及集成学习的方法应用到风机叶片损伤图像识别中,并提升了模型对复杂叶片图像的缺陷识别能力[4]。Heo 等人总结了使用无人机进行风机叶片检查的相关技术并指出无人机巡检具有高安全性和灵活性及高巡检率的优点[5]。除风机叶片外,深度学习方法也被用于多类可再生能源设备表面损伤检测,说明基于图像的智能缺陷识别方法具有较好的工程推广价值[6]。总体来看,无人机巡检与深度学习检测方法相结合,能够在降低人工巡检风险和成本的同时,为风机叶片缺陷识别、状态评估和后续维护决策提供了有效支撑。

由于深度学习目标检测算法的不断发展,风机叶片缺陷识别也有了新的方法。Faster R-CNN 等两阶段检测算法用候选区域生成、分类和边界框回归完成目标检测,一般具有较好的检测精度,但是模型结构较复杂,推理速度较慢[7]。SSD 这类单阶段检测算法可以直接在特征图上完成对目标类别的类别预测以及位置预测,因此它的检测效率比较高[8]。YOLO 系列算法把目标定位和类别预测合在一起,构成单阶段检测架构,在速度跟准确度之间找到了较为不错的均衡[9]。

在 YOLO 系列不断更新迭代下,YOLOv7 在训练过程以及网络结构上对实时检测进行了优化提升[10];YOLOv9 引入了可编程梯度来增强训练阶段的有效信息流传输,这为实时检测模型的构建提供了新方法[11];YOLOv10 面向端到端实时目标检测任务进行结构设计,在减少后处理依赖和提升推理效率方面具有代表性[12]。因此,YOLO 系列模型较适合无人机巡检、边缘端部署和实时缺陷检测等应用场景。

在风机叶片表面缺陷检测研究中,研究者们已对 YOLO 系列模型开展了多种改进。王俊等提出一种改进 YOLOv5s 的风机叶片表面缺陷检测方法,通过优化模型的结构,提升模型在复杂场景下的缺陷识别能力[13]。高文俊等针对风机叶片缺陷检测任务,对 YOLOv5 进行改进,引入图像增强、注意力机制和多尺度特征提取等策略,提高了模型对风机叶片缺陷的检测精度[14]。曾勇杰等将改进 YOLOv8 算法应用于风机叶片缺陷检测,在网络结构中引入多尺度卷积、动态检测头和改进边界框损失函数,以增强模型对不同尺度缺陷的检测能力[15]。李大华等面向无人机端部署需求,提出改进 YOLOv8n 的风机叶片表面缺陷轻量化检测网络,在降低模型参数量和计算量的同时,改善小目标缺陷检测效果[16]。李松霖等基于 YOLO11n 开展风机叶片表面缺陷检测研究,通过引入注意力机制、渐进式下采样和多尺度注意力特征金字塔网络,在检测精度和模型复杂度之间都取得了较好效果[17]。上述研究表明,围绕特征提取、注意力增强、多尺度融合和轻量化设计对 YOLO 模型进行改进,是提升风机叶片缺陷检测性能的重要方向。

然而,在航拍图像和工业缺陷检测场景中,小目标通常存在像素占比低、边缘信息弱和背景干扰强等问题,容易导致漏检和定位偏差[18]。针对低分辨率图像和小目标检测中下采样造成的空间信息损失问题,Sunkara 等提出 SPDConv 结构,通过 Space-to-Depth 操作将空间信息重排至通道维度,以减少跨步卷积或池化造成的细节丢失[19]。在复杂背景下,注意力机制能够增强网络对关键区域的特征表达能力。通道注意力机制能够通过建模不同通道特征的重要性,对特征响应进行自适应重标定,从而增强网络对关键区域的表达能力[20];空间注意力与通道注意力的联合建模能够进一步提升卷积神经网络对目标区域和有效特征的选择能力[21]。为了弥补一些复杂场景下特征响应不足的情况,Ouyang 等提出了 EMA 高效多尺度注意力模块,在不同空间上进行学习以提高网络对有效区域的特征表达能力,并在控制计算开销的同时提高多尺度的空间信息建模能力[22]。除此之外,多尺度特征融合也是提升缺陷检测性能的一个重要环节。Lin 等人提出一种特征金字塔网络结构,采用自上而下和横向链接的方法来融合不同层间的特征,并提高了模型对不同尺度的目标的检测性能[23]。在以上模型的基础上,PANet 对特征金字塔进行了扩展,在其上进一步加入了从下到上的路径聚合方法来增强由低层的位置信息向上层语义信息的传递作用[24]。EfficientDet 中 BiFPN 利用双向跨尺度连接及权重相加的方法增强了特征间的相互交流[25]。ASFF

能够利用其自适应性来学习不同尺度的特征之间的融合权重，在不同的目标尺度及特征响应下，使得检测网络能自动地调整各个尺度的信息贡献量[26]。

尽管已有方法在风机叶片缺陷检测上取得了较好的效果，但是在无人机航拍图像中仍然存在着缺陷尺度小、边缘纹理弱、背景干扰大、多尺度特征融合不好的问题，容易造成小目标漏检、复杂背景误检和缺陷定位不准。本文以 YOLO11n 为基础模型进行改进。在骨干网路下采样环节，首先用 SPDCConv 模块减少空间细节信息的丢失，提高模型对于裂纹、局部损伤等微小缺陷的感知能力；其次在 C3k2 模块里加入 EMA 注意力机制得到 C3k2_EMA 特征增强模块来加强模型对缺陷关键部位的重视程度，消除了阴影、反光和背景纹理等外界干扰因素；最后在颈部网络设计 AWFF 自适应加权特征融合模块对不同尺度的特征进行自适应加权，增强浅层细节信息与深层语义的融合。

2. YOLO11n 基础模型

2.1. YOLO11n 网络结构

YOLO11n 是一种基于单阶段的目标检测网络，整个模型分为 Backbone、Neck 和 Head 三个部分。其中 Backbone 用来提取输入图像多层级的特征表达，Neck 用来完成多尺寸特征图的传递与融合，Head 则在融合特征上实现目标类别预测及边框位置回归。YOLOv11n 网络结构见图 1。

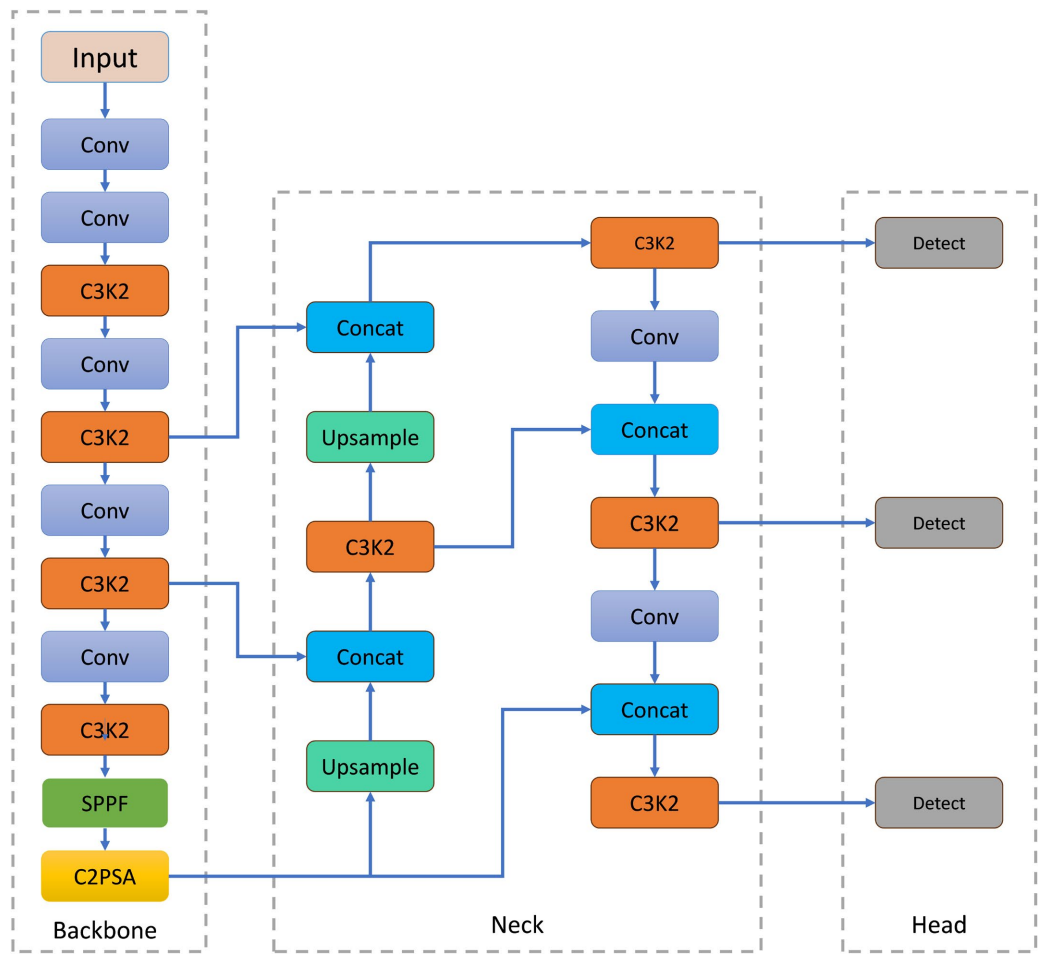


Figure 1. Network architecture of YOLO11n
图 1. YOLO11n 网络结构图

2.2. Backbone 特征提取网络

Backbone 特征提取网络对输入图像进行层次化特征提取,网络主要由卷积层、C3k2 和 SPPF 及 C2PSA 模块组成。图像进入骨干网络中,获得的特征图的空间分辨率逐级减小,但通道数量逐渐增加,让网络从较浅的层级中提取边缘、纹理等局部信息,逐步转移到更深层次的高层语义信息表征上。其中,卷积层主要是对基本特征进行提取并完成下采样处理;C3k2 模块是利用跨阶段特征链接及轻量级卷积设计实现高效地提取特征;SPPF 模块是拓展模型的感受野并对多尺度上下文信息进行融合;C2PSA 模块是对模型空间特征表达能力进行提升。YOLOv11n 采用主干网络来完成特征提取,可以对 Neck 多尺度特征融合及 Head 检测头提供不同的层级特征图。

2.3. Neck 多尺度特征融合网络

Neck 位于 Backbone 和 Head 之间,主要作用是融合主干网络输出的不同尺度特征,使检测头能够同时利用浅层细节信息和深层语义信息。YOLO11n 的 Neck 延续 YOLO 系列常用的特征金字塔融合思想,通过上采样、下采样、特征拼接和 C3k2 等操作,对来自不同层级的特征图进行逐步整合。

在多尺度特征融合过程中,高分辨率浅层特征包含较丰富的细节信息,低分辨率深层特征具有更强的语义表达能力。Neck 通过自顶向下和自底向上的路径传递与融合特征,增强了不同尺度目标的表征能力,从而为检测头提供更充分的多尺度特征输入。

2.4. Head 检测头

YOLO11n 的 Head 是最终检测输出部分,负责在多尺度特征图上完成目标类别预测和边界框回归。检测头接收来自 Neck 输出的多尺度特征图,并在不同尺度上分别进行预测,以适应大、中、小不同尺寸目标的检测需求。通过分类分支和回归分支,Head 输出目标类别、位置信息及置信度检测结果。

Head 模块是 YOLO11n 的最终检测输出层,在多个尺度上的特征图进行目标类别判断并回归边界框。检测头接受 Neck 多个层次的特征图并在每个尺度上单独预测以实现对于大、中、小不同尺度目标的检测任务。通过其内部分支的分类分支及回归分支,Head 给出目标的类别标签、空间位置及对应的目标置信度等检测结果。

2.5. YOLO11n 检测原理与损失函数

YOLO11n 属于单阶段目标检测模型,它对目标进行检测是在多个尺度上的特征图同时进行分类及定位。对于一张待检测图片来说,Backbone 和 Neck 首先提取并融合不同层级的特征信息,随后检测头在 P3、P4 和 P5 三个尺度上进行预测。设第 i 个尺度的特征图为:

$$F_i \in \mathbb{R}^{H_i \times W_i \times C_i}, i \in \{3, 4, 5\}$$

对于输入图像尺寸 $H \times W$,不同尺度特征图的空间尺寸与网络步长之间满足如下关系:

$$H_i = \frac{H}{s_i}, W_i = \frac{W}{s_i}, s_i \in \{8, 16, 32\}$$

其中, H_i 表示特征图的高度, W_i 表示特征图的宽度, C_i 表示特征图的通道数。P3 特征图分辨率较高,主要用于小目标检测;P4 和 P5 特征图包含更强的语义信息,分别适用于中尺度和大尺度目标检测。通过多尺度预测,模型能够兼顾裂纹、腐蚀、污染和损伤等不同尺寸缺陷的识别需求。

YOLO11n 检测头采用分类分支与回归分支解耦的预测方式。对于第 i 个尺度的特征图,其预测结果可表示为:

$$P_i = \{B_i, C_i\}$$

其中, B_i 表示边界框回归分支输出, C_i 表示类别预测分支输出。解耦检测头能够使位置回归和类别判别分别学习更适合自身任务的特征表达, 从而减弱分类任务与定位任务之间的相互干扰, 提高缺陷类别判断和边界定位的稳定性。

在边界框回归方面, YOLO11n 采用 anchor-free 思想, 不再依赖预设锚框宽高, 而是以特征点为参考位置预测目标框四个方向的距离。同时, 模型利用分布焦点损失思想对距离进行离散分布建模。设每个方向的距离被划分为 k 个离散区间, 网络输出为 p_k , 则连续距离预测值可表示为:

$$\hat{d} = \sum_{k=0}^{K-1} k \cdot \text{Softmax}(p_k)$$

其中, $\text{Softmax}(p_k)$ 表示第 k 个离散区间对应的概率, d 为概率加权后得到的连续距离估计值。这种分布式回归方式能够提供比直接回归单一坐标更细致的位置信息, 有助于提升缺陷边缘和细小目标的定位精度。

设特征点坐标为 (x_a, y_a) , 模型预测该点到目标框左、上、右、下边界的距离分别为 d_l 、 d_t 、 d_r 和 d_b , 则解码后的预测框可表示为:

$$\hat{B} = (x_a - \hat{d}_l, y_a - \hat{d}_t, x_a + \hat{d}_r, y_a + \hat{d}_b)$$

由于检测头在不同尺度特征图上进行预测, 解码得到的特征图坐标还需根据对应步长映射回原始图像尺度:

$$\hat{B}_{\text{img}} = s_i \cdot \hat{B}_{\text{feat}}$$

其中, B 为预测目标框坐标。该解码方式直接由特征点位置和四边距离确定目标边界, 避免了锚框匹配带来的额外先验约束。

在类别预测方面, 检测头对每个候选位置输出各缺陷类别的 logit 值, 并通过 Sigmoid 函数得到对应类别概率:

$$\hat{p}_c = \sigma(z_c) = \frac{1}{1 + e^{-z_c}}$$

其中, z_c 为类别 c 的原始输出, p_c 为类别 c 的预测概率。

训练过程中, YOLO11n 通过任务对齐方式进行正样本分配, 使分类置信度和定位质量共同参与候选样本选择。对于候选预测框, 其任务对齐指标可表示为:

$$m = s^\alpha \cdot u^\beta$$

其中, s 表示预测类别置信度, u 表示预测框与真实框之间的 IoU, α 和 β 为权重系数。本文实验所采用实现中 $\alpha = 0.5$ 、 $\beta = 6.0$, 该策略能够优先选择分类得分较高且定位质量较好的候选点作为正样本, 从而提高训练目标的一致性。

YOLO11n 的检测损失主要由边界框回归损失、分类损失和 DFL 损失组成。边界框回归损失用于衡量模型预测的预测框与真实框之间的位置偏差, 可表示为:

$$L_{\text{box}} = L_{\text{Clou}} = 1 - \text{Clou} = 1 - \text{IoU} + \frac{\rho^2(b, b^*)}{c^2} + \alpha v$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2$$

$$\alpha = \frac{v}{1 - \text{IoU} + v}$$

CIoU 综合考虑预测框与真实框的重叠面积、中心点距离和宽高比一致性。

分类损失用于约束每个候选位置对各类别缺陷的预测概率，可表示为二元交叉熵形式：

$$L_{cls} = -\sum_{c=1}^C [y_c \log(\hat{p}_c) + (1 - y_c) \log(1 - \hat{p}_c)]$$

DFL 损失则用于监督真实连续距离在相邻离散区间上的概率分布，使边界框距离预测更加精细：

$$L_{dfl} = -(w_l \log P_l + w_r \log P_r)$$

模型训练的总损失函数可表示为：

$$L = \lambda_{box} L_{box} + \lambda_{cls} L_{cls} + \lambda_{dfl} L_{dfl}$$

其中， λ_{box} 、 λ_{cls} 和 λ_{dfl} 是用于平衡边界框损失、分类损失及 DFL 损失的系数值，通过对以上损失函数的最小化可以实现对目标检测位置准确性、类别区分度和预测边界框分布准确性的共同优化，在此基础上进一步进行小目标信息保护、复杂背景信息强化和特征金字塔多尺度特征提取策略设计以获得最终的检测结果。

3. 改进 YOLO11n 算法设计

针对风机叶片无人机巡检图像中小尺度缺陷细节易丢失、复杂背景干扰明显以及多尺度缺陷特征融合不足等问题，本文以 YOLO11n 为基准模型，对其 Backbone 和 Neck 结构进行针对性改进。整体思路为：在 Backbone 的关键下采样阶段引入 SPDConv，提升小目标空间细节保存能力；在深层特征提取阶段设置 C3k2_EMA 模块，增强网络对缺陷关键区域的响应；在 Neck 多尺度融合阶段设置 AFFF 模块，使不同尺度特征可以根据检测任务需要进行贡献值自适应分配。改进模型架构及各模块嵌入位置如图 2 所示。

3.1. 改进理论依据

风机叶片缺陷图像具有明显特异性的任务，具体包括：裂纹和边缘磨损常表现为细长线状结构，像素占比小且边界灰度变化弱；污染、腐蚀和表皮脱落则形态不规则、尺度差异较大，并易与阴影、反光或叶片正常纹理混淆。以及风机叶片缺陷图像的特异性，本文改进的模型需要同时处理三个问题，包括下采样细节保留、复杂背景下的缺陷区域强化以及跨尺度信息动态融合。基于对前期问题的分析以及多种模块的调研，本文选择 SPDConv、C3k2_EMA 和 AFFF 来分别处理上述三个关键问题。

在下采样环节，普通卷积(stride=2)或池化会直接降低特征图空间分辨率，虽然计算效率较高，但容易削弱细小裂纹的连续边缘和局部缺口的位置信息。SPDConv 通过 Space-to-Depth 将相邻空间像素重排到通道维度，再采用卷积(stride=1)完成特征整合，相比直接跨步采样能够保留更多局部邻域关系。考虑到过早或过多替换会增加通道冗余和计算负担，本文仅在 Backbone 中承担尺度压缩的关键下采样位置嵌入 SPDConv，使小缺陷细节在进入深层语义提取前得到更充分保留。

在注意力机制选择上，SE 侧重通道权重建模，对裂纹这类空间方向性较强的缺陷定位帮助有限；CBAM 同时建模通道和空间注意力，但串行注意力会带来额外计算开销，并可能在轻量模型中降低推理效率。EMA 通过分组方式降低参数和计算量，并沿高度、宽度方向聚合空间上下文，能够兼顾细长裂纹的方向结构和污染、腐蚀等不规则区域的局部上下文。基于轻量化检测和复杂背景抑制的平衡，本文将 EMA 嵌入 C3k2 模块中构建 C3k2_EMA，并布置在 Backbone 较深层的特征提取单元，使其在已获得一定语义表达后对缺陷响应进行重标定。

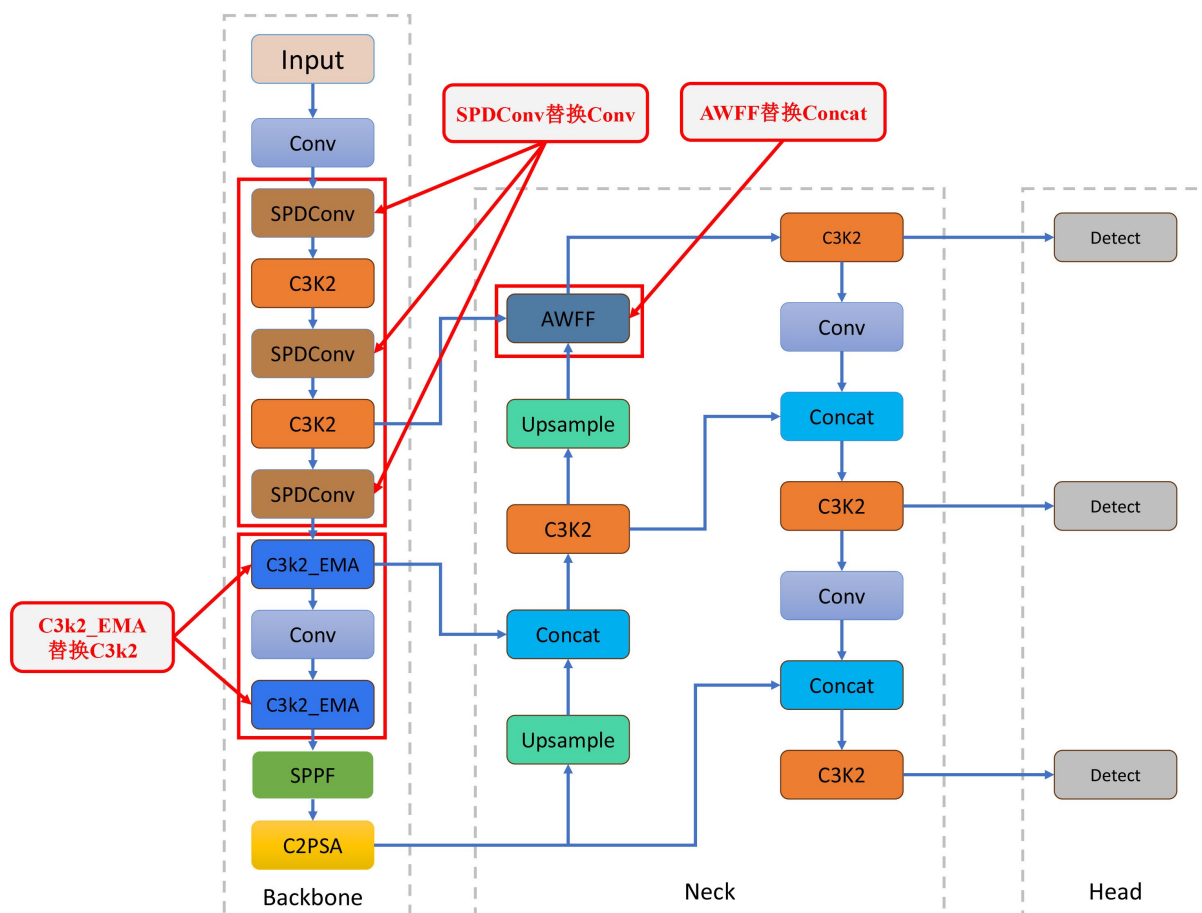


Figure 2. Network architecture of the improved YOLO11n

图 2. 改进 YOLO11n 网络结构图

在特征融合环节, 直接 Concat 虽然保留了多尺度特征, 但各分支贡献固定, 难以适应裂纹、污染、损伤等缺陷在尺度和纹理强度上的差异; 简单相加会进一步弱化分支差异, 容易造成浅层细节或深层语义被平均化。AWFF 通过可学习权重和归一化约束对不同尺度输入进行自适应分配, 使网络能够根据图像内容动态强调浅层细节或深层语义。本文将 AWFF 设置在 Neck 中由 Backbone 输出特征进入自顶向下融合的关键节点, 用以优先调节跨层信息流, 再配合后续 Concat 和 C3k2 完成特征整合, 从而在精度提升与结构复杂度之间取得较好平衡。

3.2. SPDConv 下采样模块

风机叶片表面缺陷中裂纹、边缘磨损、局部破损等目标一般都存在尺度较小、形状细长、边界模糊和纹理对比度较低的特点。原始 YOLO11n 在进行特征提取的过程中, 需要经过 $\text{stride} = 2$ 的卷积操作实现下采样, 从而缩小特征图规模、节省运算量, 但是跨步采样会导致空间分辨率下降, 容易导致小尺度缺陷的边界、纹理及位置特征信息由浅层向深层传播的过程中被弱化, 对无人机遥摄的小尺度或者低清度的缺陷图像来说, 极易造成漏检以及定位漂移现象。

为此, 本文在主干网络关键下采样位置引入 SPDConv 模块(如图 3 所示), 用于替换部分普通步长卷积。SPDConv 由 Space-to-Depth 操作和卷积($\text{stride} = 1$)组成, 其核心思想不是直接丢弃部分空间采样点, 而是先将相邻空间位置的信息重新排列到通道维度, 再通过卷积完成通道调整与局部特征融合。该设计

能够在完成下采样的同时保留更多相邻像素之间的局部关系，有利于增强细小裂纹、边缘缺口和弱纹理缺陷的表达能力，并为后续注意力增强和多尺度融合提供更完整的基础特征。

3.3. C3k2_EMA 特征增强模块

无人机巡检图像中的风机叶片表面往往存在涂层纹理、阴影、强光反射、污渍以及天空或地面背景等干扰因素，这些区域在颜色、亮度和纹理上可能与真实缺陷相似。原始 C3k2 模块具有较好的特征提取效率，但其对关键缺陷区域与干扰背景之间的区分能力仍然有限。当缺陷边缘较弱或背景纹理较复杂时，模型容易将背景异常误判为缺陷，或忽略真实缺陷中的细微结构。

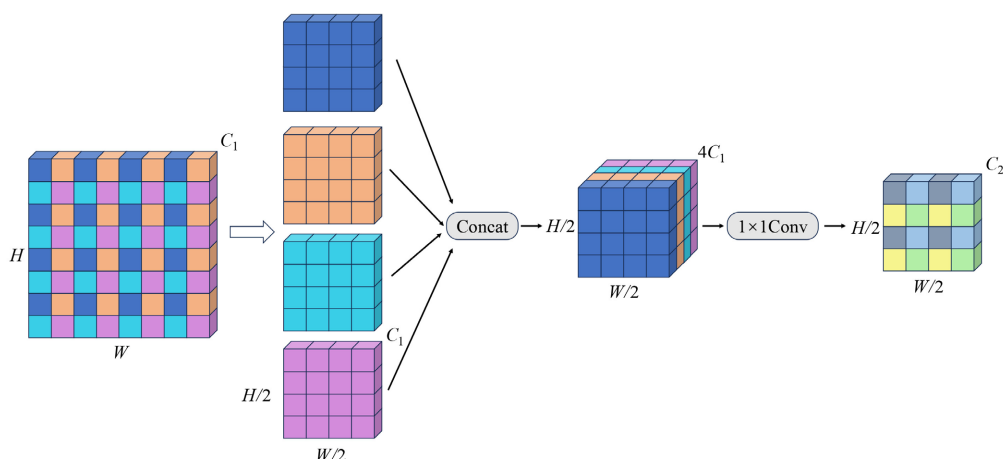


Figure 3. Structure of the SPDCnv module

图 3. SPDCnv 结构图

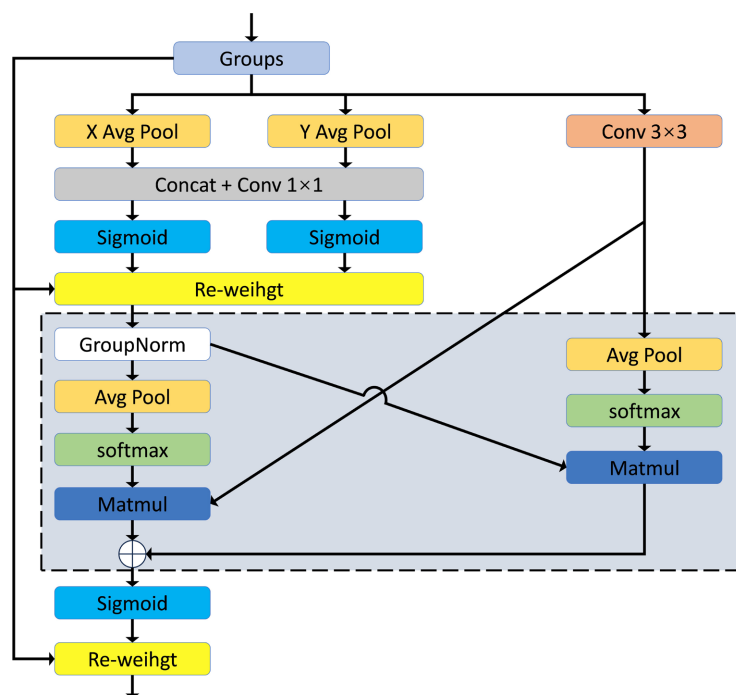


Figure 4. Structure of the EMA module

图 4. EMA 结构图

对此, 本文将 EMA 高效多尺度注意力机制(如图 4 所示)嵌入 C3k2 模块, 构建 C3k2_EMA 特征增强模块(如图 5 所示)。该模块在保持 C3k2 基本结构和轻量化优势的基础上, 将部分 Bottleneck 替换为 Bottleneck_EMA, 使特征在完成卷积提取后进一步进行注意力重标定。EMA 通过分组特征建模降低额外计算开销, 并分别沿高度和宽度方向聚合空间上下文信息, 使模型能够从不同空间方向捕获缺陷区域的结构关系。

C3k2_EMA 能够增强网络对裂纹、腐蚀、涂层脱落、污染等缺陷区域的响应, 同时抑制阴影、反光和叶片正常纹理等无关信息。对于细长裂纹类缺陷, 该模块有助于突出局部纹理差异; 对于污染、破损等面积较大的缺陷, 该模块能够结合更充分的空间上下文进行判断。C3k2_EMA 从特征增强层面提升模型在复杂背景下的鲁棒性和缺陷区域定位能力。

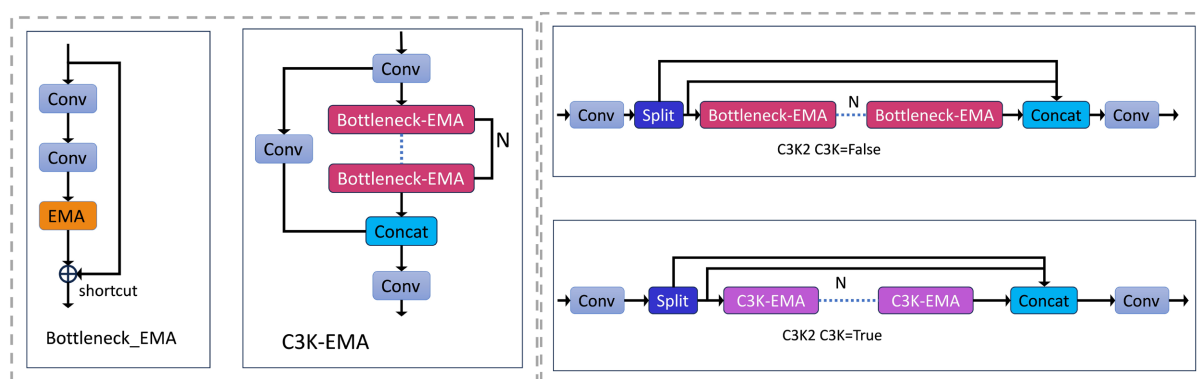


Figure 5. Structure of the C3k2_EMA module

图 5. 3k2_EMA 结构图

3.4. AWFF 自适应加权特征融合模块

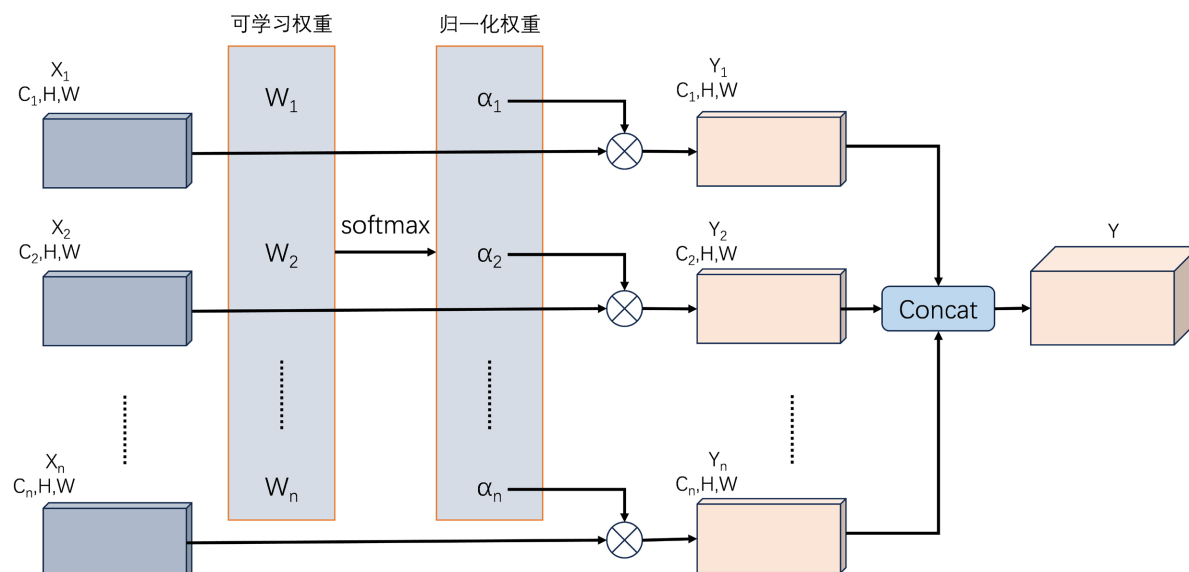


Figure 6. Structure of the AWFF module

图 6. AWFF 结构图

风机叶片表面缺陷类型多样, 不同类别在尺度、形态和纹理强度上差异明显。例如, 裂纹通常依赖

浅层特征中的边缘和纹理细节，污染或大面积损伤则更依赖深层特征中的语义信息和上下文关系。原始 YOLO11n 的 Neck 通过 Concat 操作融合不同尺度特征，虽然能够实现多层信息拼接，但这种固定融合方式默认各分支贡献相同，难以根据不同图像和不同缺陷类型动态调整特征权重。

为此，本文在 Neck 的关键融合节点引入 AWFF 自适应加权特征融合模块(如图 6 所示)，用于替换部分直接 Concat 操作。AWFF 为每个输入特征分支设置可学习权重，并通过 softmax 函数进行归一化，使网络在训练过程中自动学习不同尺度特征的重要程度。

AWFF 能让模型能够根据检测对象的尺度和特征强度自适应选择更有效的特征来源。当图像中以细小裂纹或局部磨损为主时，网络可提高浅层细节特征的贡献；当图像中存在污染、腐蚀或较大面积损伤时，网络可更多利用深层语义特征。相比固定拼接方式，AWFF 提升了特征融合的灵活性，有助于增强模型对多尺度缺陷和复杂场景变化的适应能力。

4. 实验设计

4.1. 数据集介绍

本文所用风机叶片表面缺陷数据集来自 Roboflow 网站，总共有 3594 幅图片，分辨率均为 640×640 。该数据集中包括 5 类风机叶片表面缺陷，分别是 damage、dirt、erosion、pollution 和 crack。图片标注工作全部由 LabelImg 完成。将数据集分为训练集、验证集和测试集，划分的比例为 8:1:1，其中训练集有 2875 张图片，验证集有 359 张，测试集有 360 张。

4.2. 实验环境与训练参数设置

本文实验在 Ubuntu 22.04 操作系统下进行，深度学习框架采用 PyTorch 2.5.1，CUDA 版本 12.4，编程语言为 Python 3.12。模型训练与测试均在单张 GPU 上完成，所使用的训练设备为 vGPU-32GB，显存大小为 32 GB。实验训练采用参数为：输入图像分辨率(image-size)为 640×640 ，批量大小(batch-size)为 32，训练轮数(epochs)为 200，优化器为 AdamW，初始学习率(lr0)为 0.001，动量参数(momentum)为 0.937，权重衰减系数(weight-decay)为 0.0005，预热训练轮数(warmup-epochs)为 3.0。

4.3. 评价指标

本文采用 Precision (P)、Recall (R)、mAP@0.5、mAP@0.5:0.95 等指标对模型进行综合评价。其中，Precision 表示预测为缺陷的目标中真实缺陷所占的比例；Recall 表示所有真实缺陷中被模型成功检测出的比例；mAP@0.5 表示 IoU 阈值为 0.5 时各类别平均精度的均值；mAP@0.5:0.95 表示 IoU 阈值从 0.5 到 0.95、步长为 0.05 时的平均精度。主要公式如下：

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

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

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i$$

其中， TP 表示模型预测结果为正类实际为正类的数目， FP 表示模型预测结果为正类实际为负的数目， FN 表示模型预测结果为负类实际为正类的数目， n 表示需要检测缺陷的总类别数。

5. 实验结果与分析

5.1. 消融实验

为验证本文提出的改进模块的有效性,本文以原始 YOLO11n 为基准模型,添加各改进模块,对比其性能提升。实验结果如表 1 所示。

Table 1. Ablation experiment results

表 1. 消融实验结果

模型	SPDConv	C3k2_EMA	AWFF	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
1	×	×	×	83.6	80.4	84.2	55.2
2	√	×	×	83.2	83.0	84.4	56.9
3	×	√	×	83.4	83.4	84.3	57.3
4	√	√	×	84.7	83.2	85.9	57.0
5	√	√	√	84.8	84.3	86.7	57.6

由消融实验的结果可知,在原始 YOLO11n 中 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 分别为 83.6%、80.4%、84.2%和 55.2%;在引入 SPDConv 后 Recall 提升到 83.0%,mAP@0.5:0.95 提升到了 56.9%,表明 SPDConv 能够有效地缓解普通下采样所造成的细节信息丢失的问题,进而增强模型对于小尺度缺陷的检出能力及定位精度。当引入 C3k2_EMA 模块后,Recall 和 mAP@0.5:0.95 分别为 83.4%和 57.3%,表明 EMA 注意力机制能够增强网络对于缺陷重要位置的特征激活程度,在杂乱背景下能提高检测可靠性。

当加入 SPDConv 和 C3k2_EMA 后,模型 Precision 和 mAP@0.5 分别达 84.7%和 85.9%,相较原始 YOLO11n 分别增加 1.1 和 1.7 个百分点,说明二者的细节保留以及特征增强是互补的。在此基础上再加入 AWFF,此时模型 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 分别达到 84.8%、84.3%、86.7%和 57.6%,4 个指标均为最佳。相对于原始 YOLO11n 分别提升了 1.2、3.9、2.5 和 2.4 个百分点,表明了本文提出的 3 个改进模块可以协同提高风机叶片表面缺陷检测性能。

5.2. 对比实验

为进一步验证本文方法的有效性与可行性,将本文改进模型与 YOLOv5n、YOLOv8n、YOLOv9t、YOLOv10n 和 YOLO26n 等主流轻量化检测模型进行对比。所有模型均在同一数据集、同一实验环境下进行。实验结果如表 2 所示。

Table 2. Performance comparison of different models

表 2. 不同模型性能对比

模型	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv5n	83.5	81.5	84.9	55.3
YOLOv8n	81.4	82.9	83.7	56.2
YOLOv9t	80.7	82.8	84.4	56.4
YOLOv10n	80.2	82.2	83.4	55.4
YOLO26n	80.5	79.8	81.6	56.0
改进 YOLO11n	84.8	84.3	86.7	57.6

由表 2 可以看出, 本文提出的改进 YOLO11n 模型在 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 四项指标上均取得最优结果, 分别达到 84.8%、84.3%、86.7%和 57.6%, 表明本文改进的方法具有较好的综合检测性能。与 YOLOv5n、YOLOv8n、YOLOv9t 和 YOLOv10n 等模型相比, 本文方法在 Precision 和 Recall 上均有提升, 表明模型不仅能够更准确地区分真实缺陷与复杂背景干扰, 减少误检现象, 同时也能够检测出更多真实缺陷目标, 降低漏检率。其中, Recall 的提升说明 SPDConv 在下采样过程中能够保留更多小尺度缺陷的边缘和纹理信息, C3k2_EMA 则增强了模型对裂纹、损伤、腐蚀等关键区域的关注能力, 使模型在复杂背景下具有更稳定的缺陷识别能力。

从 mAP 指标来看, 本文方法的 mAP@0.5 达到 86.7%, 相比其他对比模型表现更优, 说明改进模型在常规 IoU 阈值下具有更高的整体检测精度。同时, 本文方法的 mAP@0.5:0.95 达到 57.6%, 高于各对比模型, 表明模型在更严格的定位评价条件下仍具有较好的检测框回归能力。AWFF 自适应加权特征融合模块能对不同尺度特征贡献进行动态调整, 使浅层细节信息与深层语义信息能够更充分融合。综合来看, SPDConv、C3k2_EMA 和 AWFF 三个模块在细节保留、特征增强和多尺度融合方面形成互补, 使改进 YOLO11n 在风机叶片表面缺陷检测任务中表现出更好的检测精度、召回能力和定位质量。

6. 结论

针对无人机巡检场景下风机叶片表面缺陷小、形态差异大、背景干扰强等问题, 本文提出了一种融合 SPDConv、C3k2_EMA 和 AWFF 三大模块的改进 YOLO11n 检测方法。通过在 YOLO11n 主干网络中引入 SPDConv 模块, 使得该模型在下采样过程中能够保留更多空间细节信息; 通过构建 C3k2_EMA 特征增强模块, 使得该模型对缺陷区域的关注能力得到增强; 通过在 Neck 中引入 AWFF 自适应加权特征融合模块, 使得不同尺度特征之间的信息交叉融合更加充分。实验结果表明, 本文改进的模型在风机叶片表面缺陷数据集上的 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 分别为 84.8%、84.3%、86.7%和 57.6%, 较原始 YOLO11n 分别提升 1.2、3.9、2.5 和 2.4 个百分点。数值实验验证了本文方法能够在保持轻量化结构的同时提升缺陷识别精度、召回能力和定位质量。

本文方法虽然在提升缺陷识别精度上有一定的提升能力, 也存在一定局限性, 主要表现在两个方面。当缺陷目标极小、成像模糊或缺陷颜色与叶片表面高度接近时, 本文研究的模型仍可能出现定位偏差; 对于强反光、大面积遮挡以及样本数量较少的罕见缺陷类别, 本文的方法检测稳定性仍有进一步提升空间。后续的研究将从更细粒度的数据增强、跨场景域适应、轻量化部署优化以及结合红外或多光谱信息的多模态检测方向展开, 以提高模型在真实风场复杂巡检环境中的泛化能力和工程应用价值。

基金项目

国家自然科学基金青年科学基金项目(62101082), 重庆市自然科学基金创新发展联合基金(CSTB2023NSCQ-LZX0101), 重庆市教委科学技术研究项目(KJZD-M202301303)。

参考文献

- [1] 刘俊伶.《风能北京宣言 2.0》发布未来五年装机目标翻倍式上调[N]. 证券时报, 2025-10-21(A06).
- [2] Shihavuddin, A.S.M., Chen, X., Fedorov, V., Nymark Christensen, A., Andre Brogaard Riis, N., Branner, K., *et al.* (2019) Wind Turbine Surface Damage Detection by Deep Learning Aided Drone Inspection Analysis. *Energies*, **12**, Article 676. <https://doi.org/10.3390/en12040676>
- [3] Xu, D., Wen, C. and Liu, J. (2019) Wind Turbine Blade Surface Inspection Based on Deep Learning and UAV-Taken Images. *Journal of Renewable and Sustainable Energy*, **11**, Article No. 053305. <https://doi.org/10.1063/1.5113532>
- [4] Yang, X., Zhang, Y., Lv, W. and Wang, D. (2021) Image Recognition of Wind Turbine Blade Damage Based on a Deep

- Learning Model with Transfer Learning and an Ensemble Learning Classifier. *Renewable Energy*, **163**, 386-397. <https://doi.org/10.1016/j.renene.2020.08.125>
- [5] Heo, S.J. and Na, W.S. (2025) Review of Drone-Based Technologies for Wind Turbine Blade Inspection. *Electronics*, **14**, Article 227. <https://doi.org/10.3390/electronics14020227>
 - [6] Shihavuddin, A.S.M., Rashid, M.R.A., Maruf, M.H., Hasan, M.A., Haq, M.A.U., Ashique, R.H., *et al.* (2021) Image Based Surface Damage Detection of Renewable Energy Installations Using a Unified Deep Learning Approach. *Energy Reports*, **7**, 4566-4576. <https://doi.org/10.1016/j.egy.2021.07.045>
 - [7] Ren, S., He, K., Girshick, R., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems* **28**, 2015, 91-99.
 - [8] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., *et al.* (2016) SSD: Single Shot Multibox Detector. In: *Lecture Notes in Computer Science*, Springer International Publishing, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
 - [9] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
 - [10] Wang, C., Bochkovskiy, A. and Liao, H.M. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/cvpr52729.2023.00721>
 - [11] Wang, C., Yeh, I. and Mark Liao, H. (2024) YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T. and Varol, G. Eds., *Lecture Notes in Computer Science*, Springer, 1-21. https://doi.org/10.1007/978-3-031-72751-1_1
 - [12] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., *et al.* (2024) YOLOv10: Real-Time End-to-End Object Detection. *Advances in Neural Information Processing Systems* **37**, Vancouver, 10-15 December 2024, 107984-108011. <https://doi.org/10.52202/079017-3429>
 - [13] 王俊, 高贵兵. 一种改进 YOLOv5s 算法的风机叶片表面缺陷检测方法[J/OL]. 中国机械工程: 1-11. <https://link.cnki.net/urlid/42.1294.th.20250403.1149.002>, 2026-06-05.
 - [14] 高文俊, 张海峰. 改进的 YOLOv5 风机叶片缺陷检测方法[J]. 建模与仿真, 2023, 12(4): 3574-3586.
 - [15] 曾勇杰, 范必双, 杨涯文, 等. 改进 YOLOv8 算法在风机叶片缺陷检测上的应用[J]. 电子测量与仪器学报, 2024, 38(8): 26-35.
 - [16] 李大华, 吴超强, 高强, 等. 改进 YOLOv8n 的风机叶片表面缺陷轻量化检测网络[J]. 电子测量与仪器学报, 2025, 39(8): 145-155.
 - [17] 李松霖, 唐玲. 基于改进 YOLO11 的风机叶片表面缺陷检测研究[J/OL]. 电子测量技术: 1-17. <https://link.cnki.net/urlid/11.2175.TN.20260205.1612.011>, 2026-06-05.
 - [18] Tong, K. and Wu, Y. (2022) Deep Learning-Based Detection from the Perspective of Small or Tiny Objects: A Survey. *Image and Vision Computing*, **123**, Article 104471. <https://doi.org/10.1016/j.imavis.2022.104471>
 - [19] Sunkara, R. and Luo, T. (2023) No More Strided Convolutions or Pooling: A New CNN Building Block for Low-Resolution Images and Small Objects. In: Amini, MR., Canu, S., Fischer, A., Guns, T., Kralj Novak, P. and Tsoumakas, G. Eds., *Lecture Notes in Computer Science*, Springer, 443-459. https://doi.org/10.1007/978-3-031-26409-2_27
 - [20] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
 - [21] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
 - [22] Ouyang, D., He, S., Zhang, G., Luo, M., Guo, H., Zhan, J., *et al.* (2023) Efficient Multi-Scale Attention Module with Cross-Spatial Learning. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, 4-10 June 2023, 1-5. <https://doi.org/10.1109/icassp49357.2023.10096516>
 - [23] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2117-2125. <https://doi.org/10.1109/cvpr.2017.106>
 - [24] Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J. (2018) Path Aggregation Network for Instance Segmentation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 8759-8768. <https://doi.org/10.1109/cvpr.2018.00913>

-
- [25] Tan, M., Pang, R. and Le, Q.V. (2020) EfficientDet: Scalable and Efficient Object Detection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 10781-10790.
<https://doi.org/10.1109/cvpr42600.2020.01079>
- [26] Liu, S., Huang, D. and Wang, Y. (2019) Learning Spatial Fusion for Single-Shot Object Detection. arXiv:1911.09516.