

# 改进 YOLOv8n 的风机桨叶表面缺陷轻量化检测网络

李大华<sup>1,2</sup> 吴超强<sup>1,2</sup> 高 强<sup>1,2</sup> 于 晓<sup>1,2</sup>

(1. 天津理工大学电气工程与自动化学院 天津 300384; 2. 天津市新能源电力变换传输与智能控制重点实验室 天津 300384)

**摘要:**针对当前风机桨叶缺陷检测算法存在的检测精度不足、复杂背景下误检漏检发生率高和模型不便部署等问题,提出一种改进 YOLOv8n 的风机桨叶缺陷检测算法。首先,引入 MobilenetV2 中的关键模块加以改进提出全新的额外残差(Extra-IB)模块和 C2f-Extra-IB 模块,用于减少模型参数量实现轻量化并为后续特征融合通过高质量特征图;其次提出自适应高效多尺度空间金字塔(AEMFP)模块替换快速空间池化金字塔(SPPF)模块,该模块创新地采用融合高效多尺度注意力机制(EMA)和并行子结构的设计,提高算法的多尺度特征融合和特征自适应提取能力;最后在颈部网络引入高效局部注意力机制(ELA),削弱复杂环境对于检测效果的影响同时提高对小目标的检测精度。使用风机桨叶表面缺陷数据集进行消融实验和对比实验,所提算法平均精度均值(mAP)达到 81.7%,相较于 YOLOv8n 提升 5.1%,模型参数量和浮点计算数分别为  $2.09 \times 10^6$  和 5.4 GFLOPS,减少 22.3% 和 21.7%,模型体积减少 19.8%,检测帧速达到 45.57 fps,说明所提改进措施可以在保证算法检测精度提高的同时实现轻量化,满足使用无人机等计算资源有限的检测设备进行高效、精确的风机桨叶缺陷检测的需求。

**关键词:** 风机桨叶; 目标检测; YOLOv8n; 轻量化; 注意力机制

中图分类号: TM315; TP183; TN98 文献标识码: A 国家标准学科分类代码: 520.2060

## Improved YOLOv8n lightweight detection network for wind turbine blade surface defects

Li Dahua<sup>1,2</sup> Wu Chaoqiang<sup>1,2</sup> Gao Qiang<sup>1,2</sup> Yu Xiao<sup>1,2</sup>

(1. School of Electrical Engineer and Automation, Tianjin University of Technology, Tianjin 300384, China; 2. Tianjin Key Laboratory of New Energy Power Conversion, Transmission and Intelligent Control, Tianjin University of Technology, Tianjin 300384, China)

**Abstract:** Aiming at the problems of the current fan blade defect detection algorithm, such as insufficient detection accuracy, high incidence of false detection under complex background and inconvenient deployment of model, an improved YOLOv8n fan blade defect detection algorithm was proposed. Firstly, a new Extra-IB module and C2f-Extra-IB module are introduced to improve the key modules in MobilenetV2, which are used to reduce the number of model parameters to achieve lightweight and pass high-quality feature maps for subsequent feature fusion. Secondly, the AEMFP module is proposed to replace the SPPF module, which innovatively integrates the EMA attention mechanism and parallel substructure design to improve the multi-scale feature fusion and feature adaptive extraction capability of the algorithm. Finally, ELA attention mechanism is introduced into the neck network to reduce the influence of complex environment on the detection effect and improve the detection accuracy of small targets. Ablation experiments and comparison experiments were conducted using fan blade surface defect data set. The proposed algorithm mAP reached 81.7%, an increase of 5.1% compared with YOLOv8n. The number of model parameters and floating-point calculations were  $2.09 \times 10^6$  and 5.4 GFLOPS, respectively, decreasing by 22.3% and 21.7%. The model size is reduced by 19.8% and the detection frame speed reaches 45.57 frames. It shows that the improvement measures proposed in this paper can not only improve the detection accuracy of the algorithm, but also achieve lightweight, which can meet the demand of using the detection equipment with limited computing resources such as UAV for efficient and accurate fan blade defect detection.

**Keywords:** wind turbine blade; target detection; YOLOv8n; lightweight; attention mechanism

## 0 引言

风力发电是指通过风力发电机将风能转化为电能的新型发电方式,这是目前全球范围内应对气候恶化、进行能源转型的重要方式之一。随着国家新能源产业转型的进程加快,壮大新型清洁能源产业,加快可再生能源的循环开发利用已成为国家发展过程中的一项重大任务<sup>[1]</sup>。

风力发电机一般设立在平原、荒漠和近海位置,工作环境普遍恶劣,面临风沙、雷击、台风和潮湿空气腐蚀等问题,风机叶片作为风力发电机的核心部件,承担着将风能转化为机械能的重要功能,在长时间的运行中极易出现裂纹、油污、砂眼等缺陷<sup>[2]</sup>,这些缺陷不仅会影响发电机组的电能转化效率,严重情况下还会缩短风机的使用寿命增加不必要的维护成本。因此,对风机叶片进行定期巡查检修,及时发现并解决潜在危险,对于保障风力发电机组的正常运行、降低维修成本和发电质量具有重大意义。

近年来,计算机视觉和无人机技术的快速发展,使得将二者结合进行缺陷检测和图像采集成为各国的研究热点。现阶段基于计算机视觉的目标检测算法主要分为两种:一种是以 YOLO<sup>[3-10]</sup> (you only look once)、SSD<sup>[11]</sup> (single shot multibox detector)、RetinaNet<sup>[12]</sup>等为代表的单阶段检测方法;另一种是双阶段方法,主要有 R-CNN<sup>[13-15]</sup> (region-based convolutional neural networks) 系列。Qi 等<sup>[16]</sup>将 SSD 网络中不同层级的特征图融合,引入自注意力机制和空间稀疏卷积以提高小目标检测精度,但是在处理高分辨率图像时会增加模型复杂度。Radovic 等<sup>[17]</sup>在“YOLO”目标检测框架基础上构建目标检测算法并部署于无人机进行检测,准确率高达 97.5%,大大提升了无人机的自主飞行和决策水平。Tijtgat 等<sup>[18]</sup>将 YOLOv2 算法部署在 NVIDIA Jetson TX2 平台并进行目标检测,与传统 ACF 检测方法相比在准确性和实时性方面有巨大优势。Santos 等<sup>[19]</sup>创新地将 RetinaNet 算法搭载到无人机平台并结合高分辨率 RGB 图像进行目标检测,该算法的 focal loss 函数很好地解决了类型不平衡的问题,但是对小目标检测精度不佳、超参数难以确定且内存开销大。Zhou 等<sup>[20]</sup>使用 MobilenetV2 替换原 YOLOv3 的主干网络,并在其末端添加 SPP 单元,使用 GIoU 代替传统的均方误差 (MSE) 损失函数,实现了精度和速度的平衡。赵耘彻等<sup>[21]</sup>将 YOLOv4 网络的主干替换为 MobilenetV3,降低模型复杂度,提高检测速度,同时使用 K-means++聚类算法重新设计锚框加快网络收敛速度。石欣等<sup>[22]</sup>将浅层特征加入特征金字塔融合过程,提出基于引力的自适应特征融合方式,减少小目标特征信息损失,并采取生成对抗网络生成小目标数量增强训练

效果。杨辉羽等<sup>[23]</sup>将改进的单聚合模块嵌入 YOLOv5 主干网络中,增加小目标检测层并优化原来的 CIOU 损失函数。吴兴辉等<sup>[24]</sup>在 YOLOv7 算法的基础上替换原来的激活函数和损失函数,同时引入注意力机制,提出 YOLO-WA,提升了检测精度但是模型参数量过大。Zhai 等<sup>[25]</sup>在 YOLOv8 算法检测头部分增加了一个高分辨率检测头并在颈部特征融合网络引入全局注意力机制 (global attention mechanism, GAM),精确度 (precision, P)、召回率 (recall, R) 和平均精度 (AP) 方面分别提高了 11.9%、15.2% 和 9%,但是对小目标检测精度欠佳且模型较大。Mei 等<sup>[26]</sup>引入多头自注意力机制 (multi-head self attention, MHSA) 增强模型捕获特征的能力,用部分卷积 (partial convolution, PConv) 和分组可分离卷积 (group-wise separable convolution, GSConv) 改进 C2f 模块降低模型复杂度。Madec 等<sup>[27]</sup>将 Faster-RCNN 算法搭载到无人机平台用以麦穗计数,在光线充足和采集图像分辨率足够高的情况下取得了不错的效果。张超等<sup>[28]</sup>使用改进 Mask-RCNN 算法进行风机叶片缺陷检测,算法先进行特征提取,然后利用区域建议网络 (region proposal network, RPN) 生成感兴趣区域 (region of interest, ROI),再通过 ROIAlign 层进一步处理,最后进行边界框、掩膜和类型的预测,检测精度高但模型参数量大、实时性差。

上述研究在各领域探究了无人机和目标检测算法结合应用的可行性和发展潜力。由于风机叶片航拍图片往往存在背景复杂且干扰因素多,叶片损伤缺陷形态各异、大小不一等问题,导致在进行特征提取时存在较大困难,特别是小目标缺陷的特征提取与检测,此外由于航拍设备能力有限,一些参数量大的模型在部署和检测时无法同时满足实时性和准确性的要求。针对上述存在的问题,本文以 YOLOv8n 算法为基础,引入逆残差模块并加以改进提出额外逆残差 (extra inverted residual block, Extra-IB) 模块,在提升模型精度的同时减少模型参数;提出自适应高效多尺度空间金字塔 (adaptive efficient multi-scale feature pyramid, AEMFP) 模块替换原算法的快速空间池化金字塔模块 (spatial pyramid pooling-fast, SPPF),并融合高效局部注意力机制 (efficient local attention, ELA),改进后的模型平均精度均值 (mAP) 提升 5.1%,小目标 mAP 提升 14.6%,模型参数量减少 22.1%,满足准确性和轻量化的需求。

## 1 YOLOv8 算法简介

YOLO 由 Redmon 等<sup>[5]</sup>于 2015 年首次提出,是一种单阶段目标检测算法。在后续的 YOLO 系列算法中,YOLOv8 算法因其在速度和精度上的良好平衡,并且凭借优秀的泛化性和可拓展性,使得在目标检测、分类、实

例分割、关键点检测等计算机视觉领域表现优秀。YOLOv8 算法是在 YOLOv5 算法基础上改进提出的, YOLOv8 算法主要是由 Input、Backbone、Neck 和 Head 4 部分组成。Input 部分是网络的输入层, 用于接收输入的图像和数据预处理。Backbone 部分主要由 Conv、C2f 和 SPPF 模块组成, C2f 模块的设计参考了 YOLOv7<sup>[9]</sup>算法中高效聚合网络的设计思路, 这种跨层多分支结构能够在实现模型轻量化的同时获取多尺度多层次的梯度流特征信息。SPPF 模块延续了 YOLOv5 中的设计, 通过 3 个  $5 \times 5$  的最大池化操作串行设计, 获取感受野大小为 5、9、13 的多尺度信息。Neck 颈部特征融合部分沿用了 YOLOv5 中的路径聚合网络 (path aggregation network, PAN) 和特征金字塔网络 (feature pyramid network, FPN), 加强了模型特征融合能力, 提高了鲁棒性。Head 检测头部分换成了解耦头, 通过将分类头和检测头分离使模型

能够更好地学习到待检测目标的位置与类别信息。同时采用 Anchor-Free 结构, 这种设计无需提前设定检测框的尺寸, 减少了冗余预测框的数量, 提高了检测速度。此外 YOLOv8 还提出了新的损失函数和训练策略。通过引入了 Task Aligned Assigner 动态正样本分配策略实现了正样本根据任务进行动态分配, 配合 Distribution Focal Loss 损失函数, 提高了模型在困难样本上的识别能力。此外还引入 mosiac 等全新的训练技巧, 通过在最后几次训练迭代中关闭 mosiac 增强, 进一步优化了模型的特征学习能力, 从而显著提高了模型的识别精度。YOLOv8 算法根据不同的应用场景设计有不同尺寸和大小的网络, 考虑到风机桨叶表面缺陷识别时复杂的背景和轻量化部署的需要, 本文将以 YOLOv8n 为基础进行改进, 算法结构如图 1 所示。

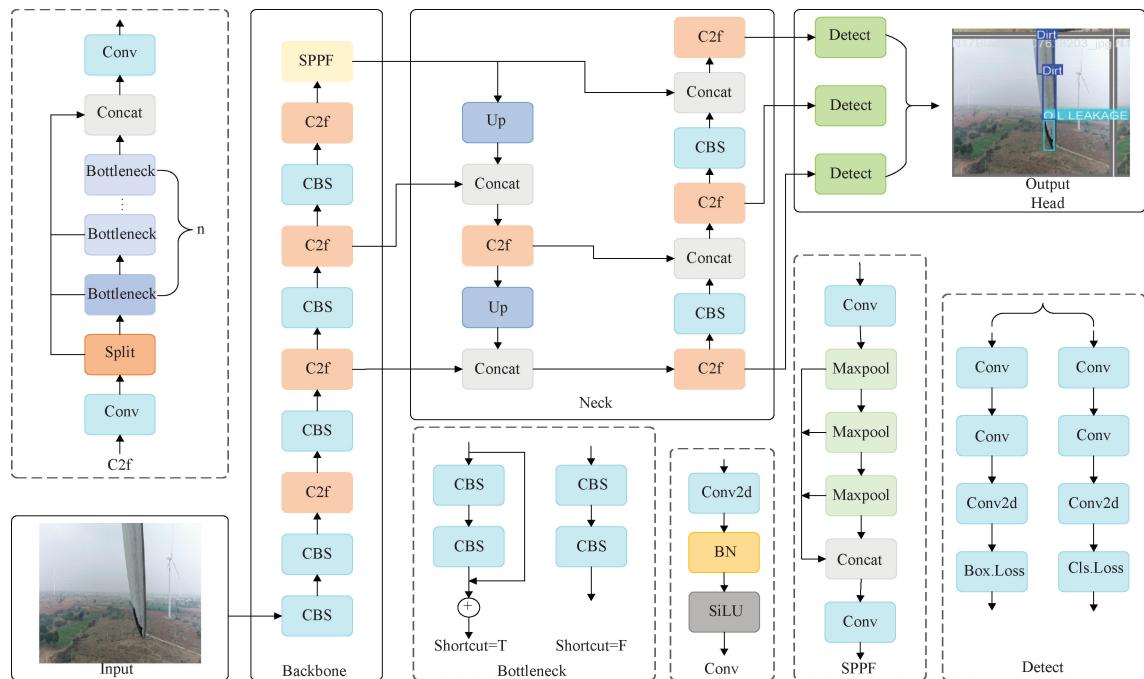


图 1 YOLOv8n 网络结构

Fig. 1 YOLOv8n network structure

## 2 改进的 YOLOv8n 目标检测算法

### 2.1 C2f-Extra-IB 模块

为将算法部署到计算资源有限的设备上, 需要对算法模型进行轻量化优化。受 MobilenetV2<sup>[29]</sup>算法的启发, 本文引入了 MobilenetV2 算法的核心模块逆残差模块, 在原本基础上进一步优化设计了 Extra-IB 模块, 并设计了全新的 C2f-Extra-IB 模块。残差结构通过引入残差连接的设计, 直接将输入特征与卷积层的输出相加, 形成一种

特征叠加的方式, 从而缓解深层网络中常见的梯度消失和退化问题。逆残差结构首先通过  $1 \times 1$  的点卷积进行升维, 使模型够提取到更多的特征信息, 然后通过  $3 \times 3$  的深度可分离卷积分别处理每个单独通道, 减少计算成本的同时保留了更多空间信息, 最后再次通过  $1 \times 1$  的点卷积进行降维, 形成倒置的瓶颈结构, 与传统的残差结构完全相反, 此外逆残差结构采用了线性单元 ReLU6 而不是非线性激活函数以减少低维空间信息损失。残差结构和逆残差结构如图 2、3 所示。

逆残差结构在进行特征提取时, 首先进行  $1 \times 1$  点卷

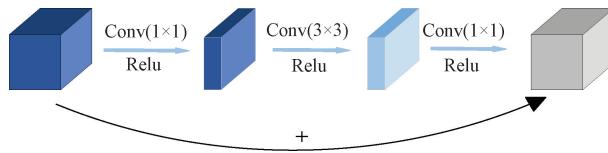


图 2 残差结构

Fig. 2 Residual structure

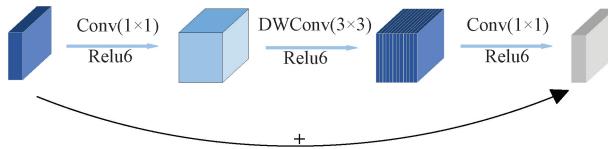


图 3 逆残差结构

Fig. 3 Inverse residual structure

积进行通道维度扩充,这一过程不涉及空间维度信息的交互,容易忽略丰富的低级空间特征(如纹理、边缘),导致空间信息的部分丢失从而影响后续的特征提取与融合效果,并且容易产生冗余通道信息影响特征提取效果。为解决这一问题,本文提出 Extra-IB 模块,在原逆残差结构前置一个深度可分离卷积块,深度可分离卷积对每个通道进行单独空间卷积,捕捉局部空间特征加强通道之间的信息交互,这样在点卷积之前,已经对空间信息进行了部分编码,减轻了由于直接进行点卷积造成的信息损失问题,并且使用高效通道注意力机制<sup>[30]</sup> (efficient channel attention, ECA) 处理升维后的特征图,通过学习通道之间的依赖关系,自适应调整不同通道对于最终结果的贡献度,并且 ECA 注意力机制通过直接操作局部通道而不是引入额外的全连接层,避免了高维特征的冗余计算,确保了算法的轻量性。这种改进增强了通道之间的建模能力,使得后续的点卷积能够更专注于通道维度的特征融合,而不是单独处理空间和通道信息,改善特征融合的效果。Extra-IB 模块结构如图 4 所示。

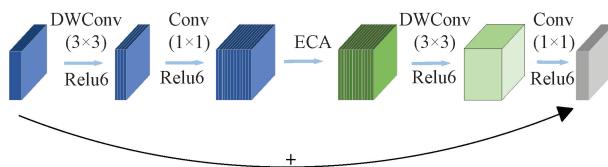


图 4 Extra-IB 模块

Fig. 4 Extra-IB module

## 2.2 AEMFP 模块

YOLOv8 的特征提取主干末端沿用了 YOLOv5 的 SPPF 模块,SPPF 先通过  $1 \times 1$  点卷积将输入降维到中间层通道数,再依次进行 3 次池化核为  $5 \times 5$  的最大池化操作最后将 3 次池化后的结果拼接,这种结构设计等效于感受野为 5、9、13 的 SPP 操作。但是这样的设计难以适

配不同分辨率的特征图并且连续的最大池化操作会丢失输入特征中的局部细节信息,此外,SPPF 是静态的池化操作,无法根据上下文信息动态自适应地调整特征权重。为了改进 SPPF 模块在特征提取方面的不足,本文参考卷积块注意力模块机制(convolutional block attention module, CBAM)的设计思想,并进一步融合高效多尺度注意力机制(efficient multi-scale attention, EMA),使模型能够更加缺陷的特征并减少冗余特征信息的干扰,同时配合多种池化策略和动态加权机制,提出 AEMFP 模块以优化特征提取能力。EMA 注意力机制结构通过对分组后的特征进行自适应加权,动态地分配特征权重。该机制通过结合全局池化和局部卷积来捕获特征的空间关系,显著增强了特征表示的灵活性和鲁棒性。此外,EMA 机制利用软权重分配的方式有效抑制无关或冗余特征,进一步提高了特征的质量。这种动态加权策略为后续特征融合提供了更加清晰且富有辨识度的特征表达。EMA 注意力机制结构如图 5 所示。

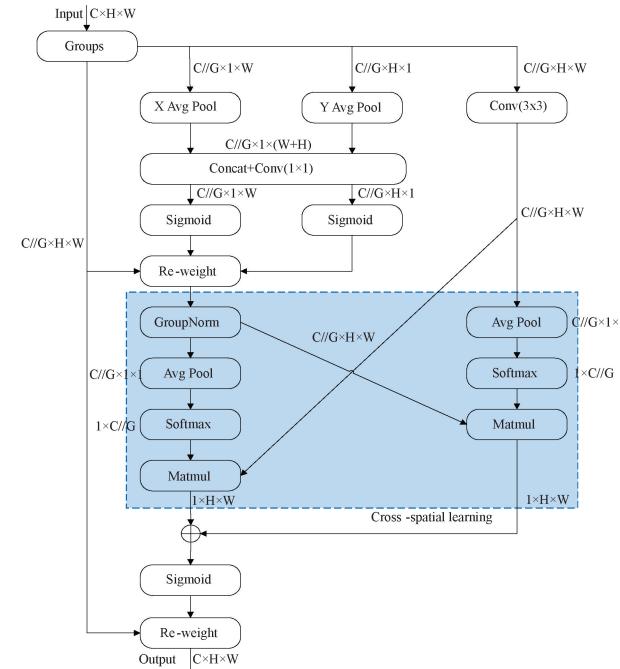


图 5 EMA 注意力机制

Fig. 5 EMA attention mechanism

EMA 主要分为以下 3 个部分。

1) EMA 模块将输入特征  $Input \in \mathbb{R}^{C \times H \times W}$  划分为  $G$  组,通过分组的方式对特征进行并行处理,学习不同子特征中的感兴趣区域。

2) 并行子结构,EMA 模块采用 3 条并行路线进行分组提取特征信息,首先分别在  $X$  和  $Y$  方向上进行自适应池化,提取水平和垂直方向上的全局上下文信息,计算公式如式(1)、(2)所示。

$$z_c^w = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (1)$$

$$z_c^h = \frac{1}{H} \sum_{0 \leq j < H} x_c(h, i) \quad (2)$$

式中:  $z_c^w$  表示宽度为  $w$  的第  $c$  个通道的输出, 仅对垂直方向做处理, 水平方向则不做变化;  $z_c^h$  表示高度为  $h$  的第  $c$  个通道的输出, 仅对水平方向做处理, 垂直方向则不做变化。

然后将  $X$  和  $Y$  方向上的特征图进行融合, 并使用  $1 \times 1$  卷积进行跨通道信息交流, 计算公式如下:

$$f = C([z^h, z^w]) \quad (3)$$

式中:  $f$  表示垂直方向和水平方向融合后的特征图, 其中  $f \in \mathbf{R}^{C/G \times (H+W)}$ ,  $G$  为分组数,  $C$  表示卷积核大小为  $1 \times 1$  的卷积, 用于变换通道数。

然后将融合后的特征图  $f$  分解为两个独立的张量, 水平方向上为  $f^w \in \mathbf{R}^{C/G \times 1 \times W}$ , 垂直方向上为  $f^h \in \mathbf{R}^{C/G \times H \times 1}$ 。然后分别使用 Sigmoid 函数进行激活得到水平和垂直方向上的中间权重。

第 3 个分支为  $3 \times 3$  的普通卷积。

3) 使用全局平均池化处理水平和垂直方向的特征图信息, 然后进行通道维度变换使之与最小分支相匹配。全局平均池化的计算公式如式(4)所示, 输入尺寸为  $H \times W \times C$ , 输出尺寸为  $1 \times 1 \times C$ 。

$$z_c = \frac{1}{H \times W} \sum_{j=1}^W \sum_{i=1}^H x_c(i, j) \quad (4)$$

式中:  $z_c$  表示第  $c$  个通道相关的输出;  $H$  和  $W$  分别表示图像的高度和宽度;  $x_c(i, j)$  表示高度是  $i$ 、宽度为  $j$  的第  $c$  个通道的输入。

水平和垂直方向上进行平均池化操作后采用 Softmax 回归, 再将回归结果与并行子结构的输出相乘得到空间注意力分布图, 对  $3 \times 3$  标准卷积的分支采取同样的操作。最后将两个分支生成的注意力权重信息进行聚合并使用 sigmoid 函数进行激活。

受 CBAM 注意力机制的启发, 将经过 EMA 模块处理后的特征图输入到最大-平均池化并行网络中, 先经过  $1 \times 1$  卷积进行降维, 减少冗余信息对于后续池化操作的影响提升权重分配的稳定性同时可以减少计算开销, 然后将分别输入池化核尺寸为  $5 \times 5$  的最大池化和平均池化分支, 最大池化分支保留了目标区域的显著性特征, 同时抑制背景噪声, 平均池化分支提取全局信息, 使模型也可以兼顾背景中微弱的特征信息, 再将池化后的结果拼接并压缩通道维度, 经过这样的操作使特征图既保留的显著细节特征同时也包含全局上下文特征, 最后配合 Sigmoid 激活函数自适应生成权重, 达到动态调整输出特征图的目的。

AEMFP 模块通过引入 EMA 注意力机制, 3 分支的

结构设计使模块可以有效提取水平和垂直方向上的全局信息,  $3 \times 3$  普通卷积分支与另外两个分支形成互补, 提升了模块的特征提取能力, 同时采用组归一化能够减少批次大小对于检测精度的影响, 增强算法的鲁棒性, 最后经过并行池化结构, 同时捕获细节和显著特征, 特征信息更加完整, 通过不同分支对不同尺度特征的感知, 提升了模块对尺度变化的适应能力, 为后续的特征融合提供更具判别性的特征, 并且经过自适应加权, 使得网络可以根据上下文信息动态调整全局和局部特征信息的权重, 提高了模型的泛化能力。AEMFP 模块结构如图 6 所示。

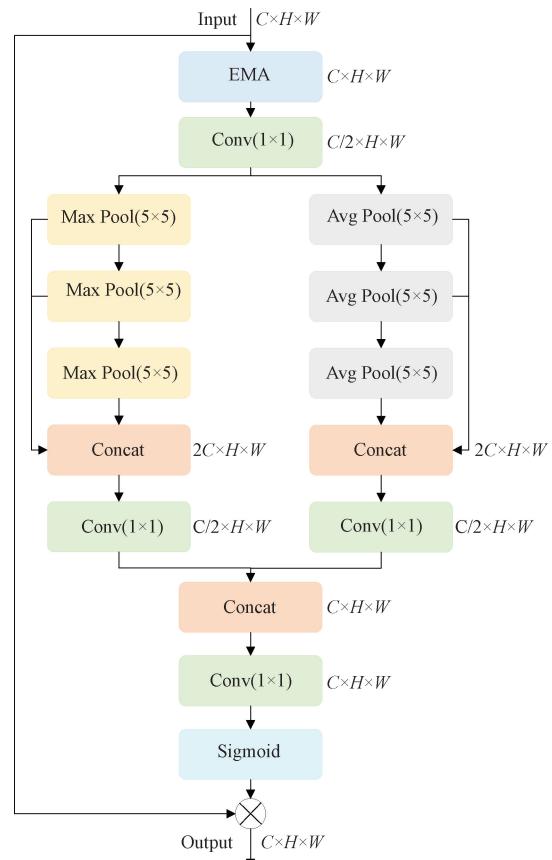


图 6 AEMFP 模块

Fig. 6 AEMFP module

## 2.3 ELA 注意力机制

为减少冗余特征信息干扰, 解决浅层和深层特征融合冲突, 平衡效率与计算成本和提高算法对小目标的检测精度。本文在颈部融合网络末端引入 ELA 注意力机制。ELA 是 Wei 等<sup>[31]</sup>于 2024 年提出, 他们分析了坐标注意力机制(coordinate attention, CA)的不足, CA 注意力机制在水平和垂直两个方向进行平均池化提取注意力权重, 然后进行  $1 \times 1$  的二维卷积降低通道维度, 这样的设计只关注水平和垂直方向上特征的空间关系, 忽略了通道值间的信息交互并且导致水平和垂直方向上的特征图信

息发生多次的提取和融合,造成局部感受野与全局信息的冲突,对小目标的检测精度欠佳。此外平均池化、二维卷积和拼接操作的存在会带来额外的计算开销,不利于模型轻量化。ELA 注意力机制在水平和垂直方向上提取注意力权重后使用一维分组卷积进行特征提取,相较于二维卷积,一维分组卷积更适合处理序列信息,保留更多的边缘特征信息,提高局部信息的感知能力,同时可以根据检测任务的精度要求调整卷积组数来调节模块的计算开销。

CA 采取批归一化方式,这种方式严重依赖批次大小,训练时计算批次内的均值和方差无法代表整个数据集从而导致累积误差的增多,并且这种计算方式与推理时使用移动平均的均值和方差不一致会使模型在部署时存在系统性偏差,所以在遇到不同的检测场景数据分布发生变化的情况时,会导致模型检测精度差、泛化能力弱。而 ELA 采用组归一化,这种归一化方式基于通道分组而不受训练批次内其他样本影响,在小批次的模型训练中表现依然稳定。组归一化只在每个组内统计信息,而不是使用全局方差和均值,表现更稳定,提高了算法鲁棒性,适合完成大型、数据分布更多样的检测任务。风机桨叶表面缺陷检测任务背景复杂,天气变化不规律,采集的图片质量不稳定,故本文采用 ELA 注意力机制改进本文模型。ELA 注意力机制结构如图 7 所示。

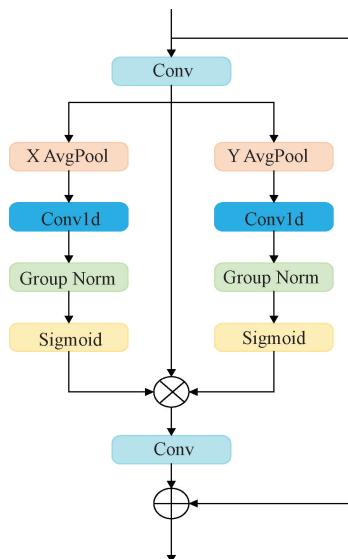


图 7 ELA 注意力机制

Fig. 7 ELA attention mechanism

对于输入特征图  $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$ , ELA 首先分别在水平和垂直方向上使用条形池化来获取特征向量,得到大小分别为  $C \times h \times 1$  和  $C \times 1 \times h$  的两个输出特征图,分别用  $z_c^h$  和  $z_c^w$  表示,计算公式如式(5)、(6)所示。

$$z_c^h(h) = \frac{1}{H} \sum_{0 \leq i < H} x_c(h, i) \quad (5)$$

$$z_c^w(w) = \frac{1}{W} \sum_{0 \leq j < W} x_c(j, w) \quad (6)$$

式中:  $z_c^h$  为水平方向上第  $c$  个通道平均池化结果;  $z_c^w$  为垂直方向上第  $c$  个通道平均池化结果;  $x_c$  为输入。

然后使用一维卷积对  $z_c^h$  和  $z_c^w$  处理,并对结果进行组归一化和 Sigmoid 激活函数后处理,得到两个方向上的位置注意力预测图,计算公式如式(7)、(8)所示。

$$y_h = \sigma(G_n(F_h(z_c^h))) \quad (7)$$

$$y_w = \sigma(G_n(F_w(z_c^w))) \quad (8)$$

最后将原始输入特征图与水平和垂直两个方向上的结果相乘得到最终的输出,计算公式如式(9)所示。

$$Y = x_c \times y_h \times y_w \quad (9)$$

基于上述改进措施,提出改进的 YOLOv8 风机桨叶缺陷检测算法,整体结构如图 8 所示。

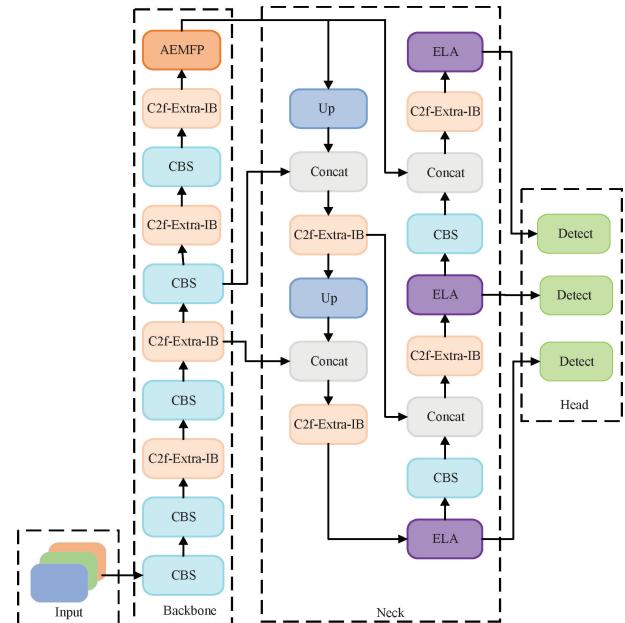


图 8 改进 YOLOv8 算法结构

Fig. 8 Improved YOLOv8 algorithm structure

### 3 实验结果与分析

#### 3.1 实验环境与参数设定

本文实验环境如表 1 所示,主要训练参数如下:训练轮次为 300,训练耐心值为 50 代,训练批次大小为 16,线程数为 8,除 AEMFP 有效性验证实验外其余实验输入图片尺寸均为  $640 \times 640$ ,优化器选用随机梯度优化器 SGD,采用线性学习率调整策略,初始学习率为 0.01,最终学习率为 0.000 1,动量为 0.937,权重衰减系数为 0.000 5,热身训练 3 次。

表 1 实验环境配置

Table 1 Experimental environment configuration

名称	参数
操作系统	Windows 11 64 位
GPU	Nvidia GTX 4090D 24 GB
CPU	Intel i9 14900K
Python	3. 10
Pytorch	1. 9
Cuda	12. 1

### 3.2 数据集

为验证本文所提改进策略的有效性,在 Roboflow 网站收集多个不同公开风机桨叶缺陷数据集的部分图片并采用随机翻转、旋转、亮度调整等数据增强技术,一共 5 065 张图片,涵盖尖端漏油、叶片砂眼和叶片表面污垢 3 种常见缺陷。按照 7:2:1 的比例划分训练集、验证集和测试集。最终训练集 3 545 张图片,验证集 1 013 张,测试集 507 张图片。数据集中包含的 3 种典型缺陷如图 9 所示。



图 9 典型风机桨叶表面缺陷

Fig. 9 Typical turbine blade surface defect

### 3.3 评价指标

本文是多目标检测任务,采用精确率、召回率、mAP 来评估模型的检测精准度和漏检率。采用参数量、浮点运算数(GFLOPs)和模型体积大小来评估模型的计算复杂度和部署难度。使用帧率验证模型检测推理的速度和实时性。其中精确度、召回率和 mAP 计算公式分别如式(10)~(12)所示。

$$P = \frac{T_p}{T_p + F_p} \quad (10)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (11)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \times 100\% \quad (12)$$

式中:  $T_p$  为真正例,表示模型预测为正的正样本的数量;  $F_p$  为假正例,表示模型预测为正的负样本的数量;  $F_N$  为假负例,表示模型预测为负的正样本的数量;  $AP$  为 P-R 曲线与坐标轴围成图形的面积;  $mAP$  为所检测 3 种缺陷  $AP$  的平均值。

### 3.4 消融实验

为评估本文所提改进策略对于模型性能提升的效果,以 YOLOv8n 为基线模型验证了 7 组不同改进策略组合的效果。第 1 组为 YOLOv8n 基线模型,第 2、3、4 组为分别单独加入 C2f-ExtraIB、AEMFP 和 ELA 后的实验结果。第 5、6、7 组为 3 种改进策略两两混合的实验结果,第 8 组(本文算法)为同时加上 3 种改进策略后的实验结果,在保持实验环境和训练器参数设定一致的情况下完成了实验,结果如表 2 所示。

表 2 消融实验

Table 2 Ablation experiment

序号	YOLOv8n	C2f-Extra-IB	ELA	AEMFP	P/%	R/%	mAP/%	参数量/( $\times 10^6$ )	计算量/GFLOPS	模型体积/MB	帧率/fps
1	√				0.806	0.697	0.766	2.69	6.90	5.50	64.25
2	√	√			0.79	0.738	0.791	1.86	4.90	3.94	63.46
3	√		√		0.817	0.722	0.791	2.77	7.10	5.51	63.23
4	√			√	0.774	0.745	0.799	2.83	7.30	5.80	63.53
5	√			√	0.864	0.71	0.795	2.92	7.40	5.96	53.9
6	√	√	√		0.788	0.724	0.784	1.95	5.1	4.12	52.9
7	√	√		√	0.808	0.706	0.788	2.00	5.2	4.23	52.37
8	√	√	√	√	0.817	0.756	0.817	2.09	5.40	4.41	45.57

实验结果表明,只添加 C2f-Extra-IB 模型的精确率有小幅降低,但是参数量和浮点运算数都有显著减少同时召回率和平均精度分别提升 4.1% 和 2.5%,说明 C2f-Extra-IB 模块不仅可以改善模型性能还可以实现模型轻量化;只添加 ELA 注意力机制以增加少量额外参数量和浮点运算数的代价带来精确率、召回率和平均精度的改

善,与 YOLOv8n 相比分别提升 1.1%、2.5%、2.5%。只添加 AEMFP 模块实验结果相较于基线模型,精确率降低 3.2%,召回率提高 4.8%,平均精度提高 3.3%,参数量增加 5.2%,浮点运算数增加 5.7%,说明 AEMFP 模块可以增强算法对于多类型目标检测的精准度和降低误检漏检率。同时添加 3 种改进策略模型精确率、召回率和平均

精度分别比基线模型提高 1.1%、4.3%、5.1%，参数量和浮点运算数分别降低 22.3% 和 21.7%，模型体积减少 19.8%，帧率减少 18.68 fps，说明本文所提改进策略可以同时提高检测精度和实现轻量化。

### 3.5 对比实验

#### 1) C2f-Extra-IB 模块替换位置对比实验

使用 C2f-Extra-IB 替换原 YOLOv8n 算法中不同位置的 C2f 模块，以探究最优的替换方案为后续其他改进策略的融合提供依据。实验以 YOLOv8 为基线模型，前两组为分别单独替换 Backbone 和 Neck 中的所有 C2f 模块，第 3 组为同时替换 Backbone 和 Neck 中的所有 C2f 模块，最终实验结果如表 3 所示。

表 3 C2f-Extra-IB 替换位置对比实验

Table 3 C2f-Extra-IB replacement position comparison experiment

模型	P/%	R/%	mAP/%	参数量/ ( $\times 10^6$ )	计算量/ GFLOPS
YOLOv8n	<b>0.806</b>	0.697	0.766	2.69	6.9
Backbone	0.788	0.719	0.788	2.26	5.7
Neck	0.785	0.705	0.778	2.28	6.1
Backbone+Neck	0.790	<b>0.738</b>	<b>0.791</b>	<b>1.86</b>	<b>4.9</b>

由实验结果可以看出，当替换 Backbone 和 Neck 中所有 C2f 时，该模型在 3 种替换方案中表现最优，召回率、平均精度 mAP@0.5 较基线模型分别提升 4.1% 和 2.5%，参数量和浮点运算数分别下降约 30.8% 和 28.9%，精确率虽然下降 1.6%，但是在 3 种替换方案中受影响最小。因此后续改进模型均采取 Backbone 和 Neck 同时替换 C2f-Extra-IB 的方案。实验结果表明 C2f-Extra-IB 模块能够有效减少算法参数量和浮点运算数的同时提高模型检测精度。

#### 2) 不同注意力机制对比实验

为验证 ELA 注意力机制在本文对提高算法检测准确度有效性和其轻量化优势，分别使用 CBAM、SE、SA(shuffle attention)、ECA 和 CA 5 种常见的注意力机制替换本文所提算法中的 ELA 注意力机制进行对比实验，实验结果如表 4 所示。

实验结果表明，每一种注意力机制加入后模型参数量均有不同程度的增加也带来了模型检测精度的提升。在这 6 种注意力机制中，SE 注意力机制凭借连续的挤压和激发，即全局池化和全连接操作，提高了模型对于显著目标的感知能力，准确率表现最佳，但是由于全局池化操作会忽略部分细节信息，因此导致召回率偏低，误检和漏检发生更频繁；ECA 采取一维深度可分离卷积，不使用全连接层，所以参数增加最少，但是提取全局信息能力弱；CBAM 将空间注意力与通道注意力结合，计算复杂，在面对复杂背景时易漏检；SA 基于通道数进行分组卷积

然后进行混洗，参数影响小，但是检测低分辨率图片时因为通道数有限，信息交互不充分，因此效果不佳；CA 使用二维卷积和拼接操作对参数量影响较大。在 Neck 末端引入 ELA 注意力机制，本文算法精确率比添加 SE 注意力时低 4%，但是优于其他 4 种注意力机制；模型召回率在 6 种注意力机制中最高，意味着本文算法在实际应用时漏检和误检发生频率更低；平均精度在 6 种注意力机制中表现最好，分别比使用 SE、ECA、CBAM、SA 注意力机制高 2.7%、2%、3.6%、0.9%，说明本文算法在检测头前选择 ELA 注意力机制进行多类型目标检测时有更高的准确率。

表 4 不同注意力机制对比实验

Table 4 Comparative experiment of different attention mechanisms

模型	P/%	R/%	mAP/%	参数量/ ( $\times 10^6$ )	计算量/ GFLOPS
YOLOv8n	0.806	0.697	0.766	2.69	6.9
C2f-Extra-IB+ AEMFP	0.808	0.706	0.788	2.00	5.2
+SE	<b>0.857</b>	0.724	0.790	2.01	5.2
+ECA	0.785	0.741	0.797	<b>2.00</b>	5.2
+CBAM	0.809	0.703	0.781	2.09	5.3
+SA	0.777	0.730	0.808	2.00	5.2
+CA	0.795	0.743	0.810	2.02	5.6
+ELA	0.817	<b>0.756</b>	<b>0.817</b>	2.01	5.4

3) AEMFP 模块有效性实验为验证 AEMFP 模块的有效性和本文所提算法检测不同尺寸图片时的鲁棒性，选取了  $416 \times 416$ 、 $640 \times 640$ 、 $800 \times 800$  三种常见分辨率的图片模拟不同规格图像输入进行实验，主要对比了基线模型 YOLOv8n、YOLOv8n+AEMFP 模型、以及本文算法 3 种模型的精确率、召回率和 mAP，实验数据如表 5 所示。

表 5 AEMFP 模块有效性实验

Table 5 Experiment on the effectiveness of AEMFP module

图片尺寸	模型	P/%	R/%	mAP/%
	YOLOv8n	78.7	<b>73.9</b>	78
416×416	+AEMFP	76.5	72.2	78.5
	本文	<b>81.6</b>	71.4	<b>79.6</b>
	YOLOv8n	80.6	69.7	76.6
640×640	+AEMFP	77.4	74.5	79.9
	本文	<b>81.7</b>	<b>75.6</b>	<b>81.7</b>
	YOLOv8n	78.1	<b>74.1</b>	79.1
800×800	+AEMFP	<b>82</b>	70.8	79.2
	本文	80.3	72.5	<b>80.5</b>

由表 5 可知，AEMFP 模块在检测不同分辨率图片时对不同指标均有一定影响。 $416 \times 416$  低分辨率输入的情况下，添加 AEMFP 模块后模型 mAP 提升了 0.5%，但是

分辨率低会减少目标的特征信息从而造成准确率和召回率下降;常见的  $640 \times 640$  尺寸输入下,AEMFP 模块显著提高了召回率和 mAP,分别提高了 4.8% 和 3.3%; $800 \times 800$  的高分辨率输入下,图片有了更丰富的特征信息,但是也引入更多背景噪声,因此 AEMFP 模块提高了模型的精确率和 mAP 的同时召回率也有所降低。

相比仅添加 AEMFP 模块,本文算法在各个分辨率输入下均获得了最优的 mAP。在  $640 \times 640$  和  $800 \times 800$  的输入下,本文算法 mAP 分别达到了 81.7% 和 80.5%,即使是  $416 \times 416$  的低分辨率输入下也能达到 79.6%。以上实验说明了 AEMFP 模块的有效性,同时体现了本文所提算法在不同分辨率输入情况下的鲁棒性。

#### 4) 常见目标检测算法对比

为了进一步评估本文所提算法在风机桨叶缺陷识别任务上的性能表现,在相同实验环境和训练器参数设定下使用多种现有主流目标检测算法进行对比实验。选择包括同系列的 YOLOv8、YOLOv10 和 YOLOv11 的不同规格算法,同时将 YOLOv8n 的特征提取骨干网络替换为 MobilenetV2、MobilenetV3、ShuffleNetV2 和 EfficientNetV2 等轻量级网络,通过对比各种缺陷的检测精度、mAP、参数量、浮点运算数、模型体积大小和帧率等 6 种指标验证本文所提算法在风机桨叶表面缺陷检测任务中的优势,最终各项实验结果如表 6 所示。

表 6 常见目标检测算法对比实验

Table 6 Comparison experiment of common target detection algorithms

序号	模型	P/%			mAP/%	参数量/ ( $\times 10^6$ )	计算量/ GFLOPS	模型体积/MB	帧率/fps
		表面油污	尖端漏油	砂眼					
1	YOLOv8n	71.1	88.8	75.3	76.6	2.69	6.9	5.50	<b>64.25</b>
2	YOLOv8s	69.5	88.6	75.9	78	9.84	23.6	19.50	48.83
3	YOLOv10n	66.8	86.3	78.1	77.1	2.7	8.4	5.64	58.07
4	YOLOv10s	69.8	87.7	79	78.8	8.06	24.8	16.16	56.64
5	YOLOv11n	<b>73.4</b>	89.3	76.3	79.7	2.59	6.4	5.36	60.24
6	YOLOv11s	72.5	88.8	80.9	80.7	9.41	21.3	18.75	60.06
7	YOLO_shufflenetV2	69.4	86.4	75	76.9	<b>1.7</b>	<b>1.1</b>	<b>3.56</b>	52.28
8	YOLO_mobilenetV2	68.4	87.6	77.5	77.8	3	7.6	6.15	59.41
9	YOLO_mobilenetV3	72	89	79.3	80.5	2.69	6	5.58	46.77
10	YOLO_efficientnetV2	70.1	87.2	75.8	77.7	2.18	2.4	4.59	46.06
11	本文	72.2	<b>89.7</b>	<b>84.6</b>	<b>81.7</b>	2.01	5.4	4.41	45.57

通过表 6 可以看出,改进后的模型能有效提高不同缺陷的检测精度,除表面油污这类大小不规则缺陷,另外两种小目标缺陷的检测精度均达到最佳,尖端漏油和砂眼的检测精度相较于基准模型分别提升 0.9%、9.3%。同时平均精度也有大幅提升,与 YOLOv8n、YOLOv8s、YOLOv10n、YOLOv10s、YOLOv11n、YOLOv11s、YOLO\_shufflenetV2、YOLO\_mobilenetV2、YOLO\_mobilenetV3、YOLO\_efficientnetV2 相比分别提升 5.1%、3.7%、4.6%、2.9%、2%、1%、4.8%、3.9%、1.2%、4%。在模型参数量和技术复杂度上,优于 3 种轻量化主干改进模型,YOLO\_shufflenetV2 虽然比本文所提算法更轻量化,但是检测性能上远不如本文算法。检测帧率达到 45.57 fps,模型体积仅比最小的 YOLO\_shufflenetV2 大 0.85 MB。通过以上数据可以看出,本文算法在保证轻量化、易部署的同时仍然能够对风机桨叶缺陷有足够高的检测精度和检测速度,具有一定的工业应用价值。

## 4 结 论

高效的检测算法,提出了一种全新的轻量化无损检测算法。针对检测设备计算资源有限的需求,设计了一种全新的 C2f-Extra-IB 模块,创新地引入深度可分离卷积和逆残差结构,降低了模型参数量;设计了 AEMFP 模块,自适应多尺度地提取目标特征,为后续的特征融合提供高质量特征信息;特征融合网络加入 ELA 注意力机制,大幅提升模型对于小尺寸显著目标的检测精度。实验结果表明改进后的算法与 YOLOv8n 算法相比,精确度提升 1.1%、召回率提升 4.3%、平均精度提升 5.1%,同时参数量减少 22.3%,计算浮点数减少 21.7%。与其他主流目标检测算法相比,本文所提算法在关键指标上也均有不同程度提升,证明本文所提算法能够在保障方便部署、设备适配性的情况下在复杂情况下也能有不错的性能,能够完成复杂背景情况下风机桨叶表面各种缺陷的检测任务,对于风机桨叶日常维护工作有一定的应用价值。未来工作中,将继续优化本文算法,考虑采用剪枝或 TensorRT 等技术降低模型深度提高检测速度,使用不同领域、类型和规模的数据集训练模型,以提高模型对于不同应用场景的检测性能,增强模型的泛化性和鲁棒性。

目前风机桨叶缺陷检测领域缺少轻量化方便部署且

## 参考文献

- [ 1 ] 乐威. 新能源背景下我国风力发电现状和未来发展方向探索 [ J ]. 绿色环保建材, 2020 ( 11 ): 165-166.
- LE W. Explore the present situation and future development direction of wind power generation in our country under the background of new energy [ J ]. Green Building Materials, 2020 ( 11 ): 165-166.
- [ 2 ] 陈丽城, 李寿清. 风机叶片缺陷识别与定位研究 [ J ]. 红水河, 2024, 43(5): 46-50.
- CHEN L CH, LI SH Q. Research on defect identification and location of fan blade [ J ]. Hongshui River, 2024, 43(5): 46-50.
- [ 3 ] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [ J ]. ArXiv preprint arXiv:2004.10934, 2020.
- [ 4 ] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications [ J ]. ArXiv preprint arXiv:2209.02976, 2022.
- [ 5 ] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [ C ]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015: 779-788.
- [ 6 ] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [ J ]. ArXiv preprint arXiv:1612.08242, 2016.
- [ 7 ] REDMON J, FARHADI A. YOLOv3: An incremental improvement [ J ]. ArXiv preprint arXiv: 1804.02767, 2018.
- [ 8 ] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [ J ]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [ 9 ] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [ J ]. ArXiv preprint arXiv:2207.02696, 2022.
- [ 10 ] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [ C ]. European Conference on Computer Vision, 2024: 1-21.
- [ 11 ] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [ J ]. Lecture Notes in Computer Science, 2016, 9905: 21-37.
- [ 12 ] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [ C ]. IEEE International Conference on Computer Cision, 2017: 2980-2988.
- [ 13 ] GIRSHICK R. Fast R-CNN [ J ]. ArXiv preprint arXiv: 1504.08083, 2015.
- [ 14 ] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [ J ]. ArXiv preprint arXiv: 1311.2524, 2014.
- [ 15 ] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [ J ]. ArXiv preprint arXiv:1703.06870, 2017.
- [ 16 ] QI G, YU Z, SONG J. Multi-scale feature fusion and context-enhanced spatial sparse convolution single-shot detector for unmanned aerial vehicle image object detection [ J ]. Applied Sciences, 2025, 15(2): 924.
- [ 17 ] RADOVIC M, ADARKWA O, WANG Q. Object recognition in aerial images using convolutional neural networks [ J ]. Journal of Imaging, 2017, 3(2):21.
- [ 18 ] TIJTGAT N, RANST W V, VOLCKAERT B, et al. Embedded real-time object detection for a UAV warning system [ C ]. IEEE International Conference on Computer Vision Workshop, 2017.
- [ 19 ] SANTOS A A D, JUNIOR J M, ARAUJO M S, et al. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs [ J ]. Sensors, 2019, 19(16): 3595.
- [ 20 ] ZHOU J, TIAN Y, YUAN C, et al. Improved UAV opium poppy detection using an updated YOLOv3 model [ J ]. Sensors, 2019, 19(22):4851.
- [ 21 ] 赵耘彻, 张文胜, 刘世伟. 基于改进 YOLOv4 的无人机航拍目标检测算法 [ J ]. 电子测量技术, 2023, 46(8): 169-175.
- ZHAO Y CH, ZHANG W SH, LIU SH W. UAV aerial target detection algorithm based on improved YOLOv4 [ J ]. Electronic Measurement Technology, 2023, 46(8): 169-175.
- [ 22 ] 石欣, 卢灏, 秦鹏杰, 等. 一种远距离行人小目标检测方法 [ J ]. 仪器仪表学报, 2022, 43(5): 136-146.
- SHI X, LU H, QIN P J, et al. A remote pedestrian small target detection method [ J ], Chinese Journal of Scientific Instrument, 2022, 43(5): 136-146.
- [ 23 ] 杨辉羽, 李海明. 单聚合 YOLO 航拍小目标检测算法 [ J ]. 国外电子测量技术, 2023, 42(4): 131-140.
- YANG H Y, LI H M. Single aggregation YOLO algorithm for small object detection in aerial photography [ J ]. Foreign Electronic Measurement Technology, 2023, 42 ( 4 ): 131-140.
- [ 24 ] 吴兴辉, 何赟泽, 周辉, 等. 改进 YOLO v7 算法下的监控水域环境人员识别研究 [ J ]. 电子测量与仪器学报, 2023, 37(5): 20-27.
- WU X H, HE B Z, ZHOU H, et al. Research on personnel identification of monitoring water environment based on improved YOLO v7 algorithm [ J ]. Journal of

- Electronic Measurement and Instrumentation, 2023, 37(5): 20-27.
- [25] ZHAI X, HUANG Z, LI T, et al. YOLO-drone: An optimized YOLOv8 network for tiny UAV object detection [J]. Electronics, 2023, 12(17):3664.
- [26] MEI J, ZHU W. BGF-YOLOv10: Small object detection algorithm from unmanned aerial vehicle perspective based on improved YOLOv10 [J]. Sensors, 2024, 24(21):6911.
- [27] MADEC S, JIN X, LU H, et al. Ear density estimation from high resolution RGB imagery using deep learning technique [J]. Agricultural and Forest Meteorology, 2019, 264:225-234.
- [28] 张超, 文传博. 基于改进 Mask R-CNN 的风机叶片缺陷检测 [J]. 可再生能源, 2020, 38(9): 1181-1186. ZHAO CH, WEN CH B. wind turbine blade defect detection based on improved Mask R-CNN [J]. Renewable Energy Sources, 2020, 38(9): 1181-1186.
- [29] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018; 4510-4520.
- [30] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [J]. ArXiv preprint arXiv:1910.03151, 2020.
- [31] XU W, WAN Y, ZHAO W N. ELA: Efficient local

attention for deep convolutional neural networks [J]. Journal of Real-Time Image Processing, 2025, 22:140.

### 作者简介



**李大华**, 2001 年于天津理工学院获得学士学位, 2004 年于广西大学获得硕士学位, 现为天津理工大学教授, 主要研究方向为人工智能、机器人与机器视觉、智能微电网。

E-mail: lidah2005@163.com

**Li Dahua** received his B. Sc. degree from Tianjin University of Technology in 2001 and his M. Sc. degree from Guangxi University in 2004. He is now a professor at Tianjin University of Technology. His main research interests include artificial intelligence, robotics and machine vision, and smart microgrids.



**吴超强**(通信作者), 2023 年于青岛理工大学获得学士学位, 现为天津理工大学硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 18090449023@163.com

**Wu Chaoqiang** (Corresponding author) obtained his B. Sc. degree from Qingdao Technological University in 2023. He is now a M. Sc. candidate at Tianjin University of Technology. His main research interests include deep learning and image processing.