

基于 YOLOv8-SPH 的防震锤缺失检测

张吉祥¹ 方东²

(1. 三峡大学电气与新能源学院 宜昌 443002; 2. 三峡大学机械与动力学院 宜昌 443002)

摘要:针对输电线路中防震锤尺寸小,图像背景复杂,防震锤缺失难以检测的问题,提出一种轻量化 YOLOv8-SPH 的防震锤缺失检测模型。通过在 YOLOv8n 网络的颈部引入 160×160 和 320×320 的浅层尺度特征图,并在检测头中融入相应尺度的目标检测模块,提升了特征图之间的上下文信息融合能力,有效扩大了模型的感受野,使得模型能够捕捉到更多防震锤缺失的特征语义信息。还创新性地提出了多尺度高效特征提取模块 (MultiFaster),通过部分卷积、多级特征提取和残差连接机制,在保持防震锤特征检测精度的同时,减少网络的计算量和参数量。此外,在颈部网络中引入动态上采样算子,提高重建特征图的分辨率,提高了该模型对防震锤缺失检测的精度,同时,将原模型解耦式检测头更换为轻量化检测头,降低了模型计算的复杂度并提升检测效率。最后对改进后的网络进行基于幅值的层自适应稀疏化剪枝,进一步减小模型参数及计算量。在针对自制防震锤缺失数据集的测试中,YOLOv8-SPH 表现卓越,其 mAP@0.5 达到了 91.51%,相比原始 YOLOv8n 提高了 6.3%,参数量减少了 80.73%,计算量减少了 48.14%,模型尺寸减少了 62.41%。该模型在计算量和参数量降低的同时,提高了检测精度,充分满足了对输电线路中的防震锤进行高效和准确检测的需求,具有实用性。

关键词:防震锤;深度学习;多尺度特征;YOLOv8n;模型剪枝

中图分类号: TP183; TM75 文献标识码: A 国家标准学科分类代码: 520.2060

Detection of missing shockproof hammers based on YOLOv8-SPH

Zhang Jixiang¹ Fang Dong²

(1. School of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;
2. School of Mechanical and Power Engineering, China Three Gorges University, Yichang 443002, China)

Abstract: To address challenges in detecting missing Shockproof Hammers on transmission lines due to their small size, complex image backgrounds, and subtle presence, this study proposes a lightweight YOLOv8-SPH model for damper absence detection. The model introduces shallow-scale feature maps of 160×160 and 320×320 within the neck of the YOLOv8n network and integrates multi-scale detection modules within the detection head. This enhances contextual information fusion across feature maps, effectively expanding the receptive field, enabling the model to capture richer semantic features related to damper absence. An innovative multi-scale high-efficiency feature extraction module (MultiFaster) is also introduced, utilizing partial convolutions, multi-level feature extraction, and residual connections. This structure maintains detection accuracy for damper features while reducing computational complexity and parameter load. Additionally, a dynamic upsampling operator is incorporated into the neck network to improve feature map resolution, improving the model's accuracy in detecting missing Shockproof Hammers. To further optimize, the original model's decoupled detection head is replaced with a lightweight detection head, reducing computational complexity and boosting detection efficiency. The enhanced network undergoes amplitude-based layer-adaptive sparse pruning, significantly reducing model parameters and computational load. Testing on a custom damper absence dataset demonstrates YOLOv8-SPH exhibited remarkable performance, achieving an mAP@0.5 of 91.51%, which marks a 6.3% improvement over the original YOLOv8n. Additionally, parameter count is reduced by 80.73%, computational load by 48.14%, and model size by 62.41%. The model achieves improved detection accuracy while reducing computational complexity and parameter size, effectively meeting the demands for efficient and precise detection of Shockproof Hammers.

in transmission lines, showcasing significant practical value.

Keywords: shockproof hammers; deep learning; multi-scale features; YOLOv8n; model pruning

0 引言

输电线路作为电力输送的主要通道,其可靠性和稳定性直接关系到电力供应的安全和经济性。然而,输电线路常年暴露在野外环境中,面临着诸如风振、冰雪负载、雷击等多种自然灾害的威胁^[1]。为了确保输电线路的安全运行,各种防护装置应运而生,其中防震锤^[2](Shockproof Hammer)是关键的防护设备之一。防震锤是一种振动减缓装置,主要作用是抑制导线在风力作用下产生的振动,以防止导线因长期振动而导致疲劳断裂^[3]。防震锤的缺失会显著增加输电线路的安全隐患,引发严重的电力事故。因此,及时发现防震锤的缺失具有重要的实际意义。

传统的输电线路巡检主要依赖于人工检查,不仅耗时耗力,而且受限于地形和气候条件,难以实现对大范围输电线路的高效监测。近年来,随着无人机巡检和深度学习的快速发展,基于深度学习的无人机巡检能够实现对图像数据的智能检测,为实现大范围输电线路的智能化检测提供了重要的基础^[4]。

周景等^[5]通过 e-fficientNet 网络提取特征并进行 BiFPN 网络融合,虽然提升了防震锤的检测精度但模型变得更加复杂。梁华刚等^[6]使用 3 个级联检测器并将非极大值抑制法替换为 Soft-NMS 算法,使得防震锤缺陷检测的误检率及漏检率均降低,但仍面临参数量大的问题。苏雨蕾等^[7]通过增加极小目标检测层,引入空间自适应特征融合检测头和可变形卷积,增强了模型的检测能力,降低遮挡重叠对检测的影响,但增大了模型的计算量。卢志博等^[8]通过改进 K-means 算法得到聚类结果,并结合深度可分离卷积。使得网络结构轻量化,并提升了检测效果。徐业东等^[9]采用 Fasternet 重构 YOLOv5s 特征提取网络,并引入 ODCConv 模块,降低了模型的计算量并保持较高的鸟巢检测精度。Dong 等^[10]通过融合 Swin-v2 和特征金字塔来增强现有的 Cascade RCNN 并采用多尺度变换的数据增强策略,提升了输电线路关键部件和缺陷检测的精度。吕游等^[11]引入双通道注意力模块(CBAM),提升了模型检测效果,但增大了模型尺寸。Yuan 等^[12]通过引入注意力机制(SEnet)和小目标检测层,提高了输电线路小目标检测效果。赵振兵等^[13]在主干提取网络中引入了具有自注意力机制的 AFF-Transformer 模块和双向加权特征融合机制,有效缓解了输电线金具在密集状态下的误检、漏检等问题。宋立业等^[14]在 BiFPN 融入小一级尺度的特征层,并在主干网络

引入坐标注意力机制,提高了输电线路元件缺陷检测的精度,但对分布密集的小目标检测效果仍不理想。Guo 等^[15]利用多尺度卷积核,并将不同尺度的信息进行融合,增大了模型的感受野,提高了防震锤缺失检测的精度,但改进后模型的计算量依然显著。

上述文献多采用引入注意力机制和改变网络融合机制的方法改进基准网络模型,取得了较好的效果,但依旧存在参数量大,计算复杂等问题,不能友好的部署到无人机等存储空间有限的边缘设备的问题。针对上述问题,提出了一种 YOLOv8-SPH 防震锤缺失检测模型。在 YOLOv8n 模型基础上,通过在颈部网络中将 C2f 替换为 C2f-MultFaster、增加浅层尺度特征融合和小目标检测头、更换动态上采样算子^[16](dynamic sampling, DySample) 和 Detect Efficient,最后对改进后的网络结构进行基于层自适应幅度的剪枝(layer adaptive magnitude based pruning, LAMP),建立 YOLOv8-SPH 防震锤缺失检测模型,使模型提高防震锤缺失检测精度的同时,变得更具轻量化。

1 原始模型

YOLOv8 是 2023 年 Ultralytics 公司推出的 YOLO 系列目标检测算法,是单阶段目标检测算法的代表,它建立在 YOLO 系列历史版本的基础上,进一步提升了性能和效率,在各种计算机视觉任务中表现出色。

YOLOv8 的结构主要包括 3 个部分,主干网络(backbone)、颈部网络(neck)和检测头(head)。虽然 YOLOv8n 在大、中尺度目标检测方面表现优异,但在防震锤缺失检测场景中,图像中的防震锤通常非常密集,且目标的尺度很小。由于小目标在图像中占据的像素比例较小,包含的特征信息有限,YOLOv8n 在这种情况下的检测能力表现相对不足,并且输电线路环境背景复杂,使得 YOLOv8n 模型容易出现漏检和误检。此外,原有的上采样机制无法完全充分重构不同尺度的目标,导致在重构特征图时小目标容易丢失。YOLOv8n 的复杂度较高,参数量大,尽管有较好的检测性能,但也带来了较大的计算开销和资源需求。

基于上述问题,以提升小目标的检测精度,降低模型的复杂度,对基准网络 YOLOv8n 进行改进。

2 改进 YOLOv8 算法网络结构

为了提升模型检测精度,降低模型参数量、计算量及

模型尺寸,从4个方面对YOLOv8n模型进行改进。改进后YOLOv8-SPH的模型结构如图1所示。YOLOv8n的原模型在检测小且密集的防震锤时精度较低,故在YOLOv8n中Neck原有的3个尺度特征图再增加浅层尺度的特征图分别是:160×160和320×320两尺度,并在Head增加相应尺度的检测头,同时将原有的上采样算子

换为DySample,提高防震锤微小特征重建质量,减少采样过程信息的丢失,从而提高检测的精度。此外将Neck中C2f换为计算量及参数量更小且能够多尺度提取防震锤信息的C2f-MultFaster,并将原有的两分支解耦头换为更为高效且轻量化检测头(detect efficient),降低模型计算的复杂度。

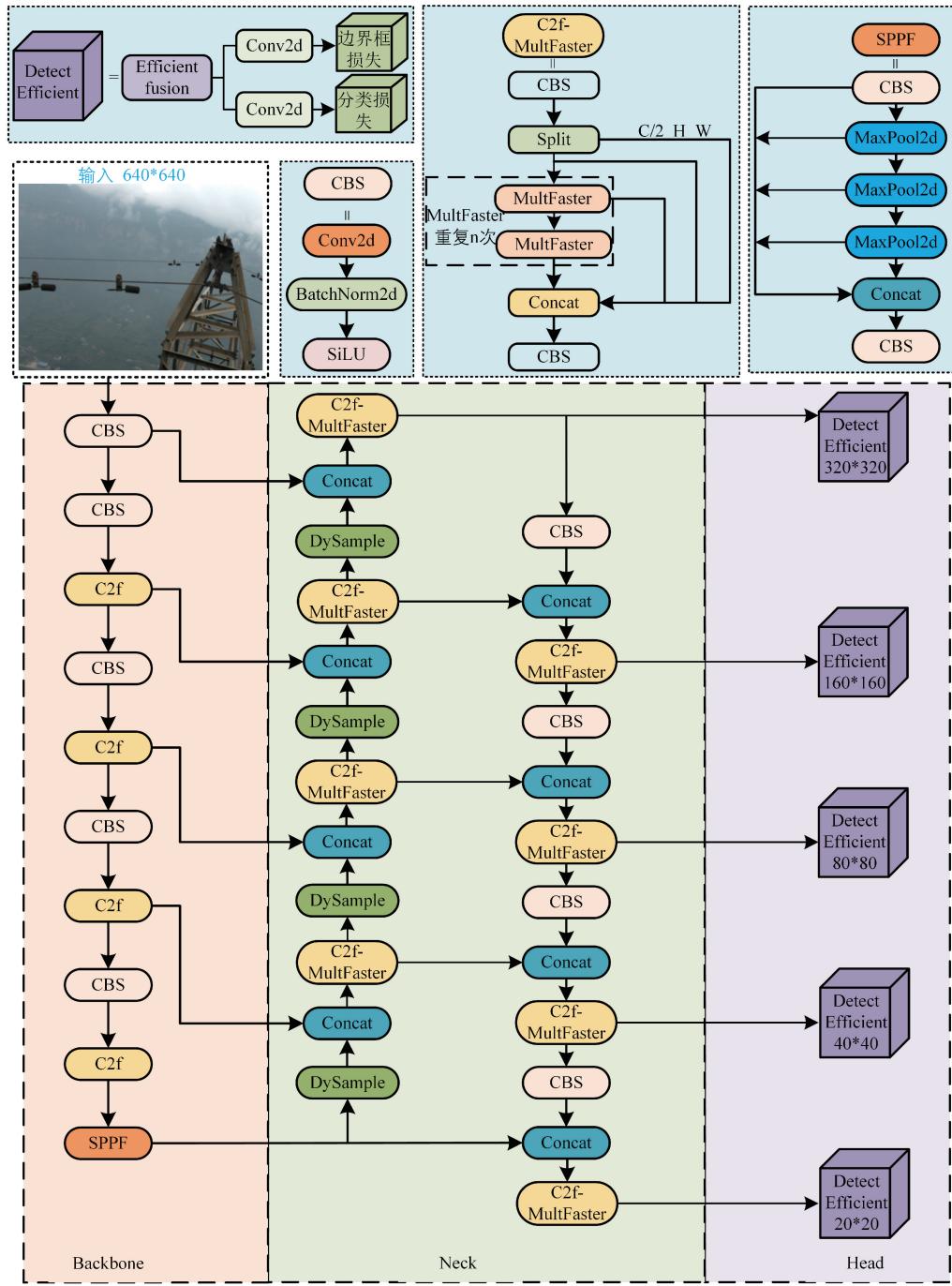


图1 改进后YOLOv8-SPH的模型结构

Fig. 1 Model structure of improved YOLOv8-SPH

2.1 增加浅层尺度特征图和小目标检测头

原有的 YOLOv8n 网络的 Neck 及 Head 中只有 20×20 、 40×40 、 80×80 这 3 个尺度, 其感受野较小, 表达的语义信息少, 这些劣势限制了对于防震锤缺失小目标的特征提取能力。小目标检测通常需要从高分辨率特征图中提取信息, 并在不同尺度上进行充分的特征融合。如果模型多尺度特征利用不充分, 小目标的信息容易在提取过程中丢失, 导致难以捕捉到小目标的细节信息, 从而影响检测精度。而防震锤在输电线路中非常密集, 尺寸很小, 且有遮挡的情况。此外防震锤的图片背景非常复杂, 使得防震锤检测难度进一步增加。

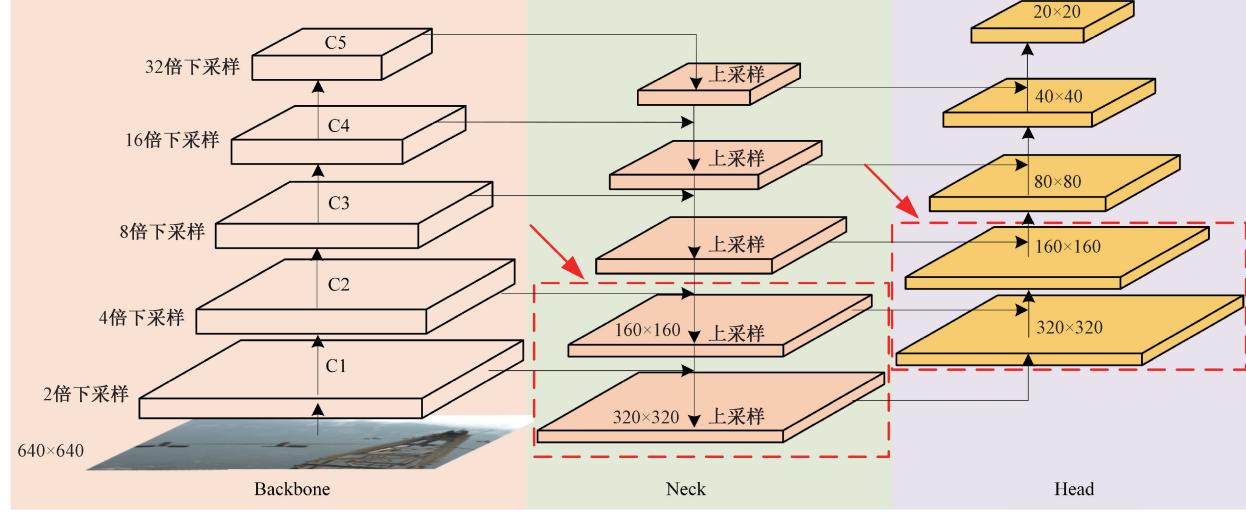


图 2 增加多尺度融合与小目标检测头

Fig. 2 Adding multi-scale fusion and small object detection heads

2.2 改进 YOLOv8n 中 Neck 的上采样算子

在原始 YOLOv8n 网络中, Neck 的上采样算子通常用于将低分辨率的特征图扩展到更高分辨率用于恢复丢失的空间信息, 从而提高模型对细节和局部特征的捕捉能力。而原模型中采用的是最近邻插值法进行上采样, 这种方法会导致上采样后的图像出现明显的锯齿状边缘和模糊的纹理, 无法捕捉到防震锤及缺失的微小特征, 会导致信息的丢失。因此引入 DySample, 使其更好地感知到防震锤缺失的特征, 提高模型对微小细节的捕捉能力。如图 3 所示基于 Pytorch 的动态上采样过程。

x 为输入的特征图, 其形状为 $H \times W \times C$, x' 为采样输出。 x 通过采样点生成器生成动态采样集 S 大小为 $sH \times sW \times 2g$ 。每个采样点由 $2g$ 个坐标(图像的像素坐标)表示, 因此有 $2g$ 维的深度。网格采样(grid sample, G)使用动态采样集对输入 x 的特征值进行插值计算, 从而生成上采样后的特征图。其过程如式(1)所示。

$$x' = \text{grid sample}(x, S) \quad (1)$$

如图 4 所示, DySample 中的采样点生成器。图 4 中

如图 2 所示, 在 YOLOv8n 中 Neck 原有的 3 个尺度特征图再增加浅层尺度特征图分别是 160×160 和 320×320 两尺度, 并在 Head 增加相应的尺度的检测头, 提升感受野及语义表达。多个检测头能够覆盖更广泛的尺度范围, 使得模型可以在更多尺度的特征层上进行融合, 这样能够更好地利用特征图中的细节和上下文信息去捕捉尺寸大小不同的目标物体, 并提高模型检测的精度。在增加浅层尺度特征图和小目标检测头后, 模型具备更强的小目标检测能力, 在应对复杂场景和变化多样的目标时更加稳定和可靠。有效地改善了防震锤检测网络对远处的小目标的误检、漏检的问题。

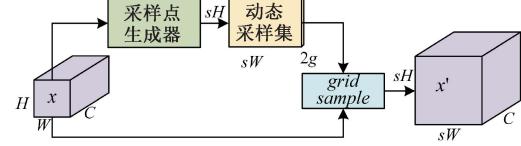


图 3 动态上采样

Fig. 3 Dynamic upsample

S 为偏移量 O 和原始网格采样进行逐元素相加。其中偏移量是用线性层(linear)生成的。其过程可表示为:

$$O = \text{linear}(x) \quad (2)$$

$$S = G + O \quad (3)$$

输入图像 x 经过两个线性变换, 输出两个特征图。这两个特征图通过逐元素乘法(element-wise multiplication)进行融合, 并乘以缩放因子 0.5α 。经过缩放后的特征图进行像素洗牌(Pixel Shuffle), 生成偏移量 O 的尺寸为 $sH \times sW \times 2g$ 。像素洗牌将特征图的深度信息重新分配到空间维度上, 以提高特征图的空间分辨率。

该方法通过动态生成采样点, 提高防震锤的微小特

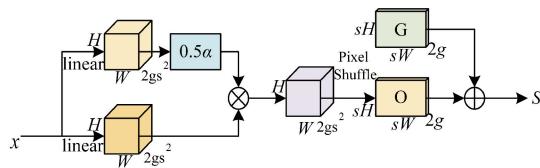


图4 DySample中的采样点生成器

Fig. 4 Sampling point generator in DySample

征重建质量。使得在检测防震锤缺失过程中能够关注到小目标的细节特征,在采样过程中减少信息的丢失,从而提高检测的精度。

2.3 改进 C2f 网络模型

在增加目标检测层之后,Neck 部分的 C2f 模块由原来的 4 个增加到了 7 个,并且上采样以及卷积层都有所增加,使得模型的大小以及参数量计算量都增大。基于以上原因,引入 FasterNet^[17] 的核心算子部分卷积(partial convolution, PConv),并设计一种全新的多尺度高效特征提取模块(multi-scale faster, MultFaster)。如图 5 所示,MultFaster 模块通过部分卷积、多级特征提取和残差连接机制,能够在减少计算复杂度的同时高效提取图像中的多尺度特征信息并进行防震锤缺失的特征学习和信息传递。

首先对输入特征 X 进行 PConv,输出特征图的通道数被减少到一半:

$$X_{Conv1} = PConv(X) \quad X_{Conv1} \in \mathbf{R}^{C/2 \times H \times W} \quad (4)$$

再次进行 PConv,输出通道再次减半:

$$X_{Conv2} = PConv(X_{Conv1}) \quad X_{Conv2} \in \mathbf{R}^{C/4 \times H \times W} \quad (5)$$

再次经过 PConv:

$$X_{Conv3} = PConv(X_{Conv2}) \quad (6)$$

将经过不同阶段的特征进行拼接,组合形成更高维度的特征图:

$$X_{Concat} = Concat(X_{Conv1}, X_{Conv2}, X_{Conv3}) \quad (7)$$

对高维度的特征图进行通道分割:

$$v, x = split(X_{Concat}), \quad v, x \in \mathbf{R}^{C/2 \times H \times W} \quad (8)$$

通过深度可分离卷积和 v 门控因子的乘法,再与输入 X 相加,可以动态调整输入的通道量,增强网络的非线性表达能力。与标准卷积相比,能够自适应地控制哪些特征被激活,从而提高模型的表达能力:

$$X = X_{DWConv}(x) \times v + X \quad (9)$$

如图 5 所示,PConv 对输入的一部分通道应用普通卷积来提取空间特征,其余通道保持不变。采用 c_p 卷积通道(c_p)作为计算整个特征映射的代表, k 是卷积核大小。因此,一个 PConv 的 FLOPs 只有:

$$h \times w \times k^2 \times c_p^2 \quad (10)$$

当取常用的取值 $r = \frac{c_p}{c} = \frac{1}{4}$ 时,一个 PConv 的

FLOPs 是一个普通 Conv 的 $\frac{1}{16}$,只有:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (11)$$

由此可见 PConv 能够同时减少计算冗余和内存访问。

将 Neck 中 C2f 的 Bottleneck 模块替换为 MultFaster,构成 C2f-MultFaster,在保持防震锤特征检测精度的同时,减小网络的计算成本。

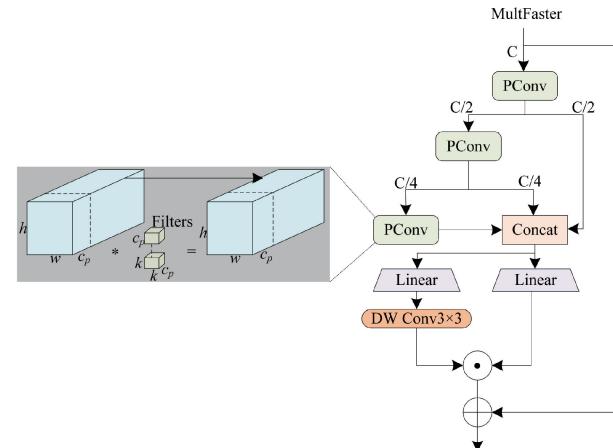


图5 MultFaster与PConv

Fig. 5 MultFaster and PConv

2.4 轻量化检测头

在实际进行防震锤的检测时,高效率且轻量化是检测模型所应具备的能力。为了提高检测头的性能,重新设计一种轻量化检测头替换原有的两分支解耦头。

如图 6 所示,在 Head 原有的 Detect 检测头进行检测时,首先并行通过两个 3×3 的 Conv 和一个 3×3 的 Conv2d,然后进行分类和回归。未改进的 YOLOv8n 模型有 3 个检测头,在 Detect 部分共要进行 12 次 3×3 的 Conv 和 6 次 3×3 的 Conv2d。在上述改进中,增加多尺度融合与多个小目标检测头,检测头增加到 5 个(要进行 20 次 3×3 的 Conv 和 10 次 3×3 的 Conv2d 的运算),使得在 Detect 部分堆叠了大量的权重参数,浪费大量的计算资源,从而导致检测效率的下降。

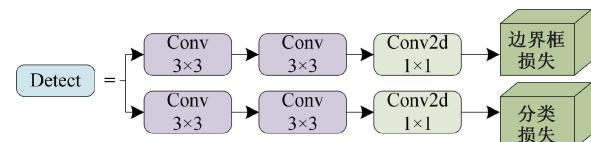


图6 原模型的 Detect

Fig. 6 Detect of the original model

根据上述中原模型检测头的不足,重新设计的 Detect Efficient 如图 7 所示。将原来并行的 3×3 的卷积

改为由一个组卷积和一个 3×3 的普通卷积串联而成的高效的特征融合模块 (Efficient fusion), 极大的简化了原来 Head 部分的复杂结构, 减少了其庞大的运算量。

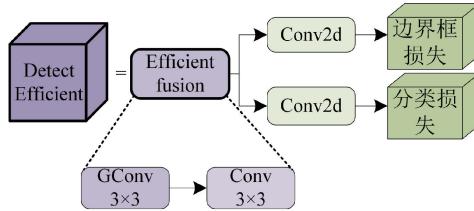


图 7 Detect efficient

Fig. 7 Detect efficient

其中, 组卷积 (convolutional groups, GConv) 如图 8 所示。

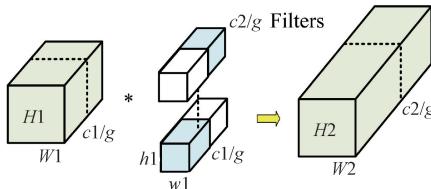


图 8 组织卷

Fig. 8 GConv

图中将输入特征图的通道分成了 g 组, 同时每组中卷积核的通道数为 $c1/g$, 卷积核的大小没有改变, 此时每组中卷积核的个数由原来的 $c2$ 变为了 $c2/g$ 个。用每组的卷积核同它们对应的输入特征图进行卷积, 输出的特征图通道数为 $c2/g$ 。整个过程并行运算 g 组相同的卷积。卷积核大小为:

$$h1 \times w1 \times c1/g \quad (12)$$

若不进行分组卷积其参数量为:

$$c1 \times h1 \times w1 \times c2 \quad (13)$$

若进行分组卷积其参数量为:

$$g \times c1/g \times h1 \times w1 \times c2/g \quad (14)$$

故参数量减小为 g 倍。因此该高效的特征融合模块能够极大的缓解原检测头所带的参数量大的问题。

3 模型剪枝

目标检测模型中并不是所有权值都是重要的, 可在精度基本保持不变的情况下, 剪除不重要的连接和冗余的权重, 以达到减小模型的尺寸, 计算量和参数。这种方法就是剪枝, 剪枝是精度无损的模型轻量化方法之一。

LAMP^[18] 剪枝算法操作简单且剪枝效果明显, 因此用该方法对改进模型进行剪枝操作。

3.1 LAMP 剪枝策略

给定一个神经网络前向传播, 它表示输入 x 经过权

重矩阵和激活函数, 直到输出:

$$f(x; \mathbf{W}^{(1:d)}) = \mathbf{W}^{(d)} \sigma(\mathbf{W}^{(d-1)} \sigma(\cdots \mathbf{W}^{(2)} \sigma(\mathbf{W}^{(1)} x))) \quad (15)$$

其中, 输入为 x , 网络层数为 d , $\mathbf{W}^{(i)}$ 表示第 (i) 层的权重矩阵, σ 是激活函数。然后分别计算每层权重张量 \mathbf{W} 的 LAMP 分数, 式为:

$$score(u; \mathbf{W}) := \frac{(\mathbf{W}[u])^2}{\sum_{v \geq u} ((\mathbf{W}[v])^2)} \quad (16)$$

$v \geq u$ 表示从权重张量 \mathbf{W} 的第 u 个权重数值到 \mathbf{W} 的最后一个权重数值。

根据权重的重要性得分, 进行权重的大小排序, 确定每层中需要保留的权重, 然后进行剪枝:

$$(\mathbf{W}[u])^2 > (\mathbf{W}[v])^2 \rightarrow score(u; \mathbf{W}) > score(v; \mathbf{W}) \quad (17)$$

4 实验结果与分析

4.1 数据集

防震锤数据集由 2 417 张高清的无人机巡检图像组成, 其中标签的比例为: 完整的防震锤 (none_defective) : 缺失的防震锤 (defective) = 4 077 : 3 592。经标注过的数据集按 8 : 1 : 1 的比例随机分为训练集, 验证集和测试集, 其中训练集由 1 936 张图片组成, 验证集 241 张图片。

4.2 实验环境与评估指标

实验硬件配置为 Xeon(R) Silver 4214R 处理器, 显卡为 NVIDIA GeForce RTX 3080 Ti, 开发语言为 python3.8, CUDA 版本为 11.3, Pytorch 版本为 1.11。模型训练参数如表 1 所示。

表 1 训练参数

Table 1 Training parameters

图像尺寸	640×640
训练轮数	240
批量大小	12
初始学习率	0.01
训练优化器	SGD
amp	True

实验评价指标有精确率 (precision, P)、平均精确率 (average precision, AP)、平均精确率均值 (mean average precision, mAP)、参数量 (Params)、计算量 (GFLOPs)、模型尺寸 (MB), 公式如式 (18) ~ (20) 所示。

$$P = \frac{TP}{TP + FP} \quad (18)$$

$$AP = \int_0^1 P(R) d(R) \quad (19)$$

$$mAP = \frac{1}{N} \sum_{i=1}^n AP_i \quad (20)$$

其中, TP 表示模型正确检测到的正例样本数量, FP 表示模型将负例预测为正例的样本数量, FN 表示将正例预测为负例的样本数量, N 指类别数量, $mAP@0.5$ 表示 $IoU=0.5$ 时的平均精确率均值。

4.3 消融实验

为了验证改进后模块对模型的提升效果,设计了 10 组消融实验,并将各个改进过程依次叠加。消融实验结果如表 2 所示。第 1 组实验将 Neck 中的 C2f 模块替换为 C2f-MultFaster, 通过 PConv 同时减少计算冗余和内存访问, 计算量减少了 0.6 G, 模型存储空间减小了 0.61 MB。第 2 组实验在 Neck 中结合动态上采样, 通过动态生成采样点生成高分辨率的特征图, 进一步提高特征图

的分辨率, 更好地还原防震锤缺失特征, 增强了模型对细节的感知能力, 使得模型在计算量及参数量不变的情况下 $mAP@0.5$ 提升 1.9%。第 3 组实验增加了两个小目标检测层, 将多尺度特征图充分融合增加了小目标提取能力, 能够捕获更多的细粒度特征, 使得 $mAP@0.5$ 提升了 6.34%。第 4 组实验将原来的解耦检测头更换成更为高效的 Detect Efficient, 通过组卷积能够减小计算量, 计算量下降 1.3 G。综合 4 种改进方案并对上述 4 种改进后的模型进行剪枝得到最终的改进模型 YOLOv8-SPH, 使得最终模型相比于原模型参数量下降了 80.73%, 计算量下降了 48.14%, $mAP@0.5$ 值相比于原模型提升了 6.3%, 模型尺寸下降了 62.41%。

表 2 消融实验

Table 2 Ablation experiment

C2f-MultFaster	DySample	多尺度特征图和小目标检测头	Detect	Efficient	LAMP	剪枝	$mAP@0.5/\%$	Params/ ($\times 10^6$)	浮点数/ GGFLOP	模型尺寸/MB
-	-	-	-	-	-	-	85.21	3.01	8.1	5.96
✓							85.13	2.67	7.5	5.35
	✓						87.11	3.01	8.1	5.99
		✓					91.55	2.89	23.5	6.35
			✓				84.16	3.12	6.8	6.2
✓		✓					90.12	2.46	21.9	5.55
	✓	✓					92.11	2.9	23.6	6.38
✓	✓	✓					91.1	2.48	22	5.57
✓	✓	✓	✓				92.41	2.24	6.9	5.09
✓	✓	✓	✓	✓	✓	✓	91.51	0.58	4.2	2.24

综上, 通过此次改进, 实现了精度上涨的同时, 参数量计算量及模型尺寸都极大的减小, 因此改进后模型可以更好地用于防震锤的检测任务。

4.4 模型可视化对比

如图 9 所示, 改进前后模型的 $mAP@0.5$ 曲线对比。从图中可以明显看到, 改进后模型的 $mAP@0.5$ 较原来的模型有所提升, 且收敛速度加快, 检测精度也有显著提高。

如图 10 所示, 改进前后混淆矩阵的对比情况。图中对角线上的元素表示模型正确预测样本的比例, 非对角线上的元素表示模型误检或漏检样本的比例, 其中 A 代表 none_defective, B 代表 defective, C 代表 background。在验证集 241 张图片中, none_defective 共有 155 个, defective 有 162 个, 如图 10(a) 所示, 其中改进前 YOLOv8n 对于 none_defective 和 defective 的识别率分别为 0.78、0.8, 漏检率较高, 正确率较低; 如图 10(b) 所示, 对于 none_defective 和 defective 的识别率分别为 0.92、0.86, 均高于图 10(a)。表明 YOLOv8-SPH 模型在降低漏检率方面有显著改善。

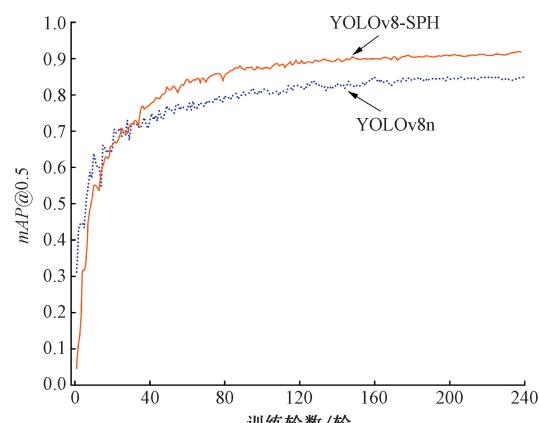


图 9 改进前后模型 $mAP@0.5$ 对比

Fig. 9 Model before and after improvement
 $mAP@0.5$ comparison chart

如图 11 所示, 各模块剪枝后及被剪去的通道数量, 并按剪枝率大小进行排序。其中, 图左边是各模块剪枝后所剩的通道数量, 图右边是各模块被剪去的通道数量。从图中可以看出, 剪枝操作显著减少了某些通道的权重

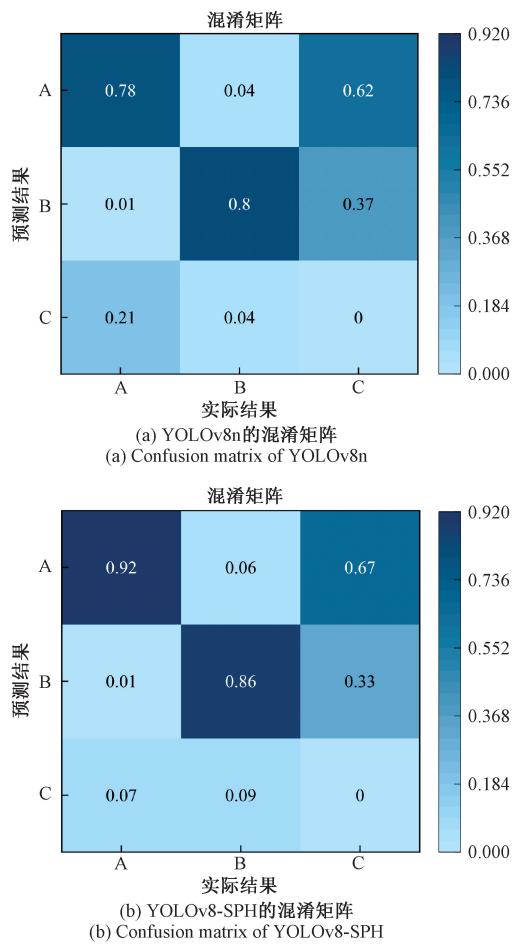


图 10 YOLOv8n 及 YOLOv8-SPH 的混淆矩阵

Fig. 10 Confusion matrix of YOLOv8n and YOLOv8-SPH

参数量,但仍保留了足够的权重以维持模型性能,实现了模型几乎无精度损失的轻量化。

如图 12 和 13 所示,改进前后算法在检测效果上的对比。原始算法在检测过程中存在误检和漏检的问题,尤其是在复杂背景下,缺失特征与背景难以区分,导致检测效果不理想。而经过改进后的 YOLOv8-SPH 模型表现出了更好的检测能力,几乎能够准确识别出所有缺失和正常的防震锤。改进后的模型不仅在复杂背景下表现更佳,而且整体检测精度也得到了显著提升,明显优于原始算法。

4.5 不同算法对比试验

将目前各主流算法与改进后算法进行对比实验,对比结果如表 3 所示。表中最优的检测结果用红色字体标出。改进后模型的 mAP@0.5 为 91.51%,是所有模型中最高的,其在防震锤缺失检测任务中的精度最高。改进后模型的参数量为 0.583×10^6 ,是所有模型中最少的,相比 Faster-RCNN^[19] 的 41.35×10^6 、Mask-RCNN^[20] 的 43.97×10^6 、YOLOv9-t^[21] 的 2.61×10^6 和 YOLOv7-tiny^[22]

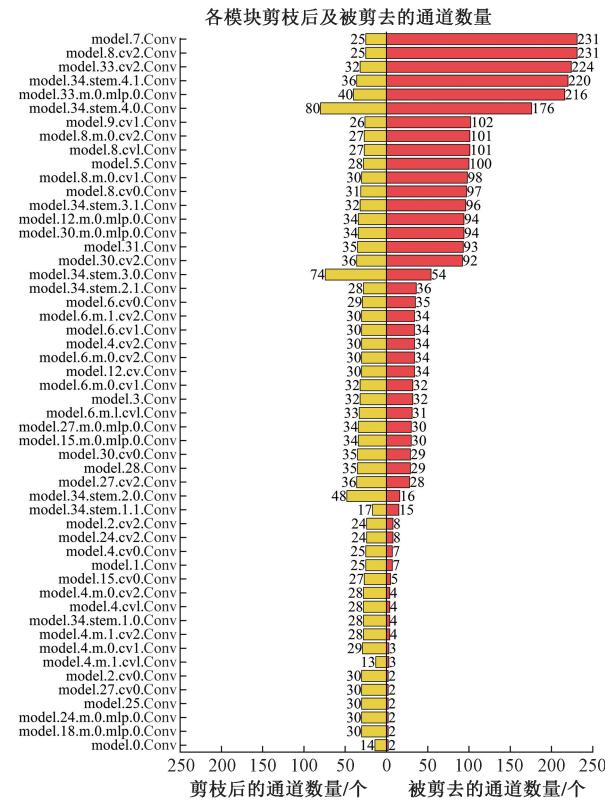


图 11 各模块剪枝后及被剪去的通道数量

Fig. 11 The number of channels pruned and pruned for each module



图 12 YOLOv8n 检测结果

Fig. 12 YOLOv8n detection results

的 6.01×10^6 ,改进后算法能够大幅减少了模型的复杂度。浮点数仅为 4.2 GGFLOPs,同样是所有模型中最低的。相比 Faster-RCNN 的 124.9 GGFLOPs 和 Mask-RCNN 的 150.4 GGFLOPs,改进后显著降低了计算成本。改进后模型的尺寸为 2.24 MB,是所有模型中最小的,使得 YOLOv8-SPH 在内存和存储空间有限的设备上更具优势。综上,改进后的模型性能优于其他算法,在保持最高检测精度的同时,显著降低了模型的参数量、计算量和模型尺寸,展现出在高效性和实用性上的明显优势,特别适合在资源有限的边缘设备中应用。

如图 14 所示,不同模型尺寸和精度的对比。不同模



图 13 YOLOv8-SPH 检测结果

Fig. 13 YOLOv8-SPH detection results

型所对应形状越大,其对应模型尺寸越大,从图中可以看到改进后模型尺寸很小,精度很高。

表 3 不同算法对比

Table 3 Comparison of different algorithms

Model	mAP@0.5/%	Params/ ($\times 10^6$)	浮点数/ GGFLOP	模型尺寸/MB
Faster-RCNN	86.82	41.35	124.9	315.0
Mask-RCNN	91.1	43.97	150.4	335.9
YOLOv8n	85.21	3.01	8.1	5.96
YOLOv5s	91.51	7.1	15.8	13.7
YOLOv9-t	80.76	2.61	10.7	5.8
YOLOv7-tiny	87.93	6.01	13.2	12.3
YOLOv8-SPH	91.51	0.58	4.2	2.24

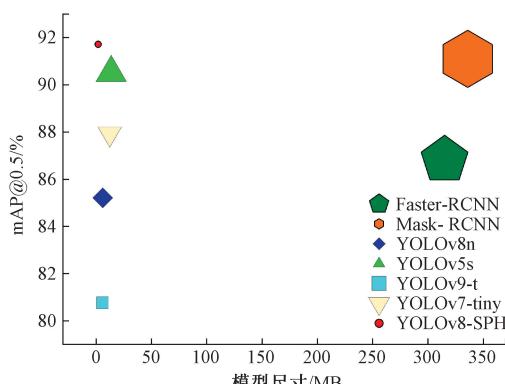


图 14 不同模型尺寸和精度的对比

Fig. 14 Comparison of different model sizes and accuracies

5 结 论

目前输电线路防震锤缺失检测缺乏有效的方法且不能友好地部署到无人机等存储空间有限的边缘设备,提出了一种新型的能兼顾模型尺寸和精度的检测算法来进行防震锤缺失检测,针对防震锤在图像中小且密集,通过增加多尺度特征融合和小目标检测头,增强多尺度融合,提升模型的感受野;针对最近邻插值法进行上采样导致的特征信息易丢失的问题,引入 DySample,提高模型对细节和局部特征的捕捉能力;针对网络模型复杂的情况提出一种新的多尺度高效特征提取模块,能够在减少计算复杂度的同时高效提取图像中的多尺度特征信息,同时结合组卷积构建了一种轻量化检测头,极大的缓解原检测头所带的参数量大的问题;对改进后的网络进行基于幅值的层自适应稀疏化剪枝,进一步减少了模型参数量和计算复杂度。通过以上创新性的改进,能够显著提升模型在防震锤缺失检测任务中的性能和实用价值。

最终改进的模型 mAP@0.5 达到了 91.51%,相比于原始模型提升了 6.3%,相比于原模型参数量下降了 80.73%,计算量下降了 48.14%,模型尺寸下降了 62.41%。在模型的计算量、参数量和模型尺寸等方面优于 YOLOv5s 及 YOLOv9-t 等,并有着很高的检测精度。综上,YOLOv8-SPH 在存储空间有限的边缘设备上有着实际应用价值,能够满足输电线路防震锤缺失检测要求。

在未来的工作中,研究者将致力于进一步优化YOLOv8-SPH模型,通过知识蒸馏技术,进一步提高模型的检测精度。同时,深入研究多源数据融合方法和迁移学习方法,将无人机采集图像与其他传感器数据(如激光雷达、红外成像)相结合,提升模型在防震锤缺失检测中的全面性和精确性。

参考文献

- [1] 赵海洋,林来鑫,全鑫,等.金具滑移对输电线路的危害分析与防治[J].自动化应用,2024,65(10):148-150,153.
ZHAO H Y, LIN L X, TONG X, et al. Analysis and prevention of hazards caused by sliding of fittings on transmission lines [J]. Automation Application, 2024, 65(10):148-150,153.
- [2] LIU Y M, DONG SH Q, CHEN ZH, et al. Research on bolt visual recognition of high voltage transmission line shockproof hammer resetting robot [C]. IEEE 6th Information Technology and Mechatronics Engineering Conference, 2022:1329-1333.
- [3] 张佳佳,袁强静,严冬,等.一种基于防震锤的新型压电复合能量收集方法[J].压电与声光,2023,45(1):56-60,65.
ZHANG J J, YUAN Q J, YAN D, et al. A novel piezoelectric hybrid energy harvesting method based on shockproof hammer [J]. Piezoelectrics & Acoustooptics, 2023,45(01):56-60,65.
- [4] LUO Y, YU X, YANG D, et al. A survey of intelligent transmission line inspection based on unmanned aerial vehicle [J]. Artificial Intelligence Review, 2023, 56(1): 173-201.
- [5] 周景,李鑫乐.基于改进型DETR的输电线路防震锤检测[J].计算机仿真,2023,40(11):101-106.
ZHOU J, LI X L. Shockproof hammer detection of transmission line based on improved DETR [J]. Computer Simulation, 2023,40(11):101-106.
- [6] 梁华刚,赵慧霞,刘丽华,等.结合跨层特征融合与级联检测器的防震锤缺陷检测[J].中国图象图形学报,2023,28(11):3485-3496.
LIANG H G ZHAO H X, LIU L H, et al. Combining cross-layer feature fusion with cascade detectors for anti-vibration hammer defects detection [J]. Journal of Image and Graphics, 2023,28(11):3485-3496.
- [7] 苏雨蕾,黄丹丹,刘智,等.无人机对地小目标检测方法研究[J].电子测量与仪器学报,2024,38(9):144-154.
SU Y L, HUANG D D, LIU ZH, et al. Research on detection methods of small targets on ground by UAV [J]. Journal of Electronic Measurement and Instrumentation, 2024,38(9):144-154.
- [8] 卢志博,徐澄宇,杨罡,等.基于改进YOLOv3的输电线路部件实时检测[J].电测与仪表,2023,60(7):138-144.
LU ZH B, XU CH Y, YANG G, et al. Real-time detection of transmission line components based on improved YOLOv3 [J]. Electrical Measurement & Instrumentation, 2023,60(7):138-144.
- [9] 徐业东,蔡亚恒,李严,等.基于YOLOv5s的轻量化架空输电线路鸟巢检测网络[J].电子测量技术,2024,47(7):138-148.
XU Y D, CAI Y H, LI Y, et al. Lightweight overhead transmission line bird nest detection network based on YOLOv5s [J]. Electronic Measurement Technology, 2024,47(7):138-148.
- [10] DONG CH, ZHANG K, XIE ZH Y, et al. An improved cascade RCNN detection method for key components and defects of transmission lines [J]. IET Generation, Transmission & Distribution, 2023,17(19):4277-4292.
- [11] 吕游,郑茜,齐欣宇,等.基于改进EfficientNet的红外图像光伏组件故障识别研究[J].仪器仪表学报,2024,45(4):175-184.
LYU Y, ZHANG Q, QI X Y, et al. A study on fault recognition of photovoltaic module with infrared images based on improved Efficientnet [J]. Chinese Journal of Scientific Instrument, 2024,45(4):175-184.
- [12] 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.
- [13] 赵振兵,王帆帆,刘良帅,等.基于注意力特征融合YOLOv5模型的无人机输电线路航拍图像金具检测方法[J].电测与仪表,2023,60(3):145-152.
ZHAO ZH B, WANG F F, LIU L SH, et al. Hardware detection method of aerial image of UAV transmission line based on attention feature fusion YOLOv5 model [J]. Electrical Measurement & Instrumentation, 2023,60(3):145-152.
- [14] 宋立业,刘帅,王凯,等.基于改进EfficientDet的电网元件及缺陷识别方法[J].电工技术学报,2022,37(9):2241-2251.
SONG L Y, LIU SH, WANG K, et al. Identification method of power grid components and defects based on improved EfficientDet [J]. Transactions of China Electrotechnical Society, 2022,37(9):2241-2251.
- [15] GUO J, XIE J H, YUAN J ZH, et al. Fault identification of transmission line shockproof hammer

- based on improved YOLOv4 [C]. International Conference on Intelligent Computing, Automation and Applications, 2021: 826-833.
- [16] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6027-6037.
- [17] CHEN J R, 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.
- [18] LEE J, PARK S, MO S, et al. Layer-adaptive sparsity for the magnitude-based pruning [J]. ArXiv preprint arXiv: 2010. 07611, 2020.
- [19] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [20] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [21] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision. Springer, Cham, 2025: 1-21.
- [22] 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 [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.

作者简介



张吉祥,2023年于安阳工学院获得学士学位,现为三峡大学硕士研究生,主要研究方向为深度学习。

E-mail: 3651068558@qq.com

Zhang Jixiang received his B. Sc. degree from Anyang Institute of Technology in 2023.

Now he is a M. Sc. candidate at China Three Gorges University. His main research interest includes deep learning.



方东(通信作者),2018年进入三峡大学工作,2017年博士毕业于武汉理工大学,获材料科学与工程专业博士学位。现为讲师,硕士生导师。主要研究方向为图像识别、高性能储热材料的开发与应用。

E-mail: 610823085@qq.com

Fang Dong (Corresponding author) joined China Three Gorges University in 2018, received a Ph. D. degree in Materials Science and Engineering major from Wuhan University of Technology in 2017. Now he is a lecturer, master's supervisor. His main research interests include image recognition and the development and application of high-performance thermal storage materials.