

DOI: 10.19650/j.cnki.cjsi.J2311233

TWRD-Net: 一种用于曳引钢丝绳缺陷的 实时检测网络算法*

高嘉, 刘涛, 王显峰, 杜宏旺, 史震

(哈尔滨工程大学智能科学与工程学院 哈尔滨 150000)

摘要: 曳引钢丝绳 (TWR) 在大型工业起重设备中发挥着重要的应用价值。在使用曳引钢丝绳作业的同时, 也需要定期对曳引钢丝绳进行缺陷诊断避免安全事故发生。传统方法是人工目测方法, 但该方法检测时间长、效率低下。因此, 本文提出了一种基于 YOLOv5 改进网络的曳引钢丝绳缺陷 (TWRD) 检测的网络算法, 简称 TWRD-Net。为了方便在算力较低的工业设备上部署, 首先设计了轻量级的 LW-C3 模块, 降低了模型的参数量以及计算开销; 其次, 改进 PAN 结构, 设计了跨层级加权特征金字塔网络 (CLW-FPN) 结构, 加强模型对于缺陷语义信息的提取、对缺陷位置信息的敏感度; 最后, 本文设计了 β -CIoU 损失函数, 相较于 CIoU 损失函数, β -CIoU 降低了边界框回归损失, 并进一步提升了检测精度。本文建立了 TWRD 数据集, 并使用 TWRD-Net 进行实验, 实验结果表明, 本文提出的 TWRD-Net 准确率可达 98%, mAP 可达 99.4%, 帧率可达 151 fps, 对比其他主流检测模型实验结果, 具有精度高、体积小和检测速度快的优势, 可为工业设备质检人员提供参考依据。

关键词: 曳引钢丝绳; TWRD; 缺陷检测; YOLO

中图分类号: TH741 TP391.41 文献标识码: A 国家标准学科分类代码: 460.40

TWRD-Net: A real-time detection network algorithm for traction wire rope defects

Gao Jia, Liu Tao, Wang Xianfeng, Du Hongwang, Shi Zhen

(College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150000, China)

Abstract: Traction wire rope (TWR) plays an important application value in large-scale industrial lifting equipment. While using the traction wire rope for operation, it is necessary to regularly diagnose the defects of the traction wire rope to avoid safety accidents. The traditional method is manual visual inspection, but this method has long detection time and low efficiency. Therefore, this article proposes a network algorithm for detecting traction wire rope defect (TWRD) based on YOLOv5 improved network, abbreviated as TWRD-Net. In order to facilitate deployment on industrial equipment with low computational power, a lightweight LW-C3 module is first designed to reduce the model's parameter count and computational overhead. Secondly, the PAN structure is improved by designing a CLW-FPN structure to enhance the model's sensitivity to defect semantic information extraction and defect location information. Finally, this article designs β -CIoU loss function. Compared with CIoU Loss function, β -CIoU reduces the loss of bounding box regression and further improves detection accuracy. This article establishes a TWRD dataset and conducts experiments by using TWRD-Net. The experimental results show that the accuracy of the proposed TWRD-Net can reach 98%, mAP can reach 99.4%, and frame rate can reach 151 fps. Compared with other mainstream detection model experimental results, it has the advantages of high accuracy, small size, and fast detection speed, which can provide reference for industrial equipment quality inspectors.

Keywords: traction wire rope; TWRD; defect detection; YOLO

0 引 言

曳引钢丝绳是工业领域内极其重要的承载构件,塔式起重机、电梯轿厢升降、起重机、索道等等都具有驱动作用。曳引钢丝绳在作业期间,一般会受到拉伸、弯曲、挤压等应力,外形上难免会产生断丝、磨损、生锈、断股,甚至突然断裂等缺陷和损伤^[1]。如果曳引钢丝绳长期作业又没有及时保养,不及时发现更换则会造成安全事故,危机人的生命安全。因此,对曳引钢丝绳的质量进行无损缺陷检测具有重要意义。

钢丝绳缺陷检测最初通过人工目测方法,但人工质检效率极低,所以为了解决这一问题,国内外有专家学者有做过相关方面的技术研究。传统的曳引钢丝绳的缺陷检测技术主要基于物理传感器进行检测,例如电磁检测技术^[2-4]利用永磁类检测原理检测钢丝绳中断丝处引起的漏磁进行损伤定性,超声波检测技术^[5-6]利用超声波对钢丝绳不同位置的反射特征进行损伤定性检测,声发射检测技术^[7-8]利用压电效应将瞬态弹性波转换为电信号,通过对电信号分析进行内部损伤定性。以上方案虽有应用,但仍然存在对缺陷难以定量分析、类别区分的问题,并且抗噪声能力弱,仪器成本高等。后来随着数字图像处理技术的快速发展,基于机器视觉的钢丝绳缺陷检测掀起热潮,主要采用高速相机采集钢丝绳表面图像,然后对钢丝绳图像进行预处理,利用钢丝绳表面的纹理信息^[9]、边缘密度等描述钢丝绳表面的特征进行前景分割,然后提取钢丝绳表面规律结构、LBP 特征^[10]、股绳边界信息等特征信息去训练线性分类器、BP 神经网络^[11]、支持向量机(support vector machine, SVM)^[10,12]等机器学习模型来实现钢丝绳缺陷的分类。Wu 等^[13]通过对曳引钢丝绳图像进行轮廓提取、直线拟合等来判断曳引钢丝绳的变形。传统的数字图像处理技术虽然在缺陷识别上有重大突破,但是仍然会受光线因素的影响造成识别干扰,并且受限于人工设计的表面特征,可识别的种类有限,泛化能力弱,鲁棒性不强,不能适用所有缺陷类型。

2012 年, Krizhevsky 等^[14]引导的小组提出了 AlexNet 模型,基于卷积神经网络的图像分类算法在 ImageNet 大赛中一举夺冠,开启了深度学习时代,随后各种不同模型层出不穷,深度学习在不同领域当中都获得了惊人的应用价值。目前,已有将深度学习方法应用到工业缺陷检测当中,例如张银胜等^[15]改进 YOLO 系列框架实现对风力涡轮机表面缺陷的检测;刘兰兰等^[16]提出一种基于超分辨率重建与多尺度特征融合的输电线路缺陷检测,对小尺寸缺陷具有更强的鲁棒性;陈仁祥等^[17]提出基于多注意力 Faster RCNN 的印刷电路板缺陷检测方法解决工业环境中噪声干扰导致的印刷电路板缺陷检测困难的问

题,验证了深度学习在缺陷检测领域的可行性;刘艳菊等^[18]改进 YOLOv4 算法网络,改进主干网络为 MobileNetv3 来提高检测速度等,在热轧钢条的表面缺陷检测上表现出良好的性能。综合以上案例,利用深度学习技术解决工业缺陷检测问题是未来趋势。

1 相关工作

YOLOv5 相较于前代 YOLOv4, YOLOv5 拥有更快的推理速度以及更小的参数体积,其主要包含 3 大部分:主干网络 Neck 网络和检测头。

主干网络主要包含 CBS 模块、C3 模块以及空间金字塔快速池化层(spatial pyramid pooling fast, SPPF)。CBS 模块是基础特征提取模块,由 Conv 层、Batch Normalization 层、SiLU 激活层串联组成,C3 模块是原跨阶段局部网络 CSPNet^[19]的变体,主要存在两种形态,分别用于主干网络和 Neck 网络。Neck 网络部分主要是对主干网络中提取的不同尺寸的特征图进行特征融合,获得更多的上下文信息,防止信息丢失,其采用 PAN 结构进行多尺度特征融合,PAN 结构是在 FPN 自顶向下进行上采样的结构基础上增加自底向上的下采样,使顶层特征图包含更多目标的位置信息,并对同一尺寸的特征图进行信息融合,使不同尺寸特征图都包含目标的强语义和强特征信息。最先进的检测器通常包含边界框回归和分类回归,边界框回归最常用的损失函数是 IoU 损失函数及其变体,用以定位对象进行分类。目前常见的 IoU 损失有 GIoU^[20]、EIoU^[21]、DiIoU^[22]、CIoU^[22]、SIoU^[23]等。YOLOv5 中采用了 CIoU, CIoU 损失函数考虑了 3 种几何参数:重叠面积、中心点距离、长宽比。CIoU 准确来讲增加了对预测框长和宽的 Loss,这种预测出的方框更加符合目标框的尺寸。

本文针对曳引钢丝绳缺陷检测的工程应用场景,拟选取 YOLO 系列作为基线框架进行改进。目前 YOLO 系列版本众多, YOLOv8 与 YOLOv7^[24]也横空出世,具有更强劲的性能。YOLOv8 系列在 COCO 数据集上带来了精度的巨大提升,但其模型体积相较于 YOLOv5 大了许多,牺牲了推理速度;YOLOv7 系列 Tiny 版本模型较小,但精度明显欠缺, YOLOv7 系列标准版精度速度表现都不错,但是同样模型体积过于庞大,资源占用率高。在工程应用层面,由于模型需要在硬件平台进行部署,考虑到硬件平台的算力问题,在精度满足正常需求的情况下,应该更注重于模型的推理速度、资源占用率,更轻量化的模型。YOLOv5 系列虽然精度略低于 YOLOv7、YOLOv8,但已然能够满足实际需求,并且与 YOLOv7、YOLOv8 同版本下的 YOLOv5 模型体积更小,占用内存更小,相较于 v7、v8 系列能够实现精度、速度、资源占用率三者的良好平衡,

另外, YOLOv8 系列目前还在更新, 不够稳定。因此综合以上考虑, 本文最终选择 YOLOv5 系列的-s 版本模型进行改进, 构建了 TWRD-Net。

2 本文方法

2.1 主干网络的改进方法

YOLOv5 的主干网络中的核心模块是 C3 模块, 在大量的实际应用中已经证实其借助 CSPNet 提取分流、ResNet^[25] 残差结构的思想, 即在 C3 模块的主分支采用堆叠 BottleNeck 模块的方法实现了模型性能的最优与平衡。但是为了模型能够部署于硬件平台进行实时的 TWRD 检测, 需要进一步对 C3 模块进行改良, 因此设计了轻量化 C3 模块 (light weighted C3, LW-C3), LW-C3 模块与 C3 模块对比图, 如图 1 所示。LW-C3 模块与 C3 模块一样, 同样利用 CSPNet 提取分流的思想, 并在主梯度流当中堆叠残差结构模块。实现 C3 模块的轻量化主要体现在以下两个方面:

一方面, 在普通卷积、批归一化、激活组合 (convolution、batch normalization、swish, CBS) 中引入深度可分离卷积^[26], 得到深度可分离卷积、批归一化、激活组合 (depthwise separable convolution batch normalization wwish, DSC-BS) 如图 1 所示深度可分离卷积 (depthwise

separable convolution, DSConv) 与标准卷积不同的是, 标准卷积是空间与通道层面同时进行卷积, 然而深度可分离卷积则是将两者分开进行, 先进行空间层面的深度卷积, 然后进行通道层面的逐点卷积, 如图 2 所示。首先, 深度卷积阶段, 输入特征图的每个通道只与一个卷积核进行卷积, 生成一个与输入特征图通道数相同的中间特征图。假设输入特征图尺寸为 $M \times H_{in} \times W_{in}$, 则卷积核尺寸为 $M \times 1 \times H_k \times W_k$, M 为卷积核个数, 输出的中间特征图尺寸为 $M \times H_{mid} \times W_{mid}$; 其次, 逐点卷积阶段与标准卷积的运算类似, 它的卷积核的尺寸固定为 $C \times M \times 1 \times 1$, M 为深度卷积输出的中间特征层的通道数, C 为卷积核个数。逐点卷积运算会将深度卷积的输出特征图在通道方向上进行加权组合, 起到融合不同通道在相同空间位置上的特征信息的作用, 其输出特征图数量取决于逐点卷积的卷积核数量, 即 C , 逐点卷积输出特征图尺寸为 $C \times H_{out} \times W_{out}$ 。

C3 模块中两条分支采用的是 1×1 的 CBS 模块进行特征图通道的降维, 本文使用 1×1 的 DSC-BS 模块替换了 CBS, 因此, 结合 DSConv 以上卷积过程的特性, 在相同输入输出、卷积核的情况下, 深度可分离卷积将一个复杂的卷积操作拆分为两个简单的步骤, 使得模型变得更轻量级和高效, 相对于标准卷积可以减少 $\frac{1}{N} + \frac{1}{D_k^2}$ 计算量, N 为输出通道数, D_k 为卷积核大小。

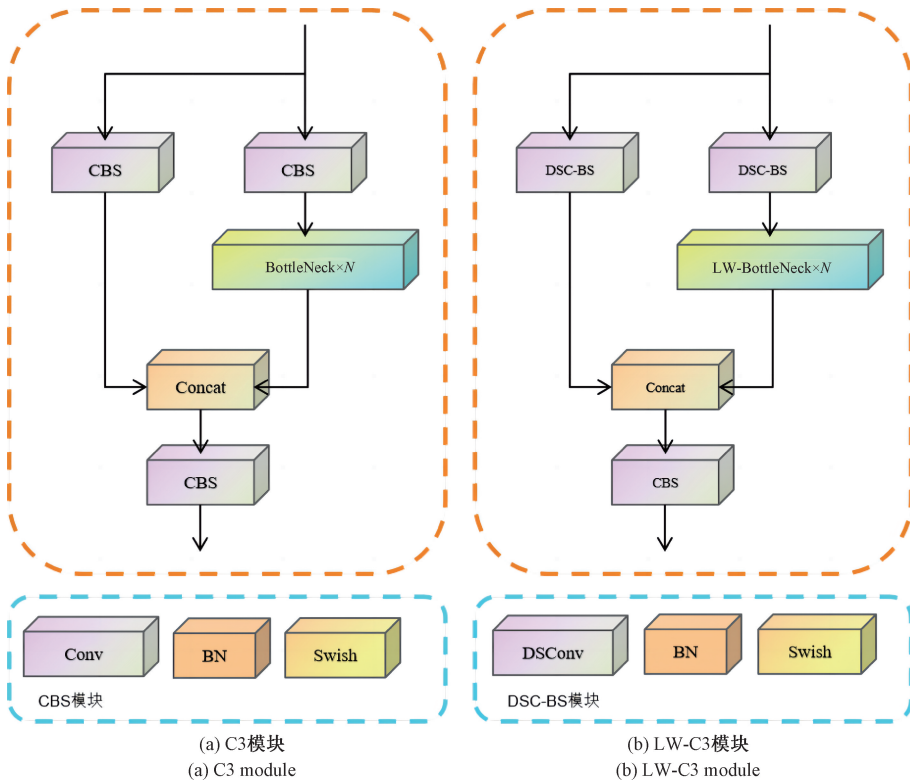


图1 C3模块与LW-C3模块对比图

Fig. 1 Comparison diagram of C3 module and LW-C3 module

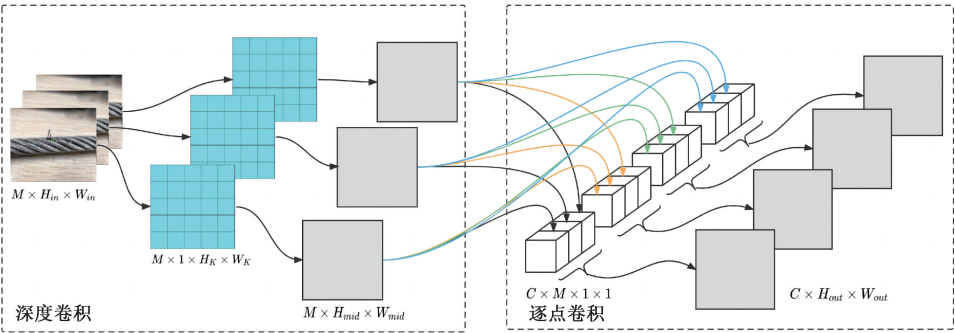


图2 深度可分离卷积结构

Fig. 2 Depthwise separable convolution structure diagram

另一方面,设计主分支梯度模块——轻量化瓶颈模块 (light weighted bottleneck, LW-BottleNeck)。LW-BottleNeck 是一个轻量级模块,其主要有 3 条梯度分支 $branch_i, i \in (1, 2, 3)$, LW-BottleNeck 模块如图 3 所示。主要流程如下:首先 branch 1 是一条级联了 1×1 、 1×3 、 3×1 CBS 卷积层的梯度分支,branch2 中是一个包含

3×3 CBS 卷积的分支,branch3 是残差分支。对于一个形状为 $C \times W \times H$ 的输入特征图 F 而言,其需要经过 3 条梯度分支处理,首先 $branch_1$ 分支的输出会与 $branch_2$ 分支的输出进行 Concat 连接,并得到中间输出 F_{mid} ,然后 F_{mid} 与 $branch_3$ 分支进行 add 操作,最终得到特征图 F_{out} 作为 LW-BottleNeck 模块的输出。

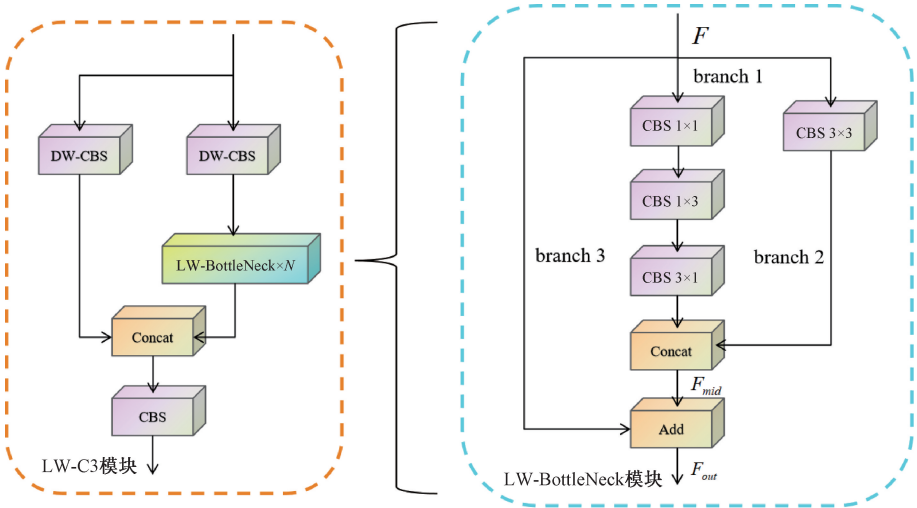


图3 LW-BottleNeck 结构

Fig. 3 LW-BottleNeck structure diagram

其中, $branch_1$ 分支中添加 1×1 卷积,其一,因为 1×1 卷积能够对特征图信息进行降维处理,降低了模型参数量以及计算量;其二,采用的 1×1 卷积在 CBS 模块中都会经过一次 Swish 激活函数的处理,加深了网络深度,增加了更多的非线性。在 1×1 卷积后接入了 1×3 卷积与 3×1 卷积,该种结构为非对称卷积^[27],通过这种级联方式组成的非对称卷积结构与标准的 3×3 卷积具有相同感受野,一方面加深了网络深度,增强了网络非线性表达的能力,并且不同形式的卷积核可以扩充特征空间,更加适应不同 width 与 height 比例的目标特征。另一方面在参数量与计算开销上能够节省 33%左右的计算开销与参数量。

$branch_2$ 分支中的 3×3 卷积核目的是为了减少深度可分离卷积与非对称卷积之后带来的精度损失。本文经实验发现,深度可分离卷积以及非对称卷积两种轻量化结构,模型的整体参数量得到了大量的减少,但是造成模型捕捉复杂特征的能力下降,因此在 LW-BottleNeck 中增加 3×3 卷积分支,通过这种增加模型宽度的方式,其在一定程度上能够增加模型的非线性表达能力,学习样本中更多不同层次的特征信息来弥补参数量减少带来的精度损失。

$branch_3$ 作为残差分支将与 $branch_1$ 和 $branch_2$ 的 Concat 进行恒等映射,一方面能够解决网络退化问题,防

止梯度消失或爆炸带来的训练困难,另一方面与 branch1 和 branch2 分支共同分割梯度流,使梯度经过不同的网络路径进行传播,最后 branch3 和 branch1 与 branch2 两个分支的 Concat 输出进行 add 融合,实现更丰富的梯度信息组合,能够在保留输入特征图语义信息的同时,增强卷积特征图当中的有效信息,最终得到输出 F_{out} 。

2.2 Neck 网络的改进方法

针对 Neck 网络进行改进,设计了跨层级加权特征金字塔网络 (cross level weighted feature pyramid network, CLW-FPN),包括跨层级加权双连接模块 (cross level weighted add 2, CLW-Add2) 和包括跨层级加权三连接模块 (cross level weighted add 3, CLW-Add3),如图 4 所示,

CLW-FPN 进行跨层级连接,增加了 P51~P31 的信息融合以及 P32~P52 的信息融合,去除 FPN 中 P41~P42 的链接,将 P4 链接到 P42 中。其中,P41、P51 和 P4、P32 两组经过 CLW-Add2 卷积模块,P51、P41、P3 和 P32、P42、P51 两组则分别经过 CLW-Add3 Up/Down 卷积模块。CLW-Add2 和 CLW-Add3 结构如图 5 所示。CLW-Add 模块沿用了 BiFPN^[28] 的融合模式,CLW-Add 将输入的两层特征图分别乘以对应权重进行加权融合,然后经过 CBS 模块输出特征图;CLW-Add3 分 UpSample 与 DownSample,由于下采样需要进行卷积操作,采用 1×1 大小的 DSC-BS 模块进行下采样和通道的升降维两种操作来避免因增加卷积操作带来的计算量巨增。

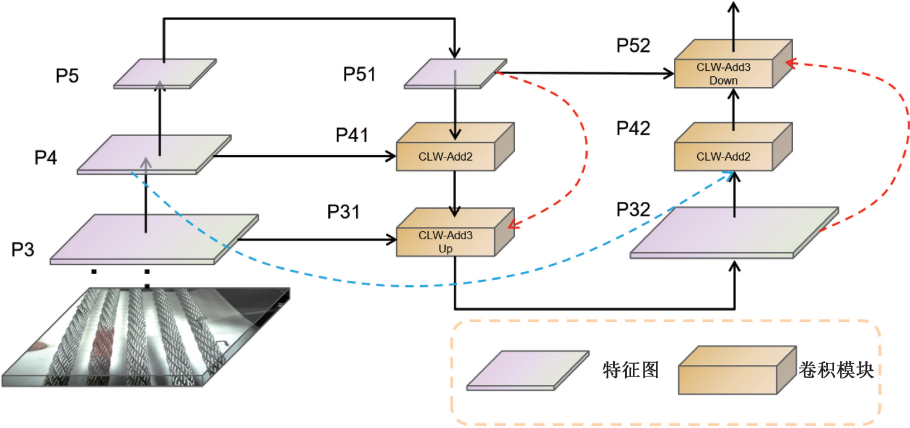


图 4 CLW-FPN 结构
Fig.4 CLW-FPN structure diagram

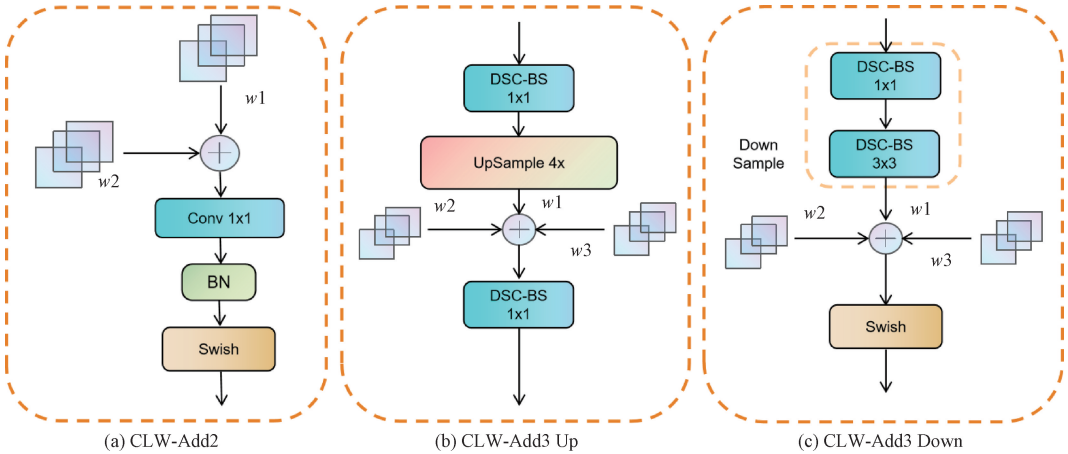


图 5 CLW-Add 模块的 3 种形式
Fig.5 Three forms of CLW-Add module

对输入的不同特征图赋予不同的权重进行加权处理,该方案沿用 BiFPN 的方案,对于多个输入的 CLW-Add 模块,权重分配方法如式 (1) 所示,通过这种方式进行处理,能够使模型学习到不同输入特征的重要性,对不

同输入特征进行有区分的融合,使模型能够减少冗余信息,降低无用信息的干扰,突出有效信息。

$$O = \sum_i \frac{w_i * I_i}{\varepsilon + \sum_j w_j}$$

(1)

其中, O 是分配权重, w_i 是可训练的权重值, $*$ 代表卷积, I_i 是输入特征图, ε 是防溢出值, 常设为 0.001。

2.3 IoU 损失函数

在目标检测任务当中, 目标检测模型的好坏常常用到指标 IoU 进行衡量, 不仅可以作为正负样本划分的依据, 也可以用作损失函数。CIoU 中提到, 对于一个优秀的损失函数应该包含重叠面积、中心点距离、长宽比, 因此最终 CIoU 公式如式(2)所示。

$$CIoU = IoU - \left(\frac{d^2}{c^2} + \alpha \times v \right) \quad (2)$$

其中, $v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2$, $\alpha = \frac{v}{(1 - IoU) + v}$ 。 $\frac{d^2}{c^2}$ 是中心点距离与两边界框对角线长度之比。本文在 CIoU 基础上进行了改进, 引入 β 系数, 提出了 β -CIoU 损失函数, 通过引入 β 系数提高对边界框回归计算的比重, 增大边界框回归的梯度来进一步降低边界框损失。具体公式如式(3)所示。

$$\beta - CIoU = IoU - \left(\frac{d^2}{c^2} \right)^\beta - (\alpha \times v) \quad (3)$$

本文认为 β 系数应为可变参数, 因为 IoU 损失中不仅包含对坐标的回归, 还有对长宽的回归, 因此当预测框与目标框距离足够相近时, 若恒为常数, 那么由于坐标回归与长宽回归的比重原因, 坐标回归的比重会始终大于长宽回归的比重, 这样可能会对长宽的回归计算造成抑制, 因此设计的 β 系数取值应取决于预测框与目标框的距离信息, 其计算公式如式(4)所示。

$$\beta = -\frac{1 + e^{-x}}{\ln(\gamma \cdot x)} \quad (4)$$

其中, $x = \frac{d^2}{c^2}$ 为预测框与目标框中心点距离与对角

线距离之比, 因此 $x \in (0, 1)$ 。另外, $\gamma \in (0, 1)$ 为缩放因子, 控制 β 系数的变化速率, β 系数随 γ 变化曲线图如图 6 所示, 由曲线图可以看出, β 系数会随着 x 的减小缓慢递减, 即当预测框与目标框距离较远时, x 趋近于 1, 此时 β 系数最大, 边界框的坐标回归损失会加大, 优先对坐标进行回归, 当预测框与目标框越来越靠近时, x 趋近于 0, β 系数会逐渐减小, 减小坐标回归损失的比重, 突出边界框的长与宽的回归比重。

3 实验细节

3.1 实验场景与设备介绍

本文针对曳引钢丝绳缺陷检测方法研究, 在实验室搭建了一套电梯曳引实验装置, 并在曳引钢丝绳上模拟了 3 种缺陷。电梯曳引实验装置如图 7(a) 所示, 曳引钢

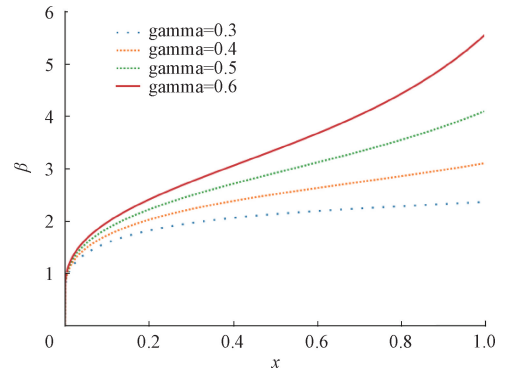
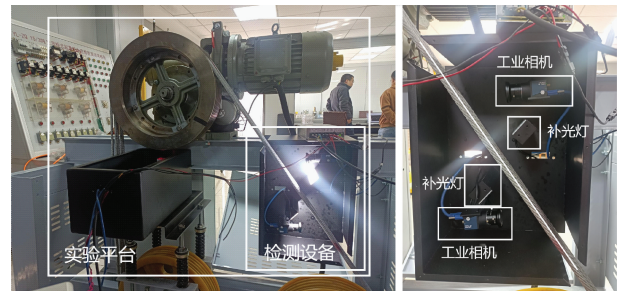


图 6 β 系数随 γ 变化曲线

Fig. 6 The graph of coefficient β with variation γ

丝绳绕曳引轮, 并由电机进行驱动牵引。TWRD 检测装置如图 7(b) 所示, 其中包含工业相机、补光灯, 用于采集曳引钢丝绳图像与补光, 对角位置为曳引钢丝绳的出入口, 即曳引钢丝绳通道。



(a) 实验环境
(a) The experimental environment

(b) TWRD检测装置
(b) The TWRD detection device

图 7 实验环境和 TWRD 检测装置

Fig. 7 The experimental environment on the left and the TWRD detection device on the right

实验过程中, 电机驱动曳引钢丝绳进行正转或翻转, 工业相机在补光灯的辅助下对移动的曳引钢丝绳缺陷表面进行捕捉与缺陷识别。

3.2 数据集介绍

本文采用的数据集是曳引钢丝绳缺陷 (traction wire rope defect datasets, TWRD) 数据集。由于工业缺陷数据集稀少, TWRD 数据集中的样本来源于实验室根据真实缺陷模拟自制的 TWRD, 缺陷类型包括 3 种: 钢绳挤出 (extrusion)、表面磨损 (wear)、表面断丝 (broken)。图片尺寸均为 640×640 , 总共拍摄收集了 1 396 张图片, 为 RGB 格式图片。图 8 选取了 TWRD 数据集中两组缺陷进行展示。

考虑到数据集会因样本过少而导致模型鲁棒性不够强, 因此对 TWRD 数据集进行了几何变换比如翻转、旋转、裁剪、变形、缩放以及颜色变换比如添加噪声、色彩空

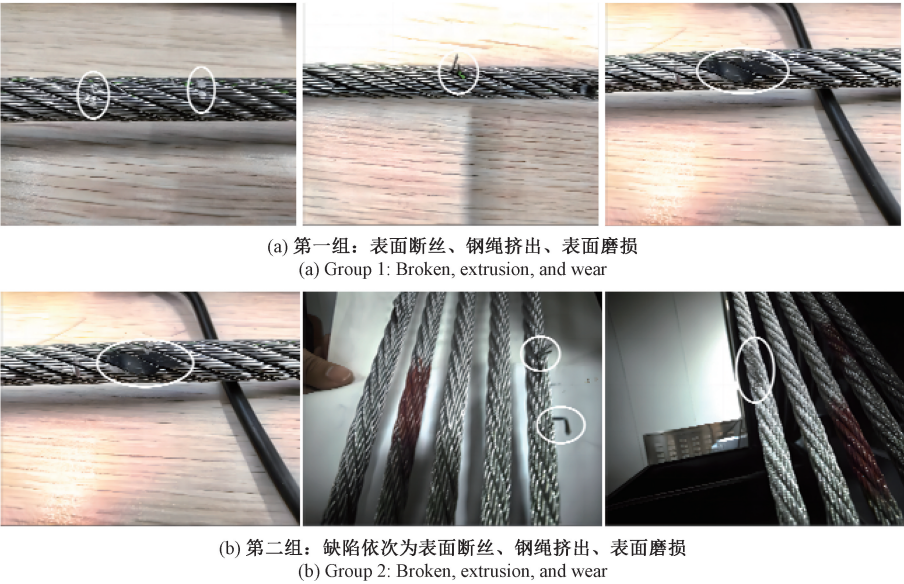


图 8 TWRD 数据集
Fig. 8 TWRD Dataset

间变换等, 扩展到 2 468 张图片。数据集划分比例 7:2:1, 数据集划分如表 1 所示。

表 1 TWRD 数据集划分
Table 1 TWRD dataset partitioning

总共	训练集	验证集	测试集
2 468	1 727	493	248

3.3 实验细节

本实验使用的软件框架是 PyTorch 1. 11. 0+cu113 版本, 并在 Ubuntu18. 04 系统下实验, CPU 是 AMD Ryzen 7 5800H, GPU 使用的是 NVIDIA GeForce RTX 3060 GPU, 运行内存为 12 G。

网络训练采用梯度下降法进行模型优化, 根据经验对超参数进行设置, 其中 Batch Size 设置 8, 初始学习率设为 0. 001, OneCycle 学习率设为 0. 01, 符合 cosine learning rate 和 warm-up 两种方法, 学习率的动量设为 0. 937, 权重衰减系数设为 0. 000 5, warm-up 轮数设为 4 轮, 总共训练迭代 300 轮。

3.4 评价指标

本文采取准确率 (precision)、召回率 (recall)、帧率 (frames per second, FPS) 作为主要的检测指标, 另外还有平均精度均值 (mean average precision, mAP)、参数量 (parameters)、浮点运算量 (floating point operations, FLOPs)。

准确率表示正确检测到的缺陷数占总缺陷的比例, 通过计算检测结果中预测正确的缺陷数和所有预测样本

数的比例得到, 其计算表达式为:

$$Precision = \frac{TP}{TP + FP}$$
 (5)

召回率表示正确检测到的缺陷数占真实缺陷数的比例, 通过计算检测结果中预测正确的缺陷数和所有真实缺陷数的比例得到, 其计算表达式为:

$$Recall = \frac{TP}{TP + FN}$$
 (6)

评判网络性能使用平均精度均值, 综合准确率、召回率, 以准确率为横坐标, 召回率为纵坐标绘制 P-R 曲线, 计算各类别 P-R 曲线下的面积作为类别的 AP, 利用各类别平均精度的均值作为 mAP, 其值越大越好。

参数量表示网络模型中需要训练的参数总数, 是衡量模型复杂度的一个重要指标, 越大越需要更多内存来存储模型参数。

浮点运算数表示运行一次网络模型需要进行浮点运算的次数, 通常衡量模型的计算效率和速度。FPS 表示每秒处理多少帧图像。

3.5 TWRD 数据集上的消融实验

本文是以 YOLOv5-s 模型作为基线进行实验, 在 YOLOv5-s 模型上进行改进。在这一节将会对 YOLOv5-s 网络结构实验结果与本文提出的 TWRD-Net 实验结果进行对比。消融实验结果表 2 所示。

方案 B 与 A 对比, 可以看出 B 方案引入了 LW-C3 模块之后, 尽管对缺陷的检测准确率降低了 3%, 但是参数量减少了 12%, 浮点运算数降低了高达 30%, mAP 有 0. 7% 的提升。

表 2 消融实验对比

Table 2 Comparison of ablation experiments

序号	方案	类别	准确率/%	召回率/%	平均精度/%	参数量/M	帧率/fps	浮点运算量/G
A	YOLOv5-s	总类别	94.5	94.5	97.9	7.01	144	15.8
		挤出	92.4	94.8	97.5			
		磨损	92.7	95.4	96.7			
		断丝	98.6	93.4	99.5			
		总类别	91.5	96.2	98.6			
B	A+LW-C3	挤出	91.8	97.8	98.7	6.11	161	11.0
		磨损	85.0	96.3	97.7			
		断丝	97.9	94.5	99.5			
		总类别	97.1	97.0	99.0			
C	B+CLW-FPN	挤出	97.8	97.8	99.2	6.72	154	12.5
		磨损	97.4	95.9	99.3			
		断丝	96.3	97.3	98.4			
		总类别	98.0	97.4	99.4			
D	C+ β -CIoU	挤出	98.6	97.8	99.4	6.72	151	12.5
		磨损	97.3	95.7	99.4			
		断丝	98.3	98.9	99.5			

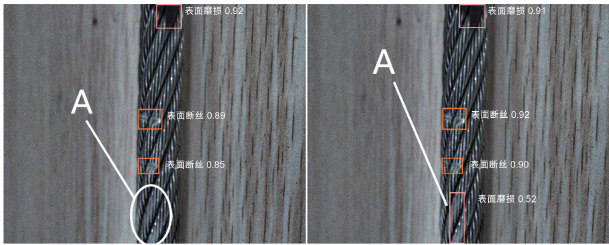
方案 C 与 B 对比,方案 C 在方案 B 基础上用 CLW-FPN 替换了 PAN 结构,从实验结果可以看出,由于增加了上采样以及升降维操作相较于方案 C 提升了部分参数量与浮点运算数,但是仍低于方案 A 的参数量和浮点运算数,对于 mAP 仍有了 0.4% 的增长,通过跨层级的特征融合方式,加强了对不同位置缺陷的敏感度,准确率、召回率也得到了提升。

方案 D 与 C 对比,方案 D 在方案 C 基础上用 β -CIoU 替换了 CIoU,从实验结果可以看出方案 D 与方案 C 参数量、浮点运算量基本持平,mAP 得到进一步提升,最终对于缺陷检测的准确率可达 98.0%,召回率达 97.4%,帧率达到 151 fps。

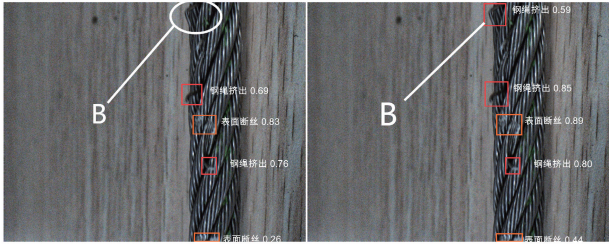
综上所述,观察以上 6 个指标,方案 D 在方案 A 基础上有了更进一步的性能提升,98% 的准确率、97.4% 的召回率、99.4% 的 mAP 以及 151 fps 足够满足需求,并且参数量、浮点运算量均低于方案 A。

3.6 TWRD-Net 与 YOLOv5-s 的检测结果

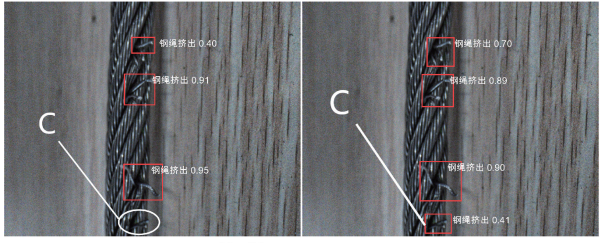
本文对 TWRD-Net 网络与 YOLOv5-s 网络在 TWRD 数据集的测试集上进行验证,TWRD-Net 与 YOLOv5s 检测结果的对比如图 9、10 所示。结合图 9、10 可以看出,在图像边缘位置的 A、B、C 3 处缺陷,YOLOv5-s 网络会出现漏检情况,并且当背景过于复杂、前景不够突出明显,即 D、E、F 3 处位置,YOLOv5-s 也会存在漏检。反观 TWRD-Net 网络能够很好对不同位置的缺陷进行识别与



(a) 第1组YOLOv5, TWRD-Net
(a) Group 1 YOLOv5, TWRD-Net



(b) 第2组YOLOv5, TWRD-Net
(b) Group 2 YOLOv5, TWRD-Net



(c) 第3组YOLOv5, TWRD-Net
(c) Group 3 YOLOv5, TWRD-Net

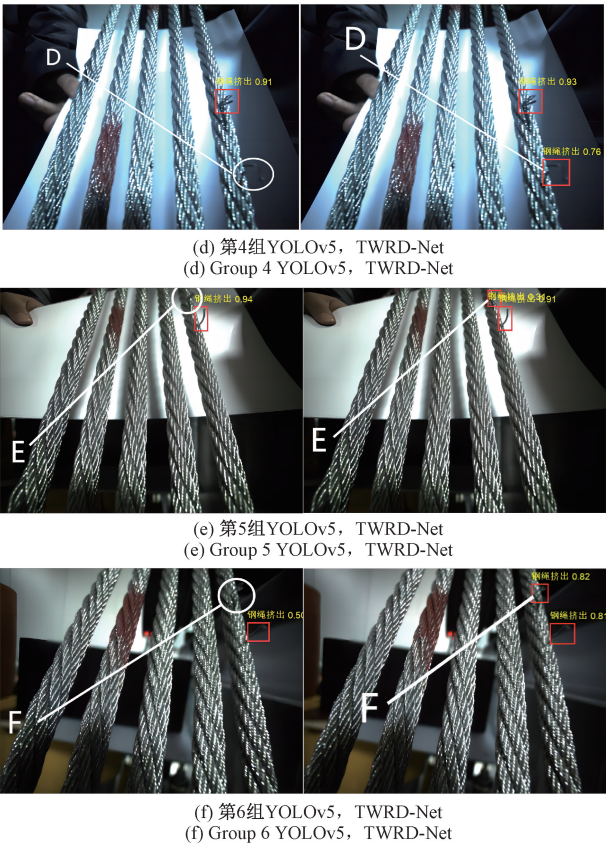


图 9 TWRD-Net 与 YOLOv5 实验结果对比

Fig. 9 Comparison diagram of TWRD-Net and YOLOv5 experimental results collected on site

定位,尤其处于 A、B、C 3 处边缘位置的缺陷,本身处在边缘位置的缺陷会存在遮挡情况,很有可能因为特征提取不全导致被漏检,但是 TWRD-Net 能够检测出,说明对边缘位置的缺陷更加敏感。另外,对于 D、E、F 3 处处于复杂背景下的缺陷也进行了识别,证明通过加权融合来

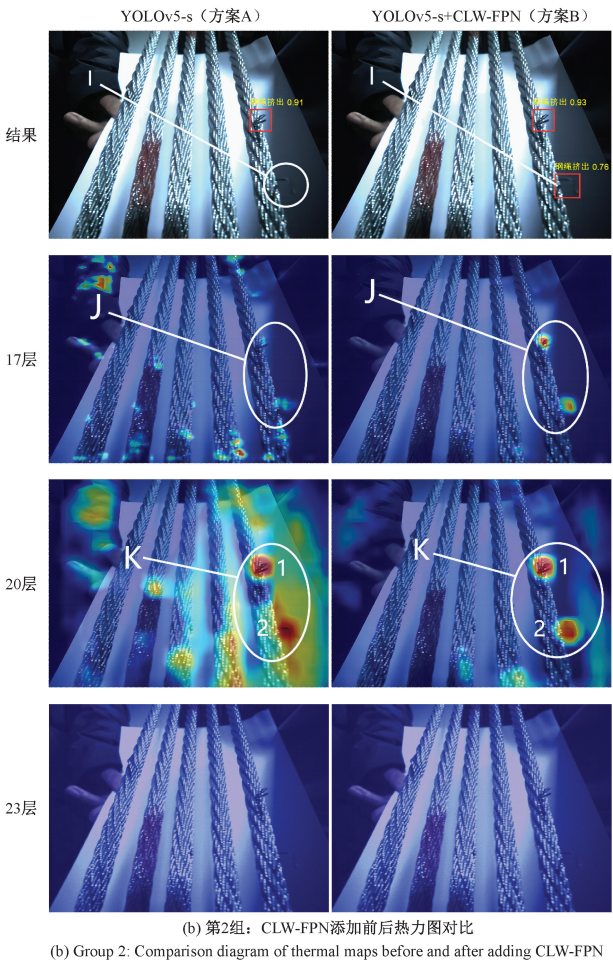


图 10 添加 CLW-FPN 模块前后的热力图效果对比

Fig. 10 Comparison diagram of thermal map effects before and after adding CLW-FPN modules in different scenarios

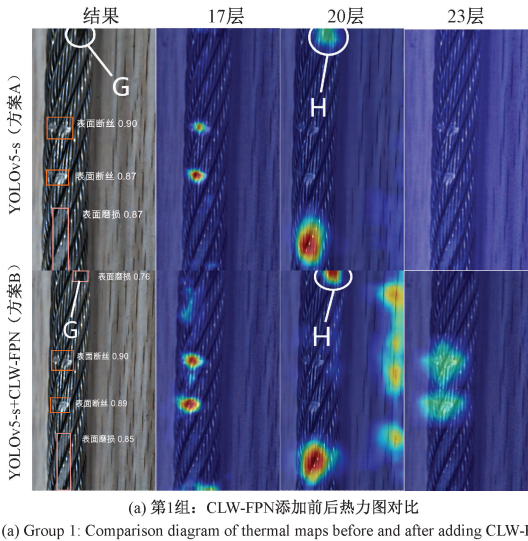
区分不同特征图的重要程度,相较于 YOLOv5 具有较强的背景抗干扰能力。

本文认为 TWRD-Net 之所以能够解决 YOLOv5-s 漏检的原因是 CLW-FPN 结构起了关键作用,通过 CLW-FPN 结构的跨层级加权连接,将更强的语义信息以及更强的位置信息直接作用于各个尺度的特征图之上,因此为了证明 CLW-FPN 的有效性,对于存在漏检情况的 TWRD 图像采用 Grad-CAM 算法^[29]进行了热力图可视化,实验条件如表 3 所示。

表 3 热力图可视化实验条件

Table 3 Experimental conditions for visualization of thermodynamic diagrams

方案	Backbone 主干网络	Neck 网络	Head 检测头
YOLOv5-s(方案 A)	不变	PAN	不变
本文(方案 B)	不变	CLW-FPN	不变



本文从 TWRD 检测结果中选取两组具有代表性的结果进行实验,分别对 YOLOv5-s (方案 A) 和 YOLOv5-s + CLW-FPN (方案 B) 网络的 17 层、20 层、23 层输出 (即对应 80×80 、 40×40 、 20×20 这 3 个尺度检测头的输出) 的热力图进行可视化。观察 YOLOv5-s + CLW-FPN 与 YOLOv5-s 网络各自的感兴趣区域以及关注程度,实验结果如图 11(a)、(b) 所示。

首先需要声明的是,热力图颜色越深的区域代表模型关注程度越高。在图 11(a) 可以看出实际存在 4 处缺陷,方案 A 仅检测出 3 处,并没有检测出 G 处的缺陷,综合 3 层热力图可以发现,针对 G 处仅有第 20 层 (对应 H 处) 有所关注,但是关注程度并不高,然而引入了 CLW-FPN 下的方案 B,与方案 A 相同的关注区域下,却有更强的关注强度,尤其针对 20 层的 H 处,以及 17 层的两处缺陷位置,并且检测出了所有的 4 处缺陷,G 处置信度也能达到 0.76;在图 11(b) 中,实际存在两处缺陷,然而有一

处缺陷,即 I 处,因与背景颜色相近,所以难以分辨。方案 A 的 17、23 层对于缺陷位置并没有很强的关注程度,在 20 层的 K 处对两处缺陷进行了关注,可以看出,对 K 区域内第 1 处与背景有鲜明对比的缺陷能够较好的独立区分,然而针对第 2 处与背景近似的缺陷却难以区分,其关注了第 2 处缺陷以及周边邻域,由此说明方案 A 并没有将缺陷与背景很好的进行区分,抗背景干扰能力较差,相反,观察方案 B 的 17、20 层热力图中的 J、K 区域,方案 B 对两处缺陷位置均有较强的关注程度,与方案 A 不同的是,20 层的 K 区域的第 2 处缺陷位置明显没有了背景的干扰,在 CLW-FPN 的跨层连接以及加权融合作用下,模型有效的去除了背景信息,并且进一步加强了缺陷区域的关注程度,成功检测出了图片中所有的 2 处缺陷。综合以上实验分析可以看出,CLW-FPN 结构加强了模型对于缺陷的定位能力,并且也能够区分有用信息与无用信息,具有一定的抗背景干扰能力,证实了 CLW-FPN 的有效性。

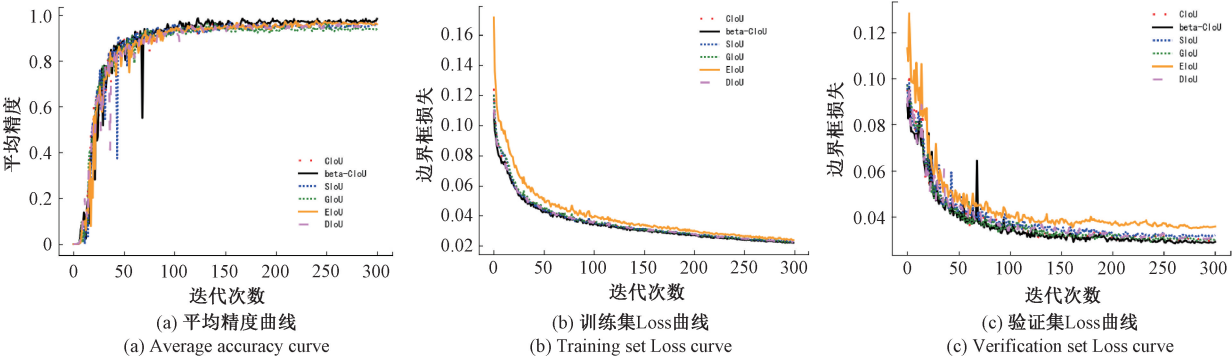


图 11 不同 IoU 损失函数下的 Loss 曲线图与 mAP 曲线
Fig. 11 loss graph and mAP graph under different IoU Loss function

3.7 不同 IoU 损失函数下的 Loss 曲线图

因为 IoU 损失函数仅与边界框损失有关,在这一节中对对比了不同 IoU 损失函数下的边界框损失以及 mAP 曲线图进行对比,如图 12 所示。实验条件以 YOLOv5-s 为基准,仅改变 IoU 损失为 ClIoU、SIoU、beta-ClIoU、EIoU、DIoU、GIoU,迭代次数为 300。通过实验结果可以看出,在 train 训练集下,不同 IoU 损失函数的边界框 Loss 差距并不大,EIoU 在最初 Loss 明显高于其他 IoU 损失函数,但最终也能收敛到较低的损失。但是在 val 验证集下,可以看出不管哪一种 IoU 损失函数,曲线都会出现波动,并且最终 Loss 收敛数值也有所差异。可以发现,本文提出的 beta-ClIoU,即图中黑色实线,相比较 EIoU、SIoU 等有更低的边界框 Loss,说明 beta-ClIoU 能在 TWRD 数据集上进一步降低边界框损失,提升目标定位精度。另外,观察 mAP 曲线图,在 beta-ClIoU 损失函数下能保持最高的数值,说明 beta-ClIoU 可以提高对 TWRD 数据集的检测精度。

3.8 主流 YOLO 模型在 TWRD 数据集的实验结果

为了进一步验证本文提出的 TWRD-Net 网络在 WRD 数据集中具有更大优势,选取了 8 种不同的主流 YOLO 模型算法进行对比试验,其中包含 YOLOv3、YOLOv3-spp、YOLOv4-s、YOLOv5-s、YOLO8-s、YOLOv7-tiny、YOLOv7、YOLOX-s。实验结果如表 4 所示。

由表 4 可以看出,针对 TWRD 的检测,YOLOv8-s、YOLOv7 模型在准确率、召回率、mAP 确实高于 YOLOv5-s,但是差距并不大,对比参数量、浮点运算量,YOLOv8-s、YOLOv7 模型却是 YOLO5-s 的几倍,尤其 YOLOv7 浮点运算量是 YOLOv5-s 的 6 倍之多,针对工程应用场景,精度、准确率等差距不大或者满足需求时,追求体量跟小更快的轻量模型,这也是选择 YOLOv5 系列做为基线改进的原因。

另外,在 YOLOv5-s 基础之上进行了改进得到了本文算法 (TWRD-Net),可以看出,准确率、召回率、mAP 有了提升,并且在准确率、mAP 上超过了 YOLOv8 系列、

表 4 主流 YOLO 模型算法实验结果

Table 4 Experimental results of mainstream YOLO model algorithm

方案	类别	准确率/%	召回率/%	平均精度/%	参数量/M	帧率/fps	浮点运算量/G
YOLOv3	总类别	95.7	93.5	96.7	9.30	109	23.1
YOLOv3-spp	总类别	96.3	95.1	96.5	9.56	105	23.3
YOLOv4	总类别	97.1	91.9	90.7	9.12	113	20.8
YOLOv5-s	总类别	94.5	94.5	97.9	7.01	144	15.8
YoloX-s	总类别	94.6	94.8	96.9	9.67	95	21.9
YOLOv7-tiny	总类别	92.8	93.9	96.5	6.02	125	13.2
YOLOv7	总类别	96.0	94.1	97.6	37.20	83	105.1
YOLOv8-s	总类别	95.2	97.6	98.2	11.12	109	28.4
本文	总类别	98.0	97.4	99.4	6.72	151	12.5

YOLOv7 系列,其召回率仅次于 YOLOv8 系列,参数量仅次于 YOLOv7-tiny,FPS 以及浮点运算量也是位居首位,可以看出改进后的本文算法(TWRD-Net)在 TWRD 上的检测具有更好的性能,本文算法(TWRD-Net)相比其余模型实现了模型精度、帧率、资源占用率的一定平衡。

4 结 论

本文针对曳引钢丝绳缺陷检测的工业应用背景,对 YOLOv5-s 模型的主干网络、Neck 网络结构以及损失函数进行了改进得到本文的 TWRD-Net,并且在 TWRD 数据集上进行了实验,对不同模块的组合、主流 YOLO 框架等的实验结果进行了对比,利用 Grad-CAM 算法进行了热力图可视化。最后利用不同的主流 YOLO 系列框架进行实验与数据对比,本文改进的 TWRD-Net 在 TWRD 检测场景上要更优于其余模型,对复杂环境下的 TWRD 具有更强的鲁棒性,很好的平衡了检测速度和准确度等,实验结果中 98%的准确率以及 151 fps 的帧率证实了深度学习应用于曳引钢丝绳缺陷检测是一种可行性方案。

参考文献

[1] ZHOU P, ZHOU G, ZHU Z, et al. A review of non-destructive damage detection methods for steel wire ropes[J]. Applied Sciences, 2019, 9(13): 2771.

[2] MAZUREK P, ROSKOSZ M, KWAŚNIEWSKI J. Novel diagnostic of steel wire rope with passive magnetic methods[J]. IEEE Magnetics Letters, 2021, 13: 1-5.

[3] TIAN J, ZHAO C, WANG W, et al. Detection technology of mine wire rope based on radial magnetic vector with flexible printed circuit[J]. IEEE Transactions

on Instrumentation and Measurement, 2021, 70: 1-10.

[4] REN L, LIU Z, ZHOU J. Shaking noise elimination for detecting local flaw in steel wire ropes based on magnetic flux leakage detection [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-9.

[5] LAGUERRE L, TREYSEDE F. Non destructive evaluation of seven-wire strands using ultrasonic guided waves[J]. European Journal of Environmental and Civil Engineering, 2011, 15(4): 487-500.

[6] RAIŠUTIS R, KAŽYS R, MAŽEIK A, et al. Ultrasonic guided wave-based testing technique for inspection of multi-wire rope structures[J]. NDT & E International, 2014, 62: 40-49.

[7] BAI W, CHAI M, LI L, et al. Acoustic emission from elevator wire ropes during tensile testing[C]. Advances in Acoustic Emission Technology: Proceedings of the World Conference on Acoustic Emission-2013. Springer New York, 2015: 217-224.

[8] ZEJLI H, GAILLET L, LAKSIMI A, et al. Detection of the presence of broken wires in cables by acoustic emission inspection[J]. Journal of Bridge Engineering, 2012, 17(6): 921-927.

[9] ZHOU P, ZHOU G, HE Z, et al. A novel texture-based damage detection method for wire ropes [J]. Measurement, 2019, 148: 106954.

[10] PLATZER E S, SÜBE H, NÄGELE J, et al. On the suitability of different features for anomaly detection in wire ropes[C]. Computer Vision, Imaging and Computer

- Graphics, 2010; 296-308.
- [11] WEN W, CHANG-YUN M, JI W, et al. Defect Recognition of X-ray steel rope cord conveyer belt image based on bp neural network [C]. 2011 International Symposium on Computer Science and Society, IEEE, 2011; 168-171.
- [12] LIU S, SUN Y, JIANG X, et al. A new MFL imaging and quantitative nondestructive evaluation method in wire rope defect detection[J]. Mechanical Systems and Signal Processing, 2022, 163: 108156.
- [13] WU P, LONG Z, LI G, et al. Research on deformation detection of traction wire rope based on non-contact[C]. 2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA), IEEE, 2022; 324-327.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [15] 张银胜,杨宇龙,吉茹,等.改进YOLOv5s的风力涡轮机表面缺陷检测[J].电子测量与仪器学报,2023, 37(1):40-49.
- ZHANG Y SH, YANG Y L, JI R, et al. Surface defect detection of wind turbine based on YOLOv5s[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 40-49.
- [16] 刘兰兰,万旭东,汪志刚,等.基于超分辨率重建与多尺度特征融合的输电线路缺陷检测方法[J].电子测量与仪器学报,2023, 37(1):130-139.
- LIU L L, WAN X D, WANG ZH G, et al. Transmission line defect detection method based on super-resolution reconstruction and multi-scale feature fusion[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1):130-139.
- [17] 陈仁祥,詹赞,胡小林,等.基于多注意力Faster RCNN的噪声干扰下印刷电路板缺陷检测[J].仪器仪表学报,2021,42(12):167-174.
- CHEN R X, ZHAN Z, HU X L, et al. An automatic detection method of aero-engine bolt installation defects based on key points detection [J]. Chinese Journal of Scientific Instrument, 2021, 42(12):167-174.
- [18] 刘艳菊,王秋霁,赵开峰,等.基于卷积神经网络的热轧钢条表面实时缺陷检测[J].仪器仪表学报,2021, 42(12):211-219.
- LIU Y J, WANG Q J, ZHAO K F, et al. Real-time defect detection of hot rolling steel bar based on convolution neural network [J]. Chinese Journal of Scientific Instrument, 2021, 42(12):211-219.
- [19] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [20] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [21] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [22] ZHENG Z, WANG P, LIU W, et al. Distance-IoU Loss: Faster and better learning for bounding box regression[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [23] GEVORGYAN Z. SIoU Loss: More powerful learning for bounding box regression [J]. ArXiv Preprint, 2022, ArXiv:2205.12740.
- [24] 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.
- [25] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [26] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv Preprint, 2017, ArXiv: 1704.04861.
- [27] DING X, GUO Y, DING G, et al. Acnet: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks [C]. Proceedings of the IEEE/CVF

International Conference on Computer Vision, 2019: 1911-1920.

[28] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.

[29] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.

作者简介



高嘉,2021 年于华北科技学院获得学士学位,现于哈尔滨工程大学攻读硕士学位,主要研究方向为深度学习、人工智能。
E-mail: 944468183@qq.com

Gao Jia received his B. Sc. degree from

North China Institute of Science and Technology in 2021, and pursuing a master's degree at Harbin Engineering University. His main research interests include deep learning and artificial intelligence.



刘涛(通信作者),分别在 2002 年,2005 年和 2009 年于哈尔滨工程大学获得学士、硕士和博士学位,现为哈尔滨工程大学副教授,主要研究方向为工业缺陷检测、工业机器视觉技术与应用。

E-mail: liutao@hrbeu.edu.cn

Liu Tao (Corresponding author) received his B. Sc. , M. Sc. , and Ph. D. degrees all from Harbin Engineering University in 2002, 2005, and 2009, respectively. He is currently an associate professor at Harbin Engineering University. His main research interests include industrial defect detection, industrial machine vision technology and applications.