

DOI: 10.13382/j.jemi.B2407532

# 面向特钢车间内物料实时跟踪的钢管目标检测算法研究

赵云涛<sup>1,2</sup> 黄哲辉<sup>1</sup>

(1. 武汉科技大学信息科学与工程学院 武汉 430081; 2. 武汉科技大学冶金自动化与检测技术教育部工程研究中心 武汉 430081)

**摘要:**在特钢企业向“灯塔工厂”转型升级中,实现钢管物料实时跟踪是其中的核心内容,由于物料多样性以及产线的复杂性使得接近式传感器无法满足物料检测可靠性要求。为此,依据车间内现有环境和需求,搭建物料跟踪摄像系统,采集了物料及产线上部分特征组成图像数据集;基于视频分析,引入了一种面向特钢车间内物料实时跟踪的钢管目标检测算法。该算法以PPYOLOE网络为基础,首先,将PPYOLOE中的CSPRepResNet主干网络替换成HGNetV2轻量级主干网络,在提升特征提取能力同时减小参数量;其次,在Neck中融合HG-Block和SPPELAN进一步减小参数提升速度;最后,在上采样阶段,运用Dysample动态上采样算子提升不同尺度特征的融合效果,提升算法的检测精度。实验结果表明,相比于原始的PPYOLOE算法,改进后的算法在检测精度上提升了1.6%达到80.5%,检测速度提升了16%达到56.4帧,GFLOPs和参数分别下降35%和33%。改进后算法有效提升了检测精度和检测速度,通过现场部署实施,满足了钢管物料实时跟踪要求。

**关键词:**钢管目标检测;PPYOLOE;HGNetV2;SPPELAN;Dysample

**中图分类号:** TP399; TN06 **文献标识码:** A **国家标准学科分类代码:** 520.2060

## Research on steel pipe target detection algorithm for real-time material tracking in special steel workshop

Zhao Yuntao<sup>1,2</sup> Huang Zhehui<sup>1</sup>(1. School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China;  
2. Engineering Research Center for Metallurgical Automation and Detecting Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** In the transformation and upgrade of special steel enterprises into “lighthouse factories”, real-time tracking of steel pipe materials is a core component. Due to the diversity of materials and the complexity of the production line, proximity sensors fail to meet the reliability requirements of material detection. Therefore, according to the existing environment and requirements of the workshop, a material tracking camera system is built, and the image data set composed of some characteristics of materials and production lines is collected. Based on video analysis, a steel pipe target detection algorithm for real-time material tracking in special steel workshops is introduced. The algorithm is based on the PPYOLOE network. Firstly, the CSPRepResNet backbone in PPYOLOE is replaced with the lightweight HGNetV2 backbone, which enhances feature extraction capabilities while reducing the number of parameters. Secondly, HG-Block and SPPELAN are integrated into the Neck, further reducing the parameters and improving speed. Finally, in the upsampling stage, the Dysample dynamic upsampling operator is employed to enhance the fusion of multi-scale features, thus improving detection accuracy. Experimental results show that compared with the original PPYOLOE algorithm, the improved algorithm enhances detection accuracy by 1.6%, reaching 80.5%, and increases detection speed by 16%, reaching 56.4 FPS, while GFlops and parameters are reduced by 35% and 33%, respectively. The improved algorithm effectively boosts both detection accuracy and speed, and through on-site deployment, it meets the real-time tracking requirements of steel pipe materials.

**Keywords:** steel pipe detection; PPYOLOE; HGNetV2; SPPELAN; Dysample

## 0 引言

随着全球制造业竞争的日益激烈,“灯塔工厂”以其智能化、数字化的特性,成为了引领未来制造业的新标杆<sup>[1]</sup>。通过数字化生产线、智能化管理系统以及精益化生产模式等先进措施,“灯塔工厂”不仅提升了生产效率和产品质量,还为企业实现了更可持续的发展。在钢铁产业领域,实现物料的数字化跟踪<sup>[2]</sup>是“灯塔工厂”建设的核心内容。

此前在特钢车间内的物料跟踪主要依赖于位置传感器定位方案。这些传感器安装在生产线上的关键位置,以捕获钢管在运输过程中的位置和状态信息,并将数据传输到中央处理单元进行分析和处理。然而,钢管的多样性以及生产线上的复杂转运过程给物料跟踪带来了极大的挑战。钢管的不同型号和尺寸也使得传感器信号的稳定性受到了严重威胁,同时,生产线上不确定性因素和人工干预使物料跟踪无法实现。在电子、汽车、造船等制造行业中实现产线的物料跟踪有多种方案。唐全军<sup>[3]</sup>应用条形码或二维码技术,在车载电子标签生产车间对产品和物料进行追踪。Yang 等<sup>[4]</sup>应用射频识别(radio frequency identification, RFID)技术将物料与互联网连接进行通信,实现物料的识别、定位、追踪。Huang 等<sup>[5]</sup>结合 RFID 区域定位和超宽带技术(ultra wide band, UWB)精确定位获得车间全部资源的位置信息。然而在特钢车间内,钢管表面不平整且转运过程中姿态不固定无法对二维码和条形码进行印刷和识别。同时产线上钢管之间的碰撞和切削、打磨等精加工操作,导致无法通过安装 RFID 定位装置进行跟踪。本文面向特钢车间内物料跟踪,通过引入基于视频分析的目标检测算法,实现对特钢车间内钢管物料的跟踪和监控。

目前基于深度学习的目标检测算法可分为单阶段和两阶段式。由于两阶段式模型比较复杂、参数量大、计算难度高导致识别速度慢一直无法实现实时检测的要求。而单阶段算法在产生候选区域的同时即可预测出物体类别和位置,在保证精度的前提下,速度到达实时性要求。在单阶段算法中,Redmon 等<sup>[6]</sup>首次提出了 YOLOv1 算法并在后续改进中提出了 YOLOv2、YOLOv3 等模型。在后续的研究中不断有其他学者对模型的速度和精度进行提升推出了一系列后续模型,其中具有代表性的有 YOLOv5、YOLOv6、YOLOv8、YOLOX 等。YOLO 系列目标检测算法已经在工业生产和生活中得到广泛的应用<sup>[7-8]</sup>。

百度团队在 PP-YOLOv2<sup>[9]</sup>的基础上通过引入无锚机制推出了 PPYOLOE<sup>[10]</sup>算法,实现对边界框的灵活精确预测,提升了复杂环境下的检测精度。为满足特钢车间内对目标检测精度需求和高密集推理任务下的速度要

求。本文提出了一种改进 PPYOLOE-s 目标检测算法。该方法在 PPYOLOE-s 的基础上使用 HGNetV2 主干网络替换原模型中的主干;同时融合 HG-Block 和 SPPELAN 提升模型的特征提取能力和速度;最后,使用 Dysample 动态上采样算子,增强模型对不同尺度目标的检测能力,提升检测精度。实验结果表明,改进后的模型精度优于原模型,模型参数小于原模型。

## 1 PPYOLOE 算法及改进

### 1.1 PPYOLOE 算法

为满足复杂多变的视觉任务需求,PPYOLOE 模型展现了一种可配置与优化的网络架构设计方法。该模型通过调控网络深度和宽度参数,定义了 s、m、l、x 共 4 个不同的配置型号,其中 s 型号参数最小推理速度最快,适用于对实时性要求较高的多线程密集计算场景,其结构如图 1 所示。

PPYOLOE 在 Backbone 设计上,融合了 CSPNet 的跨阶段部分连接策略<sup>[11]</sup>与 ResNet 的残差学习思想<sup>[12]</sup>和密集连接<sup>[13]</sup>构建了一种新型的 RepResBlock 模块。在推理阶段重参数化为 CSPRepResStage,结合 ESE(efficient spatial pyramid of attention)注意力机制,形成了增强的 CSPRepResNet 主干网络。这一设计不仅加速了网络收敛,还显著提升了特征提取的深度与广度,增强了模型对复杂场景的理解能力。在特征聚合的 Neck 部分,PPYOLOE 采用空间金字塔(spatial pyramid pooling, SPP)与路径聚合网络结构(path aggregation network, PAN)的组合。这种设计策略有效促进了多尺度特征图之间的信息交流与融合,增强了特征的表达能力。在 Head 部分,PPYOLOE 对 TOOD(task-aligned one-stage object detection)框架<sup>[14]</sup>进行了优化。它利用 ESE 模块替代了原有的层关注机制,简化了分类分支的结构。

同时引入了任务对齐学习(task alignment learning, TAL)标签分配策略,通过动态调整样本分配权重,实现了分类与回归任务的最优对齐,提升了分类与回归任务的性能。这些改进不仅简化了网络结构,在提升检测精度的同时,也保证了模型的鲁棒性。

### 1.2 改进 PPYOLOE 算法

为了进一步提升模型对钢管等目标检测精度同时增强检测速度,使用轻量级主干 HGNetV2 替换原模型的主干加强特征提取能力减小模型参数;同时在 Neck 中融合 HG-Block 和 SPPELAN,进一步减小模型参数提升检测速度;最后,使用 Dysample 动态上采样算子保留高层特征的语义信息的同时更加丰富底层特征,增强模型对不同尺度目标的检测效果。改进后模型的结构如图 2 所示。

1) 轻量级主干网络

HGNet 系列是由百度团队提出的一种神经网络架构。HGNet 主要思想是利用层次化的方法来提取特征，其中复杂的模式可以在不同的规模和抽象层次上学习，提高网络处理复杂图像数据的能力。

HGNetV2<sup>[15]</sup> 针对 GPU 设备，对目前 GPU 友好的网络做了分析和归纳，尽可能多的使用  $3\times 3$  标准卷积（计算密度最高），HG-Block 是 HGNetV2 的基本组成单元如图 3 所示，由多个 HG-Block 串联组成 HG-Stage。其中 LightConvBNAct 模块由两个 ConvBNAct（卷积层+批量归一化+激活函数）组成，第 1 个 ConvBNAct 卷积核大小为

1，只增加特征图的通道数不改变特征图大小，并且不应应用激活函数。第 2 个 ConvBNAct 通过卷积核大小为  $3\times 3$  的卷积来进行实际的特征提取，减小特征图高宽，不改变通道数。该模块进一步加深了模型的深度加强了对特征的提取能力。

HGNetV2 在 ConvBNAct 后引入了 LAB (learnable-affineblock) 模块，LAB 模块是一个可学习的仿射变换模块，添加在激活函数后，包含一个可学习的缩放参数 (scale) 和一个可学习的偏置参数 (bias) 对输入张量进行缩放和偏置来实现仿射变换。一定程度上进行了信息的传递和调整，有助于解决梯度消失和梯度爆炸等问题。

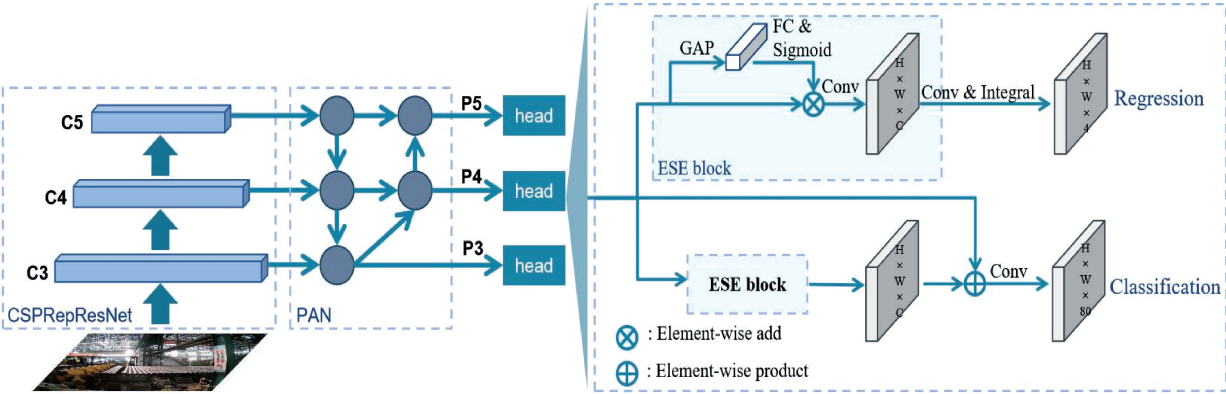


图 1 PPYOLOE 网络结构  
Fig. 1 PPYOLOE network structure

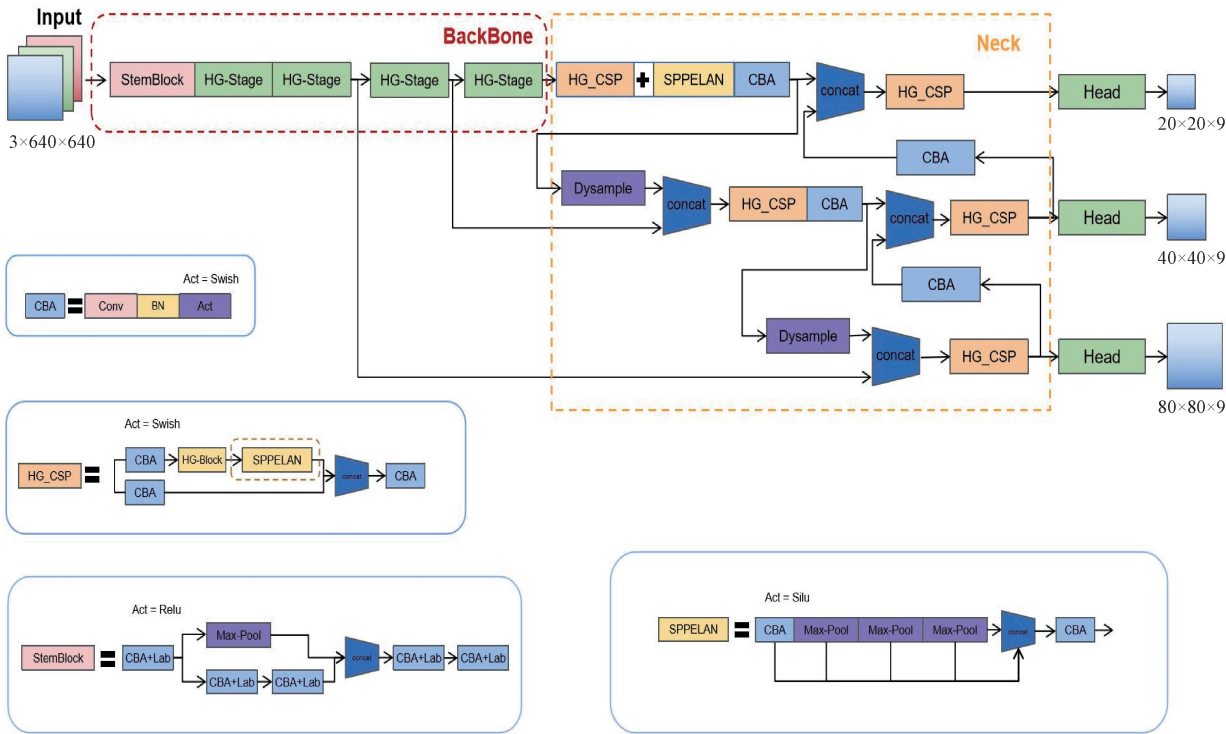


图 2 改进后模型结构  
Fig. 2 Improved model structure

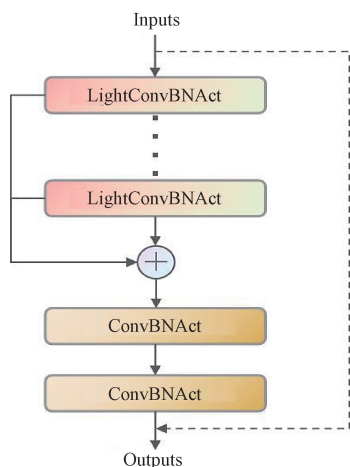


图 3 HG-Block 结构图

Fig. 3 HG-Block structure diagrams

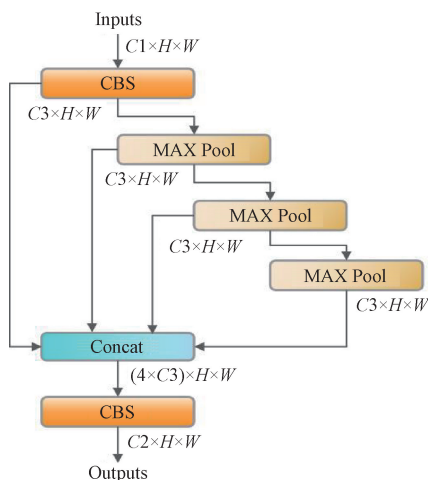


图 4 SPPELAN 结构

Fig. 4 SPPELAN structure

## 2) SPPELAN 特征金字塔

在目标检测中, SPP 模块的引入是为了丰富特征图的表达能力, 尤其是针对多目标检测任务中目标大小差异较大的情况。SPP 模块借鉴了空间金字塔的思想, 通过局部特征与全局特征的融合, 增强了模型对不同尺寸目标的检测能力。在 SPP 模块中, 最大的池化核大小要尽可能接近或者等于需要池化的特征图的大小, 以确保有效融合局部特征和全局特征。因此, SPP 模块的引入使得模型在复杂的多目标检测任务中实现了更高的精度。

SPPELAN 在 YOLOv9 中首次提出, 其结构如图 4 所示。其中 CBS 为卷积层 (conv) + 批量归一化 (batchnorm) + 激活函数 (silu),  $C1$  为输入特征的通道数,  $C2$  为输出特征的通道数,  $C3$  为中间层的通道数, 本文将中间层通道数设置为输入和输出之和的  $1/4$  即:

$$C3 = (C1 + C2) / 4 \quad (1)$$

对于输入的特征图, 先进行 CBS 操作将通道数调整为  $C3$ , 然后经过 3 个串联的最大池化层, 并将每个池化层操作后的输出和第一个 CBS 层的输出在通道维度上相加得到大小为  $(4 \times C3) \times H \times W$  的特征图。最后经过一层 CBS 将通道数调整为输出大小。SPPELAN 通过采用最大池化层的串联方式, 进一步减小了参数量。

## 3) DySample 动态上采样算子

DySample<sup>[16]</sup> 是一种高效且轻量级的动态上采样方法, 采用点采样技术以减少计算开销和延迟。该方法避免了在生成动态核时使用耗时的动态卷积和附加的子网络, 从而有效降低了整体的计算量。与传统基于内核的动态上采样器相比, DySample 在参数数量、浮点运算次数、GPU 内存占用以及延迟方面均有显著降低。在包括语义分割、目标检测等任务中, DySample 展示了优于现有上采样器的性能表现。

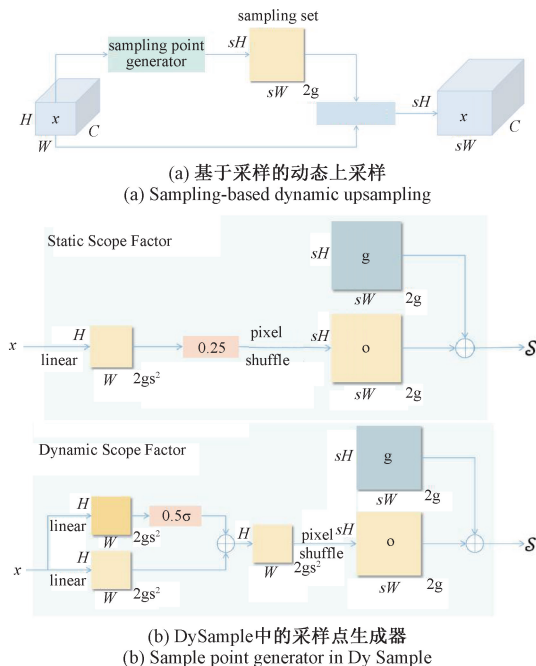


图 5 DySample 中基于采样的动态上采样及模块设计

Fig. 5 Dynamic upsampling and module design based on sampling in DySample

在图 5(a) 中展示了基于采样的动态上采样过程。对于尺寸为  $C \times H \times W$  的输入的特征  $x$ , 通过采样点生成器 (sampling point generator) 生成大小为  $2 \times H_2 \times W_2$  采样集  $\mathcal{S}$ , 利用采样集  $\mathcal{S}$  对输入特征  $x$  进行双线性插值采样得到大小为  $C \times H_2 \times W_2$  的输出  $x'$ , 整个过程可定义为:

$$x' = \text{grid\_sample}(x, \mathcal{S}) \quad (2)$$

在 DySample 中有两种静态和动态两种采样点生成器, 如图 5(b) 所示。

静态:通过线性层与像素洗牌技术<sup>[17]</sup>的结合生成固定的偏移量 $\mathcal{O}$ 。这个偏移量随后被添加到原始网络 $\mathcal{G}$ 上,以形成采样集 $\mathcal{S}$ 。这种方法利用了静态范围因子在构建偏移时的固定性,从而简化了采样过程,同时保持了对空间信息的有效利用。

动态:在静态的基础上,动态范围因子通过在线性层和像素洗牌技术中引入可变化的参数来进一步优化。首先,使用 Sigmoid 函数( $\sigma$ )生成一个动态的范围因子。该范围因子随后用于调整偏移量 $\mathcal{O}$ ,从而允许偏移量适应不同的输入特征和需求。整个过程可定义为:

$$\mathcal{O} = linear(\chi)$$

(3)

$$\mathcal{S} = \mathcal{G} + \mathcal{O}$$

(4)

本文采用静态采样点生成器生成采样集 $\mathcal{S}$ 。

2 实验与结果分析

2.1 实验环境及参数设置

1) 本实验环境如表 1 所示。

表 1 实验环境

Table 1 Experimental environment

| 项目       | 描述                          |
|----------|-----------------------------|
| 处理器(CPU) | Intel Core i9-12900         |
| 显卡(GPU)  | NVIDIA GeForce RTX3090 24 G |
| CUDA     | 11.7                        |
| 操作系统     | Windows11                   |
| 框架       | PaddlePaddle                |
| 编程语言     | Python                      |

2) 实验参数设置

训练时采用随机梯度下降策略,其中动量为 0.9,权重衰减为 0.005,并采用余弦学习率策略,总轮次为 200 轮,基本学习率为 0.01,批次为 8。

2.2 数据集建立

本文数据采集自特钢车间内真实生产环境中的 17 个摄像头,共计 2 335 张照片,包含钢管、吊车、翻钢机和其他标注 4 个类别,按照 8 : 2 的比例划分训练集和测试集,组成 Steel 数据集,数据集目标分布如图 6 所示。

对于数据集内的钢管目标,由于钢管长度过长并且摄像头视野有限无法完整的包含整根钢管,故对钢管的端部进行标注。其中钢管目标为主要部分占比达 84.2% 有 28 066 个目标。此外,还包含了实际生产中需要识别定位的吊车、翻钢机等标注类别。

鉴于数据集规模有限,可能引致模型过拟合或欠拟合,因此采取 RandomDistort 数据增强策略对数据集进行扩展。RandomDistort 数据增强策略可随机扰动图片亮度、对比度、饱和度和色相,模拟不同光照和天气条件,如图 7 所示。有助于模型车间内极端光照、灰尘、烟雾等情况。

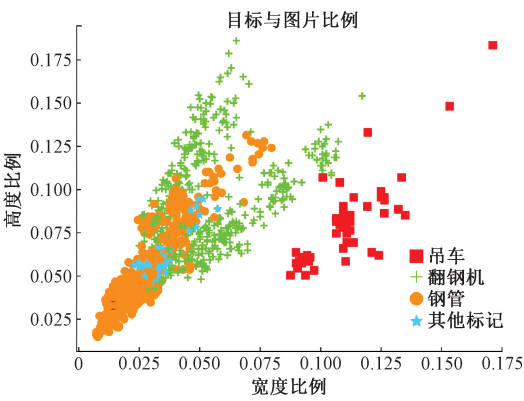


图 6 数据集目标分布

Fig. 6 Target distribution of data set

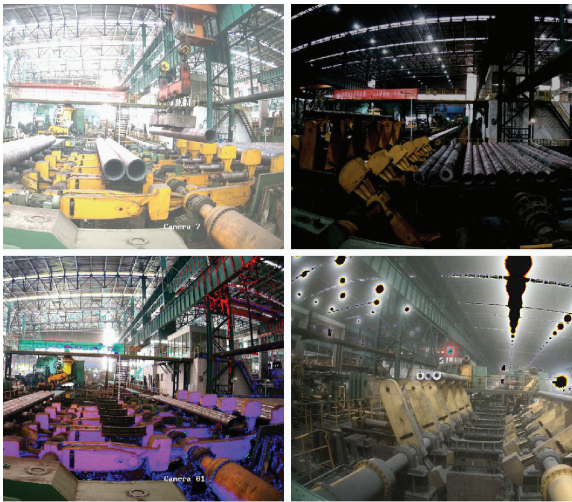


图 7 RandomDistort 数据增强

Fig. 7 RandomDistort data enhancement

2.3 评价指标

本文采用平均精度均值(mean average precision,  $mAP$ )、模型参数大小( $P/MB$ )、每秒十亿次浮点操作(giga floating-point operations per second, GFLOPs)和帧数(FPS)性能指标对模型进行评估。

精确率(Precision,  $P$ )表示正确预测的正样本占所有被检出样本的比例。 $mAP$  是平均精确率(average precision,  $AP$ )的均值,用来衡量模型在不同召回率下的精度表现。 $P$  和  $AP$  公式如式(5)、(6)所示。

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

(5)

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

(6)

其中, $TP$  为正样本预测为正的数,  $FP$  为正样本预测为负的数量,  $FN$  为负样本预测为负的数量。

模型参数大小  $P$  单位为 MB,精度为 32 位浮点数,包

含了模型的权重和偏置等参数值,以及与之相关的一些元数据,如参数名称、形状等。

GFLOPs 通常用于衡量神经网络模型的计算复杂度和计算资源的需求,以及在特定硬件上运行该模型的性能。

FPS 为在保留模型后处理检测一张分辨率为 640×640 大小图片的情况下,先进行 50 次预热再重复推理 300 次所计算出的平均值。

### 2.3 改进实验

#### 1) 轻量级主干对比实验

为提升 PPYOLOE 的检测精度并减小模型参数,本文使用 HGNetV2、PPLCNetV2<sup>[18]</sup>、GhostNet<sup>[19]</sup>、MobileNetV3、MobileOne<sup>[20]</sup>、MobileViTV3<sup>[21]</sup>、ShuffleNetV2<sup>[22]</sup> 目前较先进的轻量级主干网络替换原模型中的 CSPRepResNet 主干网络进行实验,从中选择出最适合本文应用场景下的主干网络,实验结果如表 2 所示。训练的同时每隔 10 轮对模型进行一次评估,并记录 *mAP* 最高的模型,训练时模型 *mAP* 变化曲线如图 8 所示。

表 2 轻量级主干对比实验

| Table 2 Lightweight trunk comparison experiment |               |              |          |      |
|-------------------------------------------------|---------------|--------------|----------|------|
| BackBone                                        | <i>mAP</i> /% | <i>P</i> /MB | GFLOPs   | FPS  |
| CSPRepResNet                                    | 78.9          | 29.2         | 16.230 2 | 48.6 |
| HGNetV2                                         | 79.5          | 25.8         | 13.110 6 | 51.5 |
| PPLCNetV2                                       | 78.4          | 25.7         | 12.333 0 | 54.1 |
| GhostNet                                        | 79.1          | 19.4         | 7.654 7  | 59.4 |
| MobileNetV3                                     | 78.1          | 20.3         | 7.932 8  | 54.1 |
| MobileOne                                       | 76.8          | 24.2         | 11.795 6 | 56.6 |
| MobileViTV3                                     | 77.9          | 19.8         | 8.573 2  | 49.4 |
| ShuffleNetV2                                    | 78.6          | 20.6         | 9.348 4  | 60.9 |

#### 2) HG-Block 融合

根据主干实验的结果表明,HGNetV2 主干网络在该环境下具备更强的特征提取能力同时参数更少更加轻量

表 4 消融实验

| Table 4 Ablation experiment |          |         |          |          |               |      |
|-----------------------------|----------|---------|----------|----------|---------------|------|
| Model                       | HG-Block | SPPELAN | Dysample | GFLOPs   | <i>mAP</i> /% | FPS  |
| PPYOLOE-s                   | —        | —       | —        | 16.230 2 | 78.9          | 48.6 |
| HG+PPYOLOE-s                | —        | —       | —        | 13.110 6 | 79.5          | 51.5 |
| Ours                        | ✓        | —       | —        | 10.570 8 | 79.4          | 61.2 |
| Ours                        | ✓        | ✓       | —        | 10.526 6 | 79.4          | 62.1 |
| Ours                        | ✓        | ✓       | ✓        | 10.541 4 | 80.5          | 56.4 |

由表 4 中可以看出原始 PPYOLOE 网络在 Steel 钢管数据集上的 *mAP* 为 78.9%,GFLOPs 为 16.230 2,FPS 为 48.6。在替换 HGNetV2 主干网络后 *mAP* 提升了 0.6%达

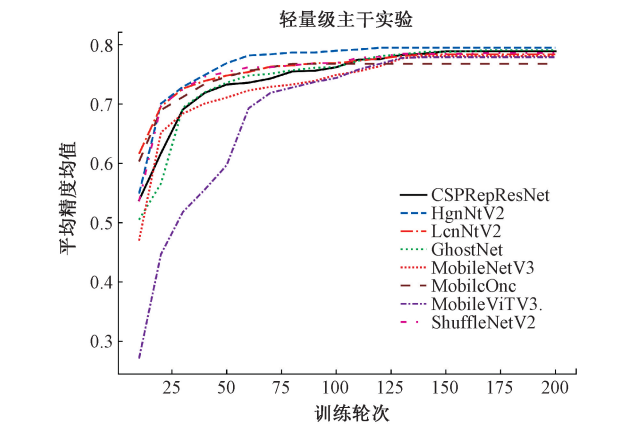


图 8 轻量级主干实验 *mAP* 变化曲线

Fig. 8 *mAP* change curve of lightweight trunk experiment

化,因此将 HGNetV2 中的 HG-Block 模块融入到 Neck 中,加强特征提取能力的同时进一步减小模型参数。在融入 HG-Block 时也对 HG-Block 中的 LightConvBNAct 层数(Layer\_num)进行实验,实验结果如表 3 所示。

表 3 HG-Block 融合实验

| Table 3 HG-Block fusion experiment |               |              |          |      |
|------------------------------------|---------------|--------------|----------|------|
| Layer_num                          | <i>mAP</i> /% | <i>P</i> /MB | GFLOPs   | FPS  |
| 1                                  | 79.2          | 19.6         | 10.570 8 | 61.2 |
| 2                                  | 79.2          | 19.8         | 10.655 2 | 60.1 |
| 4                                  | 79.3          | 20.2         | 10.824 2 | 56.9 |
| 6                                  | 79.6          | 20.6         | 10.993 2 | 55.7 |

验结果表明融合 HG-Block 后模型精度有一定的提升,但是单纯的 LightConvBNAct 层数堆叠对精度提升并不明显,为了兼顾模型的精度和速度,选择一层 LightConvBNAct 的 HG-Block 融合到 Neck 中。

#### 3) 消融实验

为了验证改进后模型的性能,以 PPYOLOE-s 为基础,在自制的 Steel 数据集上进行消融实验,实验结果如表 4 所示。

到 79.5%,GFLOPs 减小了 19.2%达到 13.110 6,FPS 提升了 6%达到 51.5,验证了 HGNetV2 在此检测环境下的有效性。在融合 HG-Block 和 SPPELAN 后 *mAP* 虽略有

下降减小了 0.1%,但模型的 GFLOPs 大幅减小,FPS 大幅提升至 62.2。验证了 HG-Block 的有效性和轻量性。在使用 Dysample 动态上采样算子后,进一步提升了模型的检测能力, $mAP$  提升至 80.5%,但速度较之前的改进有所下降达到 56.4。实验结果表明在加入 HGNetV2、HG-Block、SPPELAN 和 Dysample 后, $mAP$  较原始模型提升了 1.6%,GFLOPs 减小了 35%,FPS 提升了 16%,有效提升了模型对钢管及其他目标的检测能力同时提升了模

型大规模部署实施的检测速度。

如图 9 所示为检测效果对比图,对比 PPYOLOE 原模型检测效果图 9(b)和改进后模型检测效果图 9(c),图 9(b)中的两张图存在未检测出吊车、翻钢机目标,且对于密集放置的钢管检测效果不好。在图 9(c)中对各种目标均能正确识别。同时对比改进前目标置信度均有提升,进一步验证了模型的稳定性。



图 9 检测效果对比

Fig. 9 Comparison of detection effect

表 5 各算法性能对比

Table 5 Performance comparison of algorithms

| Model     | $mAP/\%$ | $P/\text{MB}$ | GFLOPs   | FPS  |
|-----------|----------|---------------|----------|------|
| YOLOX-s   | 71.5     | 34.2          | 26.701 2 | 46.3 |
| YOLOv5-s  | 76.6     | 26.9          | 15.922 2 | 59.5 |
| YOLOv6-s  | 71.8     | 76.9          | 49.356 8 | 40.6 |
| YOLOv8-s  | 75.4     | 42.5          | 28.595   | 59.8 |
| PPYOLOE-s | 78.9     | 29.2          | 16.230 2 | 48.6 |
| Ours      | 80.5     | 19.4          | 10.541 4 | 56.4 |

2.4 其他算法对比

为验证本文所改进的 PPYOLOE 算法的综合性能,选择 YOLOX-s、YOLOv5-s、YOLOv6-s、YOLOv8-s 目标检测算法进行对比。在 Steel 钢管数据集下使用相同的训练参数进行实验,对模型的  $mAP$ 、参数大小( $P$ )、GFLOPs 和 FPS 进行比较,评估模型性能,对比结果如表 5 所示。

由实验结果可知,本文所改进的 PPYOLOE 算法与当前较主流的其他算法相比精度有优势,模型参数为对比模型中最小。在速度方面,略慢于 YOLOv5-s 和 YOLOv8-s,精度分别高 3.9%和 5.1%。与原模型相比,速度和精度均有提升。

### 3 结 论

面向特钢车间内物料实时跟踪,为应对复杂产线上物料的检测问题,基于视频分析采用目标检测算法实现物料的实时检测。通过引入 PPYOLOE 深度学习目标检测算法,在模型基础上,通过替换轻量级主干网络 HGNetV2,增强特征提取能力;融合 HG-Block 和 SPPELAN 减小模型参数加快检测速度;使用 Dysample 动态上采样算子加强特征融合,提升不同尺度目标的检测精度。实验结果表明,改进后的算法检测精度达到了 80.5% 高于改进前的算法及其他主流的检测算法,且速度提升了 16% 更加轻量化。通过现场实施应用,满足了特钢车间内目标检测算法部署实施的需求,为实现数字化产线的物料跟踪提供了可能。

### 参考文献

- [ 1 ] HESS T, MATT C, BENLIAN A, et al. Options for formulating a digital transformation strategy [ J ]. MIS Quarterly Executive, 2016, 15(2): 123-139.
- [ 2 ] 黄新峰, 杨军, 李员妹, 等. 物料跟踪系统在高线生产过程中的应用 [ J ]. 科技创新导报, 2020, 17(36): 103-106.  
HUANG X F, YANG J, LI Y M, et al. Application of material tracking system in high wire production [ J ]. Science and Technology Innovation Guide, 2020, 17(36): 103-106.
- [ 3 ] 唐全军. 实施 ETC 产品数字化车间构建与应用 [ J ]. 智能制造. 2022, 29(2): 103-108.  
TANG Q J. Implement ETC product digital workshop construction and application [ J ]. Intelligent Manufacturing, 2022, 29(2): 103-108.
- [ 4 ] YANG ZH X, ZHANG P B, CHEN L. RFID-enabled indoor positioning method for a real-time manufacturing execution system using OS-ELM [ J ]. Neurocomputing, 2016, 174: 121-133.
- [ 5 ] HUANG SH H, GUO Y, ZHA SH SH, et al. A real-time location system based on RFID and UWB for digital manufacturing workshop [ J ]. Procedia Cirp, 2017, 63: 132-137.
- [ 6 ] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [ C ]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [ 7 ] 代啟亮, 熊凌, 陈琳国, 等. 改进 YOLOv5 的 PDC 钻头复合片缺损识别 [ J ]. 电子测量与仪器学报, 2023, 37(8): 164-172.  
DAI Q L, XIONG L, CHEN L G, et al. PDC drill bit defect recognition by improved YOLOv5 [ J ]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 164-172.
- [ 8 ] 张瀚丹, 吴一全. 基于视觉的汽车装配件缺陷检测研究进展 [ J ]. 仪器仪表学报, 2023, 44(8): 1-20.  
ZHANG H D, WU Y Q. Research progress of vehicle assembly defect detection methods based on vision [ J ]. Chinese Journal of Scientific Instrument, 2023, 44(8): 1-20.
- [ 9 ] SHI B, LI X, NIE T, et al. Multi-object recognition method based on improved YOLOv2 model [ J ]. Information Technology and Control, 2021, 50(1): 13-27.
- [ 10 ] ZHANG W, HONG Z, XIONG L, et al. Sinextnet: A new small object detection model for aerial images based on PP-Yoloe [ J ]. Journal of Artificial Intelligence and Soft Computing Research, 2024, 14(3): 251-265.
- [ 11 ] WANG CH 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.
- [ 12 ] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [ C ]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [ 13 ] HUANG G, LIU ZH, VAN DER MAATEN L, et al. Densely connected convolutional networks [ C ]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.
- [ 14 ] LI H, ZHANG J, LI R, et al. Swin-RGC: Swin-Transformer with recursive gated convolution for substation equipment non-rigid defect detection [ J ]. IEEE Access, 2023, 11: 72655-72664.
- [ 15 ] ZHAO Y AN, LYU W, XU SH L, et al. Detrs beat yolos on real-time object detection [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [ 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 ] TALAB M A, AWANG S, NAJIM S A M. Super-low resolution face recognition using integrated efficient sub-pixel convolutional neural network ( ESPCN ) and convolutional neural network ( CNN ) [ C ]. 2019 IEEE International Conference on Automatic Control and Intelligent systems ( I2CACIS ). IEEE, 2019: 331-335.

[18] OUYANG Z Q, FU X Q, ZHONG ZH B, et al. An exploration of the influence of ZnO NPs treatment on germination of radish seeds under salt stress based on the YOLOv8-R lightweight model[J]. Plant Methods, 2024, 20(1): 110.

[19] HAN K, WANG Y H, XU CH, et al. GhostNets on heterogeneous devices via cheap operations [ J ]. International Journal of Computer Vision, 2022, 130(4): 1050-1069.

[20] VASU P K A, GABRIEL J, ZHU J, et al. Mobileone: An improved one millisecond mobile backbone [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7907-7917.

[21] JU J Y, CHEN G Q, LYU ZH Y, et al. Design and experiment of an adaptive cruise weeding robot for paddy fields based on improved YOLOv5[J]. Computers and Electronics in Agriculture, 2024: 219: 108824.

[22] YANG H J, LIU J X, MEI G M, et al. Research on real-time detection method of rail corrugation based on improved ShuffleNet V2[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 106825.

作者简介



赵云涛,2010 年于北京科技大学获得博士学位,现为武汉科技大学信息科学与工程学院教授。主要研究方向为机器人学、智能视频分析。  
E-mail: zhyt@wust.edu.cn

**Zhao Yuntao** received his Ph. D. from University of Science and Technology Beijing in 2010. He is now a professor at the School of Information Science and Engineering, Wuhan University of Science and Technology. His main research interests include robotics and intelligent video analysis.



黄哲辉(通信作者),2022 年于武汉科技大学获得学士学位,现为武汉科技大学硕士研究生,主要研究方向为机器视觉、深度学习。  
E-mail: 973110924@qq.com

**Huang Zhehui** ( Corresponding author ) received his B. Sc. degree from Wuhan University of Science and Technology in 2022. He is now a M. Sc. Candidate at Wuhan University of Science and Technology. His main research interests include machine vision and deep learning.