

文章编号: 1672-8785(2024)09-0029-15

基于改进 YOLOv8 的圆形合作目标 检测定位算法研究

徐 非 林雪竹* 郭丽丽 孙 静 李丽娟

(长春理工大学光电工程学院, 吉林 长春 130022)

摘 要: 针对视觉测量中低照度或复杂背景下圆形合作目标识别精度较低或定位能力较差等问题, 提出了一种基于卷积神经网络(Convolutional Neural Networks, CNNs)对 YOLOv8 算法进行优化设计的模型。该模型共有 225 层网络、约 300 万个参数以及 8.2G FLOPs 计算量。采用不同条件下的圆形合作目标数据集进行模型训练, 同时在训练过程中实时监控模型的性能指标和计算效率, 对模型进行细致的调整和优化。实验结果表明, 本文算法具有 99% 的准确率、92% 的召回率、92% 的平均精度。相较于霍夫变换和 YOLOv3 等传统特征提取方法, 本文算法的精确率分别提升了 14% 和 4%, 召回率分别提升了 17% 和 2%, 平均精度分别提升了 10% 和 2%。该算法可以在低照度环境、复杂背景或目标形状微小变化等多变条件下, 显著提高圆形合作目标的识别定位精度。

关键词: 视觉测量; YOLOv8; 精确率; 召回率; 平均精度

中图分类号: TH74 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2024.09.005

Research on Circular Cooperative Object Detection and Localization Algorithm Based on Improved YOLOv8

XU Fei, LIN Xue-zhu*, GUO Li-li, SUN Jing, LI Li-juan

(College of Opto-Electronic Engineering, Changchun University of
Science and Technology, Changchun 130022, China)

Abstract: Aiming at problems such as low recognition accuracy or poor localization ability of circular cooperative objects in low illumination or complex backgrounds in vision measurement, a model based on CNNs is proposed in this paper to optimize the YOLOv8 algorithm. The model designed in this paper has a total of 225 layers of network, about 3 million parameters and 8.2G FLOPs of computing power. The model is trained by using the circular cooperative target data set under different conditions, and the performance index and computational efficiency of the model are monitored in real time during the training process, and the model is adjusted and optimized in detail. The experimental results show that the algorithm has a precision of 99%, a recall rate of 92% and an average accuracy of 92%. Compared with traditional feature extraction methods such as Hough transform and YOLOv3, the accuracy of the proposed algorithm is improved by 14% and 4%. Recall rates in-

收稿日期: 2024-02-08

基金项目: 吉林省科技发展计划重点研发项目(20200401063GX)

作者简介: 徐非(1997-), 男, 黑龙江黑河人, 硕士, 主要从事视觉测量方面的研究。

*通讯作者: E-mail: custsnow@163.com

crease by 17% and 2%. The average accuracy is improved by 10% and 2%. The algorithm can significantly improve the recognition and positioning accuracy of circular cooperative targets under variable conditions such as low illumination environment, complex background or small change of target shape.

Key words: vision measurement; YOLOv8; precision; recall rate; average precision

0 引言

目前,图像处理和目标检测已成为计算机视觉和人工智能领域的核心研究方向。同时,这两项技术的进步对于医疗成像、工业自动化、交通监控等领域至关重要。在医疗成像领域,精确的圆形目标检测定位可以帮助医生识别和分析各种生物结构(如肿瘤、细胞和其他组织结构)和病理变化;在工业自动化应用中,精确检测机器部件中的圆形元件可以保证质量控制和自动化装配线的有效运行;在交通监控中,通过对监控画面中图像特定目标的识别可以实现精确定位,有助于交通拥堵率和事故发生率的有效降低。

视觉信息是人类获取外部自然信息的主要形式。随着信息技术的不断发展,视觉测量技术应运而生。在传统的视觉测量任务中,利用图像处理和目标检测技术对测量合作目标进行精确识别及提取已逐渐成为当下的研究热点。通过相应算法进行圆形目标等特定形状的识别以及对测量合作目标中心进行精确提取,对于提高视觉测量系统的性能及精度至关重要,在大多数的工业视觉测量任务中具有独特的技术挑战和应用价值^[1-2]。

作为深度学习中类型不同的神经网络, CNNs^[3] 由于自动特征提取能力强大和计算效率高,适用于处理图像和视频等二维或三维数据;通过卷积操作可以捕捉输入图像数据中的边缘、形状、纹理等特征。基于 CNNs 进行结构与模型训练,可使所设计的模型具有更强大的表征学习能力,并且从大量的数据中学习更多特征信息,能够应用于视觉测量中对合作目标的检测定位,从而在不同场景中针对不同照度、不同背景或目标形状复杂微变等因素提高测量系统精度。

YOLO(You Only Look Once)算法^[4-5]在目标检测领域由于高速度和高准确性而广受欢迎。其核心在于将目标检测任务作为回归问题来解决,通过单个神经网络模型直接建立从图像像素到边界框坐标和类别概率的映射。近几年, YOLO 系列的 YOLOv8 算法^[6] 应运而生。它不仅继承了 YOLO 算法系列的高效性能,引入了多项创新,还进一步提高了目标检测的精度和速度。

虽然 YOLOv8 算法在目标检测方面的性能已得到广泛认可和验证,但其在特定形状检测方面的能力,特别是在视觉测量中合作目标的识别、定位以及精确提取等方面的应用尚未被充分探索。例如, YOLOv8 算法在智能城市的火灾检测中通过深度学习实时识别火源特定特征,展示了其在提高准确性、减少误报方面的潜力^[7]。在无人机平台上, YOLO-Drone 算法通过提升实时推理速度和平均精度,超越了传统的 YOLO 系列算法^[8]。这两项应用表明 YOLO 算法在处理速度和准确度方面具有优势,使其成为各种复杂目标检测任务的理想选择。

然而,当 YOLO 算法应用于特定的工业圆形合作目标检测场景时,其性能会受到特定条件制约。例如,在低光照条件下,尽管 YOLO 算法在识别反光点、玻璃微珠、安全眼镜等对象时可取得超过 88% 的精确度,但精确度在低于 400 lx 的光照强度下会显著下降,仅为 78.63%^[9]。此外, YOLO 算法在水下环境中也面临挑战,光线昏暗和水质浑浊等因素会严重影响算法的检测性能^[10]。

综上所述,本文在 CNNs 的基础上对 YOLOv8 算法进行优化,旨在深入评估 YOLOv8 算法在视觉测量任务(特别是工业视

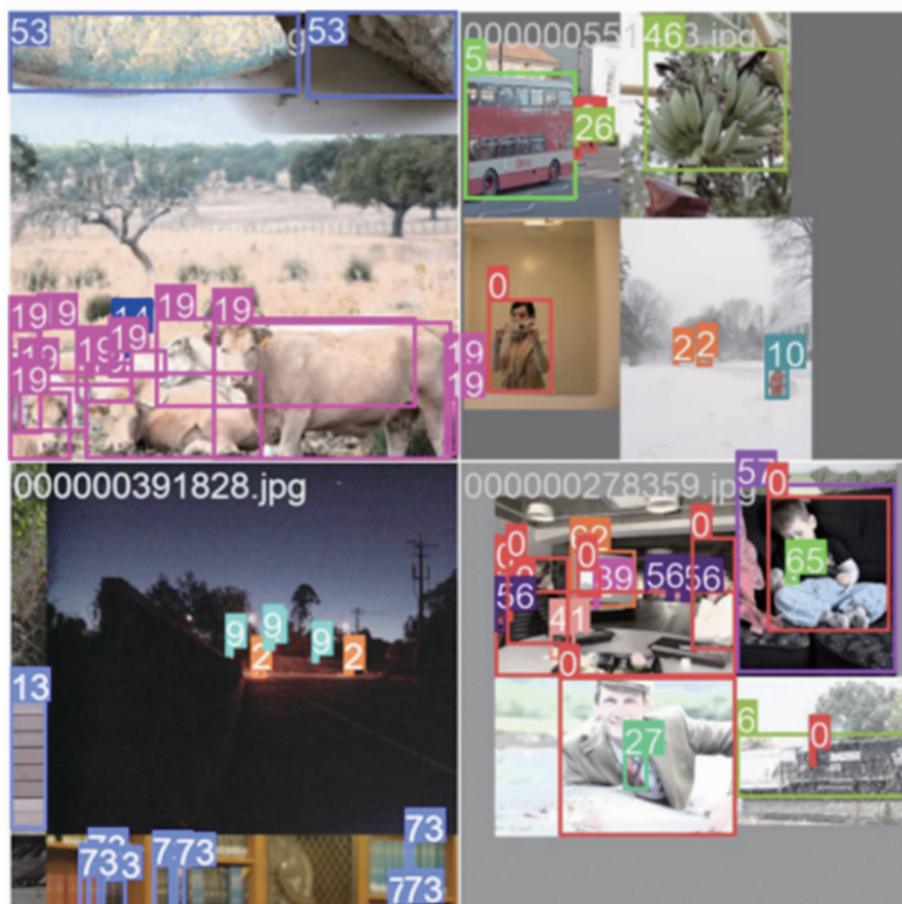


图 1 目标检测示意图

觉测量的应用场景)中图像的圆形目标检测以及定位方面的性能。通过构建包含各种条件和背景下圆形合作目标的广泛图像数据集,对本文优化设计的 YOLOv8 算法进行全面的性能评估。针对准确率、召回率和 F1 分数等标准化评估指标,将该算法与其他目标检测算法进行比较,以确保结果的可靠性和一致性。

1 相关工作

1.1 目标检测与圆形检测

在计算机视觉的演变历程中,目标检测技术始终处于核心地位,推动着早期基于知识的方法到现代的深度学习方法的显著转变。在早期,目标检测技术主要依靠基于知识的方法,通过定义特定对象相关的规则来实现目标检测与识别。图 1 所示为一个目标检测系统的示例输出。其中,不同的颜色框代表系统识别并定位的各种目标,框内的数字表示目标的类别标识。

随着相关技术的进步,特征学习阶段的兴起带来了 Viola-Jones^[11] 目标检测框架和可变形部件模型(Deformable Part Model, DPM)^[12] 等标志性成就。进入 21 世纪后,随着深度学习技术的不断进步,特别是深度学习中 CNNs 的发展,为目标检测带来了革命性变革,极大地提升了模型在复杂场景下的性能,为后续 YOLO 系列算法的发展奠定了坚实的基础^[13]。

圆形检测作为目标检测技术中的一个分支,在医疗成像、交通监控、工业检测等领域占据独特地位。对包含圆形目标的图像进行有效识别和定位一直是计算机视觉领域的研究热点。图 2 所示为一个圆形物体检测的实例。其中,红色框表示检测算法识别的目标区域。圆形检测从霍夫变换逐渐发展到基于深度学习的方法,展现出了显著进步。

在圆形检测技术的早期方法中, Canny 边缘检测算法常在图像预处理阶段用于提取图像



图 2 圆形检测示意图

中的边缘信息^[14]。它是一种含有最优化思想的算子，其最优性表现在较大的信噪比、较好的定位精度以及最小的边缘响应。该算法通过高斯滤波器平滑图像，然后计算图像的梯度强度和方向，最后通过非极大值抑制和双阈值检测确定边缘：

$$\nabla f(x, y) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1)$$

$$ag(f) = G_x^2 + G_y^2 \quad (2)$$

式中， G_x 和 G_y 分别是图像在 x 和 y 方向上的一阶导数； $ag(f)$ 表示图像中的向量幅值。

霍夫变换^[15]用于从边缘检测的结果中识别出特定形状(如圆或直线等)，也可推广至其他几何元素的提取。它将图像空间中的直线变换到参数空间中的某个点，建立图像空间到参数空间的映射关系，将图像空间中直线的检测问题转换为参数空间中某一条件点的检测问题。其中，对于圆形检测，标准霍夫变换的公式可以表示为

$$(x-a)^2 + (y-b)^2 = r^2 \quad (3)$$

式中， a 、 b 和 r 分别代表圆的中心坐标和半径。尽管霍夫变换在早期得到了广泛应用，但它在处理大规模或嘈杂图像时，计算复杂度高且容易受到噪声干扰。

1.2 卷积神经网络

CNNs 是一种规范化的前馈神经网络，通过滤波器(或核)优化自身学习特征工程，用来处理具有类似网格结构数据的神经网络(见图 3)。CNNs 通常由输入层、隐藏层和输出层组成。

隐藏层包含一个或多个执行卷积的层，通常为一个执行卷积核与层输入矩阵的点积的层。其中的点积通常是弗罗贝尼乌斯(Frobenius)内积。

随着卷积核沿着层输入矩阵滑动，卷积层操作生成一个特征图，反过来又贡献给池化层、全连接层等下一层的输入，再使用 softmax 激活函数进行多类别分类，最后经过输出层进行分类结果输出。

与其他图像分类算法相比，CNNs 至少在网络的一层中使用卷积运算替代一般的矩阵乘法运算。它通过在输入数据上滑动一个滤波器(也称为卷积核)进行特征提取：

$$y[i, j] = \sum_{m=-k/2}^{k/2} \sum_{n=-k/2}^{k/2} x[i+m, j+n] \cdot \omega[m, n] \quad (4)$$

式中， x 为输入数据； y 是卷积计算输出； ω 是卷积核(滤波器)； k 是卷积核大小；下标 i 和 j 表示输出的位置； m 和 n 表示卷积核的位置(i, j)。卷积计算过程可以简单描述如下：对于输出的每个位置，将卷积核与输入数据的对应位置进行元素乘积，然后将所有乘积结果相加得到输出 $y[i, j]$ 。

2 基于卷积神经网络的 YOLOv8 优化设计

2.1 卷积层结构设计

CNNs 的核心构件是卷积层，其主要任务是提取输入数据特征。在卷积层中，输入数据可以通过一组可学习的滤波器进行处理。这些滤波器能够捕捉到数据中的局部依赖和空间层级结构。这一过程可以通过卷积操作来进行数学表述。假设有输入 I 和滤波器 F ，卷积操作定义为

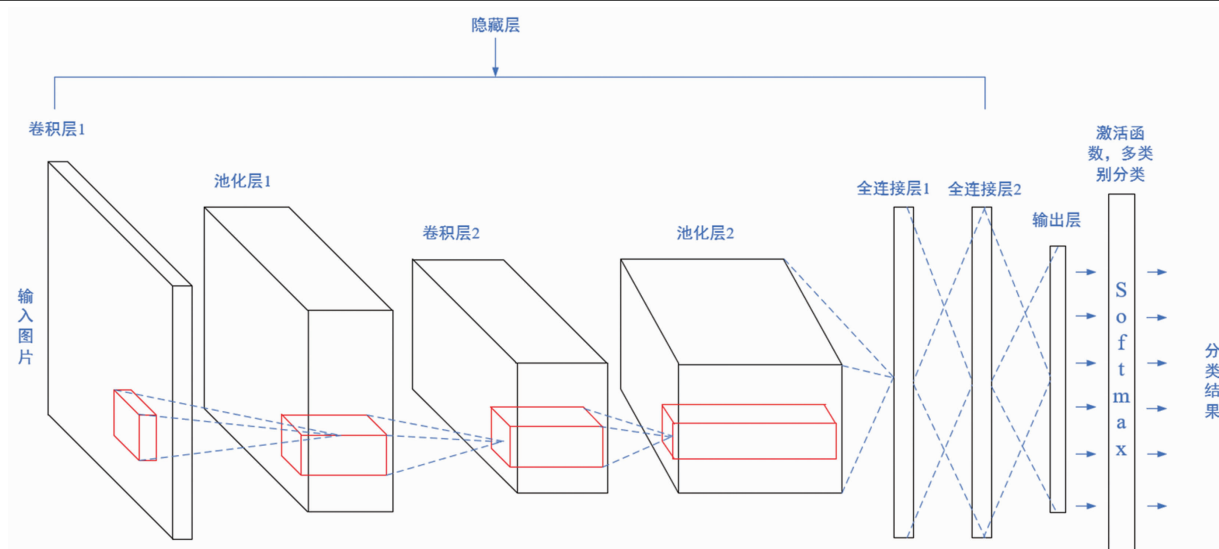


图 3 卷积神经网络模型的结构示意图

$$(I * F)(i, j) = \sum_m \sum_n I(i-m, j-n) \cdot F(m, n) \quad (5)$$

式中, i, j 表示输出特征图的空间位置; m, n 代表滤波器的维度。通过这种方式, 卷积层能够有效地提取和组合输入数据的特征。

图 4 为本文提出的卷积层结构设计的流程图。本文设计的卷积层在 CNNs 结构的基础上添加了批量归一化 (Batch Normalization, BN) 层, 同时引入了修正线性单元 (Rectified Linear Unit, ReLU) 激活函数。

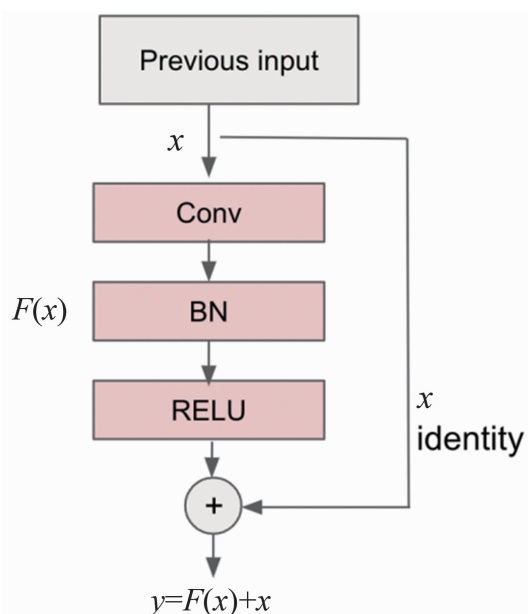


图 4 卷积层结构设计流程图

BN 层是 Ioffe S 等人^[16]提出的一种结构, 主要用于解决梯度消失问题和加速模型的收敛速度等。它通过规范化每一层的输入来减小内部协变量偏移的影响。对于一批数据 $B = \{x_1, x_2, \dots, x_m\}$, BN 层首先计算数据的均值 μ_B 和 σ_B^2 方差, 然后对每个数据进行规范化和线性变换:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (6)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (7)$$

式中, $\hat{x}_i$ 代表卷积层中输出的估计值; γ, β 表示可学习的参数量; ϵ 是一个极小的正数以防止除以零。最终输出为

$$y_m = \frac{\gamma}{\sqrt{\sigma_y^2 + \epsilon}} \left(\sum_i^N \omega_i x_i + b - \mu_y \right) + \beta \quad (8)$$

式中, μ_y, σ_y 代表 BN 层的均值; ω 是权重参数; b 是偏置参数。

通过一系列的操作实现批量归一化的过程, 使激活函数的输入分布更加稳定, 从而加速网络收敛。本文采用的激活函数为 ReLU 激活函数。此函数是目前神经网络中最经常使用的激活函数, 当输入超过阈值时才会激发:

$$RELU(x) = MAX(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (9)$$

式中, x 为来自 BN 层的输出。

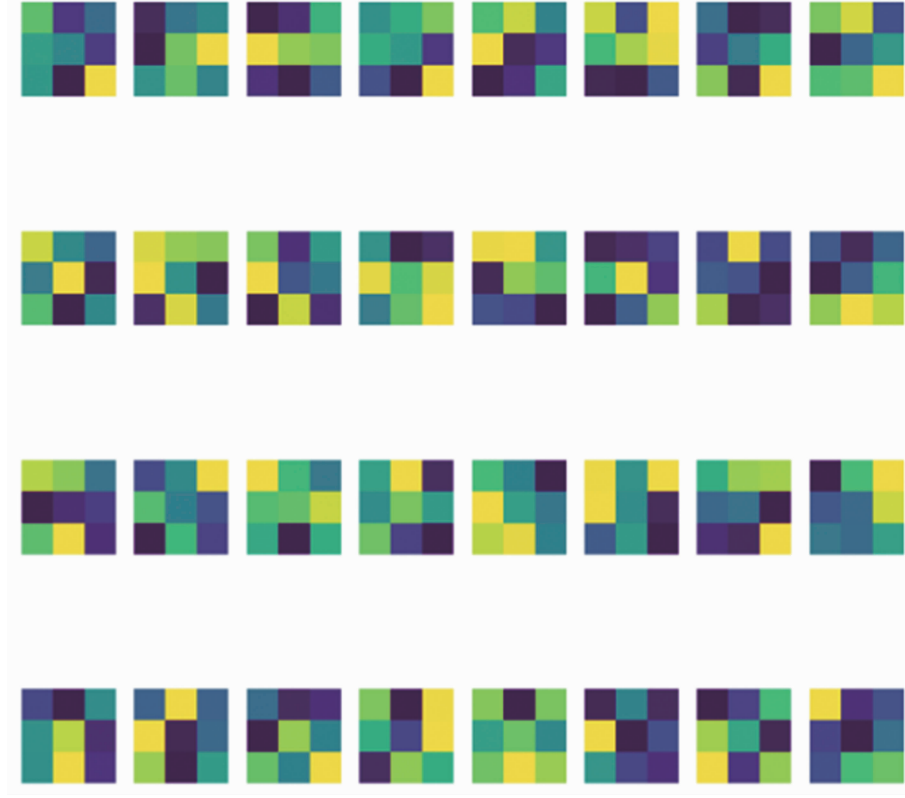


图 5 卷积层的可视化示意图

图 5 为本文设计的卷积层的可视化示意图。从中可以直观地观察卷积层如何通过不同的滤波器捕获图像中的多样化信息，从而帮助理解和分析神经网络如何学习和识别不同的视觉模式。

在卷积层添加 BN 层并采用 ReLU 激活函数后，本文通过引入残差连接来解决深层网络训练困难的问题。在残差网络中，输入图片不仅通过卷积层进行 $H(x)$ 变换，且直接与层的输出相加，形成所谓的“捷径”或“跳跃连接”，即

$$F(x) = H(x) + x \quad (10)$$

通过引入残差连接，使网络能够学习输入和输出之间的残差函数，有效避免了梯度消失或爆炸问题，提高了深层网络的训练效率和准确率。

同时，通过对卷积层的空间特征提取、BN 层的输入规范化以及残差连接的梯度流优化，以期实现本文设计的优化模型的学习能力和广泛适用性。

2.2 YOLOv8 设计

与传统目标检测方法不同，YOLO 算法直接在整个图像上划分网格，每个网格单元负责预测多个边界框和对应的置信度。这种设计使得 YOLO 算法能够在网络的单次前向传播中直接预测对象的类别和位置，从而实现快速且高效的目标检测^[17]。

YOLOv8 将损失函数设计为边界框坐标的损失、置信度的损失和类别预测的损失等多部分的组合。YOLOv8 的损失函数可表示为

$$\begin{aligned} L = & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\ & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B l_{ij}^{obj} (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} l_i^{noobj} (C_i - \hat{C}_i)^2 \\ & + \lambda_{class} \sum_{i=0}^{S^2} l_i^{noobj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2 \quad (11) \end{aligned}$$

式中， S^2 表示网格数； B 表示各网格预测的边界框数量； λ_{coord} 、 λ_{class} 用来平衡不同部分损失的系数；如果网格单元 i 中的第 j 个边界框负责预测某个对象，那么 l_{ij}^{obj} 为 1，否则为

0; 如果网格单元中没有对象, 那么 l_i^{noobj} 为 1, 否则为 0; x_i 、 y_i 、 w_i 、 h_i 是预测的边界框的中心坐标及宽高; $\hat{x}_i$ 、 $\hat{y}_i$ 、 $\hat{w}_i$ 、 $\hat{h}_i$ 是对应的真实值; $(C_i - \hat{C}_i)^2$ 是置信度损失, 表明预测的边界框与真实边界框的交并比 (Intersection over Union, IoU) 的误差; $(p_i(c) - \hat{p}_i(c))^2$ 为分类损失, 表示预测的类别概率与真实类别概率的误差^[18]。

表 1 列出了本文设计的 YOLOv8 网络层次结构的具体参数。

表 1 YOLOv8 各模型参数及 FLOPs

模型名称	参数量(M)	FLOPs(B)
YOLOv8n	3.2	8.7
YOLOv8s	11.2	28.6
YOLOv8m	25.9	78.9
YOLOv8l	43.7	165.2
YOLOv8x	68.2	257.8

图 6^[19]为 YOLOv8 深度学习模型结构的可视化示意图。其中, 卷积层和残差块被用来提取图像特征; SPPF 层和 UPSAMPLE 层用于捕捉各图像不同尺度的信息; CONCAT 层用于特征融合, 增强模型对不同大小对象的适应性; Detect 头部用于生成最终的边界框、对象置信度和分类预测。

本文设计的优化模型共有 225 层网络、约 300 万参数以及 8.2G FLOPs 的计算量, 间接反映出模型的复杂性和计算需求。在训练过程中, 通过实时监控性能指标和计算效率, 对模型进行了细致的调整和优化。

3 实验

3.1 实验设置及数据集

本文实验的硬件平台采用 Windows10 操作系统, 处理器为英特尔 i7-10870H, 运行环境为 NVIDIA 3090 显卡, 显存容量为 24G; 软件平台采用基于 Python 软件的 PyTorch 框架。本文选择基于 Python 的 PyTorch 框架作为实验平台, 旨在利用其灵活性和高效的计算能力来支持复杂模型的训练和验证。

图 7 所示为本文设计的圆形目标检测数据集。此数据集旨在评估和优化本文的目标检测算法在处理具有编码特征的圆形目标方面的性能。同时, 本文圆形目标检测的数据集的设计考虑了多样化的环境变量, 以确保算法的鲁棒性和泛化能力。数据集的图像来源于多个角度和光照条件下对多种材料和背景的拍摄, 从而增加检测任务的复杂性。在这些图像中, 圆形目标可能出现在图像的任何位置, 且存在由于遮挡、光照变化或图像质量问题而变得模糊不清的现象。

3.2 模型训练

本文将 batch size (一次训练所选取的样本数) 设置为 16 以适应硬件资源, 初始学习率设定为 0.01, 采用动态调整的策略以提高收敛速度; 模型在 640×640 大小的输入图像下进行训练, 以平衡检测精度和计算效率; 训练过程中采用自动混合精度 (Automatic Mixing Precision, AMP) 优化训练速度和减少内存使用。

本文采用 AdamW^[20] 优化器进一步提升训练效率以及模型的泛化能力。AdamW 优化器

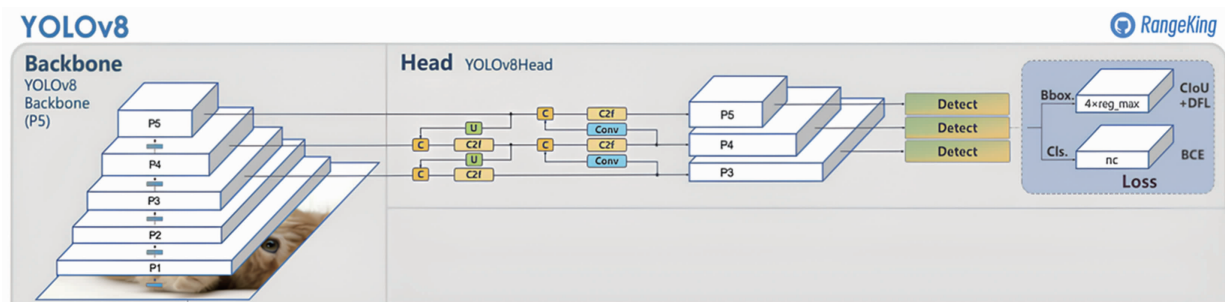


图 6 YOLOv8 深度学习模型结构的可视化示意图

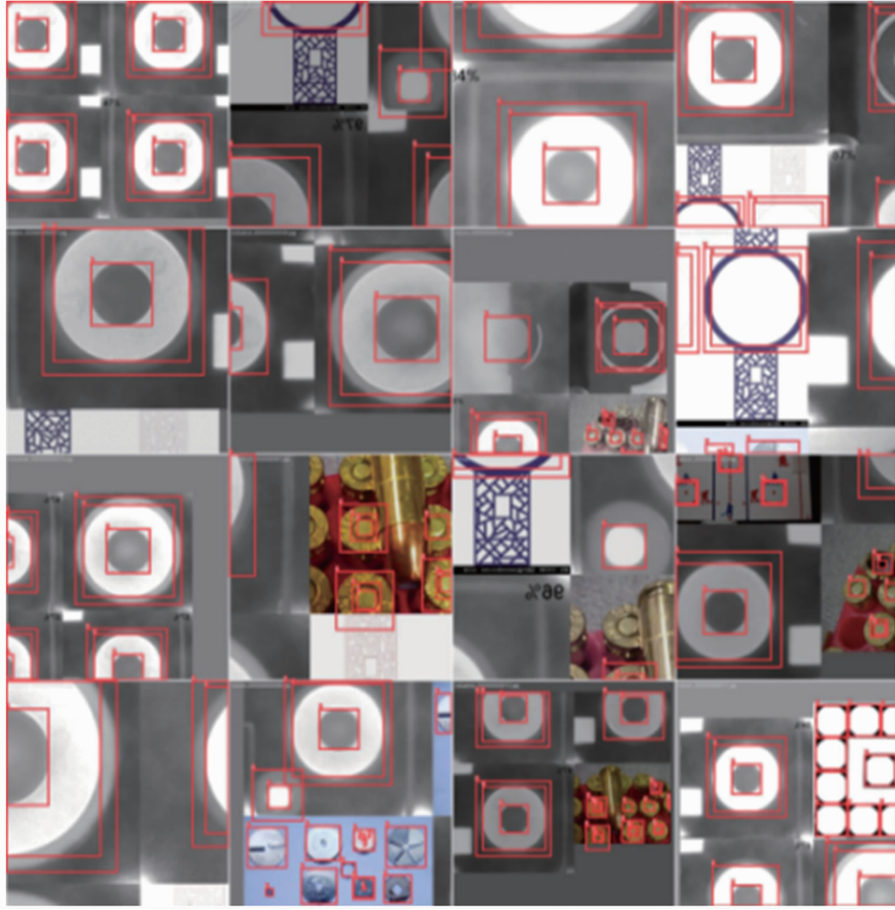


图 7 圆形目标检测数据集

是 Adam 优化器的一个变种，在原有的基础上引入了权重衰减机制。与传统的 L2 正则化相比，AdamW 优化器通过解耦权重衰减和梯度更新减轻过拟合，增强模型泛化性。AdamW 的关键更新步骤如下：

$$\begin{aligned}
 g_t &\leftarrow \nabla_{\theta} f_t(\theta_{t-1}) \\
 \theta_t &\leftarrow \theta_{t-1} - \gamma \lambda \theta_{t-1} \\
 m_t &\leftarrow \beta_1 m_{t-1} + (1 - \beta_1) g_t \\
 v_t &\leftarrow \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\
 \hat{m}_t &\leftarrow \frac{m_t}{1 - \beta_1^t} \\
 \hat{v}_t &\leftarrow \frac{v_t}{1 - \beta_2^t} \\
 \theta_t &\leftarrow \theta_t - \frac{\gamma \hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}
 \end{aligned} \quad (12)$$

式中， m 为动量； v 为梯度平方的指数加权平均； β_1 和 β_2 是动量和平方梯度的指数衰减率； $\hat{m}$ 是对动量和平方梯度进行的偏差修正。

3.3 模型评价指标

模型性能的评价提供了模型效果的量化分析，为模型的改进和优化指明方向。

本文采用四种常用的目标检测评价指标：F1 分数(F1 Curve)、精确率-召回率曲线(Precision-Recall Curve, PR Curve)、精确度-置信度曲线(P Curve)和平均精度均值(mean Average Precision, mAP)。

F1 Curve 作为精确率和召回率的调和平均，提供了一个单一指标，平衡了两者的影响：

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (13)$$

式中， $Precision$ 为模型精确率； $Recall$ 为模型召回率。在不同阈值的条件下绘制的 F1 Curve 能够直观地反映模型在精确率和召回率之间权衡的变化情况。

PR Curve 展示了评价数据不平衡时模型

在所有可能的阈值内的精确率和召回率之间的关系。其中, PR Curve 的 x 轴是召回率, y 轴是精确率。曲线下的面积(AUC)是评估模型整体性能的重要指标:

$$AUC_{PR} = \int_0^1 Precision(Recall) d(Recall) \quad (14)$$

与 PR Curve 不同, P Curve 表示模型在不同召回率水平下的精确率变化, 适用于对正样本预测精确性要求极高的场景。绘制 P Curve 时, 需要将召回率作为 x 轴, 将精确率作为 y 轴, 从而直观展示提高召回率(即降低阈值)情况下模型精确率的变化情况。

mAP 是目标检测中最常用的评价指标之一, 用于衡量模型在所有类别上的整体性能。首先, 计算每个类别的平均精度(Average Precision, AP); 其次, 计算所有类别 AP 的均值, 且单个类别的 AP 基于模型在该类别上的精确率和召回率进行计算; 同时, 对于每个类别,

需要按照置信度将检测结果进行排序, 计算每个置信度阈值下的精确率和召回率; 最后, 计算 AP:

$$AP = \int_0^1 p(r) dr \quad (15)$$

式中, $p(r)$ 是召回率为 r 时的精确率。最终, mAP 是所有类别 AP 的平均值:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (16)$$

式中, N 是类别的数量; AP_i 是第 i 个类别的平均精度。

3.4 实验结果与分析

模型的损失函数是衡量模型学习过程的直观指标, 包括边框损失、类别损失以及目标损失。其中, 边框损失衡量了模型预测边界框与实际标注边界框之间的差异; 类别损失评价了模型对目标类别的识别能力; 目标损失反映了模型对目标检测置信度的准确性。

如图 8 所示, 经过 40 个模型训练周期后,

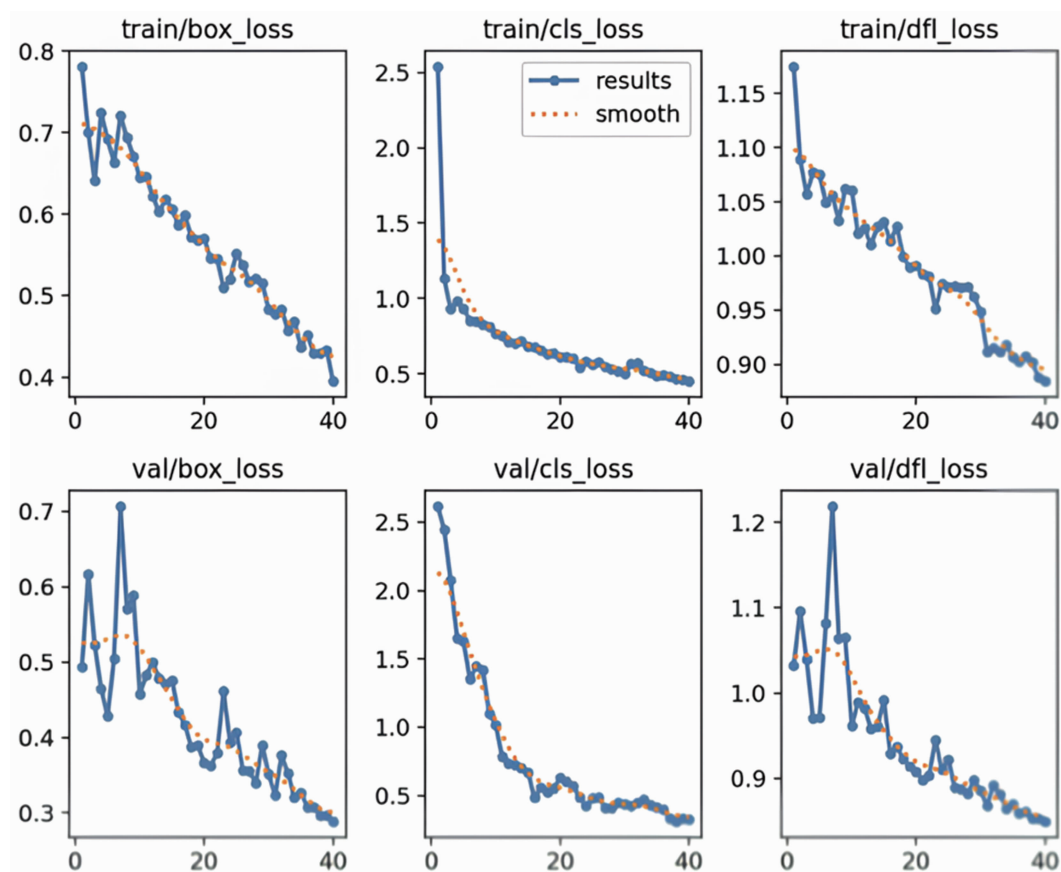


图 8 本文设计的模型损失函数的收敛图

本文所设计的模型损失函数显示出稳定下降的趋势。这表明模型在训练过程中逐渐学习到有效的特征表示。

由图 8 可知,训练阶段的边框损失从初始值稳步下降,验证阶段同样展现下降趋势,尽管中途出现了一些波动。这种波动可能源于模型在面对验证集中的某些复杂样本时的适应性问题。类别损失和目标损失在训练过程中也表现出相似的下降趋势,进一步证明模型区分不同编码类别和定位目标的能力不断增强。

图 9 所示为模型的 F1 Curve。本文模型在图 7 所示的圆形目标检测数据集上的 F1 Curve 显示,它在大多数置信度阈值下可达到良好的精确率和召回率平衡。

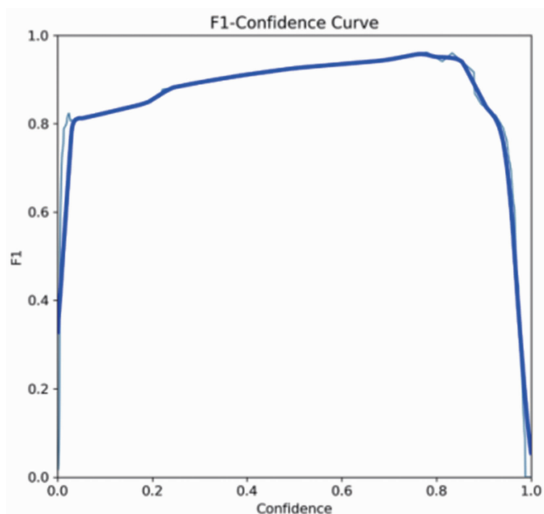


图 9 本文模型的 F1 Curve 收敛图

随着置信度阈值增加, F1 Curve 先迅速上升至高峰,在置信度较高的区间内维持稳定,最后接近置信度 1.0 时急剧下降。在置信度阈值较低的区域,模型的 F1 Curve 较低,可能是因为低置信度阈值时模型生成大量的假阳性,从而降低了精确率。但随着置信度阈值提高,模型开始排除那些不太可能是真正目标的检测,因此精确率提高,且 F1 Curve 快速上升。

图 10 所示为模型的 P Curve。当曲线在很低的置信度阈值时精确度较低,模型将很多不确切的检测结果视为正样本,使真正的正样本中含有许多误报。随着置信度阈值提高,精确

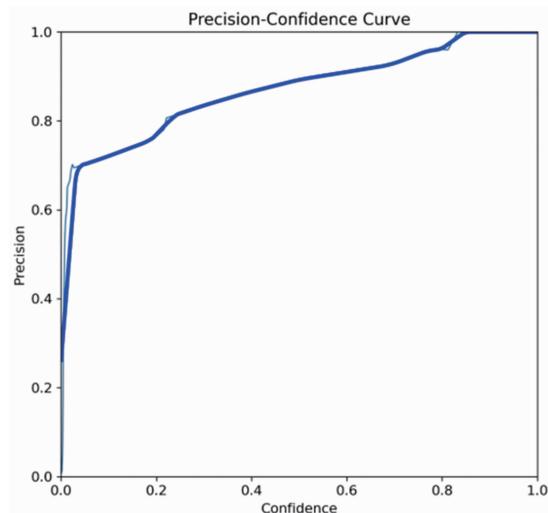


图 10 本文模型的 P Curve 收敛图

度迅速上升,模型能够去除大多数误报,保留更有可能是真实目标的检测结果。当置信度阈值增加到某一点时,精确度达到相对高值并在此区间保持稳定,表明模型对于真正目标的识别具有较高准确性。

图 11 所示为模型的 PR Curve。当召回率接近 1.0 时,曲线的精确率仍保持在非常高的水平。在目标检测任务中,这种性能是非常理想的,因为它意味着模型能够检测到几乎所有正样本(高召回率),同时保持很低的误检率(高精确率)。

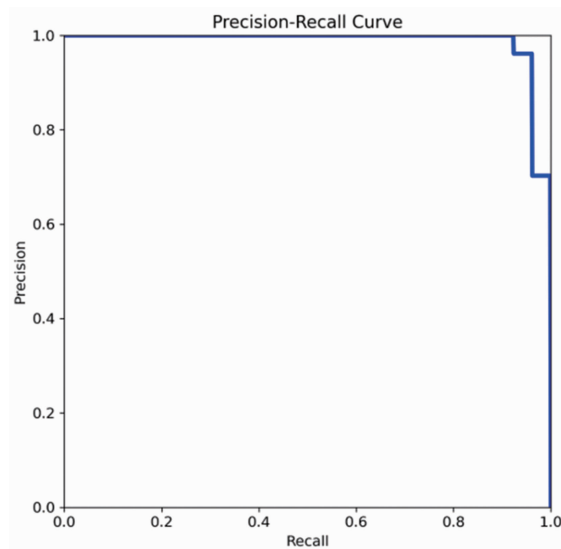


图 11 本文模型的 PR Curve 收敛图

图 12 所示为模型在不同目标检测评价指标上的性能,包括精确率、召回率、50% IoU

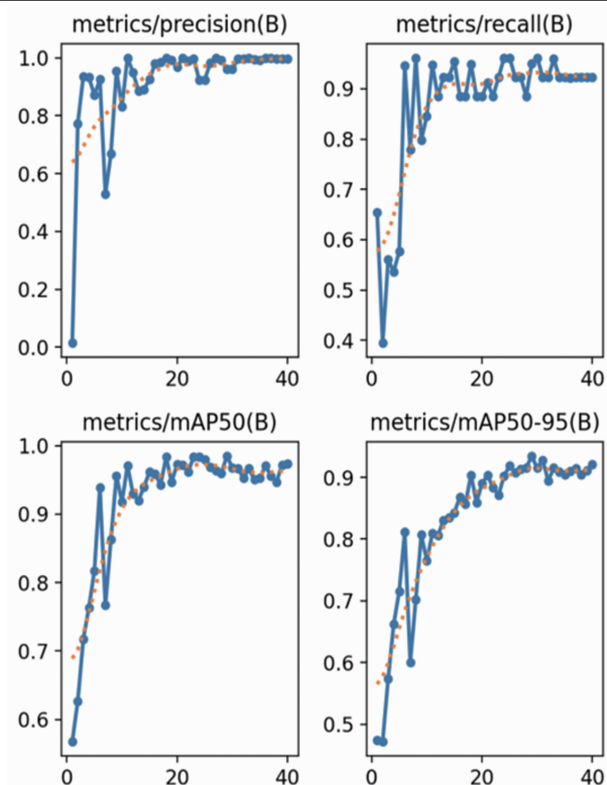


图 12 本文模型的精确率、召回率、mAP50%、mAP50%-95%

阈值下的 mAP (mAP50) 以及 50%~95% IoU 阈值下的 mAP (mAP50-95)。

在精确率图中, 曲线在前期训练阶段有较大波动后逐渐稳定, 并在训练后期达到接近 1.0 的高值, 表明模型经过充分训练后在置信度较高的情况下可以准确识别目标。在召回率图中, 曲线初期波动但在随后的训练过程中趋于稳定。召回率在训练过程中总体保持较高水平, 表明模型具有较好的全面性, 能够检测大部分正样本。

在 mAP50% IoU 图中, 模型在 50% IoU 阈值下的 mAP 值在训练初期迅速上升, 在后期训练达到高水平并维持在 0.9 以上。这表明在较宽松的 IoU 阈值下, 模型可高精度识别目标。mAP50%-95% IoU 图给出了模型在更严格的 IoU 阈值范围内的 mAP 值。模型指标在训练过程中稳步上升, 在训练后期稳定在较高水平, 表明模型在不同严格度的 IoU 阈值下可以保持良好的性能。

图 13、图 14、图 15 所示为霍夫变换、YOLOv3 和本文优化设计的 YOLOv8 三种目

标检测方法的精确度、召回率以及 mAP 值的比较。可以看出, 本文设计的 YOLOv8 优化模型在三种评价指标中均实现突破。同时, 从 YOLOv3 到本文设计的 YOLOv8 优化模型的精确度增量突显了网络架构、训练方法和泛化能力在处理各种复杂检测场景方面的持续改进。本文设计的 YOLOv8 优化模型在目标检测任务中通过监控损失函数的收敛和精确度指标的提升, 对复杂的圆形目标编码模式具有良好的学习和识别能力。

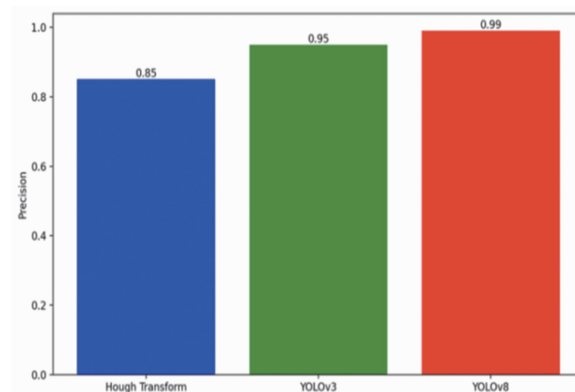


图 13 三种目标检测方法的精确度对比

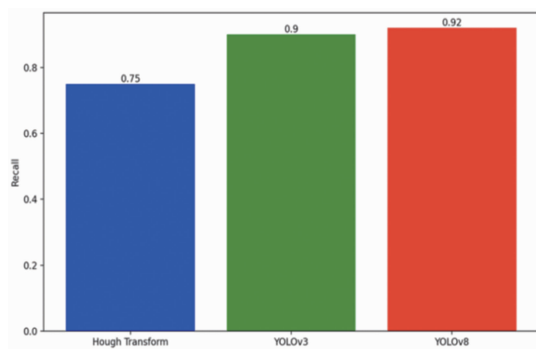


图 14 三种目标检测方法的召回率对比

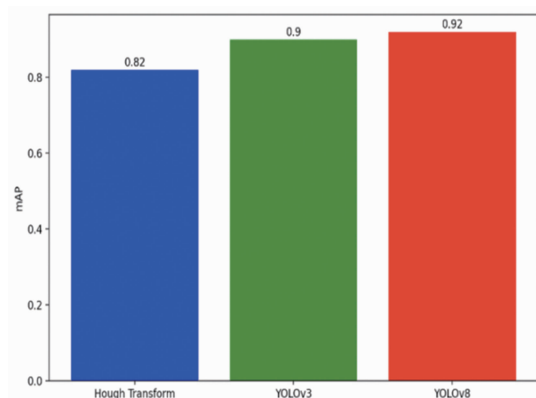


图 15 三种目标检测方法的 mAP 对比

3.5 推理应用

本文设计的 YOLOv8 优化模型的最终目的是实现视觉测量合作目标点状编码的精确定位提取，并通过图像匹配技术实现对物体运动的精确追踪。图 16 所示为本文模型在实际应用中的性能，验证了其在复杂环境下对圆形合作目标进行精确定位的能力。

由图 16 可知，本文模型在各种测试图像上都可以准确标记出圆形合作目标的边界，并给出置信度评分。即使圆形合作目标的一部分被遮挡或处于低对比度的背景中，模型仍然能够准确识别并准确标记其位置。测试结果表明，模型的检测能力不会受到环境变化的显著影响。这对于实际应用场景中的算法鲁棒性至关重要。

除了定位准确性以及算法鲁棒性，图像特征点匹配技术在合作目标的运动跟踪中也起到关键作用。在实际应用中，合作目标可能会发生移动或旋转。通过与图像匹配技术相结合，

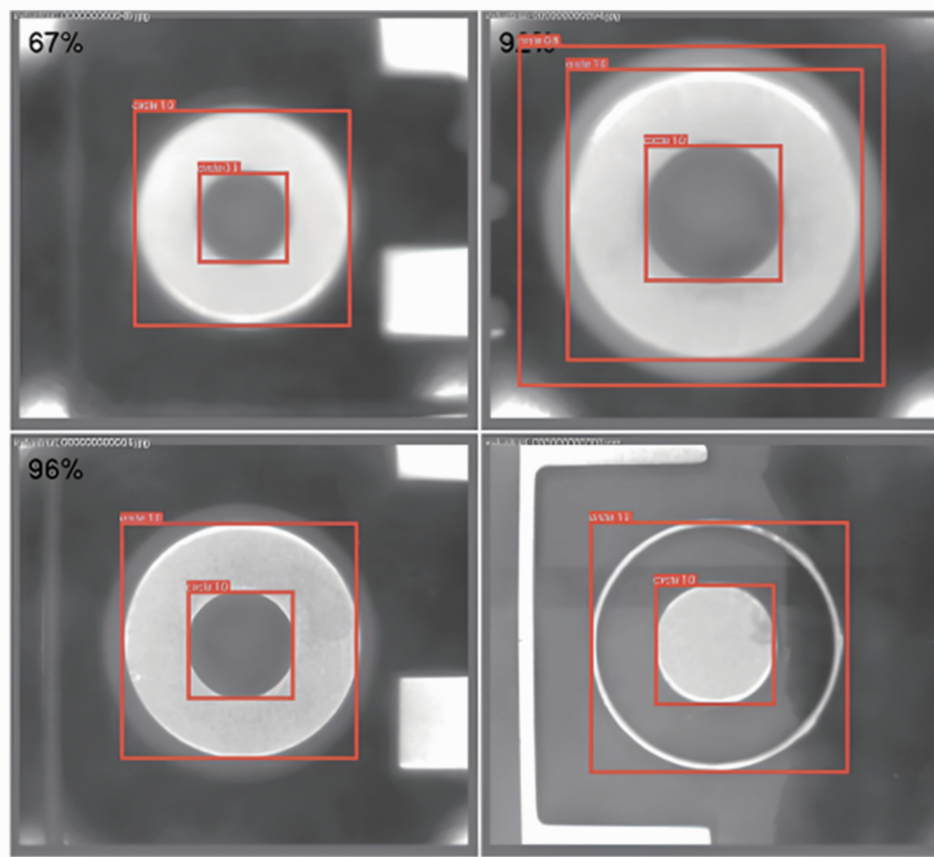


图 16 本文设计的 YOLOv8 优化算法在点状编码物体定位中的应用

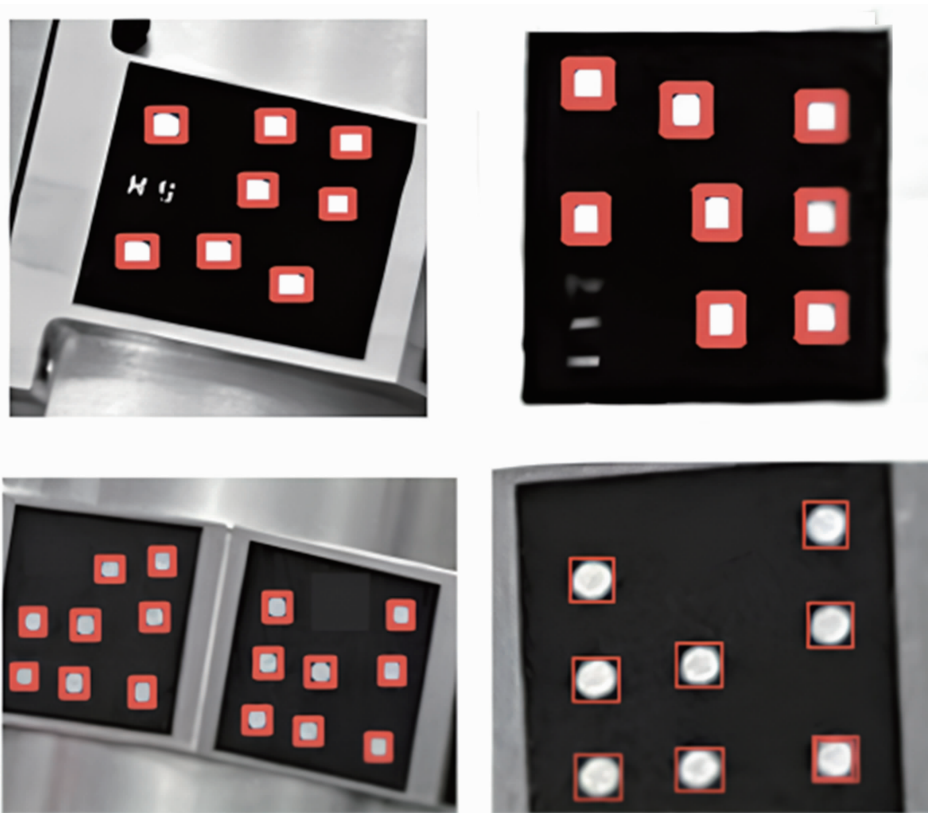


图 17 本文设计的 YOLOv8 优化算法在视觉测量所需的合作目标定位中的应用

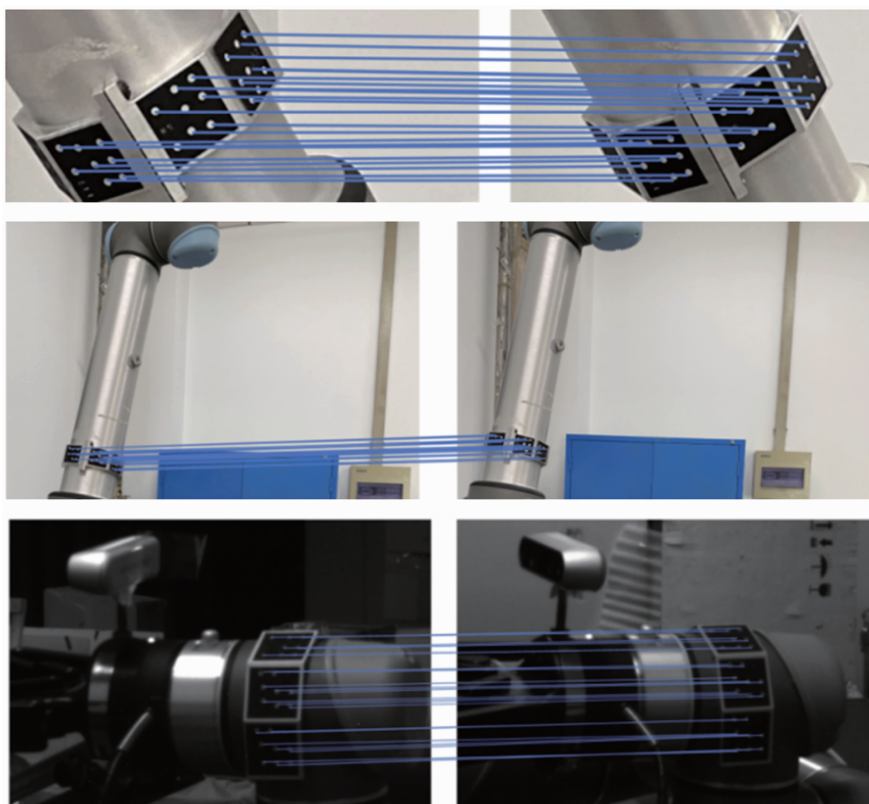


图 18 双目相机左右图像同名特征点匹配

本文的优化模型能够将合作目标识别为在连续图像序列中的运动。同时,当合作目标尺寸因远近变化而发生改变时,模型仍然能够保持对合作目标的稳定识别,在更为准确地提取圆心后,有助于实现更高精度的特征点匹配(见图 17)。

如图 18 所示,利用本文设计的 YOLOv8 模型可以对合作目标的圆心位置进行精确定位,同时可以有效提取各个圆形合作目标的中心点,从而在复杂背景或者不同尺寸和视角下出现的物体中,保持对合作目标圆心的稳定匹配。

通过在多样化的环境条件下进行全面的测试分析,本文验证了所设计模型的准确性、鲁棒性和适应性。这些特性使得该模型成为解决实际目标检测和物体追踪定位任务的理想选择。

4 结束语

本文基于卷积神经网络对 YOLOv8 算法进行优化设计,通过构建视觉测量中所需的圆形合作目标数据集,采用四种目标评价指标对所设计的模型进行评价。经实验证明,本文提出的 YOLOv8 优化算法相较于霍夫变换和 YOLOv3 等传统检测方法,具有更高的准确度、更好的召回率以及更高的平均精度。

本文设计的模型可应用于背景复杂、光照多变和目标形状微变等复杂条件下的合作目标识别与定位。同时,无论是静态图像还是动态场景,模型都能够准确识别并定位合作目标圆心。这对于需要精确操作的自动化工业应用至关重要。模型在图像特征点匹配和运动跟踪方面的适用性,为更多应用领域提供了有力支持。

尽管本文设计的模型在各种条件下都有优异的性能,但是对于遮挡严重、光照极度不均或动态快速变化的场景,仍需要进一步优化和调整。未来的工作可能包括:深入研究更高级的图像增强技术,以提高模型对复杂场景的泛化能力;探索新的网络结构,以进一步提升模型的速度和精度。

参考文献

- [1] Ma M. Infrared pedestrian detection algorithm based on multimedia image recombination and matrix restoration [J]. *Multimedia Tools and Applications*, 2019, **79**(13): 1–16.
- [2] Wahheed A, Hassan A, Saqib S. Circle detection using morphological operations [J]. *Lahore Garrison University Research Journal of Computer Science and Information Technology*, 2019, **3**(2): 31–34.
- [3] 孟祥周,李映君,王桂从,等.融合卷积块注意力模块和 Siamese 神经网络的人脸识别算法[J]. *光学精密工程*, 2023, **31**(21): 3192–3202.
- [4] 朱齐丹,李小铜,郑天昊.舰载机位姿实时视觉测量算法研究[J]. *智能系统学报*, 2021, **16**(6): 1045–1055.
- [5] 谭康霞,平鹏,秦文虎.基于 YOLO 模型的红外图像行人检测方法[J]. *激光与红外*, 2018, **48**(11): 1436–1442.
- [6] 程焕新,矫立浩,骆晓玲,等.改进 YOLOv8 的遥感图像检测算法[J/OL]. *无线电工程*, <http://kns.cnki.net/kcms/detail/13.1097.TN.2024-0118.0934.002.html>, 2024.
- [7] Talaat F M, ZainEldin H. An improved fire detection approach based on YOLO-v8 for smart cities [J]. *Neural Computing and Applications*, 2023, **35**(28): 20939–20954.
- [8] Zhu L, Xiong J, Xiong F, et al. YOLO-Drone: Airborne real-time detection of dense small objects from high-altitude perspective [J]. *Journal of Latex Class Films*, 2021, **14**(8): 1–12.
- [9] Velásquez R A, Lara J V M, Velásquez R A. Yolo algorithms to improve the detection of personal protective equipment thought industrial computer vision [C]. Lima: 2023 IEEE XXX International Conference on Electronics, Electrical Engineering and Computing (INTERCON), 2023.
- [10] 臧天琪.基于图像增强及改进 YOLO 的水下垃圾检测研究[D].上海:上海海洋大学, 2023.
- [11] 董林鹭,赵良军,杨明中,等.一种基于 Viola-Jones 的改进人脸定位算法[J]. *计算机技术与*

- 发展, 2020, **30**(10): 42-46.
- [12] 涂从刚, 毋新房, 王志颖. DPM 数字工业摄影测量系统在高压输水管道加压爆破试验变形监测中的应用 [J]. *测绘标准化*, 2023, **39**(4): 187-193.
- [13] 马啸, 邵利民, 金鑫, 等. 改进的 YOLO 模型及其在舰船目标识别中的应用 [J]. *电讯技术*, 2019, **59**(8): 869-874.
- [14] 于新善, 孟祥印, 金腾飞, 等. 基于改进 Canny 算法的物体边缘检测算法 [J]. *激光与光电子学进展*, 2023, **60**(22): 221-230.
- [15] 张志俊, 吴庆阳, 邓亦锋, 等. 基于霍夫变换的结构光场 3 维成像方法 [J]. *激光技术*, 2023, **47**(4): 492-499.
- [16] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]. *32nd International Conference on Machine Learning*, 2015.
- [17] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的 YOLO 目标检测综述 [J]. *电子与信息学报*, 2022, **44**(10): 3697-3708.
- [18] Jiang P, Ergu D, Liu F, et al. A review of Yolo algorithm developments [J]. *Procedia Computer Science*, 2022, **199**: 1066-1073.
- [19] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8 [J]. *arXiv*: 2305.09972, 2023.
- [20] Loshchilov I, Hutter F. Decoupled weight decay regularization [J]. *arXiv*: 1711.05101, 2017.