

基于深度学习的型钢建材智能计数与在线部署

陈宇韬¹, 陈隽^{2*}, 陈文豪³, 黄茜², 李洋⁴, 刘思媛¹

(1. 新疆大学 建筑工程学院, 乌鲁木齐 830047; 2. 同济大学 土木工程学院, 上海 200092;
3. 中南建筑设计院股份有限公司, 武汉 430071; 4. 中国联合工程有限公司, 杭州 310052)

摘要: 建材管控是建筑智能施工建造的重要一环, 由于型钢截面形状多样、摆放时相互堆叠遮挡且堆放背景环境复杂等因素, 其智能实时计数的研究与应用相对于钢筋、钢管等圆形截面建材仍充满挑战。对此, 本文提出了一种基于改进YOLOv7的非圆形截面建材的计数模型。首先拍摄大量工字钢、方钢管和木枋的现场照片, 结合数据增强手段获得14 950张图像, 共包含1 290 210个端面。进而在现有YOLOv7架构中新增一阶段实例分割模块, 与目标检测任务并行操作, 再通过改进骨干网络和检测头、加入注意力机制、修改损失函数和训练策略等提升模型检测精度, 最后采用轻量级卷积模块减少特征冗余以实现模型轻量化并提升检测时效。试验结果表明, 三种经过改进的建材计数模型的检测准确率均达到90%以上, 满足使用要求。最后, 将上述算法集成到小程序并已成功部署, 使用户通过手机拍照即可完成型钢的在线实时计数工作。

关键词: 目标检测; 目标分割; 施工建材; 非圆形截面; 实时计数; 模型检测精度

中图分类号: TU712; TP391

文献标志码: A

文章编号: 1671-9379(2025)06-0104-10

DOI: 10.13969/j.jzgjz.20240111002

Intelligent Counting and Online Deployment of Steel Building Materials Based on Deep Learning

CHEN Yutao¹, CHEN Jun^{2*}, CHEN Wenhao³, HUANG Qian², LI Yang⁴, LIU Siyuan¹

(1. College of Architecture and Civil Engineering, Xinjiang University, Urumqi 830047, China; 2. College of Civil Engineering, Tongji University, Shanghai 200092, China; 3. Central South Architectural Design Institute, Wuhan 430071, China; 4. China United Engineering Corporation Limited, Hangzhou 310052, China)

*Corresponding author: cejchen@tongji.edu.cn

Abstract: Building materials control is an important part of building intelligent construction, where section steel pose challenges compared with rebar and steel pipe due to its diverse cross-sectional shapes, mutual stacking and occlusion of end faces, and complex background environments in intelligent real-time counting. This study proposes a counting model for non-circular cross-section building materials based on improved YOLOv7. First, a large number of field photographs of I-beams, square steel tubes, and wooden keels were taken and augmented to 14,950 photographs, with a total of 1,290,210 end faces, using data enhancement. Then, a one-stage instance segmentation module was added to the YOLOv7 network to achieve instance segmentation and target detection tasks in parallel. Furthermore, the original YOLOv7 network was enhanced by refining the backbone network and detection head, introducing an attention mechanism, modifying the loss function, and implementing training strategy to improve the model's

收稿日期: 2024-01-11

基金项目: 同济大学2022年度学科交叉联合攻关项目(2022-3-YB-06)

作者简介:

陈宇韬(1997—), 男, 硕士研究生, 主要从事智能建造方面的研究。E-mail: 107552204528@stu.xju.edu.cn。

通信作者:

陈隽(1972—), 男, 博士, 教授, 主要从事智能建造、工程结构大数据防灾方面的研究。E-mail: cejchen@tongji.edu.cn。

detection accuracy. Finally, a lightweight convolution module was used to reduce feature redundancy, enabling effective compression of the model and improving the detection time. The results indicate that the detection accuracy of all three improved building material counting models exceeds 90%, meeting the usage requirements. The above algorithm was integrated into an applet, which has been successfully deployed so that users can complete the real-time counting work by taking photos with their mobile phones.

Keywords: target detection; instance segmentation; building material; non-circular section; real-time counting; detection accuracy of model

2022年,我国建筑业总产值超过31万亿元,其增加值占国内生产总值的6.9%,充分体现了其在国民经济中的重要地位。经济快速发展的同时,建筑业也面临着高成本、高能耗、高污染等现实问题。对此,2020年住房和城乡建设部等十三部门联合印发《关于推动智能建造与建筑工业化协同发展的指导意见》,明确指出要建立智能建造与建筑工业化协同发展的产业体系,不断提升建筑工业化、数字化、智能化水平^[1]。通过深度融合人工智能、机器视觉^[2]和大数据等新技术,智能建造正在为建筑行业带来颠覆性的变革。

施工物料数量的智能管控是建筑智能建造施工阶段的重要功能要求^[3],它能够影响成本控制、施工进度和管理精准度,对此研究者们提出了多种智能计数方法来解决人工计数低效、易错的困难。XIN等^[4]提出一种自动分割钢筋图像的方法,通过Sobel算子检测边缘和最大类间方差方法分离钢筋和背景,采用改进的霍夫变换获取钢筋位置。WU等^[5]提出了一种基于机器视觉的在线钢筋计数和分割方法,利用凹点匹配进行分割,采用K级容错进行计数,并通过视觉反馈实现自动多次分割。GHAZALI等^[6]提出适用于圆形和矩形钢材的计数方法,首先对两种钢材进行图像预处理和边缘检测的操作,接着对其中圆形钢筋进行霍夫变换和后处理,而矩形钢材则通过形态学操作来定位。WANG等^[7]提出了一种基于图像处理和机器学习的自动钢筋计数系统,尝试了多种机器学习算法并成功将原始的目标检测问题转化为更易解决的图像分类问题。FAN等^[8]提出了卷积神经网络-距离聚类(CNN-DC)框架用于钢筋计数,其中卷积神经网络(CNN)用来检测钢筋的候选中心点,距离聚类(DC)算法用于聚类候选中心点并定位真实中心。LI等^[9]在原有YOLOv3网络基础上引入了额外的特征金字塔进行特征提取、利用交并比(IoU)损失结合焦点(Focal)损失以及一系列的优化方法进行训练,有效提升了对于密集和堆叠钢筋的检测性能。LI等^[10]还提出了一种基于视觉的BMC-YOLO模型,可快速准确地对圆钢管进行计数,在大规模密集圆钢管数据集上验证了该算法的优越性。

上述研究实践表明,基于机器学习/深度学习技术的智能计数方法,具有很好的检测精度和极高的效率。然

而目前绝大多数研究针对的都是钢筋^[11]、圆钢管等圆形截面建材,对于工字钢、方钢管等非圆形截面材料的计数研究还不充分。此外,很多算法仍停留在论文纸面,缺乏普通技术人员可直接使用的便捷工具,故无法通过实践来检验成果,进而无法体现其研究价值。

针对上述问题,本文采用改进的YOLOv7模型解决非圆形截面的建材计数问题,同时开发了小程序,方便使用者直接在线进行计数操作。

1 深度学习模型训练集的建立

1.1 非圆形截面建材图片数据集

深度学习模型的训练效果直接依赖于训练样本的数量和特征丰富性^[12],目前尚未有针对非圆形截面建材的公开数据集。为此,本文在生产车间、建材市场和建筑工地等实际场景共拍摄工字钢图片501张、方钢管图片538张、木枋图片456张(表1)。图片涵盖不同光照条件、摆放角度、密集程度和堆放背景环境复杂程度。拍摄中有意选用了不同型号的相机和手机以增强未来模型的鲁棒性。

表1 各类建材数量统计

Table 1 Statistics on the quantity of all kinds of building materials

图片类别	原始图片数量/张	增强后图片/张	增强后截面数量/个
工字钢	501	5 010	185 780
方钢管	538	5 380	569 700
木枋	456	4 560	534 730

采用多边形标注框对收集到的照片中目标进行标注,这是不同于圆形截面钢材常用矩形标注框的重要变化。因为真实场景下型钢等建材常常倾斜放置,矩形框标注会产生大量的重叠区域和冗杂区域(图1),形成图像噪声,导致模型训练时难以准确捕捉目标的真实边界和特征。多边形标注框则可精准地反映出工字钢的轮廓,并让模型更好地学习真实场景中建材的形状,提高检测的准确性。

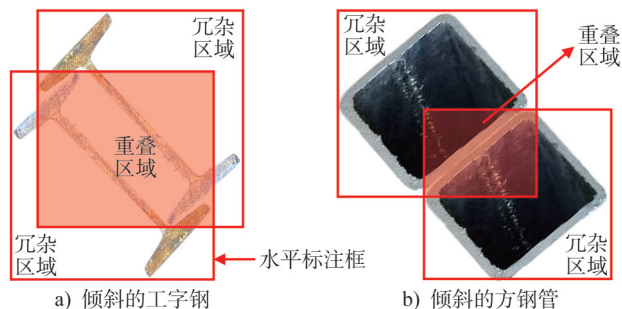


图1 水平框标注建材

Fig. 1 Using horizontal frame to label building materials

1.2 数据集增强

为了防止过度拟合并提升模型的泛化能力,对原始建材图片进行数据增强,手段包括尺度变换(如图片水平/竖直翻转、角度旋转)、色彩抖动(改变亮度、对比度、饱和度)以及添加噪声和随机点(图2)。经过数据增强后,工字钢、方钢管、木枋的图片数量分别为5 010、5 380和4 560张,包含截面总数分别为185 780、569 700和534 730个(表1)。

2 深度学习模型的建立与训练

工字钢、木枋、方钢管三类建材中,木枋和方钢管均为较易计数的矩形截面^[13],而工字钢型号繁多、截面复杂且交错堆叠,因此计数最为困难。本节以工字钢为例说明计数模型的建立过程和训练策略。

2.1 模型选择

YOLOv7作为新一代单阶段目标检测算法,相较于之前的YOLO系列,在精度和速度两个方面都有了显著提升^[14]。YOLOv7基本模型包含输入、骨干网络、检测头三部分:输入部分完成图片的读取和预处理,并转化统一格式用于后续处理;骨干网络利用残差结构,通过反复采样来提取目标的特征信息;检测头输出不同比例和数量的检测框,并给出每个检测框的类别和位置置信度预测。

2.2 工字钢网络模型的改进

图3是本文针对非圆形截面建材计数改进的YOLOv7模型结构,主要是添加了实例分割模块、CBAM^[15]注意力机制,变更激活函数为FReLU^[16]以及替换损失函数为SIoU^[17]。同时采用轻量级卷积模块GSconv^[18],实现模型轻量化,减轻计算负担并提高部署效率。本节详述各项改进目的和效果。

2.2.1 实例分割模块

在实际场景中,非圆形截面建材经常出现相互重叠和遮挡的情形,若仅依赖传统的目标检测算法会造成大

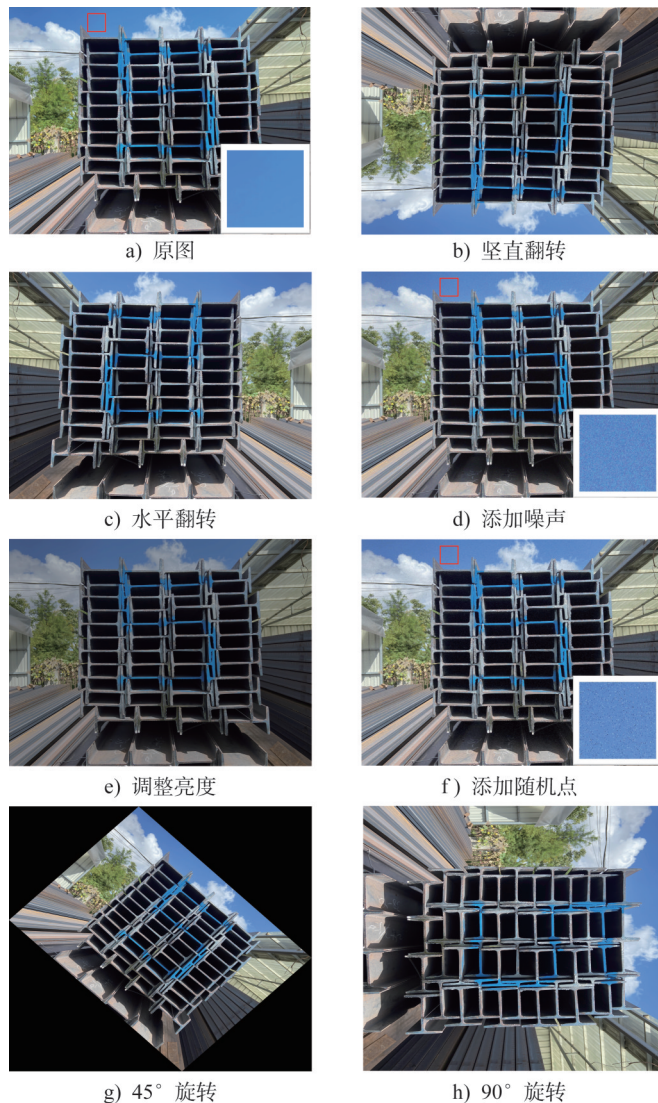


图2 数据增强后的建材图像

Fig. 2 Image of building materials after data enhancement

量漏检。对于工字钢,实例分割能够更准确地捕捉其复杂的轮廓和边界,因为它不仅仅提供了工字钢的位置信息,还提供了工字钢内每个像素的掩码。因此,在YOLOv7基础结构上,借鉴YOLACT^[19]的思想,用两个并行子网络(ProtoNet模块、Prediction Head分支)来实现分割。其中,Prediction Head分支用于生成各候选框的类别置信度、掩码系数、位置回归信息。用于生成一组“原型掩码(ProtoMask)”的ProtoNet模块,是学习到的一组关于工字钢基础的、通用的分割模式,可在不同工字钢之间共享,并可将它们与实例掩码系数线性组合来生成每一个工字钢的实例掩码。最后为了改善对小目标的分割效果,推理时会根据检测框先进行裁剪,再经阈值化后得到最终结果。相较于传统的两阶段分割方法,新方法先定位目标边框再在边框内进行分割,YOLACT的并行处理方式显著提高了计算效率,更适用于工字钢的计数场景。

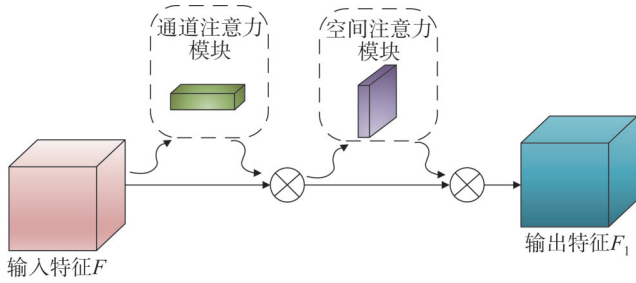


图4 CBAM模块结构图

Fig. 4 CBAM module structure

$$F^* = M_C(F) \otimes F \quad (1)$$

$$F_1 = M_S(F^*) \otimes F^*$$

式中： $M_C(F)$ 为 F 经过特征注意力的输出权值， $M_S(F^*)$ 为 F^* 经过特征注意力的输出权值， \otimes 为特征图加权乘法运算。

2.2.3 FReLU激活函数模块

工字钢是一种截面形状复杂的建材，而且由于切割原因，截面边缘存在“卷边”现象，进一步阻碍了模型对工字钢角点和边缘特征的识别。这就要求深度学习应能提高捕捉这些高频特征的敏感性。FReLU是一种专门针对视觉任务的激活函数，在分类、检测和分割任务中让模型都取得了很好的性能提升。将原本的Sigmoid函数替换为FReLU，在增加很小的内存开销下，使激活函数具备空间敏感性。FReLU公式如下：

$$f(x_{c,i,j}) = \max(x_{c,i,j}, \mathbb{N}(x_{c,i,j})) \quad (2)$$

$$\mathbb{N}(x_{c,i,j}) = x_{c,i,j}^w \cdot p_c^w$$

式中： $\mathbb{N}(x_{c,i,j})$ 为漏斗式条件， $x_{c,i,j}^w$ 表示第 c 个通道上，以2D位置 (i,j) 为中心的窗口， p_c^w 表示此窗口在同一通道中共享的参数。漏斗式条件是对ReLU的拓展，使其成为依赖于空间上下文的2D条件。进一步使网络能够为每个像素生成非线性激活中的空间条件。

式2中 $\max(\cdot)$ 函数为每个像素提供了一个是否考虑其周围空间上下文的选择，让FReLU可以拥有像素级的建模能力。以一个包含 n 个FReLU层 $\{F1, F2, F3, \dots, Fn\}$ 的网络为例，每个FReLU层都有一个 $k \times k$ 参数窗口，不考虑卷积层的影响。在第一层F1之后，每个像素的感受野集合为 $\{1, 1+r\}$ ，其中 $(r=k-1)$ ；随着层数增加到第 n 层Fn，感受野集合变为 $\{1, 1+r, \dots, 1+nr\}$ ，让每个像素拥有更多的选择，通过足够大的 n ，能够拟合任何形状。以工字钢为例，不同尺寸的正方形可以不断逼近工字钢的边界(图5)。

2.2.4 SIoU损失函数模块

YOLOv7模型采用的损失函数是CIoU(Complete Intersection over Union)^[20]，它考虑了边界框回归过程中

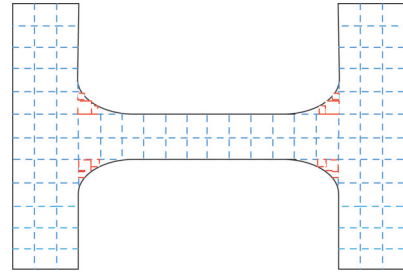
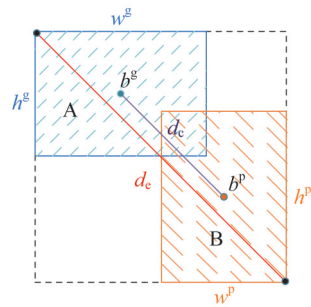


图5 对工字钢截面形状的像素级建模能力

Fig. 5 Pixel-level modeling capability for I-beam cross-sectional shapes

目标框的重叠区域、中心点的距离以及长宽比对训练的影响，如图6所示。



注： b^g 和 b^p 为预测框A和真实框B的中心点， d_c 表示两个框中心点间的距离， d_c 表示两个框相隔最远两点间的距离， (w^g, h^g) 和 (w^p, h^p) 分别代表预测框、真实框的高和宽。

图6 CIoU损失示意

Fig. 6 Diagram of CIoU loss

从图6中可以看出CIoU并没有考虑预测框与目标框之间的方向性差异，是导致工字钢计数模型收敛速度较慢和准确性较低的关键问题。本研究通过引入SIoU改进上述问题，可同时考虑预测框和真实框间的角度损失、距离损失、形状损失和IoU损失(I_{oU})。其损失函数(L_{SIoU})计算公式如下：

$$L_{SIoU} = 1 - I_{oU} + \frac{\Delta + \Omega}{2} \quad (3)$$

角度损失 Δ 示意图(图7)和计算公式如下：

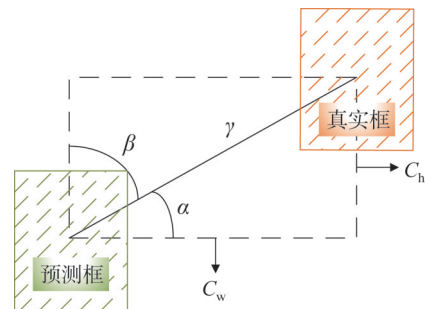


图7 角度损失示意图

Fig. 7 Diagram of angle loss

$$\Lambda = 1 - 2\sin^2(\arcsin(x) - \frac{\pi}{4}) \quad (4)$$

$$x = \frac{c_h}{\gamma} = \sin(a) \quad (5)$$

$$\gamma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \quad (6)$$

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \quad (7)$$

式中: $(b_{c_x}^{gt}, b_{c_y}^{gt})$ 为真实框的中心坐标, (b_{c_x}, b_{c_y}) 为预测框的中心坐标。

距离损失 Δ 的计算公式如下:

$$\Delta = \sum_{t=x,y} (1 - e^{-(2-\Lambda)\rho_t}) \quad (8)$$

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2 \quad (9)$$

$$\rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2 \quad (10)$$

式中: c_w 和 c_h 为连接预测框和真实框中心点对角构造的矩形的宽度和高度; Λ 为角度损失。

形状损失 Ω 的计算公式如下:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (11)$$

$$\omega_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})} \quad (12)$$

$$\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (13)$$

式中: ω 和 h 为预测框的宽度和高度, ω^{gt} 和 h^{gt} 为真实框的宽度和高度。 ω_w 和 ω_h 为对宽度和高度进行归一化处理后的结果, 目的是消除预测框与真实框的尺寸差异影响。 θ 是一个可调参数, 它用于调控形状损失占总损失的比例, 在训练时 θ 值设置为 4。

IoU 损失公式如下:

$$I_{ou} = \frac{|A \cap B|}{|A \cup B|} \quad (14)$$

综上, SIoU 首先是引入了角度损失, 有助于迅速推动模型收敛, 促使预测框快速对准 x 轴或 y 轴, 然后沿着相应轴逐步向真实框逼近。其次, SIoU 重新定义了距离和形状损失, 有助于提高模型的准确性和效率。因此, SIoU 可以更好地完成工字钢的计数任务, 提升检测结果的质量。

2.2.5 模型轻量化

通过上述改进, 工字钢计数模型已实现精准计数的目标。但为了降低前端反馈时间, 改善用户的使用体验, 本文采用轻量卷积模块 Ghostconv 替代普通卷积模块, 减少网络参数规模。

卷积神经网络的中间特征图通常包含很多冗余信息, 其中一些特征之间存在相似性。这些相似的特征通

过简单的线性变换就能得到, 无需经过复杂的非线性变换。Ghostconv 模块在运行时会使用少量卷积核生成原始特征图, 再通过线性变换生成更多 Ghost 特征图。标准卷积和 Ghostconv 模块运算的 FLOPS 计算公式为式(15)~(16), 从计算公式中可以看出采用 Ghostconv 替换传统卷积能将参量缩减为 $1/s$:

$$F_{LOPS}(C) = c \times k^2 \times n \times h \times w \quad (15)$$

$$F_{LOPS}(G) = c \times k^2 \times n/s \times h \times w + n/s \times d^2 \times (s-1) \times h \times w \quad (16)$$

式中: k 为卷积核尺寸; c 为工字钢输入特征图通道数; n, h, w 表示工字钢输出特征图通道数、高度、宽度; d 为线性操作的内核大小; s 为原始工字钢特征图线性变换生成的相似特征图数量。

2.3 模型训练策略

深度学习模型通常采用多种训练策略, 如预设先验框尺寸、迁移学习和学习率调整。

先验框(或锚点框)是目标检测任务中的一组固定大小和长宽比的框, 通过提供形状的先验信息, 引导模型更准确地定位目标。K-means 算法通过不断迭代更新聚类中心, 使每个目标标注框都分配给交并比最大的聚类中心。计算公式如下:

$$D = \min \sum_{i=1}^n \sum_{j=1}^k \left[1 - \frac{box_i \cap cen_j}{box_i \cup cen_j} \right] \quad (17)$$

通常以准确率(Accuracy, 用 A 表示)作为聚类结果的评价指标, 计算公式如下:

$$A = \frac{1}{n} \sum_{i=1}^n \max \left(\frac{box_i \cap cen_1}{box_i \cup cen_1}, \frac{box_i \cap cen_2}{box_i \cup cen_2}, \dots, \frac{box_i \cap cen_k}{box_i \cup cen_k} \right) \quad (18)$$

式中: box_i 为目标检测中的第 i 个标注框的面积, cen_j 是第 j 个聚类中心的面积, n 为检测目标的数量, k 为聚类中心个数。

通过 K-means 算法对工字钢样本包围框进行聚类, 综合考虑计算开销和准确性, 选定最优的 K 值为 9, 并把对应的聚类中心坐标输入到 YOLOv7 预测网络的特征层中, 达到加快收敛速度并提高检测精度的效果。结果如表 2 和图 8 所示。

迁移学习是借助已学习相关任务的知识来进行新任务学习的方法。在训练建材计数模型过程中直接采用 COCO 数据集上的预训练权重来进行权重初始化, 有助于模型快速学习图像中的通用特征, 提高模型的泛化能力, 加速收敛过程。

学习率是深度学习模型训练时用于控制模型权重梯度更新的重要参数。学习率设定过高可能导致训练过程振荡而不收敛, 而设定过低则可能大幅延长收敛时间。

表2 工字钢聚类结果
Table 2 I-beam clustering results

聚类中心数目	准确率	聚类中心(长像素,宽像素)
K=6	72.07%	(46,78), (66,115), (72,42), (96,151), (109,76), (171,103)
K=9	78.14%	(32,64), (56,124), (71,84), (77,42), (91,148), (104,69), (125,94), (178,126), (190,76)
K=12	79.91%	(19,44), (48,84), (60,139), (72,41), (82,105), (87,69), (98,161), (116,85), (133,58), (144,102), (189,135), (209,81)

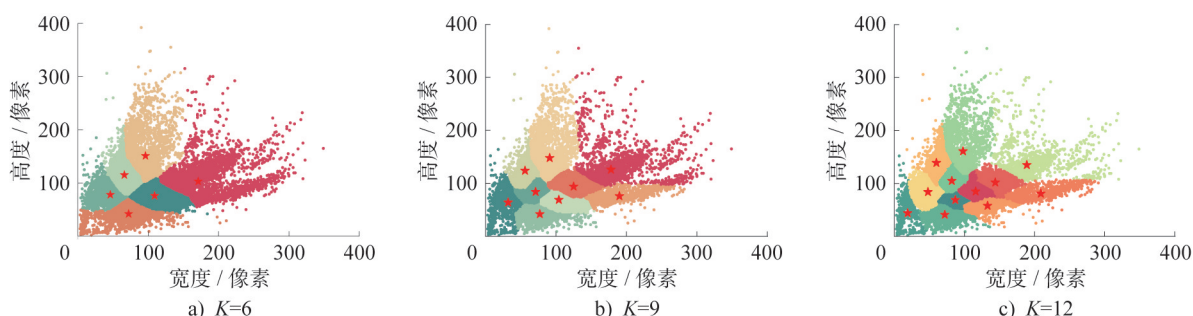


图8 工字钢样本包围框 K-means 聚类

Fig. 8 K-means clustering of I-beam samples surrounded by boxes

因此,制定合适的学习率调整策略对于模型的训练至关重要。本文采用带预热的余弦退火的方式来动态调整学习率(图9),通过前期探索性预热和后期精确调整,提升训练效能。最大和最小学习率设置为0.000 2和0。计算公式如下:

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos\left(\frac{t - t_w}{T_{\max} - t_w} \pi\right) \right) \quad (19)$$

式中: t 表示当前迭代步数; η_{\max} 和 η_{\min} 分别为该周期内学习率的最大值和最小值; t_w 是预热阶段结束时的迭代步数; T_{\max} 是整个训练周期(不含预热阶段)的迭代总数。

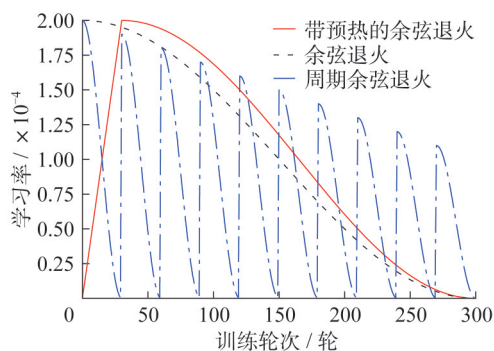


图9 学习率曲线

Fig. 9 Learning rate curves

3 模型应用效果

3.1 试验环境与评价指标

该实验硬件配置为天河二号超级计算机的 a800 显

卡,显存为 80GB。深度学习框架为 PyTorch1.10.2, CUDA11.3。CPU 为 Intel Core i7,内存为 16GB。训练集和验证集划分为 4:1。模型训练参数设置为:输入图像分辨率统一为 640×640 像素,批量大小为 8,训练轮次为 300 轮。采用 SGD 优化器,初始学习率为 0.01,动量因子为 0.937。

单一的评价指标并不能全面客观的反映模型表现。通过综合使用平均精度 A_p (average precision, 用 A_p 表示)、平均最大绝对误差 M_{ac} (mean absolute error, 用 M_{ac} 表示)和平均均方根误差 R_{mse} (root mean squared error, 用 R_{mse} 表示)这三个指标,能够更全面、准确地衡量非圆形截面建材计数模型的整体性能。其中, A_p 侧重于对建材的定位和识别能力,而 M_{ac} 和 R_{mse} 则更关注预测的建材数量是否准确。它们的计算公式如下:

$$A_p = \int_0^1 p(r) dr \quad (20)$$

式中: A_p 是 P-R (precision-recall) 曲线与坐标轴围成的面积值。在本文中使用 $A_p@0.5:0.95$, 即表示在不同 IoU 阈值(从 0.5 到 0.95, 步长 0.05)上的平均 A_p 。

$$M_{ac} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (21)$$

$$R_{mse} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (22)$$

式中: n 为测试集中的图片数量, y_i 为图片中建材的真实数量, \hat{y}_i 为模型预测的建材数量。

3.2 试验结果分析

本文针对YOLOv7模型进行了多项改进。为了评估每项改进对工字钢计数模型的具体影响,在自建数据集上进行了消融对比试验。表3比较了采用模型训练策略和在网络结构中的各项改进对工字钢模型检测和分割结果的影响。结果表明,采用模型训练策略,引入CBAM、SIoU、FReLU和Ghostconv模块在检测和分割

两方面均能对模型性能产生正向优化。对于检测, $A_p@0.5:0.95$ 差值最高达到5.24%, M_{ae} 差值最高达到1.81, R_{mse} 差值最高达到2.34。对于分割, $A_p@0.5:0.95$ 差值最高达到4.73%。值得注意的是,引入Ghostconv轻量化卷积模块不仅能在一定程度上提升模型的精度,还有效降低模型的参数量,降幅超过10%,从而抵消了其它改进带来的因参数量增加对推理速度造成的不利影响。

表3 各项改进对工字钢检测和分割性能影响结果

Table 3 Results of the effect of various improvements on the detection and segmentation performance of I-beam

Base (v7-seg)	模型训练策略	+CBAM	+SIoU	+FReLU	+Ghostconv	目标检测 $A_p@0.5:0.95$	实例分割 $A_p@0.5:0.95$	M_{ae}	R_{mse}	P_{arams}/M
✓						83.13%	57.26%	4.77	7.86	37.86
✓	✓					88.37%	61.99%	2.96	5.52	37.86
✓	✓	✓				91.33%	63.32%	2.18	3.98	37.86
✓	✓	✓	✓			94.79%	65.29%	1.89	3.09	37.87
✓	✓	✓	✓	✓		96.45%	66.35%	1.51	2.78	38.17
✓	✓	✓	✓	✓	✓	97.13%	67.03%	1.42	2.67	33.58

注:表中 P_{arams} 为模型参数量大小,M为其单位Million的简写。

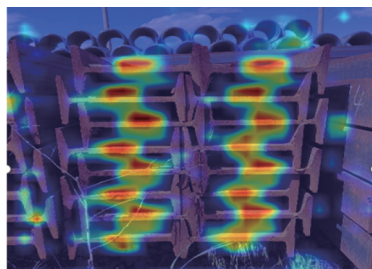
为保证识别结果的直观性和易读性,依然选择以水平矩形框来展示检测效果。同时,运用Grad-CAM++ (Gradient-weighted Class Activation Mapping Plus)^[21]对模型输出特征进行可视化(图10)。从图中可以看出,工字钢所处区域颜色较深而背景区域颜色较浅,该图清晰地反映出模型的关注点很好地聚焦在工字钢上。

表4比较了采用模型训练策略和在网络结构中的各项改进对木枋计数模型检测和分割结果的影响。结果表明,采用模型训练策略,引入SE模块、ASFF模块、SIoU模块、Ghostconv模块在检测和分割两方面均能对模型性能产生正向优化。对于检测, $A_p@0.5:0.95$ 差值最高达到2.92%, M_{ae} 差值最高达到1.85, R_{mse} 差值最高达到2.83。对于分割, $A_p@0.5:0.95$ 差值最高达到6.69%。木枋检测结果如图11。

4 实用工具开发与在线部署

考虑到深度学习模型使用的技术门槛较高,为了进一步提升非圆形截面建材计数模型的实用性,本文设计并研发了微信小程序用于工字钢、方钢管和木枋的智能识别。目前,小程序已成功上架,用户只需在微信内搜索“智能建材计数与尺寸测量”即可免下载、随时使用。

使用时,用户需先选择建材类型,随后采用拍照识别或相册选择的形式上传图片。为确保最佳识别效果,建议在光线充足环境下进行拍摄,同时确保建材轮廓清晰、无明显阴影遮挡。上传前,用户可手动涂鸦或调整裁剪框,编辑修改照片。照片上传至云端服务器后,深度模型完成数量计算,返回结果并在客户端展示(图12)。网络信号正常情况下,照片上传到返回结果的时间大约1秒左右,充分体现了本文改进的YOLOv7模型的效率。



a) 可视化热力图



b) 检测效果

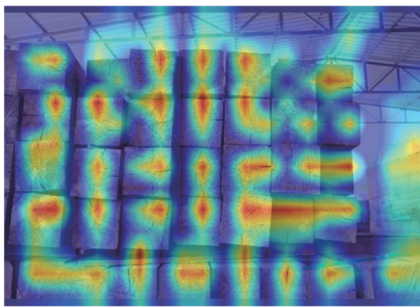
图10 工字钢检测结果

Fig. 10 I-beam detection results

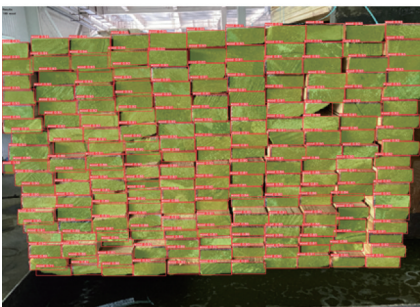
表4 各项改进对木枋检测和分割性能影响结果

Table 4 Results of the effect of various improvements on the detection and segmentation performance of wood keel

Base (v7-seg)	模型训练策略	+CBAM	+SIoU	+FReLU	+Ghostconv	目标检测 $A_p@0.5:$ 0.95	实例分割 $A_p@0.5:$ 0.95	M_{ae}	R_{mse}	P_{arans}/M
✓						85.99%	60.74%	9.87	15.99	37.86
✓	✓					88.91%	67.43%	8.02	13.52	37.86
✓	✓	✓				89.52%	69.01%	6.34	10.69	38.06
✓	✓	✓	✓			90.17%	71.61%	5.69	8.99	38.21
✓	✓	✓	✓	✓		91.32%	72.26%	4.86	8.07	38.22
✓	✓	✓	✓	✓	✓	91.90%	72.89%	4.63	7.82	33.09



a) 可视化热力图



b) 检测效果

图11 木枋检测结果

Fig. 11 Wood keel detection results

5 结论

本文针对非圆形截面建材的智能计数问题,提出了基于改进YOLOv7的检测模型,主要结论如下:

(1) 新模型实现了对工字钢、方钢管和木枋三类非圆形截面建材的高效智能计数。模型的一阶段计数方法和轻量化措施,显著提升了计数的反应时间,满足实际工程中快速计数的需求。

(2) 引入多种措施有效提升了模型的检测精度。数据增强技术提高了训练数据的丰富性,使得模型更具鲁棒性,能够适应不同的场景和光照条件。同时,通过添加分割模块、注意力机制、替换损失和激活函数,提高了计数模型的准确性。

(3) 开发并部署了实用的小程序,真正服务于实际工程建设,为智慧工地建设提供高效、准确的建材计数服务。

本文研究仍有一些不足:受数据样本覆盖面限制,计数模型在特殊或复杂场景下的计数能力仍有待提升。未来需进一步拓宽数据采集范围,包括但不限于不同地理区域、极端天气条件及光照环境,以增强模型的鲁棒性和泛化能力,从而满足实际工程的深度需求。

参考文献:

- [1] 廖玉平. 加快建筑业转型 推动高质量发展——解读《关于推动智能建造与建筑工业化协同发展的指导意见》[J]. 中国勘察设计, 2020(9): 20-21.
- LIAO Yuping. Accelerating the transformation of construction industry and promoting high-quality development: Interpretation of guiding opinions on promoting the coordinated development of intelligent construction and building industrialization [J]. China Engineering Consulting, 2020(9): 20-21. (in Chinese)
- [2] 王亦君, 蒋首超. 基于计算机视觉的钢构件防腐涂层缺陷检测[J]. 建筑钢结构进展, 2023, 25(12): 85-93, 101. DOI: 10.13969/j.cnki.cn31-1893.2023.12.009.
- WANG Yijun, JIANG Shouchao. Defect identification of anti-



图12 智能建材计数与尺寸测量小程序结果示例

Fig. 12 Results of smart building material counting and dimensioning applet

- corrosive coating of steel members based on computer vision [J]. *Progress in Steel Building Structures*, 2023, 25 (12) : 85-93, 101. DOI: 10.13969/j.cnki.cn31-1893.2023.12.009.(in Chinese)
- [3] RAN B D, LI S X, ZHOU B C, et al. Deep learning based detection and counting algorithm for rebar images[C]//2023 35th Chinese Control and Decision Conference (CCDC). New York:IEEE,2023.
- [4] XIN Y, XUE W, YUAN P X, et al. Research on an automatic counting method for steel bars' image [C]//2010 International Conference on Electrical and Control Engineering. New York:IEEE,2010.
- [5] WU Y, ZHOU X F, ZHANG Y C. Steel bars counting and splitting method based on machine vision [C]//2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). New York:IEEE, 2015.
- [6] GHAZALI M F, WONG L K, SEE J. Automatic detection and counting of circular and rectangular steel bars [C]//9th International Conference on Robotic, Vision, Signal Processing and Power Applications. Singapore:Springer Singapore,2016.
- [7] WANG H, POLDEN J, JIRGENS J, et al. Automatic rebar counting using image processing and machine learning [C]//2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER). New York:IEEE,2019.
- [8] FAN Z, LU J W, QIU B Z, et al. Automated steel bar counting and center localization with convolutional neural networks [EB/OL]. [2019-06-19]. <https://arxiv.org/abs/1906.00891v2>.
- [9] LI Y, LU Y J, CHEN J. A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector [J]. *Automation in Construction*, 2021, 124: 103602. DOI:10.1016/j.autcon.2021.103602.
- [10] LI Y, CHEN J. Computer vision - based counting model for dense steel pipe on construction sites [J]. *Journal of Construction Engineering and Management*, 2022, 148 (1) : 04021178. DOI:10.1061/(ASCE)CO.1943-7862.0002217.
- [11] 刘浩, 辛山. 面向密集型钢筋计数的 GCA-MobileNetV2-YOLOv4 算法 [J]. *电子测量技术*, 2023, 46 (9) : 166-174. DOI:10.19651/j.cnki.emt.2210228.
- LIU Hao, XIN Shan. GCA-MobileNetV2-YOLOv4 algorithm for intensive rebar counting [J]. *Electronic Measurement Technology*, 2023, 46 (9) : 166-174. DOI: 10.19651/j.cnki.emt.2210228.(in Chinese)
- [12] 方成, 于盛鑫, 李永刚, 等. 基于深度学习的土木工程计算机视觉健康监测 [J]. *同济大学学报 (自然科学版)*, 2024, 52(2):213-222. DOI:10.11908/j.issn.0253-374x.22370.
- FANG Cheng, YU Shengxin, LI Yonggang, et al. Deep learning-based computer vision for health monitoring in civil engineering [J]. *Journal of Tongji University (Natural Science)*, 2024, 52(2) : 213-222. DOI:10.11908/j.issn.0253-374x.22370.(in Chinese)
- [13] 陈隽, 陈文豪, 李洋. 基于目标检测的施工钢材物料智能实时计数 [J]. *同济大学学报 (自然科学版)*, 2023, 51(11) : 1701-1710. DOI:10.11908/j.issn.0253-374x.22157.
- CHEN Jun, CHEN Wenhao, LI Yang. Intelligent real-time counting of construction materials based on object detection [J]. *Journal of Tongji University (Natural Science)*, 2023, 51(11) : 1701-1710. DOI: 10.11908/j.issn.0253-374x.22157. (in Chinese)
- [14] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York:IEEE,2023.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Computer Vision-ECCV 2018. Berlin:Springer,2018.
- [16] MA N N, ZHANG X Y, SUN J. Funnel activation for visual recognition [C]//Computer Vision-ECCV 2020. Berlin: Springer,2020.
- [17] GEVORGYAN Z. SiO loss: More powerful learning for bounding box regression [EB/OL]. [2022-05-25]. <https://arxiv.org/abs/2205.12740v1>.
- [18] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York:IEEE,2020.
- [19] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT: Real-time instance segmentation [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). New York:IEEE, 2019.
- [20] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7) : 12993-13000. DOI:10.1609/aaai.v34i07.6999.
- [21] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-CAM: Generalized gradient-based visual explanations for deep convolutional networks [C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). New York:IEEE,2018.