

面向少儿木工教育场景的小物体目标检测方法研究

廖 员¹ 黎润霖^{1*} 杨 昌²
LIAO Yuan LI Runlin YANG Chang

摘 要

针对少儿木工教学场景中工具种类多、小目标识别难、遮挡干扰强等问题,文章提出了一种改进的YOLOv11目标检测模型,用于实现木工工具的高精度识别。首先,构建了一个覆盖五类典型工具(锤子、锯子、螺丝刀、锉刀、木工夹具)的图像数据集,涵盖多种光照、角度与遮挡情况。在模型设计方面,在YOLOv11基础上引入SimAM注意力机制以增强局部特征感知,增加高分辨率检测分支以提升小物体检测能力,并采用Focal Loss优化类别不平衡问题。实验结果表明,改进模型在mAP@0.5:0.95上达到74.1%,在所有类别上均取得90%以上的mAP@0.5,显著优于YOLOv5m、YOLOv8m等主流模型。结论表明,所提方法在精度、鲁棒性与实时性方面具有良好表现,具备在实际教学辅助系统中部署应用的潜力。

关键词

目标检测;YOLOv11;SimAM注意力机制;小物体识别;少儿木工教育;Focal Loss

doi: 10.3969/j.issn.1672-9528.2025.11.014

0 引言

近年来,随着STEAM教育理念的深入推广,少儿木工课程作为一种融合科学、技术、工程与艺术的综合实践活动,在国内外中小学教育中得到了广泛关注^[1]。该类课程通过引导儿童亲手操作锯子、锤子、螺丝刀等基础木工工具,不仅有助于提升其动手能力与空间思维能力,还能够激发创造力与工程意识。然而,由于木工活动存在一定的工具操作复杂性与安全风险,加之课堂教师资源有限,如何实现对学生使用工具行为的辅助监控与智能引导,成为当前少儿木工教学中亟须解决的问题之一。

目标检测(object detection)作为计算机视觉领域的重要研究方向,是从图像中识别并定位出预定义类别的目标物体^[2]。近年来,基于深度学习的目标检测算法在精度与速度方面取得了显著突破,尤其是YOLO(you only look once)系列算法,因其端到端训练、检测速度快、部署灵活等优势,被广泛应用于安防监控、自动驾驶、智能制造等多个实际场景^[3-5]。然而,将目标检测算法直接应用于少儿木工教育场景仍面临多项挑战。首先,木工工具种类繁多且尺寸普遍较小,属于典型的小物体检测问题,易受到图像分辨率、光照变化

及遮挡干扰的影响;其次,该类工具在外观上存在一定相似性,容易造成模型误检与漏检;最后,当前公开数据集中缺乏针对儿童木工教学环境下的目标检测数据,严重制约了相关研究的开展。

针对上述问题,本文以YOLOv11为基础,面向少儿木工教育场景中常见的小物体(如锤子、锯子、螺丝刀、锉刀等)识别任务,开展目标检测方法的改进与优化研究。本文的主要贡献包括以下三个方面:

一是结合实际教学需求与场景特征,构建了一个具有代表性的少儿木工工具图像数据集,涵盖五类典型工具,并引入多角度、不同光照与遮挡条件的样本,弥补了该领域公开数据的空白。

二是在YOLOv11模型框架下,引入SimAM^[6]注意力机制以增强空间特征表达能力,设计了小目标检测分支以提高对低分辨率工具的识别精度,并结合Focal Loss^[7]缓解类别不平衡问题,从而系统提升模型在复杂环境下的小物体检测性能。

三是本文研究不仅为教育场景下的目标检测应用提供了理论基础与工程实现路径,也为后续在增强现实、交互式教学等方向的拓展应用提供了技术支撑和方法借鉴。

1 任务分析与数据集构建

随着国家“双减”政策和素质教育理念的推进,少儿创客教育正逐步走入课堂实践,其中以木工为代表的动手类课程,因其能有效提升儿童的动手能力、工程意识和空间感知力,受到教育机构与家长的广泛重视。少儿木工教学一般依

1. 吉利学院智能科技学院 四川成都 641423

2. 四川工商学院艺术学院 四川成都 611745

[基金项目]虚拟现实技术与系统全国重点实验室开放课题基金(VRLAB2025E02);吉利学院校级科研项目(2024xzkzd028)

托专业的课程器材，如安全锤子、儿童锯、螺丝刀、木工尺及多种辅助夹具等工具，学生通过在教师指导下完成基本木工作品，学习基础结构、手工技巧与创造表达。然而，在实际教学中，由于师资人数有限、学生个体操作差异较大，教师难以全面、实时掌握学生使用工具的行为。这种情况下，若能构建一个基于视觉感知的木工工具识别系统，将显著提升教学效率与安全保障能力。

本研究探索深度学习目标检测技术在少儿木工场景中的应用。与常规目标检测任务相比，该场景下主要面临两个突出挑战：一是检测对象以小物体为主，例如锤子、螺丝刀等工具在画面中尺寸占比普遍较小，易受图像模糊、光照变化与遮挡干扰；二是工具种类虽不多但相似性较高，在形态特征上存在一定重叠，例如手锯与木工夹具可能在某些角度下产生混淆。这些特点决定了需要对检测算法在特征提取、尺度适应性等方面进行针对性优化。为此，研究的第一步即为构建一个覆盖常见木工工具种类、具备多场景多角度图像样本的数据集，以支撑模型训练与评估。

在数据采集方面，本研究以实际木工教学器材为基础，选择五种常见工具类别：锤子、锯子、螺丝刀、锉刀和木工夹具。图像采集设备选用了高分辨率手机摄像头与固定式摄像头，采集场景包括标准白底拍摄、课桌环境展示及模拟课堂操作，尽可能还原学生真实使用过程中的视觉环境。在光照方面，涵盖自然光、室内顶灯、背光等多种条件，保证图像数据具有足够的泛化能力。采集完成后，经人工筛选与清洗，获得有效图像共计 1 512 张，每张图像中包含 1~5 个目标对象，目标分布均衡。所有图像均采用 YOLO 系列算法通用的标注格式进行人工标注，具体包括目标类别编号、中心点坐标以及目标框宽高的归一化值。标注工作使用 Labellmg 工具完成，标注人员接受了统一的类别识别与标注规范培训，确保标注结果的一致性与准确性。图像分辨率主要集中在 1 280 px × 720 px 与 1 920 px × 1 080 px 两个级别之间，能够满足主流检测模型对输入尺寸的要求。

数据集划分方面，本文采用图像为单位进行划分，遵循 8:1:1 的比例分别构建训练集、验证集与测试集，保证各类别在不同子集中的分布均衡，并在划分过程中使用随机种子控制过程的可复现性。训练集用于模型的主干学习与参数拟合，验证集用于中期模型调整与调参指导，测试集用于评估最终模型在未知样本上的性能表现。通过对划分结果的统计分析，发现各类别工具在不同子集中的数量差异控制在合理范围内，未出现样本过度集中或类别缺失的问题。

为了提升模型在复杂场景中的鲁棒性，本文对训练集图像应用了多种数据增强策略^[8]。具体包括图像水平翻转、亮度增强、高斯模糊、仿射变换、随机遮挡与对比度调整等方式。这些增强方法能够模拟教学环境中常见的图像干扰因素，

如不同光照条件、学生手部遮挡、背景杂物等，从而增强模型对多样化输入的适应能力。在增强过程中，所有目标框信息同步调整，确保标签与图像保持一致。

图 1 展示了其中一张原始图像与对应的三种典型增强方式：水平翻转、遮挡模拟与颜色抖动。可以看出，在增强后的图像中，目标边界依然清晰可辨，模型训练能够借此学习更加稳健的特征表达。增强样本的引入不仅提升了模型对少样本场景的适应能力，也在一定程度上缓解了小目标在深层网络中被忽略的问题。



图 1 数据集示例

2 模型设计与改进方法

在少儿木工教学场景中，目标检测系统需应对目标尺寸小、类别相似度高、遮挡频繁等问题。尽管当前主流检测算法如 YOLOv8、YOLOv5 在多个公开数据集上均取得了良好性能，但其在小物体密集识别与细粒度分类任务中的表现仍存在一定局限性^[9]。随着 Ultralytics 于 2024 年底发布的 YOLOv11，其在模型结构设计、特征融合机制与训练策略方面相较上一代 YOLOv8 实现了全面升级。YOLOv11 不仅大幅提升了检测精度与推理速度，还特别加强了对小目标的感知能力，展现出更强的通用性与部署灵活性。因此，本文选择 YOLOv11 作为基础检测框架，并结合少儿木工工具的任务需求，从特征提取、注意力机制、检测分支与训练策略等方面进行了针对性优化，以实现复杂场景下小型目标的高效检测。

YOLOv11 整体架构依旧遵循经典的主干网络、颈部网络、检测头三段式结构。其中，主干网络采用新设计的 C3K2 模块作为主要构建单元，该模块通过使用较小卷积核与跨阶段连接结构，实现了对细节特征的高效提取，同时显著降低了参数量。在特征融合阶段，YOLOv11 引入了改进的 SPPF 模

块,通过对不同尺度特征的空间池化操作,增强模型对上下文信息的理解能力。此外,YOLOv11还内置了轻量级的C2f模块与C2PSA模块,使得网络在保持低延迟的同时提升了空间注意力能力。在检测头部分,YOLOv11采用无锚框设计,输出中心点、边界框尺度与类别概率,无需先验框设定,有利于提升模型在非结构化数据下的适应能力。

尽管YOLOv11在整体性能方面已优于前代模型,但在少儿木工工具识别场景中,由于目标普遍尺寸较小、图像遮挡复杂,仅依赖YOLOv11的默认结构仍难以获得理想精度。因此,本文对YOLOv11进行了结构级与策略级的定向改进。首先,在特征提取阶段,为进一步增强模型对小目标的敏感性,与边缘特征的感知能力,本文在主干网络中引入SimAM (simple attention module) 注意力机制。SimAM是一种基于神经元能量最小化思想的轻量级无参注意力模块,其在不显著增加计算开销的前提下,能够对图像中空间位置显著性区域进行强调。该机制基于神经元能量最小化原理,计算每个像素点的重要性分数,其能量函数定义为:

$$E(x_i) = (x_i - \mu_i)^2 + \lambda \cdot \sigma_i^2 \quad (1)$$

式中: μ_i 与 σ_i^2 表示当前通道中其余神经元的均值与方差; λ 表示常数调节因子。SimAM通过对能量的抑制,实现对关键区域的显著增强,尤其适用于锯齿、尖端等细节特征强化。

此外,考虑到原YOLOv11的多尺度检测头设计中主要使用了8倍与16倍下采样特征图,对于占图像面积不足5%的小工具仍存在分辨率不足的问题,本文新增了一个基于4倍下采样特征图的小目标检测分支。该分支通过直接接入主干网络中浅层高分辨率特征图,联合SimAM输出进行融合处理,能够更有效地保留工具边缘、细节纹理等关键线索,进而提高模型对遮挡场景下部分可见目标的识别准确率。同时,颈部网络结构中的SPPF输出也被同步传递至该小目标分支,以补充上下文信息支持,提高目标定位的稳定性。

在损失函数设计方面,YOLOv11原生采用了DFL (distribution focal loss) 与CIoU Loss作为主干损失组合,具有良好的收敛特性与回归性能。为进一步适配木工工具类别不平衡与遮挡目标的检测需求,本文在分类损失中引入Focal Loss机制,通过对易错类别(如木工夹具与螺丝刀)增强权重,提升模型在难分类样本上的判别能力,Focal Loss加权机制为:

$$FL(p_i) = -\alpha_i(1-p_i)^\gamma \log(p_i) \quad (2)$$

式中: p_i 为预测概率; α_i 为类别权重系数; γ 为调节难易样本的指数项。

本文设置 $\gamma = 2.0$, 增强模型对小样本类如“木工夹具”的学习关注。

同时,定位损失从CIoU切换为GIoU (generalized IoU) 损失,以减缓目标边界框在遮挡或非完整情况下所带

来的训练不稳定问题^[10]。配合改进后的损失函数,本文还采用了Warm-up结合余弦退火的学习率调度策略,使得训练过程更加平滑,减少早期梯度震荡,提高最终模型泛化能力。

改进后的YOLOv11模型在结构上实现了对浅层信息的有效补强,在训练策略上加强了对小目标、难分类目标的关注能力,整体体系更加适配少儿木工教学中的实际图像环境。图2展示了该模型的总体架构示意图,图中标注了SimAM插入位置、小目标检测分支路径、特征融合与多尺度输出结构。通过这些结构的协同作用,模型能够更准确地识别尺寸不一、遮挡复杂、分布密集的木工工具,为后续教学辅助系统的构建提供了可靠的视觉感知基础。

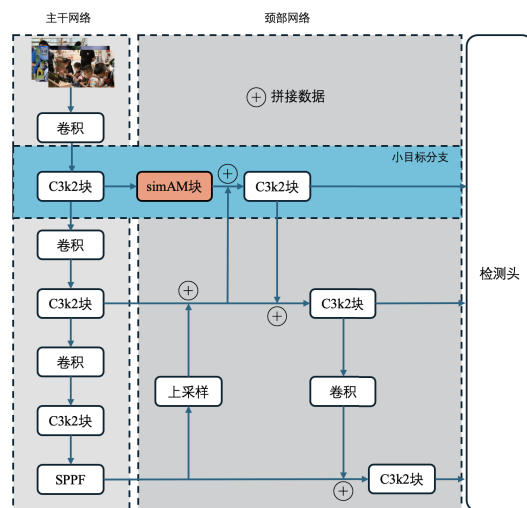


图2 改进后YOLOv11模型结构示意图

3 实验

为全面验证本文所提出的改进YOLOv11模型(简称YOLOv11-Ed)在少儿木工工具目标检测任务中的有效性与先进性,本文设计并实施了一系列系统化实验。实验内容包括:与现有主流模型的横向对比、不同改进模块的消融实验、各类别检测精度分析。通过上述实验,旨在从准确性、鲁棒性与部署可行性等多角度评估改进方案的综合性能。

本节所述全部实验均在同一软硬件环境中执行,以确保对比结果的公平性与可复现性。实验环境配置如下:操作系统为Ubuntu 20.04, GPU为NVIDIA RTX 4090, 配套主机CPU为Intel Core i9-13900K, 系统内存为128 GB, 深度学习框架采用PyTorch 2.7.1, CUDA版本为12.8。所有模型均加载COCO预训练权重后,在本研究自建数据集上进行微调训练,训练轮数设置为100, batch size设为32, 优化器选用AdamW, 初始学习率为0.001, 并采用Cosine Annealing策略进行动态调整。

为准确衡量检测性能,本文选取如下常用目标检测评估指标:

(1) mAP@0.5: IoU阈值为0.5时的平均精度,反映整

- 体检测准确率；
- (2) **mAP@0.5:0.95**：在 IoU 从 0.5 至 0.95 的多个阈值下计算平均精度，更能反映模型对边界框定位的整体能力；
- (3) **Precision**（精确率）：衡量模型对正样本预测的准确性；
- (4) **Recall**（召回率）：反映模型识别所有目标的能力；
- (5) **FPS**：表示模型在推理阶段的处理速度；
- (6) **Params** 与 **FLOPs**：分别衡量模型的参数量与浮点计算复杂度，反映其部署代价。

为验证 YOLOv11-Ed 的检测性能，本文将其与当前主流检测模型 YOLOv5m、YOLOv8m 及 YOLOv11 原始模型（YOLOv11-Base）进行对比实验。表 1 为各模型在本研究测试集上的检测性能对比结果。

表 1 不同模型在测试集上的总体性能比较

模型	mAP@0.5/%	mAP@0.5:0.95/%	Precision/%	Recall/%	FPS / (帧 · s ⁻¹)	Params/10 ⁶	FLOPs/10 ⁹
YOLOv5m	90.2	65	89.5	87.1	135	21.2	49
YOLOv8m	92.1	67.8	91.4	88.9	150	25.8	78.7
YOLOv11-Base	93	70.2	92.3	89.7	155	23.1	68
YOLOv11-Ed	95.4	74.1	94.2	92.3	148	24.7	74.2

由表 1 可见，YOLOv11-Ed 在所有关键指标上均优于其他对比模型，尤其是在 **mAP@0.5:0.95** 方面达到 74.1%，相较于 YOLOv8m 提升了 6.3 个百分点，相较于 YOLOv5m 提升了 9.1 个百分点，表明本文模型在多尺度目标检测上的精度更高，泛化能力更强。尽管 YOLOv11-Ed 在推理速度上略低于 YOLOv11-Base，但其仍保持了 148 帧/s 的实时检测能力，足以满足教学辅助场景对响应速度的要求。

为进一步分析模型对不同类型木工工具的检测能力，本文分别统计了 YOLOv11-Ed 在 5 个类别目标上的 **mAP@0.5** 值，如表 2 所示。

表 2 本文模型在各工具类别上的检测精度

工具类别	mAP@0.5/%
锤子	97.10
锯子	94.30
螺丝刀	92.80
锉刀	95.00
木工夹具	90.20

从结果可以看出，YOLOv11-Ed 在所有类别上均保持较高的检测精度，其中对锤子与锉刀的识别性能最为优越，均超过 95%。对于结构更为复杂或尺寸更小的螺丝刀与木工夹

具，尽管存在部分遮挡与姿态变化，但模型依然保持了 90% 以上的精度，验证了 SimAM 注意力机制与小目标检测分支的有效性。相比 YOLOv11-Base，该模型在细粒度辨别能力上表现更强，尤其在边界模糊和背景干扰条件下，检测性能更为稳定。

为验证模型结构中各改进组件的实际贡献，本文设计了 3 组消融实验，分别剔除 SimAM 注意力机制、小目标检测分支以及 Focal Loss 分类损失模块，并记录模型性能的变化情况，结果如表 3 所示。通过对比发现，3 个改进模块均对模型性能具有显著影响。其中 SimAM 注意力机制对 **mAP@0.5:0.95** 的提升最为显著，说明其在特征选择与遮挡区域判别中的积极作用；小目标检测分支主要改善了模型对小

尺寸目标的定位精度；Focal Loss 则有效缓解了训练阶段的类别不平衡问题，提升了模型对难分类样本的区分能

力。综合而言，3 个模块间具有良好的互补性，共同构成了本文提出模型性能优势的基础。

表 3 改进模块消融实验结果

模型配置	SimAM	小目标分支	Focal Loss	mAP@0.5	mAP@0.5:0.95
YOLOv11-Ed（完整）	√	√	√	95.40%	74.10%
移除 SimAM	×	√	√	92.70%	70.90%
移除小目标分支	√	×	√	93.20%	71.40%
移除 Focal Loss	√	√	×	93.90%	72.20%
YOLOv11-Base（全无）	×	×	×	93.00%	70.20%

4 结语

本文针对少儿木工教育场景中存在的工具种类多样、目标尺寸小、遮挡干扰强等问题，提出了一种基于 YOLOv11 的小物体目标检测方法。通过引入 SimAM 注意力机制、小目标检测分支及 Focal Loss 优化策略，有效提升了模型对微小目标的感知能力与类别判别能力。

在构建的数据集中，改进模型在 **mAP@0.5:0.95** 上达到 74.1%，在所有类别上均实现 90% 以上的 **mAP@0.5**，显著优于 YOLOv5m、YOLOv8m 与 YOLOv11 原始模型。消融实验进一步验证了各模块的性能贡献。

尽管如此,模型仍受限于数据规模和静态图像输入,未来可在视频检测、行为识别与AR交互系统方向展开进一步研究。总体而言,本文方法在精度与实用性之间取得了较好平衡,为教育场景下的视觉感知系统提供了有益探索。

参考文献:

- [1] 李安琪. 生发·建构·探索·拓维: 基于STEAM理念的木工坊活动实践的创新探索[J]. 中国信息技术教育, 2025(5): 105-107.
- [2] 李承烨, 张震, 梁哲恒, 等. 目标检测模型综述[J/OL]. 计算机研究与发展, 1-35[2025-06-13]. <http://kns.cnki.net/kcms/detail/11.1777.TP.20250506.1645.012.html>.
- [3] 张晓雷, 蒋源昊, 徐慧英, 等. BVD-YOLO: 基于YOLOv5改进的无人机小目标检测算法[J/OL]. 浙江师范大学学报(自然科学版), 1-11[2025-06-13]. <https://doi.org/10.16218/j.issn.1001-5051.2025.055>.
- [4] 张志豪, 厉小润, 陈淑涵. 基于改进YOLO11的无人机航拍图像小目标检测算法[J]. 液晶与显示, 2025, 40(6): 915-930.
- [5] 韩科立, 王振坤, 余永峰, 等. 基于改进YOLO11n模型的棉花田间复杂环境障碍物检测方法[J]. 农业机械学报, 2025, 56(5): 111-120.
- [6] YANG L X, ZHANG R Y, LI L D, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]//International conference on machine learning. New York: JMLR.org, 2021: 11863-11874.
- [7] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense

object detection[C]//Proceedings of the IEEE international conference on computer vision. Piscataway: IEEE, 2017: 2980-2988.

- [8] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of big data, 2019, 6(1): 1-48.
- [9] TERVEN J, CORDOVA-ESPARZA D M, ROMERO-GONZÁLEZ J A. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS[J]. Machine learning and knowledge extraction, 2023, 5(4): 1680-1716.
- [10] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 658-666.

【作者简介】

廖员(1982—), 女, 四川成都人, 硕士研究生, 讲师, 研究方向: 人工智能、虚拟现实, email: liaoyuan@guc.edu.cn.

黎润霖(1993—), 通信作者(email: arnold0824@hotmail.com), 男, 四川成都人, 硕士研究生, 助教, 研究方向: 自动驾驶、计算机视觉。

杨昌(1979—), 男, 湖南宁乡人, 硕士研究生, 副教授, 研究方向: 工业设计、少儿木工, email: 112932889@qq.com.

(收稿日期: 2025-06-15 修回日期: 2025-11-04)

(上接第57页)

- [3] 冯玉光, 徐望, 顾钧元, 等. 导弹装备健康管理及其关键技术研究[J]. 兵器装备工程学报, 2017, 38(1): 7-11.
- [4] 徐望, 冯玉光, 奚文骏. 导弹发射系统的CBM体系结构研究[J]. 飞航导弹, 2016(7): 54-58.
- [5] 杜涛, 阮爱武, 汪鹏, 等. 一种基于BP神经网络的集成电路PHM模型[J]. 计算机工程与科学, 2017, 39(1): 55-60.
- [6] 王林秋, 郑磊, 张永武. 基于PHM的复杂机电系统健康管理技术研究[J]. 计算机测量与控制, 2018, 26(1): 27-30.
- [7] 周奇才, 沈鹤鸿, 赵炯, 等. 基于深度学习的机械设备健康管理综述与展望[J]. 现代机械, 2018(4): 19-27.
- [8] 张雪英, 栾忠权, 刘秀丽. 基于深度学习的滚动轴承故障诊断研究综述[J]. 设备管理与维修, 2017(18): 130-133.
- [9] 杨剑锋, 乔佩蕊, 李永梅, 等. 机器学习分类问题及算法研究综述[J]. 统计与决策, 2019, 35(6): 36-40.
- [10] PRABU S, VELAN B, CHRISTALIN N S, et al. Mobile

technologies for contact tracing and prevention of COVID-19 positive cases: a cross-sectional study[J]. International journal of pervasive computing and communications, 2020, 18(2): 226-235.

- [11] 孟宏鹏. 基于机器学习的飞行器舵面系统故障诊断研究[D]. 南京: 南京航空航天大学, 2017.
- [12] ZHONG Y J, ZHANG Y M, ZHANG W, et al. Actuator and sensor fault detection and diagnosis for unmanned quadrotor helicopters[J]. IFAC-papersonline, 2018, 51(24): 998-1003.

【作者简介】

王杰(1998—), 男, 甘肃天水人, 硕士, 工程师, 研究方向: 飞行器控制与故障诊断。

(收稿日期: 2025-05-16 修回日期: 2025-11-11)