

基于改进 YOLOv10 的学生课堂行为检测模型

李英豪¹ 徐建¹ 魏迎澳¹
LI Yinghao XU Jian WEI Ying'ao

摘要

针对在复杂的课堂环境中,传统的目标检测算法出现的漏检和误检问题,文章结合学生课堂行为监控的实际需求,基于 YOLOv10s 目标检测模型进行优化,提出了 3 项改进措施。首先将普通 C2f 替换为 C2f_TripletAt 模块,在原有基础上添加了自注意机制,让网络能够更全面地理解图像内容,使模型更有效地区分目标,提高检测精度,特别是在复杂场景下的目标检测能力。然后在网络中加上 SEAM 注意力机制,用于补偿被遮挡面部的响应损失,通过增强未遮挡面部的响应来实现,同时可以增强特征提取能力,提高检测精度,保持计算效率。此外,通过采用 ADown 模块,可以大幅降低参数值,进一步提升模型对复杂场景中多目标行为的辨识能力。结果表明,与原始 YOLOv10s 模型相比,改进 YOLOv10s 模型的精确率、召回率、平均精确率均值分别上升了 2.8%、0.9%、1.6%。有效提升了课堂监控系统的智能化水平,为智能教育提供了更可靠的数据支持。

关键词

目标检测; YOLOv10; 课堂行为; 注意力机制

doi: 10.3969/j.issn.1672-9528.2025.12.001

0 引言

学生课堂行为的有效监控直接关系到教学质量和课堂秩序的稳定,随着班级规模扩大和学生行为多样化,传统的人工观察方式难以实时准确发现异常情况,如低头、使用手机等,容易影响教学效果。将目标检测技术与课堂监控相结合,可以提升行为识别效率,降低管理负担,并为教育机构提供数据支持,从而为教师提供实时、客观地反馈结果^[1]。

近年来,深度学习模型在教学中的应用成为备受关注的研究领域,众多研究员将其灵活运用到教学场景中,推动了该领域的进步。其中,谭斌等^[2]利用 Faster R-CNN 通过基于 ZFNet 预训练网络模型来提取学生课堂行为特征,实现对学生学习、玩手机等行为的检测。刘经纬等^[3]通过 CNN-GRU 模型来分析学生在课堂中的行为。张杰等^[4]通过轻量化的 CBW-YOLOv5s 模型,简化模型复杂度,提高检测精度。这些研究方法相较于传统算法在检测效果上有所改进,但依然面临计算复杂、漏检严重以及检测精度降低等问题。

虽然当前的研究领域中已然涌现出各式各样的研究方法,然而纵观现有的诸多成果,不难发现它们普遍存在着一定的局限性,诸如泛化能力欠佳,难以应对大规模数据的处理需求等问题。在这样的背景下,为切实达成高精度以及实

时检测的目标,满足实际应用场景的迫切需要,本文提出了用于学生课堂行为检测的改进 YOLOv10s 模型。

1 改进 YOLOv10 模型

YOLOv10 是 YOLO 系列的新一代目标检测算法,在实时性和准确性方面取得了显著进展,具有高性能的检测能力,同时降低了计算成本^[5],减少了推理延迟。从效率和准确性的角度来看,YOLOv10 全面优化了多个组件,包括轻量级分类头、空间-通道解耦下采样和基于秩引导的块设计等,这些改进提升了模型的性能。改进 YOLOv10s 模型的总体结构如图 1 所示,图中虚线框为改进的模块。

由图 1 可知,改进 YOLOv10s 模型首先将全部的 C2f 替换为 C2f_TripletAt,引入了 Triplet Attention 三重注意力机制,通过三重注意力机制,对高度、宽度和通道维度的建模,增强了模型对关键特征的关注能力^[6]。其次,该注意力机制显著提升了小目标检测能力^[7-8]。C2f_TripletAt 能通过关注细节特征并增强其表达,帮助模型更精准地定位和识别这一目标,从而提升模型在密集场景和远距离目标检测任务中的识别精度。

然后,添加了 SEAM 注意力机制,其核心作用在于通过分离通道注意力与空间注意力,对特征进行增强处理,来提升模型的特征提取和表达能力^[9]。在 YOLOv10 加入 SEAM 模块能够通过动态调整特征通道权重,增强模型对重要特征的聚焦能力,从而提升目标检测精度。

1. 湖北民族大学智能科学与工程学院 湖北恩施 445000

[基金项目] 国家自然科学基金项目(12464004、11964009)

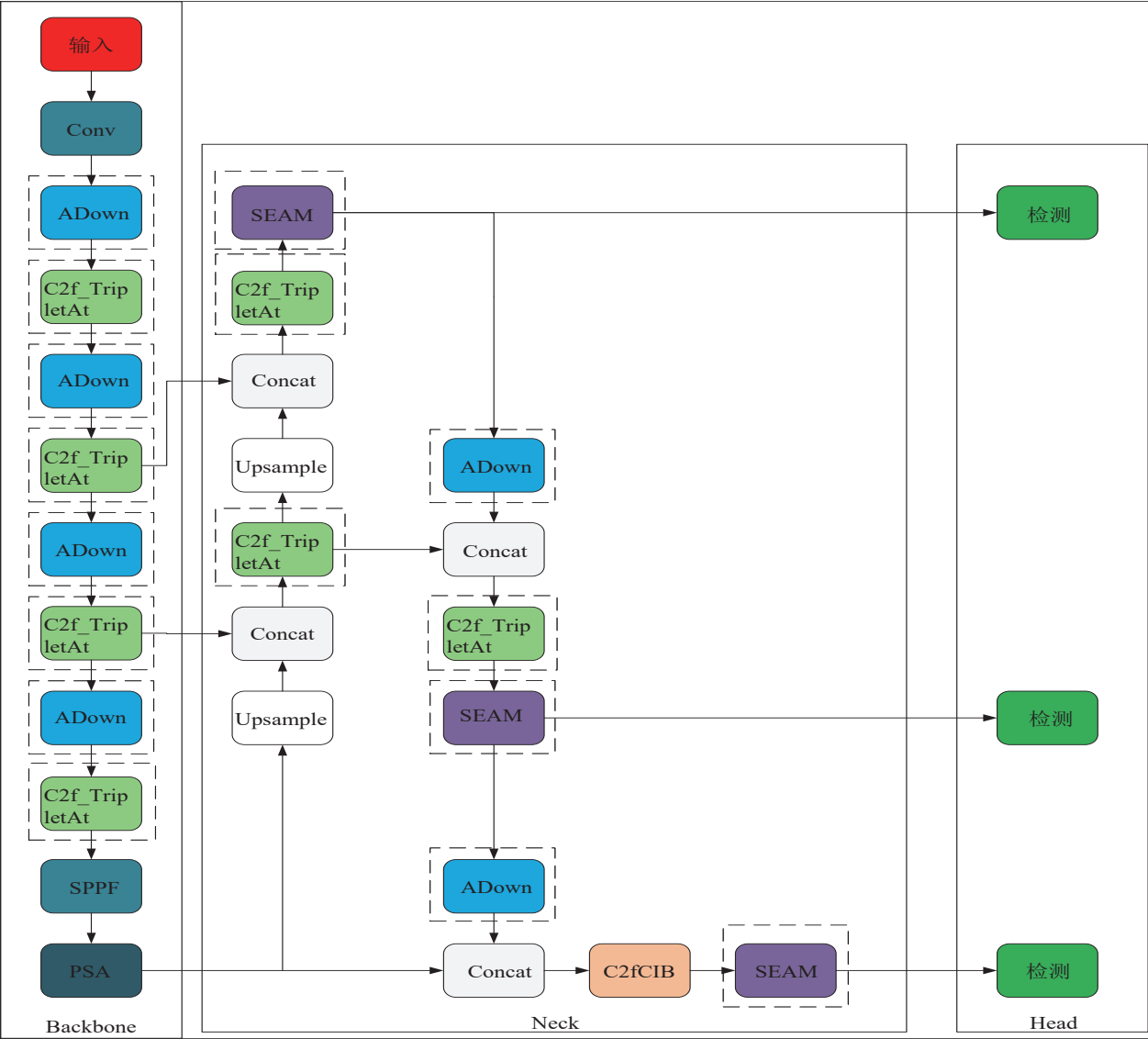


图 1 YOLOv10s 改进网络模型

最后将全部下采样模块全部替换为 ADown 模块，包括 SCDwn。这一改进显著优化了特征提取的效率和信息保留能力。相比传统的 SCDwn 更有效地减少了特征信息的损失，尤其在小目标检测上表现更优^[10]。

1.1 C2f_TripLetAt 模块

C2f_TripLetAt 模块结合了 C2f 模块和 Triplet Attention 机制，通过引入 3 个不同的注意力分支，分别建模特征图在通道、高度和宽度三个维度上的依赖关系，从而计算注意力权重，同时利用 Triplet Attention 捕捉图像在不同维度上的显著特征，生成注意力权重以增强特征表达能力，从而提高模型对图像内容的理解和特征提取能力^[11-13]。

1.2 SEAM 注意力模块

SEAM 模块旨在提升模型对重要特征区域的关注能力，在目标被遮挡或背景复杂的场景中表现优越^[14-16]。该模块使

用深度可分离卷积来提高计算效率，同时通过残差连接保留原始特征信息，以增强梯度传播的稳定性。

1.3 ADown 模块

ADown 模块是一种轻量化下采样模块，其目的是在保证模型检测精度的同时，减少计算量和参数量，实现特征保留与计算效率的平衡^[17]。该模块通过优化结构设计，有效减少了信息丢失，在目标检测任务中实现了参数量减少、计算效率提升和精度优化的多重目标。其核心优势在于兼顾特征保留和计算效率。

2 实验环境与数据

2.1 学生课堂行为数据集

实验中使用的数据集按 7:2:1 的比例划分为训练集、验证集和测试集，其中训练集、验证集、测试集分别包括 2 888、827、413 张图片。数据集主要包含多种行为类别，具

体检测类别 6 类，适合环境复杂、行为种类较多的检测模型的研究。数据集分类如表 1 所示。

表 1 数据分类

行为类别	编号	行为名称	标签名称
行为检测	1	举手	hand-raising
	2	阅读	reading
	3	写作	writing
	4	打电话	using phone
	5	低头	bowing the head
	6	靠在桌子上	leaning over the table

2.2 实验平台

本实验代码采用 Python 编程语言开发，使用 PyCharm 作为主要集成开发工具。深度学习框架选用了 PyTorch，并通过 Anaconda 进行项目环境和依赖库的管理，配置了基于 CUDA11.8 的计算框架。实验所用计算机运行 Windows11 操作系统，搭载英特尔 i7-12700H 处理器和英伟达 GeForce RTX 4050 显卡。

3 系统测试

3.1 评估指标

本实验采用 4 个评估指标，分别是：精确率（precision, P），即正确检测目标的比例；召回率（recall, R），即数据集中实际检测出目标的比例；平均精确率均值（mean average precision, mAP），指所有类别平均精确率的平均值，实验中采用交并比阈值为 0.5 的 mAP 表示；检测速度通常用帧率（frames per second, FPS）来衡量。分别用公式表示为：

$$S_P = [N_{TP} / (N_{TP} + N_{FP})] \times 100\% \quad (1)$$

$$S_R = [N_{TP} / (N_{TP} + N_{FN})] \times 100\% \quad (2)$$

$$S_{AP} = \int_0^1 S_{P(R)} dS_R \quad (3)$$

$$S_{mAP} = (\sum_{i=1}^n S_{AP_i}) / n \quad (4)$$

式中： S_P 表示精确率； S_R 表示召回率； N_{TP} 表示实际为正，且预测为正的样本数量； N_{FP} 表示实际为负，但预测为正的样本数量； N_{FN} 表示实际为正，但预测为负的样本数量； S_{AP} 表示第 i 类平均精确率； $S_{P(R)}$ 表示召回率为 S_R 时的精确率； S_{mAP} 表示平均精确率均值； n 表示类别数。

3.2 消融实验

为验证改进 YOLOv10s 算法各个改进点的有效性，在数据集上进行消融实验。对基于 YOLOv10s 改进模型进行 4 组消融实验，并通过评估指标公式计算评价指标，实验结果如表 2 所示。

表 2 消融实验结果

实验编号	SEAM	C2f_TripletAt	ADown	$S_P/\%$	$S_R/\%$	$S_{mAP}/\%$	检测速度/(帧·s ⁻¹)
1				89.7	85.7	89.0	172.4
2	√			89.1	87.4	89.7	156.3
3	√	√		88.1	87.0	89.7	131.6
4	√	√	√	92.2	86.5	90.4	128.2

注：“√”表示应用此模块。

由表 2 可知，实验 2 在实验 1 的基础上引入 SEAM 注意力模块，其召回率上升了 2%，平均精确率均值提高了 0.8%；实验 3 在实验 2 的基础上加上了 C2f_TripletAt 模块，相比实验 1，在召回率、平均精确率均值方面提升了 1.5%、0.8%；最终改进的实验 4 在实验 3 的基础上引入 ADown 模块，相比实验 1 在精确率、召回率和平均精确率均值上提升了 2.8%、0.9% 和 1.6%。

3.3 对比实验

为了更好地验证改进 YOLOv10s 模型的有效性和可行性，本研究与几种主流目标检测模型进行了对比实验，包括 Faster R-CNN、SSD、YOLOv3-tiny、YOLOv12n，实验结果如表 3 所示。

表 3 不同算法对比实验

模型	$S_P/\%$	$S_R/\%$	$S_{mAP}/\%$	检测速度/(帧·s ⁻¹)
SSD	80.1	21.3	51.2	88.4
Faster R-CNN	55.7	81.5	73.5	7.0
YOLOv3-tiny	84.2	72.8	83.1	238
YOLOv12n	86.3	83.5	87.1	185.2
改进 YOLOv10s	92.2	86.5	90.4	128.2

由表 3 可知，经过系统性实验验证与对比分析可得出如下结论，改进的 YOLOv10s 的综合性能最为出色。在其他检测模型中，SSD 模型的召回率太差，Faster R-CNN 的精确率和检测速度远低于其他模型，YOLOv12n 和 YOLOv3-tiny 相较于 YOLOv10s 各方面性能都较差。综合来看改进的 YOLOv10s 的性能优于其他几种算法。

不同算法检测效果对比如图 2 所示。通过对部分具有代表性的图像进行可视化分析，对比改进 YOLOv10s 模型和其他算法模型。结果如图 2 所示，图中的方框表示漏检，箭头表示误检，检测到的行为类别则用方框加上标签概率显示出来。

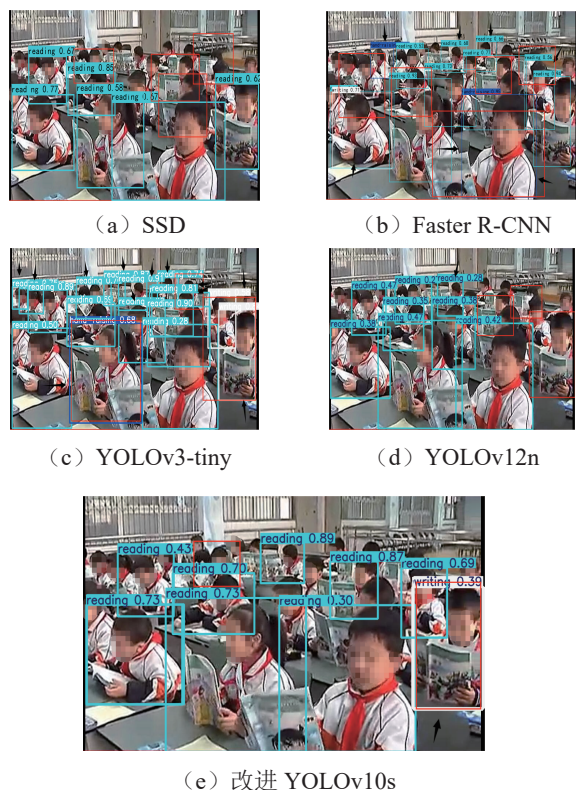


图2 不同算法检测效果对比

由图2可知,5个对比模型在复杂环境中均出现了不同程度的漏检和误检情况,而改进YOLOv10s模型表现更加优秀,漏检和误检情况较少,且精度较高。由此可见,改进YOLOv10s模型可以满足学生课堂行为检测要求,但检测效果仍有巨大提升空间。

4 结语

本研究以YOLOv10s目标检测算法为核心,通过改进模型架构,引入注意力机制,显著提升了算法在复杂课堂场景下的检测精度和实时性。YOLOv10s的高效性和准确性为课堂监控提供了坚实的技术保障,使其能够精准识别学生行为,为学生课堂行为分析提供更加高效、实用的解决方法。

参考文献:

- [1] 张小妮,杨萌萌,张军锋,等.基于改进YOLOv7算法的学生课堂行为识别研究[J].现代信息科技,2025,9(4):69-73.
- [2] 谭斌,杨书烔.基于Faster R-CNN的学生课堂行为检测算法研究[J].现代计算机(专业版),2018(33):45-47.
- [3] 刘经纬,王则渊.基于CNN-GRU模型的学生课堂行为与教学质量分析研究[J].中国管理信息化,2024,27(13):195-197.
- [4] 张杰,魏艳龙,薛红新.基于CBW-YOLOv5s轻量化模型的学生课堂行为快速检测[J].现代计算机,2024,30(21):1-7.

- [5] 凌旭,梁玉亮.基于YOLOv10-s的变电站智能门禁系统研究[J].电器工业,2025(4):48-52.
- [6] 李希,潘誉.基于改进YOLOv8的探地雷达管线目标检测方法[J].计算机与现代化,2025(1):94-99.
- [7] 高建安.基于YOLOv5的绝缘子检测与缺陷识别方法研究[D].兰州:兰州理工大学,2022.
- [8] 周雯晴,代素敏,王阳萍,等.基于注意力机制的多尺度手部分割方法[J].液晶与显示,2024,39(11):1506-1518.
- [9] 阳纯旭,阳奕,黄心燕,等.改进YOLOv8算法在高密度人群检测中的应用研究[J].信息与电脑(理论版),2024,36(17):6-9.
- [10] 李芳,危疆树,王玉超,等.基于改进YOLOv8n模型的辣椒病害检测方法[J].江苏农业学报,2025,41(2):323-334.
- [11] 刘京富,李玉丽.基于改进YOLOv8的疲劳驾驶检测技术[J].电脑与电信,2024(11):74-78.
- [12] TANG B Y, ZHOU J P, ZHAO C J, et al. Using UAV-based multispectral images and CGS-YOLO algorithm to distinguish maize seeding from weed[J]. Artificial intelligence in agriculture, 2025, 15(2): 162-181.
- [13] 柯冬冬.基于轻量注意力机制YOLO模型的警务审讯不规范行为识别[D].杭州:杭州电子科技大学,2024.
- [14] 谢建斌,唐俊峰,卿鄰波,等.改进YOLOv8的轻量级航拍目标检测算法[J].无线电工程,2024,54(12):2880-2887.
- [15] 蔡忠祺,林珊玲,林坚普,等.基于改进YOLOv8n-Pose的疲劳驾驶检测[J].液晶与显示,2025,40(4):617-629.
- [16] 田青,王颖,张正,等.改进YOLOv8n的选通图像目标检测算法[J].计算机工程与应用,2025,61(2):124-134.
- [17] WU Z Q, TOHTI G, GENI M, et al. Wind turbine rotor blade encoding marker recognition method based on improved YOLOv8 model[J]. Signal, image and video processing, 2024, 18: 6949-6960.

【作者简介】

李英豪(2001—),男,湖北荆州人,硕士研究生,研究方向:深度学习, email:1752445145@qq.com。

徐建(1981—),男,湖北英山人,博士,副教授,研究方向:嵌入式系统与智能控制的研究。

魏迎澳(1999—),男,河南安阳人,硕士研究生,研究方向:电气工程。

(收稿日期:2025-06-18 修回日期:2025-12-01)