DOI:10.19651/j.cnki.emt.2517803

# 改进 YOLO11 的学生课堂行为检测算法*

曹 倩[1] 曹 燚[2] 钱承山[1,2]
(1.南京信息工程大学自动化学院 南京 210044;2.无锡学院物联网工程学院 无锡 214105)

**摘 要**:针对 YOLO11 在课堂行为检测中存在复杂细节丢失、多尺度感知能力不足、计算效率低以及检测精度低的问题,提出了一种改进的 ATDW-YOLO 算法。首先,在颈部网络中构建了自适应极化特征融合模块,提升特征语义融合能力,更好地捕捉复杂细节。其次,设计了任务动态对齐检测头模块,提高模型在多尺度目标上的识别能力。然后,在主干网络中引入动态分组卷积混洗转换模块,增强特征表示能力,实现网络轻量化。最后,采用 Wise-IoU 函数替代 CIoU 损失函数,改善边界框的拟合能力,提高检测精度。实验结果表明,与 YOLO11n 模型相比,ATDW-YOLO 的 mAP0.5 和 mAP0.5:0.95 分别提高了 3.1%和 4.0%,而模型参数量、计算量和模型大小分别降低了 23.1%、9.5%和 23.6%,显著提升了检测精度,实现网络轻量化。

**关键词**:YOLO11;课堂行为检测;目标检测;智慧教育
**中图分类号**:TP309.2;TN40 **文献标识码**:A **国家标准学科分类代码**:520.2040

# Improved YOLO11 algorithm for student classroom behavior detection

Cao Qian[1] Cao Yi[2] Qian Chengshan[1,2]
(1. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China;
2. School of Internet of Things Engineering, Wuxi University, Wuxi 214105, China)

**Abstract**: In response to the issues of complex details loss, insufficient multi-scale perception, low computational efficiency, and low detection accuracy in YOLO11 for classroom behavior detection, an improved ATDW-YOLO algorithm is proposed. Firstly, an Adaptive Polarized Feature Fusion module is constructed in the neck network to improve feature semantic fusion capabilities and better capture complex details. Secondly, a task dynamic align detection head module is designed to enhance the model's recognition ability across multi-scale targets. Subsequently, a dynamic group convolution shuffle transformer module is introduced into the back-bone network to improve feature representation and achieve network lightweight. Finally, the Wise-IoU function replaces the CIoU loss function to improve the bounding box fitting capability and detection accuracy. Experimental results demonstrate that compared to the YOLO11n model, ATDW-YOLO improves mAP0.5 and mAP0.5:0.95 by 3.1% and 4.0%, respectively, while reducing model parameters, computational complexity, and model size by 21.6%, 7.4%, and 20.6%, respectively, significantly enhancing detection accuracy and achieving model lightweight.

**Keywords**: YOLO11;classroom behavior detection;target detection;intelligent education

## 0 引 言

在 2024 年 1 月 26 日的教育部新闻发布会上,教育部科学技术与信息化司明确指出:发展数字教育、推进教育现代化是大势所趋,也是改革必然方向。因此,应探索新方法,在有限资源下提升教学质量,以满足学生的多样化需求。

为实现这一目标,课堂管理技术也在不断进步。然而,传统课堂管理依赖教师观察,存在工作量大和主观偏见的问题。使用动作[1]和生理传感器[2]可以实时监测学生状态,但面临成本高和解读难的挑战。音频分析技术虽然能捕捉学生发言时的情感状态,但无法监测不发言的学生。相比之下,计算机视觉技术[3]能够提供一种客观、精确且成本较低的课堂行为分析方法,无需学生佩戴设备,且能持续

收稿日期:2025-01-03
*基金项目:无锡市"太湖之光"科技攻关(基础研究)项目(K20241046)、国家传感网工程技术研究中心开放课题基金(2024YJZXKFKT02)、江苏高校哲学社会科学研究一般项目(2023SJYB0919)、无锡学院引进人才科研启动专项经费(2022r043)项目资助

监测所有学生的上课状态。

作为计算机视觉领域的关键技术之一，目标检测技术在课堂行为监测中的应用逐渐成为教育技术研究[4]中的重要方向。目前，目标检测算法主要分为两大类：二阶段(two-stage)检测算法和一阶段(one-stage)检测算法。二阶段检测算法，如 R-CNN、Faster R-CNN，以其高精度获得广泛应用。例如，Zhou 等[5]采用 CNN-10 算法成功检测了举手、低头和听讲等多种课堂行为，准确率达到了 97.92%；Zhang 等[6]改进的 SlowFast(3D CNN)算法用于检测 7 种常见课堂行为，准确率为 75.67%。尽管这些算法在精度上表现突出，但由于采用了两阶段框架，模型结构较为复杂，计算效率低，难以满足实际应用中的资源限制需求。与之相比，一阶段检测算法，如 YOLO 和 SSD，具有较高的检测精度和更低的计算复杂度，适合实际应用。董琪琪等[7]改进的 SSD 算法用于检测 5 种学生行为，取得了 95.4%的平均精度，同时模型每秒能够处理 29 帧图像；谭暑秋等[8]采用改进的 YOLOv3 算法检测睡觉和玩手机两种学生异常行为，mAP 达到 94.9%，且检测速率为每秒 20 帧；Pan 等[9]基于 YOLOv5 框架提出的 G-ODNet 算法用于学生异常行为检测，将 mAP 提高至 94.6%，同时将模型大小压缩至 7 MB；Wang 等[10]提出了基于 YOLOv8 改进的 SBD-Net 算法，能够检测 7 种学生行为，mAP 达到 82.4%，且计算复杂度仅为 9.8 G。较于二阶段算法，一阶段算法在保证较高检测精度的同时，显著降低了计算资源的消耗，具有更好的应用潜力。

综上所述，YOLO 系列目标检测算法在课堂行为监测研究中得到了广泛应用。YOLO11 作为 YOLO 系列中的最新版本，在之前版本的基础上进行了改进。然而，针对课堂学生行为检测任务，YOLO11 仍然存在一些不足，如复杂细节的丢失、多尺度感知能力不足、计算效率较低以及检测精度较低等问题。因此，本研究基于 YOLO11 算法进行了优化，旨在提升其在课堂环境下的适用性，解决上述挑战。主要贡献包括：

1) 针对 YOLO11 中 C3k2 模块特征融合不充分的问题，构建了自适应极化特征融合(adaptive polarized feature fusion，APFF)模块。本模块注重信息极性，自适应调整输入，使网络更聚焦于不同区域的关键信息。从而更好地捕捉和融合上下文信息，减少复杂背景下的细节特征丢失。

2) 为解决 YOLO11 检测头对多尺度目标处理不佳、参数冗余、任务交互不足的问题，针对性地设计了任务动态对齐检测头(task dynamic align detection head，TDADH)模块。本模块旨在增强检测头的多尺度感知能力，提升检测头在定位和分类任务上的性能。

3) 针对 YOLO11 在处理高分辨率特征图时计算效率降低的问题，引入动态分组卷积混洗转换(dynamic group convolution ahuffle transformer，DGCST[11])模块。该模块结合 ShuffleNet v2[12]和 Vision Transformer[13]的优势，通过提升特征图的分辨率和信息丰富度，有效减少计算量和参数量，从而在保持检测精度的同时提高计算效率，实现网络轻量化。

4) 为应对低质量训练样本标注导致检测精度下降的问题，采用 Wise-IoU[14]函数替代传统损失函数 CIoU[15]。这种方法可以增强边界框损失的拟合能力，使检测头能够兼顾不同质量的锚框，从而提高检测精度。

## 1　基于 YOLO11 改进的课堂行为检测算法

在 YOLO11 框架的基础上，本文提出了一种改进的课堂行为检测模型 ATDW-YOLO。该模型在保证高检测精度的同时，也实现了轻量化。图 1 展示了 ATDW-YOLO 模型的架构，其中红色虚线框表示算法改进的部分。

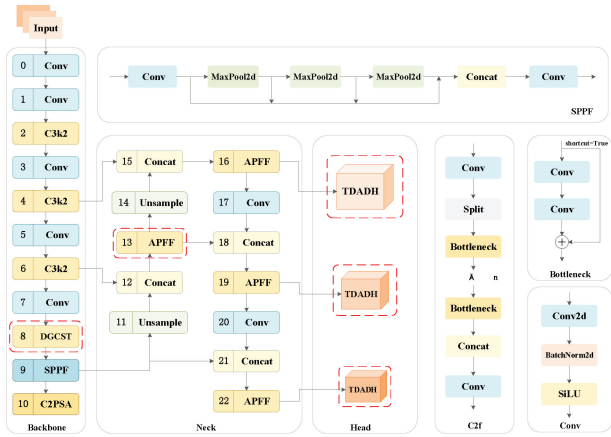



图 1　ATDW-YOLO 网络结构

Fig. 1　Network structure of ATDW-YOLO

ATDW-YOLO 模型的改进主要体现在以下 4 个方面。首先，构建了 APFF 模块，以增强不同特征之间语义信息的融合能力，减少细节信息丢失，提高识别准确性。其次，设计了 TDADH 模块，以提高检测头的多尺度感知能力，获得更高的检测精度和效率，从而提升检测头在定位和分类任务上的性能。然后，通过引入 DGCST 模块，提升特征表示的分辨率和信息丰富度，提高计算效率，实现网络轻量化。最后，采用 Wise-IoU 函数替代传统损失函数 CIoU，使检测头兼顾不同质量的锚框，提高模型的检测精度。

### 1.1　APFF 模块设计

本文提出了一种自适应极化特征融合——APFF 模块，旨在实现高效的特征融合和目标变化的自适应，其结构如图 2 所示。APFF 模块通过 1×1 标准卷积(standard convolution，SC)和动态卷积(low-parameter dynamic convolution，LDConv[16])串联设计，使网络在特征融合过程中能够聚焦不同区域的关键信息，提升特征表达的灵活性。同时，模块中融入了极化自注意力机制(polarized self-attention，PSA[17])，以减轻降维操作带来的信息丢失，从而强化模型对目标细粒度特征的捕捉能力。最终，APFF 模块通过将 LDConv、SC 和 PSA 有机融合，替代原颈部网

络 C3k2 中的 Bottleneck 模块,构建出更加高效的特征融合单元。

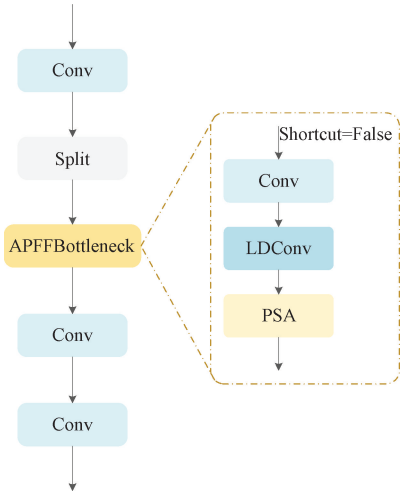



图 2　APFF 结构

Fig. 2　APFF structure

1) LDConv 替换标准卷积

在学生课堂行为识别任务中,光线变化、肢体差异以及遮挡会导致目标特征发生变化,仅使用 SC 存在明显局限。首先,SC 在捕捉细粒度目标细节方面能力不足,特别是面对姿态多样的目标时,其感受野调整能力有限,容易导致细节信息的丢失或模糊。其次,SC 缺乏对目标变化的适应能力,无法动态调整卷积核的采样位置,从而降低了目标定位的准确性,增加了误检的风险。

针对学生行为检测场景的特点以及标准卷积的局限性,本文引入 LDConv。LDConv 能够根据输入图像的特性动态调整感受野,在保持低计算量的同时,更精准地捕捉目标的细粒度细节,并自适应目标的位姿变化,从而提升模型对特征图中关键区域的感知能力。LDConv 结构如图 3 所示,与常规卷积中的单一卷积核不同,LDConv 通过动态聚合多个并行卷积核生成动态卷积核,以非线性方式聚合特征,从而获得更强的特征表示能力。此外,LDConv 利用 SE Block 计算卷积核的注意力权重。具体而言,SE Block 通过全局平均池化压缩全局空间信息输入,然后通过两个全连接层和中间的非线性层对特征降维,最后使用 Softmax 生成 $k$ 个卷积核的归一化注意力权重 $\pi_k$。将权重 $\pi_k$ 与对应的静态卷积核 $Conv_k$ 进行广播相乘并加权求和,得到最终的动态卷积 LDConv。这种设计有效增强了模型对目标细节的适应能力和不同特征之间的语义融合能力。

2) 融入 PSA 注意力

在学生课堂行为检测任务中,由于摄像头画质限制,学生的位姿估计对图像分辨率高度敏感。为实现更精准的目标位姿估计并减少背景干扰,本文提出将 PSA 嵌入模块中。PSA 由通道自注意力和空间自注意力两部分组成,如图 4 所示。具体而言,在通道维度上,PSA 保持原始特征

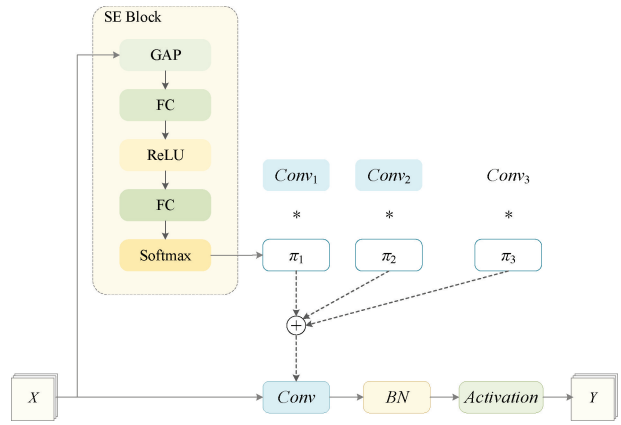



图 3　LDConv 结构

Fig. 3　LDConv structure

的一半维度;在空间维度上,PSA 则保持原始特征的完整维度。这一设计有效地在通道和空间维度上保留了较高的分辨率,减少了降维带来的信息损失,帮助模型更好地聚焦于目标的细粒度特征并提升识别准确性。

如图中 PSA 结构所示,输入特征 $X(\boldsymbol{C}\times\boldsymbol{H}\times\boldsymbol{W})$ 中,$\boldsymbol{C}$ 表示特征的通道数,$\boldsymbol{H}$ 表示特征的高度,$\boldsymbol{W}$ 表示特征的高度。空间极化注意力首先通过两个 1×1 卷积操作将输入特征 $\boldsymbol{X}$ 转换为特征 $\boldsymbol{Q}$ 和 $\boldsymbol{V}$。随后,通过全局池化将 $\boldsymbol{Q}$ 的空间维度压缩为 1×1,而 $\boldsymbol{V}$ 的空间维度保持为 $\boldsymbol{H}\times\boldsymbol{W}$。由于 $\boldsymbol{Q}$ 的空间维度被压缩,通过 Softmax 函数强化其信息,并与 $\boldsymbol{K}$ 进行矩阵相乘,再通过尺寸变换和 Sigmoid 函数将特征值归一化至 0～1 范围[18]。通道极化注意力与此类似,先通过两个 1×1 卷积将输入特征 $\boldsymbol{X}$ 转换为 $\boldsymbol{Q}$ 和 $\boldsymbol{V}$,其中 $\boldsymbol{Q}$ 的通道被完全压缩,$\boldsymbol{V}$ 的通道维度被压缩为原来的一半。基于极化滤波理论,通过 Softmax 函数增强 $\boldsymbol{Q}$ 的特征信息,再与 $\boldsymbol{V}$ 进行矩阵乘法。最后,通过 1×1 卷积和层归一化操作恢复通道维度,并使用 Sigmoid 函数将特征值归一化至 0～1。这种模块设计有效增强了模型对特征图中关键信息的捕捉能力,显著提高了行为识别的准确性。

### 1.2　TDADH 模块设计

YOLO11 原有的检测头在多尺度目标处理方面存在一定的局限性。其传统单尺度预测结构仅从特征图的一个尺度进行预测,忽略了其他尺度特征的贡献,导致目标检测的精度下降。同时,任务分支由两个 3×3 的一维卷积和一个 1×1 的二维卷积构成,参数数量大,计算冗余明显,传统卷积在学生目标特征聚合上表现不足。此外,此外,检测头的解耦设计将分类与定位任务完全分离,缺乏任务间的交互,可能导致空间错位,进一步降低了预测精度。

针对上述问题,本文提出了 TDADH 模块。其设计包括以下创新点:首先,模块引入分组归一化(group normalization, GN[19])代替 BN,通过通道维度进行归一化,消除显存限制,在训练和推断中表现更加稳定。GN 的引入不仅提升了多尺度感知能力,还增强了轻量化特征提

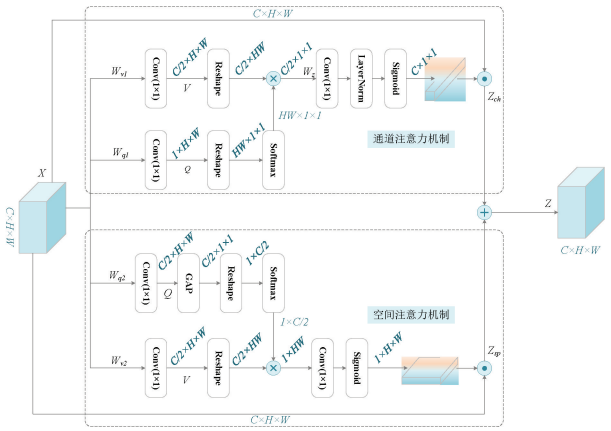



图 4　PSA 结构

Fig. 4　PSA structure

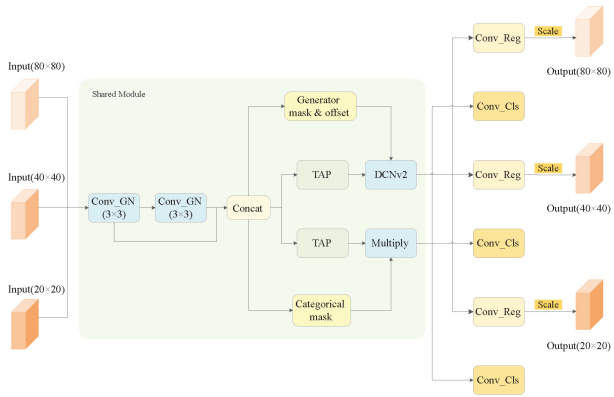



图 5　TDADH 结构

Fig. 5　TDADH structure

取的效果，从而改善了分类和定位性能。其次，改进共享参数检测头结构，将原有的双分支设计替换为共享参数化设计，构建更加轻量的检测头，显著减少了参数数量并降低了计算冗余。最后，借鉴 TOOD[20] 的设计理念，本文创新性地在检测头上定制了任务对齐结构，使检测头能够通过特征提取器从多个卷积层中学习到任务交互特征，增强了任务之间的交互能力。

在具体实现中，TDADH 模块包括任务交互特征提取、任务拆解模块（TAP）和多尺度检测增强 3 个部分，其结构如图 5 所示。首先，通过两个共享的 3×3 卷积对 3 个不同尺度的特征图进行特征提取，并在通道维度上将提取出的特征拼接，生成任务交互特征。然后，交互特征通过 TAP 模块拆解为分类特征和定位特征。其中，定位分支采用可变形卷积（deformable convnets version 2，DCNv2[21]）进行定位特征处理，并通过交互特征生成 DCNv2 所需的偏移量（offset）和掩膜（mask），以增强对目标形变的适应能力和定位精度；分类分支则借鉴 DyHead 的动态思想，通过交互特征生成分类掩码，并与分类特征相乘实现分类对齐，从而实现动态特征选择，强化分类与定位任务的交互能力。

最后，为应对检测头输出目标尺度不一致的问题，加入 Scale 层对特定输出进行尺度调整，进一步增强多尺度检测能力。通过将 TDADH 模块替换 YOLO11 原有检测头模块，不仅实现了轻量化设计，还显著提升了多尺度感知能力和任务交互能力，从而在分类与定位任务上获得了更高的检测精度和更优的整体性能。

### 1.3　DGCST 模块设计

YOLO11 主干网络在下采样和池化过程中，容易导致末端高层次特征图的分辨率显著降低，从而引发特征信息丢失，尤其是在检测较小目标或细节丰富目标时影响检测效率。为解决这一问题，本文引入了 DGCST 模块，替换主干网络 SPPF 层前的 C2f 模块，以提升特征图质量和模型性能。

DGCST 模块通过创新性结构设计，将 ShuffleNet v2 和 Vision Transformer 的优点相结合。首先，模块引入 ShuffleNet v2 中的通道分组卷积（group convolution，GConv），对输入特征图进行分组卷积操作，有效减少参数量和计算量；同时，通过通道混洗（channel shuffle）技术，增强跨通道信息交换能力，从而在保持特征图分辨率的同时提高特征表达能力。这一设计缓解了下采样和池化导致的特征信息损失。其次，结合 Vision Transformer 的设计思想，DGCST 模块能够捕捉全局信息，更好地理解特征图中不同位置的关联性，提升特征表达的丰富性和准确性。

DGCST 模块的结构如图 6 所示，其具体过程如下：输入特征图首先通过一个 1×1 卷积层调整通道数，并按 3∶1 的比例分割为两部分。一部分（75%）保留原始特征，另一部分（25%）通过 3×3 分组卷积和通道混洗操作，以增强特征混合性和表达能力。随后，将两部分特征图进行拼接并进入卷积前馈网络（ConvFFN）。ConvFFN 借鉴 Vision Transformer 的设计理念，由两个连续的 1×1 卷积层完成特征融合和压缩，并通过残差连接（skip connection）将输入特征与处理后的特征相加，增强特征表达能力，同时保持梯度传递的稳定性。

通过分组卷积、通道混洗技术和视觉 Transformer 思想的结合，DGCST 模块在显著提升特征表达质量的同时实现了网络的轻量化改进，有效提高了检测性能和计算效率。

### 1.4　Wise-IoU 损失函数

YOLO11 模型使用 DFL Loss 和 CIoU Loss 作为边框回归损失函数。DFL 损失函数帮助检测网络更快地聚焦到学生目标位置以及其邻近区域，而 CIoU 损失函数则用于衡量边界框预测的准确性。然而，CIoU 损失函数仅考虑了预测框与真实值之间的距离、纵横比等几何因素，却未能充分关注标注样本的质量。这可能导致收敛速度较慢和模型泛化能力不足。

为了弥补 CIoU 损失函数的不足，本文提出采用 Wise-IoU 作为边框回归损失的改进方案。Wise-IoU 在目标边框与锚框重合较好时，能够减弱几何因素的影响，从而增强模型的泛化能力，提升其稳定性和精度。同时，Wise-IoU

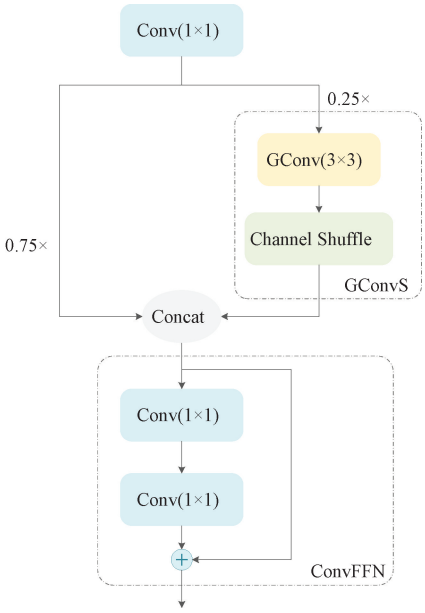



图 6　DGCST 结构

Fig. 6　DGCST structure

结合了注意力机制和梯度增益的思想,优化了目标检测模型的训练过程。其核心特点是引入了动态非单调聚焦机制,通过评估锚框的质量来调整训练重点。当锚框的质量较差时,其"异常值"较大,模型会为该锚框赋予较小的梯度增益,从而使边界框回归更加聚焦于质量更高的锚框。这一机制显著提高了训练过程的效率,进而提升了模型的整体性能。Wise-IoU 的具体公式如下:

$$L_{Wise-IoU} = r \cdot \exp\left[\frac{(x-x_{gt})^2+(y-y_{gt})^2}{(W_g^2+H_g^2)^*}\right] L_{IoU} \quad (1)$$

$$L_{IoU} = 1 - I_{IoU} \quad (2)$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (3)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \quad (4)$$

式中:$x$ 和 $y$ 分别为边界框的中心点在图像中的水平和垂直位置;$x_{gt}$ 和 $y_{gt}$ 分别为真实边界框的中心点坐标;$W_g$ 和 $H_g$ 分别为真实边界框的宽度和高度;$(W_g^2+H_g^2)^*$ 中的 $*$ 表示将 $W_g$, $H_g$ 从计算图中分离,从而有效地消除阻碍收敛的因素;$I_{IoU}$ 为真实框与预测框的重叠程度;$r$ 是非单调聚焦系数,其中 $\beta$ 为锚框离群度,$\alpha$、$\delta$ 为超参数,当 $\beta=\delta$ 时,$r=1$;$L_{IoU}^*$ 是单调聚焦系数。

Wise-IoU 参数的示意图如图 7 所示。

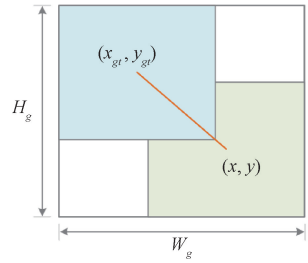



图 7　Wise-IOU 参数示意图

Fig. 7　Schematic diagram of Wise-IOU parameters

## 2　实验与结果分析

### 2.1　实验细节

1) 数据集

实验使用公开可用的学生课堂行为数据集(POCO)来评估本文提出的课堂行为检测方法的有效性。基于原始数据集,本文通过数据增强处理,将 POCO 数据集扩展至 5 709 张图像。增强后的数据集将学生上课状态划分为积极、消极和中立 3 类,涵盖多种上课行为,具体如表 1 所示。其中有 3 种中立状态:身体坐直,学生可能在听课;手机放在桌面上,手不接触,学生可能在看手机;书放在桌面上,手不接触,学生可能不在看书。

**表 1　学生行为说明**

**Table 1　Description of the student behavior dataset**

| 标签类别 | 行为名称 | 学生状态 |
|---|---|---|
| front face | 听课 | 积极 |
| bowed head | 低头 | 消极 |
| side head | 转头 | 消极 |
| upright | 坐直 | 中立 |
| body desk | 趴在桌上 | 消极 |
| phone | 手机 | 中立 |
| phone hands | 玩手机 | 消极 |
| book | 书 | 中立 |
| book hands | 阅读/书写 | 积极 |
| head arms | 睡觉 | 消极 |

本实验将数据以 7∶2∶1 的比例随机划分为训练集、验证集和测试集,投入模型训练、验证和测试,从而评估整个框架的性能。

2) 评价指标

为了客观地评估学生行为检测模型的性能,本文采用了多种评价指标,包括准确率(precision,P)、召回率(recall,R)、平均精度(average precision,AP)、均值平均精度(mean average precision,mAP)、每秒千兆浮点运算次数(GFLOPs)、模型大小、参数量(Params)。其中,GFLOPs 用于衡量模型的时间复杂度,参数量则反映模型的空间复杂度。P、R 和 mAP 用于评估模型的检测精度,值越大表示模型检测准确率越高;GFLOPs、模型大小和参数则衡量模型的轻量化程度,值越小表示模型越轻量,对硬件的性能要求也相应降低。具体计算公式如式(5)~(8)所示。

$$P = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{6}$$

$$AP = \int_0^1 P(R)\mathrm{d}R \times 100\% \tag{7}$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{8}$$

其中，$TP$ 是正确预测的正样本的数量，$TN$ 是正确预测的负样本的数量，$FP$ 是被划分为正样本的负样本的数量，$FN$ 是被划分为负样本的正样本的数量，$AP_i$ 是类别 $i$ 的平均精度，$N$ 是行为类别的数量。

3）实验设置

实验训练阶段使用的硬件平台和环境参数配置如表 2 所示。训练时，设置输入图片大小为 640×640 像素，训练迭代次数为 300，批次大小为 16，初始学习速率取 0.01。

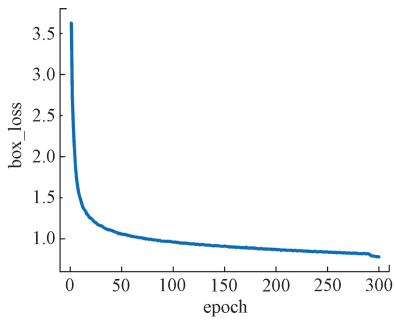
## 2.2　改进 ATDW-YOLO 模型评估

在训练用于学生行为检测的网络模型时，使用随机梯度下降（stochastic gradient descent，SGD）优化器对模型进行优化，并在最后 10 个训练阶段使用 mosaic 数据增强技术，这种调整旨在提高模型在检测学生行为时的鲁棒性能。训练结果如图 8 所示。

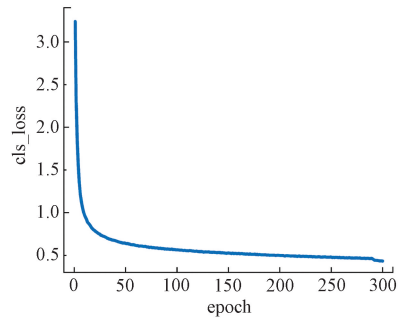
**表 2　训练环境和参数配置表**

**Table 2　Training environment and parameter configuration sheet**

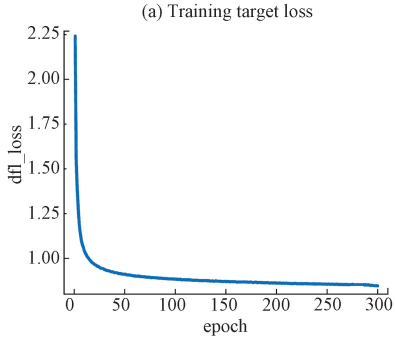
| 环境 | 参数配置 |
|---|---|
| 操作系统 | Windows11 |
| CPU | Intel(R) Core(TM) i7-14650HX |
| GPU | NVIDIA GeForce RTX4090 |
| 开发环境 | PyCharm |
| 编译环境 | Python 3.8 |
| 深度学习框架 | Pytorch 1.11.0 |



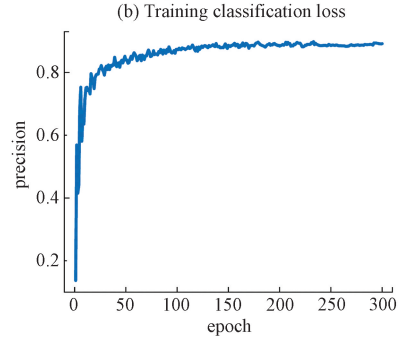
(a) 训练的目标损失
(a) Training target loss



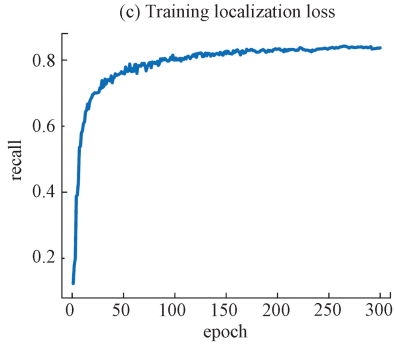
(b) 训练的分类损失
(b) Training classification loss



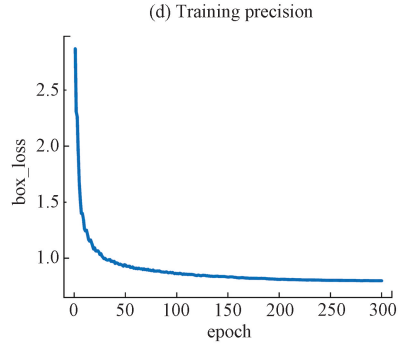
(c) 训练的定位损失
(c) Training localization loss



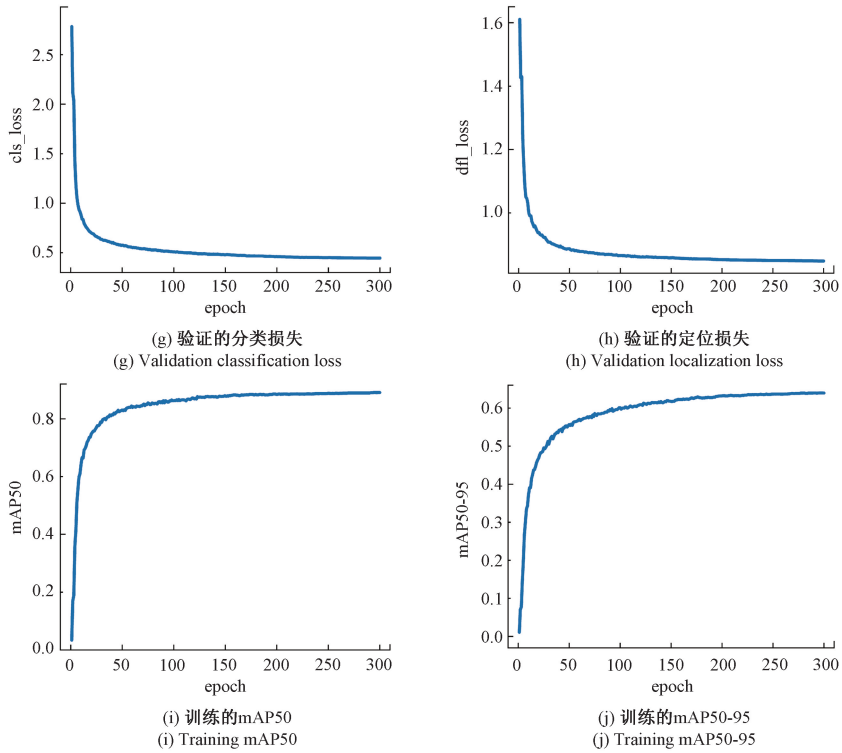
(d) 训练的精确度
(d) Training precision



(e) 训练的召回率
(e) Training recall



(f) 验证的目标损失
(f) Validation target loss




(g) 验证的分类损失
(g) Validation classification loss

(h) 验证的定位损失
(h) Validation localization loss

(i) 训练的mAP50
(i) Training mAP50

(j) 训练的mAP50-95
(j) Training mAP50-95

图 8　改进模型的训练结果

Fig. 8　Improved model training results

这些曲线图展示了改进模型 ATDW-YOLO 在训练和验证过程中的损失和评估指标变化。每张图表示不同的损失或评估指标随着训练轮数的变化情况。以下是对每个曲线图的分析:图 8(a)和(f)展示了改进模型训练和验证的目标损失,图 8(b)和(g)展示了改进模型训练和验证的分类损失,图 8(c)和(h)展示了改进模型训练和验证定位损失。这些曲线描绘了损失趋势,其中 $X$ 轴代表 Epoch,$Y$ 轴代表总体损失值。从曲线中可以看出,随着训练的进行,总体损失值逐渐下降并趋于稳定。这些结果表明,所提出的 ATDW-YOLO 表现出良好的拟合性能、高稳定性和准确性。图 8(d)、(e)、(i)和(j)展示了模型的重要评估指标,其中 $X$ 轴表示训练时间,$Y$ 轴表示精确度、召回率和不同 IoU 阈值下的 mAP 变化。从曲线中可以看出,随着训练轮数增加,曲线值逐渐上升并趋于稳定,模型的性能逐渐提升并达到收敛。从图 8 可以看出,所提出的 ATDW-YOLO 是有效的。

### 2.3　改进模块有效性分析

1) APFF 模块有效性分析

APFF 模块实现了在减少权重参数的同时,提高对目标变化的自适应能力,使网络更聚焦不同区域的关键信息,减少细节信息丢失。为了测试 APFF 模块加入模型不同位置后对模型性能提升最大化问题,本文设计如表 3 所示的有效性实验。

**表 3　APFF 模块有效性实验**

**Table 3　APFF module effectiveness experiment**

| 模型 | P/% | R/% | mAP50/% | mAP50-95/% | Params/M | GFLOPs/B | 模型体积/MB |
|---|---|---|---|---|---|---|---|
| YOLO11n | 87.0 | 79.8 | 86.4 | 60.6 | **2.6** | 6.3 | **5.5** |
| APFF(Backbone) | 87.4 | 81.2 | 86.5 | 60.9 | 3.3 | **6.2** | 6.6 |
| APFF(Neck) | **89.0** | **81.5** | **87.6** | **61.8** | 3.1 | 6.3 | 6.6 |

注:加粗表示最优值。

从表中数据可以看出,与基础 YOLO11 n 模型相比,无论是应用到 Backbone 还是 Neck 层,APFF 模块在召回率(R)和 mAP50-95 指标上均有所提升。尤其在 Neck 层应用 APFF 模块时,mAP50 和 mAP50-95 分别提高至 87.6%和 61.8%,表现出最佳性能。尽管模型的参数量和体积略有增加,但这一增加换来了显著的性能提升,证明了 APFF 模块在增强目标检测效果方面的有效性。

2) DGCST 模块有效性分析

YOLO11 主干网络在执行下采样和池化操作时，可能会降低高层次特征图的分辨率，会影响对模型的检测性能。为了解决这个问题，本文引入了 DGCST 模块，以替代主干网络中的 C2f 模块。表 4 展示了 DGCST 模块在 Backbone 的不同位置替换 C2f 模块后，对模型性能的具体影响。

**表 4　DGCST 模块有效性实验**

**Table 4　DGCST module effectiveness experiment**

| 模型 | P/% | R/% | mAP50/% | mAP50-95/% | Params/M | GFLOPs/B | 模型体积/MB |
|---|---|---|---|---|---|---|---|
| YOLO11n | 87.0 | 79.8 | 86.4 | 60.6 | 2.6 | 6.3 | 5.5 |
| DGCST（替换 Backbone 最后一个 C2f） | **87.6** | **80.0** | **86.5** | **60.9** | 2.3 | 6.1 | 4.7 |
| DGCST（替换 Backbone 最后两个 C2f） | 87.4 | 80.2 | 86.4 | **60.9** | **2.2** | **5.8** | **4.8** |

注：加粗表示最优值。

实验结果显示，DGCST 模块在替换 YOLO11 主干网络中的 C2f 模块后，显著提升了模型性能。特别是当替换主干网络中最后两个 C2f 模块时，参数量减少了 0.4 M，GFLOPs 降至 7.6 B，模型体积缩小至 5.6 MB。尽管替换单个 C2f 模块略微提高了检测精度，但在检测速度上不如替换最后两个 C2f 模块显著。为实现模型轻量化，本文选用 DGCST 替代了主干网络中的最后两个 Cf 模块。综上所述，DGCST 模块在提升检测准确率的同时，显著减少了模型的参数量和计算复杂度，实现了模型的轻量化。

3) Wise-IoU 模块有效性分析

为验证不同 IoU 损失函数对模型的影响情况，在 YOLO11 模型的基础上引入不同的 IoU 损失函数进行对比实验，实验结果如表 5 所示。

**表 5　Wise-IoU 模块有效性实验**

**Table 5　Wise-IoU module effectiveness experiment**

| 模型 | P/% | R/% | mAP50/% | mAP50-95/% | Params/M | GFLOPs/B | 模型体积/MB |
|---|---|---|---|---|---|---|---|
| YOLO11n | 87.0 | 79.8 | 86.4 | 60.6 | 2.6 | 6.3 | 5.5 |
| ＋EIoU | 86.8 | 80.1 | 86.6 | 61.4 | 2.6 | 6.3 | 5.5 |
| ＋Inner-IoU | 87.2 | 79.7 | 86.8 | 61.2 | 2.6 | 6.3 | 5.5 |
| ＋ShapeIoU | 87.8 | 79.8 | 87.0 | 61.0 | 2.6 | 6.3 | 5.5 |
| ＋Wise-IoU | **88.0** | **81.0** | **87.4** | **61.6** | 2.6 | 6.3 | 5.5 |

注：加粗表示最优值。

对比实验显示，4 种 IoU 损失函数（EIoU、Inner-IoU、ShapeIoU、Wise-IoU）在计算量、参数量和模型体积上未发生变化。与基本 YOLO11 n 模型相比，引入不同的 IoU 变体后，模型的准确率（P）、召回率（R）和 mAP 均有所提升。特别是使用 Wise-IoU 损失函数时，模型在各项指标上表现最佳：准确率达到 88.0%，召回率提高至 81.0%，mAP50 和 mAP50-95 分别达到 87.4%和 61.6%。综合分析，Wise-IoU 损失函数在提升检测性能方面效果最为显著。

### 2.4　消融实验

为验证本文提出的新模块和改进模型在学生课堂行为检测中的性能优势，对 ATDW-YOLO 进行消融研究。如表 6 所示，本文在 YOLO11 n 基础模型之上进行改进，在提升模型检测精度的同时实现了模型轻量化。

在 YOLO11 中，将原始特征融合阶段的 C2f 模块替换为 APFF 模块，增强了模型对信息极性的关注，提高了对目标变化的自适应能力，使得网络能够更好地聚焦于不同区域的关键信息，从而减少了细节信息的丢失。与原 YOLO11 n 模型相比，引入 APFF 模块后，精确度、召回率、mAP50 和 mAP50-95 分别提升了 2.0%、1.7%、1.2%和 1.2%，表明改进后的模型在检测精度上有了显著提升。

进一步地，将 TDADH 模块替换原 YOLO11 n 的检测头模块后，增强了检测头的多尺度感知能力，同时加强了定位和分类任务之间的交互性。相比于 YOLO11 n，改进模型的精确度、召回率、mAP50 和 mAP50-95 分别提升了 1.5%、3.3%、2.3%和 2.8%；此外，模型参数量减少了 0.4 M，模型大小压缩了 0.8 MB。尽管模型的计算量略有提升，但其体积大幅减小，检测精度得到明显改善，表明 TDADH 模块在提升定位与分类性能方面起到了关键作用。

表 6　消融实验结果对比

Table 6　Comparison of ablation experiment results

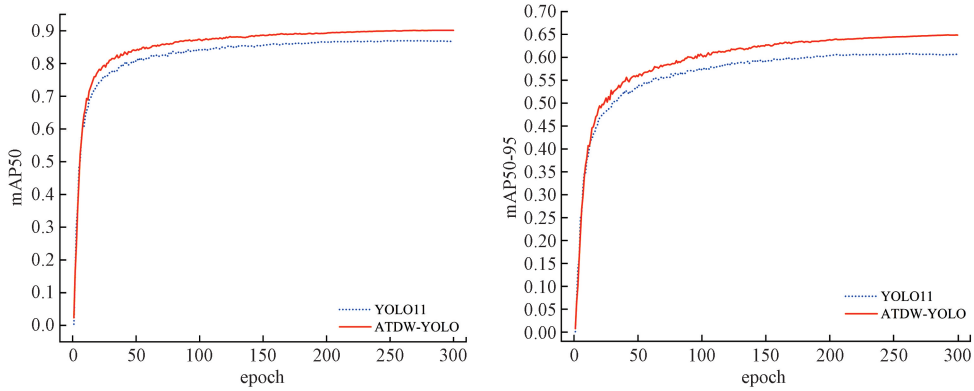
| APFF | TDADH | DGCST | Wise-IoU | P/% | R/% | mAP50/% | mAP50-95/% | Params/M | GFLOPs/B | 模型体积/MB |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 87.0 | 79.8 | 86.4 | 60.6 | 2.6 | 6.3 | 5.5 |
| √ | | | | 89.0 | 81.5 | 87.6 | 61.8 | 3.1 | 6.3 | 6.6 |
| | √ | | | 88.5 | 83.1 | 88.7 | 63.4 | 2.2 | 7.9 | 4.7 |
| | | √ | | 87.4 | 80.2 | 86.4 | 60.9 | 2.2 | 5.8 | 4.8 |
| | | | √ | 88.0 | 81.0 | 87.4 | 61.6 | 2.6 | 6.3 | 5.5 |
| √ | √ | | | 88.8 | 83.4 | 88.8 | 64.2 | 2.3 | 6.2 | 4.9 |
| √ | √ | √ | | 90.1 | 83.6 | 88.7 | 64.1 | **2.0** | **5.7** | **4.2** |
| √ | √ | √ | √ | **90.6** | **84.2** | **89.5** | **64.6** | **2.0** | **5.7** | **4.2** |

注:加粗表示最优值。

为进一步实现轻量化,在特征提取阶段引入了 DGCST 模块,改进后的模型在保持检测精度的前提下,参数量减少了 0.4 M,计算量降低了 0.5 B,模型大小压缩了 0.7 MB。

此外,模型还引入了 Wise-IoU 函数替代了 CIoU 函数,在不改变计算量、参数量和模型体积的情况下,精确度、召回率、mAP50 和 mAP50-95 分别提升了 1.0%、1.2%、1.0%和 1.0%。

图 9(a)和(b)更直观地展示了原始模型与最终改进模型之间的精度对比。通过消融实验结果可见,改进的 4 个模块不仅有效提高了检测精度,还在实现轻量化的同时保持了良好的耦合性。因此,ATDW-YOLO 相比原 YOLO11n 模型在整体性能上有了显著提升。




(a) mAP50对比
(a) Comparison of mAP50 curve

(b) mAP50-95对比
(b) Comparison of mAP50-95 curve

图 9　ATDW-YOLO 模型与原模型的 mAP 值对比图

Fig. 9　Comparison of mAP values of the ATDW-YOLO model with original model

## 2.5　对比实验

为了进一步验证本文提出的 ATDW-YOLO 模型在精度和轻量化方面的优势,本文选择了 9 个具有代表性的网络模型,与 ATDW-YOLO 进行对比实验。所有模型均在相同数据集上进行训练,并确保实验环境的一致性。为全面评估各模型性能,本文选择了精确度、召回率和 mAP 作为模型精度的评价指标,同时选取计算量、参数量和模型大小作为模型速度的评价指标。实验结果如表 7 所示。

与上述 9 种模型相比较,本文提出的模型 ATDW-YOLO 在各项指标上均表现出最优的综合性能。与 Faster R-CNN、SSD、YOLOv5n、YOLOv6n、YOLOv8n、YOLO11 n、SBD-Net、VWM-YOLO 和 SimAM-YOLO 相比,ATDW-YOLO 的精确度分别提高了 16.3%、20.4%、4.1%、1.9%、2.6%、3.6%、1.9%、1.0%和 1.8%;召回率分别提高了 13.6%、10.8%、13.0%、5.6%、3.8%、4.4%、1.3%、1.3% 和 2.8%;mAP50 分别提高了 14.2%、21.6%、9.7%、4.3%、2.9%、3.1%、1.0%、0.9% 和 2.0%;mAP50-95 分别提高了 17.5%、23.3%、13.9%、5.3%、3.9%、4.0%、1.5%、1.6% 和 3.3%。此外,ATDW-YOLO 在模型参数量、计算量和模型大小方面也表现出显著优势,参数量分别减少了 135.1、16.9、0.5、2.2、1.0、0.6、1.6、22.8 和 1.0 M;计算量分别降低了 362.6、29.6、1.4、6.1、2.4、0.6、4.0、66.1 和 2.4 B;模型大小分别压缩了 101.9、61.5、1.1、4.5、2.1、1.3、3.3、45.9 和 2.1 MB。综上所述,ATDW-YOLO 模型不仅在教室场景下的学生行为检测中展现出更优的精度和召回率,还在检测速度和模型轻量化方面具有一定优势,成功实现了精度与效率的平衡,具有较高的实用性。

表 7　不同模型性能比较

Table 7　Performance comparison of different models

| 模型 | P/% | R/% | mAP50/% | mAP50-95/% | Params/M | GFLOPs/B | 模型体积/MB |
|---|---|---|---|---|---|---|---|
| Faster R-CNN | 74.3 | 70.6 | 75.3 | 47.1 | 137.1 | 368.3 | 106.1 |
| SSD | 70.2 | 73.4 | 67.9 | 41.3 | 18.9 | 35.3 | 65.7 |
| YOLOv5n | 86.5 | 71.2 | 79.8 | 50.7 | 2.5 | 7.1 | 5.3 |
| YOLOv6n | 88.7 | 78.6 | 85.2 | 59.3 | 4.2 | 11.8 | 8.7 |
| YOLOv8n | 88.0 | 80.4 | 86.6 | 60.7 | 3.0 | 8.1 | 6.3 |
| YOLO11n | 87.0 | 79.8 | 86.4 | 60.6 | 2.6 | 6.3 | 5.5 |
| SBD-Net[10] | 88.7 | 82.9 | 88.5 | 63.1 | 3.6 | 9.7 | 7.5 |
| VWM-YOLO[22] | 89.6 | 82.9 | 88.6 | 63.0 | 24.8 | 71.8 | 50.1 |
| SimAM-YOLO[23] | 88.8 | 81.4 | 87.5 | 61.3 | 3.0 | 8.1 | 6.3 |
| ATDW-YOLO | **90.6** | **84.2** | **89.5** | **64.6** | **2.0** | **5.7** | **4.2** |

注：加粗表示最优值。

### 2.6　实际检测效果

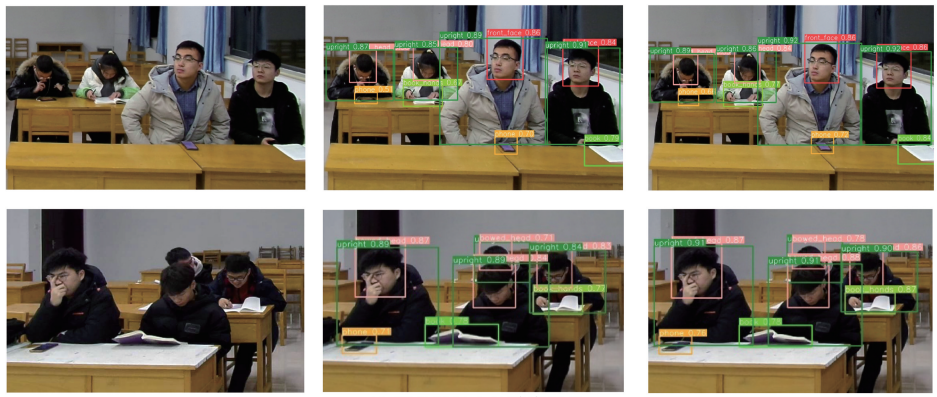
为了直观展示 ATDW-YOLO 模型相较于 YOLO11 n 模型在真实课堂场景中的优越性，本文对两种模型在实际课堂环境中捕捉的学生行为进行了对比分析。具体结果如图 10 和图 11 所示，展示了不同课堂场景下的行为识别能力。

在简单课堂场景中，YOLO11 n 模型面临重复检测、漏检和误检等问题，且整体检测精度较低。如图 10(a)～(d)所示，YOLO11 n 模型在学生行为识别上存在误判，尤其是在学生间的遮挡和小目标检测方面。而 ATDW-YOLO 在此类问题上表现出显著的优势，成功减少了误检、漏检和重复检测的发生，同时提高了预测准确度。这表明，ATDW-YOLO 能够更好地处理真实课堂环境中的简单场景，提供更加准确的行为识别。在复杂课堂场景下，YOLO11 n 模型难以有效捕捉复杂背景下的细节特征，导致目标漏检问题严重，尤其是在远距离目标的识别上，漏检问题显著。如图 11(a)～(f)所示，YOLO11 n 未能准确识别远离监控的学生，导致大量目标未被检测到。相比之下，ATDW-YOLO 模型在复杂课堂场景中取得了显著改进，能够有效识别远距离目标，并且在捕捉复杂背景中的细节特征方面表现突出。这表明，ATDW-YOLO 在复杂场景中的表现更加优越，能够有效应对密集的学生群体和复杂的背景干扰。


(a) 简单场景中的重复检测
(a) Repeat detection in simple scenarios

(b) 简单场景中的漏检
(b) Missed detection in simple scenarios

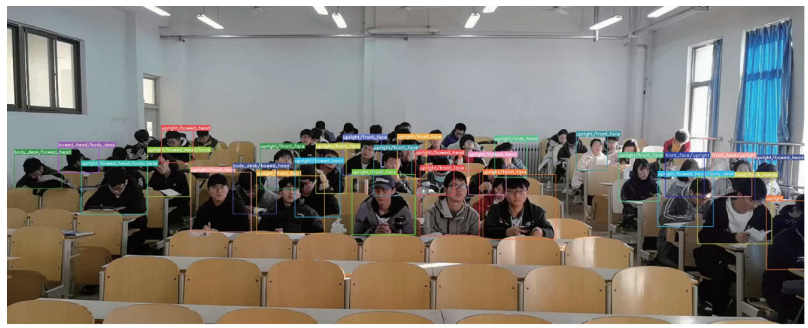
(c) 简单场景中的误检
(c) False detection in simple scenarios



(d) 简单场景中的精度提升
(d) Accuracy improvement in simple scenarios

图 10　简单场景的检测效果对比

Fig. 10　Comparison of detection effectiveness in simple scenarios



(a) 复杂场景1中的YOLO11检测效果
(a) YOLO11 detection effect in complex scenario 1

(b) 复杂场景1中的ATDW-YOLO检测效果
(b) ATDW-YOLO detection effect in complex scenario 1

(c) 复杂场景2中的YOLO11检测效果
(c) YOLO11 detection effect in complex scenario 2

(d) 复杂场景2中的ATDW-YOLO检测效果
(d) ATDW-YOLO detection effect in complex scenario 2

(e) 复杂场景3中的YOLO11检测效果
(e) YOLO11 detection effect in complex scenario 3

(f) 复杂场景3中的ATDW-YOLO检测效果
(f) ATDW-YOLO detection effect in complex scenario 3

图 11　复杂场景的检测效果对比

Fig. 11　Comparison of detection effectiveness in complex scenarios

综上所述，ATDW-YOLO 算法在应对课堂行为检测中的多样性和复杂背景时，展现出了更强的性能，尤其在复杂和动态的课堂环境中。与 YOLO11 n 相比，ATDW-YOLO 有效减少了误检和漏检，提高了模型在复杂场景下的适应性，为实际课堂行为识别任务提供了更加可靠的解决方案。

## 3　结　　论

本研究针对现有 YOLO11 模型在课堂行为检测任务中的局限性，提出了一种改进型算法 ATDW-YOLO。该算法通过一系列创新性优化，显著提升了检测精度，并有效降低了模型的计算复杂性。具体而言，提出的自适应极

化特征融合(APFF)模块加强了特征之间的语义信息融合,减少了关键信息的丢失。任务动态对齐检测头(TDADH)的设计显著提升了模型对多尺度目标的感知能力,并增强了定位与分类任务的交互性。动态分组卷积混洗转换(DGCST)模块优化了特征的分辨率和丰富度,同时实现了模型轻量化。Wise-IoU 损失函数的应用增强了模型对不同质量锚框的适应性,从而进一步提高了检测精度。

实验结果表明,ATDW-YOLO 的精确度达到了 90.6%,召回率为 84.2%,mAP50 为 89.5%,mAP50-95 为 64.6%;计算量为 2.0 M,参数量为 5.7 B,模型大小为 4.2 MB。与原 YOLO11 n 模型相比,ATDW-YOLO 在精确度、召回率、mAP50 和 mAP50-95 方面分别提高了 3.6%、4.3%、3.1%和 4.0%;同时,模型的参数量、计算量和模型大小分别降低了 0.6 M、0.6 B 和 1.3 MB。综合对比实验表明,ATDW-YOLO 不仅具有更高的检测精度,还实现了轻量化,展现出优越的实用性。

综上所述,本文提出的 ATDW-YOLO 在提高检测精度的同时,兼顾了模型的轻量化,为课堂行为检测领域提供了一种新技术途径。未来的研究可进一步优化模型的实时检测能力,并扩展其在其他教育场景中的应用,以提升其实用性和普适性。

## 参考文献


[1] 孙梓誉, 顾晶. 基于雷达时频变换和残差网络的人体行为检测[J]. 电子测量技术, 2024, 47(10): 27-33.
SUN Z Y, GU J. Human activity detection based on radar time-frequency transformation and residual network[J]. Electronic Measurement Technology, 2024, 47(10): 27-33.

[2] CAI Y, LI X, LI J. Emotion recognition using different sensors, emotion models, methods and datasets: A comprehensive review[J]. Sensors, 2023, 23(5): 2455.

[3] 刘雨萌, 桑海峰. 基于关键帧定位的人体异常行为识别[J]. 电子测量与仪器学报, 2024, 38(3): 104-111.
LIU Y M, SANG H F. Human abnormal behavior recognition based on keyframes localization[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(3): 104-111.

[4] 刘清堂, 何皓怡, 吴林静, 等. 基于人工智能的课堂教学行为分析方法及其应用[J]. 中国电化教育, 2019(9): 13-21.
LIU Q T, HE H Y, WU L J, et al. Classroom teaching behavior analysis method based on artificial intelligence and its application[J]. Intelligent Lead & Smarter Education, 2019(9): 13-21.

[5] ZHOU J, RAN F, LI G, et al. Classroom learning status assessment based on deep learning [J]. Mathematical Problems in Engineering, 2022(1): 7049458.

[6] ZHANG SH W, LIU H, SUN CH, et al. MSTA-SlowFast: A student behavior detector for classroom environments[J]. Sensors, 2023, 23(11): 5205.

[7] 董琪琪, 刘剑飞, 郝禄国, 等. 基于改进 SSD 算法的学生课堂行为状态识别[J]. 计算机工程与设计, 2021, 42(10): 2924-2930.
DONG Q Q, LIU J F, HAO L G, et al. Student action recognition based on improved SSD algorithm [J]. Computer Engineering and Design, 2021, 42(10): 2924-2930.

[8] 谭暑秋, 汤国放, 涂媛雅, 等. 教室监控下学生异常行为检测系统[J]. 计算机工程与应用, 2022, 58(7): 176-184.
TAN SH Q, TANG G F, TU Y Y, et al. Classroom monitoring students abnormal behavior detection system[J], Computer Engineering and Applications, 2022, 58(7): 176-184.

[9] PAN X, CHEN T, ZHAO X, et al. A study of lightweight classroom abnormal behavior recognition by incorporating ODConv[C]. 2023 5th International Conference on Frontiers Technology of Information and Computer(ICFTIC), 2023: 491-500.

[10] WANG ZH F, WANG M H, ZENG CH Y, et al. SBD-Net: Incorporating multi-level features for an efficient detection network of student behavior in smart classrooms[J]. Applied Sciences, 2024, 14(18): 8357.

[11] GONG W K. Lightweight object detection: A study based on YOLOv7 integrated with ShuffleNetv2 and Vision Transformer [J]. ArXiv preprint arXiv: 2403.01736, 2024.

[12] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design[J]. Computer Vision-ECCV 2018, 2018, 11218: 122-138.

[13] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image is Worth 16x16 Words: Transformers for image recognition at scale [J]. ArXiv preprint arXiv:2010.11929, 2020.

[14] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. ArXiv preprint arXiv:2301.10051, 2023.

[15] ZHENG ZH H, WANG P, LIU W, et al. Distance-IoU Loss: Faster and better learning for bounding box regression [J]. AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

[16] HAN K, WANG Y H, GUO J Y, et al.

ParameterNet: Parameters are all you need for large-scale visual pretraining of mobile networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:15751-15761.

[17] LIU H J, LIU F Q, FAN X Y, et al. Polarized self-attention: Towards high-quality pixel-wise regression[J]. ArXiv preprint arXiv:2107.00782, 2021.

[18] SHAFIQ M, GU ZH Q. Deep residual learning for image recognition: A survey[J]. Applied Sciences, 2022, 12(18): 8972.

[19] TIAN ZH, SHEN CH H, CHEN H, et al. FCOS: A simple and strong anchor-free object detector[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 44(4): 1922-1933.

[20] FENG C, ZHONG Y, GAO Y M, et al. TOOD: Task-aligned one-stage object detection[C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2021: 3490-3499.

[21] ZHU X ZH, HU H, LIN S, et al. Deformable ConvNets V2: More deformable, better results[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9308-9316.

[22] 曾钰琦, 刘博, 钟柏昌, 等. 智慧教育下基于改进YOLOv8的学生课堂行为检测算法[J]. 计算机工程, 2024, 50(9): 1-17.
ZENG Y Q, LIU B, ZHONG B CH, et al. An improved YOLOv8 algorithm for student behavior detection in classrooms within the context of smart education[J]. Computer Engineering, 2024, 50(9): 1-17.

[23] XU Q, WEI Y T, GAO J, et al. ICAPD framework and simAM-YOLOv8n for student cognitive engagement detection in classroom[J]. IEEE Access, 2023, 11: 136063-136076.

## 作者简介

**曹倩**,硕士研究生,主要研究方向为计算机视觉、目标检测。

E-mail:c1519958281@163.com

**曹燚**(通信作者),博士,讲师,主要研究方向为计算机视觉、车联网安全。

E-mail:caoyi@cwxu.edu.cn

**钱承山**,博士,教授,主要研究方向为智能终端与物联网应用、自动检测技术、非线性系统控制。

E-mail:qianchengshan@163.com