

基于 YOLOv8s 改进的居家跌倒行为检测算法

叶智国¹ 熊盛佳¹ 吴章云¹ 杜远飞¹ 吴传宇¹ 陈仕国^{1*}

YE Zhiguo XIONG Shengjia WU Zhangyun DU Yuanfei WU Chuanyu CHEN Shiguo

摘要

室内环境中,因摄像机拍摄角度多变、光线条件不稳定,导致老年人跌倒状况难以精准识别,为此,文章提出了一种基于 YOLOv8s 改进的 RRS-YOLO 算法。在 Backbone 层,采用基于 GELAN 的模块改进 C2f 模块,以获取更丰富特征信息,同时减少计算量和内存消耗。在 Neck 层,采用重参数加权双向特征金字塔网络,融合不同特征层,提升检测准确性和鲁棒性。检测头部分引入 DWConv 和 SaE 自编码器,关注小目标和遮挡目标,提高整体性能。实验显示,RRS-YOLO 算法参数量降低 19.4%,计算量降低 9.2%,mAP@0.5 提升 2.1%。该算法不仅降低了模型参数量,还显著提升了跌倒检测速度与精度,有助于完善养老服务体系,提升老年人生活安全性和幸福感。

关键词

跌倒检测; YOLOv8s; RGELAN; RSBiPAN; SD

doi: 10.3969/j.issn.1672-9528.2025.05.001

0 引言

根据国家统计局网站发布的《2023 年国民经济和社会发展统计公告》,截至 2023 年年底,60 岁及以上人口占全国人口的 21.1%^[1]。当前,我国正面临着人口老龄化趋势加剧及空巢老人群体迅速膨胀的双重挑战,而与之配套的养老服务体系尚处于构建和完善阶段,整体服务水平尚待提升。张田田等人^[2]研究发现,60 岁以上的老人摔倒发生率达到 23.4%,跌倒成为我国 ≥65 岁老年人因伤死亡的首要原因。因此,通过技术手段,特别是在算法层面实现高效、精准的跌倒检测,并确保能够及时触发应对措施,对于有效提升独居老人的日常生活安全性与幸福感具有重大意义。

近年来,跌倒检测领域的研究已经取得了显著的进展与成果,这些方法在提升检测效率、准确性和适应性方面均有所突破。陈明祥等人^[3]通过惯性传感器采集人体各部分的质心坐标,通过对质心位置、速度、加速度的实时检测,建立了以静态、动态判据为主,结合瞬时加速度方向来综合实现跌倒检测,但由于设备监测准确性差和老人抗拒穿戴且设备需要频繁充电等弊端而使用受限。李伟等人^[4]用 MIMO 体制的毫米波雷达芯片采集人体动作的回波信号,并基于外部注意力机制的 PointLSTM 网络实现点云在时空中的提取特征和分类识别,以此来实现室内跌倒检测,但使用场景受限,只能用于单人使用场景,无法同时监测多人跌倒,对人和动物无法有效区分。Wang 等人^[5]提出了一种基于视觉的双通

道特征集成跌倒检测方法,利用目标检测模型(YOLO)和人体姿态检测模型(OpenPose)进行预处理,获取人体关键点和位置信息,应用多层感知器和随机森林分别对动态和静态特征数据进行分类。李春华等人^[6]提出一种基于 YOLOv5 的跌倒检测系统,选择 K-means++ 算法对所用数据集进行聚类优化,以提高摔倒检测系统在现实环境中的泛化性。王小鹏等人^[7]提出一种改进型 YOLOv8 结合人体骨骼关键点的跌倒检测算法,将人体质心下降速度、人体躯干与地面间的夹角变化速度和人体的高宽比三者共同作为跌倒行为的判决依据。但是这些模型的计算复杂度高,响应时间长,难以完成实时检测。

为解决室内环境复杂导致检测率低,模型响应时间长的问題,本文提出了一种用于室内跌倒状态检测的改进 YOLOv8s 的跌倒检测算法,该方法结合互联网充分发挥信息的实时传递的作用,将跌倒的画面实时传输到终端,并确保及时触发应对措施,从而提升独居老人的日常生活安全性与幸福感。

1 跌倒数据集采集与图像预处理

本研究的数据集分为两类,一类是国外的跌倒开源数据集 UR Fall Detection Dataset^[8]、Fall-detection Dataset^[9]和 Multiple cameras fall Dataset^[10];另一类是自建数据集,因开源的数据集大部分都是单人场景,而实际场景中多人的场景居多,因此自建数据集主要是多人场景,图片来源于网络。

为确保所制作的跌倒数据集具有多样性,在划分训练集、测试集与验证集之后,对数据集中的原始图像进行裁剪,并

1. 福建农林大学 福建福州 350002

[基金项目] 国家重点研发计划项目(2022YFF1202400)

对裁剪后的图像进行亮度增强、亮度变暗、模糊、垂直方向水平方向翻转等图像处理方法进行处理。同时利用 Labelme 软件对图片进行标注,根据跌倒动作特征不同,标注3个类别:up(站立)、bending(弯腰、蹲下)、down(跌倒),标注为 YOLO 格式。

对数据集按照 8:1:1 进行划分,数据集共 4 000 张图片,其中训练集 3 200 张图片,测试集 400 张图片,验证集 400 张图片,在跌倒数据集中,每张图有 1 个及以上的标签,其中站立标签数为 2 690,弯腰、蹲下标签数为 1 611,跌倒标签数为 1 349,共计 5 650 个标签。

2 模型改进

2.1 C2f-RGELAN

室内环境中,由于环境状况复杂,各种家具可能会一定程度上遮挡人体,导致目标检测算法无法正确检测出人的跌倒状态,因此需要使用基于 GELAN^[11] 的 C2f-RGELAN 模块对输入图像的特征进行高效率提取和捕捉上下文信息。

本研究引入了表现优异的 GELAN 模块,并将其与 C2f 模块相结合,衍生了全新的 C2f-RGELAN 模块,如图 1 所示。该模块融合了 CSPNet Block^[12] 模块、RepConv^[13] 模块以及 ELAN (efficient long-range attention network)^[14] 模块,CSPNet Block 通过跨阶段局部网络(CSPNet)的设计,增强了特征的学习能力,RepConv 模块则通过重复利用特征图,提高了特征的利用率,ELAN 模块有两条分支。第一条路径通过一个 1×1 的卷积层来改变通道数量。第二条路径则首先通过一个 1×1 的卷积模块调整通道数,随后依次通过 4 个 3×3 的卷积模块进行特征提取,最终将 4 个特征图合并以获得最终的特征提取结果。ELAN 的设计还考虑到梯度路径的优化,能够减少梯度消失和梯度爆炸的问题,提高模型的检测性能和推理速度。

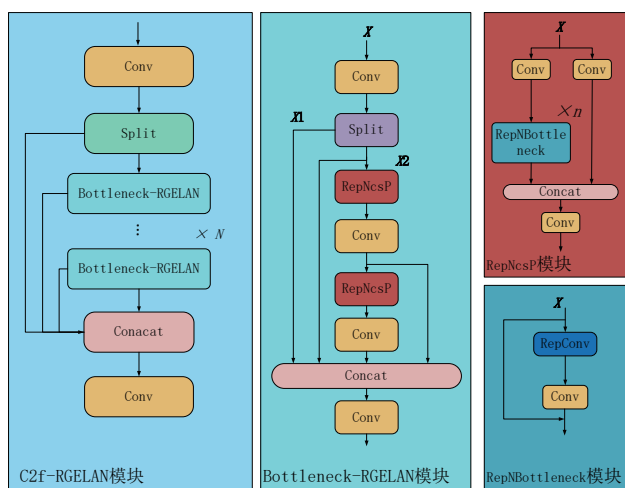


图 1 C2f-RGELAN 模块

2.2 重参数加权双向特征金字塔网络 RepBiPAN

实践证明,多尺度特征融合是目标检测中至关重要且有

效的组成部分。特征金字塔网络(FPN)通过自上而下的路径聚合高级语义特征和低级特征,从而提供更准确的定位。同时,为增强分层特征表示的能力,PANet^[15]在 FPN 上添加了一个额外的自底向上路径,以缩短低级和顶层特征的信息路径,有助于从低级特征传播准确的信号。BiFPN^[16]为不同的输入特征引入可学习的权重,并简化 PAN 以实现更好的性能和更高的效率。PRB-FPN^[17]通过具有双向融合和相关改进的并行 FP 结构来保留高质量的特征以进行精确定位。在本文算法 Neck 层中引入 RepBiPAN。如图 2 所示,RepBiPAN 采用了双向融合的方式,即在自上而下的传输路径中引入自底向上的信息流,使得浅层特征能够以更高效的方式参与多尺度特征融合,同时引入可重参数化思想,通过对网络结构的通道宽度和深度进行调整,实现了在保持模型复杂度不变的情况下,提升模型的表征能力。还设计一个双向联结用于在特征融合过程中保留更准确的定位信号,并解决信息流不畅和特征丢失的问题,还使得特征在不同层级之间得到了充分的交互和融合,有利于提取更丰富的特征表示。在特征融合的过程中,引入了一个可训练的权重参数,以优化融合效果。具体公式表达为:

$$O = \sum_{i=0} \frac{w_i}{w_j + \varepsilon} \cdot I_i \quad (1)$$

式中: O 表示加权特征融合后的输出结果; w 表示可训练权重,可以是一个单独的数值(标量)或一个包含多个数值的数组(向量),以适应不同维度的特征融合需求。为了确保计算过程中的数值稳定性,特意加入了一个常数项 ε ,其值设定为 0.000 1;此外 I_i 表示加权融合的原数据输入。

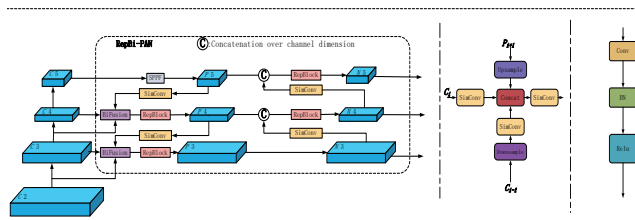


图 2 RepBiPAN 结构示意图

2.3 SD-Head 模块

头部网络位于模型的末端,主要负责根据主干网络和颈部网络提取的特征信息,头部网络在目标检测模型中发挥着至关重要的作用。主要功能涵盖:预测目标的类别和边界框(即目标所在的位置和大小),并参与损失函数的计算以指导模型的训练和优化。原始的检测头设计对于复杂场景下的目标检测效果有限,特别是在处理多尺度、遮挡或背景复杂的图像时。且缺乏一种自适应的空间特征融合方式,无法有效地过滤掉冲突信息并增强尺度不变性,会影响模型对不同大小目标的检测性能。因此,本文中在头部网络中融入了

DWConv^[18] 和 SAE^[19] 自编码器。选择在 Decoupled-Head 解耦头的分类分支中引入 DWConv，使模型可以在保持相对较高性能的同时，降低模型的大小和计算需求，特别是在移动设备等资源受限的环境中；在检测的分支中引入 SaE 自编码器，因其在特征提取和重构方面具有良好的性能，可能有助于提取更为丰富和有效的特征、优化检测头对输入数据的处理。YOLOv8s 改进后的网络结构如图 3 所示。

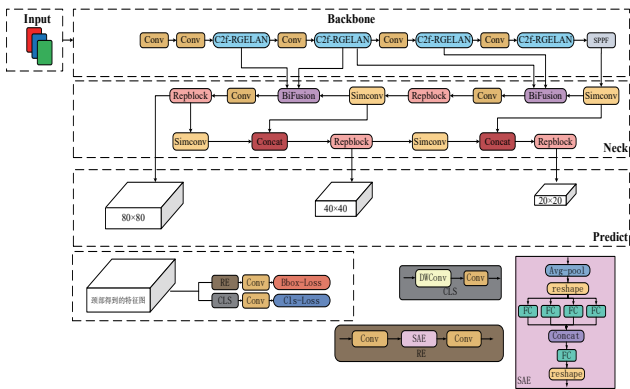


图 3 RRS-YOLO 模型结构图

3 实验与结果分析

3.1 实验环境

本文使用的环境：实验平台的操作系统为 Windows10（专业版），处理器为 Inter(R)Core(TM)i5-12490F @3.0 GHz，显卡为 Nvidia GeForce RTX 4070，显存为 12 GB，CUDA 版本为 12.1，内存为 32 GB。编程语言为 python3.11，PyTorch-GPU 2.1.2 深度学习框架。

3.2 评估指标

本文主要采用 3 个指标进行评估：mAP@.5、mAP@0.5-0.95、参数量（Params）和计算量（GFLOPs）。

mAP@.5 表示在 IoU（Intersection over Union）阈值为 0.5 时所有类别的平均准确率的平均值，mAP@0.5-0.95 则是在多个 IoU 阈值（从 0.5 到 0.95，步长为 0.05）下计算得到的平均准确率的平均值。其中，IoU 是交并比，是衡量两个边界框重叠程度的指标，是预测框与真实框交集区域与并集区域的比值，值越高表明预测框与真实框的匹配要求越高。

参数量（Params）指模型在训练过程中所涉及的全部参数的数量，被用来评估模型的空间复杂性和大小，参数数量的多少直接反映了模型的复杂程度和所需的存储空间，参数量越多，表明模型的构建更为精细和复杂，通常也意味着所需的存储容量会相应增长。

GFLOPs（giga floating point operations）代表模型每秒钟能够完成的十亿次浮点运算次数。这一指标常被用作衡量模型计算复杂性和效率的重要标准。通过 GFLOPs，可以对模型的计算能力有一个量化的认识，进而评估其在处理数据时的速度和效率。

3.3 消融实验

为验证本文提出的模型改进针对跌倒检测数据集是否有效，在相同的试验条件下，在本文跌倒检测数据集上进行试验，试验结果如表 1 所示。

表 1 消融实验结果

C2f-RGELAN	RepBi-PAN	SD-Head	mAP@.5	mAP@0.5-0.95	FLOPs /10 ⁹	Params /10 ⁶
×	×	×	0.897	0.781	28.4	11.1
√	×	×	0.910	0.789	29.8	9.1
×	√	×	0.910	0.784	31.8	9.5
×	×	√	0.912	0.784	25.2	12.2
×	√	√	0.915	0.791	31	10.1
√	√	√	0.918	0.790	25.8	8.9

从表 1 的消融实验可知，仅替换 C2f-RGELAN 模块对于原始 YOLOv8s 在 mAP@.5 提升了 1.3 个百分点，mAP@0.5-0.95 提升了 0.8 个百分点；仅在颈部网络中引入了 RepBiPAN 模块后 mAP@.5 提升了 1.3 个百分点，mAP@0.5-0.95 提升了 0.3 个百分点；仅在头部网络中引入 SD-Head 模块时，mAP@.5 提升了 1.5 个百分点，mAP@0.5-0.95 提升了 0.3 个百分点，计算量降低了 3.2×10^9 。当采用 RepBiPAN 模块和 SD-Head 模块进行改进时，mAP@.5 提升了 1.8 个百分点，mAP@0.5-0.95 提升了 0.1 个百分点，参数量降低了 1×10^6 。同时使用 C2f-RGELAN 模块、RepBiPAN 模块和 SD-Head 模块进行改进，在 mAP@.5 提升了 2.1 个百分点，mAP@0.5-0.95 提升了 0.9 个百分点，计算量下降了 2.6×10^9 ，参数量下降了 2.2×10^6 。在 300 轮迭代训练后，实验数据分析证明了本文提出的改进模型在本文跌倒数据集上的有效性，同时兼顾了轻量化。

3.4 不同模型对比实验

为验证改进模型的优势，这一小节在保持各实验环境不变的情况下，将 RRS-YOLO 与现阶段主流的目标检测算法：YOLOv5s^[20]、YOLOv7-tiny^[21]、PP-YOLOE^[22]、RT-DETR^[23]、EfficientDet^[24]、YOLOv10s^[25] 和 YOLOv11s^[26] 进行性能上的对比，实验结果如表 2 所示。

表 2 主流算法对比实验

模型	mAP@.5	mAP@0.5-0.95	FLOPs /10 ⁹	Params /10 ⁶
YOLOv8s (baseline)	0.897	0.781	28.4	11.1
YOLOv5s	0.908	0.696	16.0	7.0
RT-DETR	0.854	0.714	108	32.8
PP-YOLOE	0.909	0.706	13.9	8.4
EfficientDET	0.846	0.535	11.517	6.5
YOLOv10s	0.901	0.775	21.4	7.2
YOLOv11s	0.904	0.779	21.3	9.4
(RRS-YOLO) Ours	0.918	0.790	25.8	8.9

根据表2所示, YOLOv5s模型、RT-DETR模型、PP-YOLOE模型、EfficientDet模型、YOLOv10s和YOLOv11s模型的平均精度 $mAP@0.5-0.95$ 均低于YOLOv8s模型, 平均精度 $mAP@0.5$ 基本高于YOLOv8s模型, 但在模型权重大小方面, YOLOv8s模型对比上述6种模型位居第二。相比之下, 由于网络结构的变化, RRS-YOLO模型与YOLOv8s模型相比, 权重大小减少了19.8%且计算量也减少了9.2%, 而平均精度均优于试验中的模型。

综上所述, RRS-YOLO在精度上表现最优, 同时保持了较低的计算复杂度和模型大小, 这表明该模型在精度、计算复杂度和模型大小等方面都取得了良好的权衡。

3.5 应用场景结果分析

深入开展针对数据集中多种跌倒场景的对比测试。这涵盖光线明暗变化的环境、目标物体被遮挡以及非标准视角以及包含多人的复杂场景, 旨在全面评估RRS-YOLO模型在这些具有挑战性的条件下的表现。通过这些细致的对比试验, 能够更准确地了解模型在识别跌倒行为时的稳健性和适应性。如图4所示, 左侧呈现了原始图像, 以便观察实际场景; 中间展示了YOLOv8s算法的检测效果; 而右侧则展示了RRS-YOLO算法在相同场景下的检测结果。



图4 检测效果对比图

图4中自然光环境和暗光环境的图片源自 Fall-detection Dataset, 物体遮挡的图片来自 Multiple cameras fall Dataset, 视角遮挡的图片来自 UR Fall Detection Dataset, 多人的图片来自自建的数据集。由图4可知, 在标准光照条件以及低光照环境下, RRS-YOLO算法相较于YOLOv8s, 其检测结果的置信度都有了1%的提升。在视角遮挡的场景下, RRS-YOLO算法相较于YOLOv8s, 其检测结果的置信度有了2%的提升。在物体遮挡的场景下, YOLOv8s出现了错检, 相比之下, RRS-YOLO正确检测出了部分被遮挡的人体的跌倒状态。在人群密集的环境中, 由于个体间的相互遮挡, YOLOv8s在其检测结果的左下角未能准确识别出所有人体, 导致了漏检现象。相比之下, RRS-YOLO成功检测出了部分被遮挡的人体, 这表明在类似遮挡严重的场景下, RRS-YOLO算法展现出了更高的鲁棒性和准确性。

4 总结与展望

本文介绍了一种基于YOLOv8优化而来的跌倒检测算法RRS-YOLO。实验数据充分展示了RRS-YOLO在多种环境场景下的卓越表现, 不仅在准确性和泛化能力上有所提升, 而且在模型参数量、计算复杂度、检测精度及速度等多个维度均展现出优异性能。这一算法显著增强了室内跌倒检测的可靠性, 成功在轻量级模型与高精度要求之间找到了理想的平衡点, 为跌倒检测技术在终端设备上的部署奠定了坚实基础。

尽管RRS-YOLO已经取得了显著成就, 但其性能仍有进一步提升的空间。通过探索新的模型架构、采用更高效的优化算法或借助硬件加速技术, 可以期待模型性能的进一步优化。此外, 未来研究可以着眼于将RRS-YOLO应用于更广泛的场景, 比如户外环境或高风险工作场所, 以适应更多元化的检测需求。同时, 结合目标跟踪技术, 为室内人员分配唯一ID并进行实时跟踪, 将有望为跌倒检测提供更加精准的辅助信息。

综上所述, 未来的研究方向可以聚焦于跌倒检测模型的性能优化、应用场景的拓展以及与时跟踪技术的结合, 以期为该领域提供更加全面、高效的解决方案, 推动跌倒检测技术的持续进步与应用深化。

参考文献:

- [1] 国家统计局. 中华人民共和国2023年国民经济和社会发展统计公报[J]. 中国统计, 2024(3):4-21.
- [2] 张田田, 丰志强, 王婉晨, 等. 中国老年人跌倒现状及影响因素研究[J]. 中华疾病控制杂志, 2022, 26(5):502-507.
- [3] 陈明祥, 王钰, 刘环宇, 等. 基于人体稳定性的实时跌倒检测系统[J]. 传感器与微系统, 2023, 42(6):17-20.
- [4] 李伟, 李丹丹, 丁奇宁, 等. 基于毫米波雷达三维点云的室内跌倒检测[J]. 电子技术应用, 2024, 50(9):59-66.
- [5] WANG B H, YU J, WANG K, et al. Fall detection based on

- dual-channel feature integration[J]. IEEE access, 2020,8: 103443-103453.
- [6] 李春华, 王玲玲, 左珺, 等. 基于优化 YOLOv5s 的老人跌倒检测算法[J]. 河北科技大学学报, 2023,44(5):459-467.
- [7] 王小鹏, 石欢. 改进型 YOLOv8 融合关键点的跌倒检测算法[J]. 西安电子科技大学学报, 2024,51(5):149-164.
- [8] KWOLEK B, KEPSKI M. Human fall detection on embedded platform using depth maps and wireless accelerometer[J]. Computer methods and programs in biomedicine, 2014,117(3): 489-501.
- [9] AUVINET E, ROUGIER C, MEUNIER J, et al. Multiple cameras fall data set[R/OL]. (2010-07-08)[2024-10-20]. <https://www.researchgate.net/publication/267693256>
- [10] ADHIKARI K, BOUCHACHIA H, NAIT-CHARIF H. Activity recognition for indoor fall detection using convolutional neural network[C]//2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). Piscataway: IEEE, 2017: 81-84.
- [11] WANG C Y, YE H I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[C]//Computer Vision-ECCV 2024: 18th European Conference. New York: ACM, 2024:1-21.
- [12] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 390-391.
- [13] LI C Y, LI L L, JIANG H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07)[2024-05-23]. <https://doi.org/10.48550/arXiv.2209.02976>.
- [14] WANG C Y, LIAO H Y M, YE H I H. Designing network design strategies through gradient path analysis[EB/OL]. (2022-11-09)[2024-05-16]. <https://doi.org/10.48550/arXiv.2211.04800>.
- [15] LIU S, QIN H F, SHI J P, et al. Path aggregation network for Instance segmentation[C/OL]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018[2024-06-13]. <https://ieeexplore.ieee.org/document/8579011>. DOI: 10.1109/CVPR.2018.00913.
- [16] LI C Y, LI L L, GENG Y F, et al. YOLOv6 v3. 0: a full-scale reloading[EB/OL]. (2023-01-13)[2024-09-12]. <https://doi.org/10.48550/arXiv.2301.05586>.
- [17] CHEN P Y, CHANG M C, HSIEH J W, et al. Parallel residual bi-fusion feature pyramid network for accurate single-shot object detection[J]. IEEE transactions on image processing, 2021, 30: 9099-9111.
- [18] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 1251-1258.
- [19] YUAN X F, HUANG B, WANG Y L, et al. Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE[J]. IEEE transactions on industrial informatics, 2018, 14(7): 3235-3243.
- [20] YOLOv5[EB/OL]. (2020-06-26)[2024-11-26]. <https://github.com/ultralytics/YOLOv5>.
- [21] WANG C Y, BOCHKOVSKIY A, LIAO H Y 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). Piscataway: IEEE, 2023: 7464-7475.
- [22] XU S L, WANG X X, LÜ W Y, et al. PP-YOLOE: an evolved version of YOLO[EB/OL]. (2022-12-12)[2024-05-13]. <https://doi.org/10.48550/arXiv.2203.16250>.
- [23] ZHAO Y, LÜ W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[EB/OL]. (2024-04-03)[2024-09-11]. <https://doi.org/10.48550/arXiv.2304.08069>.
- [24] TAN M M, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 10781-10790.
- [25] WANG A, CHEN H, LIU L H, et al. YOLOv10: real-time end-to-end object detection[EB/OL]. (2024-10-30)[2024-05-11]. <https://doi.org/10.48550/arXiv.2405.14458>.
- [26] YOLOv11[DB/OL]. (2024-09-30)[2024-11-26]. <https://github.com/ultralytics/YOLOv11>.

【作者简介】

叶智国 (1999—), 男, 江西上饶人, 硕士研究生, 研究方向: 人工智能, email: 71972144@qq.com。

熊盛佳 (2000—), 女, 湖南益阳人, 硕士研究生, 研究方向: 人工智能, email: 1483551312@qq.com。

吴章云 (1999—), 男, 湖南郴州人, 硕士研究生, 研究方向: 人工智能, email: 741124643@qq.com。

杜远飞 (1998—), 男, 贵州遵义人, 硕士研究生, 研究方向: 农业机械化, email: 1907069320@qq.com。

吴传宇 (1976—), 男, 福建建瓯人, 博士, 高级工程师, 研究方向: 农业机械化, email: 360626964@qq.com。

陈仕国 (1986—), 通信作者 (email: 117796657@qq.com), 男, 福建连江人, 硕士, 实验师, 研究方向: 深度学习、机电一体化。

(收稿日期: 2025-01-07 修回日期: 2025-05-12)