

基于改进YOLOX的行人检测与跟踪算法

赵士博¹, 李 鹤^{2*}, 李世博¹, 庞博阳¹

¹沈阳航空航天大学电子信息工程学院, 辽宁 沈阳

²沈阳工学院信息与控制学院, 辽宁 抚顺

收稿日期: 2026年6月18日; 录用日期: 2026年7月16日; 发布日期: 2026年7月23日

摘 要

行人多目标跟踪(MOT)在智能交通、自动驾驶及公共安全监控领域应用广泛。针对现有跟踪系统存在的小目标检测精度低、密集场景易漏检、遮挡条件下身份频繁切换、轻量化设计与检测性能难以兼顾等问题, 本文提出一种基于YOLOX与ByteTrack的改进型行人检测跟踪框架SSA-YOLOX。该框架引入SPHead小目标检测分支, 强化浅层高分辨率特征提取, 提升远距离行人与密集目标的感知能力; 在特征融合阶段嵌入无参SimAM注意力机制, 增强特征判别性并抑制背景干扰; 采用轻量化ADown下采样模块, 降低计算复杂度且保留关键特征信息; 将优化后的检测器与ByteTrack集成, 构建完整的多目标跟踪系统。在MOT17数据集上的实验结果表明, SSA-YOLOX以8.92 M参数量实现94.1%的AP@0.5, 较原始YOLOX提升2个百分点; 跟踪指标达到84.5%的MOTA与80.4%的IDF1, 身份切换次数显著降低。所提方法在检测精度、跟踪鲁棒性与模型轻量化之间实现良好平衡, 可满足实时行人跟踪的工程应用需求。

关键词

YOLOX, SimAM注意力, ByteTrack, 行人检测, 多目标跟踪

Pedestrian Detection and Tracking Algorithm Based on Improved YOLOX

Shibo Zhao¹, He Li^{2*}, Shibo Li¹, Boyang Pang¹

¹Electronic Information Engineering College, Shenyang Aerospace University, Shenyang Liaoning

²College of Information and Control, Shenyang Institute of Technology, Fushun Liaoning

Received: June 18, 2026; accepted: July 16, 2026; published: July 23, 2026

Abstract

Pedestrian Multi-Object Tracking (MOT) is widely applied in intelligent transportation, autonomous

*通讯作者。

文章引用: 赵士博, 李鹤, 李世博, 庞博阳. 基于改进 YOLOX 的行人检测与跟踪算法[J]. 计算机科学与应用, 2026, 16(7): 92-101. DOI: 10.12677/csa.2026.167244

driving and public security monitoring. To address the problems of low small-object detection accuracy, frequent missed detections in dense scenes, identity switches under occlusion and the trade-off between lightweight design and detection performance, this paper proposes SSA-YOLOX, an improved pedestrian detection and tracking framework based on YOLOX and ByteTrack. The framework introduces the SPHead small-object detection branch to strengthen shallow high-resolution feature extraction and improve the perception of distant pedestrians and dense targets; embeds the parameter-free SimAM attention mechanism into the feature fusion stage to enhance feature discrimination and suppress background interference; adopts the lightweight ADown downsampling module to reduce computational complexity while preserving key feature information; and integrates the optimized detector with ByteTrack to construct a complete multi-object tracking system. Experiments on the MOT17 dataset show that SSA-YOLOX achieves 94.1% AP@0.5 with 8.92 M parameters, which is 2% higher than the original YOLOX. The tracking metrics reach 84.5% MOTA and 80.4% IDF1, and the number of identity switches is significantly reduced. The proposed method achieves a favorable balance among detection accuracy, tracking robustness and model lightweight, and can meet the engineering application requirements of real-time pedestrian tracking.

Keywords

YOLOX, SimAM Attention, ByteTrack, Pedestrian Detection, Multi-Object Tracking

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

行人多目标跟踪(Multi-Object Tracking, MOT) [1] [2]是智能安防、自动驾驶、交通流量监测等领域的核心支撑技术,其性能直接决定了复杂监控场景下系统的可靠性与鲁棒性。然而,现有跟踪算法在实际落地中仍面临三大核心挑战:其一,密集人群与远距离小目标场景下检测精度不足,易引发轨迹断裂与目标漏检;其二,目标遮挡与群体交叉运动时,特征判别性不足导致目标身份频繁切换;其三,高精度检测模型参数量与计算开销大,难以在算力受限的边缘设备上实现实时部署。

当前主流的行人多目标跟踪方案普遍采用“检测-跟踪”串行范式[3],系统整体性能高度依赖前端检测器的输出质量。YOLOX [4]作为轻量化无锚框检测器,凭借优异的实时性与基础检测性能被广泛应用,但该模型的连续下采样操作易造成小目标特征衰减,且复杂背景下的特征判别能力不足,难以适配密集人群、目标遮挡等典型复杂场景。ByteTrack [5]作为当前性能优异的跟踪算法,通过高低置信度检测框二次关联策略提升了轨迹稳定性,但其跟踪效果对检测器输出质量高度敏感,与原生 YOLOX 搭配时,仍无法从根本上解决漏检引发的轨迹断裂与身份跳变问题。

针对上述问题,本文从三个维度对 YOLOX 检测器进行系统性改进,构建了 SSA-YOLOX 检测框架。首先,在网络结构中增设 SPHead 小目标检测分支,充分挖掘骨干网络输出的浅层高分辨率特征,强化模型对远距离行人与小尺寸目标的捕捉能力;其次,在特征融合环节嵌入无参 SimAM [6]注意力机制,在不额外增加模型计算开销的基础上,提升目标特征与背景特征的区分度;最后,采用轻量化 ADown 模块替换原始卷积下采样结构,在降低模型参数量与运算量的同时保留图像细节特征,实现检测精度与推理效率的协同优化。完成检测器优化后,本文将 SSA-YOLOX 与 ByteTrack 跟踪器深度集成,搭建端到端行人多目标跟踪系统,旨在改善复杂场景下轨迹碎片化、身份频繁切换等问题,适配边缘设备的实时工程应用需求。

2. SSA-YOLOX 模型构建

2.1. YOLOX 基准模型框架

YOLOX 是 2021 年提出的无锚框目标检测框架，凭借无锚框设计与多尺度特征融合结构，在检测精度与推理效率间实现了良好平衡，被广泛应用于实时目标检测任务。本研究选取其作为基线模型，后续与 ByteTrack 跟踪器集成，构建完整的行人多目标跟踪系统。

YOLOX 的整体结构如图 1 所示，由三部分核心模块构成：主干特征提取网络采用 CSPDarknet 结构 (图 1 绿色框)，通过多阶段下采样与 CSP、SPP 等单元优化特征学习，生成 C_3 、 C_4 、 C_5 三层多尺度特征图；颈部采用 PAFPN 结构(图 1 蓝色框)，结合 FPN 与 PAN 实现双向特征融合，强化特征的语义表达与定位能力；解耦检测头(图 1 红色框)将分类与回归任务分离，兼顾模型收敛速度与检测精度，输出目标检测结果。

尽管 YOLOX 具备优异的基础性能，但在密集行人、远距离小目标与遮挡场景下，仍存在小目标特征衰减、复杂背景下特征判别性不足等问题，难以适配本研究的复杂监控场景需求，因此本文针对其不足进行系统性改进，构建 SSA-YOLOX 检测框架。

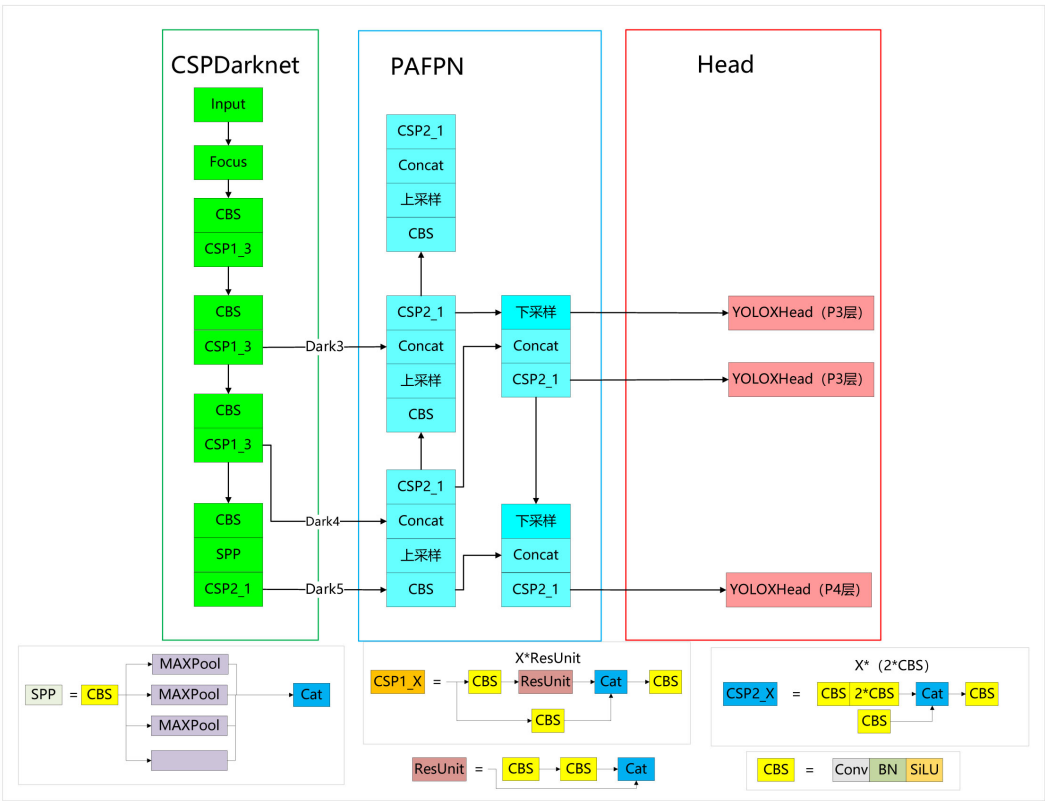


Figure 1. Baseline YOLOX
图 1. 基线 YOLOX

2.2. SSA-YOLOX 检测模型

针对现有行人多目标跟踪方法在密集人群、小目标及遮挡场景下的性能瓶颈，本研究以 YOLOX 为基线模型，从多尺度特征增强、轻量化下采样优化与特征判别性能提升三个维度对检测器进行系统性改

进，构建 SSA-YOLOX 检测框架，后续与 ByteTrack 跟踪器集成形成完整的跟踪系统，改进后的整体结构如图 2 所示。

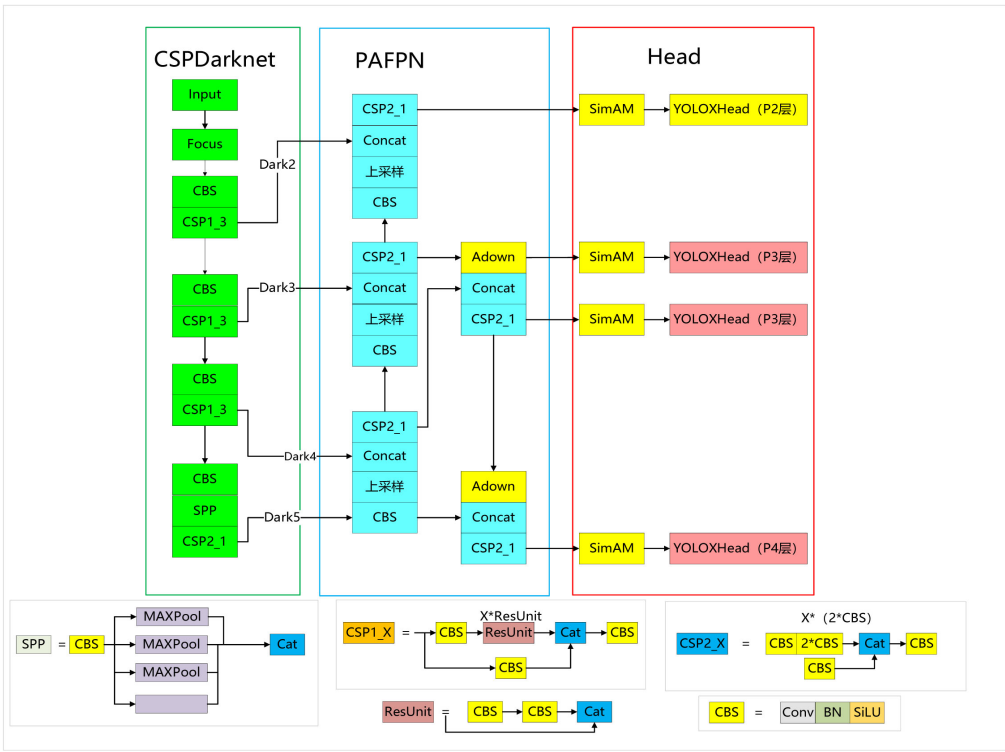


Figure 2. SSA-YOLOX
图 2. SSA-YOLOX

为强化小目标与密集场景下的特征提取能力，本文新增 P2 小目标检测分支，利用 CSPDarknet 骨干网络输出的 Dark2 层高分辨率特征，保留更多细粒度空间信息，增强网络对远距离行人、小尺寸目标及密集场景弱特征目标的感知能力；同时结合改进的 PAFPN 多尺度特征融合结构，实现浅层细粒度特征与深层语义特征的充分交互，缓解传统 YOLOX 因下采样导致的小目标特征丢失问题，提升小目标检测的召回率与定位精度。

在 PAFPN 特征融合阶段，本文采用轻量化 ADown 模块替换传统卷积下采样结构，优化特征压缩与信息传递方式，在降低特征图分辨率的同时减少关键信息损失。相比标准卷积下采样，ADown 模块在保证特征提取能力的基础上，有效降低了模型参数量与计算复杂度，提升了网络轻量化程度与推理效率，为边缘设备的实时部署提供支撑。

为增强特征判别性、抑制复杂背景干扰，本文在 P2~P4 所有尺度的检测头前嵌入无参 SimAM 注意力机制，通过自适应权重分配强化目标区域特征，在不引入额外可学习参数的前提下，提升目标与背景特征的区分度，缓解密集人群、遮挡及复杂背景下的漏检与误检问题，提升模型检测鲁棒性。

完成检测器改进后，将 SSA-YOLOX 与 ByteTrack 跟踪框架集成，利用高质量检测结果为目标关联与轨迹更新提供支撑。优化后的检测器能够为 ByteTrack 提供更稳定、高精度的目标检测输出，有效改善复杂场景下因漏检、误检导致的轨迹断裂与身份切换问题，提升目标遮挡、快速运动及密集交互场景下的轨迹连续性与跟踪稳定性。

2.3. SPHead 小目标检测分支

小尺寸行人在多次下采样中易丢失边缘与纹理特征，为此本文增设 SPHead 检测分支。SSA-YOLOX 基于骨干网络输出的 160×160 高分辨率特征图，新增并行 P2 检测头，与原 P3~P5 检测头协同推理，强化对小目标与远距离行人的检测灵敏度。

SPHead 沿用 YOLOX 的解耦头结构，分类与回归分支独立优化、不共享参数，可从高分辨率 P2 特征中提取细粒度空间信息，提升远距离行人定位精度。特征融合阶段，P2 浅层特征经 PAN/FPN 与深层语义特征融合，训练时所有检测头采用联合监督策略。通过保留浅层细粒度特征，SPHead 显著提升密集人群、远距离行人与弱特征目标的召回率与定位精度。

2.4. ADown 轻量化下采样模块

传统卷积下采样在特征压缩过程中易丢失细节，降低小目标检测性能。本文以轻量化 ADown 模块替换标准卷积下采样结构。ADown 先通过平均池化保留局部结构信息，再沿通道维度将特征分为双分支：一支采用卷积实现高效特征提取，另一支融合最大池化与 1×1 卷积保留显著特征，最终通过通道拼接完成多尺度特征融合。

相较于标准卷积下采样，ADown 通过轻量化分支设计减少计算冗余，将参数量从 9.57 M 降至 8.92 M，GFLOPs 从 29.54 降至 29.12，同时保持高精度检测。此外，ADown 缓解下采样过程中的过度特征压缩，提升特征保留效果并稳定深层梯度传播，优化模型推理效率，适配边缘设备实时部署。

2.5. SimAM 注意力机制

密集行人场景与动态遮挡常会弱化目标特征、增强背景干扰。为提升特征判别性，本文在检测头前嵌入无参 SimAM 注意力机制。与 SE [7]、CBAM [8] 等注意力机制不同，SimAM 不引入额外可学习参数。受神经科学空间抑制理论启发，SimAM 依据同一通道内神经元间的线性可分性评估神经元重要性。

对于特征图 $X \in R^{C \times H \times W}$ ，首先计算每个通道内神经元的均值与方差，再通过能量函数估计神经元重要性，经自适应权重分配实现特征增强。增强后的特征图可表示为：

$$\tilde{X} = \text{Sigmoid}\left(\frac{1}{E}\right) \odot X \quad (1)$$

式中，Sigmoid 为 Sigmoid 激活函数， $\odot$ 为哈达玛积。

SimAM 通过联合优化空间与通道特征，强化行人目标与复杂背景间的区分度。由于未引入额外参数，该模块在提升遮挡与密集人群干扰鲁棒性的同时，保留实时推理能力。

2.6. 改进 YOLOX 融合 ByteTrack 多目标跟踪算法

本文采用检测-跟踪范式，将优化后的 SSA-YOLOX 检测器与 ByteTrack 跟踪器集成，构建端到端行人多目标跟踪系统。ByteTrack 摒弃传统跟踪器仅使用高置信度检测框的策略，采用双阈值关联机制：先通过匈牙利算法与卡尔曼滤波完成高置信度框的轨迹匹配，再将低置信度框用于未匹配轨迹的二次关联，有效缓解遮挡、小目标漏检导致的轨迹断裂问题。

该集成方案中，SSA-YOLOX 负责输出高精度、高连续性的行人检测框，为 ByteTrack 提供可靠的输入；ByteTrack 则通过置信度分层关联策略，最大化利用检测信息，提升轨迹连续性与身份一致性。二者结合可有效解决密集遮挡场景下的身份切换、轨迹碎片化难题，充分发挥检测优化与跟踪策略的协同优势。

3. 实验结果与分析

3.1. 实验环境

实验基于 PyTorch2.1.2、CUDA11.8 搭建环境,硬件配置为 RTX2070GPU 与 IntelCorei7-10750HCPU,采用 MOT17 数据集完成检测与跟踪评估。所有实验遵循 MOTChallenge 公开检测协议,检测结果由 SSA-YOLOX 自主生成,再经 ByteTrack 完成轨迹关联。

训练超参数如表 1 所示,检测指标包括 AP@0.5:0.95、小目标 AP@0.5:0.95、AP@0.5、参数量(M)、GFLOPs;跟踪指标包括 MOTA、IDF1、身份切换次数(IDS)、假阳性(FP)、假阴性(FN)。

Table 1. Training hyperparameters

表 1. 训练超参数

Hyperparameters	Configuration
Image Size	640 * 640 * 3
Epoch	200
Batch size	4
Learning Rate	6.25×10^{-4}
Momentum	0.9
Optimizer	SGD
Weight Decay	5×10^{-4}
Close mosaic	15

3.2. 不同注意力机制对比

为验证不同注意力机制在行人检测任务中的性能差异,本文在 YOLOX + ADown 网络结构基础上分别引入 SE、CBAM、ECA、CA 和 SimAM 注意力机制,并在相同训练配置下进行对比实验,实验结果如表 2 所示。

从表 2 可以看出,引入注意力机制后,各模型在检测性能方面均有所提升,但提升幅度存在一定差异。SE、CBAM 和 ECA 等经典注意力机制的 AP@0.5:0.95 分别达到 73.1%、73.3%和 73.2%,小目标检测指标 APs 分别为 43.8%、44.0%和 43.9%,相较于 YOLOX + ADown 基线模型(AP@0.5:0.95 为 72.9%,APs 为 44.0%)仅取得有限提升。同时,这些方法均引入了额外的可学习参数,导致模型参数量和计算量出现不同程度增长。相比之下,CA 与 SimAM 均表现出更优的特征增强能力。其中,CA 将 AP@0.5:0.95 提升至 73.5%,APs 提升至 44.3%;而 SimAM 在无需引入额外可学习参数的情况下取得了最佳检测性能,其 AP@0.5:0.95 达到 75.4%,较基线模型提升 2.5 个百分点,较其他注意力机制提升约 2 个百分点;小目标检测指标 APs 达到 47.3%,较基线模型提升 3.3 个百分点,表明其能够更加有效地增强小目标特征表达能力。

此外,SimAM 的参数量仅为 8.29M,GFLOPs 为 12.96,与基线模型保持一致,未对模型复杂度造成额外负担。综合来看,SimAM 在检测精度、小目标检测能力以及模型轻量化之间取得了更优的平衡,验证了其在行人检测任务中的有效性和适用性。因此,本文最终选择 SimAM 作为 SSA-YOLOX 的注意力增强模块。

Table 2. Performance comparison of attention mechanisms
表 2. 注意力机制性能对比

	AP@0.5:0.95 (%)	APs (%)	AP@0.5 (%)	Para (M)	GFLOPs
CA	73.5	44.3	92.1	8.32	12.97
SE	73.1	43.8	92.1	8.35	13.01
CBAM	73.3	44.0	92.1	8.38	13.03
ECA	73.2	43.9	92.2	8.31	12.98
SimAM	75.4	47.3	92.2	8.29	12.96

3.3. 检测模块消融实验

为验证各模块有效性，开展消融实验，结果如表 3 所示。

Table 3. Ablation results of detection modules
表 3. 检测模块消融实验

YOLOX	SPHead	SimAM	ADown	AP@0.5:0.95 (%)	APs (%)	AP@0.5 (%)	Para (M)	GFLOPs
√	√			73.1	43.7	91.1	8.94	13.38
√		√		73.3	45.2	92.2	8.94	13.38
√			√	72.9	44.0	92.2	8.29	12.96
√		√	√	75.4	47.3	92.2	8.29	12.96
√	√	√		76.6	49.5	93.9	9.57	29.54
√	√		√	76.7	49.3	93.9	8.92	29.12
√	√	√	√	78.2	51.1	94.1	8.92	29.12

首先，引入 SPHead 小目标检测分支，模型性能显著提升：AP@0.5:0.95 由基线 73.1% 升至 76.7%，AP@0.5:0.95 (small) 由 43.7% 升至 49.3%，AP@0.5 由 92.1% 升至 94.1%。该结果表明，SPHead 通过浅层高分辨率特征提取，有效解决小目标下采样特征衰减问题，大幅提升小目标检测召回率。其次，引入 SimAM 后，AP@0.5:0.95 由 76.7% 变化为 76.6%，整体保持稳定；与此同时，小目标检测指标 APs 由 49.3% 提升至 49.5%。说明 SimAM 对整体精度提升有限，但能够进一步增强小目标与复杂背景场景下的特征表达能力，为后续模块协同优化提供基础。最后，引入 ADown 轻量化下采样模块，模型参数量由 9.57 M 降至 8.92 M (减少 0.65 M)，GFLOPs 由 29.54 降至 29.12；同时检测性能回升至最优：AP@0.5:0.95 = 78.2%、AP@0.5:0.95 (small) = 51.1%、AP@0.5 = 94.1%。实验证实，ADown 在降低计算开销的同时，可保留关键特征信息，实现轻量化与精度的平衡。

综合消融结果，最终 SSA-YOLOX 模型仅以 8.92 M 参数量、29.12GFLOPs，实现 AP@0.5:0.95 = 78.2%、AP@0.5:0.95 (small) = 51.1%、AP@0.5 = 94.1% 的最优性能，较基线模型在小目标检测、遮挡鲁棒性、轻量化三方面均实现显著提升。

3.4. 跟踪性能评估

跟踪消融实验结果如表 4 所示。SPHead 提升小目标召回率、减少漏检，为跟踪提供连续检测输入；SimAM 增强遮挡场景特征判别性，改善轨迹连续性。

Table 4. Ablation results of tracking performance
表 4. 跟踪性能消融实验

YOLOX	SPHead	SimAM	ADown	MOTA (%)	IDF1 (%)	IDS	FP	FN
√	√			83.3	79.2	1196	16,249	38,887
√	√	√		84.2	80.6	1083	14,396	37,857
√	√	√	√	84.5	80.4	992	14,922	37,327

跟踪消融实验结果表明，SPHead 通过减少小目标漏检，为 ByteTrack 提供连续检测输入，使 MOTA 达到 83.3%；SimAM 通过提升特征判别性，进一步将 MOTA 提升至 84.2%，身份切换次数降至 1083 次；全模块集成后，SSA-YOLOX + ByteTrack 的 MOTA 达 84.5%，身份切换次数仅 992 次，较基线降低 17.1%。

值得强调的是，ByteTrack 仅依赖检测置信度完成轨迹关联，并未引入外观重识别(ReID)模块。因此，身份切换次数的下降主要来源于 SSA-YOLOX 检测质量的提升，即漏检与轨迹断裂现象的减少。这表明本文提出的检测改进策略能够有效提升后续跟踪性能。

3.5. 对比实验

将 SSA-YOLOX 与 YOLOv5、YOLOv8、YOLOv10、原生 YOLOX 对比可知：本文所提 SSA-YOLOX 的 AP@0.5=94.1%，为所有模型中最高；参数量仅 8.92 M，与原生 YOLOX (8.94 M)基本持平，远低于 YOLOv8 (11.2 M)；尽管 GFLOPs 因 SPHead 分支升至 29.12，但 ADown 模块抵消部分计算开销，仍保持轻量化优势。实验表明，SSA-YOLOX 在精度与轻量化之间取得最优平衡，更适合边缘端实时部署。对比结果如表 5 所示。

Table 5. Comparison with mainstream detection models
表 5. 主流检测模型对比

	AP@0.5 (%)	Para (M)	GFLOPs
YOLOv5	87.5	7.2	16.5
YOLOv8	89.6	11.2	28.6
YOLOv10	88.1	8.5	21.8
YOLOX	92.1	8.94	13.38
SSA-YOLOX	94.1	8.92	29.12

跟踪算法对比如表 6 所示，本文方法在 MOTA、IDF1 与 IDS 指标上均显著优于 SORT、DeepSORT 与原始 ByteTrack，跟踪鲁棒性与身份保持能力优势突出。

Table 6. Comparison of tracking algorithms on MOT17
表 6. MOT17 数据集上跟踪算法对比

	MOTA (%)	IDF1 (%)	IDS	FP	FN
Sort	77.8	72.1	2420	23,820	39,650
DeepSort	79.6	74.5	2396	21,155	39,122
ByteTrack	82.6	76.5	2133	19,418	37,141
Ours	84.5	80.4	992	14,922	37,327

3.6. YOLOX 与 ByteTrack 结合验证

本研究选取 MOT17-03-SDP 测试序列中目标密集、行人交互频繁的典型片段，将所提 SSA-YOLOX + ByteTrack 算法与原生 YOLOX + ByteTrack 基线方案进行可视化跟踪对比，结果如图 3 所示。



Figure 3. Visualized comparison of tracking results for MOT17-03 test sequence

图 3. MOT17-03 测试视频可视化跟踪对比结果

在基线算法的跟踪结果中，连续帧的跟踪过程暴露出明显缺陷：一是目标漏检与检测延迟问题，画面左侧穿红衣的行人在第 131 帧未被基线算法检测到，直至第 153 帧才被识别并分配 ID 80，同时第 131 帧中画面边缘的弱特征目标(ID 49)在第 153 帧出现目标丢失，未被持续跟踪；二是密集场景下特征鲁棒性不足，行人间的特征干扰导致基线算法易出现 ID 分配混乱，影响轨迹的连续性与稳定性。

而本文所提算法在相同场景下表现出显著的性能提升：对于基线算法漏检的红衣行人，所提算法在第 131 帧即可稳定检测并分配 ID27，且该 ID 在第 153 帧中保持不变，未出现轨迹断裂；同时，所提算法对画面边缘的弱特征目标、密集场景下的所有行人目标均实现了稳定检测，帧间目标 ID 未出现跳变或丢失，有效保障了目标轨迹的连续性。

这一可视化结果与前文定量实验中 APs 指标提升、IDS 指标下降的结论完全吻合，上述可视化结果与表 3 中 APs 提升 7.4 个百分点以及表 4 中 IDS 下降 17.1%的定量结果相一致，从视觉层面验证了 SSA-YOLOX 在小目标检测与身份保持方面的优势。

4. 结论

本文针对复杂场景行人多目标跟踪的小目标检测差、遮挡易漏检、精度与轻量化失衡问题，提出融合 SPHead、SimAM、ADown 模块的 SSA-YOLOX 检测算法，并与 ByteTrack 跟踪框架结合，构建完整的行人多目标跟踪系统。实验在 MOT17 数据集上，SSA-YOLOX 的 AP@0.5 = 94.1%，较原生 YOLOX (92.1%)提升 2 个百分点；集成 ByteTrack 后，跟踪指标达 MOTA = 84.5%、IDF1 = 80.4%，身份切换次数仅 992 次，较原生 ByteTrack 降低 17.1%，参数量仅 8.92 M；跟踪指标 MOTA 达 84.5%、IDF1 达 80.4%，身份切换次数大幅降低，跟踪性能优于 SORT、DeepSORT、原生 ByteTrack 及主流 YOLO 系列算法，在密集遮挡场景下具备更优的检测与跟踪鲁棒性，且满足实时应用要求。未来将进一步优化模型轻量化程

度与泛化能力，推进算法在边缘设备的部署落地。

参考文献

- [1] Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W. and Kim, T. (2021) Multiple Object Tracking: A Literature Review. *Artificial Intelligence*, **293**, Article 103448. <https://doi.org/10.1016/j.artint.2020.103448>
- [2] Chan, S., Jia, Y., Zhou, X., Bai, C., Chen, S. and Zhang, X. (2022) Online Multiple Object Tracking Using Joint Detection and Embedding Network. *Pattern Recognition*, **130**, Article 108793. <https://doi.org/10.1016/j.patcog.2022.108793>
- [3] Divyaprabha, D. and Seebaiah, G. (2024) A Cost-Effective, Reliable and Accurate Framework for Multiple-Target Tracking by Detection Approach Using Deep Neural Network. *International Journal of Electrical and Computer Engineering (IJECE)*, **14**, 5681-5690. <https://doi.org/10.11591/ijece.v14i5.pp5681-5690>
- [4] Zhu, L., Zhao, H. and Wang, W.L. (2023) An Object Detection Method in Foggy Weather Based on Novel Feature Enhancement and Fusion. *Journal of Xi'an Polytechnic University*, **37**, 106-113. (In Chinese)
- [5] Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., *et al.* (2022) ByteTrack: Multi-Object Tracking by Associating Every Detection Box. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M. and Hassner, T., Eds., *Lecture Notes in Computer Science*, Springer, 1-21. https://doi.org/10.1007/978-3-031-20047-2_1
- [6] Guo, L., Yu, H.Y. and Zhou, Z.Q. (2023) Nearshore Ship Detection Method Based on SimAM Attention Mechanism. *Journal of Harbin Institute of Technology*, **55**, 14-21.
- [7] Zhang, Y., Liu, X. and Zhao, X. (2024) YOLO-SE: A Universal Real-Time Multi-Scale Object Detection Framework. *2024 IEEE International Conference on Imaging Systems and Techniques (IST)*, Tokyo, 14-16 October 2024, 1-7. <https://doi.org/10.1109/ist63414.2024.10759243>
- [8] Ma, Q. (2024) YOLOv5-CBAM: A Small Object Detection Model Based on YOLOv5 and CBAM. *2024 6th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI)*, Nanjing, 6-8 December 2024, 618-623. <https://doi.org/10.1109/ricai64321.2024.10911839>