Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2025, 14(6), 1489-1498
Published Online November 2025 in Hans. https://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2025.146139

基于改进YOLOv8的复杂环境行人识别模型

韦剑锋

广东工业大学物理与光电工程学院，广东 广州

收稿日期：2025年10月16日；录用日期：2025年11月12日；发布日期：2025年11月21日

摘　要

针对现有行人识别模型在复杂场景下，存在小目标及密集目标误检漏检率高等问题，提出了基于YOLOv8改进的行人检测模型Swin-YOLO。首先设计了SE-Conv模块，该模块通过将卷积操作和注意力机制进行融合，有效提升了模型的特征选择能力；其次，引入Swin Transformer (ST)模块，利用ST模块大感受野特性提升模型的长距离建模能力；最后，在输出端构建了基于深度卷积与空洞卷积的多尺度卷积模块，该模块旨在以较低的计算复杂度实现高效的多尺度特征融合。实验结果表明，在小目标场景数据集TinyPerson中，Swin-YOLO模型相比YOLOv8模型在检测精度指标mAP@0.5和mAP@0.5:0.95上分别提升了6.5和2.5个百分点，有效减少了误检和漏检，为解决智能辅助驾驶及道路状况预警提供一种有效的改进YOLOv8行人识别模型。

关键词

行人图像识别，多尺度，注意力，YOLO

Pedestrian Recognition Model for Complex Environment Based on Improved YOLOv8

Jianfeng Wei

School of Physics and Optoelectronic Engineering, Guangdong University of Technology, Guangzhou Guangdong

Received: October 16, 2025; accepted: November 12, 2025; published: November 21, 2025

Abstract

To address the high false positive and false negative rates of existing pedestrian detection models in complex scenes, particularly for small and densely clustered targets, we propose Swin-YOLO, an improved pedestrian detection model based on YOLOv8. First, the SE-Conv module was designed, which effectively enhances the model's feature selection capability by integrating convolutional

文章引用: 韦剑锋. 基于改进 YOLOv8 的复杂环境行人识别模型[J]. 人工智能与机器人研究, 2025, 14(6): 1489-1498.
DOI: 10.12677/airr.2025.146139

operations with attention mechanisms. Second, the Swin Transformer (ST) module was introduced to leverage its large receptive field for improved long-range modeling. Finally, a multi-scale convolutional module based on deep convolutions and dilated convolutions was constructed at the output layer to achieve efficient multi-scale feature fusion with reduced computational complexity. Experimental results demonstrate that on the TinyPerson dataset for small-object scenarios. The Swin-YOLO model outperforms the YOLOv8 model by 6.5 and 2.5 percentage points in mAP@0.5 and mAP@0.5:0.95 respectively. This effectively reduces false positives and false negatives, to address intelligent driving assistance and road condition warnings, an effective improved YOLOv8 pedestrian detection model is provided.

Keywords

Pedestrian Image Recognition, Multi-Scale, Attention, YOLO

1. 引言

行人识别是计算机视觉任务中的一个热门研究技术，其目的在于让计算机自动在图像或视频中检测、定位并识别出行人。这些年来，得益于深度学习的高速持续发展，卷积神经网络(Convolutional Neural Networks, CNN)凭借其优异的特征提取能力，极大地促进了计算机视觉领域的跨越式发展[1]，基于深度学习的行人图像检测方法在准确率方面取得了显著提升，在视频监控、自动驾驶和交通管理等领域得到了广泛应用。其核心优势在于端到端的特征习得机制，不需要人为地设计抽取模型，就可以利用神经网络自主地从海量数据中挖掘与任务相关的有效特征[2]。在目标检测领域，深度学习算法总体上被划分为二阶段算法和一阶段算法两大类。由于二阶段算法实时性较差，且对小目标的检测能力不足，因此单阶段行人图像识别方法逐步成为了主流研究方向。

一阶段算法直接借助 CNN 在全图开展密集预测，能够一步完成行人的定位与分类任务，检测速度也更快，在小目标检测上的效率更高。代表算法有 SSD (Single Shot MultiBox Detector) [3]、YOLO (You Only Look Once) [4]系列。高嘉振等[5]在 YOLOv3 算法中，引入空间金字塔池化(SPP, Spatial Pyramid Pooling)模块强化特征表达能力，在此基础上研究基于多尺度特征的融合方法，实现对检测物体的空间与语义两方面的信息有效融合，以克服目标像素占比小造成的检漏问题。Fu 等[6]提出了一种基于 YOLOv5s 改进的 YOLOv5s-DC 算法，通过两个不同的卷积对 YOLO 头部进行解耦操作，以此得到回归与分类结果，并且使得回归及分类任务的关注度有效提升。

虽然上述方法有效的提升了行人图像识别任务的精度，但是当面临密集目标、小目标等复杂情况时，由于卷积核尺寸的约束导致模型感受野有限，难以有效捕捉全局上下文信息，且现有架构缺乏对关键特征的关注程度，模型在实际检测中容易出现误检和漏检等问题。

针对上述问题，本文提出了一种改进 YOLOv8 的行人检测模型 Swin-YOLO，Swin-YOLO 在最大程度保持 YOLOv8 原有结构的基础上进行了改进：首先，本文提出了一种基于通道注意力机制的 SE-Conv 模块，该模块通过将通道注意力 SE-Net [7]与标准卷积块进行深度融合，有效增强了模型对关键特征的关注能力。其次，为了提升模型长距离建模能力，使其充分融合上下文信息，引入了 Swin Transformer [8]并优化，提出了 ST (Swin Transformer, ST)模块。最终，在模型的输出端构建了基于深度卷积[9]和空洞卷

积[10]的多尺度卷积模块(Multi-scale Convolution Module, MCM)，通过增加网络宽度并集成不同尺寸的卷积核，实现高效多尺度特征融合。该设计在保持较低计算复杂度的同时，增强了模型对多尺度特征的提取能力，提升了特征表达的多样性和鲁棒性。

2. 改进方法

2.1. YOLOv8

YOLO (You Only Look Once)是一种基于单阶段检测思想的目标检测算法，该算法采用端到端的回归网络设计，将目标定位与分类任务统一在一个卷积神经网络(CNN)架构中。通过直接回归预测物体的边界框坐标和类别概率，有效避免了传统检测算法中复杂的候选区域生成过程，显著提升了检测效率。经过多代技术迭代，YOLO 系列已发展出包括 YOLOv5 [11]、YOLOv8 [12]等在内的完整算法体系，其中 YOLOv8 凭借其优异的性能平衡性，在计算机视觉基础研究和行人识别等[13]实际应用中展现出强大的实用价值，其核心特点在于高性能与智能化的平衡设计。

2.2. Swin-YOLO 网络模型结构

本节详细介绍 Swin-YOLO 网络模型的结构，Swin-YOLO 由骨干网络、颈部网络、 头部网络三部分组成，其结构如图 1 所示。

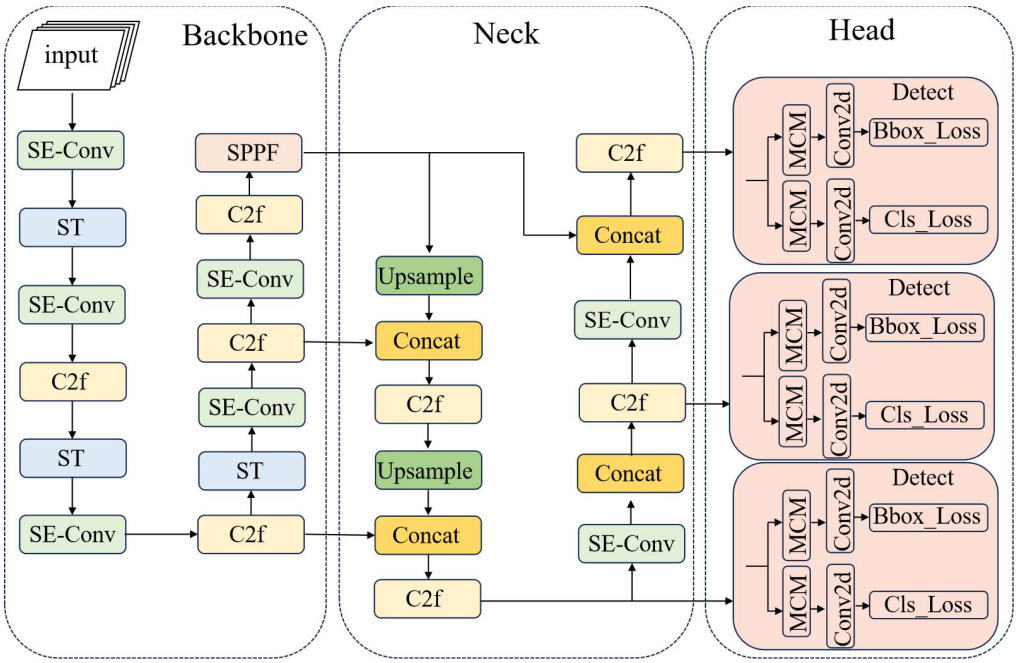

Figure 1. Structure of Swin-YOLO network model
图 1. Swin-YOLO 网络模型的结构

2.2.1. 骨干网络

(1) SE-Conv 模块

SE-Conv 模块由 3 × 3 卷积层、批量归一化(BN)、SiLU 激活函数和 SE 注意力机制组成，其结构如图 2 所示。该模块首先利用 3 × 3 卷积提取局部空间特征，随后通过 SE 注意力机制动态校准通道间的特征响应。SE 模块的 Squeeze 阶段通过全局平均池化压缩空间信息，获取通道级的全局表征，Excitation 阶段

则通过门控机制学习各通道的权重分布，自适应地强化关键特征通道并弱化冗余信息。这种设计显式地建立了跨通道的长程依赖关系，使模型能够更有效地捕捉复杂场景中的判别性特征，从而显著提升整体性能。

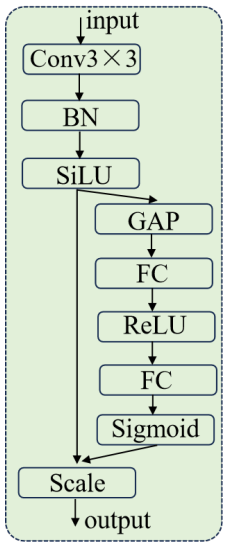

Figure 2. Structure of SE-Conv module
图 2. SE-Conv 模块结构

(2) Swin Transformer 模块

针对传统卷积神经网络因局部感受野限制而难以建模长距离依赖关系的问题，本章在 YOLOv8 骨干网络中引入了 Swin Transformer (ST)模块，由于 SE-Conv 模块中集成了 SE 注意力机制，因此为了避免信息冗余和计算效率低下的问题，本章移除原始 Swin Transformer 模块中的 MLP，使模块更加轻量化，改进后 Swin Transformer 模块的结构如图 3 所示。该模块通过移动窗口机制在保持计算效率的同时实现了全局信息的高效交互。此外，ST 模块利用 Transformer 的自注意力机制实现查询－键－值(QKV)的动态交互，使网络能够自适应地聚焦于不同位置的关键特征，从而捕捉丰富的上下文特征。

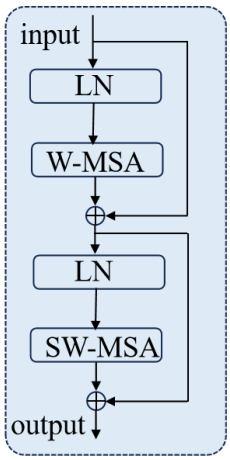

Figure 3. Improved ST module structure
图 3. 改进后的 ST 模块结构

改进后的 ST 模块的主要组成包括窗口多头自注意力机制(Windows Multi-head Self-Attention, W-MSA)和移动窗口多头自注意力机制(Shifted Windows Multi-head Self-Attention, SW-MSA)两部分。其中，W-MSA 的作用是将输入特征图划分为多个不重叠的局部窗口，在每个窗口内独立计算自注意力，通过限制注意力计算的范围，显著降低了自注意力的计算复杂度，从图像尺寸的平方复杂度($O(n^2)$)降至线性复杂度($O(n)$)。SW-MSA 既保持了 W-MSA 的计算效率，又增强了模型的全局建模能力，同时避免了传统滑动窗口方法的冗余计算。W-MSA 与 SW-MSA 的交替使用，使 Swin Transformer 能够以层次化的方式逐步融合局部和全局特征，显著提升了模型在视觉任务中的性能，ST 模型计算公式(2.1)和(2.2)所示。

$$z^{l} = \text{W - MSA}\left(\text{LN}\left(z^{l-1}\right)\right) + z^{l-1} \tag{2.1}$$

$$z^{l+1} = \text{SW - MSA}\left(\text{LN}\left(z^{l}\right)\right) + z^{l} \tag{2.2}$$

其中，z^{l-1} 表示输入特征，LN 表示层归一化，z^{l} 表示 W-MSA 的输出特征，z^{l+1} 表示 SW-MSA 的输出特征。

2.2.2. 颈部网络

本文提出的 Swin-YOLO 网络模型的颈部网络主要包括 Upsample 模块、C2f 模块和 SE-Conv 模块。SE-Conv 模块的结构与骨干网络中的 SE-Conv 模块完全一致。Upsample 模块主要用于特征图的上采样操作，将低分辨率特征图放大到更高分辨率，以便与浅层特征图进行多尺度融合。该模块不包含可学习参数，计算高效，默认采用 2 倍上采样，直接复制邻近像素值来扩展特征图尺寸。上采样后的特征图通常会与浅层高分辨率特征图拼接或相加，再经过后续卷积层(如 C2f 模块)进一步融合，从而增强模型对小目标的检测能力。

2.2.3. 头部网络

改进后的头部网络采用多尺度卷积模块(MCM)替换了原有的两层 3 × 3 卷积结构，实现了更高效的多尺度特征提取，其结构如图 4 所示。其中分类分支输出每个特征点的类别概率，回归分支预测目标框的坐标偏移和尺寸。为了提高小目标检测能力，头部采用逐像素预测形式，并通过分布式焦点损失优化边界框回归精度，同时使用交叉熵损失处理分类任务。改进后的头部网络采用多尺度卷积模块(MCM)替换了原有的两层 3 × 3 卷积结构，该模块包含三个并行分支：左侧分支采用 3 × 3 深度卷积进行通道内特征提取，这种设计大幅降低了参数量和计算复杂度；中间分支结合了 3 × 3 深度卷积与空洞卷积(空洞率 r = 2)，使该层卷积具有 5 × 5 卷积核等效感受野的同时显著减少了参数规模；右侧分支则通过 7 × 7 深度卷积捕获更大范围的上下文信息。三个分支提取的特征经过拼接后，再通过 1 × 1 卷积进行跨通道特征融合，有效整合了不同尺度的特征信息。这种结构不仅实现了模型的轻量化设计，还通过多尺度特征融合增强了网络的特征表达能力，在保证检测精度的同时显著提升了计算效率。

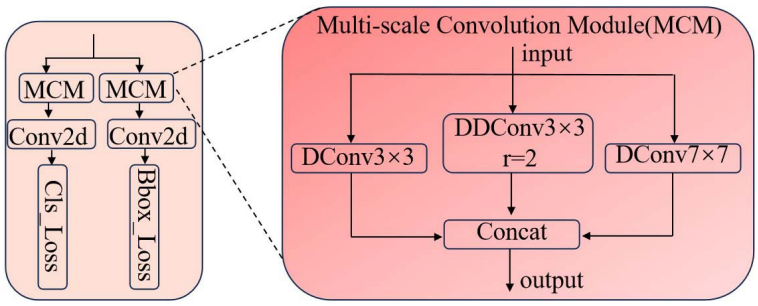

Figure 4. Improved head network structure
图 4. 改进后的头部网络结构

3. 实验及分析

3.1. 数据集与评价指标

为了验证 Swin-YOLO 在行人识别任务中的性能，本文基于两个公开的数据集即 TinyPerson [14]、CrowdHuman [15]进行实验，数据集详细信息如表 1 所示。

Table 1. Data set details
表 1. 数据集详细信息

数据集	训练集样本量	验证集样本量	测试集样本量	针对场景
TinyPerson	794	/	816	小目标检测
CrowdHuman	15,000	4730	5000	密集行人场景

为了验证模型的性能，本文采用精确率(P)、召回率(R)和平均精度均值(mAP@0.5， mAP@0.5:0.95)作为评价指标，其结果越大表示模型新能越好，但通常需要结合不同评价指标才能对模型性能做出全面的评价。精确率、召回率和平均精度均值计算公式如(3.1)、(3.2)、(3.3)和(3.4)所示。

$$P = \frac{TP}{TP + FP} \tag{3.1}$$

$$R = \frac{TP}{TP + FN} \tag{3.2}$$

$$AP = \int_0^1 P(R)\,dR \tag{3.3}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{3.4}$$

平均精度均值可进一步划分为 mAP@0.5 和 mAP@0.5:0.95，其主要区别如表 2 所示。

Table 2. Main differences between mAP@0.5 and mAP@0.5:0.95
表 2. mAP@0.5 和 mAP@0.5:0.95 主要区别

特点	mAP@0.5	mAP@0.5:0.95
IoU 阈值	固定单点(0.5)	多点平均(0.5~0.95，步长 0.05)
严格程度	宽松	严格
对定位误差敏感度	低(允许较大偏差)	高(需精细定位)
数值范围	通常较高	通常更低

3.2. 对比实验

为了验证本文提出的网络模型 Swin-YOLO 的性能，本节将 Swin-YOLO 与现有先进方法进行了对比。

首先，为了验证 Swin-YOLO 在小目标行人识别任务中的性能优势，在 TinyPerson 数据集上进行了对比实验。由表 3 实验结果可知，本文提出的模型 Swin-YOLO 相较于 YOLOv8s 模型，Swin-YOLO 的召回率(Recall)和 mAP@0.5 分别提升 7.4 个百分点和 6.5 个百分点，验证了其在小目标识别任务中的优越性。相比 YOLOv10s 和 YOLOv11s 等最新变体，Swin-YOLO 在所有评估指标上均保持领先优势。与当前最新方法 LMFI-YOLO [17]相比，Swin-YOLO 在精确率(Precision)上提升 0.4 个百分点，mAP@0.5 提高

0.4 个百分点，mAP@0.5:0.95 提升 0.2 个百分点，虽然召回率出现轻微下降，但模型的整体性能更优。

Table 3. Comparison experiment of TinyPerson data set
表 3. TinyPerson 数据集对比实验

模型	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv8s	36.2	17.0	15.6	5.4
YOLOv10s	37.8	20.6	18.9	6.7
YOLOv11s	37.0	18.8	16.5	5.9
ADE-YOLO [16]	/	24.3	18.6	6.1
LMFI-YOLO [17]	38.3	**24.6**	21.7	7.7
Swin-YOLO (ours)	**38.7**	24.4	**22.1**	**7.9**

Swin-YOLO 与 YOLOv8 在 TinyPerson 数据集上的实际检测效果对比如图 5 所示。从对比图中可以看出，两种模型在陆地人员(earth person)的目标检测中无显著差异，但在身体 50%以上浸入水中的海上人员(sea person)目标检测中，YOLOv8 完全无法检测该目标，而 Swin-YOLO 则能准确地检测出图片中央的半身浸入水中目标，检测性能明显增强。

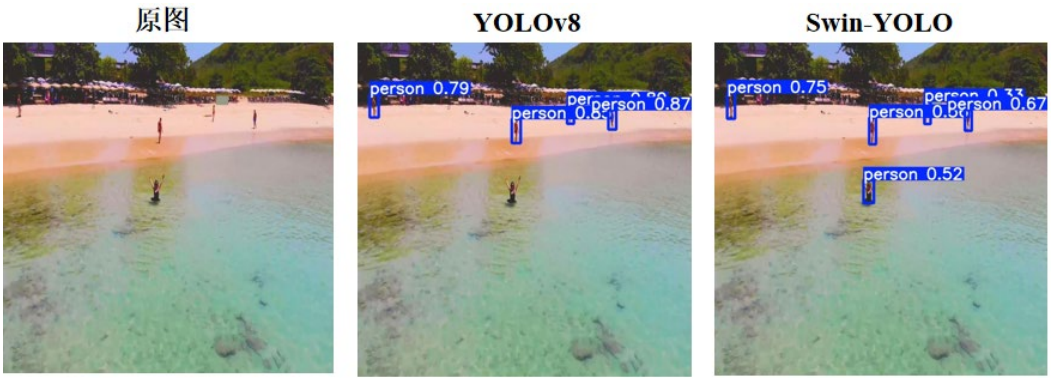

Figure 5. Detection comparison on TinyPerson data set
图 5. 在 TinyPerson 数据集上的检测对比

其次，为了验证 Swin-YOLO 模型在密集行人场景中的性能，在 CrowdHuman 数据集上进行了对比实验，其结果如表 4 所示。相较于基准模型 YOLOv8n，Swin-YOLO 取得了显著提升，其中 mAP@0.5 和 mAP@0.5:0.95 分别大幅提高了 8.6 个百分点和 10.9 个百分点。与当前最先进的文献[20]方法相比，Swin-YOLO 在精确率、召回率和平均精度均值等关键指标上均实现了不同程度的性能提升。虽然与文献[20]方法相比 mAP@0.5 指标出现轻微下降，但综合各项指标来看，Swin-YOLO 的整体性能仍然保持着一定的领先优势。

Table 4. Comparison experiment of CrowdHuman data set
表 4. CrowdHuman 数据集对比实验

模型	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv8n	81.6	63.1	75.0	42.4
YOLOv10n	83.6	69.4	81.5	50.6
YOLOv11n	85.1	69.7	81.6	51.0
YOLOv5-GB [18]	N/A	N/A	**84.8**	N/A

续表

MER-YOLO [19]	N/A	N/A	82.1	52.9
文献[20]	83.4	69.1	79.3	52.7
Swin-YOLO (ours)	**85.2**	**70.8**	83.6	**53.3**

Swin-YOLO 与 YOLOv8 在 CrowdHuman 数据集上的实际检测效果对比如图 6 所示。从图可以看出，在 CrowdHuman 数据集上，YOLOv8 仅能检测到前排及侧边的目标，对于远端的目标检测效果不佳，而 Swin-YOLO 在前者基础上检测到了队伍中间部分遮挡目标及右侧远端密集目标，并且检测的精度密度更高。以上这些实验结果充分证明了 Swin-YOLO 模型在密集人群场景下较强的鲁棒性和泛化能力。

Figure 6. Detection comparison on CrowdHuman data set
图 6. 在 CrowdHuman 数据集上的检测对比

3.3. 消融实验

为验证本文提出改进模块方法的有效性，并探索各模块对模型最终性能的影响程度，本小节进行了消融实验，分别移除了 ST 模块、使用标准卷积代替 SE-Conv 模块，使用标准的两层串联的卷积代替多尺度卷积模块(MCM)。由表 5 可知，SE-Conv 模块对模型性能具有关键性影响，当该模块被移除时，Swin-YOLO 在 TinyPerson 数据集上的检测性能出现显著下降，其中精确率分别降低 3.3 个百分点和 2.1 个百分点。ST 模块的移除导致模型在 CrowdHuman 数据集上性能下降，但在 TinyPerson 数据集上却出现精确率和召回率同时提升的反常现象。而 MCM 模块的移除则导致所有数据集上的性能全面下降，其中在 TinyPerson 数据集上的性能衰减最为显著，精确率和召回率分别衰减降 3.1 个百分点和 2.1 个百分点。

Table 5. Ablation experiments
表 5. 消融实验

数据集	SE-Conv	ST	MCM	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%
TinyPerson	×	√	√	35.4	22.8	20.9	6.3
	√	×	√	**38.8**	**24.6**	22.0	7.8
	√	√	×	35.6	22.3	21.0	6.4
	√	√	√	38.7	24.4	**22.1**	**7.9**
CrowdHuman	×	√	√	84.5	70.0	82.7	52.3
	√	×	√	84.7	70.5	82.2	52.1
	√	√	×	84.9	70.2	82.4	52.6
	√	√	√	**85.2**	**70.8**	**83.6**	**53.3**

4. 结论

本文提出了一种新的行人识别模型 Swin-YOLO，对模型的各个改进部分进行了详细的介绍，在两个常用的数据集上将 Swin-YOLO 与现有先进方法进行了比较实验，验证改进后的模型整体性能优势，最后通过消融实验验证各个模块对模型整体性能的影响。其中，在小目标场景数据集 TinyPerson 中，Swin-YOLO 模型相比 YOLOv8s 模型在检测精度指标 mAP@0.5 和 mAP@0.5:0.95 上分别提升了 6.5 和 2.5 个百分点；在密集目标场景数据集 CrowdHuman 中，Swin-YOLO 模型相比 YOLOv8n 模型在检测精度指标 mAP@0.5 和 mAP@0.5:0.95 上分别提升了 8.6 和 10.9 个百分点，有效减少了误检和漏检，为解决智能辅助驾驶及道路状况预警提供一种有效的改进 YOLOv8 行人识别模型。

参考文献

[1] 王晓路, 李晓婷, 谭永辉. 基于深度融合 3D 与 2D 卷积网络的步态识别方法[J]. 现代电子技术, 2025, 48(9): 109-115.

[2] 宋冬影. 基于深度学习的电子信息网络异常行为检测技术研究[J]. 信息记录材料, 2025, 26(5): 47-49+52.

[3] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., *et al*. (2016) SSD: Single Shot Multibox Detector. In: *Computer Vision-ECCV* 2016: 14*th European Conference*, Springer International Publishing, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2

[4] 徐彦威, 李军, 董元方, 等. YOLO 系列目标检测算法综述[J]. 计算机科学与探索, 2024, 18(9): 2221-2238.

[5] 宋宇博, 高嘉振. 改进 YOLOv3 算法的交通多目标检测方法[J]. 北京邮电大学学报, 2022, 45(5): 103-108.

[6] Fu, P., Zhang, X. and Yang, H. (2023) Answer Sheet Layout Analysis Based on YOLOv5s-DC and MSER. *The Visual Computer*, **40**, 6111-6122. https://doi.org/10.1007/s00371-023-03156-7

[7] Hu, J., Shen, L., Sun, G., *et al*. (2017) Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 2011-2023.

[8] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., *et al*. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. 2021 *IEEE*/*CVF International Conference on Computer Vision* (*ICCV*), Montreal, 10-17 October 2021, 10012-10022. https://doi.org/10.1109/iccv48922.2021.00986

[9] Chollet, F. (2017) Xception: Deep Learning with Depthwise Separable Convolutions. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 1251-1258. https://doi.org/10.1109/cvpr.2017.195

[10] Yu, F., Koltun, V. and Funkhouser, T. (2017) Dilated Residual Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, 21-26 July 2017, 472-480. https://doi.org/10.1109/cvpr.2017.75

[11] Zhu, X., Lyu, S., Wang, X. and Zhao, Q. (2021) TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. 2021 *IEEE*/*CVF International Conference on Computer Vision Workshops* (*ICCVW*), Montreal, 11-17 October 2021, 2778-2788. https://doi.org/10.1109/iccvw54120.2021.00312

[12] Varghese, R. and Sambath, M. (2024) YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. 2024 *International Conference on Advances in Data Engineering and Intelligent Computing Systems* (*ADICS*), Chennai, 18-19 April 2024, 1-6. https://doi.org/10.1109/adics58448.2024.10533619

[13] Li, J., Wu, W., Zhang, D., Fan, D., Jiang, J., Lu, Y., *et al*. (2023) Multi-Pedestrian Tracking Based on KC-YOLO Detection and Identity Validity Discrimination Module. *Applied Sciences*, **13**, Article No. 12228. https://doi.org/10.3390/app132212228

[14] Yu, X., Gong, Y., Jiang, N., Ye, Q. and Han, Z. (2020) Scale Match for Tiny Person Detection. *Proceedings of the IEEE*/*CVF Winter Conference on Applications of Computer Vision*, Snowmass, 1-5 March 2020, 1257-1265. https://doi.org/10.1109/wacv45572.2020.9093394

[15] Shao, S., Zhao, Z., Li, B., *et al*. (2018) Crowdhuman: A Benchmark for Detecting Human in a Crowd.

[16] 江旺玉, 王乐, 姚叶鹏, 等. 多尺度特征聚合扩散和边缘信息增强的小目标检测算法[J]. 计算机工程与应用, 2025, 61(7): 105-116.

[17] 袁婷婷, 赖惠成, 汤静雯, 等. LMFI-YOLO: 复杂场景下的轻量化行人检测算法[J]. 计算机工程与应用, 2025, 61(15): 111-123.

[18] 黄俊杰, 胡畅, 包嘉琪, 等. 轻量型密集行人检测算法研究[J]. 计算机仿真, 2024, 41(5): 183-188.

[19] 王泽宇, 徐慧英, 朱信忠, 等. 基于 YOLOv8 改进的密集行人检测算法: MER-YOLO [J]. 计算机工程与科学, 2024, 46(6): 1050-1062.

[20] 姚聪, 方逍, 郭星浩. 改进 YOLOv8 的轻量化密集行人检测方法[J]. 计算机工程与应用, 2025, 61(13): 138-150.