

基于YOLOv5的电梯行人遮挡检测

孟健伟, 王 谦

沈阳工业大学信息科学与工程学院, 辽宁 沈阳

收稿日期: 2025年3月11日; 录用日期: 2025年4月30日; 发布日期: 2025年5月14日

摘 要

本文基于YOLOv5对手扶电梯行人的遮挡进行检测, 针对电梯场景中行人遮挡导致的检测难的问题, 本文在YOLOv5算法的基础上进行了改进。首先, 为了解决因遮挡造成的预测框高度重合问题, 本文采用了软非极大值抑制(Soft NMS)算法替代传统的NMS算法, 该算法通过先对预测框的得分进行衰减处理, 然后再进行过滤, 从而减少因遮挡导致的人物漏检的情况。其次, 针对遮挡导致行人身体可见区域减小、易被忽略的问题, 本文在检测端增加了一个微小目标检测头, 专门用于检测因遮挡而只部分可见的行人目标, 从而提高了模型对这类目标的检测能力。实验结果表明, 通过采用以上两种的改进方法对YOLOv5算法进行改进, 在电梯行人遮挡检测中能够取得较好的效果。

关键词

自动扶梯, YOLOv5, 行人遮挡, Soft NMS

Elevator Pedestrian Occlusion Detection Based on YOLOv5

Jianwei Meng, Qian Wang

School of Information Science and Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: Mar. 11th, 2025; accepted: Apr. 30th, 2025; published: May 14th, 2025

Abstract

This article is based on YOLOv5 to detect pedestrian occlusion in escalators. To address the problem of difficulty in detecting pedestrian occlusion in elevator scenes, this article improves the YOLOv5 algorithm. Firstly, in order to solve the problem of high overlap of predicted boxes caused by occlusion, this paper adopts the Soft Non Maximum Suppression (Soft NMS) algorithm instead of the traditional NMS algorithm. This algorithm attenuates the scores of predicted boxes first, and then filters them to reduce the occurrence of missing persons caused by occlusion. Secondly, in response

to the problem of occlusion causing a reduction in the visible area of pedestrians and being easily overlooked, this paper adds a small object detection head at the detection end, specifically designed to detect pedestrian targets that are only partially visible due to occlusion, thereby improving the model's detection ability for such targets. The experimental results show that by using the above two improvement methods to improve the YOLOv5 algorithm, good results can be achieved in elevator pedestrian occlusion detection.

Keywords

Escalator, YOLOv5, Pedestrian Obstruction, Soft NMS

Copyright © 2025 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

行人检测一直是计算机视觉领域的关键研究方向,广泛应用于自动驾驶、视频监控、机器人、无人机等多个领域[1][2]。目前,主流的行人检测技术主要基于深度学习方法,这些算法可以分为双阶段(two-stage)和单阶段(one-stage)两类[3][4]。双阶段检测算法以 Faster-RCNN 为代表,而单阶段检测算法[5]-[7]则以 YOLO 系列为代表。手扶电梯作为现代公共交通系统中不可或缺的设施,广泛应用于地铁、商场、机场等人员密集的场所[8]。由于其能够有效解决垂直运输问题,手扶电梯在提供便捷服务的同时,也带来了人流管理的挑战。特别是在高密度人群聚集的环境下,电梯上行人的遮挡现象尤为突出,这不仅影响了电梯的正常使用,也为行人检测和行为分析带来了严峻的挑战。行人检测作为计算机视觉和人工智能领域的重要研究课题,已经在智能交通、安防监控等领域取得了显著进展[9]。然而,在手扶电梯这种动态环境中,行人往往因电梯的运动、相对位置的变化以及其他行人之间的遮挡而难以准确检测。现有的行人检测方法包括基于背景建模的遮挡检测方法和基于图像分割的遮挡检测方法,虽然这两种方法都能实现对遮挡行人的检测,但在面对复杂的遮挡和动态变化时,前者存在背景建模容易受到光照变化、阴影和动态背景的影响的问题后者存在对于边界模糊、物体形状复杂或小目标的遮挡效果不好,分割算法本身的准确度会影响最终结果的问题,且处理复杂场景时效果较差检测精度和实时性都面临着巨大的挑战。

2. YOLOv5 算法

YOLOv5 是一种高效的实时目标检测算法,由 Ultralytics 团队开发,该算法在保持 YOLO 系列快速推理优势的基础上,针对准确性和计算效率进行了多项优化[10]。YOLOv5 采用 CSPDarknet53 作为主干网络,通过引入 CSP 结构提高了特征提取的效率,并利用 PANet 优化了特征融合,进一步提升了对小物体和复杂场景的检测能力[11]。此外,YOLOv5 支持多种模型尺寸,可以根据计算资源需求和精度要求选择合适的模型。在目标检测任务中,YOLOv5 通过端到端训练框架、自动化的超参数调整和高效的数据增强策略,能够在多种应用场景中实现高精度和低延迟的目标检测。

3. 改进 YOLOv5 算法

3.1. NMS 算法改进

在目标检测的过程中,预测阶段往往会产生大量的候选检测框。为了精简这些检测框,会设定一个

置信度阈值来初步筛选, 但即便如此, 还是会有一些重叠或重复的框被保留下来。为了确保每个目标只对应一个最精确的检测框, 多数目标检测算法会采用非极大值抑制(NMS)技术。NMS 其实是一种寻找局部最优解的策略: 先挑选出置信度最高的检测框, 然后计算其他框与这个框之间的交并比(IOUS)。接下来, 根据预设的 IOU 阈值来决定是否保留这些框——如果 IOU 超过了这个阈值, 就认为它们是重复的, 于是予以剔除; 反之, 如果 IOU 低于阈值, 则予以保留。通过这样就能有效地去除冗余的检测框, 只保留最准确的那个。交并比 IOU 计算公式如公式(1)所示, S_1 和 S_2 为两个预测框, 分子为 S_1 和 S_2 的交集, 分母为 S_1 和 S_2 的并集。

$$IOU = \frac{S_1 \cap S_2}{S_1 \cup S_2} \quad (1)$$

但由于 NMS 算法并没有考虑到目标由于遮挡导致二者的预测框之间高度重合的问题, 进而剔除掉了一些本该保留的预测框导致行人漏检, 于是提出了 Soft NMS 算法来替代 NMS。Soft NMS 算法在处理预测框重叠问题时, 采用了一种更为柔和的策略来减少重复检测, 同时尽可能保留更多有用的预测信息。当某个预测框与置信度最高的预测框之间的 IOU 值超过预设的 IOU 阈值时, 传统 NMS 算法会直接丢弃这个预测框。但 Soft NMS 算法则不同, 它首先会根据 IOU 值的大小对该预测框的置信度进行衰减, IOU 值越大, 表示两个预测框重叠程度越高, 因此置信度衰减得越厉害。

完成置信度衰减后, Soft NMS 算法会进一步根据设定的置信度阈值来筛选预测框。那些置信度低于这个阈值的预测框会被过滤掉, 而置信度较高的预测框得以保留, Soft NMS 算法流程图如图 1 所示。

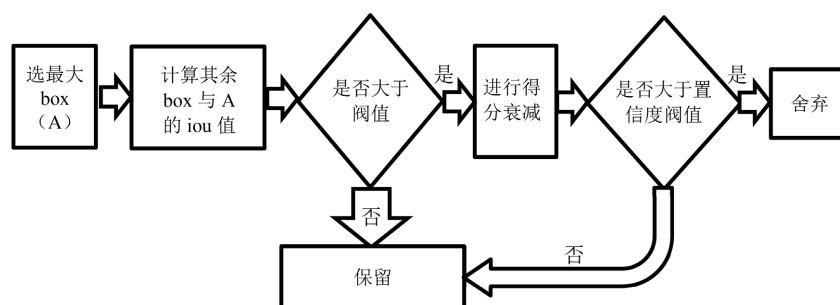


Figure 1. Soft NMS flowchart
图 1. Soft NMS 流程图

通过这种方式, Soft NMS 算法在降低重复检测的同时, 也能够一定程度上保留更多的预测框, 特别是那些可能因为遮挡而置信度不高但仍有价值的预测框。这样, 算法在检测遮挡目标时就能够减少漏检的数量, 提高检测的准确性和完整性。

Soft NMS 算法的得分衰减采用高斯加权形式的 Soft NMS 算法, 得分衰减公式如公式(2)所示。

$$S_i = S_i e^{-\frac{IOU(M, b_i)^2}{\sigma}}, \forall b_i \notin D \quad (2)$$

从公式可以看出, 高斯加权形式的 Soft NMS 算法具有连续性, 其中新增了一个超参数 σ , 用于控制置信度衰减程度, D 表示最优框集合。

3.2. 增加微小目标检测头

本文针对 YOLOv5 算法在检测微小尺度行人目标时容易出现的漏检问题, 提出了一种针对性的网络结构改进策略, 即微小目标检测头改进。这一改进的主要目的是通过增加一个专门用于检测微小目标的检测头, 来提升 YOLOv5 算法对微小尺度行人目标的检测能力。原始的 YOLOv5 算法包含三个检测头,

而引入微小目标检测头改进策略后算法的检测头数量增加至四个。新增的微小目标检测头特别设计为使用更大尺度的 160×160 特征图进行检测, 这样能够更好地捕捉到由于遮挡等原因导致的微小尺度行人目标的信息。

在改进后的网络架构中, 依旧保留了 YOLOv5 原有的 FPN 与 PAN 相结合的结构, 但针对颈部(Neck)和检测头部分进行了重要的调整和优化。

对于颈部部分, 采用了微小目标检测头改进策略。这一策略在 YOLOv5 的 FPN 结构上新增了一次上采样操作扩大特征图的尺寸, 使其包含更多细节信息。随后, 利用 Concat 操作将上采样后的特征图与主干网络 Backbone 中更浅层的特征进行融合, 这样做可以进一步增强特征图中深层特征的高级语义信息, 为后续的微小目标检测提供更有力的支持。

在 FPN 的输出端设计了两个分支, 第一个分支经过 C3 模块和卷积模块的处理后, 直接输出尺寸为 160×160 的特征图, 这个特征图专门用于微小尺度目标的检测。第二个分支则进入 PAN 结构, 将特征融合后的特征图继续向上传递, 这样做的目的是进一步增强深层特征图中包含的浅层特征强定位信息, 从而提高整体检测的准确性。最后, 在检测头部分增加了一个专门用于微小目标检测的检测头。这个检测头使用尺寸为 160×160 的特征图, 能够更精准地捕捉到微小尺度行人目标的信息, 从而有效减少漏检情况, 提升模型的检测性能, 改进的 YOLOv5 结构图如图 2 所示。

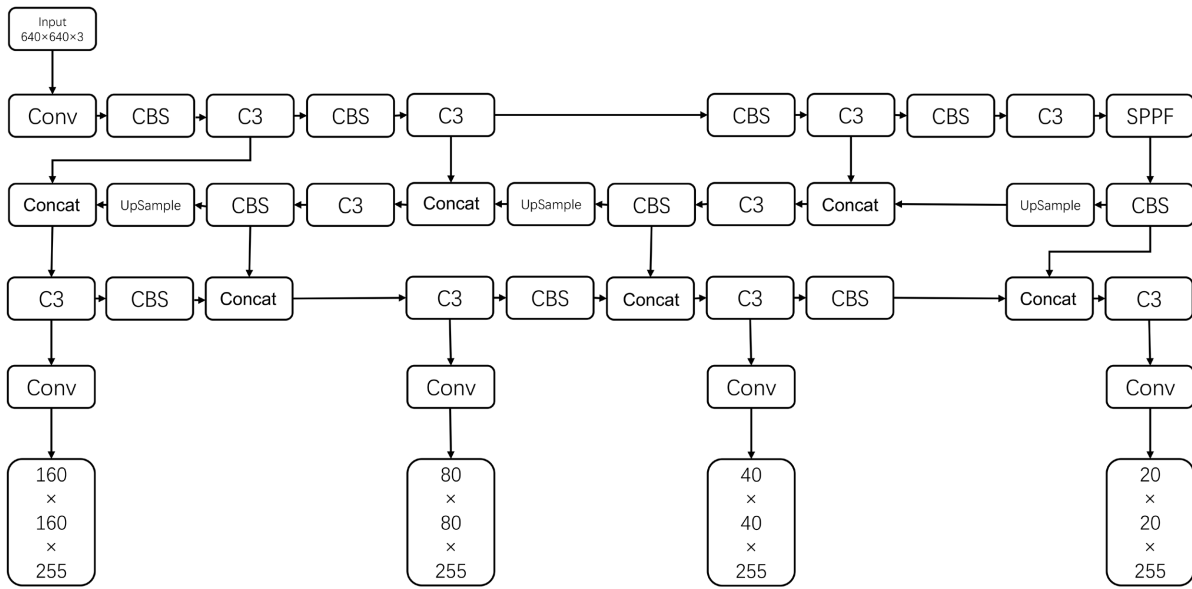


Figure 2. Improved YOLOv5 structure diagram
图 2. 改进的 YOLOv5 结构图

在引入微小目标检测头后, YOLOv5 算法的检测部分从原本三个检测头变为四个。因此, 先验框 Anchors 的设置也需要调整为四种不同尺寸的 Anchors。本文采用 K-means++ 聚类算法来初始化 Anchors, 以获得更适合该数据集的尺寸。通过对自定义数据集进行聚类, K-means++ 算法为微小目标检测头生成了三个适合的 Anchors 尺寸, 分别为 7×11 、 9×22 和 20×25 。

4. 实验结果与分析

4.1. 数据集

本研究使用了一个自定义的电梯行人数据集并命名为“Elevator-pedestrian”, 此数据集专门用于训练

和评估改进版 YOLOv5 模型。该数据集包含了 10,253 张电梯环境中的行人图像, 模拟电梯人流的多样性和复杂性。图像内容涵盖了不同角度、光照条件以及行人行为, 体现了电梯的实际使用情境。为了确保模型的泛化能力和评估的可靠性, 数据集被划分为训练集、验证集和测试集, 比例为 7:1:2。其中所有图像中训练集包含 7177 张图像, 验证集包含 1025 张图像, 测试集包含 2051 张图像。该数据集的设计用来推动行人检测算法在特定环境下的表现, 特别是在电梯这一特殊场景中的应用, 检测实验结果如图 3 所示, 其中图 3(a)和图 3(b)分别为不同检测结果。

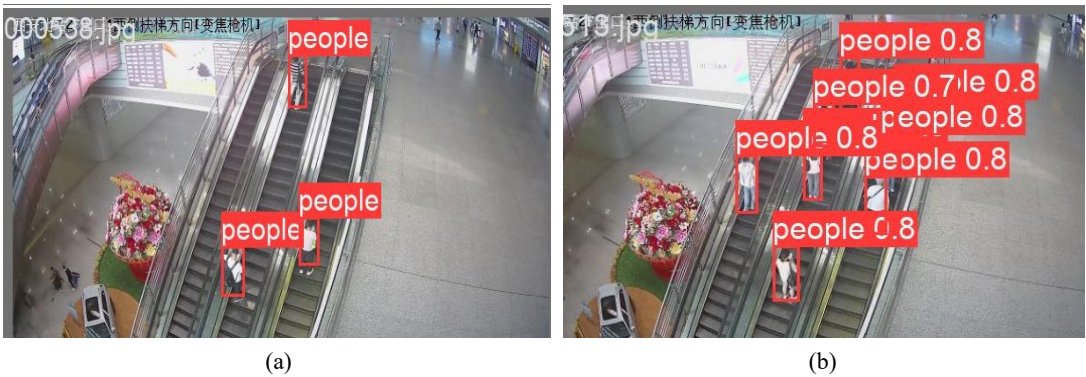


Figure 3. Test result chart
图 3. 检测结果图

4.2. 实验环境

实验环境配置见表 1, 实验程序采用 Python 3.8 语言编写, 深度学习框架采用 Pytorch 1.7.1。

Table 1. Experimental environment configuration
表 1. 实验环境配置

名称	配置
操作系统	Ubuntu 20.04
CPU	Intel(R) Core (TM) i7-12700H 2.30 GHz
内存	16 GB
GPU	NVIDIA GeForce RTX 3060, 10 GB
GPU 加速库	CUDA 11.0, CUDNN8.0.5

4.3. 训练结果

4.3.1. YOLOv5 + Soft NMS 训练结果

由于 Soft NMS 算法无需额外的训练, 只需在推理和验证阶段将传统 NMS 替换为 Soft NMS, 即可实现优化, 而不会增加额外的参数量和计算量。因此, 本节仅使用 Precision、Recall 和 mAP 来对模型性能进行分析, 训练结果图如图 4 所示。

从图中可以观察到, 三种评价指标都基本趋于稳定。加入 Soft NMS 算法后, 检测精度达到 91.4%, Recall 达到 82.7%, mAP 达到 45.6%, 各项指标与原 YOLOv5 相比, Precision 提高了 0.2 个百分点, Recall 提高了 2.4 个百分点, mAP 提高了 4.4 个百分点。其中, YOLOv5 + Soft NMS 算法在 mAP 指标上的提升最为明显。

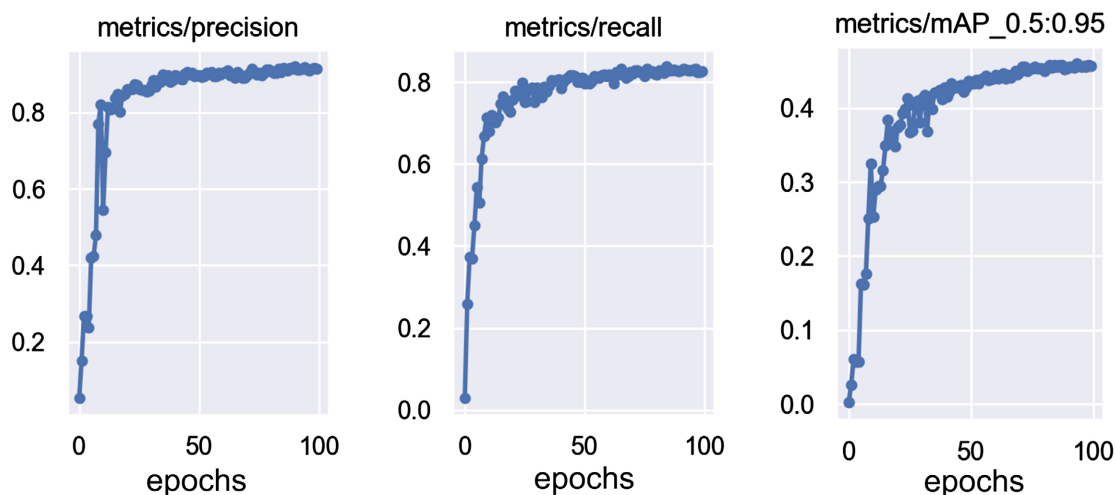


Figure 4. YOLOv5 + Soft NMS

图 4. YOLOv5 + Soft NMS

4.3.2. 微小目标检测头训练结果

由于微小目标检测头的改进增加了网络模型的参数和计算量, 因此本次实验选用了精度(Precision)、召回率(Recall)和平均精度(mAP)这三个指标进行综合分析, 并与 YOLOv5 算法进行了对比。从图 5 可以看出, 模型基本已达到收敛状态, 且在 YOLOv5 算法基础上加入微小目标检测头改进策略后, 检测效果明显优于原 YOLOv5 算法。

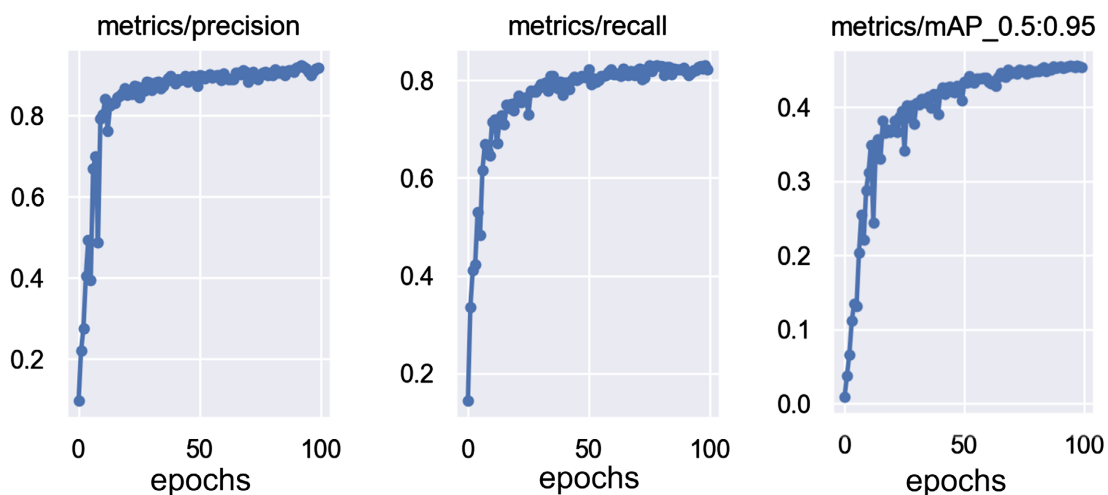


Figure 5. YOLOv5 + Small target detection head

图 5. YOLOv5 + 微小目标检测头

在 YOLOv5 算法中引入微小目标检测头之后, 模型在多个评估指标上取得了显著提升。其中, Precision 达到了 91.6%, 提高了 0.4 个百分点; Recall 为 82.1%, 增加了 1.8 个百分点; 而 mAP 则上升到了 45.3%, 提升幅度为 4.1 个百分点。这表明通过改进微小目标检测头, 模型对目标的检测能力有所增强。尽管该改进增加了网络模型的参数计算量, 但相比于原始的 YOLOv5 算法, 整体性能在精准度、召回率和平均精度等关键指标上都得到了明显的提升, 表明微小目标检测头在提高目标检测效果的同时, 仍保持了较高的性能。

4.4. 消融实验

提出的 ours-YOLOv5 算法, 将微小目标检测头和 Soft NMS 算法两种改进策略进行融合, 并从检测精度和检测速度两个方面评估其整体效果。同时, 通过消融实验分析了这两种改进策略对性能提升的贡献。由于 Soft NMS 算法仅在推理阶段替代原有的 NMS 算法, 且不需要修改网络结构, 因此 ours-YOLOv5 算法的网络结构与微小目标检测头改进策略中的网络结构保持一致, 结果如表 2 所示。

Table 2. Ours-YOLOv5 Ablation experiment
表 2. Ours-YOLOv5 消融实验

Model	Precision (%)	Recall (%)	mAP (%)
YOLOv5	91.2	80.3	41.2
YOLOv5 + Soft NMS	91.4	82.7	45.6
YOLOv5 + 微小目标检测头	91.6	82.1	45.3
ours-YOLOv5	93.2	84.4	47.8

综上所述, ours-YOLOv5 算法的综合检测性能优于其他三种算法, 表明本文提出的两种改进策略能够相互促进, 通过结合 YOLOv5 + Soft NMS 算法和增加 YOLOv5 微小目标检测头这两种改进, 从而实现了更好的检测效果。

5. 结语

本文提出了一种改进的 YOLOv5 算法, 旨在解决手扶电梯场景下行人检测中的遮挡问题, 尤其是由于遮挡导致可见区域较小而引发的漏检现象。通过将传统的 NMS 算法替换为 Soft NMS 算法, 并结合微小目标检测头, 我们有效地提升了模型对微小目标的检测能力, 同时优化了预测框的处理过程, 增强了模型在复杂环境下的检测精度。

实验结果表明, 所提出的 ours-YOLOv5 算法在自制手扶电梯数据集上取得了显著的性能提升。相比于原始 YOLOv5 算法, mAP、Recall 等指标均有明显提高。通过消融实验进一步验证了各项改进策略的有效性, 证明了 Soft NMS 算法在处理预测框时的优势, 以及微小目标检测头在提升检测精度方面的重要作用。

参考文献

[1] Boudjit, K. and Ramzan, N. (2021) Human Detection Based on Deep Learning YOLO-v2 for Real-Time UAV Applications. *Journal of Experimental & Theoretical Artificial Intelligence*, **34**, 527-544. <https://doi.org/10.1080/0952813x.2021.1907793>

[2] 高英, 吴玉虹. 基于 YOLOv5s 的轻量化行人检测算法[J]. 现代电子技术, 2023, 46(22): 151-158.

[3] 郭庆梅, 刘宁波, 王中训, 等. 基于深度学习的目标检测算法综述[J]. 探测与控制学报, 2023, 45(6): 10-20+26.

[4] Shao, X., Wei, J., Guo, D., Zheng, R., Nie, X., Wang, G., et al. (2021) Pedestrian Detection Algorithm Based on Improved Faster RCNN. 2021 *IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chongqing, 12-14 March 2021, 1368-1372. <https://doi.org/10.1109/iaeac50856.2021.9390882>

[5] Wang, C., Bochkovskiy, A. and Liao, H.M. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/cvpr52729.2023.00721>

[6] Chen, Q., Wang, Y., Yang, T., Zhang, X., Cheng, J. and Sun, J. (2021) You Only Look One-Level Feature. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13039-13048. <https://doi.org/10.1109/cvpr46437.2021.01284>

- [7] Adarsh, P., Rathi, P. and Kumar, M. (2020) YOLO V3-Tiny: Object Detection and Recognition Using One Stage Improved Model. 2020 *6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, 6-7 March 2020, 687-694. <https://doi.org/10.1109/icaccs48705.2020.9074315>
- [8] 黄光积. 手扶电梯乘客客流统计和异常行为识别算法研究[D]: [硕士学位论文]. 广州: 华南理工大学, 2023.
- [9] 张建军. 基于手扶电梯监控视频的危险行为检测及研究[D]: [硕士学位论文]. 合肥: 安徽大学, 2021.
- [10] 陈斌, 陈丽. 基于改进 YOLOv5s 算法的行人检测方法[J]. 智能物联技术, 2023, 6(3): 34-40.
- [11] Kannadasan, B. and Yogeswari, K. (2024) ANN-Based Traffic Planning Model Using Adaptive YOLOv5-Aided Object Counting from Surveillance Videos. *Asian Journal of Civil Engineering*, **25**, 4771-4788. <https://doi.org/10.1007/s42107-024-01079-9>