

本文引用格式: 林创鲁,叶亮,李刚,等.基于深度学习的自动扶梯乘客异常行为识别方法研究[J].自动化与信息工程,2022,43(6): 1-6.

LIN Chuanglu, YE Liang, LI Gang, et al. Research on identification method of escalator passengers' abnormal behavior based on deep learning [J]. Automation & Information Engineering, 2022,43(6):1-6.

基于深度学习的自动扶梯乘客异常行为识别方法研究*

林创鲁 叶亮 李刚 李丽宁

(广州特种机电设备检测研究院, 广东 广州 510180)

摘要: 自动扶梯乘客异常行为识别方法的研究对保障乘客安全具有重要的意义。针对自动扶梯出入口拥堵、长时间停留等乘客异常行为缺乏有效识别和预警手段的不足,提出一种基于深度学习的自动扶梯乘客异常行为识别方法。该方法采用 YOLOv4 算法对自动扶梯使用场景的视频进行特征提取,识别检测区域的乘客信息;结合 DeepSORT 算法对检测到的乘客进行追踪和统计,构建乘客异常行为识别模型,实现乘客异常行为的识别。对 4 段自动扶梯监控视频的实验结果表明,该方法检测平均准确率为 95.09%,能准确地识别自动扶梯出入口拥堵、长时间停留等乘客异常行为。

关键词: 深度学习; 目标检测; 目标跟踪; 异常行为识别; YOLOv4 算法; DeepSORT 算法

中图分类号: TH391.4

文献标识码: A

文章编号: 1674-2605(2022)06-0001-06

DOI: 10.3969/j.issn.1674-2605.2022.06.001

Research on Identification Method of Escalator Passengers' Abnormal Behavior Based on Deep Learning

LIN Chuanglu YE Liang LI Gang LI Lining

(Guangzhou Academy of Special Equipment Inspection & Testing, Guangzhou 510180, China)

Abstract: The research on the identification method of escalator passengers' abnormal behavior is of great significance to ensure the safety of escalators. Aiming at the shortage of effective identification and early warning methods for abnormal behaviors of escalator passengers, such as congestion and long stay at the entrance and exit of escalator, this paper proposes a method for identifying abnormal behaviors of escalator passengers based on deep learning. In this method, YOLOv4 algorithm is used to extract the features of the escalator scene video and identify the passenger information in the sensing area; Combined with DeepSORT algorithm, the detected passengers are tracked and counted, and the identification model of abnormal behavior is built to realize the identification of abnormal behavior of passengers. The experimental results of four escalator surveillance videos show that the average detection accuracy of this method is 95.09%, and it can accurately identify the abnormal behaviors of passengers such as congestion at the entrance and exit of the escalator and long stay.

Keywords: deep learning; object detection; target tracking; identification of abnormal behavior; YOLOv4 algorithm; DeepSORT algorithm

0 引言

随着我国城市化进程的不断深入,城镇人口也在快速增长。自动扶梯因在持续输送大量人流方面具有独特的优势,故在公共聚集场所得广泛使用。公共聚集场所的自动扶梯负荷高、持续工作时间长,乘客

异常行为(出入口拥堵、长时间停留等)状况较为普遍。因乘客异常行为而引发的摔倒、夹伤和踩踏等事故给安全管理带来巨大挑战^[1]。快速地识别这些乘客异常行为,是管控安全风险的重中之重。传统的自动扶梯使用管理方法多为使用单位在客流高峰时段,安

* 基金项目: 广州市市场监督管理局科技项目(2022kj18)

排工作人员在自动扶梯出入口值守,疏导客流,引导乘客安全乘梯,耗费了大量的人力和物力。随着深度学习技术的发展和运用,利用其进行自动扶梯乘客异常行为检测是发展趋势。

近年来,国内外学者对基于深度学习的自动扶梯乘客异常行为检测开展如下研究:田联房等^[2]通过对自动扶梯乘客人体骨架序列进行提取,识别摔倒、探头、探手等异常行为;杜启亮等^[3]采用 YOLOv3 检测自动扶梯乘客的位置,使用 MobileNetv2 结合反卷积的方法来提取乘客骨架,利用匈牙利匹配算法对乘客进行跟踪,识别乘客的异常行为;吉训生等^[4]提出一种改进的 Tiny YOLOv3 自动扶梯乘客异常行为检测网络结构,提高了乘客异常行为的检测速度;张鑫鑫等^[5]开发了基于 OpenPose 算法的自动扶梯乘客异常行为智能监控系统,并验证了该系统对模拟的摔倒等异常行为检测的有效性;LI Y 等^[6]研究一种基于加权卷积自动编码器和 YOLO 网络的异常行为检测方法,

提高了其在 UCSD 数据集上的检测性能;AZHAR M I H 等^[7]提出基于 YOLOv3 和 DeepSORT 的人流跟踪方法,实现了人员实时检测和运动路径的有效跟踪;HOSTK 等^[8]运用 YOLOv3 目标检测器和 DeepSORT 算法来检测和跟踪手球运动员,并分析特定球员在场上的表现。然而现有的研究大多集中在乘客个体姿态的异常动作识别和人流跟踪,少有涉及自动扶梯使用场景的出入口拥堵、长时间停留的研究。为此,本文针对自动扶梯乘客使用场景,结合深度学习的目标检测算法和多目标跟踪算法,提出基于 YOLOv4 和 DeepSORT 的自动扶梯乘客异常行为识别方法,实现自动扶梯出入口拥堵、长时间停留状况的识别。

1 自动扶梯乘客异常行为识别方法

基于深度学习的自动扶梯乘客异常行为识别方法包括:1) 基于 YOLOv4 的实时自动扶梯乘客检测;2) 基于 DeepSORT 的乘客动态跟踪;3) 乘客异常行为识别,其框图如图 1 所示。

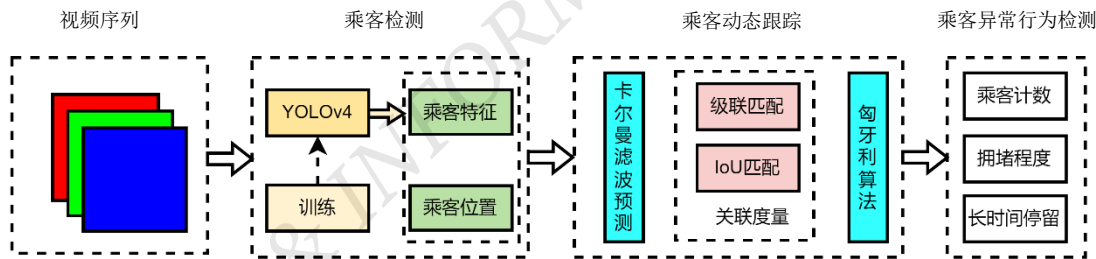


图 1 基于深度学习的自动扶梯异常行为识别方法框图

利用视觉传感器采集自动扶梯使用场景的视频作为乘客检测的输入;通过目标检测算法 YOLOv4 获取乘客特征,并对乘客进行定位;通过多目标跟踪算法 DeepSORT 动态跟踪检测区域内乘客的实时情况;利用乘客异常行为识别算法进行异常行为分类,输出识别结果。

2 乘客检测

乘客检测是乘客动态跟踪和异常行为识别的基础。考虑到后续边缘侧部署及实时性的需要,采用单

阶段目标检测算法 YOLOv4 进行乘客检测。YOLOv4 网络结构主要由 Input、Backbone、Neck、Prediction 等 4 部分组成^[9],如图 2 所示。

Input 对自动扶梯使用场景检测区域图像或训练数据集进行统一处理,形成尺寸相同的图像,作为 Backbone 的输入;Backbone 采用 CSPDarkNet53 网络对输入的图像进行特征提取;Neck 由空间金字塔池化(spatial pyramid pooling, SPP)与路径聚合网络(path aggregation network, PANet)组成,分别用于分离上下文特征和特征融合;Prediction 即输出端,对 Neck 处

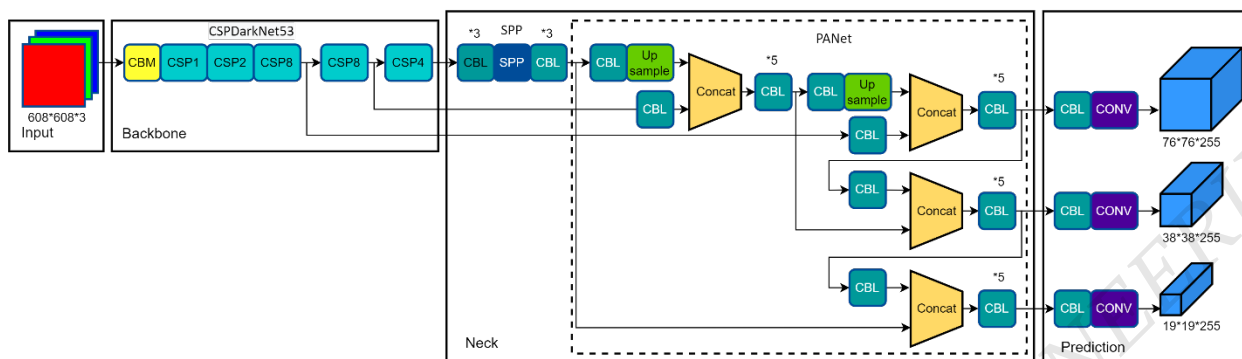


图2 YOLOv4 网络结构

3 乘客动态跟踪

本文采用多目标跟踪算法 DeepSORT^[10]进行乘

客动态跟踪。DeepSORT 算法利用卡尔曼滤波和匈牙利算法分别处理乘客动态跟踪的运动预测和数据关联部分, 算法流程如图 3 所示。

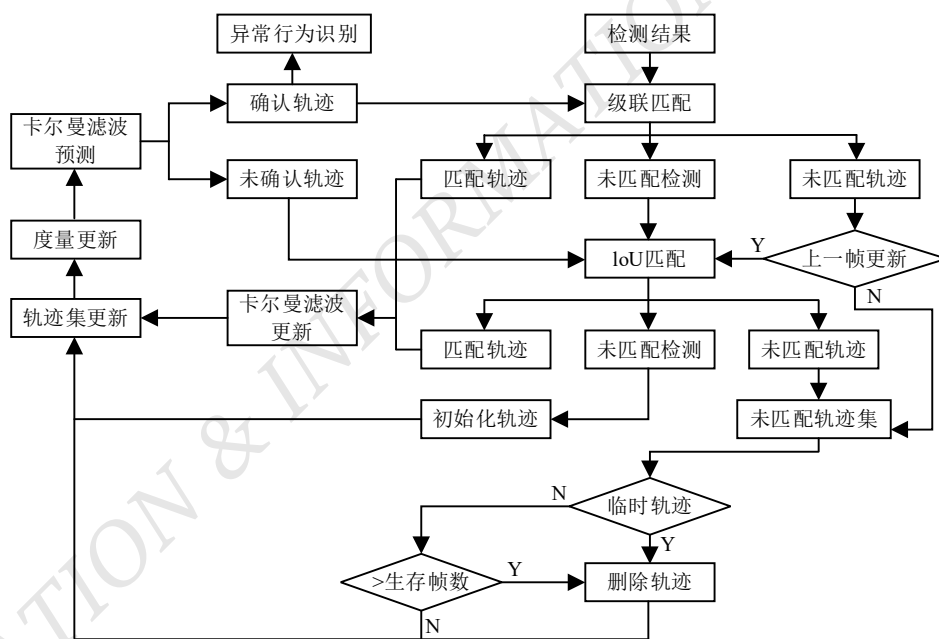


图3 基于 DeepSORT 的自动扶梯乘客动态跟踪算法

基于 DeepSORT 的自动扶梯乘客动态跟踪算法包括乘客状态预测、交并比匹配 (intersection over union, IoU)、级联匹配和跟踪器更新 4 个步骤。

3.1 乘客状态预测

为跟踪 YOLOv4 模型检测到的每一位乘客, 用 8 维状态向量 $\mathbf{x} = (u, v, \gamma, h, \dot{u}, \dot{v}, \dot{\gamma}, \dot{h})$ 标注乘客的运

动状态。其中, (u, v) 为乘客检测框的中心坐标; γ 为乘客检测框的长宽比; h 为乘客检测框的高度; $(\dot{u}, \dot{v}, \dot{\gamma}, \dot{h})$ 为乘客 (u, v, γ, h) 对应的参数在图像坐标系中的速度。

本文利用标准卡尔曼滤波进行乘客轨迹预测。对于每个乘客轨迹 k ，定义上次卡尔曼滤波预测框与检

测框成功匹配的帧图像与当前检测的帧图像之间的间隔帧数为 a_k 。在卡尔曼滤波预测期间 a_k 递增，当预测框与检测框成功匹配时， a_k 置0；当 a_k 超过最大阈值 $A_{\max}$ 时，认为乘客已离开场景，删除该轨迹；当出现与现有乘客轨迹无法匹配的检测框时，生成一条临时轨迹，如果临时轨迹在3帧中连续无法匹配，则删除。

3.2 交并比匹配

在自动扶梯使用场景的视频中，帧与帧之间同一乘客正确关联且编码一致，是识别乘客异常行为的基本前提。通常，利用卡尔曼滤波根据最近检测框信息对乘客轨迹进行预测，计算新检测框与轨迹预测框的IoU，采用匈牙利算法实现多个乘客ID的最优分配。IoU的计算公式为

$$IoU = \frac{S_C}{S_A + S_B - S_C} \quad (1)$$

式中：

IoU ——检测框与预测框的重叠率；
 S_A ——预测框的面积；
 S_B ——检测框的面积；
 S_C ——检测框与预测框的相交面积。

若在第 t 帧检测到 J 个乘客的检测框，卡尔曼滤波预测跟踪 I 个乘客的预测框，此时各个检测框与预测框之间的IoU为 s_{ij} ，并构成 $I \times J$ 大小的IoU度量矩阵。匈牙利算法的任务是寻找一种方案，使检测框与预测框匹配后的距离最小，即

$$\min S = \sum_{i=1}^I \sum_{j=1}^J x_{ij} s_{ij} \quad (2)$$

$$x_{ij} = \begin{cases} 1, & \text{第}j\text{检测框与第}i\text{预测框匹配} \\ 0, & \text{其他} \end{cases} \quad (3)$$

由于IoU的计算仅根据当前帧和之前帧的信息对乘客轨迹进行预测，而不需要一次性处理整个视频，可获得较好的实时性和准确性。但当被跟踪乘客受到遮挡或运动状态变化较大时，会产生多于实际的乘客ID，降低了识别准确度。

3.3 级联匹配

为解决乘客被遮挡或运动状态变化较大时，导致的乘客ID切换问题，引入运动信息和外观描述信息作为匹配度量。检测框 j 分配给轨迹 i 的代价函数 $c_{i,j}$ 为

$$c_{i,j} = \lambda d^{(1)}(i,j) + (1-\lambda) d^{(2)}(i,j) \quad (4)$$

式中：

λ ——空间度量 $d^{(1)}(i,j)$ 和外观度量 $d^{(2)}(i,j)$

的权重系数。

若 $c_{i,j}$ 处于阈值范围内，则认为第 j 个乘客检测框与第 i 个乘客轨迹正确关联。

在空间度量方面，利用马氏距离来度量卡尔曼滤波预测框与检测框的偏离程度，马氏距离计算公式为

$$d^{(1)}(i,j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (5)$$

式中：

$d^{(1)}(i,j)$ ——第 i 个乘客轨迹和第 j 个乘客检测

框的马氏距离；

d_j ——第 j 个乘客的检测框；

y_i ——第 i 个乘客轨迹的预测边界框；

S_i ——第 i 个乘客轨迹预测在当前检测空间的协

方差矩阵。

当马氏距离小于阈值时，则认为第 j 个乘客检测框与第 i 个乘客轨迹空间匹配成功。

在外观度量方面，利用卷积神经网络来提取乘客的外观特征，使用最小余弦距离来度量当前帧中各个乘客的特征向量与历史轨迹中所有乘客特征向量之间的外观相似度。最小余弦距离的计算公式为

$$d^{(2)}(i,j) = \min \{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in \mathcal{R}_i\} \quad (6)$$

式中：

$d^{(2)}(i,j)$ ——第 i 个乘客轨迹和第 j 个乘客检测

框的最小余弦距离；

r_j ——第 j 个乘客检测框 d_j 的外观描述符；

$r_k^{(i)}$ ——第 i 个乘客轨迹的外观描述符，保存最近 G_k 条匹配成功的外观描述符；

$\mathcal{R}_i$ ——每条确认乘客轨迹的外观描述符集合。

当 $d^{(2)}(i, j)$ 小于卷积神经网络训练的阈值时, 判定第 j 个乘客检测框与第 i 个乘客轨迹外观匹配成功。

3.4 跟踪器更新

乘客数据关联匹配后, 需要对跟踪器进行更新, 以便持续跟踪乘客。跟踪器更新主要包括:

1) 对于级联匹配成功的跟踪器, 被检测的乘客继承与其匹配成功的跟踪器编码, 并利用匹配成功的边界框信息预测下一帧的乘客位置;

2) 对于级联匹配未成功的跟踪器, 进行 IoU, 若匹配成功, 则继承跟踪器编码; 若匹配不成功, 考虑检测器漏检或短时遮挡的情况, 如跟踪轨迹的标记为不确定, 则删除轨迹; 如跟踪轨迹的标记为确定, 则为其设置生存帧数, 若在生存帧数期间仍匹配失败, 则移除轨迹;

3) 对级联匹配未成功的乘客, 进行 IoU, IoU 未成功的乘客, 为其建立一个新的跟踪器并分配编码, 标记为临时轨迹, 如接下来 3 帧连续匹配成功, 则标记为确定轨迹。

4 乘客异常行为识别

自动扶梯使用场景的视频经 YOLOv4 目标检测后获得乘客框信息; 新的或匹配的乘客轨迹经卡尔曼滤波预测获得乘客状态估计; 通过级联匹配和 IoU 实现数据关联匹配, 获得乘客轨迹, 作为异常行为识别算法的输入。

设第 t 帧自动扶梯使用场景的视频图像中跟踪到 I 名乘客, 其集合为 $P_t = \{H_t^1, H_t^2, \dots, H_t^i, \dots, H_t^I\}$, 其中每个乘客信息为 $H_t^i = (d, x_{\min}^{(t)}, y_{\min}^{(t)}, x_{\max}^{(t)}, y_{\max}^{(t)})$ 。以矩形感兴趣区域 (region of interest, ROI) 为例, 设 ROI 的坐标范围为 $R = (X_{\min}, Y_{\min}, X_{\max}, Y_{\max})$, 则 ROI 内的乘客满足:

$$\begin{cases} X_{\min} \leq \frac{x_{\min}^{(t)} + x_{\max}^{(t)}}{2} \leq X_{\max} \\ Y_{\min} \leq \frac{y_{\min}^{(t)} + y_{\max}^{(t)}}{2} \leq Y_{\max} \end{cases} \quad (7)$$

判断在 ROI 内的乘客数量为 J , 乘客集合为 $A_t = \{K_t^1, K_t^2, \dots, K_t^j, \dots, K_t^J\}$, 其中 $K_t^j = (d, T_{\text{total}}^{(t)})$ 为第 t 帧检测到在 ROI 内的乘客监测信息, $T_{\text{total}}^{(t)}$ 为 ID = $K_t^j \cdot d$ 的乘客截至图像第 t 帧时进入 ROI 区域的总时间, 计算公式为

$$T_{\text{total}}^{(t)} = T_{\text{total}}^{(t-1)} + t_{\text{FRAM}} \quad (8)$$

式中:

t_{FRAM} ——第 t 帧图像的处理时间。

当 $K_t^j \cdot T_{\text{total}}^{(t)} \geq T_{\text{max}}$ 时, 则判断为 ID = $K_t^j \cdot d$ 的乘客在 ROI 内长时间停留。

考虑到乘客的拥堵程度与短时的乘客数量、持续时间密切相关, 引入权重系数 λ_1 和 λ_2 来度量乘客拥堵程度, 计算函数为

$$f(t) = \lambda_1 \cdot (\text{set}(\{A_{t-TP}, A_{t-(TP-1)}, A_{t-(TP-2)}, \dots, A_{t-2}, A_{t-1}, A_t\})) + \lambda_2 \cdot (\text{set}(A_t)) \quad (9)$$

式中:

$\text{set}(A_t)$ ——第 t 帧的 ROI 内乘客的并集元素

数量;

TP ——预设的跟踪周期;

$$\text{set}(\{A_{t-TP}, A_{t-(TP-1)}, A_{t-(TP-2)}, \dots, A_{t-2}, A_{t-1}, A_t\})$$

——第 $t-TP$ 帧到第 t 帧的 ROI 内乘客并集元素数量。

对 $f(t)$ 设定正常、较拥堵、拥堵、长时间停留等阈值, 即可识别自动扶梯乘客拥堵程度。当自动扶梯乘客拥堵程度加剧或有乘客长时间停留时, 可输出预警信息进行警示, 以便使用单位及时采取安全措施, 化解安全风险。

5 实验

5.1 实验环境

本实验所用的数据集源自某市地铁自动扶梯使用场景的视频数据。以 Tensorflow 为软件框架, 模型训练、测试硬件环境为 AMD Ryzen 53600 型 CPU, 内存为 24 GB; NVIDIA RTX 3080 型 GPU, 显存为

12 GB；操作系统为 Windows 10。

5.2 乘客异常行为识别

本实验利用 YOLOv4 模型对自动扶梯使用场景视频中的乘客进行检测，将检测结果作为 DeepSORT 跟踪算法的输入。通过对检测结果与跟踪器的关联与匹配，实现目标区域内的乘客动态跟踪；通过对乘客动态跟踪结果的行为识别，实现自动扶梯乘客异常行为的识别。选取正常、较拥堵、拥堵、长时间停留等典型的自动扶梯使用场景，开展异常行为识别方法的实验验证，实验结果如表 1 所示。

表 1 自动扶梯典型使用场景的异常行为识别

使用场景描述	实际乘客数/人	检测结果/人	准确率/%	识别结果
客流稀疏	31	33	93.97	正常
短时遮挡	62	64	96.88	较拥堵
遮挡较多	105	101	96.19	拥堵
出口滞留	15	14	93.33	长时间停留

由表 1 可以看出，自动扶梯乘客异常行为识别方法对 4 种典型使用场景都有较好的检测效果，平均准确率达到 95.09%，可较好地适应乘客被遮挡的情况，并有效判别 4 种典型使用场景及异常行为。

6 结语

本文提出基于深度学习的自动扶梯乘客异常行为识别方法，利用 YOLOv4 模型进行乘客检测；结合 DeepSORT 算法对目标区域内的乘客进行动态跟踪；利用乘客异常行为识别算法实现自动扶梯使用状况和乘客异常行为的识别。实验结果表明，本文的识别方法对自动扶梯典型的使用场景有较好的检测效果，平均准确率达到 95.09%，可识别正常、较拥堵、拥堵和长时间停留等 4 种使用场景。该方法为公共聚集场所自动扶梯的使用安全管理提供一种新的技术手段，

可有效支持自动扶梯精准化管理，减轻自动扶梯安全管理的压力。下一步将考虑对识别算法进行轻量化设计，开展边缘侧部署，以实现自动扶梯的实时监测和预警。

参考文献

- [1] 卢文刚,黄小珍.城市地铁踩踏事件应急管理研究——基于 2008-2017 年典型案例的分析[J].中国应急救援,2017(4):4-9.
- [2] 田联房,吴啟超,杜启亮,等.基于人体骨架序列的手扶电梯乘客异常行为识别[J].华南理工大学学报(自然科学版),2019,47(4):10-19.
- [3] 杜启亮,黄理广,田联房,等.基于视频监控的手扶电梯乘客异常行为识别[J].华南理工大学学报(自然科学版),2020,48(8):10-21.
- [4] 吉训生,滕彬.基于深度神经网络的扶梯异常行为检测[J].激光与光电子学进展,2020,57(6):140-149.
- [5] 张鑫鑫,杨超宇,杨成.基于 OpenPose 的自动扶梯智能监控系统应用研究[J].黑龙江工业学院学报(综合版),2020,20(10):49-55.
- [6] LI Y, DAI Z. Abnormal behavior detection in crowd scene using YOLO and Conv-AE[C]//2021 33rd Chinese Control and Decision Conference (CCDC), 2021:1720-1725.
- [7] AZHAR M I H, ZAMAN F H K, TAHIR N M, et al. People tracking system using DeepSORT[C]//2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2020:137-141.
- [8] HOST K, IVASIC-KOS M, POBAR M. Tracking handball players with the DeepSORT algorithm[C]. 9th International Conference on Pattern Recognition Applications and Methods (ICPRAM), 2020:593-599.
- [9] 邵延华,张铎,楚红雨,等.基于深度学习的 YOLO 目标检测综述[J].电子与信息学报,2022,44(10):3697-3708.
- [10] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//IEEE International Conference on Image Processing (ICIP), 2017:3645-3649.

作者简介：

林创鲁，男，1983 年生，硕士，高级工程师，主要研究方向：特种设备智能监测与预警。E-mail: linter0663@163.com

叶亮，男，1983 年生，硕士，高级工程师，主要研究方向：特种设备智能监测与预警。

李刚，男，1988 年生，硕士，工程师，主要研究方向：特种设备风险监测与预警。

李丽宁，女，1988 年生，本科，助理工程师，主要研究方向：特种设备智能监测与检测。