

DOI:10.19344/j.cnki.issn1671-5276.2024.04.042

# 基于改进 YOLOv8 的电梯内电动车识别方法研究

路成龙,冯月贵,庆光蔚

(南京市特种设备安全监督检验研究院,江苏南京 210002)

**摘要:**针对电梯内电动车识别存在效率低下、精度不佳的问题,提出一种结合 AUGMIX 图像增强技术和改进 YOLOv8 模型的电动车识别方法。将变形卷积层和动态稀疏注意力机制融入 YOLOv8,识别更精确和高效。实验结果表明:改进后算法模型的精确率、召回率和平均精度均值分别达到了 94.5%、93%和 82.4%,电动车识别准确率达到 95.8%,为电梯内电动车智能识别提供了理论基础。

**关键词:**电动车识别;AUGMIX;YOLOv8;变形卷积层;动态稀疏注意力机制

**中图分类号:**TP183 **文献标志码:**A **文章编号:**1671-5276(2024)04-0219-05

## Research on Identification Method of Electric Vehicles in Elevators Based on Improved YOLOv8

LU Chenglong, FENG Yuegui, QING Guangwei

(Nanjing Special Equipment Safety Supervision Inspection and Research Institute, Nanjing 210002, China)

**Abstract:** A new electric vehicle identification method combining AUGMIX and improved YOLOv8 model is proposed to address the issues of low efficiency and poor accuracy in identifying electric vehicles in elevators. The YOLOv8 model incorporates DCNv3 and BRA to identify electric vehicles with better accuracy and efficiency. The experimental results show that the precision, recall, and mean average precision of the improved algorithm model reach 94.5%, 93%, and 82.4% respectively. And the accuracy of electric vehicle identification reaches 95.8%, providing a theoretical basis for intelligent recognition of electric vehicles in elevators.

**Keywords:** machine vision; AUGMIX; YOLOv8; DCNv3; Bi-level routing attention

### 0 引言

随着我国城市化的快速推进,商场、居民楼及城市轨道交通等场所都安装了电梯,在带来了极大乘坐便利的同时,也存在着重大安全隐患。电动车进入电梯轿厢,不仅占用了乘客乘用空间,而且一旦发生自燃会产生大量的火焰和高温有毒气体,短时间内充满整个轿厢,造成人员疏散和逃生困难。此外,电动车进出轿厢如果操作不当,很容易使层轿门碰撞脱轨或者轿厢撞击变形,造成安全隐患且缩短了电梯的使用寿命<sup>[1]</sup>。因此,电动车驶入电梯轿厢存在很大的安全隐患。

目前电梯内电动车识别手段主要采用 3 种方式:人工监控识别、物理阻车和视觉识别<sup>[2]</sup>。人工监控识别是物业人员通过摄像头监控轿厢内乘客情况,当发现有电动车进入轿厢时通过语音提醒电动车退出轿厢。物理阻车是在轿厢入口和轿厢内壁设置阻拦栏杆。视觉识别是通过轿厢内摄像

头采集、分析视频数据,通过计算机视觉技术进行识别。人工监控识别对人力和物力消耗较大,需要相关人员不间断监控,效率低下且无法及时制止电动车驶入轿厢的行为。物理阻拦方式在阻拦电动车的同时也限制了轮椅、儿童车等工具的驶入,在客流量较大的场所容易产生人员绊倒、摔倒和踩踏事故。视觉识别是基于机器视觉系统的电动车识别系统,根据摄像头数据信息进行识别,采集电动车图像信息,进行算法匹配来实现电动车识别<sup>[3]</sup>,具有精度高、响应迅速的优点,具有广阔的应用前景。

### 1 整体框架设计

基于 AUGMIX 和改进 YOLOv8 的电梯内电动车识别方法,通过在轿厢内安装或共用现有监控摄像头采集轿厢内行人、电动车、自行车等图像信息,进行算法匹配来实现电动车识别。通过 AUGMIX 技术为电梯轿厢中的电动车识别引入增

**基金项目:**国家市场监督管理总局科技计划项目(2022MK156);江苏省市场监督管理局科技计划项目(KJ2023039)

**第一作者简介:**路成龙(1990—),男,山东齐河人,工程师,硕士,从事特种设备检验和研究工作,luchenglong88@126.com。

强的图像数据集,提高了模型的鲁棒性,并通过适应不同情况的变化,显著降低了过拟合的可能<sup>[4]</sup>。YOLOv8 算法的端到端识别方法与 DCNv3 的变形卷积层和动态稀疏注意力机制相结合,优化了对电梯轿厢内电动车形状和结构的适应性,为轿厢式电梯电动车识别提供了实时、精准且计算效率高的识别解决方案。整体框架设计图如图 1 所示。

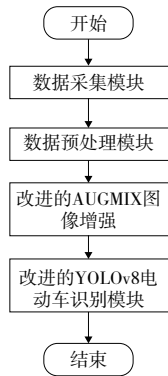


图 1 整体框架设计图

## 2 算法设计

### 2.1 AUGMIX 图像增强算法

AUGMIX 结合多种不同的图像增强技术来生成新的训练样本,提高模型在面对自然界图像分布中共同变形时的泛化能力,从而增加模型训练数据的多样性。在 AUGMIX 中,模型不仅在原始图像上训练,而且还要在增强后的图像上进行训练。这些增强后的图像是通过原始图像应用一系列随机选择的图像处理操作(称为“增强操作”)生成的。然后,这些增强操作生成的图像以一定方式与原始图像混合,形成最终的训练样本。

AUGMIX 的关键步骤可以概括为以下几点。

1) 选择原始图像  $x$ , 进行  $k$  个增强操作, 每个增强操作表示为  $o_{pk}$ 。这些操作是针对图像的特性变化, 例如旋转、剪切、颜色变换等, 本技术中仅对电动车识别率进行改进。

2) 将这些操作应用于原始图像  $x$ , 生成一组增强图像  $\{x_1, x_2, \dots, x_k\}$ 。每个增强图像  $x_i$  是通过将相应的增强操作  $o_{pi}$  应用于原始图像  $x$  得到的, 即  $x_i = o_{pi}(x)$ 。

3) 将这些增强图像线性混合, 生成最终的训练样本。这个混合过程可以表示为以下公式:

$$x' = w_0 \times x + \sum_{i=1}^k w_i \times x_i \quad (1)$$

式中  $(w_0, w_1, \dots, w_k)$  是混合权重, 它们是随机抽取的, 满足  $\sum_{i=0}^k w_i = 1$ 。

混合后的图像  $x'$  将作为新的训练样本。通过这种方式, AUGMIX 可以在训练过程中引入大量的图像变化, 从而提高模型的泛化能力。

### 2.2 改进的 YOLOv8 算法介绍

YOLOv8 是一种高效的单步目标识别模型, 可以实现复杂的目标识别以及定位功能, 对图像框中的多个目标进行标注, 得到识别结果<sup>[5]</sup>。相较于两步式识别方法, YOLOv8 则通过神经网络一次性提取图像的特征图, 基于此直接推断锚框的位置与类别。相较于两步法, 单步识别流程的速度更快, 特别是在需要快速响应的实时应用中更显优势。

本文基于 YOLOv8 的改进模型针对电梯轿厢内电动车识别任务进行了针对性改进, 改进模型通过应用 AUGMIX 图像增强技术提升模型泛化能力, 采用端到端的检测方法简化系统设计并提高识别效率, 实施多尺度检测策略以增强对不同大小目标的鲁棒性。同时, 集成 DCNv3 算子优化了模型对目标形状的自适应能力, 而动态稀疏注意力机制则降低了计算复杂性, 使得改进后的 YOLOv8 不仅识别准确, 而且适应实时性和资源受限的应用环境, 为电梯轿厢中电动车的精确识别提供了有效的技术支持。改进的 YOLOv8 算法架构图如图 2 所示。

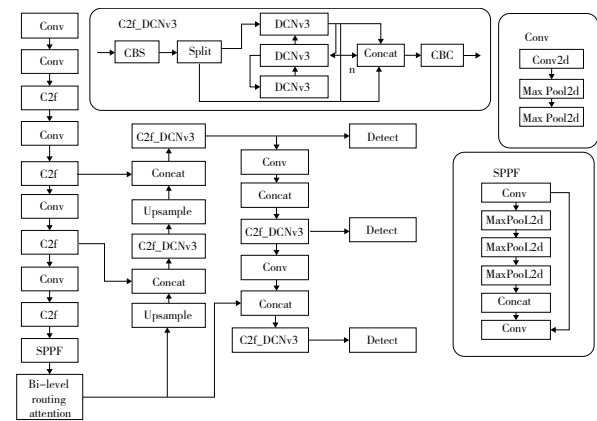


图 2 改进的 YOLOv8 算法架构图

1) 结合 C2f 和可变性卷积网络的 C2f\_DCNv3 模块

YOLOv8 相较于同系列的其他算法, 在主干网络结构中采用了 C2f (cross stagepartial network bottleneck with 2 convolutions) 模块, 该模块能够有

效提高网络计算速度同时降低计算复杂度。DCNs (deformable convolutional networks) 是卷积神经网络 (CNNs) 的扩展, 引入了可变形卷积层, 使得网络能够更好地适应图像中的几何和形状变化<sup>[6]</sup>。这种类型的网络特别适用于处理图像中存在的非刚性变形, 如姿态变化、视角变换或物体内部的变形等。采用最新的 DCNv3 对 C2f 模块进行扩展, 得到 C2f\_DCNv3 模块, 能够自适应地调整感受野大小, 从而更好地符合目标对象的形状和结构变化, C2f-DCNv3 模块网络结构如图 3 所示。

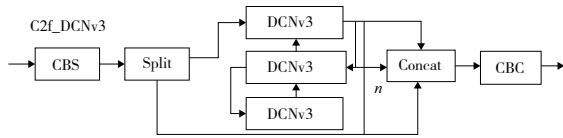


图 3 C2f\_DCNv3 模块网络结构

通过引入共享投射权重的策略, 将与位置无关的权重代替独立的投射权重, 从而有效减少参数和内存复杂度。此外, 采用了多组机制, 将空间聚合过程划分为不同组, 每组具有独立的采样偏移量, 以提高特征的多样性。为缓解模型容量扩大时的不稳定问题, 采用了逐采样点的 Softmax 归一化作为归一化模式, 这不仅提升大规模模型的训练稳定性, 还构建了所有采样点之间的连接关

系。这些改进旨在提高模型效率、降低参数量、增强特征的多样性, 并改善模型训练的鲁棒性, 其公式如下所示。

$$y(p_0) = \sum_{g=1}^G \sum_{k=1}^K w_g m_{gk} x_g(p_0 + p_k + \Delta p_{gk}) \quad (2)$$

式中:  $G$  为聚合组总数; 对于第  $g$  个组,  $w_g \in R^{C \times C'}$  表示与位置无关的投影权重, 其中  $C' = C/G$  表示组维度;  $m_{gk} \in R$  表示第  $g$  个组中第  $k$  个采样点的调制标量, 由沿  $k$  维的 Softmax 函数归一化;  $x_g \in R^{C' \times H \times W}$  表示切片后的输入特征图;  $\Delta p_{gk}$  为第  $g$  组网络采样位置  $p_k$  对应的偏移量。

## 2) 动态稀疏注意力

动态稀疏注意力机制作为一种新型注意力机制, 通过两个层次实现输入的动态稀疏化处理, 这样不仅能过滤掉粗糙区域大部分不相关交互信息, 加强有效信息间交互, 减弱无效信息间干扰, 而且针对细粒度的注意力应用于路由区域的联合之中, 能够获取更多有效特征信息。在大幅降低了模型的计算复杂度的同时保持了高水平的性能。相对于其他注意力机制, 动态稀疏注意力机制能够更灵活地调整注意力分布, 从而更适应不同尺度和复杂的目标识别, 在小目标捕捉中精更高。动态稀疏注意力机制作用过程如图 4 所示, 图中 gather 表示基于模型子集的收集/聚集操作, mm 表示矩阵乘法 (matrix multiplication)。

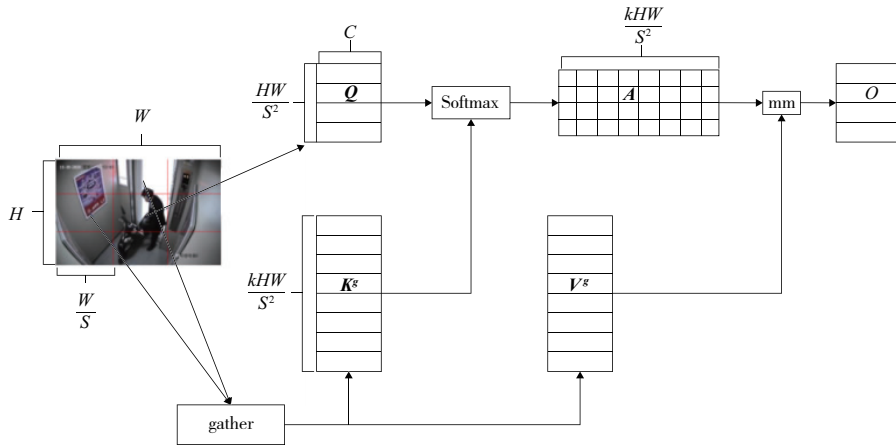


图 4 动态稀疏注意力机制作用过程

给定一个二维输入特征映  $X \in R^{H \times W \times C}$ , 首先将其划分为  $S \times S$  个非重叠区域, 使得每个区域包含  $\frac{HW}{S^2}$  个特征向量。这一步通过 reshape  $X$  变为  $X \in R^{S^2 \times \frac{HW}{S^2} \times C}$  来完成。再经过线性映射得到  $Q$ 、 $K$ 、 $V$  张量:

$$Q = X' W^q, K = X' W^k, V = X' W^v \quad (3)$$

然后, 通过有向图来对区域到区域的注意力关系进行关联。具体来说, 首先通过分别在  $Q$  和  $K$  上应用每个区域平均值来导出区域级  $Q^r$ 、 $K^r \in R^{S^2 \times C}$ 。其次, 通过  $Q^r$  和转置  $K^r$  之间的矩阵乘法推导出区域到区域亲和度图的邻接矩阵  $A^r \in R^{S^2 \times S^2}$ :

$$A^r = Q^r (K^r)^T \quad (4)$$



邻接矩阵  $A^r$  中的条目衡量两个区域在语义上的相关性。接下来执行的核心步骤是通过为每个区域只保留前  $k$  个关系最密切的区域。 $I^r \in N^{S^2 \times k}$  为一个路由索引矩阵,具有逐行 topk 算子:

$$I^r = \text{topkIndex}(A^r) \quad (5)$$

第  $i$  行  $I^r$  包含第  $i$  个区域的  $k$  个最相关区域的索引。

有了区域到区域路由索引矩阵  $I^r$ ,在粗粒度过滤了最不相关的令牌,可以应用细粒度的令牌到令牌关注。对于区域  $i$  中的每个  $Q$ ,它将关注以  $I^r_{(i,1)}, I^r_{(i,2)}, \dots, I^r_{(i,k)}$  为索引的  $k$  个最关注的路由区域,并 gather 这  $k$  个区域中所有  $K$  和  $V$ 。因此首先收集  $K^g, V^g$  张量:

$$K^g = \text{gather}(K, I^r), V^g = \text{gather}(V, I^r) \quad (6)$$

式中  $K^g, V^g \in R^{S^2 \times \frac{HW}{S^2} \times C}$ 。将注意力应用于收集到的

$K^g, V^g$ :

$$O = \text{Attention}(Q, K^g, V^g) + \text{LCE}(V) \quad (7)$$

引入了一个局部上下文增强项  $\text{LCE}(V)$ 。函数  $\text{LCE}(V)$  使用深度卷积进行参数化。

### 3 实验与分析

#### 3.1 数据集介绍

数据集在南京某小区进行,采集周期覆盖了春夏秋冬四个季节,以确保数据具有时间上的多样性。此外,考虑到电梯的使用高峰期和非高峰期,在工作日和周末的不同时间段进行了采集,以确保获得充分的数据集,反映出电梯内电动车出现的各种情况。然后,将收集到的图片利用 AUGMIX 图像增强,共得到 9 000 张高分辨率照片,部分数据集如图 5 所示。



图 5 电梯内电动车数据集

模型训练期间,模型的迭代曲线和准确性评估将会通过数据可视化工具进行监控,确保训练的有效进展。测试结果通过混淆矩阵和识别效果图来展示,从而对模型在实际应用中的表现给予直观的评价。图 6 为本文的数据长宽比分析图。

#### 3.2 模型测试评估与分析

AUGMIX - YOLOv8 - org 为本文提出的将 AUGMIX 与改进的 YOLOv8 相结合的预测模型。利用改进后的算法模型对数据集进行测试,将数据集按 8 : 1 : 1 的比例分为训练集、验证集和测试集进行实验。迭代次数 200,初始学习率设置为 0.001,优化器采用 Adam,动量参数值为 0.8,置信度阈值为 0.5,模型的训练结果及迭代曲线如图 7 所示。实验选取平均精度均值 (mAP)、准确率 (precision)、召回率 (recall) 及单项识别准确率作为评价指标,其照片中包括电动车、乘客、自行

车,试验测试结果混淆矩阵如图 8 所示。

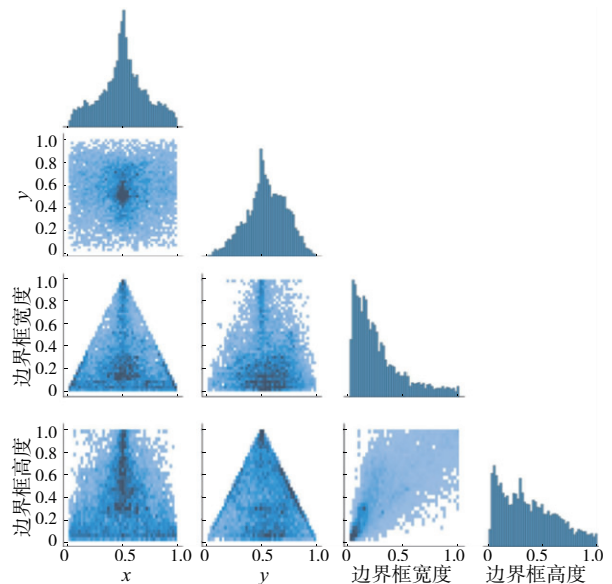


图 6 数据集长宽比分析图

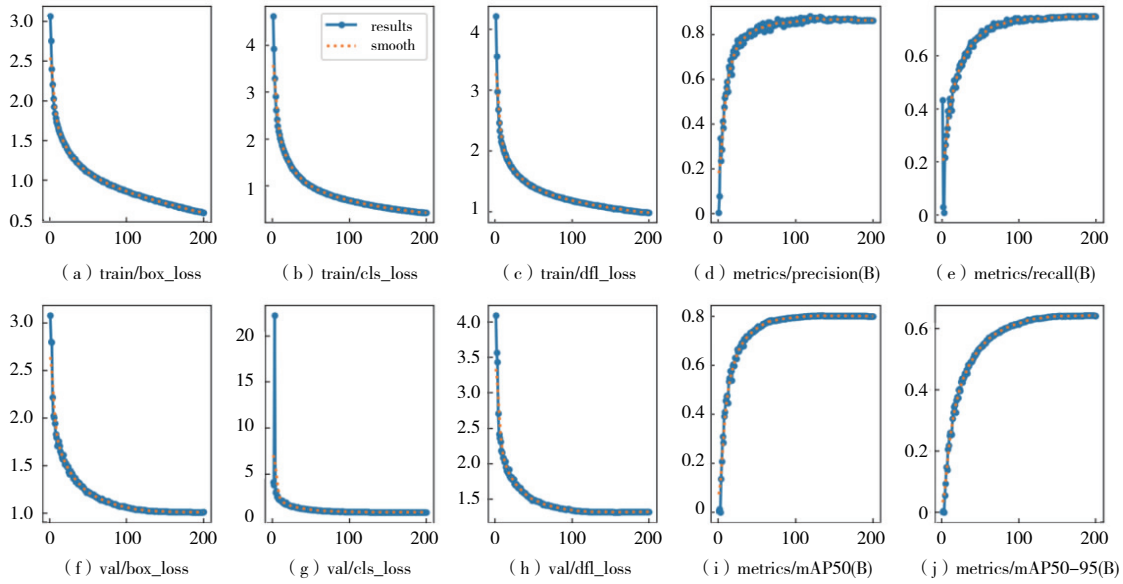


图 7 模型训练及验证集迭代曲线图

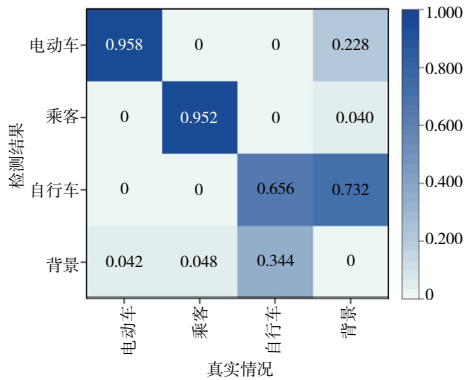


图 8 识别结果混淆矩阵

准确率进行对比,图 9 为实验各项评价指标对比图。由图 9 可知, AUGMIX-YOLOv8-org 模型准确率达到 94.5%,召回率达到 93%,平均精度均值达到 82.4%,相比原始的 YOLOv8 模型分别提升了 2.5、5.0 和 2.3 个百分点。在具体的识别任务中,电动车、乘客和自行车的识别准确率, AUGMIX-YOLOv8-org 分别实现了 95.8%、95.2% 和 65.6% 的高水平,与原始 YOLOv8 模型的 91.7%、91.2% 和 50.8% 相比,有显著的性能提升。这不仅凸显了 AUGMIX 技术在提升模型对复杂场景下目标识别能力的有效性,也说明了 AUGMIX-YOLOv8-org 模型在实际应用中对于确保电梯安全运行的潜在价值。新算法模型识别结果如图 10 所示。

对电动车、乘客和自行车的电梯轿厢识别准

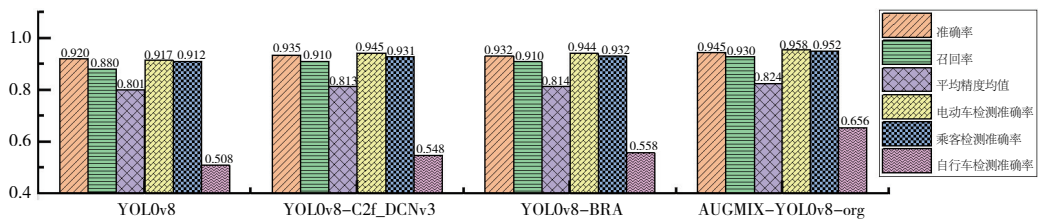


图 9 实验评价指标对比图



图 10 电动车识别效果图

(下转第 229 页)

现远程连接,支持学生在异地对设备进行编程和控制。该设备已在疫情期间支持学生进行线上实物实验,突破了传统在线仿真实验的局限。设备已申请国家发明专利,并在哈尔滨工业大学本科生课程“无人机系统工程应用”中进行了实践教学,获得了学生的好评。

### 参考文献:

- [1] 薛晶勇,杨波,朱永乐,等. 四旋翼不确定干扰前馈补偿与反步滑模姿态控制[J]. 计算机集成制造系统, 2022, 28(2): 345-354
- [2] 时来富,沈建新,王启盛,等. 四旋翼飞行器的模糊自抗扰姿态控制[J]. 机械制造与自动化, 2021, 50(3): 157-162.
- [3] 杨亚,李玉爽,徐杰. 基于自抗扰控制器的变桨距控制系统[J]. 辽东学院学报(自然科学版), 2022, 29(1): 36-41.
- [4] 张钊,杨忠,段雨潇,等. 主动变形四旋翼自抗扰飞行控制方法[J]. 控制理论与应用, 2021, 38(4): 444-456.
- [5] 何紫云,陶玫玲,陈强,等. 基于两相幂次趋近律的四旋翼飞行器自适应控制[J]. 高技术通讯, 2022, 32(3): 294-302.

- [6] 卢艳军,张晓东,纪鹏飞,等. 多旋翼飞行器螺旋桨升力特性测试实验系统[J]. 实验室研究与探索, 2017, 36(1): 5-9.
- [7] 齐浩然,齐晓慧. 基于 LADRC 的四旋翼无人机自适应姿态跟踪控制[J]. 飞行力学, 2018, 36(3): 42-46.
- [8] 刘建龙,窦银科,左广宇. 倾转旋翼无人机动力学建模与控制系统设计[J]. 电子技术应用, 2021, 47(5): 59-63.
- [9] 赵超轮,赵国荣,戴邵武,等. 基于改进型 PID 算法的四旋翼控制研究[J]. 计算机仿真, 2020, 37(11): 48-54.
- [10] 胡改玲,桂亮,权双璐,等. 串联模糊 PID 控制的四旋翼无人机控制系统设计[J]. 实验技术与管理, 2019, 36(3): 132-135.
- [11] 万慧,齐晓慧,李杰,等. 基于 SADRC 的四旋翼飞行器姿态解耦控制方法[J]. 火力与指挥控制, 2020, 45(4): 83-88.
- [12] 范云飞,任小洪,袁文林. 基于并联 PID 的四旋翼飞行控制策略设计[J]. 软件, 2015(4): 37-39.

收稿日期:2022-10-11

(上接第 223 页)

## 4 结语

针对电梯内电动车识别存在精度不佳的问题,基于 YOLOv8 算法模型提出改进方案。数据预处理阶段,利用 AUGMIX 图像增强方法引入多样化的图像变换,提高了模型的泛化性能,减少了过拟合的风险,为模型的鲁棒性训练提供了有力支持。其次,通过将可变性卷积网络算子(DCNv3)和动态稀疏注意力融入 YOLOv8 模型,进一步提升了识别性能。改进后算法模型的精确率、召回率和平均精度均值分别达到了 94.5%、93%和 82.4%,电动车识别准确率达到了 95.8%,为电梯内电动车智能识别提供了理论基础。算法模型仍有改进空间,未来将考虑在保证精度的前提下,轻量化算法模型,提高识别速度和泛化能力,使算法模型更好地胜任电梯内电动车识别任务。

### 参考文献:

- [1] 周永强,王万军,杨浩,等. 基于图像识别与信息融合的电梯智能阻车系统研究[J]. 机电工程技术, 2023, 52(9): 141-144.
- [2] 李信芳,李艳. 浅析电梯智能阻车系统的加装属性[J]. 中国特种设备安全, 2022, 38(9): 46-48, 57.
- [3] 马子喻. 基于神经网络与证据理论的电动车进入轿厢识别研究[D]. 郑州:郑州大学, 2021.
- [4] 王佳宁. 面向自动驾驶的图像与点云融合及实时性研究[D]. 哈尔滨:哈尔滨工业大学, 2021.
- [5] 于飞,徐斌,王荣浩,等. 基于改进 YOLOv8 的旋转链板检测算法[J]. 制造业自动化, 2023, 45(9): 212-216, 220.
- [6] 刘奇,赵丽霞,郑曙光,等. 基于 DYOLO 神经网络的超声图像肾脏检测[J]. 计算机工程, 2021, 47(7): 307-313.

收稿日期:2024-02-26