

改进 YOLOv8 的电动自行车电池检测算法^{*}

帅勃宇 张雅丽

(中国人民公安大学信息网络安全学院 北京 100038)

摘要: 针对电动自行车电池样式单一,特征信息少,应用场景单一问题,重点关注模型轻量化,提出了一种基于视频监控的改进 YOLOv8 的电动自行车电池检测算法——PSPG-YOLO。在网络中特征提取方面,设计了一种多分支的 PStarblock 结构优化 C2f 模块,在 Starblock 的基础上进一步降低模型复杂度,增强特征信息表达能力;在多尺度融合特征金字塔方面,应用具有共享参数的空洞卷积改进 SPPF 结构,有效增大了感受野,保留了更多的细节特征信息;在检测头方面,提出了一种超轻量化共享卷积检测头 GSPH,应用可共享参数的部分卷积,大幅降低模型复杂度的同时能够动态调整锚点和步长,自动调整内部参数,从而提高对不同尺度特征图的适应能力。在专门针对电动自行车电池的自建数据集上实验表明:PSPG-YOLO 相较于基线模型 YOLOv8n 在计算量、参数量分别下降 57% 和 43% 的同时 mAP50 值提高 0.8,在与其他主流检测模型的对比中,综合性能最佳,为目前电动自行车电池违规入户管理提供了一种有效的解决方式。

关键词: 电动自行车电池检测;YOLOv8;轻量级;深度学习

中图分类号: TN911 **文献标识码:** A **国家标准学科分类代码:** 520.2

Improved YOLOv8 electric bicycle battery detection algorithm

Shuai Boyu Zhang Yali

(School of Information Network Security, People's Public Security University of China, Beijing 100038, China)

Abstract: In order to solve the problems of single battery style, less feature information and limited academic scenario, we focus on lightweight detection algorithm, and propose an electric bicycle battery detection algorithm based on an improved version of YOLOv8, termed PSPG-YOLO. In terms of feature extraction within the network mesh, have designed a multi-branch PStarblock to optimize C2f module enhancing the feature-information representational capacity with a simpler module. For multi-scale fusion feature pyramid, dilated convolution with shared parameters is employed to refine the SPPF structure, increasing receptive fields while preserving more detailed feature information. regarding detection head design, we introduce a GSPH. By leveraging shared convolution alongside partial convolution techniques, we reduce model complexity, additionally, both anchor point and stride can be dynamically adjusted; internal parameters are automatically calibrated to enhance adaptability across various scale feature maps. Experiments conducted on a self-constructed data-set specifically targeting electric bicycle batteries demonstrate that, compared to the baseline model YOLOv8n, PSPG-YOLO achieves reductions in FLOPs by 57% and in params by 43%, and improved mAP50 values by 0.8, yielding superior overall performance relative to other mainstream detection models, providing an efficient management method for electric bicycle battery illegal entry into households.

Keywords: electric bicycle battery detection; YOLOv8; lightweight class; Deep learning

0 引言

电动自行车作为当下最为常见的交通工具,因其节能环保、轻便等优点,已经逐步成为城市出行的重要方式。有关数据显示,目前我国电动自行车的保有量至少已有 3.5

亿,也就是说,平均每 4 个人中就有 1 人拥有电动自行车。不少居民为了方便进行电动自行车电池充电,将电池违规携带入户,这种行为极易引发火灾,存在极大安全隐患。

全国各级机关单位对严禁电动自行车电池入户充电行为都发布了相关规定,但仍有部分居民将电池违规携带入

户。作者实际调研发现,针对电动车违规入户行为,大多数小区已经实现了较好的监管效果,但在电动自行车电池的违规入户行为监管中仍效果甚微,大大增加了基层管理者的劳动量和管理方的额外支出。一些小区通过加装智能电表,监测用电量的方式,对电动自行车电池充电的行为进行监控;亦有小区在电池上加装 RFID 电子标签,同时在电梯和楼道内加装 RFID 信号采集器,对电池上楼进行监控,但此种方法因需要进行射频信号传输受环境影响较大,同时该种方法所需的电子标签和信号接收设备均需管理者自行购买提供分发,增加了额外的支出,尤其是在一些老旧小区物业管理较为薄弱甚至没有物业,因此目前普及率较低,除此之外该种方法需要居民自觉保持电子标签的张贴位置(如将电子标签贴在电动车而非电池上则无法检测住户只携带电池入户行为)和完好性,受限于住户的配合程度和自觉性,在实际应用中将管理者置于劣势方,难以保证管理效果;随着摄像头的不断普及,一些小区开始采用基于视频监控系统的电动自行车或电动自行车电池智能检测手段,在电动车检测中,目前以大华和海康为主的企业已推出智能检测摄像头,只需相应安装即可实现电动车的检测,但在电动自行车电池的检测领域,市面上已有的智能摄像头还未涉及此种功能。在电动自行车电池检测的实现中,目前存在两种情况:1)由管理方聘请第三方公司在已有的可编程智能摄像头上进行二次编程直接实现与上述电动车检测智能摄像头相同的效果,需要额外购买可编程智能摄像头并支付开发费用,价格不菲;2)在现有的摄像头基础上进行额外软件的开发,但企业为了自身利益往往不会只针对电动车电池进行算法开发,而是与其他功能一起开发集成出售,这就导致了系统对电动车电池检测的效果并不能完全达到预期。此外目前电动自行车电池检测算法均是以电梯内场景为背景进行开发的,无法对低楼层或未安装电梯的小区住户进行有效监督。

YOLO 系列算法是目前目标检测算法中发展最快、使用效果最好的算法之一,相比于 RCNN^[1] 系列和 SSD^[2] 系列算法,YOLO 系列算法拥有更好的实时检测效果,对软硬件更为友好,因此在工业等实际应用领域得到了更加广泛的应用。YOLO 系列算法从 YOLOv1 和 YOLOv9000^[3] 开始,由世界范围内不同的团队进行迭代创新,目前已迭代至 YOLOv11。其中,YOLOv8 发布于 2023 年初,相比于其他几个更新的版本,YOLOv8 目前在工业等实用领域使用情况较为成熟,应用面积更广具有更好的兼容性,是目前主要大范围研究和应用使用的 YOLO 系列算法,因此本文选用基础模型为 YOLOv8 算法模型。基于 YOLO 系列算法,在不同的检测任务中,张奕博等^[4]在轻量化卷积和特征金字塔融合方面进行改进,提出一种流对齐交叉融合特征金字塔网络,有效解决夜间车辆异常行为的检测问题;梁天添等^[5]利用可扩张残差结构优化 C2f 结构,并在主干网络的 SPPF 模块前置层引入了高效多尺度注意力机制,用以

解决恶劣天气车辆行人检测问题;李志星等^[6]在骨干网络中引入改进型双层链路注意力机制(BiFormer)^[7]加强模型对图像的分析能力和信息融合能力,使模型更适合钢丝绳损伤检测;杨永跃等^[8]结合 NAS-FPN^[9]思想,提出了改进的 PANet^[10]多特征融合结构,实现高效的双向跨尺度连接和加权特征层融合,实现对 PCB 缺陷检测的精度改进。

在电动自行车检测中,路成龙等^[11]提出一种结合 AUGMIX 图像增强技术并将变形卷积层和动态稀疏注意力机制融入 YOLOv8,有效解决了电梯内电动车识别问题。

在电动自行车电池检测中,由于企业开发算法并不对外公开,从可查询到的论文来看学术界目前研究重点为防止电池高温起火、电池性能提升等硬件方面研究,没有用于违规入户管理的专门针对性检测算法。因此,本文提出了一种改进 YOLOv8 的电动自行车电池检测算法,从特征提取、特征金字塔、检测头 3 个方面对 YOLOv8 算法进行针对性改进,有效解决了电动自行车电池样式单一,特征信息少,应用场景单一问题,重点关注模型轻量化,使其能够更好的适用于不同的软硬件环境 and 应用场景。

1 PSPG-YOLO

针对电动自行车电池检测中存在的重难点问题,本文以 YOLOv8n 为基线模型,提出了适用于多元场景监控视频图像下的电动自行车电池检测算法——PSPG-YOLO,主要改进有以下几个方面:

1)在特征表达上,本文基于 StarNet 中的 Starblock 模块,设计了一种多分支部分 Starblock——PStarblock (partical starblock),替换 C2f 模块中的 bottleneck 结构,增强特征信息表达能力,同时降低模型复杂度。

2)在特征金字塔融合方面,本文提出使用共享参数的空洞卷积替换最大池化,有效增大了感受野,增强特征表达能力,保留了更多细节信息,有效解决了背景信息干扰。

3)在检测头的设计上,本文提出了一种新的超轻量化共享卷积检测头 (super lightweight group convolution and shared partial convolutional detection head, GSPH),应用可共享参数的部分卷积,在降低模型参数数量的同时强化多尺度交流,有效解决了特征图计算冗余问题。

改进后的模型网络结构如图 1 所示。

1.1 PStar_C2f

在 YOLOv8 原始网络中,C2f 模块的 bottleneck 单元仅为两个普通卷积的堆叠,导致模型的计算量较大,检测速度较慢,不利于特征信息的高效利用和模型整体轻量化。结合电动自行车电池检测中的具体问题,相较于电池的固有大小,图像的背景信息较多、影响较大,bottleneck 模块在进行特征提取时不仅存在大量计算冗余还不利于特征信息表达。

StarNet^[12]是微软公司在 2024 年 3 月发布的一种超强

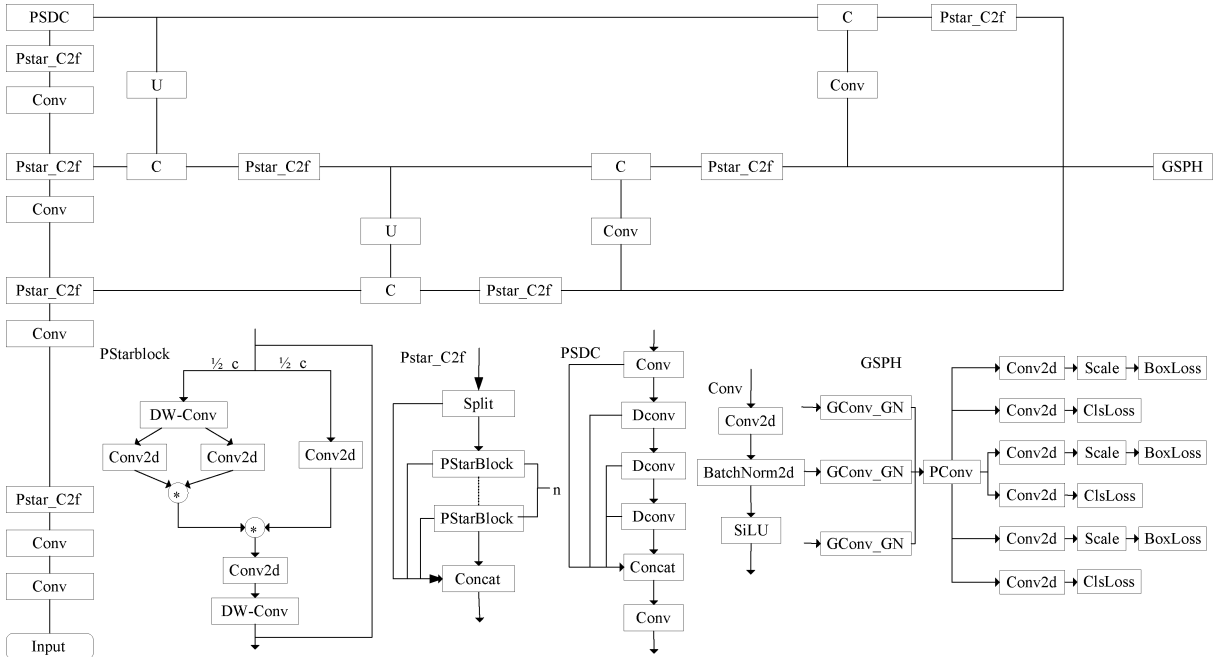


图 1 PSPG-YOLO 网络结构

Fig. 1 PSPG-YOLO network structure

轻量化网络结构,利用一种简单的逐元素相乘操作,实现多通道特征信息提取后的融合,可有效降低模型复杂度同时提高模型表达能力。本文基于 StarNet 中的 StarBlock 结构设计了一种更适用于电动自行车电池检测的多分支轻量化星形结构(partical starblock, PStarblock),替换原 C2f 中的 Bottleneck 结构,在无需拓展网络结构的同时将输入映射到高维、非线性特征空间,从而提高对输入特征信息的表达,降低模型复杂度,将改进后的结构命名为 PStar_C2f,其结构如图 2 所示。

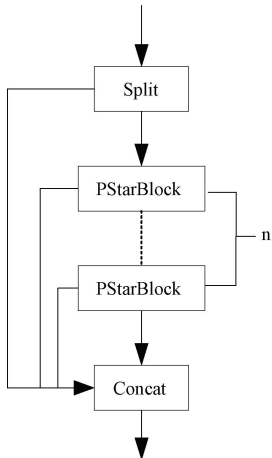


图 2 PStar_C2f 结构

Fig. 2 Structure of PStar_C2f

StarNet 中的 Starblock 模块主要由深度卷积^[13] (depthwise convolution, DWConv),普通卷积和星形操作

组成。输入特征首先经过一个深度卷积,随后分别用两个普通卷积对全部输入进行卷积使其拓展到高维空间,其中一个分支使用 Rule6 函数进行激活,然后将两个分支的结果进行逐元素相乘即星形操作,并使用一个普通卷积将高维空间信息映射到最原始的输入的维度,最后再通过一个深度卷积对特征信息进行进一步表达,其结构如图 3 所示。

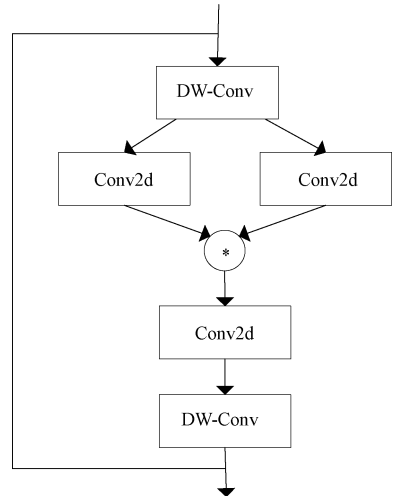


图 3 StarBlock 结构

Fig. 3 Structure of StarBlock

图 3 中,DW_Conv 为深度卷积操作,Conv2d 为标准卷积操作用于将特征信息在高度维度空间中的转换,“*”代表星形操作即逐元素相乘。设输入图像尺寸为 $h \times w$,通道数为 c ,隐空间的高维映射倍数为 mr ,深度卷积卷积核

为 k , 标准卷积卷积核为 1, 步长均为 1, 则整个过程的参数量可由式(1)表示。

$$Pa_{Star_all} = k \times k \times 1 \times c \times 2 + 1 \times 1 \times c^2 \times mr \times 3 \quad (1)$$

其中, Pa_{Star_all} 表示 StarBlock 的总参数量, 每个深度卷积的参数量为 $k \times k \times 1 \times c$, 上层两个标准卷积的参数量为 $1 \times 1 \times c \times mr \times c$, 即输入通道数为 c , 输出通道数为 mr 倍的 c , 下层标准卷积参数量为 $1 \times 1 \times mr \times c \times c$, 即输入通道数为 mr 倍的 c , 输出通道数为 c 。

计算量可由式(2)~(4)表示。

$$G_{Star_Conv} = k^2 \times c \times h \times w \times 2 + 1 \times 1 \times h \times w \times c^2 \times mr \times 3 \quad (2)$$

$$G_{Star_*} = O(c^2) + O(c^2) + O(c) = O(2c^2 + c) \approx O(c^2) \quad (3)$$

$$G_{Star_all} = G_{Star_Conv} + G_{Star_*} \quad (4)$$

其中, G_{Star_Conv} 表示 StarBlock 中所有卷积操作的计算量, G_{Star_*} 表示 StarBlock 中星形操作的计算量, G_{Star_all} 表示 StarBlock 总计算量, O 用于描述算法的最坏时间复杂度, 即在最坏的情况下, 算法中某一操作被执行的次数, 本文用此作为反映星形操作计算量的参数。每个深度卷积的计算量为 $k^2 \times c \times h \times w$, 上层两个普通卷积的计算量为 $1 \times 1 \times h \times c \times mr \times c$, 下层普通卷积的加算量为 $1 \times 1 \times h \times mr \times c \times c$, 星形操作的计算量可表示为 $O(c^2) + O(c^2) + O(c)$, 因为这里考虑的是增长的速率, 而不是常数系数, 因此可化简为 $O(c^2)$ 。

在上述的操作中, 深度卷积和普通卷积的高维空间映射均是对所有的输入通道进行的操作, 没有充分考虑到特征图相似冗余的问题, 这一点在 GhostNet^[14] 论文中已有所证明, 因此, 本文所提出的改进方法将输入特征先进行通道分割, 其中一半的通道进行深度卷积并进行高维映射, 而另一半输入通道则直接进行高维映射, 再将两个结果通过星形操作逐元素相乘, 最后使用普通卷积将维度映射到原始输入维度并使用深度卷积对特征信息进行进一步表达, 其结构如图 4 所示。

图 4 中各部分操作与图 3 中相同, 在输入图像尺寸、通道数、隐空间维度和各卷积参数均相同的情况下, PStarblock 的参数量可由式(5)~(8)表示:

$$Pa_{DW_Conv} = k \times k \times 1 \times \frac{1}{2}c \quad (5)$$

$$Pa_{Conv2d_up} = 1 \times 1 \times \frac{1}{2}c \times mr \times \frac{1}{2}c \quad (6)$$

$$Pa_{Conv2d_down} = 1 \times 1 \times mr \times \frac{1}{2}c \times \frac{1}{2}c \quad (7)$$

$$Pa_{PStar_all} = 2Pa_{DW_Conv} + 3Pa_{Conv2d_up} + Pa_{Conv2d_down} \quad (8)$$

其中, Pa_{DW_Conv} 表示深度卷积参数量, Pa_{Conv2d_up} 表示上层 3 个普通卷积参数量, Pa_{Conv2d_down} 表示下层普通卷积参数量, Pa_{PStar_all} 表示 PStarBlock 的总参数量。

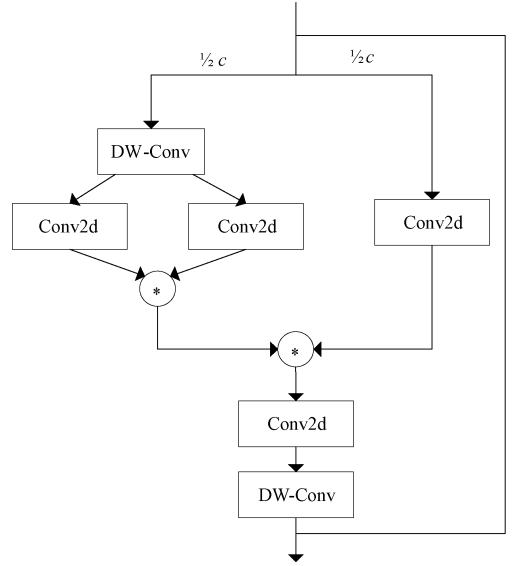


图 4 PStarBlock 结构

Fig. 4 Structure of PStarBlock

计算量可由式(9)~(13)表示。

$$G_{DW_Conv} = k^2 \times \frac{1}{2}c \times h \times w \quad (9)$$

$$G_{Conv2d_u} = 1 \times 1 \times h \times w \times \frac{1}{2}c \times mr \times \frac{1}{2}c \quad (10)$$

$$G_{Conv2d_d} = 1 \times 1 \times h \times w \times mr \times \frac{1}{2}c \times \frac{1}{2}c \quad (11)$$

$$G_{PStar_*} = O(mr \times \frac{1}{2}c \times mr \times \frac{1}{2}c) \approx O(c^2) \quad (12)$$

$$G_{PStar_all} = 2G_{DW_Conv} + 3G_{Conv2d_u} + G_{Conv2d_d} + 2G_{PStar_*} \quad (13)$$

其中, G_{DW_Conv} 表示 PStarBlock 中深度卷积计算量, G_{Conv2d_u} 表示 PStarBlock 中上层 3 个普通卷积计算量, G_{Conv2d_d} 表示 PStarBlock 中下层普通卷积计算量, G_{PStar_*} 表示 PStarBlock 中星形操作的计算量, G_{PStar_all} 表示 PStarBlock 总计算量。

综上所述, 当输入图像尺寸、通道数、隐空间高维映射倍数、深度卷积卷积核大小均相同时, PStarBlock 结构的总参数量比 StarBlock 结构的总参数量少 $k^2c + 2mr \times c^2$, 总计算量少 $k^2c \times h \times w + 2mr \times h \times w \times c^2$ 。因此, 本文所提出的 PStarBlock 结构相较于 StarBlock 结构有效减少了计算冗余。

1.2 PSDC 结构

SPP(spatial pyramid pooling)是由 He 等^[15]在 2015 年提出的空间金字塔结构, 有效避免了因剪裁、缩放等操作造成的图像失真问题, 同时解决了卷积神经网络对特征图相关特征重复提取问题, 大大提高了产生候选框的速度。在 YOLOv8 原始网络中沿用了 YOLOv5 中提出的 SPPF (spatial pyramid pooling-fast) 结构, 该结构在 SPP 基础上, 将多个并行的最大池化层改成串行方式, 进一步加快了产

生候选框的速度,降低了计算成本。SPPF 结构如图 5 所示。

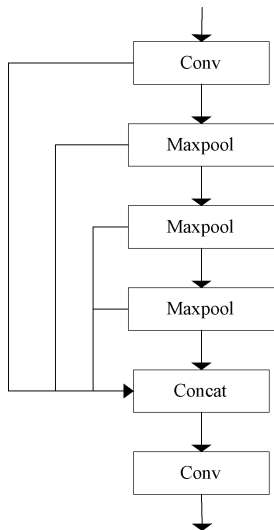


图 5 SPPF 结构

Fig. 5 Structure of SPPF

图 5 中,SPPF 结构使用最大池化层 maxpool 对特征图进行特征提取,并使用 concat 操作将多种不同的维度信息进行融合,最大池化的操作是将区域内的最大值作为该区域的值进行保留,从而实现降低维度的效果,这种操作只保留了最大值,丢失了许多细节信息,同时对抗噪能力表现不佳。在实际的电动自行车电池检测中,由于电动自行车电池形状较为规则,样式较为简单,易与其他物品混淆,如果使用最大池化,覆盖细节信息不够,极可能导致出现较多误检错检现象,同时基于现阶段监控摄像头画面情况,大多数图像中仍存在较多噪声,会影响最大池化的特征信息表达。

针对上述问题,本文提出了一种金字塔共享空洞卷积(pyramid shared and dilated convolution, PSDC),该结构使用卷积层替换最大池化层能够有效的保留特征图中的细节信息,实现更好的特征表达,但因卷积层的参数量和计算量明显高于最大池化层,因此在设计该结构时使用共享卷积思想控制参数量,使用空洞卷积^[16](dilated convolution, Dconv)实现特征维度的降低,从而实现了以较小的参数量和计算量的代价提高识别精度,同时 PSDC 结构相较于 SPPF 结构拥有更大的感受野,能够有效捕捉上下文信息,在实际检测中具有更好的表现。PSDC 结构如图 6 所示。

图 6 中,Conv 表示普通卷积,卷积核大小为 1,Dconv 表示空洞卷积,3 个空洞卷积卷积核大小为 3,三者同时共享参数,空洞率分别是 1、3、5。

1.3 超轻量化共享卷积检测头

在电动自行车电池的检测中,由于电池样式较为单一,在前序的特征提取过程中,多通道内的特征信息相似度较大,如果均用于计算会造成计算冗余,为解决上述问题,本

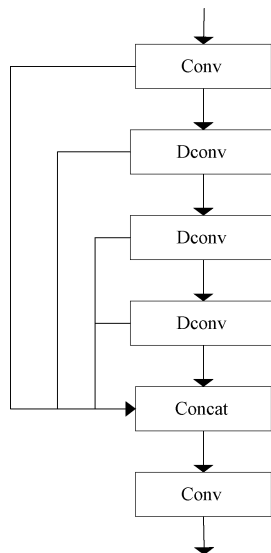


图 6 PSDC 结构

Fig. 6 Structure of PSDC

文基于共享卷积、部分卷积和分组卷积思想,提出了一种新的超轻量化共享卷积检测头(super lightweight group convolution and shared partial convolutional detection head, GSPH)。

GSPH 首先通过分组卷积对输入特征进行特征固定,并使用 GroupNorm^[17](Group Normalization)提升检测头定位和分类的性能,随后将不同尺度的特征信息通过一个部分卷积 PConv^[18](Partial Convolution)进行参数共享,以实现超轻量化,同时使用 Scale 层对特征进行缩放,用于解决每个检测头所检测的目标尺度不一致的问题。这种设计有效减少了模型参数数量,降低了模型的内存占用和计算需求,同时使用 PConv 作为共享卷积层使得模型能在保持高效处理能力的同时,减少资源消耗,有利于在资源受限或边缘计算设备上部署,能够更好的适用于不同的硬件环境。

在实际检测过程中,GSPH 也能够做到动态调整锚点和步长,根据输入图像的尺寸和特征图的变化自动调整其内部参数,从而提高对不同尺寸输入的适应能力和检测精度。GSPH 超轻量化共享卷积检测头如图 7 所示。

图 7 中,GConv_GN 表示组卷积、GroupNorm 和 SiLU 的顺序排列,即输入先经过一个组卷积再经过一个 GroupNorm 最后经过一个 SiLU 激活函数,PConv 表示部分卷积,Conv2d 表示普通标准卷积,Scale 表示尺度对齐,ClsLoss 表示类别分类损失,BoxLoss 表示回归框损失。

2 实验与分析

2.1 实验环境

本文基于 PyTorch 搭建深度学习框架,所有实验都基于表 1 中软硬件环境进行。网络训练使用 SGD 优化器,初

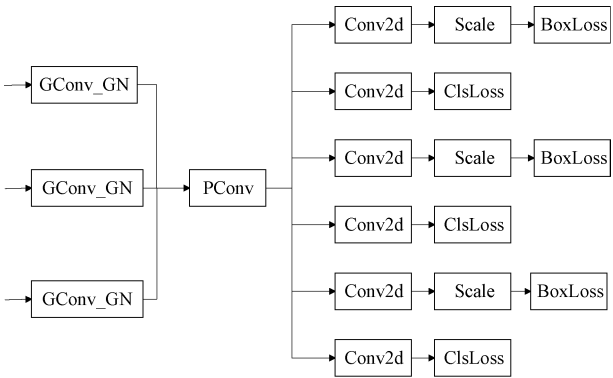


图 7 超轻量化共享卷积检测头结构

Fig. 7 Structure of super lightweight group convolution and shared partial convolutional detection head

始学习率设置为 0.01,初动量为设置 0.937,权重衰减系数为设置 0.000 5;batchsize 设置为 32,以 8 个工作线程加载数据,最大训练轮数为 300 轮,早停设置为 30。

表 1 实验环境

Table 1 Experimental environment

操作系统	ubuntu20.04
Python	3.8
Cuda	11.8
图形处理器	NVIDIA GeForce RTX 4090 (24 G)
中央处理器	Intel(R) Xeon(R) Platinum 8352V

2.2 电动自行车电池数据集

目前没有专门针对电动自行车电池的数据集,为了契合实际应用,基于真实视频监控图像,选取电梯内,楼道内,进户门等多种应用场景下的电动自行车电池图像,并添加部分背景信息与电池极度相似的图像,构建了真实监控视频下的电动自行车电池数据集(e-bike battery under surveillance video),该数据集截取至 6 个场景的 16 个角度的共计 28 段约 1 340 min 的真实监控视频,充分涵盖了在实际工作中可能遇到的电动自行车电池识别场景,共计 1 150 张内容形式各异的原始图片。

由于电池样式单一,数据复杂性不高,为了更好的适应不同情况下的检测需求,使用数据增强的手段,如随机旋转、翻转、随机遮挡、随机裁剪、随机噪声等方式将数据集扩充至 4 569 张,并随机划分成 3 426 张训练集图像和 1 143 张验证集图像。由于电动自行车电池样式大多为黑色规则的长方体,特征不明显,在实际识别过程中易于其他物体进行混淆,因此在上述数据集的基础上还构建了 e-bike batteries 电动自行车电池预训练数据集,由 472 张样式各异的单一电动自行车电池(即整幅图像中仅有电动自行车电池一种物体)组成,作为预训练数据集。e-bike battery under surveillance video 数据集部分原始图像如图 8 所示,

部分进行数据增强后的图像如图 9 所示,E-bike batteries 电动自行车电池预训练数据集部分图像如图 10 所示。

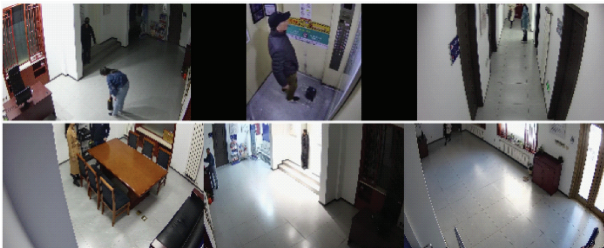


图 8 数据集部分原始图像

Fig. 8 Part of the original image of the dataset

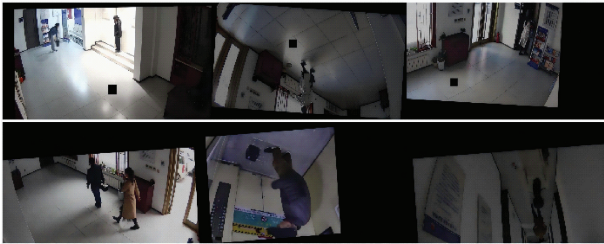


图 9 数据集部分数据增强图像

Fig. 9 Part of the dataset data augmented images



图 10 部分预训练数据集图像

Fig. 10 Part of the pre-training dataset images

2.3 评价指标

为了更好的评估模型的性能并与基线模型进行对比,本文使用目标检测中常用的 P (precision)、R (recall)、mAP50、GFLOPs、参数量(Params)作为评价指标。其中 P 代表正确检测出数量与总检测数量的比值,反映模型的准确率;R 代表正确检测出数量与需要检测出的总数量的比值,反应模型的检出率;mAP50 表示预测框与真实框交并比为 0.5 时,所有类别的平均检测精度;GFLOPs 为每秒万亿次浮点运算量,反应模型的计算量和运行速度;Params 反应模型复杂度。具体公式如式(14)~(17)所示。

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$AP = \int_0^1 P(R) dR \tag{16}$$

$$mAP = \frac{\sum_{j=1}^C AP_j}{C} \tag{17}$$

其中, TP 表示真正例(true positive)即正确预测为正的样本数量; FN 表示假反例(false negative),即模型未能正确预测为正的样本数量; FP 表示假正例(false positive),即模型错误预测为正的样本数量, $TP + FN$ 表示总检测数量, $TP + FP$ 表示需要检测出的总数量; AP 表示单一类别的准确率; C 为类别数。

2.4 预训练模型

因电动自行车电池样式较为简单,在训练过程中容易出现过拟合现象,因此在本文实验中首先使用 YOLOv8 原始模型在该数据集上进行训练得到预训练模型,并在后续实验中加载该预训练模型,使得改进模型对电动自行车电池特征更敏感,增加电动自行车电池的识别效果。预训练模型效果如表 2 所示。后续实验中全部默认加载此预训练模型。

表 2 预训练数据集结果
Table 2 Experimental environment

评价指标	模型(YOLOv8n)
P	0.964
R	0.978
$mAP50$	0.993
$FLOPs/G$	8.1
参数量/ M	3.0

2.5 消融实验

为了评估各个改进点的性能和对整体模型的贡献度,

本小节针对各改进点进行了消融实验。为方便观察,本小节实验将 3 种不同的改进方式 PStar_C2f、PSDC、GSPH 分别用 A、B、C 代替。实验中 GSPH 模块分组卷积参数设置为 4,部分卷积参数设置为 8,实验结果如表 3 所示。

从表 3 中可以看出,PStar_C2f、PSDC、GSPH 分别对模型复杂度和精度做出了不同贡献,

PStar_C2f 主要用于降低模型复杂度,PSDC 则是主要用于提高模型的检测精度,而 GSPH 则是在降低模型复杂度的同时提高检测精度,使最终改进模型性能达到最优。具体来看,单独进行 PStar_C2f 改进时,模型得到了有效的轻量化,计算量和参数量分别下降了 22%和 23%,而精度仅下降了 0.3,在此基础上进行 GSPH 改进,模型精度增幅为 0.5,计算量和参数量基本保持不变。在 PStar_C2f 改进的基础上不进行 PSDC 改进而是 GSPH 改进时,模型复杂度降至最低,相较于基线模型,计算量和参数量分别降低了 58%和 50%,但模型精度下降明显;单独进行 PSDC 改进时,由于特征信息表达更充分,感受野更大,模型精度相较于基线模型提高了 0.6,同时参数量和计算量基本不变;单独进行 GSPH 改进时,模型效果在 3 种单独改进中最好,精度提升了 0.9 的同时,计算量和参数量分别下降了 35%和 23%;当 3 种改进方式同时添加到原始网络中时,模型整体性能达到最好,相较于基线模型,计算量和参数量分别下降了 57%和 43%,同时精度提高了 0.8。因此本文所提出的 3 种改进方式均具有一定效果,最终的改进模型 PSPG-YOLO 相较于基线模型性能得到了全面的提升。

表 3 消融实验
Table 3 Ablation experiments

基线	A	B	C	P	R	$mAP50$	$FLOPs/G$	参数量/ M
YOLOv8n				0.853	0.885	0.928	8.1	3.0
YOLOv8n	√			0.847	0.878	0.925	6.2	2.3
YOLOv8n		√		0.894	0.862	0.934	8.1	3.2
YOLOv8n			√	0.882	0.859	0.937	5.3	2.3
YOLOv8n	√	√		0.868	0.889	0.93	6.2	2.4
YOLOv8n	√		√	0.848	0.871	0.919	3.4	1.5
YOLOv8n	√	√	√	0.861	0.886	0.936	3.5	1.7

2.6 对比实验

为了更全面的评价本文所提算法的优越性,本小节实验选取了目前主流的多种目标检测算法进行对比实验,包括 YOLOv3-tiny^[19]、YOLOv5n、YOLOv5s、YOLOv7-tiny^[20]、YOLOv9s^[21]、YOLOv10n^[22]、YOLOv11 n、RT-DETR-r18^[23],实验结果如表 4 所示。

从表 4 中可以看出,相较于表中其他主流实时检测模型,本文所提出的 PSPG-YOLO 算法在电动自行车电池检

测任务中具有碾压性的效果,从模型复杂度达到了最好水平,相较于精度最高的 RT-DETR,PSPG-YOLO 虽然在精度上低 0.5,但计算量和参数量分别降低了 83%和 97%,模型复杂度方面差距巨大;相较于最新的 YOLOv11 n 模型,本文所提 PSPG-YOLO 在参数量和计算量上仍具有明显优势,因此相比于其他主流检测模型,本文提出的 PSPG-YOLO 综合表现最佳。

表 4 对比实验

Table 4 Comparison experiments

算法	<i>P</i>	<i>R</i>	<i>mAP</i> 50	参数量/ <i>M</i>	<i>FLOPs</i> / <i>G</i>
YOLOv3-tiny	0.813	0.829	0.880	8.7	12.9
YOLOv5n	0.852	0.870	0.915	1.8	4.1
YOLOv5s	0.880	0.868	0.924	7.0	15.8
YOLOv7-tiny	0.839	0.839	0.908	6.0	13.0
YOLOv8n	0.853	0.885	0.928	3.0	8.1
YOLOv9s	0.885	0.882	0.929	7.2	26.7
YOLOv10n	0.831	0.871	0.916	2.3	6.5
YOLOv11n	0.881	0.860	0.938	2.6	6.3
RT-DETR-r18	0.890	0.867	0.941	20.0	56.9
本研究	0.861	0.886	0.936	1.7	3.4

3 检测结果

本文提出的 PSPG-YOLO 算法在电动自行车电池检测中能够胜任包括单元楼进户门、房间内、走廊、空旷室

内、电梯内等多种不同场景,能够有效的避免背景信息给电动自行车电池检测带来的干扰,检测置信率平均超过 80%,具有良好的检测效果。效果如图 11 所示。



图 11 PSPG-YOLO 检测效果

Fig. 11 Rendering of PSPG-YOLO detection

4 结 论

本文从实际应用出发,针对电动自行车电池检测问题中面临的挑战,提出了一种改进 YOLOv8 的电动自行车电池检测算法——PSPG-YOLO,通过消融实验和对比实验表明,本文所提出的算法在轻量化上具有明显优势,能够更好的适用于各种终端设备,对硬件具有较高的兼容性,符合实际应用需要。同时本文所提出的算法在对电动自行车电池的识别精度上相较于原始模型具有一定的提高,实现了轻量化的同时提高检测精度。

参考文献

[1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.

[2] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, 2016: 21-37.

[3] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]. IEEE Conference on Computer

- Vision and Pattern Recognition, 2017: 7263-7271.
- [4] 张奕博,张雅丽,帅勃宇. 交叉融合特征的夜间车辆异常检测算法[J]. 科学技术与工程, 2024, 24 (27): 11869-11877.
ZHANG Y B, ZHANG Y L, SHUAI B Y. Vehicle anomaly detection in night based on FaFPN [J]. Science Technology and Engineering, 2024, 24 (27): 11869-11877.
- [5] 梁天添,杨淞淇,钱振明. 基于改进 YOLOv8s 的恶劣天气车辆行人检测方法[J]. 电子测量技术, 2024, 47 (9):112-119.
LIANG T T, YANG S Q, QIAN ZH M. Improved YOLOv8s method for vehicle and pedestrian detection in adverse weather [J]. Electronic Measurement Technology, 2024, 47(9):112-119.
- [6] 李志星,杨啸龙,李天昊,等. 基于 YOLOv8 的煤矿用钢丝绳损伤检测算法[J]. 电子测量技术, 2024, 47(9): 120-128.
LI ZH X, YANG X L, LI T H, et al. YOLOv8-based surface damage detection of mine wire rope[J]. Electronic Measurement Technology, 2024, 47(9):120-128.
- [7] ZHU L, WANG X, KE Z, et al. Biformer: Vision transformer with bi-level routing attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323-10333.
- [8] 杨永跃,夏远超. PCB 缺陷检测深度学习算法的精度改进[J]. 电子测量与仪器学报, 2023, 37(5):11-19.
YANG Y Y, XIA Y CH. Accuracy improvement of deep learning algorithm for PCB defect detection[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 11-19.
- [9] GHIASI G, LIN T Y, LE Q. NAS-FPN: Learning scalable feature pyramid architecture for object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 7036-7045.
- [10] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [11] 路成龙,冯月贵,庆光蔚. 基于改进 YOLOv8 的电梯内电动车识别方法研究[J]. 机械制造与自动化, 2024, 53 (4):219-223, 229.
LU CH L, FENG Y G, QING G W. Research on identification method of electric vehicles in elevators based on improved YOLOv8[J]. Machine Building & Automation, 2024, 53(4): 219-223, 229.
- [12] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5694-5703.
- [13] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [14] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [15] HE K M, ZHANG X Y, REN SH Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9): 1904-1916.
- [16] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [J]. ArXiv preprint arXiv: 1511.07122, 2015.
- [17] WU Y X, HE K M. Group normalization [C]. European Conference on Computer Vision (ECCV), 2018: 3-19.
- [18] CHEN J R, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [19] FARHAD A, REDMON J. YOLOv3: An incremental improvement[C]. Computer Vision and Pattern Recognition, Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1-6.
- [20] WANG J Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [21] WANG J Y, YE H I, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [J]. ArXiv preprint arXiv: 2402.13616, 2024.
- [22] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.
- [23] ZHAO Y, LYU W Y, XU SH L, et al. Detrs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

作者简介

帅勃宇, 硕士研究生, 主要研究方向为安全防范技术。

E-mail:1627151981@qq.com

张雅丽(通信作者), 副教授, 主要研究方向为安全防范技术与智能视频技术。

E-mail:zhangyl_mail@163.com