

文章编号: 2095-2163(2023)09-0172-05

中图分类号: TP249

文献标志码: A

基于改进 YOLOv6 电动单车违法停放的检测方法研究

汪燕超, 胡旭晓

(浙江理工大学 机械工程学院, 杭州 310018)

摘要: 日常电动单车存在违法停放现象, 为消除安全隐患本文提出了一种改进的 YOLOv6 目标检测模型。在原 YOLOv6 模型的基础上, 在 neck 部分引入 Ghost 幻影卷积模块, 提升网络模型的检测速度; 在检测部分嵌入 CBAM 卷积注意力模块以突出所检测目标信息, 提升网络模型的检测精度。实验证明改进后的 YOLOv6 模型性能得到了一定的提升, 其精确率提升了 1.5 个百分点, 对电动单车违法停放的检测具有较大的实际意义。

关键词: 电动单车; YOLOv6 模型; 卷积注意力模块; 幻影卷积模块

Research on detection method of illegal parking of electric bicycles based on improved YOLOv6

WANG Yanchao, HU Xuxiao

(Faculty of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

【Abstract】 Daily electric bicycles are illegally parked to eliminate safety hazards. This paper proposes an improved YOLOv6 target detection model. Based on the original YOLOv6 model, the Ghost phantom convolution module is introduced in the neck part to improve the detection speed of the network model; the CBAM convolutional attention module is embedded in the detection part to highlight the information of the detected targets and improve the detection accuracy of the network model. The experiments demonstrate that the performance of the improved YOLOv6 model has been improved to a certain extent, and its accuracy rate has been increased by 1.5 percentage points, which is of greater practical significance for the detection of illegal parking of electric bicycles.

【Key words】 electric scooter; YOLOv6 model; convolutional block attention module; Ghost convolution model

0 引言

电动单车是生活中常见的交通工具, 据相关资料显示, 中国居民电动单车拥有数已超过 3 亿, 而大量电动单车随意停放的现象对公众的人身和财产安全造成了巨大威胁^[1]。

以往对于电动单车违法停放的监测主要依靠在场所附近安装监控并拍摄照片, 而后交给监控中心的工作人员进行监控和预警。这种方法消耗了大量的人力和物力, 对于电动单车违法停放的识别准确率也不高。近年来人工智能技术快速发展, 有学者提出了基于深度学习的目标检测模型, 深度学习模型主要分为两大类: 一类是基于候选区域的二阶段目标检测模型, 另一类是基于回归的一阶段目标检测模型, 两者分别依据候选框卷积和模型回归对目标进行检测, 代表算法有 Faster-RCNN (Faster

Region with CNN feature)、SSD (Single Shot MultiBox Detector) 和 YOLO 系列等, 但这些算法常常存在对小目标漏检和误检等问题。

针对传统检测方法的不足, 本文提出了改进的 YOLOv6 模型, 在 YOLOv6 模型检测部分嵌入 CBAM (Convolutional Block Attention Module) 卷积注意力模块, 并引入 Ghost 幻影卷积模块, 以提高模型的精确度和检测速度。改变外接摄像头所拍摄电动单车停放图片的亮度、遮挡范围和目标个数等属性制作相应数据集, 据此评估改进后的模型的鲁棒性、识别速度和精度, 进一步提升改进后模型的性能。

1 YOLOv6 模型简介

YOLO 系列模型采用直接回归的方法, 与传统的目标检测模型相比, YOLO 系列模型计算效率高, 能够方便地进行端到端的训练, 因此能快速检测

作者简介: 汪燕超(1999-), 男, 硕士研究生, 主要研究方向: 图像处理; 胡旭晓(1965-), 男, 博士, 教授, 主要研究方向: 图像处理、机器视觉。

收稿日期: 2022-09-26

哈尔滨工业大学主办 ◆ 科技创见与应用

出目标,在许多实际的应用场景都取得了较好的效果。

YOLOv6 模型的网络结构主要由输入端 (input)、骨干网络 (backbone)、颈部 (neck) 以及输出端 (output) 构成。YOLOv6 在输入端采用了 HSV (Hue, Saturation, Value) 数据增强方式,将输入端所输入的 RGB 图片拆分成 3 个通道即色度,饱和度和明度,在此基础上设置 Hgain 色度增值、Sgain 饱和度增值和 Vgain 明度增值系数,将 3 个通道各自的增值系数与输入端产生的范围在-1~1 的 3 个随机系数相乘,加一可求得随机的增益系数,借此可重新获得映射增强后的 RGB 图片,不仅可以减少原始图像中的噪声干扰,还避免了在图像增强的过程中造成失真。

在颈部部分,YOLOv6 模型引入了重参数化的视觉几何组结构,提出了效率更高的高效重参数化

方法 EfficientRep,将颈部中步长为 2 的卷积层替换为了步长为 2 的重参数化卷积层,降低了模型对内存的占用率,并且将信号处理块改为重参数化块结构,加快了模型的推理速度^[2]。

YOLOv6 模型在颈部部分的特征融合上同样引入了重参数化的结构,将像素聚合网络与重参数化块两者结合,降低了模型在硬件上的延时。而在输出端部分对检测头进行了解耦,基于 Hybrid Channels 的策略重新设计出了一个更高效的解耦头结构,避免了 YOLOvx 在解耦头中新增两个额外的 3×3 的卷积,降低了网络运算的复杂度,此外 YOLOv6 模型还分开了边框回归与类别分类的过程,提高了模型的性能,且由于重新定义了算法的距离损失,导致 YOLOv6 模型加快了收敛的速度并降低了检测头的复杂度^[3]。

YOLOv6 模型网络结构如图 1 所示。

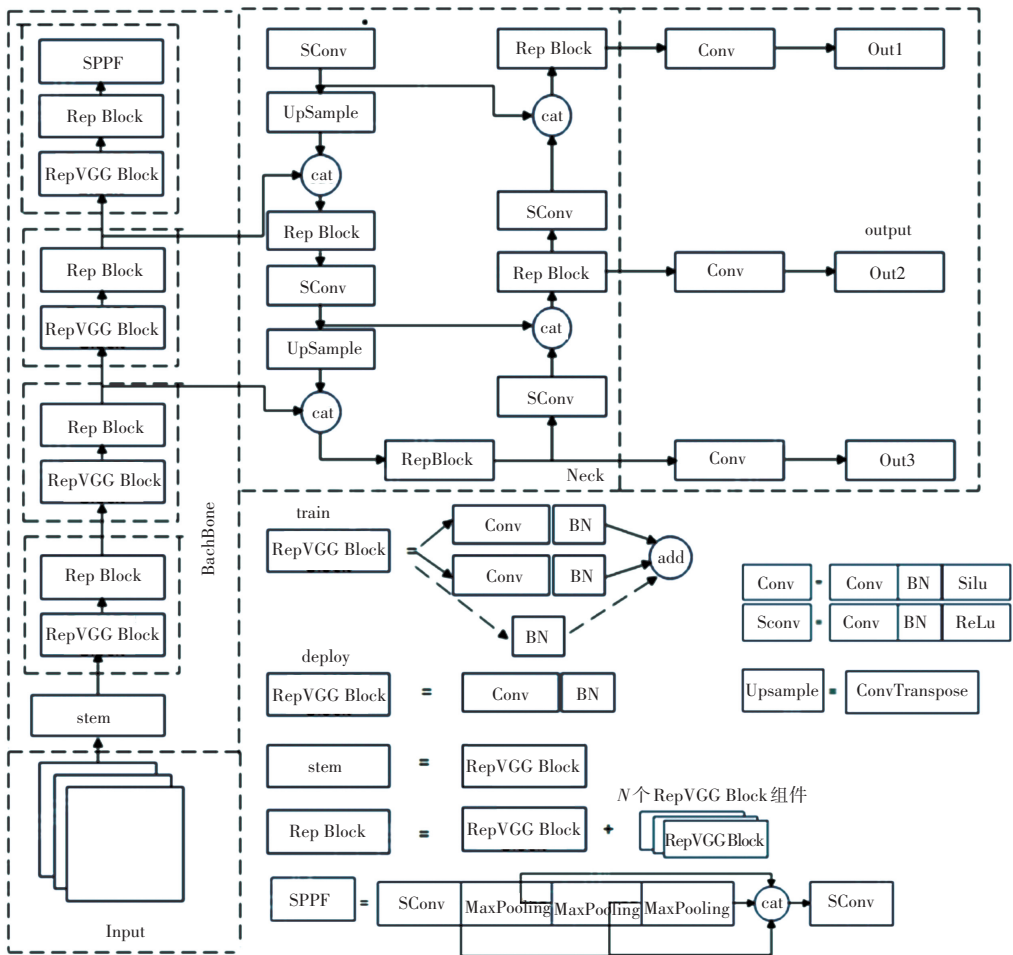


图 1 YOLOv6 网络结构图
Fig. 1 YOLOv6 network structure

2 改进的 YOLOv6 模型

(1) 在 neck 结构中将普通卷积替换为轻量级的 Ghost 幻影卷积模块,如图 2 和图 3 所示;

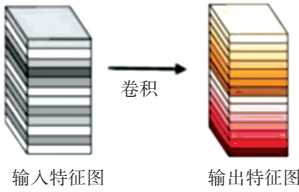


图 2 普通卷积

Fig. 2 Ordinary convolution

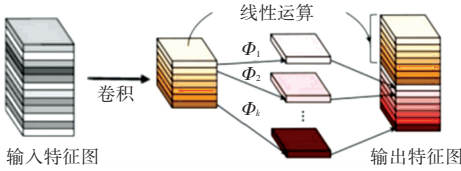


图 3 Ghost 幻影卷积

Fig. 3 Ghost convolution

传统的卷积运算为了得到较为全面信息,采用了较大的卷积核和通道数会产生较多相似的特征图,假设输入特征图的大小为 $w * h * c$,经过 n 个卷积核,每个卷积核大小为 $k * k * 1$,根据卷积运算规则可知总体计算量为 $w * h * c * k * k * n$,其中 h 和 w 是输入特征图的长度和宽度, c 是输入特征图通道数, k 是采用卷积核的长度和宽度, n 指卷积的次数。

普通卷积过程所产生相似的特征图,可以直接通过线性变换得到,而不需要进行复杂的非线性变换得到。于是引入了 Ghost 模块,Ghost 模块将传统的卷积模块分成了两个部分,第一步仍然进行普通卷积,但减少特征图的输出数量,第二步在此基础上进行线性变换生成相似特征图,最终所得特征图数量与传统卷积运算一致,其总体计算量为 $n/s * h * w * k * k * c + (s - 1) * n/s * h * w * d * d$,其中 s 和 d 分别是线性变换的次数以及线性变换过程中卷积核的大小。

由上述结论可得出传统卷积运算核与 Ghost 卷积运算所耗时间比值,式(1):

$$r_s = \frac{n \cdot h \cdot w \cdot c \cdot k \cdot k}{\frac{n}{s} \cdot h \cdot w \cdot c \cdot k \cdot k + (s - 1) \cdot \frac{n}{s} \cdot h \cdot w \cdot d \cdot d} \quad (1)$$

由于 s 是个常量,上述公式比值约为 s 即线性变换的次数,故可知 Ghost 卷积运算相比与传统卷积运算减少了网络模型计算量,降低了所耗时间成本。

(2) 卷积注意力模块 CBAM

CBAM 模块是一种结合了空间和通道的注意力机制模块,常在深度学习网络结构中被用于提升网络特征提取的性能,与通道注意力模块相比增加了空间的注意力模块,因此能够取得更好的目标识别效果,因此能够取得更好的目标识别效果。CBAM 模块具体的网络结构如图 4 所示。

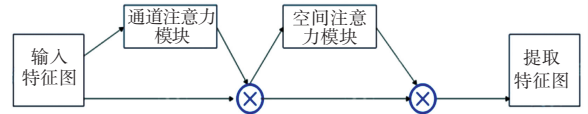


图 4 CBAM 模块网络结构

Fig. 4 Network structure of CBAM module

由图 4 可知, CBAM 模块主要由通道注意力模块和空间注意力模块所组成, CBAM 模块中的通道注意力模块与传统通道注意力模块相比采取了全局平均池化和全局最大池化,两种不同的池化意味着提取的高层次特征更加丰富,并且平均池化和最大池化共同与多层共享感知机连接以减少学习参数。通道注意力模块网络结构如图 5 所示。

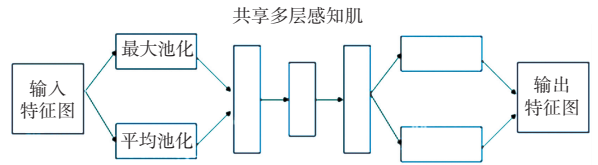


图 5 通道注意力模块网络结构

Fig. 5 Channel attention module network structure

在通道注意力模块中的多层共享感知机是一种 3 层结构的人工神经网络,包括输入层,输出层和隐含层 3 个部分,其中隐含层可以由多个隐层所构成,并且多层共享感知机中层与层是全连接的, w_1, w_2, w_3 则是权重, $\&$ 为偏置, X 为输入参数,可知隐藏层输出为 $w_1 X + \&$, 其结构如图 6 所示。

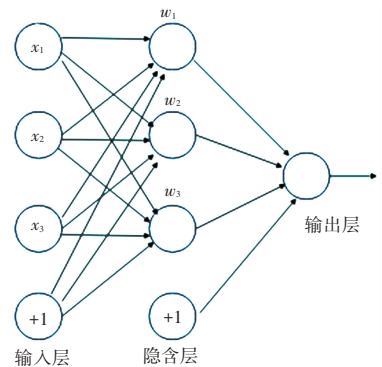


图 6 多层共享感知机结构图

Fig. 6 Structure of multi-layer shared perceptron

空间注意力模块同样采用了最大池化和平均池

化,不同的是在通道这个维度上进行的操作,即把所有的输入通道池化成 2 个实数,接着通过一个 7×7 的卷积核得到空间注意力矩阵,在此矩阵基础上与输入特征图相计算可得最终的输出特征图,其结构如图 7 所示。

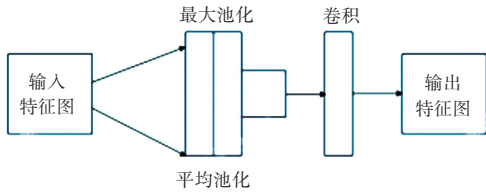


图 7 空间注意力模块网络结构

Fig. 7 Spatial attention module network structure

本文采用的 CBAM 注意力模块尽可能提取所识别目标的特征信息,忽略目标背景等非必要信息,从而提升改进后的网络模型检测精度。

3 数据集制作及实验过程

本文对于电动自行车违法停放的现象标注了在电梯和消防通道两个场景下的数据,并结合开源的 coco 数据制作了数据集,但由于数据格式不匹配或图像文件损坏等问题需要筛掉不符合要求的图像文件,采用 HSV 数据增强方法对所得数据集进行数据预处理,增强电动自行车图像的色彩深度和对比度,经处理后得到符合条件的图像数据共计 1 030 幅,并将所得图像数据使用 LabelImg 工具标注。

将标注好的数据集按 8:2 的比例划分为训练集和测试集,得到训练集图片共计 824 幅,测试集图片 206 幅。

本次实验在 Ubuntu 22.04 操作系统上完成,硬件配置采用 NVIDIA GFore GTX1660Ti 显卡和 Intel i5-7300H 处理器。网络模型采用 pytorch1.10.2 搭建,在训练过程中,设置初始学习率为 0.01,训练周期上限设置为 500,批次大小设置为 32,即每次参与训练的样本数量。

4 评价标准和实验结果

查准率 (*Precision*) 指正确预测为正的占全部预测为正的比例,式(2):

$$Precision = TP / (TP + FP) \quad (2)$$

其中, *TP* 指的是将正例判定为正例,即将图中消防通道和电梯等场所背景中的电动自行车检测为电动自行车的样本,而 *FP* 指的是将反例判定为正例,即将图中消防通道和电梯等场所背景检测为电动自行车的样本。

查全率 (*Recall*) 指即正确预测为正的占全部实际为正的比例,式(3):

$$Recall = TP / (TP + FN) \quad (3)$$

其中, *FN* 指的是将正例判定为反例,即将图中消防通道和电梯等场所背景中的电动自行车检测为背景。

综合查准率和查全率可得到 *P-R* 曲线, *P-R* 曲线代表了网络模型的预测效果, *P-R* 曲线所围成的面积大小被称为 *AP* 平均精度,而针对检测的每个目标都有单个的平均精度 *AP* 值,由此可引入全类平均准确率 *mAP*,即对所有的目标类别的 *AP* 值再次求和取平均,式(4):

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (4)$$

此外,工程的实时性也十分重要,因此 *FPS* (*Frames Per Second*) 帧率也是衡量目标检测算法性能的重要指标,代表了网络模型每秒中图像的检测速度。

本文将改进后的网络模型与原 YOLOv6 模型在平均准确率、精确率和帧率这 3 个方面进行比较,实验结果如图 8,图 9 和表 1 所示,改进后的网络模型准确率提升了 1.5 个百分点。



图 8 YOLOv6 模型检测结果

Fig. 8 YOLOv6 model detection results

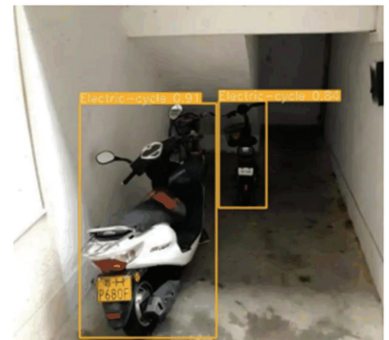


图 9 改进后的 YOLOv6 模型检测结果

Fig. 9 Improved YOLOv6 model detection results