

# 夜间行驶车辆远光灯检测方法



龚航 刘培顺

中国海洋大学信息科学与工程学院 山东 青岛 266100

(201612735@qq.com)

**摘要** 有效地对夜间车辆违规使用远光灯的行为进行管理,可以降低夜间交通事故的发生,但目前缺乏高效的远光灯检测方法,相关交通法规无法得到有效执行。针对此问题,文中提出了一种夜间车辆远光灯检测深度学习算法。该算法基于 YOLOv3 进行设计,通过降低各层卷积层维数的方式,来减少整体网络的参数量,提高算法的运行速度;然后对网络的残差组件进行改进,使用标准的残差组件,同时设计了一个空洞卷积模块来加强网络局部和全局特征的融合,增强了网络的特征表达能力;接着对 YOLOv3 的损失函数进行了改进,优化小尺寸目标对坐标损失的贡献,增强了小尺度目标的检测能力;最后对 YOLOv3 先验框聚类算法和个数进行优化,提高模型的表达能力和检测速度。实验结果表明,所设计的算法的平均准确率(mAP)达到了 99.09%,相比 YOLOv3 提升了 30%,满足了实用化要求,能够有效地检测违规行为。

**关键词:** 交通管理;远光灯检测;深度学习;YOLOv3;残差网络;空洞卷积

**中图分类号** TP391

## Detection Method of High Beam in Night Driving Vehicle

GONG Hang and LIU Pei-shun

College of Information Science and Engineering, Ocean University of China, Qingdao, Shandong 266100, China

**Abstract** Managing the illegal use of high beams can reduce the occurrence of night traffic accidents. However, at present, there is no efficient detection method for high beam of night, and relevant traffic regulations cannot be effectively implemented. In order to solve this problem, an algorithm to detect the high beam at night is proposed in this paper. Based on YOLOv3, this algorithm optimizes the network structure of YOLOv3, accelerates its operation speed, uses standard residual components and dilates convolution to enhance the feature expression ability of the network, and then the loss function of YOLOv3 is improved to optimize the contribution of small-scale target to coordinate loss, which enhances the detection ability of small-scale target, finally YOLOv3 prior frame clustering algorithm and number are optimized to improve the expression ability and detection speed of the model. The experimental results show that the mean average precision (mAP) of the algorithm designed in this paper is 99.09%, and 30% higher than that of YOLOv3. The algorithm satisfies the practical requirement and can detect the violation effectively.

**Keywords** Traffic management, High beam detection, Deep learning, YOLOv3, ResNet, Dilated convolution

### 1 引言

汽车前照灯分为近光模式和远光模式两种情况,正确使用远光灯可以带来方便,但是违规使用远光灯会带来极大的安全隐患。据公安部数据统计,夜间发生的交通事故中,约 30%与违规使用远光灯有关<sup>[1]</sup>。另据相关网络调查显示<sup>[2]</sup>,夜间滥用远光灯的行为已成为广大网民心中的十大开车恶习之一。

使用传统技术对夜间车辆远光灯进行检测,当外部因素干扰较少时,检测准确率较高。但是,当路况复杂,或者遇到雨、雪和道路积水等外部环境时,由于信息量大、冗余信息多,对远光灯检测造成干扰,大大降低了远光灯检测的准确率。

目前针对夜间车辆远光灯检测算法的研究成果较少。2018 年, Su 等<sup>[3]</sup>根据夜间各光源的成像差异和车灯的帧间运动规律,提出了一种结合随机森林判断远光灯的方法。但该方法是在传统基于固定阈值分割和先验知识的车灯检测方法之上的,适用性较差。2018 年,安谐智能科技有限公司提出一种利用低曝光相机提取车灯区域,并结合特定车型的远光灯形态学特征对车灯区域进行分类的远光灯检测专利,但该方法只能对双灯筒灯型的远光灯进行识别,检测率较低。

近年来,随着深度学习技术的不断发展,许多基于深度学习的目标检测算法被提出。将基于深度学习的目标检测算法应用到远光灯识别上,可以有效解决传统远光灯识别方法检测率低和误警率高等问题。

收稿日期:2020-07-03 返修日期:2021-01-31 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家重点研发计划基金资助项目(2017YFC0806200)

This work was supported by the National Key Research and Development Project(2017YFC0806200).

通信作者:刘培顺(liups@ouc.edu.cn)

目前基于深度学习的目标检测算法主要分为两类:基于区域的目标检测算法和基于回归的目标检测算法。基于区域的目标检测算法,典型代表有 R-CNN<sup>[4]</sup>,Fast R-CNN<sup>[5]</sup>,Faster R-CNN<sup>[6]</sup>,R-FCN<sup>[7]</sup>和 Mask R-CNN<sup>[8]</sup>等,此类算法先产生候选区域信息,然后对候选区域目标进行分类和识别。但由于候选框提取阶段的计算复杂,限制了此类算法的检测速度,使其难以满足实时检测的要求。基于回归的目标检测算法,典型代表有 YOLO<sup>[9]</sup>,SSD<sup>[10]</sup>,YOLO9000<sup>[11]</sup>,RetinaNet<sup>[12]</sup>,YOLOv3<sup>[13]</sup>和 RFBNet<sup>[14]</sup>等,此类算法直接通过神经网络来获取物体的类别概率和位置信息,是一种快速的、端到端的目标检测算法,其中 YOLOv3 算法是该类别中的优秀算法,但该算法识别物体位置的精准性差、召回率低。为加强 YOLOv3 网络的特征提取,文献[15-17]针对 YOLOv3 特征提取网络提出了多种改进方法,这些方法不同程度地提高了 YOLOv3 的平均准确率。针对 YOLOv3 目标框回归策略不合理的问题,Rezatofighi 等<sup>[18]</sup>和 Xu 等<sup>[19]</sup>用多种算法对 YOLOv3 的 IoU(Intersection-over-Union)损失函数进行了改进,解决了 IoU 算法无法对非重叠区域进行优化的问题,提高了 YOLOv3 的平均准确率。但是改进算法也带来了新的问题,有的增加了原有网络的计算量,使模型的训练变得更加复杂,有的在噪声较大的环境下存在鲁棒性较差、虚警率较高等问题。

为实现夜间车辆远光灯的高效检测,本文首先使用 YOLOv3 目标检测算法对夜间交通摄像机拍摄的车辆远光灯照片进行了训练和检测,实验表明原算法的检测效率和准确率都不能满足实时工作的需要,因此本文基于 YOLOv3 设计了一个夜间车辆远光灯检测算法。首先为降低 YOLOv3 下采样造成的信息丢失,本文对原网络的输入尺度进行了扩大,并通过降低各卷积层的维数来提高算法的运行速度;然后对降维后网络的残差组件进行改进,改用标准的残差<sup>[20]</sup>(Res-Net)组件来增强网络的特征表达能力,并设计了一个空洞卷积<sup>[21]</sup>(dilated convolution)模块来加强网络局部和全局特征的融合;接着对 YOLOv3 的损失函数进行了改进,优化小尺度目标对边框损失的贡献;最后对 YOLOv3 先验框聚类算法和个数进行优化,提高模型的表达能力和检测速度。实验表明,本文设计的算法在准确率和效率上都有很大提高,满足了实时工作的需求。

## 2 夜间远光灯检测

### 2.1 实验数据采集

现阶段还没有公开的夜间车辆远光灯数据集,为此我们与交警队合作录制了一批夜间行车录像,从中选出了适合的训练集和测试集。

目前交通摄像机为了清晰地抓拍卡口车辆,安装角度普遍压得非常低,以摄像机安装所在的立杆(6 m 高)为起点,摄像机的成像下边沿一般距离立杆 10~15 m,如图 1(a)所示,摄像头安装角度过低,使得近处车辆的远、近光灯区分不明显,即使成功抓拍,也难以形成证据。本文在数据采集和实际应用上对摄像头安装角度进行了调整,通过大量实验最终设计为摄像机的成像下边沿距离立杆 20 m,如图 1(b)所示,使

150 m 外车辆的远光灯照射光能直射进入摄像头,形成耀眼的光斑。

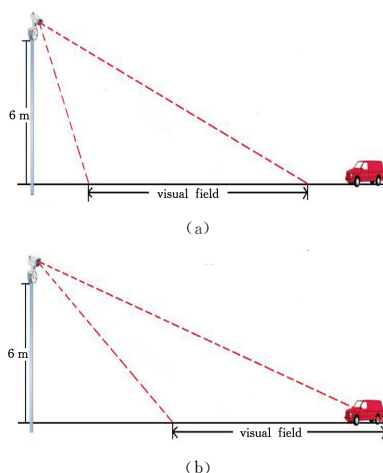


图 1 监控摄像机示意图

Fig. 1 Visual field of surveillance camera diagram

如图 2 所示,调整摄像机安装角度后,从图像上看,车辆远光灯视觉效果明显,远、近光灯“炫目”程度对比清晰。基于此方法,本文利用海康 900 万像素摄像机在多地 11 个路口共采集了 5000 张车辆远光灯图片,分辨率为 1920×1080。为了防止训练过程中的过拟合,本文对数据进行了增强,通过调整图片的饱和度和色度和曝光量来生成更多的训练样本。同时采用多尺度训练策略,随机选取尺寸从 384×384 到 736×736 之间的数据进行训练,以增强模型对不同尺寸图像的鲁棒性。



图 2 监控摄像机调整后抓拍对比示意图

Fig. 2 Comparison of captured results of surveillance camera before and after adjustment

### 2.2 数据标注

本文使用 LabelImg 工具对数据集进行标注,直接生成 YOLO 格式的 txt 标签文件。为保持训练样本和测试样本的独立性,使实验的结果更具说服力,本文对采集的图片数据进行了随机划分,其中训练集 4000 张,测试集 1000 张。

本文首先使用现有深度学习目标检测算法进行检测实验,实验结果如表 1 所列,其中  $mAP@50$  表示平均准确率,MT 表示单张图片的平均检测时间。可以看出,YOLOv3-512 的检测效果最好,但是性能上还不能满足实时检测的需要。YOLOv3 是为了解决多目标检测的问题而设计的算法,远光灯检测属于单目标检测,对于单类目标检测而言,YOLOv3 深层次的网络过于复杂和冗余,导致算法效率不高,为此本文基于 YOLOv3 设计了一个高效的检测算法。

表1 现有目标检测算法性能对比

Table 1 Performance comparison of existing target detection algorithms

Method	Backbone	$mAP@50/\%$	$MT/ms$
Faster R-CNN	VGG16	89.18	248
Faster R-CNN	ResNet101	90.35	289
SSD300	VGG16	87.51	27
SSD512	VGG16	90.72	68
YOLOv2-512	Darknet-19	90.43	29
YOLOv3-416	Darknet-53	92.86	54
YOLOv3-512	Darknet-53	96.56	61

## 2.3 检测算法设计

### 2.3.1 调整网络参数

在目标检测任务中,模型不仅需要识别出目标的类别,还需要对目标的位置进行定位。YOLOv3 采用了多尺度特征融合的方式进行预测,网络输入经过 5 次下采样和 2 次特征融合操作之后,可得到 3 个不同尺度的特征图。虽然 YOLOv3 下采样操作能增大网络的感受野,使深层网络获取更加丰富的语义信息,但随着特征图尺度的减小,深层网络会不可避免地丢失大尺度目标的位置信息;而且在特征融合问题上,小尺度目标的语义信息在深层网络中也有可能变得不可见,导致融合后的大尺度特征图无法获取足够的语义信息来对小尺度目标进行检测。

在远光灯检测任务上,由于目前交通摄像机视频主码流分辨率通常为  $1920 \times 1080$ ,对于整个画面而言,车辆远光灯区域只占了画面的一小部分。而且 YOLOv3 在训练和检测时,都需要将输入图片压缩成固定大小,在一定程度上也缩小了远光灯区域,使得模型从面积较小的远光灯区域中无法提取到足够的信息传递给后面的目标检测器,导致模型检测时出现漏检的情况。如图 3 所示,后面车辆的远光灯均未被识别。



图3 漏检示意图

Fig. 3 Missed detection diagram

因此,为保持模型特征图的分辨率足够大,避免远处小目标在经过多次下采样之后出现特征消失的问题,提高复杂场景下目标的检测率,本文将输入图像的尺寸增大至  $512 \times 512$ 。实验结果证明这一改进提高了目标检测的准确率和定位的精度。

此外,为了实现快速的车辆远光灯检测,本文对 YOLOv3 网络各卷积层的维数进行了调整,将网络的宽度减小为原来的  $1/2$ ,使得改进后的 YOLOv3 网络的计算开销大幅度减少。

### 2.3.2 网络设计

YOLOv3 网络是一个全卷积网络,在特征提取部分大量使用了简单的 bottleneck 结构的残差组件,保证了网络结构

在很深的情况下仍能收敛。YOLOv3 残差组件在第一层采用  $1 \times 1$  卷积进行降维,虽然该  $1 \times 1$  卷积能大量地减少每次卷积的维数和参数量,加速算法的运行速度,但对于浅层低纬度的卷积再次使用  $1 \times 1$  卷积进行降维,会导致部分特征信息丢失,使浅层的卷积无法提取到足够的信息往下传递。同时,受限于 YOLOv3 残差组件的深度和卷积核的大小,YOLOv3 的感受野有限。因此,本文改用标准的残差组件代替原来 bottleneck 结构的残差组件,设计了一个全  $3 \times 3$  卷积的特征提取网络。

如图 4 所示,改进后的 YOLOv3 残差组件采用了两层相同维数的  $3 \times 3$  卷积层代替原来的  $1 \times 1$  卷积层和  $3 \times 3$  卷积层(维数为  $1 \times 1$  卷积的 2 倍),使得网络的特征提取能力大幅度提升。同时, $3 \times 3$  卷积能够在不改变特征图尺寸的基础上,获取更大的感受野和更丰富的语义信息,有效提高了网络分类的准确性。

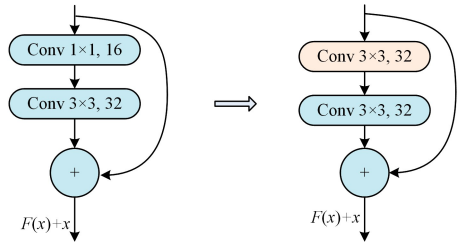


图4 改进的 YOLOv3 残差组件示意图

Fig. 4 Improved YOLOv3 residual component diagram

### 2.3.3 加强特征融合

有效感受野在目标检测过程中具有非常重要的作用。受 SPP-Net 的启发,YOLOv3-SPP 提出在 YOLOv3 小尺度特征图上引入 SPP 模块,实现局部特征和全局特征的 Feather Map 级别的融合,丰富最终特征图的表达能力,从而提高模型的平均检测率。其中 YOLOv3-SPP 添加的 SPP 模块如图 5 所示,它由 4 个并行的分支构成,分别是 kernel size 为  $1 \times 1, 5 \times 5, 9 \times 9$  和  $13 \times 13$  的最大池化。

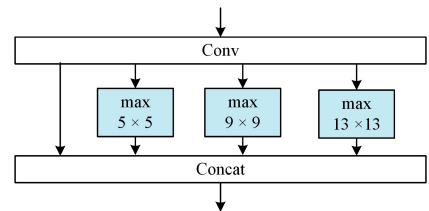


图5 YOLOv3-SPP 模块示意图

Fig. 5 YOLOv3-SPP module diagram

但是 SPP 模块中的最大池化容易造成特征图内部数据结构和空间层级化信息的丢失,且池化层参数无法学习更新。为了更好地利用有效感受野和增加模型对目标的多尺度特征表达能力,本文采用空洞卷积来替代 SPP 模块中的最大池化,设计了一个 DC(Dilated Convolution)模块来增大感受野,实现了局部特征和全局特征的特征图级别融合。如图 6 所示,改进后的 DC 模块采用 3 个扩张率分别为 1, 3 和 5 的  $3 \times 3$  卷积来获得比 SPP 模块更全面的空间覆盖。

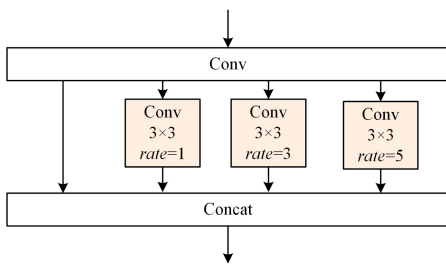


图6 DC模块示意图

Fig. 6 DC module diagram

结合上述3点改进策略,本文设计的检测算法的网络结构如图7所示。特征提取网络第一层首先使用了一个16维的 $3 \times 3$ 卷积对尺寸为 $512 \times 512$ 的输入图像进行过滤;其次第二层使用了一个步长为2的32维 $3 \times 3$ 卷积进行下采样;然后执行5组分别包含1个、2个、8个、8个及4个标准残差组件的残差模块,分别获取 $64 \times 64$ 、 $32 \times 32$ 和 $16 \times 16$ 3种不同尺度的特征图;接着在 $16 \times 16$ 特征图上执行DC模块;最后结合多尺度融合的方式,分别在不同尺度的特征融合图上独立进行预测。与原YOLOv3特征提取网络相比,本文的特征提取网络各尺度特征图维数仅为原来的1/2,但由于输入图像尺寸的增大和采用了包含两层 $3 \times 3$ 卷积的残差组件,实际模型的准确率反而有所提高,最终模型的参数量只有原来的70%。

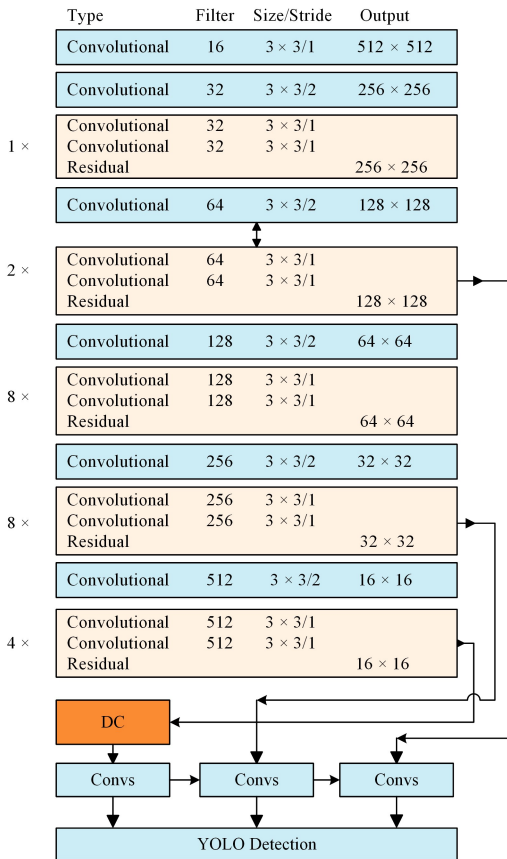


图7 检测算法的网络结构示意图

Fig. 7 Network structure of detection algorithm diagram

### 2.3.4 损失函数设计

损失函数是用于表征网络模型的预测值与真实值的误差

程度,是评估一个网络模型好坏的关键。在YOLOv3网络结构中,YOLOv3的损失函数可总结为:

$$L_{loss} = L_{box\_loss} + L_{confidence\_loss} + L_{class\_loss} \quad (1)$$

其中, $L_{box\_loss}$ 为边界框坐标预测误差, $L_{confidence\_loss}$ 为目标置信度误差, $L_{class\_loss}$ 为分类预测误差。具体计算公式为:

$$L_{loss} = \lambda_c \sum_{i=0}^{S^2} \sum_j^B I_{ij}^{obj} [(x_i - x_i^\wedge)^2 + (y_i - y_i^\wedge)^2] + \lambda_c \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [(\omega_i - \omega_i^\wedge)^2 + (h_i - h_i^\wedge)^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [c_i^\wedge \log(c_i) + (1 - c_i^\wedge) \log(1 - c_i)] + \lambda_{mobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{mobj} [c_i^\wedge \log(c_i) + (1 - c_i^\wedge) \log(1 - c_i)] + \sum_{i=0}^{S^2} \sum_{c \in C} I_{ij}^{obj} [\hat{p}_i(c) \log(p_i(c)) + (1 - \hat{p}_i(c)) \log(1 - p_i(c))] \quad (2)$$

其中, $\lambda_c$ 和 $\lambda_{mobj}$ 分别表示边界框坐标和未检测到目标的置信度误差权重; $S^2$ 为特征图单元格个数, $B$ 为每个单元格的预测边界框数, $c_i$ 为置信度误差, $C$ 为总分类数; $I_{ij}^{obj}$ 和 $I_{ij}^{mobj}$ 表示第 $i$ 行、 $j$ 列单元格是否存在被检测目标; $(x, y, \omega, h)$ 表示预测边界框的中心坐标、宽度和高度, $(x^\wedge, y^\wedge, \omega^\wedge, h^\wedge)$ 表示真实框的中心坐标、宽度和高度; $p_i$ 和 $p_i^\wedge$ 分别表示预测类别概率和实际类别概率。在实际训练过程中, $L_{loss}$ 越小,表明网络模型的鲁棒性越好。

在模型检测过程中,大尺度目标的预测框与真实框的细微偏差,可能并不影响检测效果,但对于小尺度目标而言,细微偏差可能会导致预测框与真实框在视觉效果上发生严重偏移,而且在边界框坐标损失计算中,YOLOv3采用了均方误差(MSE)损失函数,对于大尺度和小尺度目标产生的边框坐标损失值并没有严格区分。实际上,MSE计算对目标尺度非常敏感, $L_{box\_loss}$ 会随着目标尺度的增大而增大。与小尺度目标相比,大尺度目标预测边界框会产生更大的损失值,使得模型在训练过程中会倾向于大尺度目标的检测而忽略小尺度目标对坐标损失值的贡献。

因此,为平衡不同尺度的目标对坐标损失的贡献,本文对预测边界框的坐标损失值进行了加权,引入了真实框宽和高的乘积作为判断真实框尺度大小的依据,并以此构建了一个加权系数。改进后的边界框坐标损失函数为:

$$L_{box\_loss} = \lambda_c \sum_{i=0}^{S^2} \sum_j^B I_{ij}^{obj} \lambda_i [(x_i - x_i^\wedge)^2 + (y_i - y_i^\wedge)^2] + \lambda_c \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \lambda_i [(\omega_i - \omega_i^\wedge)^2 + (h_i - h_i^\wedge)^2] \quad (3)$$

$$\lambda_i = 2 * (1 - \omega_i^\wedge h_i^\wedge) \quad (4)$$

由式(3)、式(4)可知,改进后的边界框坐标损失函数平衡了不同尺度产生的坐标损失,使损失函数的计算更加合理,加快了模型的收敛速度,有效提高了模型的检测率,特别是小尺度目标的检测率。

### 2.3.5 优化先验框聚类

在目标检测中,先验框的数量和大小直接影响到最终模型的检测速度和准确率,因此如何选取合适的算法产生合适的先验框就显得尤为重要。YOLOv3使用K-means聚类算法为每种尺度的特征图设定了3种不同尺寸的先验框,但由于K-means聚类算法初始聚类中心是随机从训练集中产生的,一旦初始值选取不当,就可能无法得到有效的聚类结果。

为降低分类结果的最终误差,本文采用了改进的 K-means++ 对训练集上的边界框进行聚类。与 K-means 相比, K-means++ 受初始值的影响较小,其分类结果更加合理。

在计算先验框的过程中,为降低边界框大小带来的欧氏距离误差<sup>[22]</sup>,采用重叠度交并比 (IoU) 来替代 K-means++ 算法中的欧氏距离,其计算公式为:

$$D(box, centroid) = 1 - IoU(box, centroid) \quad (5)$$

其中,  $box$  为数据集的边界框,  $centroid$  为簇类中心,  $IoU$  为两个框的交并比。通过对数据集中的边界框进行聚类,可以得到平均交并比 ( $Avg\_IoU$ ) 与簇类个数 ( $K$ ) 的关系,其变化曲线如图 8 所示。

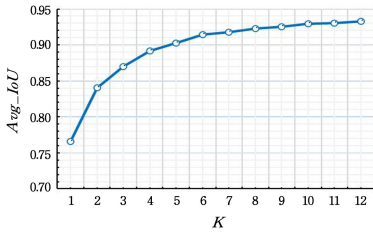


图 8 平均交并比变化曲线

Fig. 8 Avg\_IoU change curve

由图 8 可知,随着簇类个数  $K$  值的增加,平均交并比逐渐增大,当  $K=6$  时,平均交并比曲线开始趋于平滑,此时若增加簇类个数 ( $K>6$ ),会使聚类结果出现冗余,降低算法的检测速度。当先验框与真实框的交并比相同时,所需的先验框个数越少,模型的表达能力就越强,训练越容易。因此本文选取的簇类个数为 6。

### 3 实验结果分析

本文改进的算法是在开源的 Darknet 框架上实现的,训练所用的计算机 CPU 配置为 Intel xeon E5-2630 v4, GPU 型号为 TITAN xp, 显存为 12 GB, 搭配的操作系统为 Ubuntu 16.04; 测试所用的计算机 CPU 配置为 Intel i7-8750H, GPU 型号为 NVIDIA GeForce 1050Ti, 显存为 4 GB, 搭配的操作系统为 Ubuntu 18.04。

#### 3.1 训练方法

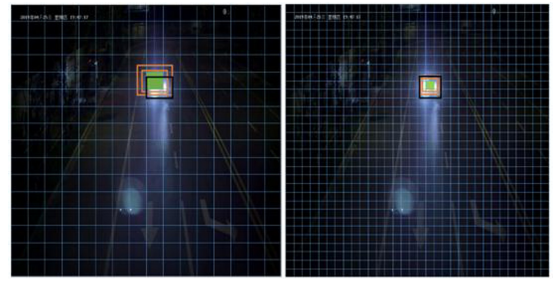
由于改进后的算法输出特征图尺度和先验框个数发生了变化,相应的先验框尺寸也要做出调整,因此实验之前必须先计算先验框尺寸。表 2 列出了本文在训练集上计算得出的 6 个先验框尺寸。

表 2 特征图与先验框

Table 2 Feature map and prior box

特征图	16×16	32×32	64×64
感受野	大	中	小
先验框	(40,38),(46,42)	(30,31),(35,34)	(20,20),(26,26)

分配上,随着特征图尺寸的增大,其具有的感受野逐渐减小,因此 16×16 特征图适合检测较大的对象,32×32 特征图适合检测中等大小的对象,而 64×64 特征图则适合检测较小的对象。图 9 为 16×16 和 32×32 特征图对应的先验框尺寸,其中绿色框为目标中心点所在的网格,黑色框为真实框,橙色框为先验框。



(a) 16×16

(b) 32×32

图 9 预测边界框示意图(电子版为彩色)

Fig. 9 Schematic diagram of prediction bounding box

在训练阶段,本文设置的实验参数如表 3 所列,在前 100 次迭代时,训练所采用的学习率为:

$$learning\_rate = learning\_rate * (iter/100)^2 + 10^{-6} \quad (6)$$

其中,初始  $learning\_rate$  为 0.001,  $iter$  为当前迭代次数。当训练迭代到 4000 次和 4800 次时,学习率在前一个学习率的基础上衰减十分之一。

表 3 相关实验参数

Table 3 Relevant experimental parameters

参数名	参数值
batch	64
width×height	416×416,512×512
momentum	0.9
decay	0.0005
max_batches	5500

#### 3.2 算法性能评价

##### 3.2.1 模型评价

为验证改进算法的性能,本文将改进后的网络命名为 YOLOv3-DC。对训练过程中的平均损失值和平均交并比进行记录,随着迭代次数的增加,对比的平均损失曲线和平均交并比曲线如图 10、图 11 所示。图 10 中横坐标为迭代次数,纵坐标为平均损失值;图 11 中横坐标为迭代次数,纵坐标为平均交并比值。

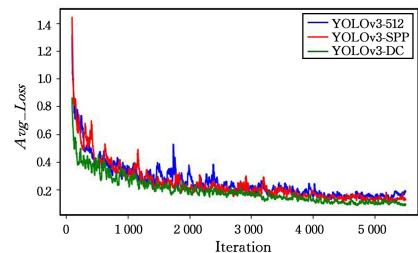


图 10 平均损失值曲线

Fig. 10 Average loss curve

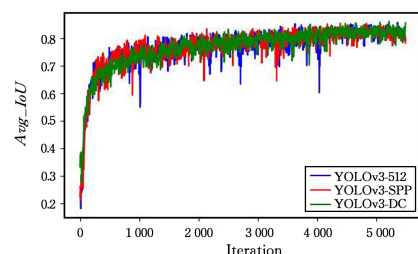


图 11 平均交并比曲线

Fig. 11 Average intersection ratio curve

由图 10 可以看出,随着迭代次数的增加,3 种算法的平均损失值逐渐下降,迭代了 4000 次之后开始趋于收敛,最终稳定在 0.10~0.20 之间,其中 YOLOv3-512 的平均损失值稳定在 0.13~0.14 之间, YOLOv3-SPP 稳定在 0.11~0.12 之间,而 YOLOv3-DC 稳定在 0.08~0.09 之间。这表明 YOLOv3-DC 算法相比于 YOLOv3-512 和 YOLOv3-SPP 算法效果更佳,在参数较少的情况下,能取得更低的平均损失值,而且算法的收敛速度更快,说明本文改进的网络模型是有效的。

由于远光灯区域的边界难以界定,算法的平均交并比并不高。如图 11 所示,随着迭代次数的增加,3 种算法的平均交并比均稳定在 0.8 左右。这说明本文优化先验框聚类对收敛后的平均交并比并无明显影响,反而加快了模型的收敛速度,使模型的训练变得更加容易。

### 3.2.2 与 YOLOv3 性能的对比如

本文将 YOLOv3, YOLOv3-SPP 和 YOLOv3-DC 算法的性能进行对比,使用测试集对算法性能进行评估,具体评价指标包含分类准确率(Precision, P)、召回率(Recall, R)、F1 分数(F1-score, F1)、IoU=0.5 时的平均准确率(mAP@50)以及单张图片的平均检测时间(MT)。

$$Precision = TP / (TP + FP) \tag{7}$$

$$Recall = TP / (TP + FN) \tag{8}$$

$$F1\text{-score} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{9}$$

式(7)中, TP 表示正确预测的正样本个数, FP 表示负样本被预测为正样本的个数。式(8)中, FN 表示正样本被预测为负样本的个数。测试实验结果如表 4 所列。

表 4 算法性能对比

Table 4 Comparison of algorithm performance

Net	TP	FP	FN	P/%	R/%	F1/%	mAP/%	MT/ms
YOLOv3-416	1018	67	189	93.82	84.34	88.83	92.86	54
YOLOv3-512	1099	52	108	95.48	91.05	93.21	96.56	61
YOLOv3-SPP-512	1095	20	112	98.21	90.72	94.32	98.06	62
YOLOv3-DC	1162	21	45	98.22	96.27	97.24	99.09	40

由表 4 可知,本文算法 YOLOv3-DC 在各个类别上的数值均优于 YOLOv3 及其改进算法,其平均准确率(mAP)达到了 99.09%,说明本文设计的 YOLOv3-DC 算法性能更好。与 YOLOv3-512 算法相比, YOLOv3-DC 单张图片的平均检测时间缩短了 21ms,而且对夜间车辆远光灯的检测能力也大大提高,其分类准确率、召回率和平均准确率分别比 YOLOv3-512 提高了 2.74%, 5.67% 和 2.53%。与 YOLOv3-SPP-512 算法相比,采用空洞卷积加强特征融合的 YOLOv3-DC 算法的召回率提高了 5.55%,平均准确率也提高了 1.03%。实验结果表明,本文设计的算法能够兼顾实时性和准确性,实现实时高效的车辆远光灯检测。

### 3.2.3 特征图对比

为了验证本文算法相比原 YOLOv3 算法具有更强的特征提取能力,本文提取了实验过程中 YOLOv3, YOLOv3-SPP 和 YOLOv3-DC 第一个残差模块得到的特征图进行分析,结果如图 12 所示。

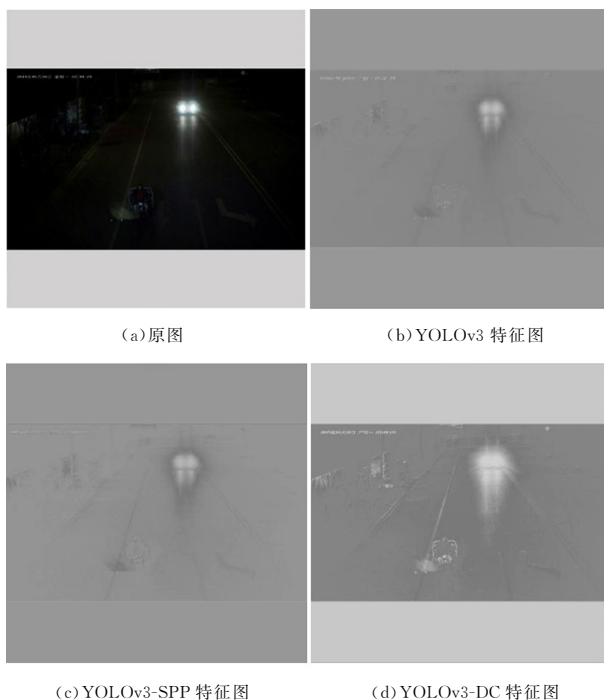


图 12 网络特征图对比示意图

Fig. 12 Comparison of network feature map

从图 12 可以看出,本文改进的算法对于特征的提取效果更好,提取的特征图更加清晰,说明本文改进的算法更有利于目标的检测。

### 3.2.4 实验结果分析

图 13 为 YOLOv3-DC 模型和 YOLOv3 模型在不同场景下的检测结果,其中第一列为原图,第二列为 YOLOv3 模型检测结果,第三列为 YOLOv3-DC 模型检测结果。

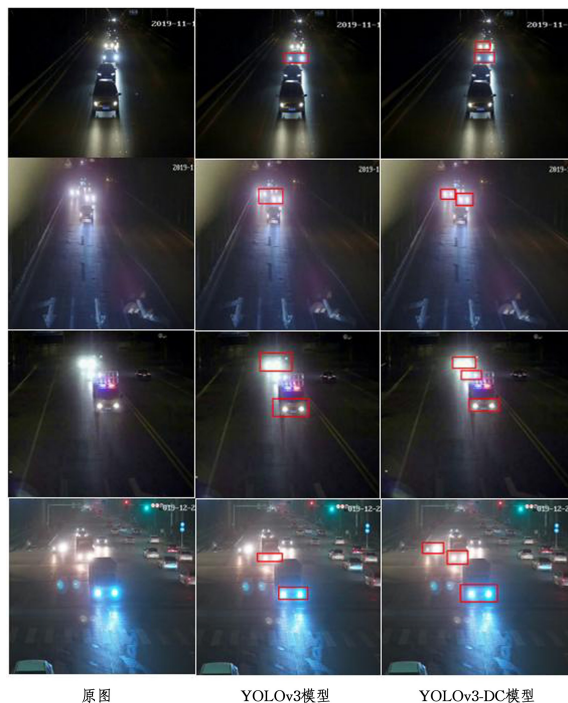


图 13 3 种模型在不同场景下的检测结果对比

Fig. 13 Comparison of detection results of three models in different scenarios

从对比结果来看, YOLOv3-DC 模型具有更好的鲁棒性, 在远光灯区域较小、远光灯骤亮、远光灯区域重叠以及多车道的场景中仍能做出正确的判断。

由于本文对不同尺度目标的坐标损失进行了加权, 虽然该方法在一定程度上提高了模型对小尺度目标的检测率, 但在一些目标样本局部被遮挡的场景下, 模型的边框回归效果并不是很好。如图 14 所示, 当车辆远光灯左(右)车灯被前车遮挡时, YOLOv3-DC 模型未能精确推理出车辆远光灯边框位置, 出现远光灯漏检的情况。而在一些噪声环境复杂(雨、雪天气等)的场景下, 由于地面反射光与车辆远光灯相似, 模型可能会将地面反射光误检为车辆远光灯, 如图 15(a)所示, 这也是车辆远光灯检测的难点问题; 对于远光灯光团的相互叠加而形成的超大光团, 模型也可能会出现漏检的情况, 如图 15(b)所示。

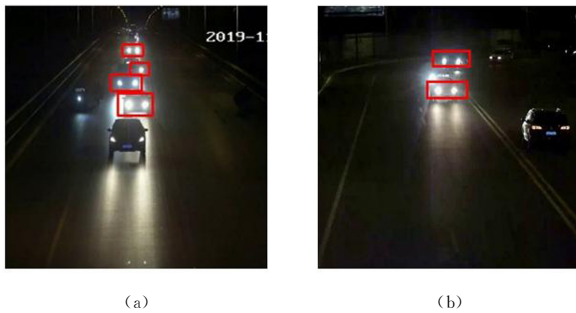


图 14 局部遮挡目标检测

Fig. 14 Partially-occluded objects detection

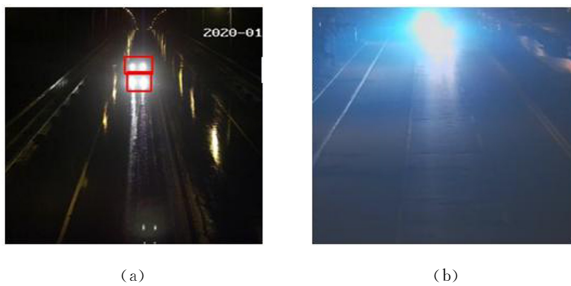


图 15 复杂环境下的误检和漏检

Fig. 15 False detection and missed detection in complex environment

上述出现的漏检与误检情况, 可以归结为夜间车辆远光灯困难样本的识别问题, 一方面可通过数据增强的方式来生成更多的样本, 以平衡简单样本与困难样本之间的比例; 另一方面, 也可通过 GHM<sup>[23]</sup> 等算法增大困难样本对梯度的贡献, 这也是我们后续研究的方向。而在现阶段应用中, 本文采用了持续跟踪检测的方法来解决该问题, 即对某个区域内的车辆灯光状态进行多次检测与验证, 以规避由偶然因素造成的远光灯漏检与误检情况。

目前基于本文改进 YOLOv3 模型的远光灯抓拍系统已经在国内几个城市开始试用。从实际测试效果来看, 本文改进的 YOLOv3 模型具有抓拍率高、适用范围广等特点, 图 16 为实际抓拍的评判结果界面, 其中车辆号码已屏蔽。



图 16 实际工作场景

Fig. 16 Practical work scene

### 3.2.5 与其他模型性能的对比

本文使用其他目标检测算法模型在相同数据集上进行训练和测试, 实验结果如表 5 所列。

表 5 不同算法性能的对比

Method	Backbone	$mAP@50/\%$	MT/ms
Faster R-CNN	VGG16	89.18	248
Faster R-CNN	ResNet101	90.35	289
R-FCN	ResNet101	90.68	376
SSD300	VGG16	87.51	27
SSD512	VGG16	90.72	68
FSSD300	VGG16	90.47	40
FSSD512	VGG16	90.84	92
YOLOv2-512	Darknet-19	90.43	29
YOLOv3-512	Darknet-53	96.56	54
RFBNet300	VGG16	89.33	42
RFBNet512	VGG16	94.60	94
YOLOv3-SPP-512	Darknet53+SPP	98.06	62
YOLOv4-512	CSPDarkNet53	98.39	65
YOLOv3-DC	—	<b>99.09</b>	<b>40</b>

由表 5 可以看出, 本文设计的模型 YOLOv3-DC 的综合性能更好。与基于区域的目标检测算法相比, 本文模型不仅在  $mAP@50$  值上优于基于 ResNet101 主干网的 Faster R-CNN 和 R-FCN 模型, 而且在检测速度上能达到实时检测的要求。与基于回归的目标检测算法相比, 本文模型与 2020 年 4 月开源的 YOLOv4<sup>[24]</sup> 模型的  $mAP$  值相近, 但参数规模和  $MT$  值明显优于 YOLOv4。而与 SSD 模型及其改进模型 FSSD<sup>[25]</sup>, RFBNet 相比, 本文模型在检测精度和速度上达到了很好的平衡。

**结束语** 为解决夜间车辆远光灯治理难的问题, 本文将深度学习算法应用到夜间车辆远光灯检测中。针对 YOLOv3 算法对单类目标检测速度慢、训练复杂和定位精度不高等问题, 本文通过增大特征图尺度的方式来保持目标信息的完整性, 通过降低各卷积层维数的方式来加快算法的运行速度, 并提出运用标准残差组件来替代 YOLOv3 残差组件和采用空洞卷积模块以加强局部特征与全局特征的融合, 在保证实时检测速度的前提下, 提高了夜间车辆远光灯的检测率。实验结果表明, 改进的算法取得了较好的效果, 具体体现在  $mAP$  值、分类准确率、召回率和平均交并比的提高, 可以满足智能交通的需要。但其仍存在提升空间, 下一步的研究方向是针对复杂场景下的样本进行更精确的识别, 以及对模型进行进一步的压缩以提高检测准确率和效率。

## 参 考 文 献

- [1] WU Y Q, HUA J F, CHEN Y, et al. Difficulties and Countermeasures for illegal use of driving beam headlamp of motor vehicles[J]. China Public Security (Academy Edition), 2018(1): 78-81.
- [2] GUAN Y. Look at the top ten “most offensive traffic habits” from a technical perspective [J]. Auto & Safety, 2014(7): 56-63.
- [3] SU J Y, ZHANG W W, WU X C, et al. The driving beam recognition of vehicle based on multi-structure feature extraction and trajectory tracking[J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(10): 103-110.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440-1448.
- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]//International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 91-99.
- [7] DAI J, LI Y, HE K, et al. R-FCN: Object detection via region-based fully convolutional networks[C]//Advances in Neural Information Processing Systems. 2016: 379-387.
- [8] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy, 2017: 2980-2988.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE Computer Society, 2017: 6517-6525.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017(99): 2999-3007.
- [13] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. arXiv:1804.02767, 2018.
- [14] LIU S, HUANG D, WANG Y. Receptive field block net for accurate and fast object detection[C]//Computer Vision-ECCV. 2018: 404-419.
- [15] DAI W C, JIN L X, LI G N, et al. Real-time airplane detection algorithm in remote-sensing images based on improved YOLOv3 [J]. Opto-Electronic Engineering, 2018, 45(12): 84-92.
- [16] ZHANG F K, YANG F, LI C. Fast Vehicle Detection Method Based on Improved YOLOv3 [J]. Computer Engineering and Applications, 2019, 55(2): 12-20.
- [17] ZHANG G S, GE G Y, ZHU R H, et al. Research of Gear Defect Detection Based on Improved YOLOv3 Network [J/OL]. <http://kns.cnki.net/kcms/detail/31.1690.TN.20191107.1704.010.html>.
- [18] REZATOFI G H, TSOI N, GWAK J Y, et al. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 658-666.
- [19] XU Y L, HE P. Yolov3 vehicle detection algorithm with improved loss function [J]. Information & Communications, 2019(12): 4-7.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, 2016: 770-778.
- [21] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[J]. arXiv:1511.07122, 2015.
- [22] CUI W L, WANG Y J, KANG S Q, et al. The detection method of highway lane line based on improved yolov3 algorithm [J/OL]. <https://doi.org/10.16383/j.aas.c190178>.
- [23] LI B, LIU Y, WANG X. Gradient Harmonized Single-stage Detector[J]. arXiv:1811.05181, 2018.
- [24] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. arXiv:2004.10934, 2020.
- [25] LI Z X, ZHOU F Q. FSSD: feature fusion single shot multibox detector[J]. arXiv:1712.00960, 2017.



**GONG Hang**, born in 1995, postgraduate, is a member of China Computer Federation. His main research interests include information security and object detection.



**LIU Pei-Shun**, born in 1975, Ph.D, lecturer, is a member of China Computer Federation. His main research interests include information security and deep learning.