

文章编号: 2095-2163(2021)06-0056-07

中图分类号: TP391.4

文献标志码: A

基于改进先验框和损失的交通标志多尺度检测

李丝缘, 张荣芬, 刘宇红

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 针对交通检测中存在的交通标志尺寸较小、距离较远、实际道路场景复杂等问题, 提出一种基于改进先验框和目标定位损失的交通标志多尺度检测算法。以 darknet-62 为特征提取网络, 针对小型交通标志检测, 融合基于 FPN 的五尺度预测网络, 进一步提升输出特征层的分辨率和丰富语义信息; 改进先验框设计方法, 以更精准地检测位于特征图网格拐角处或相邻的小型交通标志; 通过引入基于 GIoU 指标的目标定位损失函数, 以提高网络的整体识别精度。本文算法在中国交通标志违规概率 top30 数据集取得了 43.6% 的 *mAP* 和 14 的 *FPS*, 相较于现有主流算法收获了更精准的检测效果。

关键词: 交通标志检测; 改进先验框; 五尺度预测; GIoU

Multi scale traffic sign detection based on improved prior frame and loss

LI Sitao, ZHANG Rongfen, LIU Yuhong

(School of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] Aiming at the problems of small size, long distance and complex road scene in traffic sign detection, a multi-scale traffic sign detection algorithm based on improved prior frame and target location loss is proposed. Taking darknet-62 as the feature extraction network, aiming at small traffic sign detection, the five scale prediction network based on FPN is integrated to further improve the resolution of output feature layer and rich semantic information, and the design method of a prior bounding box is improved to more accurately detect small traffic signs located at the corner or adjacent to the grid of feature map. The loss function of target location based on GIoU index is introduced to improve the overall recognition accuracy of the network. This algorithm achieves 43.6% *mAP* and 14 *FPS* in the top 30 highest violation probability data set of traffic signs in China, which is more accurate than the existing mainstream algorithms.

[Key words] traffic sign detection; improved prior bounding box; five scale prediction; GIoU

0 引言

随着交通问题的日益严峻, 智能化交通的发展是当今社会发展的一大趋势, 交通标志检测与识别作为计算机视觉中一个重要的领域, 是智能化交通发展中的重要一环。由于自然场景的复杂多变性及交通标志本身在尺寸与距离上的特殊性, 增加了目标检测和识别的难度。

深度学习的快速发展为交通标志识别带来了新的思路, 相比于过去基于颜色或形状的检测方法^[1], 基于卷积神经网络(CNN)的检测方法能够对交通标志进行准确的定位和分类。2012年 Alex 等人提出了 AlexNet 网络^[2], 在 ImageNet 竞赛中获得冠军。该网络使用双 GPU 并行处理, 并采用卷积加全连接的网络结构对目标进行检测, 在交通标志类

的大规模数据集的识别过程中, 该网络 top-5 错误率仅为 15.3%; 2015年 Redmon J^[3]提出了用于目标检测的 YOLOv3 网络模型, 其在交通标志检测与识别领域有着很宽泛的应用, 通过多尺度单阶段对交通标志进行检测; Mingyu Gao^[4]等人提出了基于高斯混合模型和类别聚焦损失的交通标志目标检测网络, 并将类别比例因子纳入网络中, 在提高识别精度的同时以解决数据集类别不平衡的问题; Chang Sun^[5]等人采用单一对象检测框架, 并构建稠密链接传输模块对低像素交通标志图像进行检测并获得上下文信息, 从而进行更为准确的识别。目标联合交集(IoU)是目标检测中最为常用的评估指标之一, Hamid Rezaatofghi^[6]等人提出了通用 IoU, 并将这一指标以损失函数的改进方式添加到经典的目标检测算法(例如 Fast R-CNN、Mask R-CNN)中, 改

基金项目: 贵州省科技计划项目(黔科合平台人才[2016]5707)。

作者简介: 李丝缘(1994-), 女, 硕士研究生, 主要研究方向: 深度学习; 张荣芬(1977-), 女, 博士, 教授, 主要研究方向: 嵌入式通信系统、云计算与大数据应用; 刘宇红(1963-), 男, 硕士, 教授, 主要研究方向: 嵌入式通信系统、云计算与大数据应用。

通讯作者: 刘宇红 Email: 1693623574qq.com

收稿日期: 2021-01-25

进后算法在检测与识别精度上都获得了显著的提升;Khaled Bayoudn^[7]等人提出了基于迁移学习的混合 2D-3D 的 CNN 模型,并通过上采样和反卷积的操作组成空间语义检测的架构,从而提升交通标志检测及识别的精度。

本文以 darknet-62 为特征提取网络,以基于 FPN 的五尺度预测网络、先验框中心定位、基于 *GIoU* 的损失函数为改进点,用以提高在交通标志检测上的精确度。

1 模型分析

交通标志检测模型由特征提取网络 Darknet62 和五尺度预测网络组成,并在最浅输出特征层添加先验框改进设计方法,改进后的模型结构如图 1 所示。基于 FPN 的五尺度预测网络获取丰富的上下文信息,以检测尺寸较小的交通标志;改进的先验框设计从特征图网格角度,解决与小目标检测有关的预测框定位问题;通过引入 *GIoU* 作为目标定位损

失函数,优化常用的目标定位损失与最大化目标检测评估指标 *IoU* 值之间的偏差问题,进一步提高网络的整体识别精度。

Darknet62 中,在卷积层后添加批次归一化 (Batch Normalization) 层^[8]和 Mish 激活函数,防止过度拟合并增加网络的非线性表达能力,以 CBM(卷积+批次归一化+Mish 激活函数)卷积作为 Darknet62 的基本单位;为了增加特征提取网络的深度,同时避免训练过程梯度爆炸和梯度消失问题,使用 1×1 CBM 卷积和 3×3 CBM 卷积执行残差连接^[9],以形成 res 残差单元结构,并将 resN 残差块结构设计为由 3×3 CBM 卷积和 N res 残差单元串联连接。Darknet62 首先将输入图像的大小调整为 512×512×3,然后使用 3×3 CBM 卷积来过滤输入图像,依次使用 res1、res2、res8、res8、res4、res4 对特征图进行下采样,获得的 128×128×64、64×64×128、32×32×256、16×16×512、8×8×1024 这 5 个特征图层将通过特征融合^[10],输出 5 个不同比例的特征图,用于五尺度预测。

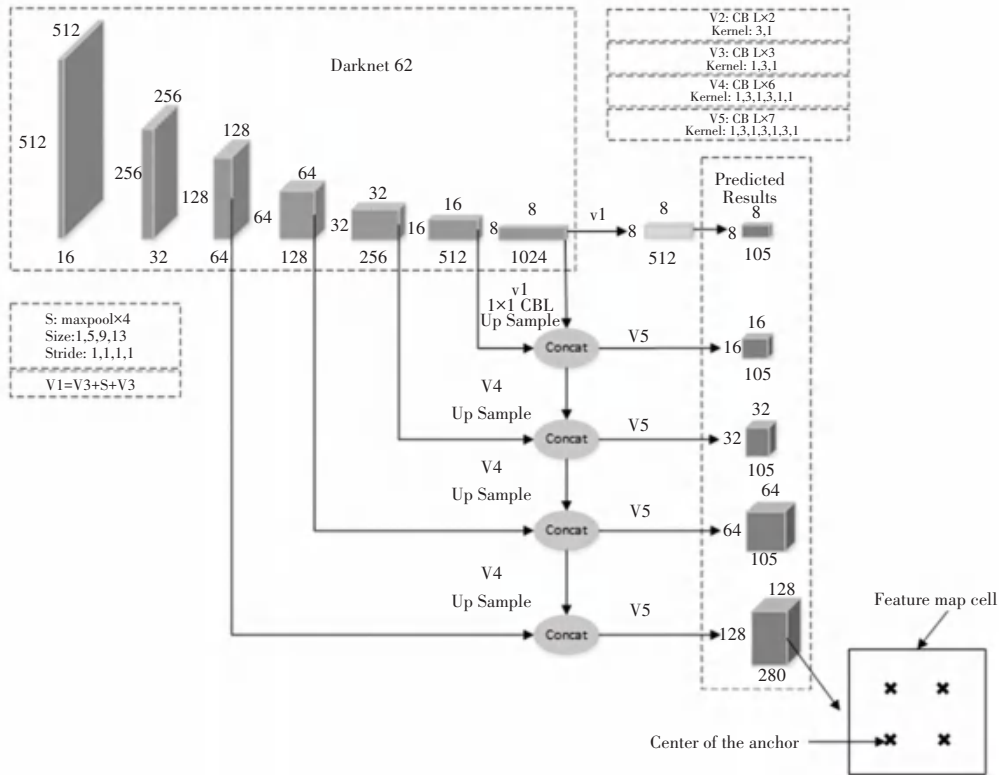


图 1 交通标志检测网络模型结构

Fig. 1 Structure of traffic sign detection network model

1.1 基于 FPN 的五尺度预测网络

浅层特征图层的特征语义信息较少,不利于检测目标的分类,但像素丰富、分辨率高、目标信息准确;高深层特征图层特征语义信息比较丰富,但像素较少、分辨率低、目标位置不准确。特征金字塔网

络 (Feature Pyramid Networks, FPN) 通过将低分辨率但强语义深层特征图上采样与高分辨率但弱语义浅层特征图融合^[11],构建在所有层次上共享丰富语义的特征金字塔,极大提高了具有高分辨率、用以检测小目标的浅层输出特征图的目标检测精度。

在五尺度预测网络中,由卷积层、批次归一化和线性整流 ReLU 激活函数组成 CBL (卷积+批次归一化+ReLU 激活函数) 卷积,作为五尺度预测网络的基本单元。同时,使用 CBL 卷积形成 res 残差单元和 resN 残差块结构。其次,为了增加预测网络中的感受野,在 Darknet62 之后添加了 SPP 块(由 4 个具有不同比例(1、5、9、13)的 maxpool 并行构成并将输出串联在一起),SPP 块在不影响网络识别速度的同时,可以显著分离出最重要的上下文特征,并提高网络识别精度。在 SPP 块前后添加 3 个 CBL 卷积(V3: CBL×3 Kernel: 1, 3, 1)用以更好地匹配要素图的大小和通道数,并替换了基础网络的 5 个 CBL 卷积。

为了充分利用网络中浅层特征所包含的小目标像素信息,更好检测图像中像素较小的交通标志。如图 1 所示,本文基于 FPN 思想,将 Darknet62 的 $8 \times 8 \times 1024$ 特征图经过 V1(CBL×3 + maxpool×4 + CBL×3)和 V2(CBL×2)操作,获得 $8 \times 8 \times 105$ 的输出特征图;另将 $8 \times 8 \times 1024$ 特征图,通过 V1、 1×1 CBL 卷积和上采样(Up Sample)操作获得新的特征图,并与 $16 \times 16 \times 512$ 特征图进行特征融合。融合后的新图层可以通过 V5(CBL×7)操作获得 $16 \times 16 \times 105$ 的输出特征图,通过 V4(CBL×6)和上采样(Up Sample)操作进一步与更浅层特征图 $32 \times 32 \times 256$ 特征融合。以此类推,分别在 Darknet62 的其余 $64 \times 64 \times 128$ 、 $128 \times 128 \times 64$ 特征图上执行类似操作。最终,预测网络可以获得五尺度预测结果,分别为 $128 \times 128 \times 280$ 、 $64 \times 64 \times 210$ 、 $32 \times 32 \times 105$ 、 $16 \times 16 \times 105$ 和 $8 \times 8 \times 105$ 。本次改进将结合丰富语义信息的输出图层,分辨率提高到了 128×128 ,因此预测网络对于图像中的较小较远的交通标志将具有更好的性能。

在前向过程中,预测网络执行边界框回归过程,将先验框(anchors)转换为相应的预测边界框(predicted bounding box),依次通过设定的阈值和非

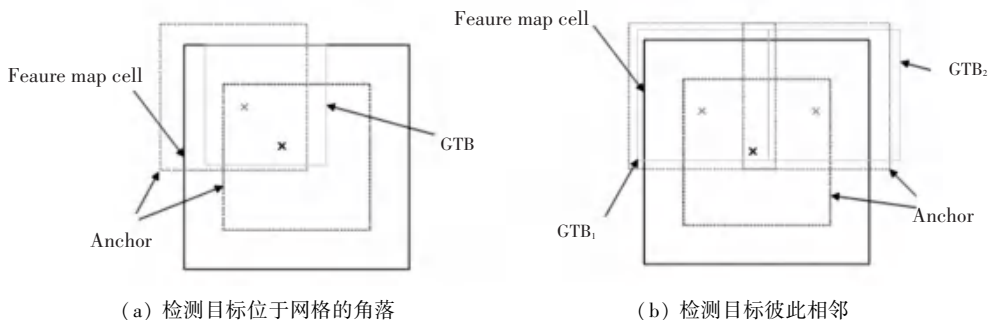
极大值抑制(NMS)算法删除不符合的预测边界框^[12]。最后,在图像或视频的相应位置标记,并输出预测框和交通标志的检测类别。

1.2 先验框改进设计方法

先验框是一组预设的边框,用于定位输入图像中的目标对象,一个先验框可以由边框的纵横比和边框的面积(尺度)来定义,相当于一系列预设边框的生成规则。在图像划分的每一个网格(grid cell),生成一系列的边框,需要在训练过程之前进行设计,通常使用 K 均值(K-means)聚类得到先验框形状和尺寸。对于每个真实标注框(GTB),将选择具有最大交并比(IoU)值,且 IoU 值高于阈值(通常为 0.5)的先验框作为匹配项。

通过研究发现,为了更好适用于小目标检测,原始的先验框设计方法并不是最佳的。因为所有先验框的中心通常固定于每个网格的中心,即相对坐标(0.5, 0.5)。对于具有较小尺寸并且位于彼此相邻的两个网格或小型对象交界处的交通标志,可能导致小目标对象与 GTB 丢失匹配,难以检测精准^[13]。

图 2(a)、(b)中的实线框 feature map cell 为最浅输出特征图(128×128)中的一个网格,居中虚线框代表未改进的先验框,其中心与网格中心重合(0.5, 0.5),角落虚线框代表改进后的先验框。在图 2(a)中检测目标位于网格的角落,改进框的中心坐标为(0.25, 0.75),明显可以看出位于网格角落的小交通标志(GTB)和改进框之间的 IoU 要高于同一 GTB 和未改进框之间的 IoU,因此改进先验框更为适合;在图 2(b)中检测目标彼此相邻,改进框的中心坐标分为(0.25, 0.75)和(0.75, 0.75),这类情况下,改进框可以同时匹配两个彼此相邻的小型交通标志(GTB1 和 GTB2),而未改进框只能匹配其中一个,且匹配度不高。



(a) The detection target is located in the corner of the grid (b) The detection target is adjacent to each other

图 2 原始先验框设计的局限性

Fig. 2 The limitation of original prior bounding box design

基于上述研究, 本文对先验框的设计方法做出以下改进:

(1) 应用 k 均值聚类获得先验框形状

原始网络配置中, 聚类获得的先验框形状适合于具有不同纵横比的 GTB(针对 VOC 和 COCO 数据集进行聚类), 但现实世界中的交通标志通常具有相似的纵横比。为了更适用于本文交通标志检测, 提高目标检测精度, 通过对训练集的真实标注框进行 K 均值聚类, 获得新的先验框形状。其中, k 设置为 14(2 组尺度分配给最浅输出特征图, 其余 4 层输出特征图每层分配 3 组尺度)。

(2) 确定先验框中心坐标

改进模型将输出 5 个不同尺度的特征图作为预测(五尺度预测网络)。在步骤 1 中获得的先验框形状对应于目标对象的不同比例。其中最浅输出特征图(128×128)使用小型先验框, 其具有小的感受野, 针对小目标具有更好的检测效果。将两组先验框尺度应用于该特征图上, 并将先验框的中心坐标分别设置为每个网格相对位置(0.25, 0.25)、(0.25, 0.75)、(0.75, 0.25)和(0.75, 0.75), 使最浅的特征图网格的先验框总数为 $2 \times 4 = 8$, 其余 4 层输出特征图先验框中心位置保持(0.5, 0.5)不变, 每个网格分配 3 个尺度的先验框。

1.3 改进目标定位损失函数

IoU (交并比, Intersection over Union), 是比较 2 个任意形状之间相似度的最常用指标^[13-14]。通过对目标的形状属性进行编码(例如, 将 2 个比较的边界框宽度、高度、位置放到 region 属性中。), 然后计算归一化度量。在目标检测中, 通常使用预测框与真实标注框 IoU 评价模型检测性能。而目前常用的目标定位损失函数则是预测框与真实标注框角点坐标距离的 1-范数 Loss 或者 2-范数 Loss。

通常在优化常用的目标定位损失与最大化 IoU 值之间是存在差距的。如图 3(a), 深色框代表预测框, 浅色框代表真实标注框。由于这 3 个预测框的第二角都位于以真实标注框的第二角为中心的固定半径圆上, 因此具有相同的 2-范数 Loss, 但是其 IoU 值却各不相同, 显而易见 IoU 值最高的第 3 个检测结果精确度更好; 同理, 如图 3(b), 当 1-范数 Loss 相同时, IoU 值和检测效果也各不相同。直观的是, 针对 1-范数 Loss 和 2-范数 Loss 的良好局部最优不一定是评价指标 IoU 的局部最优, 而要使检测性能最优选择就是指标本身^[15]。

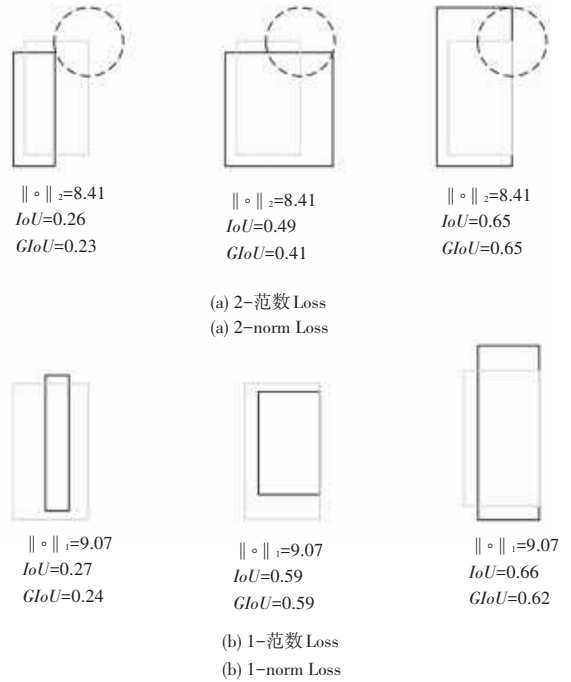


图 3 相同范数 Loss 的不同 IoU 和 GIoU

Fig. 3 Different IoU and GIoU with the same norm loss

但是 IoU 作为损失函数仍然存在 2 个问题:

- (1) 如果真实标注框与预测框不重叠, 则 IoU 值将为 0, 并且反映两个框之间的距离。如果将 IoU 用作 Loss, 则其梯度将为零且无法优化。
- (2) IoU 对于预测框与真实标注框之间不同的对齐方式, 是无法进行辨认的。

基于以上问题, 本文通过引入 $GIoU$ ^[16](Generalized Intersection over Union, 通用 IoU) 作为损失函数, 以解决 IoU 存在的非重叠问题。 $GIoU$ 的计算过程如下:

- (1) 对于预测框 A 和真实标注框 B , 找到最小的包围凸对象 C , 并将 A 和 B 都包围起来。
- (2) 计算 IoU :

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

- (3) 以 C 占用的面积(不包括 A 和 B) 的部分为分子, 以 C 占用的总面积为分母, 代表着重于 A 和 B 之间的空白部分的标准化度量。最后, 通过从 IoU 值中减去此比率来获得 $GIoU$:

$$GIoU = IoU - \frac{|C - (A \cup B)|}{|C|} \quad (2)$$

- (4) 与 IoU 类似, $GIoU$ 作为距离度量^[17], loss 计算公式如下:

$$L_{GIoU} = 1 - GIoU \quad (3)$$

$GIoU$ 具有以下属性:

(1) 具有与 IoU 相同的定义,即将比较对象的形状属性编码为 $region$ 属性。

(2) 保持了 IoU 的尺度不变性,以及对目标大小的不敏感性。

(3) $GIoU$ 始终作为 IoU 的下界,即 $\forall A, B, GIoU_{(A, B)} \leq IoU_{(A, B)}$ 。同时在预测框与真实标注框重叠度增加的情况下,确保了与 IoU 的强相关性。即 $\lim_{A \rightarrow B} GIoU_{(A, B)} = IoU_{(A, B)}$,在 A 和 B 形状完全重合时,则有 $GIoU = IoU = 1$ 。

(4) 由于引入了最小包围凸对象 C ,即使 A, B 不重合时, L_{GIoU} 依然可以进行优化。当 A 和 B 不重合度越高, $GIoU$ 越趋近于 -1 。

总之, $GIoU$ 既保留了 IoU 的主要特性,同时解决了 IoU 存在的问题。

2 实验与结果分析

2.1 交通标志数据集

本文采用的数据集为中国交通标志违规概率 top30 数据集,由 Gao^[4] 等人整理贡献。该数据集包含 Go Straight Slot、No Entry、Bus-only Lane、Speed Limit 120、Left Turn Slot、No Trucks、One-way Road 等 30 类交通标志,共 10 000 张图片,其中每张包含 1~6 个交通标志。在数据集中,将划分 7 000 个样本作为训练集,3 000 个样本作为测试集。

2.2 实验环境和参数设置

硬件环境:CPU 为 Intel® Core™ i7-7800X 6 核,GPU 为 GTX1080Ti,显存大小 11 GB,运行内存 64 GB。

软件环境:操作系统 Ubuntu16.04、OpenCV3.4.1、Tensorflow1.6.0、Python 3.6、Keras 2.1.5。

使用 k -means++ 聚类算法对交通标志数据集的真实标注框进行聚类操作,得到对应 5 个输出特征图的 14 组 anchor,尺寸分别为:

128×128 : (9×10) , (17×20) ; 64×64 : (28×32) , (33×34) , (36×48) ; 32×32 : (48×104) , (53×61) , (65×68) ; 16×16 : (87×110) , (76×82) , (100×153) ; 8×8 : (113×134) , (162×192) , (368×389) 。

此外,训练阶段通过更改网络配置文件参数,对训练集进行旋转、饱和度调整、曝光度调整、色调调整、噪声调整等数据增强预处理,在原有数据集基础上生成更多训练样本,提高模型泛化能力。

基于迁移学习的方法,使用在 coco 数据集上的 darknet-62 预训练模型作为基础特征提取网络。

输入图像大小调整为 512×512 ,图像通道数为

3。在迭代过程中,batchsize 设置为 16,优化器选择 SGD 算法^[18],初始学习率设置为 0.001,动量和权重衰减分别配置为 0.9 和 0.000 5,最大迭代次数设置为 50 200,并在迭代次数达到 40 000 和 45 000 时,学习率分别更改为 0.000 1 和 0.000 01。在训练阶段,网络参数会迭代更新,直到迭代次数达到预设值,或者损失函数的变化小于阈值,获得最终训练模型。

2.3 实验结果分析

为了有效说明算法对交通标志检测的性能优势,将本文算法与几种最新交通标志识别算法在相同条件下进行了性能比较实验。评价指标使用模型大小(Size),每秒帧数(FPS);判断网络识别速度的指标,其值越大,网络的识别速度越快;平均精度均值(mAP, mean average precision);30 类交通标志平均精度(AP)的平均值;平均精确率(AP, average precision);每一类别被正确检测出来的交通标志占全部检出区域的百分比, AP_S 、 AP_M 、 AP_L 分别代表小型交通标志、中型交通标志、大型交通标志的平均精确率。实验结果见表 1。其中,方法 A-F 分别代表基于改良的 FasterR-CNN^[19]、SSD^[20]、RefineDet^[21]、RetinaNet50^[22],基于 GMM 和类别质量焦点损失优化算法^[4]及基于 YOLOv3 交通标志识别算法^[23]。以加粗字体代表最佳结果。

表 1 不同交通标志检测算法的性能对比

Tab. 1 Objective evaluation of the performance of different traffic sign recognition algorithms

Methed	Size	FPS	mAP	AP_S	AP_M	AP_L
A ^[19]	109.3M	4	39.1%	18.4%	43.2%	50.0%
B ^[20]	100.4M	9	29.4%	10.5%	34.1%	43.6%
C ^[21]	78.1M	8	33.2%	15.4%	38.3%	44.9%
D ^[22]	145.6M	6	33.6%	14.4%	37.1%	47.8%
E ^[4]	245.3M	15	40.1%	24.1%	43.1%	52.2%
F ^[23]	237.4M	16	34.7%	18.6%	39.7%	43.8%
ours	240.6M	14	43.6%	29.3%	45.4%	53.3%

从表 1 中可以看出,与两阶段算法改进的 Faster R-CNN(A)相比,本文算法的检测速率高于其 3 倍多,同时在 mAP、 AP_S 、 AP_M 、 AP_L 均取得更好的检测效果。与一阶段算法中表现较优的 YOLOv3(F)和基于 GMM 和类别质量焦点损失优化算法(E)相比,本文算法在模型尺寸和检测速率上略占劣势。但本文算法通过选用相较于 YOLOv3 的特征提取网络层数更深的 darknet-62 网络,结合损失函数改进提升了模型的整体检测精度,同时融合基于 FPN 的五尺度预测网络、改进先验框设计,显著优

化了小型交通标志的检测性能,在 mAP 、 AP_s 、 AP_M 、 AP_L 相较于检测精度第二的结果分别提高了 3.5%、5.2%、2.2%、1.1%,其中小型交通标志检测性能提升显著。

本文算法与改良的 FasterR-CNN 算法(A)、基于 GMM 和类别质量焦点损失优化算法(E)的检测效果对比如图 4 所示。在背光环境下的图 4(a)~(c)中,本文算法能够较为精准的检测出不同远近的交通标志,而 A 算法漏检了远处较暗的 Strictly No

Parking 交通标志,E 算法检测精度较低,对于较远且相邻的小交通标志的定位也不够精准;在遮挡环境下图 4(d)~(f)中,本文算法对部分遮挡的交通标志识别置信度达 0.86,而 A 算法和 E 算法仅有 0.62 和 0.74;在多目标复杂环境下图 4(g)~(i)中,本文算法能将图中多达 6 个不同尺寸和角度交通标志无遗漏检测,而 A 算法和 E 算法均存在不同程度的漏检和置信度降低。

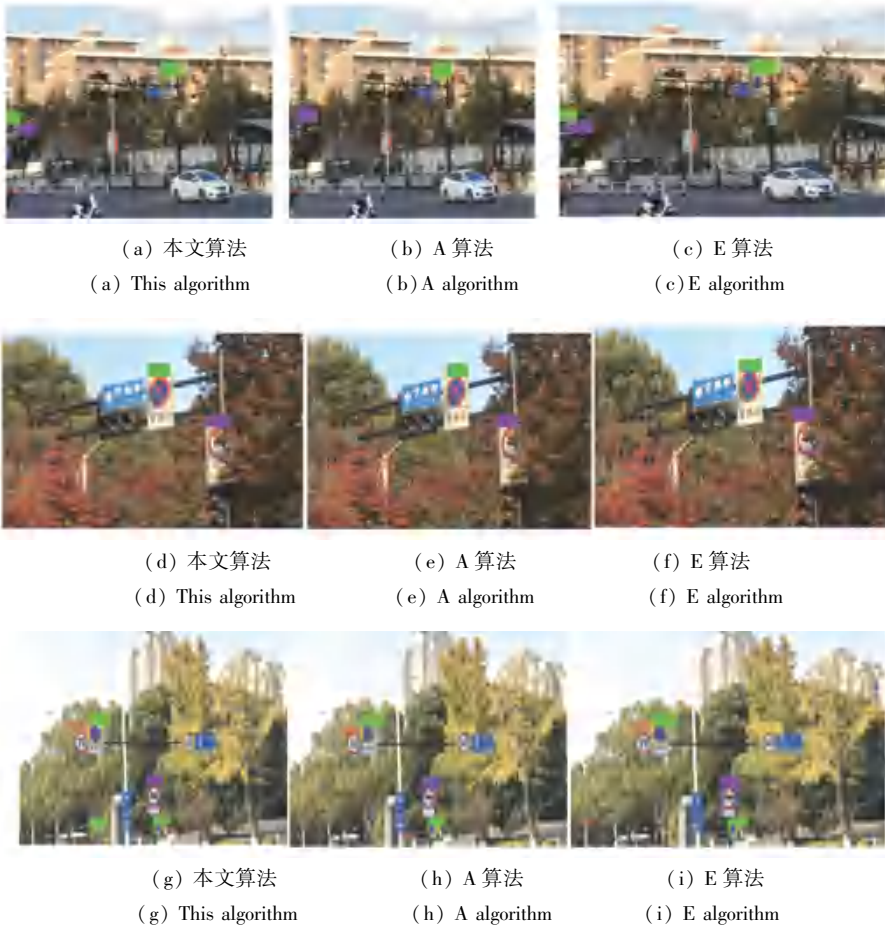


图 4 各算法检测效果对比

Fig. 4 Comparison of detection effect of each algorithm

3 结束语

本文融合基于 FPN 的五尺度预测网络,将输出特征层的分辨率提高到 128×128 ,为小型交通标志生成丰富的上下文信息,改进了先验框设计方法,更精准的检测位于特征图单元拐角处的小型交通标志或相邻的小型交通标志。由于度量的最佳损失是度量本身,通过引入 GIoU-Loss 用作损失函数,提升模型整体检测精度。结合上述改进,本文相比较最新

算法,在检测小型交通标志(<322 像素)方面具有显著提升效果,并且在平均检测精度、中型(322 像素 $<$ 中 <962 像素)和大型交通标志(962 像素 $<$ 大 $<4\ 002$ 像素)方面也获得了更好的性能。但是由于模型参数较多,仍然占用过多空间,同时图像检测时间仍有待提高,如何在保持甚至进一步优化检测精度的基础上,简化网络结构、降低计算量将是未来的主要研究方向。

参考文献

- [1] Intra color-shape classification for traffic sign recognition [C]// International Computer Symposium, 2010:642-647.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[C]//NIPS, 2012.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2015:779-788.
- [4] GAO M, CHEN C, SHI J, et al. A Multiscale Recognition Method for the Optimization of Traffic Signs Using GMM and Category Quality Focal Loss [J]. Sensors, 2020, 20(17):4850.
- [5] SUN C, AI Y, WANG S, et al. Dense-RefineDet for Traffic Sign Detection and Classification [J]. Sensors, 2020, 20, 6570.
- [6] Rezatofighi, Hamid Tsoi, Nathan Gwak. Generalized Intersection over Union: A Metric and A Loss for Bounding Box [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
- [7] BAYOUDH K, HAMDAOUI F, MTIBAA A. Transfer learning based hybrid 2D-3D CNN for traffic sign recognition and semantic road detection applied in advanced driver assistance systems [J]. Applied Intelligence, 2020(1).
- [8] SERGEY I, CHRISTIAN S. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J]. arXiv 2015, arXiv:1502.03167.
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [10] ZHANG J, JIN X, SUN J, et al. Spatial and semantic convolutional features for robust visual object tracking [J]. Multimedia Tools and Applications, 2018:1-21.
- [11] LUO Y, QIN J, XIANG X, et al. Coverless real-time image information hiding based on image block matching and dense convolutional network [J]. J. Real-Time Image Process. 2019, 17, 125-135.
- [12] LIU Z, HU J, WENG L, et al. Rotated region based CNN for ship detection [C]//International Conference on Image Processing (ICIP), 2017:900-904.
- [13] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (voc) challenge [J]. International Journal of Computer Vision, 2010, 88(2):303-338.
- [14] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C]// European conference on computer vision, Springer, 2014:740-755.
- [15] YU J, JIANG Y, WANG Z, et al. Unitbox: An advanced object detection network [C]//2016 ACM on Multimedia Conference, ACM, 2016:516-520.
- [16] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2019:658-666.
- [17] KOSUB S. A note on the triangle inequality for the jaccard distance [J]. arXiv preprint arXiv:1612.02696, 2016.
- [18] PARAS. Stochastic gradient descent [J]. Optimization, 2014.
- [19] LIN Xiu, WU Houjie, et al. Traffic sign detection method based on Faster R-CNN [C]// 2018 计算机科学与工程国际研讨会. 2018.
- [20] JIN Y, FU Y, WANG W, et al. Multi-Feature Fusion and Enhancement Single Shot Detector for Traffic Sign Recognition [J]. IEEE Access 2020, 8, 38931-38940.
- [21] ZHANG S, WEN L, BIAN X, et al. Single-shot refinement neural network for object detection [C]//Conference on Computer Vision and Pattern Recognition (CVPR), 2018; 4203-4212.
- [22] RAJENDRAN S P, SHINE L, PRADEEP R, et al. Fast and Accurate Traffic Sign Recognition for Self Driving Cars using RetinaNet based Detector [C]//2019 International Conference on Communication and Electronics Systems, 2019; 784-790.
- [23] REDMON J, FARHADI A. YOLOv3: an incremental improvement [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- (上接第55页)
- [2] 易靖国, 程江华, 库锡树. 视觉手势识别综述 [J]. 计算机科学, 2016, 43(S1): 103-108.
- [3] 赵宝琚. 中小學生握筆姿勢問題的調查研究 [J]. 現代教育科學, 2014 (6): 147-148.
- [4] 薛俊韜, 縱蘊瑞, 楊正瓚. 基於改進的 YCbCr 空間及多特徵融合的手勢識別 [J]. 計算機應用與軟件, 2016, 33(1): 151-155.
- [5] 謝鋒桂. 基於改進的卷積神經網絡的手勢識別的研究 [J]. 計算機應用與軟件, 2019, 36(3): 192-195.
- [6] CAO Z, SIMON T, WEI S E, et al. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 1302-1310.
- [7] MAZHAR O, NAVARRO B, RAMDANI S, et al. A real-time human-robot interaction framework with robust background invariant hand gesture detection [J]. Robotics and Computer-Integrated Manufacturing, 2019, 60:34-48.
- [8] WU, LIONEL, PIGOU, et al. Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition [J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(8):1583-1597.
- [9] 郭雷. 手勢識別中手分割算法 [J]. 軟件, 2015(4):81-83.
- [10] 程文山. 基於肤色分割和 Camshift 的手勢識別研究 [D]. 武漢: 華中師範大學, 2009.
- [11] PINTO R F, BORGES C, ALMEIDA A, et al. Static Hand Gesture Recognition Based on Convolutional Neural Networks [J]. Journal of Electrical and Computer Engineering, 2019, 2019:1-12.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks [C]// NIPS 2012: 1106-1114.
- [13] 郭書杰. 基於改進 AlexNet 的複雜背景手勢識別方法研究 [J]. 大連交通大學學報, 2020, 41(6):97-101.