

基于改进 YOLOv2 模型的多目标识别方法

李珣¹, 时斌斌^{1*}, 刘洋², 张蕾¹, 王晓华¹

¹西安工程大学电子信息学院, 陕西 西安 710048;

²西安计量技术研究院, 陕西 西安 710068

摘要 在 YOLOv2 算法的基础上, 根据实际道路环境的变化对 YOLOv2-voc 的网络结构进行改进, 基于 ImageNet 数据集和微调技术得到分类训练网络模型, 根据训练结果与车辆目标特征的分析, 对算法参数进行修改, 获得改进的车型识别分类网络结构模型 YOLOv2-voc_mul。为验证所提模型的有效性, 分别对简单背景和复杂背景下的样本进行检测, 并与 YOLOv2、YOLOv2-voc 和 YOLOv3 模型在迭代 70000 次后的检测结果进行了对比。实验结果表明: 在简单背景下, YOLOv2-voc_mul 模型的精度可达 99.20%, 不同车型的平均精度均值达到了 89.03%; 在复杂背景下, YOLOv2-voc_mul 模型对 4 种车型在单目标和多目标的检测下平均准确率达到了 92.21% 和 89.44%, 具有较高的精确度、较小的误检率和良好的鲁棒性。

关键词 图像处理; 智能交通; 多目标识别; YOLOv2; 深度学习

中图分类号 U491 文献标志码 A

doi: 10.3788/LOP57.101010

Multi-Target Recognition Method Based on Improved YOLOv2 Model

Li Xun¹, Shi Binbin^{1*}, Liu Yang², Zhang Lei¹, Wang Xiaohua¹

¹School of Electronics and Information, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China;

²Xi'an Metrological Technology Research Institute, Xi'an, Shaanxi 710068, China

Abstract Based on the YOLOv2 algorithm, the YOLOv2-voc network structure is improved according to the actual road-scene change. The classification training model is obtained based on ImageNet data and fine-tuning technology and in accordance with the analysis of the training results and target vehicle characteristics. Consequently, the improved vehicle identification classification network structure YOLOv2-voc_mul is obtained. Using samples from simple and complex backgrounds, experiments are conducted to verify the validity of the detection method. Further, the proposed model is compared with the YOLOv2, YOLOv2-voc, and YOLOv3 models after 70000 iterations. Results show that under simple background, the improved YOLOv2-voc_mul model has an accuracy of 99.20% and the mean average precision of different models achieves 89.03%. Under complex background, the improved YOLOv2-voc_mul model has average accuracies of 92.21% and 89.44% for the single- and multi-target detection of four different models, respectively. The proposed model shows excellent accuracy, small false detection rate, and good robustness.

Key words image processing; intelligent traffic; multi-target recognition; YOLOv2; deep learning

OCIS codes 100.2000; 100.4996; 150.0155

1 引言

在现实交通场景中, 车辆目标检测^[1]受到诸多因素的影响, 例如光照、角度、形变、遮挡等^[2]。快速

准确地对交通场景下的车辆目标进行检测, 一直是图像处理与交通工程交叉领域的研究内容, 在此基础上的车型识别更是重要的研究课题, 国内外学者对此进行了大量卓有成效的研究。

收稿日期: 2019-08-28; 修回日期: 2019-10-08; 录用日期: 2019-10-18

基金项目: 国家自然科学基金(51607133)、陕西省自然科学基金基础研究计划项目(2019JM567)、中国纺织工业联合会科技指导性项目(2018094)、西安工程大学控制科学与工程建设经费资助项目(107090811)

* E-mail: 734931099@qq.com

最早使用的检测方法是基于机器学习的方法,屈治华等^[3]提出的基于图像关键点统计变换(MCT)特征的 Adaboost 集成算法、张敦凤等^[4]提出的基于分块局域二值模式(LBP)融合特征和支持向量机(SVM)的人脸识别算法、郭健^[5]提出的基于局部特征的图像匹配算法等,均是通过对目标进行区域选择,然后进行特征提取,并将提取到的特征输入到 SVM^[6]、迭代器^[7]等分类器进行分类识别。但传统的目标检测方法主要是针对特定目标的检测,受限于多分类的目标,而且目标的区域选择过程较复杂、检测效率较低。在选择对象时,其特征提取存在主观性强、鲁棒性差、泛化能力弱等缺点,在实际应用中难以获得精准的识别效果。

将深度学习^[8]技术应用于目标检测领域是近年来使用最广的目标检测方法。卷积神经网络相比传统方法更容易提取出图像的深层特征,通过卷积神经网络就能把特征提取、选择和分类融合^[9]在一起,这种方法可在很大程度上提升检测准确度。Girshick 等^[10]基于卷积神经网络提出了 R-CNN(region-conventional neural network)目标检测模型。该模型将传统机器学习和深度学习结合起来,通过卷积神经网络提取目标特征,再使用 SVM 进行分类,最后得到的平均精度均值(mAP)明显增大^[11]。但是 R-CNN 在选择候选区域时,算法比较复杂且耗时较长,不能达到实时检测的目的。由于 R-CNN 存在许多不足,人们又提出许多新的目标检测算法:Fast R-CNN^[12]、Faster R-CNN^[13]、YOLO(you only look once)^[14]和 SSD(single shot multibox detector)^[15]等。其中 YOLO 的性能较强,其网络结构简单,检测速度快,基本可以满足视频检测^[16]的要求。同时,深度卷积神经网络对形变、光照、几何变换具有一定的不变性^[17],有效克服了车辆外观多变给目标检测识别带来的困难^[18]。

本文针对道路车辆多目标车型识别检测中,车辆外形、结构、颜色以及现实场景等特征导致的识别率低的问题,利用 YOLOv2 实时目标检测算法^[19],创建实验室自有的构建车辆目标的 VOC 数据集,用于 YOLOv2 和 YOLOv2-voc 模型的训练;根据实验中发现的多目标车辆特征,通过对网络模型的参数进行多次调整后再次进行训练,获得了适用于车辆多目标车型实时识别的改进 YOLOv2 模型;同时,在实际交通环境下对本文模型进行了实验分析,并与 YOLOv2、YOLOv2-voc 和 YOLOv3 等模型进行了对比,验证了本文模型对实际的交通环境具

有适用性和先进性。

2 基于实证数据的 YOLOv2 改进模型

YOLO 是一个端到端的目标检测网络,能够实现实时检测且其检测性能稳定。相比于 R-CNN^[20]、Fast R-CNN 等采用选择性搜索的方法来产生候选框和 Faster R-CNN 等利用区域候选的方法来提取候选框,YOLO 算法直接利用回归方法在输出层提取出候选框的位置和所属的类别,从而提高了目标检测的速度,但存在检测准确率不高的问题。YOLOv2^[21]是对 YOLO 算法的改进,对每个候选框预测一个独立的类别,提高了网络对于多目标的检测能力,因此在保持原有较高检测速度的基础上,提高了目标检测的准确率。YOLOv3 在检测远距离的小目标时有较高的检测精度和较好的识别效果;而在检测近距离的大目标时,其识别效果相比于 YOLOv2 有所下降。

2.1 基于 YOLOv2 模型的目标检测算法

YOLOv2 是实时的物体检测算法,该算法把图片作为输入,直接将目标位置及其对应位置的置信度评分作为输出。YOLOv2 模型中去掉了分类器,不再使用基于滑动窗口的方法进行特征提取。YOLOv2 算法遵循端到端训练和实时检测的设计理念:把输入图片分割成 $n \times n$ 个区域,如果一个标注对象的中心落在某个区域上,那么这个区域负责预测这个物体,每个区域预测所需要的 bounding box 的位置和置信度,每一个 bounding box 都会得到 5 个预测值,分别为 (t_x, t_y) 、 (t_w, t_h) 和对应的置信度 C_{on} (图 1)。 b_x 、 b_y 分别表示每个 bounding box 的中心点距离边界的距离, b_w 、 b_h 分别表示 anchor 的实际宽和高; σ 表示激活函数 sigmoid, $\sigma(t_x)$ 、 $\sigma(t_y)$ 是每个 bounding box 的中心点距离其所在边界的偏移量, x 和 y 是 bounding box 中心点相对于对应网格的偏移量比例; (t_w, t_h) 分别是相对于整幅图像比例的目标真实宽和高, w 和 h 是 bounding box 相对于整幅图片尺寸的比例;网络距离图像左上角的边距为 (c_x, c_y) ,每个区域对应的 bounding box 长和宽分别为 p_w 、 p_h ,则 bounding box 的真实位置可表示为

$$\begin{cases} b_x = \sigma(t_x) + c_x \\ b_y = \sigma(t_y) + c_y \\ b_w = p_w \exp(tw) \\ b_h = p_h \exp(th) \end{cases} \quad (1)$$

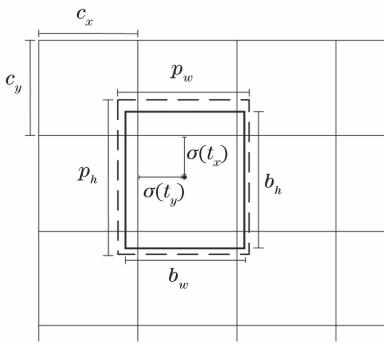


图1 目标位置

Fig. 1 Target position

C_{on} 表示候选边框的置信度,通过 bounding box 及其对应的待检测目标概率及该 bounding box 和真实位置区域的交并比 (IOU, I_{pred}^{truth}) 的关系对该 bounding box 预测位置的精度进行表征,即

$$C_{on} = P_r \times I_{pred}^{truth}, \quad (2)$$

式中: P_r 表示网格中存在目标的概率值。若一个网格中出现了目标, $P_r=1$;若没有出现目标,则 $P_r=0$,那么 $C_{on}=0$ 。 $A_{inter}(A_{pred} \cap A_{truth})$ 表示预测目标框和真实目标框的交集部分面积, $A_{sum}(A_{pred} \cup A_{truth})$ 表示预测目标框和真实目标框的并集部分面积, I_{pred}^{truth} 表示预测目标框与真实目标框的面积交并

比,其计算公式为

$$I_{pred}^{truth} = \frac{A_{inter}(A_{pred} \cap A_{truth})}{A_{sum}(A_{pred} \cup A_{truth})}. \quad (3)$$

当网格中检测到目标时,需要进一步对目标的类别进行预测,并且用 $P_{r-class|object}$ 表示预测得到的概率值。将类别预测得到的概率与置信度 C_{on} 相乘,得到某个类别 M 的置信度 $C_{on}(M)$:

$$C_{on}(M) = P_{r-class|object} \times P_{r-object} \times I_{pred}^{truth} = P_{r-classM} \times I_{pred}^{truth}. \quad (4)$$

2.2 基于 YOLOv2-voc 多目标检测识别模型的改进

YOLOv2 利用了 anchor boxes 来预测 bounding boxes,去掉最后的全连接层部分,网络结构采用卷积层和池化层,图片尺寸由原来的 448 pixel×448 pixel 调整为 416 pixel×416 pixel,利用 YOLOv2 对图像进行了 32 倍的降采样,最终输出的特征图尺寸是 13 pixel×13 pixel,这样可以产生一个中心栅格,用这个中心栅格去预测落在图像中心的物体。为了得到多目标检测识别效果更好的网络,本文在 YOLOv2-voc 网络结构的基础上,调整网络中的参数(表 1)并进行多次实验,得到了不同的检测结果,如图 2 所示。

表 1 网络框架

Table 1 Network framework

Model	Learning rate	Layer number of network structure			Activation function
		Convolution layer	Maximum pooling layer+ average pooling layer	BN layer	
Model 1	0.001	23	5+0	22	22 leaky+1 linear
Model 2	0.0001	23	5+0	20	22 leaky+1 linear
Model 3	0.01	23	5+0	22	22 leaky+1 linear
Model 4	0.001	20	5+1	22	19 leaky+1 linear
Model 5	0.001	20	5+1	22	19 leaky+1 ReLU
Model 6	0.001	20	5+0	20	19 leaky+1 linear



图 2 不同模型的测试结果。(a)模型 1;(b)模型 2;(c)模型 3;(d)模型 4;(e)模型 5;(f)模型 6

Fig. 2 Test results of different models. (a) Model 1; (b) model 2; (c) model 3; (d) model 4; (e) model 5; (f) model 6

模型 1 采用的是 YOLOv2-voc 模型,该网络结构包括 5 个最大池化层和 23 个卷积层,使用 linear 激活函数,并且设置初始学习率为 0.001;模型 2 在模型 1 的基础上调整初始学习率为 0.0001;模型 3 则把学习率设置为 0.01;模型 4 在模型 1 的基础上加入了 1 个平均池化层,并且去掉 3 个卷积层;模型

5 在模型 4 的基础上将最后一层激活函数改为 ReLU;模型 6 则是在模型 1 的基础上去掉了 3 个卷积层。

从图 2 所示的试验模型测试结果可以看出:图 2(a)在检测远处小目标时出现了漏检现象;图 2(b)没有检测出任何目标;图 2(c)~(e)出现了严重

3.3 实验结果及分析

3.3.1 模型验证结果分析

对典型的 YOLOv2、YOLOv2-voc、YOLOv3 模型和改进的 YOLOv2-voc_mul 模型进行对比,使用预训练得到的网络模型对 YOLOv2、YOLOv2-voc 和 YOLOv2-voc_mul 网络模型参数分别进行初始化。在训练时,将初始学习率设为 0.001,在迭代 10000 次、20000 次和 40000 次时,以之前的 10% 来改变学习率,得到权重不同的多目标检测模型。数据集中的样本车辆目标为简单背景下的大目标,导致 YOLOv3 模型在训练过程中是“Nan”,不适用于本文所制作的数据集,该模型更适用于小目标的检测。最终用验证集比较所得到的多目标检测效果。

1) 损失曲线分析

图 5 为 YOLOv2、YOLOv2-voc、YOLOv3 和 YOLOv2-voc_mul 模型在训练时的损失曲线图。

图 5(a)~(d) 分别是 YOLOv2、YOLOv2-voc、YOLOv3 和 YOLOv2-voc_mul 模型以 0.001 的初始学习率在迭代 2000 次时的损失曲线图,训练开始时由于学习率过小发散比较严重,其中 YOLOv3 在训练开始时的损失值超过了 1800,相比其余 3 种模型明显较高;但在收敛速度上 YOLOv3 模型在迭代 200 次左右开始收敛并趋于 0,其余 3 种模型则均在迭代接近 350 次开始收敛,在迭代 1200 次之后开始无限趋近于 0。从图 5 来看,除 YOLOv3 外的 3 种模型的损失曲线只有微小的差别,在实验过程中发现,训练前期微小的收敛速度对识别效果没有太大影响。因此,对于总体的车型识别模型,YOLOv3 相对于本文模型不具有绝对优势。

2) 准确性分析

本文进一步对验证集进行测试,结果如表 2 所示。 N_{total} 表示待检测的实际目标数; C_{correct} 表示将图片输入后,网络检测出的 bounding box 目标数,每一个 bounding box 都有对应的置信度,置信度大于阈值的 bounding box 要进行 IOU 的计算,从中找到 IOU 最大的 bounding box,当最大 IOU 的 bounding box 大于预设的 IOU 阈值,那么 C_{correct} 值就增加 1; P_{proposal} 表示在检测出的所有 bounding box 中置信度大于阈值的 bounding box 数量; $P_{\text{precision}}$ 表示精确度; R_{recall} 表示召回率,是检测目标的个数与验证集中所有目标个数的比值; F_1 表示的是 F_1 分数(F_1 -score),又称平衡 F 分数(balanced F score),它被定义为精确率和召回率的调和平均数,同时兼顾了模型的召回率和精确度,其取值在

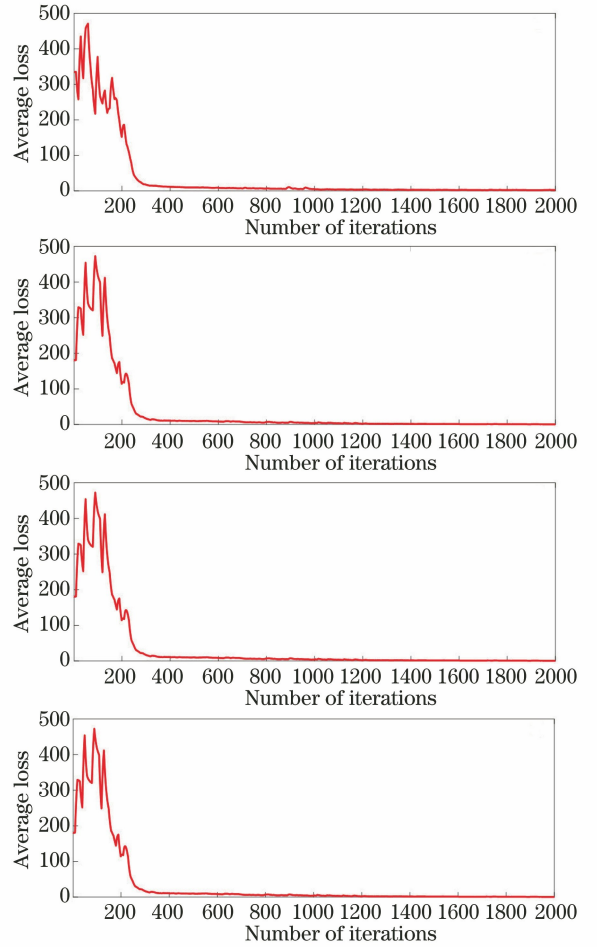


图 5 不同模型的损失曲线图。(a) YOLOv2; (b) YOLOv2-voc; (c) YOLOv3; (d) YOLOv2-voc_mul

Fig. 5 Loss graphs of different models. (a) YOLOv2; (b) YOLOv2-voc; (c) YOLOv3; (d) YOLOv2-voc_mul

0~1 之间, F_1 越大,说明效果越好。

$$P_{\text{precision}} = \frac{C_{\text{correct}}}{P_{\text{proposal}}}, \quad (5)$$

$$R_{\text{recall}} = \frac{C_{\text{correct}}}{N_{\text{total}}}, \quad (6)$$

$$F_1 = 2 \times \frac{P_{\text{precision}} \times R_{\text{recall}}}{P_{\text{precision}} + R_{\text{recall}}}. \quad (7)$$

由表 2 可知:在验证 152 个目标时,以 0.001 的初始学习率得到的 YOLOv3 模型可准确检测出 85 个车辆目标,但其 $P_{\text{precision}}$ 只有 56.67%, R_{recall} 值为 55.92%, F_1 值为 69.68%,效果比较差;YOLOv2 模型可以准确检测出 146 个目标, $P_{\text{precision}}$ 达到了 96.69%, R_{recall} 值达到 96.05%,且 F_1 的值为 96.36%;YOLOv2-voc 模型的 $P_{\text{precision}}$ 提升到了 98.63%;改进模型 YOLOv2-voc_mul 训练后得到的

表 2 测试结果
Table 2 Test results

Model	N_{total}	$C_{correct}$	$P_{proposal}$	$P_{precision}/\%$	$R_{recall}/\%$	$F_1/\%$
YOLOv2	152	146	151	96.69	96.05	96.36
YOLOv2-voc	152	144	146	98.63	94.74	96.64
YOLOv3	152	85	150	56.67	55.92	69.68
YOLOv2-voc_mul	152	145	146	99.20	95.39	97.26

验证结果显示,其 $P_{precision}$ 提高到了 99.20%, R_{recall} 值提高到了 95.39%, F_1 则增加到 97.26%, 三者均有不同程度的提高, 因此认为改进后的模型达到了一个较好的折中效果。

为了将表 2 更加直观地体现出来, 将 YOLOv2、YOLOv2-voc、YOLOv3 和 YOLOv2-voc_mul 模型的 IOU 值、 R_{recall} 值和 $P_{precision}$ 值绘制成曲线, 如图 6 所示。

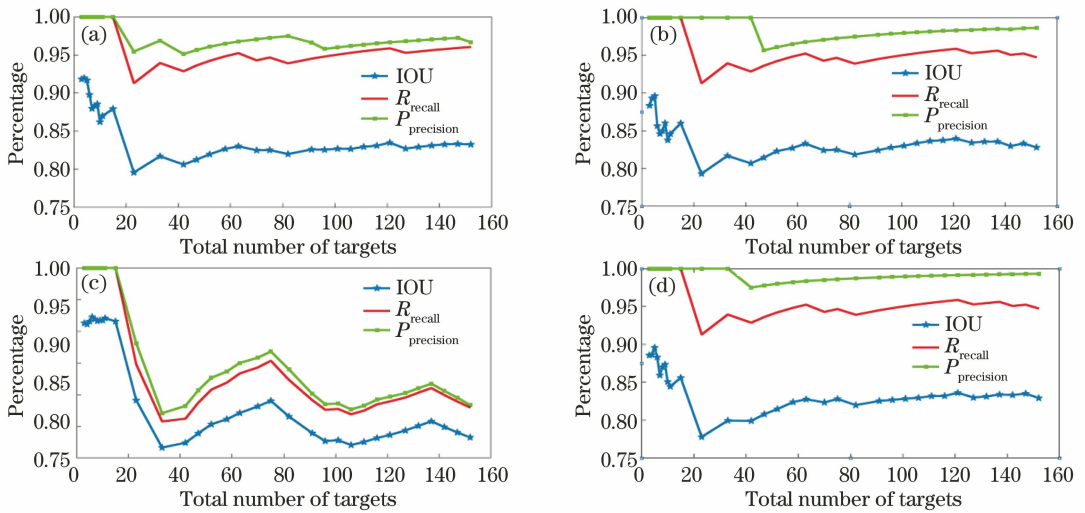


图 6 不同模型的验证结果。(a) YOLOv2; (b) YOLOv2-voc; (c) YOLOv3; (d) YOLOv2-voc_mul

Fig. 6 Verification results of different models. (a) YOLOv2; (b) YOLOv2-voc; (c) YOLOv3; (d) YOLOv2-voc_mul

图 6 所示为 4 个模型的检测结果, 可以看出: 4 个模型的召回率在初始时都出现了较大波动, 但当检测的目标数增加时, YOLOv2 模型的召回率逐渐稳定在 96% 左右, YOLOv2-voc 趋于 94.5%, YOLOv2-voc_mul 模型召回率稳定在 95.5% 左右, 而 YOLOv3 模型的召回率在 40%~60% 之间波动, 说明在简单背景下 YOLOv2、YOLOv2-voc、YOLOv2-voc_mul 模型均能得到较高的正确率, 而 YOLOv3 的正确率较低; 在精度曲线方面, YOLOv2 模型的精度曲线出现较大的波动, YOLOv2-voc 模型的精度曲线则在目标数增加时出现跳变, 之后逐渐稳定在 98.6% 左右, YOLOv2-voc_mul 模型的精度曲线在发生小幅度波动后稳定在 99.2% 左右, 保持了很高的精确度和稳定性, 而 YOLOv3 模型的曲线出现了很大的跳变, 最终的精度值在 60% 左右波动; 同时对比 4 种模型的交并比曲线可以看出, YOLOv2 模型交并比值在 0.75~

0.83 之间波动, 即稳定性较低, YOLOv2-voc 模型和 YOLOv2-voc_mul 模型的交并比较 YOLOv2 有所提高, 能够保持在 0.8~0.83 之间, YOLOv2-voc_mul 模型的交并比值在目标数增加的情况下在 0.83 上下浮动, 而 YOLOv3 的交并比只在 0.4~0.7 间波动, 相较于其他 3 种模型, YOLOv3 的稳定性最差。

3) 检测结果分析

为进一步验证模型的效果, 训练不同权重模型对检测结果的影响。本文将 YOLOv2-voc_mul 模型以 0.001 的初始学习率在经过训练后得到了迭代 60000 次和 70000 次的权重模型, 对不同的权重模型进行车辆多目标的检测, 对检测结果进行直观的对比, 结果如图 7 所示。

在验证集样本上测试不同权重的检测模型, 发现: 该模型用训练 60000 次的权重进行检测时, 将面包车误检为小汽车, 将小汽车误检为大卡车, 出现了明显的误检现象; 而当使用训练 70000 次



图 7 不同迭代次数时的测试结果。(a) 60000 次;(b) 70000 次

Fig. 7 Test results at different number of iterations. (a) 60000; (b) 70000

的权重进行检测时,则该误检现象被消除。由以上实验可得:YOLOv2-voc_mul 模型训练 70000 次学习到的特征更全面,消除了误检现象,可以正确识别出大卡车、公交车、面包车、小汽车 4 种车型,识别效果良好。

3.3.2 实验对比结果及分析

在车型识别检测的实验中,选择用 YOLOv2、YOLOv2-voc、YOLOv3 和 YOLOv2-voc_mul 模型的平均精度均值(mAP)进行对比分析。使用 YOLOv2、YOLOv2-voc、YOLOv3 和 YOLOv2-voc_mul 模型迭代 70000 次后得到相同初始化后的车型识别检测模型,检测验证集样本结果如表 4 所示。

表 4 不同模型检测 mAP 的比较

Table 4 Comparison of mAP value in different models

Model	mAP / %			
	Truck	Bus	Van	Car
YOLOv2	86.45	84.16	82.37	86.88
YOLOv2-voc	87.61	85.22	83.02	87.31
YOLOv3	83.35	81.83	77.96	83.92
YOLOv2-voc_mul	88.72	88.56	86.64	89.03

检测结果表明:YOLOv2 检测 4 种车型的 mAP 为 84.97%,YOLOv2-voc 检测 4 种车型的 mAP 为 85.79%,YOLOv3 检测 4 种车型的 mAP

为 81.77%,YOLOv2-voc_mul 模型检测 4 种车型的 mAP 则达到 88.24%;由于面包车的特征比较难区分,YOLOv2 检测的 mAP 为 82.37%,YOLOv2-voc 检测的 mAP 为 83.02%,YOLOv3 的 mAP 只有 77.96%,YOLOv2-voc_mul 模型的 mAP 则提高到了 86.64%。可见,YOLOv2-voc_mul 模型对不同车型具有更高的识别率和更好的分类效果。

3.4 多目标识别检测结果及分析

在深度学习中,训练过程所使用的数据样本对训练结果的影响很大,简单背景下的单目标特征明显,在训练过程中容易学习,识别率高;而实际道路目标的背景非常复杂且存在许多干扰因素,特征的学习相对较难,识别率自然较低,因此需要样本丰富多样且具有代表性。只采用互联网收集的简单背景下的车型不符合实际情况,因此本研究对训练样本进行了扩充,增加 500 幅在实际交通环境下拍摄的不同车型样本图像,部分训练样本如图 8 所示。

3.4.1 实际道路目标检测结果

对增加实际道路目标样本之后的数据集训练 70000 次后,4 种模型在验证集上得到的检测结果如图 9 所示。

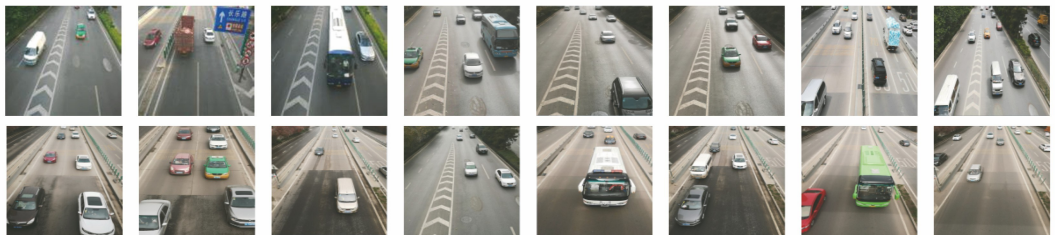


图 8 增加的数据样本类型

Fig. 8 Added model samples

YOLOv2 模型在检测时漏检了图 9(a1)左上角的面包车,图 9(a3)中无法检测到距离较远的小目标,出现严重漏检现象;YOLOv2-voc 模型的

图 9(b1)中没有出现漏检现象,图 9(b3)中同样出现小目标的严重漏检,图 9(b2)中则出现了面包车的误检和重检现象,将面包车误检为小汽车;

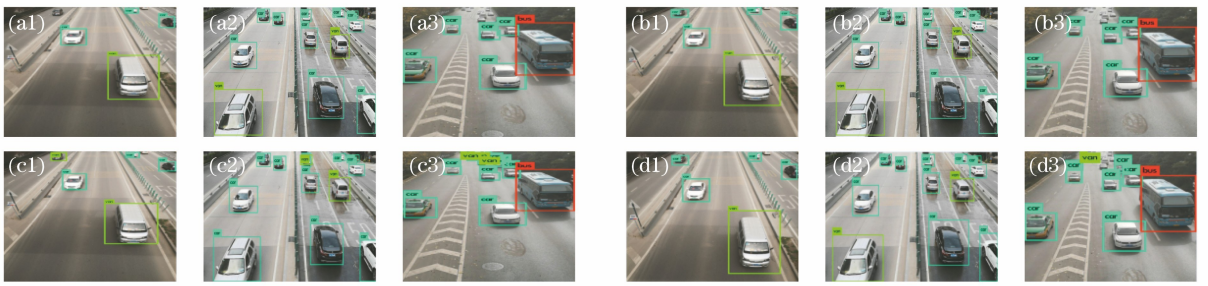


图9 不同模型的检测结果。(a) YOLOv2模型;(b) YOLOv2-voc模型;(c) YOLOv3模型;(d) YOLOv2-voc_mul模型
Fig. 9 Test results of different models. (a) YOLOv2 model; (b) YOLOv2-voc model; (c) YOLOv3 model;
(d) YOLOv2-voc_mul model

YOLOv3模型的图9(c1)左上角出现错检现象,图9(c2)左下角出现错检,中间部分出现重检现象;YOLOv2-voc_mul模型在图9(a2)的基础上消除了图9(d2)中的误检和重检现象,且在图9(d3)中检测出后面的面包车,可见改进后的模型在实际道路目标中具有更好的效果。

3.4.2 单目标和实际道路多目标检测结果对比分析

为进一步证明YOLOv2-voc_mul模型的适用性,本文用增加后的数据集进行迭代70000次的单目标和多目标分类检测实验,实验结果的平均准确率如表5所示。

表5 YOLOv2-voc_mul模型的车型识别平均准确率
Table 5 Average accuracy of vehicle identification in the YOLOv2-voc_mul model

Type	Accuracy / %				Average accuracy / %
	Truck	Bus	Van	Car	
Simple target	92.14	91.89	90.08	94.72	92.21
Multiple target	89.01	88.76	88.20	91.71	89.44

在对单目标和多目标的不同车型进行检测时,两者的平均准确率分别为92.21%和89.44%。从表5可以看出:在单目标检测中,YOLOv2-voc_mul模型对大卡车的平均准确率为92.14%,对公交车的平均准确率为91.89%,对面包车的平均准确率为90.08%,对汽车的平均准确率为94.72%;而在多目标检测中,YOLOv2-voc_mul模型对大卡车的平均准确率为89.01%,对公交车的平均准确率为88.76%,对面包车的平均准确率为88.20%,对汽车的平均准确率为91.71%。由于面包车车型相比于其他车型特征不明显,因此在单目标检测对面包车的平均准确率只有90.08%。在检测不同车型时,单目标的检测准确率比多目标高,对于学习到的特征有一定的记忆性,由于车型数据集中单目标样本比较多,因此该模型对单目标车型特征的学习会更

精确,检测效果更好。

4 总结

针对实际交通环境下车辆的多目标车型检测问题,利用深度学习Darknet框架下的YOLOv2目标检测算法,通过对YOLOv2、YOLOv2-voc和YOLOv3模型参数的分析,根据车型的特征及车辆的运动特征,进行不同网络模型的训练和对比,提出了一种检测精确度较高的YOLOv2-voc_mul模型。通过对所建立的模型进行大量实验,发现改进的YOLOv2-voc_mul模型对简单背景下的单目标实验正确率可以达到92.21%,在复杂背景下的正确率也达到了89.44%。相比于传统的机器学习方法,该方法提高了检测的准确率和运行效率。

参考文献

- [1] Wang D C, Chen X N, Zhao F, et al. Vehicle detection algorithm based on convolutional neural network and RGB-D images [J]. Laser & Optoelectronics Progress, 2019, 56(18): 181003. 王得成, 陈向宁, 赵峰, 等. 基于卷积神经网络和RGB-D图像的车辆检测算法[J]. 激光与光电子学进展, 2019, 56(18): 181003.
- [2] Hua X, Wang X Q, Wang D, et al. Multi-objective detection of traffic scenes based on improved SSD [J]. Acta Optica Sinica, 2018, 38(12): 1215003. 华夏, 王新晴, 王东, 等. 基于改进SSD的交通大场景多目标检测[J]. 光学学报, 2018, 38(12): 1215003.
- [3] Qu Z H, Shao Y M, Deng T M, et al. Traffic sign detection and recognition under complicated lighting conditions [J]. Laser & Optoelectronics Progress, 2019, 56(23): 231009. 屈治华, 邵毅明, 邓天民, 等. 复杂光照条件下的交通标志检测与识别[J]. 激光与光电子学进展, 2019,

- 56(23): 231009.
- [4] Zhang D F, Gao N H, Wang H, et al. Face recognition algorithm based on block LBP fusion feature and SVM[J]. *Transducer and Microsystem Technologies*, 2019, 38(5): 154-156, 160.
张敦凤, 高宁化, 王姮, 等. 基于分块 LBP 融合特征和 SVM 的人脸识别算法[J]. *传感器与微系统*, 2019, 38(5): 154-156, 160.
- [5] Guo J. Research of image matching algorithm based on local features[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2018.
郭健. 基于局部特征的图像匹配算法研究[D]. 南京: 南京邮电大学, 2018.
- [6] Wang W X, Fu Y T, Dong F, et al. Infrared ship target detection method based on deep convolution neural network[J]. *Acta Optica Sinica*, 2018, 38(7): 0712006.
王文秀, 傅雨田, 董峰, 等. 基于深度卷积神经网络的红外船只目标检测方法[J]. *光学学报*, 2018, 38(7): 0712006.
- [7] Jin L S, Wang Y, Liu J H, et al. Front vehicle detection based on Adaboost algorithm in daytime[J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2014, 44(6): 1604-1608.
金立生, 王岩, 刘景华, 等. 基于 Adaboost 算法的日间前方车辆检测[J]. *吉林大学学报(工学版)*, 2014, 44(6): 1604-1608.
- [8] Zheng W X, Wang C P, Fu Q. Improved KCF infrared aerial target tracking method[J]. *Laser & Infrared*, 2017, 47(12): 1553-1558.
郑武兴, 王春平, 付强. 改进的 KCF 红外空中目标跟踪方法[J]. *激光与红外*, 2017, 47(12): 1553-1558.
- [9] Zhao H, An W S. Image salient object detection combined with deep learning [J]. *Laser & Optoelectronics Progress*, 2018, 55(12): 121003.
赵恒, 安维胜. 结合深度学习的图像显著目标检测[J]. *激光与光电子学进展*, 2018, 55(12): 121003.
- [10] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014. Columbus, OH, USA. New York: IEEE, 2014: 8-13.
- [11] Li Y P, Hou L Y, Wang C. Vehicle object detection in complex scene based on YOLOv2 [J]. *Video Engineering*, 2018, 42(5): 100-106.
李云鹏, 侯凌燕, 王超. 基于 YOLOv2 的复杂场景下车辆目标检测[J]. *电视技术*, 2018, 42(5): 100-106.
- [12] Zhang Q, Hu G D, Li Y S, et al. Binocular vision vehicle detection method based on improved Fast-RCNN[J]. *Journal of Applied Optics*, 2018, 39(6): 832-838.
张琦, 胡广地, 李雨生, 等. 改进 Fast-RCNN 的双目视觉车辆检测方法[J]. *应用光学*, 2018, 39(6): 832-838.
- [13] Feng X Y, Mei W, Hu D S. Aerial target detection based on improved faster R-CNN [J]. *Acta Optica Sinica*, 2018, 38(6): 0615004.
冯小雨, 梅卫, 胡大帅. 基于改进 Faster R-CNN 的空中目标检测 [J]. *光学学报*, 2018, 38(6): 0615004.
- [14] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016. Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [15] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[M] // *Computer Vision-ECCV 2016*. Cham: Springer International Publishing, 2016: 21-37.
- [16] Song H S, Zhang X Q, Zheng B F, et al. Vehicle detection based on deep learning in complex scene[J]. *Application Research of Computers*, 2018, 35(4): 1270-1273.
宋焕生, 张向清, 郑宝峰, 等. 基于深度学习方法的复杂场景下车辆目标检测[J]. *计算机应用研究*, 2018, 35(4): 1270-1273.
- [17] Zhu M M, Xu Y L, Ma S P, et al. Airport detection method with improved region-based convolutional neural network [J]. *Acta Optica Sinica*, 2018, 38(7): 0728001.
朱明明, 许悦雷, 马时平, 等. 改进区域卷积神经网络的机场检测方法[J]. *光学学报*, 2018, 38(7): 0728001.
- [18] Lu X J, Guo L, Chen Z R, et al. Study on vehicle detection and tracking algorithms based on appearance and motion [J]. *Computer Engineering*, 2014, 40(8): 152-157.
陆星家, 郭璘, 陈志荣, 等. 基于外观和运动的车辆检测和追踪算法研究[J]. *计算机工程*, 2014, 40(8): 152-157.
- [19] Li X, Liu Y, Li P F, et al. Vehicle multi-target detection method based on YOLO v2 algorithm under

- darknet framework [J]. Journal of Traffic and Transportation Engineering, 2018, 18(6): 142-158.
- 李珣, 刘瑶, 李鹏飞, 等. 基于 Darknet 框架下 YOLO v2 算法的车辆多目标检测方法[J]. 交通运输工程学报, 2018, 18(6): 142-158.
- [20] Li M, Jing J F, Li P F. Yarn-dyed fabric defect detection based on GAN and Faster R-CNN [J]. Journal of Xi'an Polytechnic University, 2018, 32(6): 663-669.
- 李明, 景军锋, 李鹏飞. 应用 GAN 和 Faster R-CNN 的色织物缺陷识别[J]. 西安工程大学学报, 2018, 32(6): 663-669.
- [21] Li X, Liu Y, Zhao Z F, et al. A deep learning approach of vehicle multitarget detection from traffic video[J]. Journal of Advanced Transportation, 2018 (11):1-11.