

文章编号: 2095-2163(2021)09-0170-05

中图分类号: TP183;TP391.41

文献标志码: J

# 基于改进 SSD 算法的骑车人识别研究

马佳峰, 陈凌珊

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

**摘要:** 在智能网联汽车蓬勃发展的大背景下, 目标识别作为智能驾驶的关键技术能够提高公路环境的安全性, 本文采用 SSD 算法对公路环境下骑车人检测识别技术进行研究, 发现 SSD 算法的小目标检测效果和平均识别精度都不太理想, 故通过参考 YOLOv3 算法的跨层链接思想, 在网络上引入 FPN 结构, 进而提高识别效果。在 TDCB 数据集上的实验结果表明, 平均检测精度和对小目标检测效果均有所提高, 精度上提高约为 2.2%, 检测速度虽略微减缓, 仍符合实际应用需求, 改进后的 SSD 算法对提高公路环境下骑车人安全有着重要意义。

**关键词:** 骑车人检测; 深度学习; SSD; FPN; 跨层链接思想

## A cyclist identification study based on improved SSD

MA Jiafeng, CHEN Lingshan

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**【Abstract】** Under the background of the fast development of intelligent networked vehicles, target recognition, as a key technology of intelligent driving, can improve the safety of highway environment. In this paper, SSD algorithm is used to study the cyclist detection and recognition technology in highway environment. It is found that the small target detection and average recognition accuracy of SSD algorithm are not ideal. Therefore, by referring to the cross layer link idea of YOLOv3 algorithm, this paper introduces FPN structure into the network to improve the recognition effect. The experimental results on TDCB dataset show that the average detection accuracy and small target detection effect are improved, and the accuracy is improved by about 2.2%. Although the detection speed is slightly slowed down, it still meets the requirements of practical application and the improved SSD algorithm is of great significance to improve the safety of cyclists in highway environment.

**【Key words】** cyclist identification; deep learning; SSD; FPN; cross layer connection

## 0 引言

在智能网联汽车和无人驾驶快速发展的环境下, 基于深度学习的目标识别作为其中的关键技术显得愈加重要, 而公路环境中的骑车人作为目标识别当中一个种类, 自然也被考虑并加以研究。骑车人作为公路环境中易受伤的人, 在公路上的行车可能不规范, 或者汽车驾驶员的驾驶不规范, 加上日渐复杂的公路交通场景, 多种因素叠加, 骑车人群体的自身安全受到了严重的威胁, 由此引发的问题不容小觑, 该群体的安全应当受到重视。现今智能网联汽车的目标识别检测技术在大力发展但并没有完善, 有关公路环境下骑车人识别的研究相对于其它目标研究来说较少, 提高骑车人的安全十分有必要, 因此对公路环境下骑车人识别技术的研究具有重要意义<sup>[1]</sup>。

当下目标识别的技术主要是基于机器视觉技术, 分为传统机器学习和深度学习两种方式。传统的机器学习对于目前的实际应用场景来说, 其检测速度并

不符合实际需求, 目前无人驾驶中的主要目标识别技术为深度学习<sup>[2]</sup>。智能网联汽车和无人驾驶成为大趋势, 安全行作为其中一个重要考量指标, 如何降低公路安全事故发生的概率成为越来越热的话题, 对公路骑车人的识别研究也被广泛讨论<sup>[3]</sup>。深度学习算法分 One-stage 类和 Two-stage 类算法, 目前主流采用的是 One-Stage 类, 该类算法检测速度较快, 具有重要代表性的直接基于回归算法为 SSD 算法, 该算法在精度上和速度上都有较好的检测效果, 其汲取 Two-stage 算法中的特点, 借鉴了 RPN 多窗口技术, 在此基础上在几个不一样的尺度上的特征图上都采用了框的坐标回归和多种尺度的检测, 使 SSD 算法既达到了对小目标检测效果, 同时还能保证检测的速度。本文采用深度学习算法对其进行研究, 采用的算法是 SSD 算法, 并对其尝试性的改进。

## 1 SSD 算法的网络结构

从网络结构上: 其主干网络是 VGG-16, 并对其进

行了相应的改善。VGG-16 网络的结构如图 1 所示。

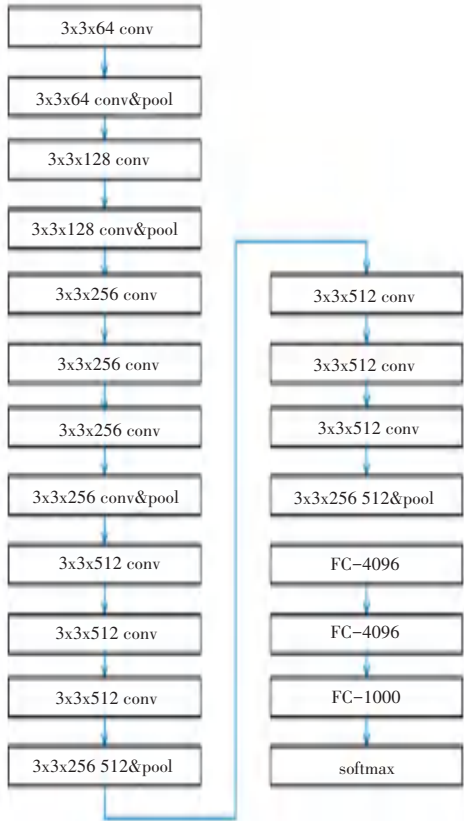


图 1 VGG-16 网络结构

Fig. 1 VGG-16 network structure

引用和参考 Alexnet, 对网络的深度进行了提高, 能够更好的描述目标特征, 表达能力更强, 在复杂场景的应用上效果更好, 总的来说 VGG-16 用于特征提取效果不错, 适用于目标检测和识别领域。

SSD 算法的网络模型, 在使用 VGG-16 网络模型的基础上结合了 YOLO 算法的跨层连接思想和 FasterRCNN 算法思想<sup>[4]</sup>, 使整体的检测效果有了极大的加强, 其网络结构如图 2 所示。

从图 2 可知, SSD 算法在主干网络 VGG-16 的改进为: 加入了一定数目的卷积层, VGG16 包含 5 个卷积块, 一般命名从 conv1~conv5, 改进后的 VGG-16 网络, 不含 conv5 卷积块, 用的是前 4 卷积块。SSD 算法输入端为 300x300x3 的图片, 靠 VGG-16 提取特征后, 得出尺寸为 38x38x512 的特征层, 每一个卷积特征层是靠每单次的下采样得出的, 经过上述步骤再对得出的卷积特征层采取目标检测步骤, 通过下采样 5 次, 得出尺寸为 1x1x256 的卷积特征层, 得出 5 个尺寸不一的卷积特征层, 和 VGG 网络的特征层加在一起, 最终有 6 个卷积层。一般采用一组的卷积滤波器对每个卷积特征层实行检测。在大小为 MxN 且包含 P 通道的特征层, 以 3x3xp 的卷积核为检测的基本单元, 生成单个类别的边框坐标还有置信度。

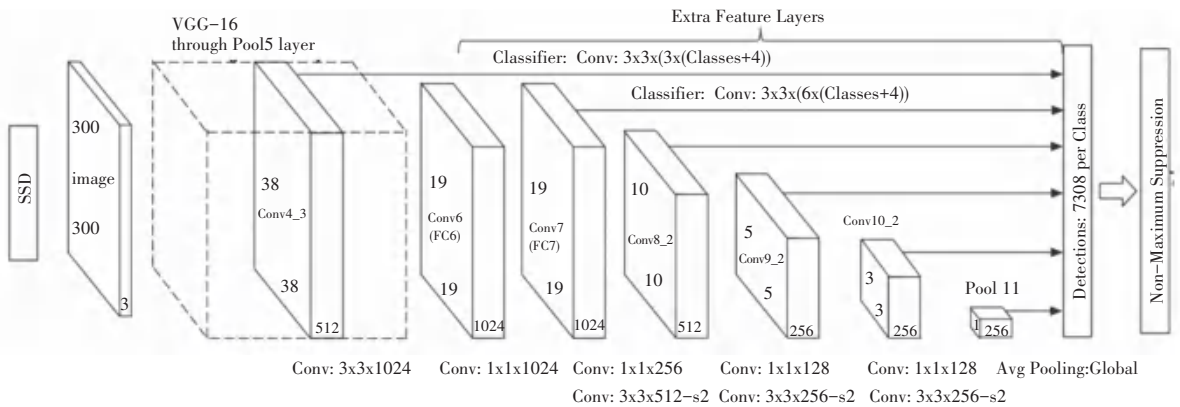


图 2 SSD 算法的网络结构图

Fig. 2 Network structure of SSD algorithm

## 2 基于 SSD 算法的改进优化

SSD 问世以来就广受研究人员的喜爱, 用 Python 语言编写的开源算法, 使用方便, 故而对于 SSD 算法的改进和优化络绎不绝<sup>[5]</sup>。SSD 算法检测小目标效果较差, 平静检测精度也有些偏低, 针对这些问题, 本文对该算法进行了研究, 并对其进行优化, 用以解决上述问题, 提升该算法的检测效果。针

对 SSD 算法存在的缺陷, 分析其原理上可能造成此结果的原因, 该算法将多个不同特征层没有跨层连接起来, 只是检测多种不一样尺度的特征图, 高层网络的语义信息并没有被底层网络更多的学习到。

YOLO 系列算法中引用了 FPN 思想, 参考这些理念, 对 SSD 算法进行改进, 在其网络结构上加入 FPN 网络<sup>[6]</sup>, 从而得出一种网络掺入了跨层连接理念 (cross layer connection), 命名为 SSD-CLC 网络。

## 2.1 SSD-CLC 网络概述

常态下,网络结构中,位于底层的特征图其分辨率高于其目标的位置精确,不过蕴有较少的特征语义信息,另一方面在高层方面,其包含的语义信息多,目标的位置却不怎么精确。特征金字塔的核心理念就在于跨层连接多种尺寸大小的卷积层,这样连接的卷积层兼并了低层网络的特征还有高层网络的特征,让网络对更多的特征进行学习,嵌入 FPN 网络,能让 SSD 算法的检测效果更进一步。

## 2.2 SSD-CLC 的网络结构模型

引入跨层连接思想能够让被改进的网络多个特征层的语义信息被更多的学习到,SSD 算法里采取

相似的跨层连接于多种不同尺度的卷积层,也能实现相近的结果。把 FPN 嵌入到 SSD 算法的结构中,其网络模型 SSD-CLC 如图 3 所示。

从图 3 可知,SSD-CLC 网络结构中  $19 \times 19$  规格的特征图是由  $10 \times 10$  规格的特征图经过上采样后得出的;将上采样得出的结果和原来的  $19 \times 19$  规格的相叠加,得到合成后的  $19 \times 19$  特征图;最终将新得到的 2 个特征图加上原有的 4 个特征图实行目标检测流程。采样方法一般有双线性插值、均值插值、中值插值、最邻近插值等,这里选取的是最邻近插值,这种方法对运算量要求不高,把图片中所有点的像素值赋成原图片里和其最挨着的点的像素值。

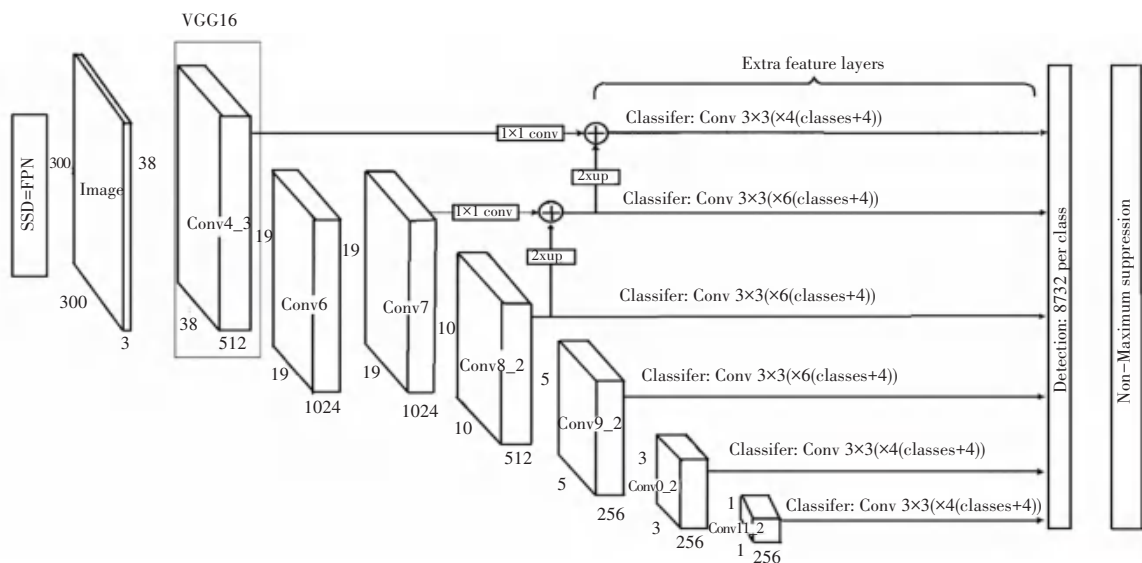


图 3 SSD-CLC 网络模型

Fig. 3 SSD-CLC network model

## 3 实验及其结果分析

本文实验环境为:惠普暗影精灵 5Plus 笔记本 win10 操作系统,模型使用 Pytorch 搭建,采用 GPU 加速训练,显卡型号为 NVIDIA 2080Ti,最终对实验结果用 3 个参数做为评价指标:  $AP$  (平均检测精度)、 $Recall$  (召回率) 和检测速度  $DR$  (Detection Rate)。

### 3.1 神经网络训练

数据集在网络模型训练中起着重要的作用,数据集的质量对模型训练的结果起着至关重要的作用,本文识别的对象为骑车人,目前针对于骑车人识别研究的数据集较少,经过筛选,本文选用的数据集为 TDCB 数据集,该数据集由清华大学 Li 的团队和

戴姆勒司合作制作而成<sup>[7]</sup>,经过图像的采集和标注后制作而成。

该数据集主要包含训练集和测试集,比例约 3 : 1,训练集包含 9 741 张图片且只标注骑车人,测试集共有 2 914 张图片,内含 6 种标注目标,分别为:骑车人、行人、骑摩托人、轮椅骑手、骑电瓶车人和骑三轮车人。目前主流数据集标注文件格式为 VOC 数据集标注文件格式和 COCO 数据集标注文件格式, VOC 数据集标注文件格式为 xml 格式,由于数据集标注文件为 json 格式,故 TDCB 数据集并不能被直接使用,编写相应的标注文件格式更改程序,将其转化为 xml 格式,使其能够被直接使用。

SSD-CLC 算法在上述的 TDCB 数据集进行训练后,损失函数曲线如图 4 所示。

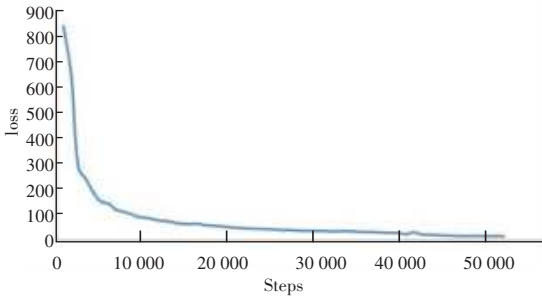


图 4 训练完成后的损失函数曲线  
Fig. 4 Loss function curve after training

### 3.2 实验结果

通过仿真实验可以看出,在 SSD 算法引入了跨层连接,识别效果有一定程度的提高,如图 5 所示。图 5(a)和图 5(c)是 SSD 引入跨层连接前的识别效果,其误把夹在人群间的行人识别成了骑车人,引入之后,误检消除,如 5(b)所示。图 5(c)为 SSD 的识别效果,其把远处的一个行人误识别成了骑车人,引入跨层连接后同样消除了误识别,结果为图 5(d)。从图 5(b)和图 5(d)结果还可以看出,在消除了误检的同时,还一定程度的提高了检测精度。



图 5 SSD-CLC 识别效果

Fig. 5 SSD-CLC recognition effect

### 3.3 SSD-CLC 仿真结果评估分析

SSD 和 SSD-CLC 在 TDCB 数据集上对所有测试样本进行试验后,得出 PR 曲线如图 6 所示。

由图 6 可以看出,相对于 SSD 算法,改进的 SSD-CLC 算法在平均精度上有所提高,其识别精度在召回率低的时候也有所提高,通过此结果可以看出,网络结构引入跨层链接后,能学习到更多的特征,识别的结果统计见表 1。

表 1 SSD 与 SSD-CLC 识别结果对比

Tab. 1 Comparison of recognition results between SSD and SSD-CLC

算法	测试样本个数	平均检测精度 (AP)	平均检测时间/s
SSD	2 914	0.607	0.12
SSD-CLC	2 914	0.629	0.14

实验结果表明,对 SSD 算法改进后,在检测小目标上效果更进一步,减少了误检的情况,其平均检 (下转第 183 页)

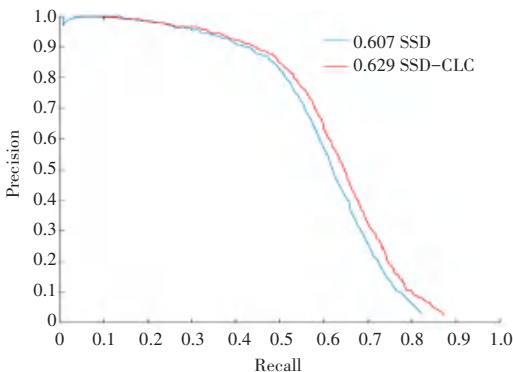


图 6 SSD 和 SSD-CLC 实验的 PR 曲线

Fig. 6 PR curves of SSD and SSD-CLC experiments