

DOI:10.19651/j.cnki.emt.2209423

## 基于改进 SSD 算法的交通标识检测方法研究\*

詹华伟<sup>1</sup> 邹昊好<sup>1</sup> 刘旭<sup>1</sup> 史水娥<sup>1</sup>

(1.河南师范大学电子与电气工程学院 新乡 453007; 2.河南师范大学河南省光电传感集成应用重点实验室 新乡 453007)

**摘要:** 针对目前 SSD 算法对小目标检测精确度低,泛化能力弱,且存在误检、漏检等问题,提出一种基于 SSD 网络的交通标识检测方法。为增加对目标的检测精度,使用 ResNet-50 网络作为 SSD 算法的骨干网络,在额外添加层中加入 BN 层,提高训练速度;使用 sub-pixel 来代替上采样,提高识别目标分辨率,并加入 MFPN 模型融合低层与高层特征信息,避免出现漏检问题。实验结果表明与现有的 SSD 算法相比,改进的 SSD 算法在公开数据集 CCTSDB 和 GTSDb 数据集上 mAP 值分别提高 4.2% 和 3.1%,FPS 保持在 87.2 f/s,检测精度显著提升。满足对交通标识实时检测的要求,在无人驾驶领域具有广泛的应用前景。

**关键词:** SSD;卷积网络;MFPN 模型;sub-pixel;CCTSDB

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6240

Research on traffic sign recognition method based on improved  
SSD algorithmZhan Huawei<sup>1</sup> Zou Haohao<sup>1</sup> Liu Xu<sup>1</sup> Shi Shui'e<sup>1</sup>

(1. College of Electronic and Electrical Engineering, Henan Normal University, Xinxiang 453007, China;

2. Henan Key Laboratory of Optoelectronic Sensing Integrated Application, Henan Normal University, Xinxiang 453007, China)

**Abstract:** This paper proposes a traffic sign detection method based on an SSD network. This method improves the existing SSD algorithm, which has low detection accuracy and weak generalization ability for small targets, and has problems such as false detection and missed detection. The ResNet-50 network is used as the backbone network of the SSD algorithm, and the BN layer is added to the additional layer to improve the training speed. Sub-pixel is used instead of upsampling to improve the resolution of the recognition target, and the MFPN model is added to fuse the low-level and high-level feature information to avoid the problem of missed detection. The experimental results show that the improved SSD algorithm improves the mAP value by 4.2% and 3.1% on the public datasets CCTSDB and GTSDb datasets, respectively, the FPS remains at 87.2 f/s, and the detection accuracy is significantly improved. This work meets the requirements for real-time detection of traffic signs and has broad application prospects in the field of unmanned driving.

**Keywords:** SSD;convolutional networks;MFPN model;sub-pixel;CCTSDB

## 0 引言

现如今无人驾驶正在逐步兴起,而在无人驾驶中,先进的无人驾驶辅助系统(advanced driver assistance system, ADAS)<sup>[1]</sup>发挥着重要作用。道路交通标志的检测和识别为无人驾驶辅助系统提供实时的道路标志感知信,由于交通标识通常由特定的形状(圆圈、正方形和三角形)和颜色(黄色、蓝色和红色)组成,这些标识在道路环境中具有显著的视觉影响<sup>[2]</sup>。目前的交通标识检测可分为基于颜色、基

于形状的方法,例如 HSI、CIElab、HSL 方法等。交通标识对安全起着重要的作用,但是其在图像中的比例较小,且分辨率较低所以很难进行识别,在实际的应用场景中对交通标识的检测还存在一定的困难。

近年来,深度卷积神经网络已成功应用于目标检测与识别,而主流的检测算法主要分为两个类型:One-stage 和 two-stage 方法,two-stage 方法如 Ross Girshich 等在 2014 年提出了 R-CNN<sup>[3]</sup>算法,该模型使用候选区域的选择性搜索算法,然后在候选区域使用 R-CNN 模型从中提取特征,

收稿日期:2022-03-28

\* 基金项目:河南省高等学校重点科研项目基础研究项目(19B510006)资助

并在向量机 (support vector machines, SVM)<sup>[4]</sup> 中完成分类,模型设计了边界框回归算法来计算候选区域的坐标并在坐标上进行测试,平均准确率比非神经网络算法高约 20%。

One-stage 方法中典型算法如 YOLO<sup>[5-6]</sup> 算法和 SSD<sup>[7]</sup> 算法, YOLO 是 Joseph Redmon 在 2016 年提出,改革了基于建议框的检测框架,处理图像非常快,适合实时处理,但是其对于交通标志及邻近目标检测效果差,且定位不准确。2016 年, Liu 等将 YOLO 网络与 Faster R-CNN 网络相结合提出了 SSD 目标检测算法。SSD 算法在整个图像上生成不同大小的目标边界框,并使用非极大抑制 (non-maximum suppression, NMS)<sup>[8]</sup> 将高度重叠的边界框合并为一个边界框,将候选区域转为一个线性回归问题,找到最接近目标的预测框,以提高计算速度和精度。在整个过程只用一步,提高了对交通标识的检测速度。这类算法将分类过程与回归网络合并为一个阶段进行,检测速度虽大幅提升,但精度也有所损失。而在实际的道路交通中,交通标识背景复杂,像素很低,特征少,而且难以识别<sup>[9-10]</sup>。One-

stage 方法虽然在检测目标方面取得了较好的效果,但对交通标识的检测还存在一定的缺陷。

针对现有的问题,为了提高对交通标识的检测精度和速度,本文提出了一种改进的 SSD 算法,在原有的 SSD 算法基础上,使用 Resnet-50 网络<sup>[11]</sup> 作为 SSD 算法的 backbone 网络,在额外层中添加 BN 层提高训练速度,用 MFPN 网络结构来对特征层进行处理,将低层特征和高层特征进行融合,并改变采样方式,在不改变原有模型计算量的基础上,增加训练样本,大幅度提升了对交通标识的检测性能。

### 1 SSD 目标检测算法

在原始 SSD 算法中,使用 VGG-16 模型<sup>[12]</sup> 作为 SSD 网络的 backbone 网络,原始的 SSD 算法中,用 Conv\_7 层代替全连接层 fc7,同时又添加 4 个额外层: Conv\_8、Conv\_9、Conv\_10、Conv\_11。通过卷积的方式为目标检测生成 10×10、5×5、3×3、1×1 的特征图,并将这些特征图输出到后续的分类与回归中。其中低层特征图对小目标进行检测,高层特征图对大目标进行检测。原始的 SSD 网络结构如图 1 所示。

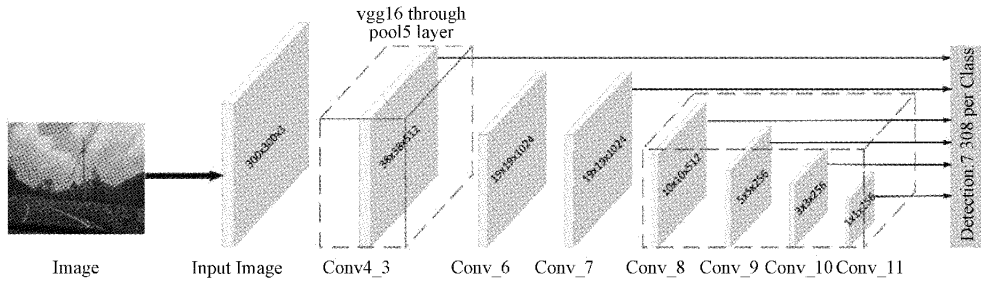


图 1 原始 SSD 网络

SSD 算法的目标函数分为两个部分: 计算相应的 default box 与目标类别的置信损失 (confidence loss) 以及对应的定位损失 (location loss), 总体 Loss 计算为:

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (1)$$

其中,  $N$  是 match 到 Ground Truth 的 default box 数量; 而  $\alpha$  参数用于调整 confidence loss 和 location loss 之间的比例, 默认  $\alpha=1$ :

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k smooth_{L1}(l_i^m - \hat{g}_i^m) \quad (2)$$

对于置信度误差, 其采用 softmax loss, 定义如下:

$$L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_j^p) - \sum_{i \in Neg} \log(\hat{c}_i^o) \quad (3)$$

### 2 改进 SSD 目标检测算法

原始 SSD 算法使用 VGG-16 作为骨干网络, 而 VGG 在内存和时间的计算要求很高, 且卷积通道很大, VGG 模型并不高效。所以采用 Resnet-50 模型作为 SSD 算法的 backbone 网络如图 2 所示, Resnet 网络中的残差模块将输

入和输出直接通过简单的恒等映射完成, 不需要引入额外的参数, 减少计算负担, 并且降低了过拟合风险, 准确度比 VGG 网络要高。使用 sub-pixel 采样代替上采样对特征图进行处理, 并通过嵌入深度可分离卷积层<sup>[13]</sup> (depthwise separable convolution) 避免了不同尺度特征图在融合过程中所产生的噪声淹没特征信息。使用 MFPN 网络代替传统的 FPN 金字塔网络, 使特征图能够获取全局信息, 增强主干网路对低层特征图的信息挖掘, 避免丢失边缘信息。将生成的预测特征层通过 MFPN 网络结构进行处理, MFPN 模型增加了对于特征映射的分辨率, 获得更多关于交通标志的有用信息, 使得 SSD 算法对小目标检测的鲁棒性更高。

#### 2.1 Resnet 残差网络结构

在改进的 SSD 算法中, 使用 ResNet-50 作为算法的骨干网络, 首先将 300×300 图像输入到 Resnet-50 分类网络中如图 3 所示, 而在 Resnet-50 网络模型中, 只保留 Conv4\_x 之前的层结构, 改动 Conv4\_x 中的 block1 残差结构, 将 3×3 卷积核的步距修改为 1, 特征层经过 Conv4\_x 所对应的一系列残差结构后, 得到 38×38×1024 的特征矩阵。再添加 5 个额外层结构, 每一个额外都通过 CN<sup>[14]</sup>、Batch

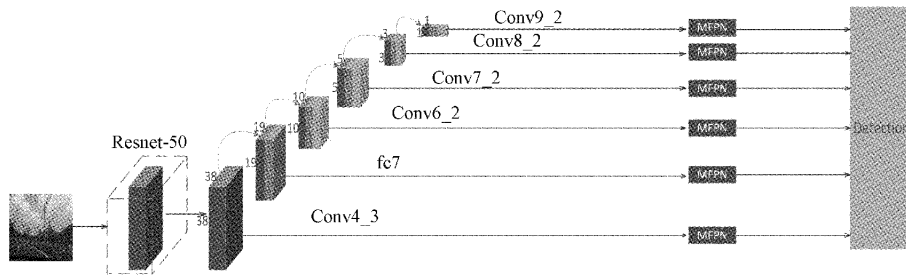


图 2 改进的 SSD 网络

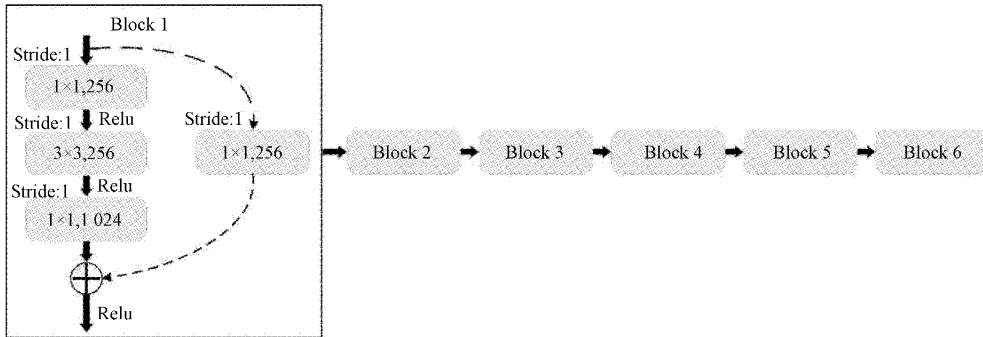


图 3 Resnet 残差网络结构

Normalization(归一化)、RELU 来进行生成如图 4 所示,为防止交通标志目标信息在下采样过程中由于噪声而导致某些特征信息丢失,同时保证轻量化,故采用  $3 \times 3$  深度可分离卷积层代替常规  $3 \times 3$  卷积层来进行下采样。深度可分离卷积将提取过程分为逐通道卷积和逐点卷积。假设使用  $3 \times 3$  的卷积层将大小为  $H \times W$ 、维度为  $x$  的特征图提取为维度  $y$ ,计算量为  $9 \times H \times W \times x \times y$ 。而在深度可分离卷积层中,使用  $C$  个  $3 \times 3 \times 1$  卷积核对  $H \times W \times C$  的图像进行单通道特征提取后,再通过  $N$  个  $1 \times 1 \times C$  卷积核对经过深度卷积的图像进行通道堆叠,计算量为  $9 \times x \times H \times W + x \times y \times H \times W$ ,计算量只有普通卷积层的  $1/9$ 。减少了改进网络的计算量,提高了检测速度。

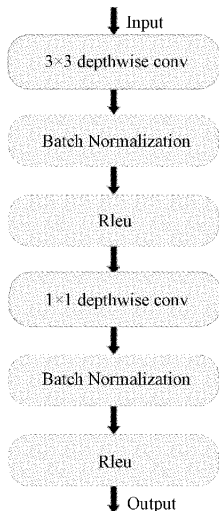


图 4 深度可分离卷积层结构

添加 BN 层的目的是为了解决在训练过程中,中间层生成数据时发生的问题,防止产生梯度消失梯度爆炸,而且可以使用更高的学习率,提高训练速度。

### 2.2 特征图调整方法

主干网络生成不同尺寸的特征图后,需要使用采样方法对特征图进行调整。一般情况下会使用上采样对特征图进行放大处理,从而可显示在更高分辨率的设备上。常用的上采样方法有:最近邻算法、双线性插值算法、双三次插值算法等,即在原有图像的基础上在像素点之间采用合适的插值算法插入新的元素。用下采样方法对特征图进行缩小,使图像符合显示区域的大小。

本文采用 pathway layer 来代替下采样,使用 sub-pixel convolution layer 来代替上采样。sub-pixel convolution layer 主要是将低分辨率的特征图  $N \times (C \times r \times r) \times W \times H$  通过卷积和多通道间的重组从而得到高分辨率的特征图  $N \times C(H \times r) \times (W \times r)$ ,其中  $N$  为图像的个数, $C$  为通道, $W$  为宽度, $H$  为高度, $r$  为上采样或下采样的倍数。而 pathway layer 则是将高分辨率的特征图转化为低分辨率的特征图,特征图调整方式如图 5 所示。

### 2.3 MFPN 模型

现有的 SSD 改进算法大多是基于特征融合的,利用这些特征图对检测目标进行识别、分类和定位。传统的 FPN 模型通过自底向上移动的图像中提取特征,构建一组金字塔结构,利用上采样实现特征融合,但是使用 FPN 模型会导致网络注重特征的优化,从而对目标的检测精度降低。因此,在本文改进的 SSD 网络中生成一系列的特征图后加入优化的 MFPN 模型如图 6 所示,MFPN 模型包括:一个

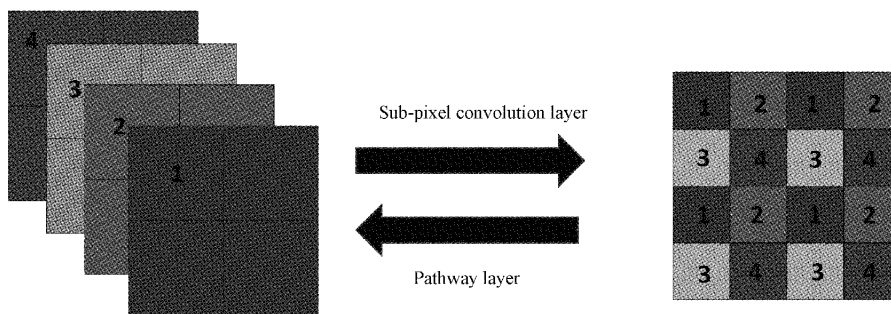


图 5 特征图调整过程

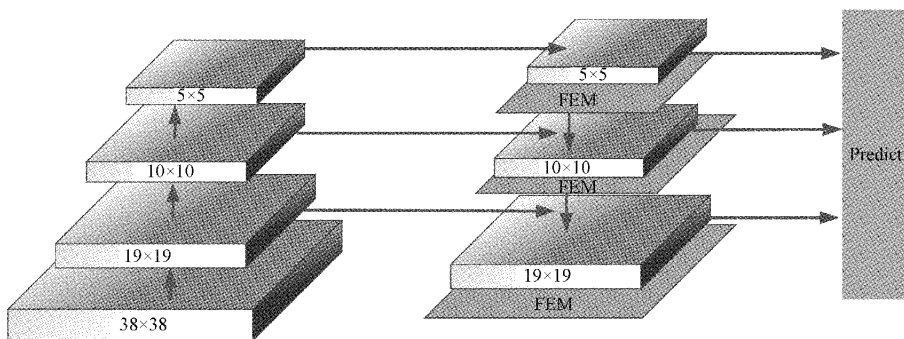


图 6 MFPN 网络模型

自底向上和一个自顶向下的线路横向连接、FEM 模块如图 7 所示。自底向上的过程在经过卷积核计算后,特征图尺寸会越来越小。而自顶向下是将高层特征图进行上

采样处理,特征图经过采样处理后和自底向上的特征图进行融合即为横向连接过程,且两层特征在空间中尺寸相同。

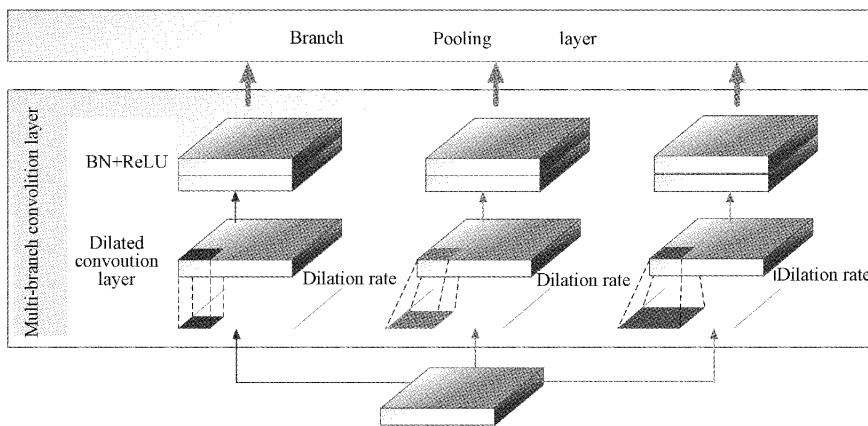


图 7 FEM 网络模块

FEM 模块分为多分支卷积层 (multi-branch convolution layer) 和分支池化层 (branch pooling layer), 主要是根据在检测过程中不同尺度的交通标志, 利用扩张卷积自适应地学习每个特征图中不同的感受野, 从而提高对交通标志的检测精度。多分支卷积层包括扩张卷积、BN 层和 ReLU 激活层。多分支卷积层通过扩张卷积 (dilated convolution layer) 为输入的特征图提供不同的感受野, 利用平均池化层融合来自 3 个支路的不同感受野的交通标志信息, 提高多尺度的精度预测。在 3 个支路中, 扩张卷积使用相同大小的卷积核, 扩张速率 (dilation rate) 却有所

不同。扩张卷积的感受野定义为:

$$r_1 = d \times (k - 1) + 1 \tag{4}$$

$$r_n = d \times (k - 1) + r_{n-1} \tag{5}$$

其中,  $k$  表示内核尺寸,  $r_i$  表示膨胀率,  $d$  表示卷积的步距。

分支池化层用于融合不同分支的信息, 避免引入附加参数。定义为:

$$y_p = \frac{1}{B} \sum_{i=1}^B y_i \tag{6}$$

$y_p$  表示分支池化层的输出,  $B$  表示并行分支的数量,



在本文中设置  $B=3$ 。

### 3 数据集和实验设置

#### 3.1 图像数据集

本实验采用从中国交通标志资料库 CTSDB 中选取的交通面板数据集 TPD<sup>[15]</sup>, 中国交通标志分为如图 8 所示的 43 类, 数据类别分为: mandatory、prohibitory、warning。其数据集包括了各种天气环境下记录的自然交通情景以及许多的天气情况如图 9 所示。随机选取数据集中的 10 000 张图片作为实验对象, 以 7:3 随机划分训练集和测试集, 其中 7 000 张图片作为训练集, 另外 3 000 张图片作为测试集, 以验证改进 SSD 算法的有效性。

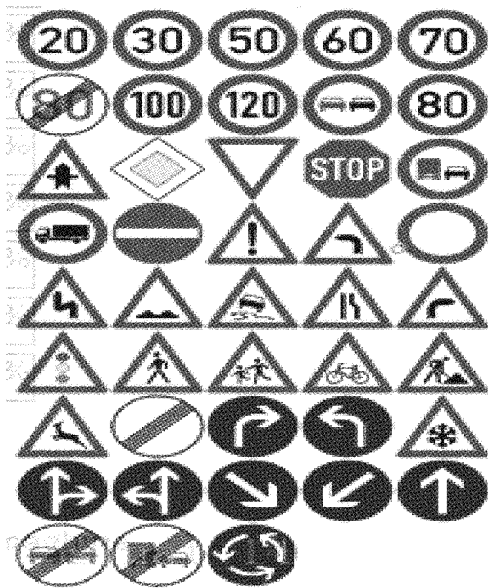


图 8 中国交通标志

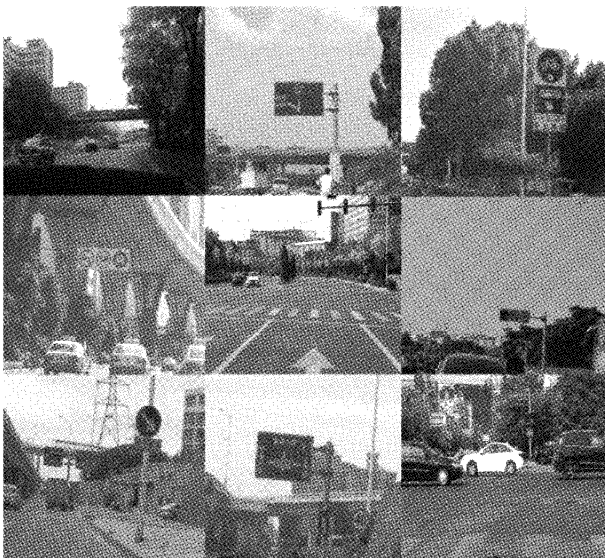


图 9 中国交通标识检测数据集

#### 3.2 硬件环境

实验中选取了 Intel(R) Core(TM) i7-10700 CPU@2.90 GHz 处理器, 16GRAM 和 Nvidia GTX 3060。所有实验都是在 pytorch1.9、cuda11.1、python3.7 和 win10 环境下进行的。SSD 算法是在实验环境下使用 pytorch 环境进行搭建的。

#### 3.3 训练参数

通过设置下列参数来训练 SSD 算法, 输入图像为  $300 \times 300$ , 将 batch\_size 设置为 32, 学习率 learning\_rate 设置为 0.0005, 共进行 200 次迭代训练。

#### 3.4 评价指标

主要通过召回率、精确度、mAP 等参数对 SSD 算法进行评价。在目标检测中, FPS 即每秒处理的图片数量。被正确检测的正类样本总数为 TP, 被误检的正类总数为 FP, 被正确检测的负类样本总数为 TN, 被误检的负类样本为 FN。因此把精确度 (precision)、召回率 (recall) 定义为:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$mAP = \frac{1}{C} \sum_{c \in C} AP(c) \quad (10)$$

AP 值代表某一类的综合评价, AP 值越高代表某一类别的精度越高, mAP 值是对整体网络的评估, 其中 C 代表数据集所包含类别数, c 表示单个类别。

## 4 结果分析

#### 4.1 训练结果分析

实验通过将本文改进的 SSD 算法与原始的 SSD 算法进行比较, 在训练时保持所有模型参数一致。模型的总损失变化如图 10 所示。随着迭代次数的增加, 损失值逐步减小, 在 40 轮后没有产生过拟合或欠拟合等现象, 证明模型的识别精度在不断地提高。

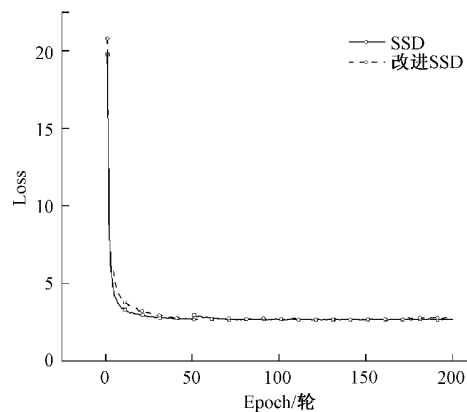


图 10 训练损失曲线

4.2 消融实验结果与分析

本文重点研究实时检测交通标识算法的实现,为实际的应用提供可能性。消融实验是以 SSD 算法为基础结

合不同的创新方法进行训练和性能测试,在保证实时检测的前提下,从而提高算法的检测精度,实验结果如表 1 所示。

表 1 消融实验结果

实验	VGG	ResNet	Sub-pixel	MFPN	mAP/%	FPS	时间/s
1	√				86.3	88.5	0.001 0
2		√			87.5	85.2	0.002 2
3		√	√		88.2	84.2	0.001 8
4		√		√	88.7	84.5	0.001 5
5		√	√	√	90.5	87.2	0.001 0

从表 1 中对每项改进策略的有效性进行分析,可以看出每个改进模块都对模型有不同程度的贡献,将原始的 VGG-16 骨干网络换成 ResNet-50 骨干网络,且在额外添加的特征层中加入 BN 层,可将检测精度提高 1.2%,但检测速度会有所下降;使用 Sub-pixel 采样方法可使特征图分辨率增强,有效增强交通标志的特征表达能力,可使检测精度达到 88.2%,在 ResNet-50 网络的基础上加入了 FPN 模型,使得对于小目标的检测更为有效,检测精度也达到 88.7%。结合 3 种不同的改进方法,相较于原始的 SSD 检测算法,本文所提出的改进 SSD 算法将检测精度提高了 4.2%,且检测速度也相同,表明本文提出的改进 SSD 算法有较高的使用价值。

4.3 与其它算法性能比较

将本文改进的 SSD 算法与其它性能较好的目标检测算法进行比较,其中包括:YOLOv3、YOLOv4、Faster RCNN、FSSD、Alexnet、MF-SSD、SSD。

实验结果如表 2 所示,可以看出改进的 SSD 算法有较好的检测结果。Faster RCNN 的检测精度为 83.2%,而检测速度为 17.2,耗费时间较长。AlexNet 的检测精度稍逊色与其它检测算法,检测速度也相对较慢。YOLOv3 和 YOLOv4 在准确率方面达到了 85.7%和 86.8%,检测速度也较快。改进后的 SSD 算法与表现较好的 MF-SSD 检测算法相对比,改进的 SSD 算法准确率得到了提升,mAP 值达到了 90.5%,且检测速度也达到了 87.2,表明模型有较高的检测速度,满足移动端实时检测。

表 2 不同算法性能对比

模型	mAP/%	FPS	时间/s
Faster RCNN	83.2	17.2	0.007 8
YOLOv3	85.7	82.4	0.001 7
YOLOv4	86.8	87.6	0.001 3
FSSD	87.2	72.7	0.001 8
AlexNet	82.5	22.7	0.005 6
SSD	86.3	88.5	0.001 0
MF-SSD	88.7	86.5	0.001 2
改进 SSD	90.5	87.2	0.001 1

4.4 GTSDDB 数据集实验结果

为了进一步检验改进的 SSD 检测模型对其它交通标志的检测效果,在 GTSDDB 数据集上进行实验,实验结果如表 3 所示。通过对比表 3 中的检测结果,改进后的 SSD 算法在 GTSDDB 数据集中 mAP 值达到了 89.4%,相较于原始的 SSD 算法提高了 3.1%,表明改进后的 SSD 算法对其它交通标志也具有较好的检测精度。

表 3 检测性能对比

模型	mAP/%	FPS	时间/s
Fast RCNN	83.2	15.6	0.067 0
YOLOv4	87.6	87.0	0.001 2
SSD	86.3	88.5	0.001 0
改进 SSD	89.4	85.2	0.001 2

图 11 为在 CCTSDB 数据集的部分检测结果,左边为

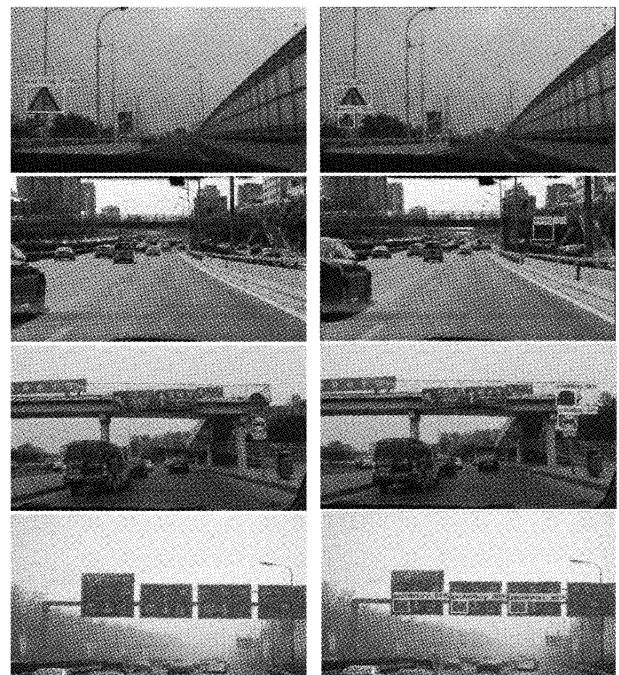


图 11 图片检测结果

原始 SSD 算法的检测结果,右边为改进后的 SSD 算法的检测结果,在不同道路环境下,本文所提出的改进的 SSD 检测算法在检测精度上均有所提升,极少出现漏检、误检等情况。

## 5 结 论

本文提出了一种基于 SSD 算法的交通标识检测模型,首先使用 resnet-50 网络作为 SSD 模型的 backbone 网络,又对特征图进行 sub-pixel 采样处理,后将低层特征与高层特征相融合,构建 MFPN 模型金字塔结构,使得对小目标检测精度更高,实验结果表明本文 SSD 算法检测精度达到了 90.5%。此外,本文算法对不同背景目标也具有很强的适应性,且在查准率和查全率都优于原始 SSD 算法。虽然该方法在检测小目标图像方面有了很大的改进,但是在实际应用方面的精度还有较大的提升空间。

## 参考文献

- [1] GUDIGAR A, CHOKKADI S, RAGHAVENDRA U. A review on automatic detection and recognition of traffic sign[J]. *Multimedia Tools and Applications*, 2016, 75(1):333-364.
- [2] LI H, SUN F, LIU L, et al. A novel traffic sign detection method via color segmentation and robust shape matching[J]. *Neurocomputing*, 2015, 169(2): 77-88.
- [3] 陈玉敏,李森,房晓丽. 基于时空融合加速的 Fast RCNN 运动车辆检测算法[J]. *电子测量技术*, 2020, 43(3):139-145.
- [4] 张熙,许浩,李沁程. 基于 SVM 的飞机驾驶舱头顶板字符识别方法研究[J]. *电子测量技术*, 2020, 43(22): 82-87.
- [5] 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.
- [6] REDMON J, FARHADI A. YOLOv3: An incremental improvement[C]. *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)*, 2018: 1-4.
- [7] LIU W, ANGUELOY D, ERHAN D, et al. SSD: Single shot multibox detector[C]. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016: 21-37.
- [8] 张长伦,张翠文,王恒友,等. 基于注意力机制的 NMS 在目标检测中的研究[J]. *电子测量技术*, 2021, 44(19):82-88.
- [9] MAHMOUD M, GUO P, XIN X. A novel method for traffic sign recognition based on DCGAN and MLP with PILAE algorithm[J]. *IEEE Access*, 2019(99):1-1.
- [10] GUDIGAR A, CHOKKADI S, RAGHAVENDRA U, et al. An efficient traffic sign recognition based on graph embedding features[J]. *Neural Computing and Applications*, 2019, 31(2): 395-407.
- [11] 于波,刘畅. 基于改进 SSD 算法的行人检测方法[J]. *电子测量技术*, 2021, 44(12):24-28.
- [12] 何其霖,穆平安. VGG 网络与多特征融合的遮挡人脸检测[J]. *电子测量技术*, 2021, 44(18):150-154.
- [13] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [J]. *ArXiv Preprint*, 2017, ArXiv:10.48500.
- [14] HU P, RAMANAN D. Finding tiny faces [C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017: 951-959.
- [15] ZHANG J, WANG W, LU C, et al. Lightweight deep network for traffic sign classification[J]. *Annals of Telecommunications*, 2020, 75(7): 369-379.

## 作者简介

詹华伟,工学硕士,副教授,主要研究方向为射频通信技术、无线传感器网络,人工智能。

E-mail: zhanhw@126.com

邹昊好(通信作者),硕士研究生,主要研究方向为目标检测。

E-mail: 434808476@qq.com

刘旭,硕士研究生,主要研究方向为目标检测。

E-mail: 781820258@qq.com

史水娥,工学硕士,副教授,主要研究方向为电子电路设计,可编程逻辑 FPGA 技术。

E-mail: 021044@htu.edu.cn