



# 计算机科学

COMPUTER SCIENCE

## 基于优化 YOLO-V4 的交通标志检测识别方法

潘惠苹, 王敏琴, 张福泉

引用本文

潘惠苹, 王敏琴, 张福泉. [基于优化 YOLO-V4 的交通标志检测识别方法](#)[J]. 计算机科学, 2022, 49(11): 179-184.

PAN Hui-ping, WANG Min-qin, ZHANG Fu-quan. [Traffic Sign Detection and Recognition Method Based on Optimized YOLO-V4](#)[J]. Computer Science, 2022, 49(11): 179-184.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [高分辨率斜视聚束 SAR 回波仿真加速算法研究](#)

Study on Acceleration Algorithm for Raw Data Simulation of High Resolution Squint Spotlight SAR

计算机科学, 2022, 49(8): 178-183. <https://doi.org/10.11896/jsjcx.210600066>

### [基于 DNGAN 的磁共振图像超分辨率重建算法](#)

Super-resolution Reconstruction of MRI Based on DNGAN

计算机科学, 2022, 49(7): 113-119. <https://doi.org/10.11896/jsjcx.210600105>

### [基于向量注意力机制 GoogLeNet-GMP 的行人重识别方法](#)

Person Re-identification Method Based on GoogLeNet-GMP Based on Vector Attention Mechanism

计算机科学, 2022, 49(7): 142-147. <https://doi.org/10.11896/jsjcx.210600198>

### [基于改进 YOLOv3 的机坪工作人员反光背心检测研究](#)

Study on Reflective Vest Detection for Apron Workers Based on Improved YOLOv3 Algorithm

计算机科学, 2022, 49(4): 239-246. <https://doi.org/10.11896/jsjcx.210200119>

### [基于元迁移的太阳能电池板缺陷图像超分辨率重建方法](#)

Super Resolution Reconstruction Method of Solar Panel Defect Images Based on Meta-transfer

计算机科学, 2022, 49(3): 185-191. <https://doi.org/10.11896/jsjcx.210100234>

# 基于优化 YOLO-V4 的交通标志检测识别方法

潘惠苹<sup>1,2</sup> 王敏琴<sup>3</sup> 张福泉<sup>4</sup>

1 广东工商职业技术大学计算机学院 广东 肇庆 526000

2 菲律宾克里斯汀大学国际学院 马尼拉 1004

3 华南理工大学计算机科学与工程学院 广州 510000

4 北京理工大学计算机学院 北京 100081

**摘要** 交通标志检测识别是自动驾驶系统的核心功能,为了实时准确地识别交通标志,在 YOLO-V4 的基础上进行改进,并结合了空间金字塔池化(Spatial Pyramid Pooling,SPP)模块。首先,为了提高分辨率和增大感受野,将原特征图 3 个尺度的分辨率更改为  $26 \times 26$  和  $52 \times 52$ ;然后,在连接层中添加 SPP 模块,消除网络对固定尺度的约束,在最大池化层中得到最优特征,改善网络性能。实验中,利用行车记录仪采集各种交通标志图像,与其他优秀方法相比,所提方法取得了更优的性能,其平均检测识别准确率达 99.0%,平均检测时间为 0.449 s,达到了实时检测的要求。

**关键词**: 交通标志识别;感受野;YOLO-V4;最大池化;空间金字塔池化;分辨率

中图法分类号 TP391

## Traffic Sign Detection and Recognition Method Based on Optimized YOLO-V4

PAN Hui-ping<sup>1,2</sup>, WANG Min-qin<sup>3</sup> and ZHANG Fu-quan<sup>4</sup>

1 College of Computer Science, Guangdong Business and Technology University, Zhaoqing, Guangdong 526000, China

2 Center for International Education, Philippine Christian University, Manila 1004, Philippine

3 College of Computer Science and Engineering, South China University of Science and Engineering, Guangzhou 510000, China

4 College of Computer Science, Beijing Institute of Technology University, Beijing 100081, China

**Abstract** Traffic sign detection and recognition is the core function of automatic driving system. In order to identify traffic signs in real time and accurately, a method is improved on the basis of YOLO-V4 and combined with the spatial pyramid pool(SPP) module. Firstly, to increase the resolution and receptive field, the resolution of the three scales of the original feature map is changed to  $26 \times 26$  and  $52 \times 52$ . Then, SPP module is added to the connection layer to eliminate the constraints of the network on the fixed scale, obtain the optimal characteristics in the maximum pooling layer and improve the network performance. Experiment uses the tachograph to collect various traffic sign images, compared with other excellent methods, the proposed method achieves better performance. The average detection and recognition accuracy of the proposed method is 99.0%, and the average detection time is 0.449 s, which meets the requirements of real-time detection.

**Keywords** Traffic sign recognition, Receptive field, YOLO-V4, Maxpooling, Spatial pyramid pool, Resolution

交通标志识别<sup>[1]</sup>指利用机器视觉技术进行标志识别,其应用领域广泛,如自动驾驶辅助系统<sup>[2]</sup>(Automatic Driving Assistance System, ADAS)、自动驾驶、交通控制、公路网维护等。随着深度学习的兴起,目标检测和识别任务<sup>[3-4]</sup>得到了极大发展,但目前尚无可用的交通标志检测数据库或分析系统。影响识别的因素非常多,如部分遮挡、多视点、光照和温度条件等,因此交通标志识别是一项具有挑战性的任务。

交通标志通常根据不同国家的要求,使用基本样式和标准颜色,这些标志大多相似。文献[5]运用图像分割技术,并

结合卷积神经网络(Convolutional Neural Networks, CNN)模型,提出一种交通标志识别方法,称为(My REsidual-Squeeze and Excitation, MRESE)。然而,在测试过程中,决策制定延迟是关键问题,会受到环境和实际因素的影响。文献[6]将种群概念引入人工蜂群算法中,然后将其应用于德国交通标志数据集的检测,但该方法局限于圆形交通标志。文献[7]提出了一种级联的 R-CNN 来获取金字塔中的多尺度特征,通过点积和 softmax 来获得加权多尺度特征,以细化特征,突出交通标志的特征。为了消除尺度和旋转的影响,文献[8]中的

到稿日期:2022-03-27 返修日期:2022-05-19

基金项目:国家自然科学基金面上项目(61871204);福建省科技厅引导性项目(2018H0028)

This work was supported by the General Project of the National Natural Science Foundation of China of the(61871204) and Guiding Project of Fujian Provincial Department of Science and Technology(2018H0028).

通信作者:潘惠苹(15588022@qq.com)

成像是在立体空间中完成的,使用捕获的图像对生成虚拟图像,用于识别标志。该方法对数据库的构建具有严格的要求。

目前很多交通标志识别以 SSD<sup>[9]</sup>, R-CNN<sup>[10]</sup>, YOLO v3<sup>[11-12]</sup> 等检测方法为基础,但这些方法的局限是固有的网络约束,检测的适应性较差。因此,本文对其缺点进行改善,着眼于禁止符号的识别和理解。其主要创新工作总结如下:1)基于 SPP 原理对 YOLO-V4 构建的特征提取原始架构进行增强,消除网络对固定尺度的约束,在最大池化层中得到最优特征,通过 SPP 层建立网络层,以改善 YOLO-V4 的性能;2)利用空间金字塔技术,对相同网络层中不同等级的局部特征进行组织和聚合,对多尺度目标函数进行更全面的训练。实验结果表明,使用大量尺度时,所提方法可有效加快检测速度,并具有较高的识别准确性。

## 1 优化的 SPP-YOLO-V4 网络

### 1.1 SPP 模块

空间金字塔池化 (Spatial Pyramid Pooling, SPP) 可以避免图像由于缩放、拉伸、剪裁等操作造成的信息失真,提供不受输入大小影响的输出,这是滑动窗口池化技术无法实现的<sup>[13]</sup>,另外 SPP 可使用多个尺度的池化,而滑动窗口池化仅使用一个窗口尺度。SPP 模块的基本结构如图 1 所示。可以看出,由于输入尺寸是灵活的,SPP 可以将不同尺度处得到的特征进行合并,转换的特征向量与全连接层的维度相同,同时缓解图像特征的重复提取问题。其具体说明请参考文献<sup>[14]</sup>。

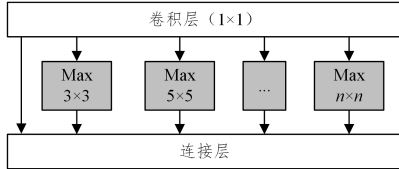


图 1 SPP 模块的基本结构

Fig. 1 Basic structure of SPP module

### 1.2 简化的 SPP-YOLO-V4 结构

YOLO-V4 是集成了 YOLO v1, YOLO v2 和 YOLO v3 特征的高精度实时单阶段目标检测算法,其在残差模块中构建了跨阶段部分网络 (Cross Stage Partial Network, CSPNet),特征层为输入,更高层的特征信息为输出。这表明在输出和输入之间,模型在 ResNet 模块中的学习目标存在差异,由此实现了残差学习,同时降低了模型参数,增强了特征学习。

本文中,416 × 416 大小的输入图像上采用 3 个尺度,得出的 3 个尺度分别为 13 × 13, 26 × 26, 52 × 52。这样的设置不利于低质量交通标志识别图像的检测。尺寸为 13 × 13 的特征图可能会使得语义信息严重丢失,从而造成识别错误。考虑到交通标志识别的应用环境,在原网络的基础上作了一些改动,将原特征图 3 个尺度的分辨率 13 × 13, 26 × 26 和 52 × 52 更改为了 26 × 26 和 52 × 52,这样的分辨率变化可以让特征图的信息更多地保留下来,同时 2 个尺度有利于提高运行效率,且并不影响检测识别的准确率。最终的网路结构如图 2 所示。

首先将特征层卷积 3 次,其次利用不同大小的最大池化核,对输入特征层进行最大池化,由此改善网络感受野。在

卷积和上采样后,将特征层与从 CSPDarkNet53 得到的特征层串联在一起,以实现特征融合。然后,执行下采样,压缩高度和宽度,最终与之前的特征层堆栈存储在一起,以实现更多的特征融合(5 次)。预测模块利用从网络中提取出的特征进行预测,以 13 × 13 网格为例,其等于将输入图像分割为 13 × 13 的方格,则每个方格将被预设 3 个先验框。网络的预测结果将会调整这 3 个先验框的位置,最后通过非极大值抑制 (Non Maximum Suppression, NMS) 算法进行过滤,从而得到最终的预测帧。

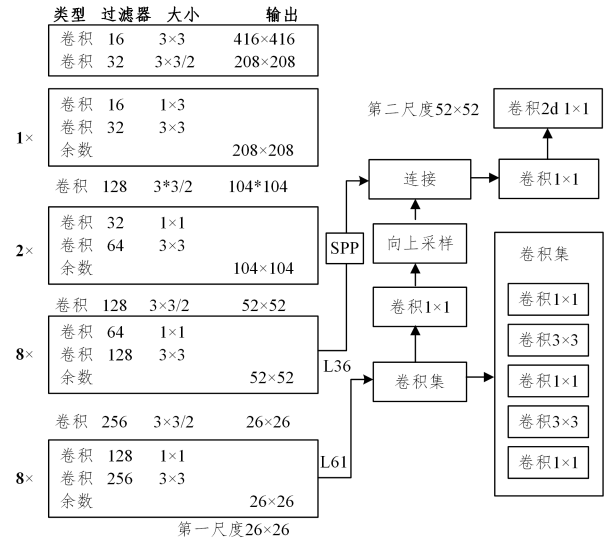


图 2 改进后的 YOLO-V4 网络的结构图

Fig. 2 Structure diagram of improved YOLO-V4 network

## 2 基于优化 YOLO-V4 的交通标志检测识别方法

### 2.1 坐标转换操作

为了使用 YOLO-V4 检测并识别交通标志,首先利用 BBox 标签工具制定每个标志的包围框以进行训练,本文利用 4 种类型的标签完成标记,一张图像包含多个包围框。该阶段中,使用单类检测器模型,对一个交通符号使用单个模型进行训练。目标坐标形式为  $(x_1, y_1, x_2, y_2)$ ,作为包围框标记工具的输出值。为了使输出格式是 YOLO 目标坐标格式,必须对包围框坐标进行转换,转换过程如下:

$$dw = 1/w, x = \frac{(x_1 + x_2)}{2} \times dw, dh = 1/h \quad (1)$$

$$y = \frac{(y_1 + y_2)}{2} \times dh \quad (2)$$

$$w = (x_2 - x_1) \times dw \quad (3)$$

$$h = (y_2 - y_1) \times dh \quad (4)$$

其中,  $w$  为图像宽度,  $dw$  为图像绝对宽度,  $h$  为图像高度,  $dh$  为图像总高度,  $(x, y, w, h)$  表示目标的中心点、宽度和高度。

### 2.2 检测与识别

利用改进 SPP-YOLO-V4 进行交通标志识别的过程如图 3 所示。初始阶段,将图像输入分割为  $S \times S$  的网格,每个网格生成  $K$  个边界。然后,实施 CNN,从图像和预测中提取目标特征,将  $K$  个包围框的最大置信度  $IoU_{pred}^{truth}$  与交并比阈值  $IoU_{th}$  进行比较。若  $IoU_{pred}^{truth} > IoU_{th}$ ,则说明包围框中包含目标,否则包围框不包含目标。最后,系统选择具有最高预测

概率的类别作为目标类别,执行最大局部探索。为了抑制冗余包围框,采用了 NMS,NMS 的具体描述如算法 1 所示。

#### 算法 1 非最大值抑制(NMS)

输入:包围框候选 P\_B[0,⋯,n]

输出:最终包围框 F\_B

1. 通过得分降序的方式对 P\_B 排序;
2. while P\_B 中要检查的包围框非空 do
3.   max\_box←得分最高的包围框;
4.   for each box∈P\_B[0,⋯,n] do
5.     if IOU(max\_box,box)>既定阈值 then
6.       包围框得分←0
7.   移除所有得分为 0 的包围框;
8.   将 max\_box 插入 F\_B;
9.   从 P\_B 中移除 max\_box.

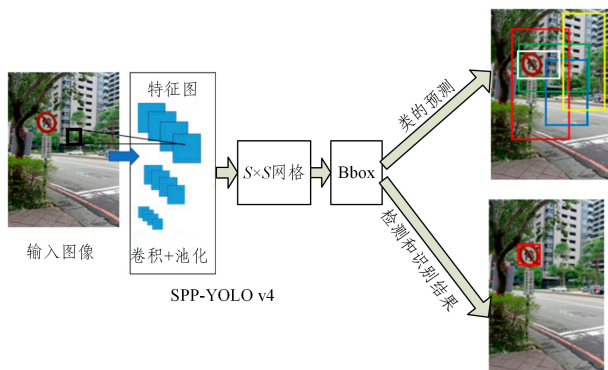


图 3 交通标志识别框架图

Fig. 3 Frame diagram of traffic sign identification

在本文方案中,SPP-YOLO-V4 利用卷积层采样实现最大池化层的最优性能。创建的特征图为固定长度表征,在高斯零均值分布中对 Softmax 分类层和包围框回归进行初始化,标准偏差分别为 0.01 和 0.001。全局学习率为 0.001,动量为 0.9,参数衰减率为 0.0005。学习率参数决定了使用最近一批数据进行学习的进展情况。

#### 2.3 损失函数

包围框的损失函数包括 3 个部分<sup>[15]</sup>,可以表示为:

$$Loss = E\_Coord + E\_BBox + E\_Conf \quad (5)$$

其中, $E\_Coord$  为预测中心坐标的损失, $E\_BBox$  为预测包围框的宽度和高度的损失, $E\_Conf$  为预测置信度的损失。

这些度量的计算式如下:

$$E\_Coord = \beta' \sum_{i=0}^{\xi} \sum_{j=0}^{\beta} I_{ij}^{obj} [(x_i - x_i^*)^2 + (y_i - y_i^*)^2] \quad (6)$$

$$E\_BBox = \beta' \sum_{i=0}^{\xi} \sum_{j=0}^{\beta} I_{ij}^{obj} [(\sqrt{w_i} - \sqrt{w_i^*})^2 + (\sqrt{h_i} - \sqrt{h_i^*})^2] \quad (7)$$

$$E\_Conf = \sum_{i=0}^{\xi} \sum_{j=0}^{\beta} I_{ij}^{obj} (c_i - c_i^*)^2 + \beta'' \sum_{i=0}^{\xi} \sum_{j=0}^{\beta} I_{ij}^{obj} (c_i - c_i^*)^2 \quad (8)$$

其中, $(x_i, y_i)$  为预测包围框的位置; $(x_i^*, y_i^*)$  表示从训练数据中得到的实际位置; $w_i$  和  $h_i$  分别表示预测包围框的宽度和高度; $\beta'$  用于控制预测包围框的预测位置损失; $\beta''$  用于控制单个网格中无目标的损失; $c_i$  为置信度得分; $c_i^*$  表示预测包围框和实际包围框的交集部分。此外,YOLO-V4 利用 sigmoid 函数作为预测激活函数的工具。当同一个目标有两个标签时,sigmoid 函数能够高效求解问题。

### 3 实验结果与分析

本文的硬件平台是英特尔酷睿 i3-M350 CPU@ 双核 2.20GHz,8GB 的 DDR2 内存,GPU 为 Nvidia RTX2080Ti,显存为 11GB。开发环境为 Python 3.5.0 语言的 PyCharm 集成开发工具。利用 Python 编写的 YOLO 注释框架进行数值格式转换,使其可被 YOLO 读取。对比方法有文献[5]提出的 MRESE 方法、文献[7]提出的 R-CNN 方法以及文献[12]提出的 YOLO-V4 方法。

#### 3.1 数据集

本文从行车记录仪中采集了车辆在晴天和夜间行驶时的交通标志图像,并进行了手动处理。相机图像分辨率为 1920×1080 像素,从中提取出交通标志图像。本文还使用华为荣耀手机相机,采集交通标志图像。在使用交通标志图像进行训练之前,先对其进行裁剪和注释。此外,图像中主要为交通禁止标志,包括 235 个禁止驶入图像( $D_1$ )、280 个禁止长时间停留图像( $D_2$ )、190 个限速图像( $D_4$ )以及 220 个禁止停车图像( $D_3$ )。将数据集的 75% 用于训练,25% 用于测试。表 1 列出了各标志和解释。

表 1 交通禁止标志

Table1 Traffic sign of NO

ID	解释	标志
$D_1$	禁止进入	
$D_2$	禁止停留	
$D_3$	禁止停车	
$D_4$	限速	

#### 3.2 评价指标

平均精度均值(mean Average Precision,mAP)的计算式如下<sup>[16]</sup>:

$$mAP = \int_0^1 p(\tau) d\tau \quad (9)$$

其中, $p(\tau)$  为禁止标志检测的准确度。

精度和召回率的定义分别为:

$$Pr = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

其中,TP 为正警率,FP 为误警率,FN 为漏警率。

F1 得分是精度和召回率的调和值,值越高,表示性能越佳,其定义为:

$$F1 = 2 \frac{Recall \times Pr}{Recall + Pr} \quad (12)$$

#### 3.3 训练结果分析

数据增广对改善深度学习模型的性能非常重要,相关研究表明,数据增广能够显著改善图像分类的性能。因此,本文应用预处理步骤,在训练阶段中纳入数据增广。在数据增广过程中,系统执行多个操作,例如针对每张图像进行概率为 0.5 的旋转或最大 20° 的旋转,缩放范围为 10%,宽度和高度的变化范围为 0.2。此外,利用包围框标签工具<sup>[17]</sup> 识别交通

标志,以提供目标的坐标位置。输出为位置坐标系上的4个点。针对YOLO系列方法,在进行训练前,系统先将标签转换为YOLO格式。

表2列出了在8000轮训练后,所有类别的训练损失值、mAP等结果。可以看出,所提方法的训练模型能够有效地识别目标。在精度、召回率和F1得分方面具有一定优势,其

损失值也是各种方法中的最低值,只有0.0122,因此所提方法的稳定性和精确性更佳。这主要得益于本文改进后方法的分辨率高和感受野大,在L36到连接层中添加了SPP模块,保留了SPP所带来的优势,且本文方法的训练时间仅稍微长于原YOLO-V4。MRESE方法由于需要训练非常多的残差网络,其训练时间更长。

表2 各方法针对不同类别的性能比较

Table 2 Performance comparison of different methods for different categories

模型	损失数值	训练时间 ±10 min	ID	mAP / %	TP	FP	精度	召回率	F1得分
R-CNN	0.0143	2h30 min	D <sub>1</sub>	97.5	77	0	0.98	0.97	0.98
			D <sub>2</sub>	98.81	83	0			
			D <sub>3</sub>	99.92	62	1			
			D <sub>4</sub>	98.74	76	2			
YOLO-V4	0.0151	2h	D <sub>1</sub>	97.53	78	0	0.98	0.96	0.97
			D <sub>2</sub>	100	83	0			
			D <sub>3</sub>	99.85	61	3			
			D <sub>4</sub>	98.01	75	3			
MRESE	0.0131	2h40 min	D <sub>1</sub>	97.50	77	0	0.98	0.97	0.97
			D <sub>2</sub>	98.81	83	0			
			D <sub>3</sub>	99.92	59	0			
			D <sub>4</sub>	97.75	72	5			
本文方法	0.0122	2h10 min	D <sub>1</sub>	97.51	78	0	0.99	0.99	0.99
			D <sub>2</sub>	98.82	83	0			
			D <sub>3</sub>	99.90	62	1			
			D <sub>4</sub>	98.94	79	3			

图4给出了本文SPP-YOLO-V4训练过程的可靠性,其中,横坐标为迭代次数,蓝色点状线条为损失函数值,红色实线条为mAP值,最大迭代次数为8000。由于初始时未掌握任何先验知识,必须在训练开始阶段使用高学习率。但随着神经网络中可用数据的增加,权重随之调整,学习率随着时间推移而变低。开始时,学习率为0.001,并在6400次迭代中保持不变。由图4可以看出,所提方法在训练过程中的稳定性较佳。

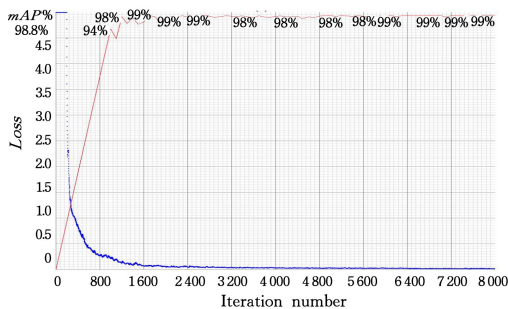


图4 训练过程中的可靠性情况(电子版为彩图)

Fig. 4 Reliability during training

### 3.4 测试结果分析

为了研究不同尺度所带来的影响,在同类方法中进行比较,以YOLO-V4为例,比较的尺度有(0.1,0.1),(0.2,0.2)和(0.3,0.3)。针对不同尺度的测试准确度比较如表3所列。这里的检测准确度区别于3.2节的指标,本文采用重叠度(Intersection over Union, IoU)数值代表检测准确度,将其视为检测结果和真实结果之间的关系。

$$IoU = \frac{S_{pr} \cap S_{gt}}{S_{pr} \cup S_{gt}} \quad (13)$$

其中, $S_{pr}$ 表示实际检测结果的区域, $S_{gt}$ 表示真实结果的区域。

文献[12]提出的方法使用YOLO-V4的基础配置,使用了DarkNet提供的最优权重, $scale=(0.1,0.1)$ 。虽然优化了准确率,但未考虑检测时间。从表3中可发现,本文方法的整体准确度高于原YOLO-V4,与原始方法相比,本文的SPP包含更多层,因此处理时间也增加了,尺度 $scale=(0.1,0.1)$ 时的平均检测时间为0.459s,当 $scale=(0.2,0.2)$ 时,检测时间为0.450s,当 $scale=(0.3,0.3)$ 时,检测时间为0.445s。结果表明,与使用 $scale=(0.1,0.1)$ 相比,使用 $scale=(0.3,0.3)$ 可加快检测速度。基于实验结果可得出以下结论:

(1)若希望识别系统得到更高的准确度,则可使用原始尺度 $scale=(0.1,0.1)$ ;

(2)若希望系统缩短检测时间,则可使用尺度 $scale=(0.3,0.3)$ 。

表3 针对不同尺度的测试准确度的比较

Table 3 Comparison of test accuracy for different scales

模型	尺度	准确度				均值
		D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	
YOLO-V4	(0.1,0.1)	0.931	0.914	0.901	0.933	0.919
	(0.2,0.2)	0.889	0.961	0.894	0.880	0.906
	(0.3,0.3)	0.921	0.910	0.900	0.913	0.911
本文方法	(0.1,0.1)	0.951	0.969	0.961	0.959	0.956
	(0.2,0.2)	0.889	0.982	0.911	0.889	0.918
	(0.3,0.3)	0.937	0.989	0.919	0.938	0.945

后续使用更高准确度的标准,YOLO-V4和本文方法均

使用尺度参数  $scale=(0.1,0.1)$ 。实验随机使用 12 个不同尺度的禁止标志图像,对 4 种方法进行了测试比较。表 4 列出了测试准确度和检测时间的数值,可以看出,本文的平均准确度达到 99.0%,平均检测时间为 0.449s,因此所提方法在对比的 4 种方法中取得了更优的性能。可以

推断出,所提方法的上采样和最大池化方面带来了显著收益,最大池化从相邻区域中选择最大值,以小幅删除目标图像中的某些最大频噪声,由此可以在后续采样层中对卷积子采样进行更好的操作,通过这些优点,SPP 能够改善骨干网络的性能。

表 4 不同方法的测试准确度和检测时间

Table 4 Test accuracy and detection time of different methods

No.	R-CNN		YOLO-V4		MRESE		本文方法	
	准确度	时间/s	准确度	时间/s	准确度	时间/s	准确度	时间/s
1	0.960	0.459	0.987	0.448	0.973	0.456	0.994	0.454
2	0.791	0.453	0.963	0.442	0.826	0.459	0.993	0.452
3	0.991	0.459	0.991	0.435	0.994	0.448	1.000	0.451
4	0.970	0.473	0.990	0.448	0.994	0.457	0.997	0.443
5	0.951	0.445	0.990	0.447	0.931	0.452	0.982	0.457
6	0.900	0.462	0.922	0.448	0.963	0.443	0.981	0.439
7	0.907	0.460	0.870	0.449	0.901	0.451	0.982	0.448
8	0.787	0.453	0.880	0.462	0.841	0.460	0.996	0.449
9	0.869	0.457	0.982	0.464	0.947	0.447	0.993	0.433
10	0.863	0.447	0.961	0.442	0.866	0.459	0.992	0.450
11	0.957	0.453	0.960	0.443	0.990	0.459	0.991	0.451
12	0.881	0.452	0.632	0.471	0.961	0.466	0.989	0.459
均值	0.902	0.456	0.927	0.450	0.933	0.491	<b>0.990</b>	<b>0.449</b>

一些交通标志检测识别图如图 5 和图 6 所示,图 5 中只有单个交通标志,图 6 中有 3 个交通标志。可以看出,针对图 5 所示的交通标志检测,本文方法的检测识别准确率为 0.9793,可以看出,包围框完全框住了禁止停车的标志。R-CNN 和原始 YOLO-V4 的包围框包含了较大的多余区域,或者没有完全包含交通标记。图 6 中,R-CNN 和原 YOLO-V4 仅能检测识别出单个标志,MRESE 可检测出两个标志。而本文方法可以检测出所有标志,这说明 SPP 可以对多尺度目标特征进行更加全面的学习,从而增强骨干网络的性能。从另一个方面可以看出,NMS 抑制冗余包围框的效果也较为明显。



(a)R-CNN



(b)YOLO-V4



(c)MRESE



(d)本文方法

图 6 多个交通标志的检测识别

Fig. 6 Detection and recognition of multiple traffic signs

**结束语** 本文提出了利用 SPP 改进 YOLO-V4 网络架构,利用 SPP 在相同卷积层的不同尺度上选择局部区域,以学习多尺度系统的特征。同时,在 YOLO-V4 网络中针对交通标志识别,将 3 个尺度分辨率改为 2 个较大尺度分辨率,提高了 YOLO-V4 的分辨率并增大了感受野。实验结果表明,所提方法能够提升交通标志的检测识别性能,且在不同尺度上性能也较优。

未来的工作将对不同环境下的交通禁止标志检测进行研究,包括遮挡、多视点变化以及雨雪不同天气状况。此外,可通过生成对抗网络对数据集进行扩展,创建合成图像,以得到更好的结果。

参考文献

[1] WANG J M,WEI Y,ZHOU Y,et al. Speed limit sign recogni-



(a)R-CNN



(b)YOLO-V4



(c)MRESE



(d)本文方法

图 5 单个交通标志的检测识别

Fig. 5 Detection and recognition of single traffic sign

- tion based on LeNet-5 CNN and color feature[J]. *Computer Science*, 2021, 48(S2): 345-350.
- [2] ÁLVARO A G, ÁLVAREZ-GARCÍA JUAN A, SORIA-MORILLO L M. Evaluation of Deep Neural Networks for traffic sign detection systems[J]. *Neurocomputing*, 2018, 316(Nov. ): 332-344.
- [3] TAO X, GONG Y, SHI W, et al. Object Detection with Class Aware Region Proposal Network and Focused Attention Objective[J]. *Pattern Recognition Letters*, 2018, 130(5): 353-361.
- [4] SHI J, ZHOU Y L, ZHANG Q Z. Service robot item recognition system based on improved mask R-CNN and Kinect[J]. *Chinese Journal of Science Instrument*, 2019, 40(4): 216-228.
- [5] DI L, HE R B, LIANG J Z. Road traffic sign recognition algorithm based on possibility clustering and convolution neural network[J]. *Journal of Nanjing University(Science Edition)*, 2019, 55(2): 238-250.
- [6] DONG N, LIU X Y, WU A G. Detection method for circular traffic signs based on artificial bee colony[J]. *Journal of Tianjin University(Natural Science and Engineering Technology Edition)*, 2019, 52(6): 585-593.
- [7] ZHANG J, XIE Z, SUN J, et al. A Cascaded R-CNN With Multiscale Attention and Imbalanced Samples for Traffic Sign Detection[J]. *IEEE Access*, 2020, 51(8): 29742-29754.
- [8] YAZDAN R, VARSHOSAZ M. Improving traffic sign recognition results in urban areas by overcoming the impact of scale and rotation[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 171(3): 18-35.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]// *European Conference on Computer Vision*. 2016: 21-37.
- [10] CHEN P D, HUANG L, XIA Y, et al. Detection and recognition of road traffic signs in UAV images based mask R-CNN[J]. *Remote Sensing for Land & Resources*, 2020, 32(4): 64-70.
- [11] PAN W G, LIU B, CHEN Y H, et al. Traffic sign detection and recognition based on YOLO v3[J]. *Transducer and Microsystem Technologies*, 2019, 38(11): 147-150.
- [12] CHEN Y L, LU Y R, BAISEN W, et al. Applying deep learning technology to image recognition of traffic signs[C]// *International Conference on Technologies and Applications of Artificial Intelligence*. Berlin, Germany; 2019, IEEE: 1-4.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// *Proceedings of the Advances in Neural Information Processing Systems*. Lake Tahoe, USA, 2012: 1097-1105.
- [14] JOSEPH R, ALI F. Yolov3: An incremental improvement[J]. *arXiv*: 1804. 02767, 2018.
- [15] WU F, JIN G, GAO M, et al. Helmet detection based on improved YOLO v3 deep model[C]// *IEEE 16th International Conference on Networking, Sensing and Control*. Ban, Canada: IEEE, 2019: 363-368.
- [16] LI W G, YE X, ZHAO Y T, et al. Strip steel surface defect detection based on improved YOLOv3 algorithm[J]. *Acta Electronica Sinica*, 2020, 48(7): 1284-1292.
- [17] DEWI C, CHEN R C, YU H. Weight analysis for various prohibitory sign detection and recognition using deep learning[J]. *Multimedia Tools and Applications*, 2020, 79(43/44): 1-19.



**PAN Hui-ping**, born in 1982, Ph.D, associate professor. Her main research interests include image processing, artificial intelligence and digital media technology.

(责任编辑: 喻黎)