

# 基于改进卷积神经网络的实时交通标志检测方法

童英\*, 杨会成

安徽工程大学电气工程学院, 安徽 芜湖 241000

**摘要** 提出了一种基于改进卷积神经网络的交通标志检测方法。预训练模型产生否定;使用难分类负样本采集将负样本输入到网络中,提高模型的判别能力;使用多尺度训练过程中的特征级联策略来进一步提升模型的性能。利用 TensorFlow 框架在德国交通标志检测数据集上对所提方法的有效性进行了仿真。研究结果表明,与现有技术相比,所提方法能够获得更快的检测速率,处理每幅图像仅需 0.016 s。

**关键词** 图像处理;卷积神经网络;交通标志检测;特征拼接;难分类负样本采集;多尺度训练

中图分类号 TP181

文献标识码 A

doi: 10.3788/LOP56.071003

## Real-Time Traffic Sign Detection Method Based on Improved Convolution Neural Network

Tong Ying\*, Yang Huicheng

College of Electrical Engineering, Anhui Polytechnic University, Wuhu, Anhui 241000, China

**Abstract** A detection method of traffic signs is proposed based on a modified convolutional neural network. The model is pre-trained to produce the negatives, and hard negative mining is used to add the negative samples into the network to improve the discriminating ability of the model. A feature concatenation strategy during the multi-scale training process is employed to further enhance the performance of the model. On the German traffic sign detection dataset, the effectiveness of the proposed method is simulated in the TensorFlow framework. The research results show that compared with the existing methods, the proposed method can be used to obtain a high detection rate and processing time of only 0.016 s for each image.

**Key words** image processing; convolution neural networks; traffic sign detection; feature concatenation; hard negative mining; multi-scale training

**OCIS codes** 100.4996; 100.5010; 100.2000; 100.3008

## 1 引言

道路交通事故已成为全球最常见的死因之一<sup>[1]</sup>。而高级驾驶员辅助系统可以帮助驾驶员了解交通标志位置以及沿道路行驶的规则,并通知驾驶员标志的含义,降低事故的发生概率。交通场景分析是计算机视觉和智能系统中非常重要的课题<sup>[2-5]</sup>。因此,设计一个自动实时驾驶辅助系统来检测和识别交通标志具有重要的意义。

目前的交通标志检测方法分为 3 类:基于颜色、基于形状和基于滑动窗口。基于颜色的方法通常使

用颜色信息对像素进行分类以提取候选符号<sup>[6-7]</sup>,将候选交通标志提取为最大稳定极值区域<sup>[8]</sup>。文献[9]提出了基于 Ohta 空间的颜色概率模型,以计算交通标志颜色概率图。基于形状特征的方法通常会为每种类型设计一组模板以匹配形状<sup>[10]</sup>。文献[11]采用了定向梯度直方图特征和支持向量机分类器来检测交通标志并取得了良好的效果。文献[12]同时利用颜色与形状特征,在图像的红色边缘位图中采用圆 Hough 变换检测圆形。其他类似 Hough 的方法包括顶点和平分线变换<sup>[13]</sup>、双边中式变换<sup>[14]</sup>、直立三角形单目标投票<sup>[15]</sup>,以及直立椭圆单

收稿日期: 2018-09-25; 修回日期: 2018-10-17; 录用日期: 2018-10-22

基金项目: 安徽省高校自然科学研究重点项目(KJ2018A0122)

\* E-mail: 864844537@qq.com

目标投票<sup>[15]</sup>。还有一些研究提出用智能交通标志分类系统来对基于表意文字的交通标志进行实时分类<sup>[16-18]</sup>。基于滑动窗口的方法已被广泛应用于目标检测,主要使用级联 AdaBoost 分类器,其中弱学习者经常使用 Haar-like 特征<sup>[19]</sup>。

近年来,深度卷积神经网络在物体检测中取得了巨大的成功。空间金字塔池化网络(SPPnet)通过共享卷积特征映射来提高速率<sup>[20]</sup>。由于 SPPnet 方法仅计算整幅输入图像的卷积特征映射,并且培训过程为多阶段,而不是端到端的方法,导致 SPPnet 具有明显的缺点。更快卷积神经网络(Faster R-CNN)能够在共享的端到端检测器上进行培训卷积特征,并显示出更优的准确性和速率<sup>[21]</sup>。快速神经网络仍然使用选择性搜索方法生成 2000 个区域提案<sup>[22]</sup>。Faster R-CNN 引入了一种训练方案,对区域提议的微调和对对象检测的微调之间交替,并产生两个任务之间共享的具有卷积特征的统一网络<sup>[23]</sup>。Zeiler 等<sup>[24]</sup>提出的 ZF(Zeiler Fergus)网络通过使用可视化技术展示了神经网络各层的结构和用处,并获得了 2013 年大规模视觉识别挑战赛第一名的佳绩。YOLO(You Only Look Once)网络将物体检测作为回归问题求解<sup>[25]</sup>。

但上述方法在检测速率上无法达到实时性的要求,针对此问题,本文基于 Faster R-CNN 方法提出了一种新快速卷积神经网络(NF R-CNN)模型,以实现实时处理交通标志检测的目的。

## 2 基本原理

### 2.1 交通标志识别方案

所提方法遵循类似深度学习方法 Faster R-CNN,该方法已被证明是最先进的对象检测通用学习方案<sup>[23]</sup>。它由两个步骤组成:1) 一个区域提案网络(RPN)生成可能包含对象的区域提议列表,或称为感兴趣区域(RoI);2) 利用 Fast R-CNN 网络将图像区域分为对象(和背景)与对象区域边界。这两部分共用特征提取中的卷积层参数,使得此架构以极具竞争力的速度完成对象检测任务。本文建议扩大用于交通标志检测的 Faster R-CNN 体系结构,以提高召回率和精确度,并通过如图 1 所示的流程训练交通标志检测模型。

首先,基于 Faster R-CNN 模型,使用德国交通标志检测基准(GTSDB)测试预先训练好的模型,以便产生强烈的否定。然后,将难分类负样本(HNS)作为培训的第二步,把负面信息输入到网络中。最

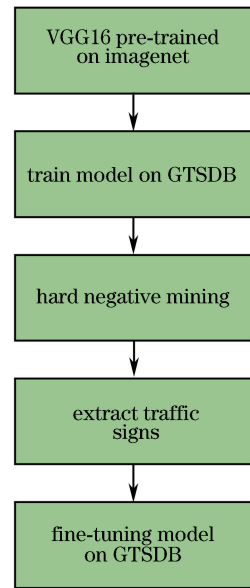


图 1 程序流程图

Fig. 1 Flow chart of program

后,在微调过程中,本文应用多尺度训练过程,并且采用特征级联策略来进一步提升模型的性能。对于整个培训过程,加入类似端到端培训策略。

### 2.2 NF R-CNN 网络结构

传统的 Fast R-CNN 网络是在最后的特征映射层上执行 RoI 池以生成该区域的特征,然后由网络的分类部分进一步分析。这种巧妙的设计使得分类网络可以利用 RPN 计算的特征,节省了计算量。然而,更深卷积层输出中的特征具有更宽的感受域,导致一些重要特征被忽略,所以这种方法存在缺陷,并不是最优的<sup>[26]</sup>。为了捕获更精细的 RoI 细节,建议通过组合多个卷积层的特征映射来改进 RoI 池,包括低级和高级特征。连接多个卷积特征映射的池化结果,以生成检测任务的最终池化特征。首先,将这些结果特征连接起来并重新缩放以匹配特征的原始比例;然后,应用  $1 \times 1$  卷积以匹配原始网络的信道数,与 RPN 中的最终功能图组合在一起产生最终的池化功能;接着,将这些结果特征串联;最后,把  $1 \times 1$  卷积用于匹配原始网络的信道数量。详细架构如图 2 所示,其中 conv 表示卷积层,fc 表示全连接层,concat 为函数。

### 2.3 改进模型的方法

介绍了应用于模型的 4 种方法,将这些方法进行融合,可以提高模型的各项性能。

#### 1) 难分类负样本采集

HNS 采集已被证明是一种提高深度学习表现的有效策略。将 HNS 再次加入网络,可以减少误

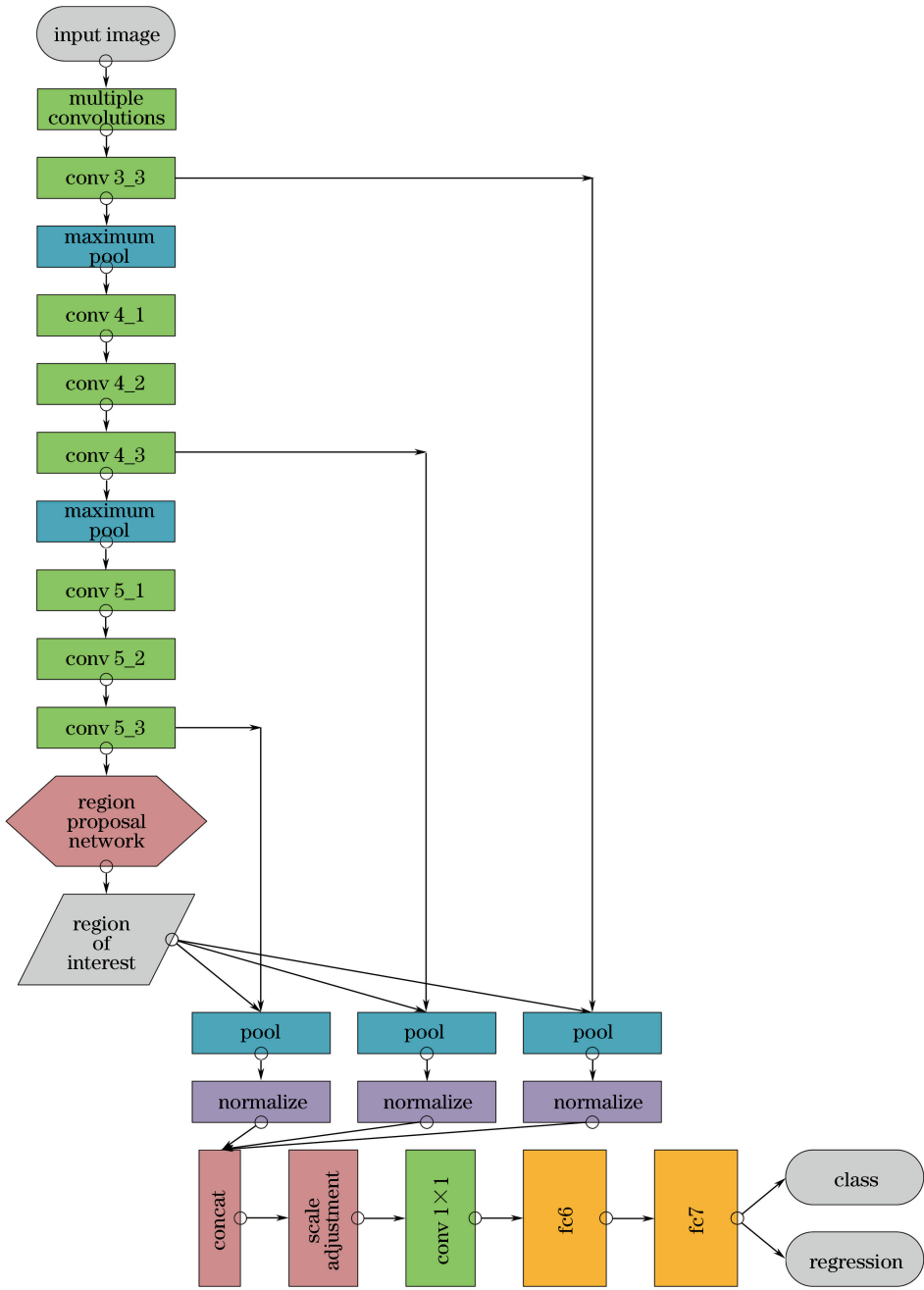


图 2 NF R-CNN 神经网络结构

Fig. 2 Structural diagram of NF R-CNN neural network

报,加强培训效果。本文明确添加一些难以消除的内容进入 RoI 来微调模型,并平衡比率前景和背景约为 1:3。算法如下。

设  $N_t$  为负样本集,  $P_t$  为正样本集,其中下标  $t$  表示在训练过程中可以在多个时间步骤收集样本。设  $D_0$  为原始训练集,由原始的一组正例  $P_0$  和一组负例  $N_0$  组成:

$$D_0 = P_0 \cup N_0. \quad (1)$$

一旦训练过程完成,将网络运行到图像集  $N$ ,从中

收集假正  $F_1$  的子集,将其添加到原始的负集  $N_0$  中:

$$N_1 = N_0 \cup F_1. \quad (2)$$

在每轮训练期间,根据其分数对误报进行排序,并选择预定数量的样本。为了在每轮训练后保持相同的正负样本比例,本文增加了正样本的数量。包含交通标志的一组新样本集  $T_1$  被添加到原始的正例  $P_0$  中:

$$P_1 = P_0 \cup T_1. \quad (3)$$

在完成  $D_t$  训练后,重复上述训练和增加训练

样本的过程:

$$N_{t+1} = N_t \cup F_{t+1}, \quad (4)$$

$$P_{t+1} = P_t \cup T_{t+1}, \quad (5)$$

$$D_{t+1} = P_{t+1} \cup N_{t+1}. \quad (6)$$

因此,集合  $N_{t+1}$ 、 $P_{t+1}$  包含比集合  $N_t$ 、 $P_t$  更多数量的负样本和正样本。

## 2) 锚的调整

研究发现,本文算法最关键的是数量 RPN 中的锚点数量。传统的 Faster R-CNN 使用 9 个锚无法召回小物件。然而,对于交通标志检测任务,目标的尺寸往往较小,特别是在不受约束情况下的交通标志,其检测更加困难。因此,本文添加一个  $64 \times 64$  的尺寸组,增加了锚的数量。仿真表明,该调整显著地增加了模型检测的精确度和召回率。

## 3) 模型预训练

为了使 Faster R-CNN 适应交通标志检测,选择在交通标志数据集上微调预先训练好的模型。作为一个被广泛确认的数据集(包含无约束的交通标志),德国交通标志识别基准(GTSRB)可以用于验证模型,首先对本文模型进行预训练交通标志数据集以及 HNS。另外,HNS 采集是预训练前必不可少的步骤,可以使误报最小化。这种预训练的策略,加上 HNS 的采集,再次提升了模型的准确率。

## 4) 多尺度训练

Faster R-CNN 结构通常采用固定比例训练所有的图像。通过将图像大小调整为随机比例,从而提高尺度不变性方面的性能。本文为每幅图像在输入网络之前随机分配 3 个比例之一。研究表明,多尺度的训练使得本文模型适用于不同大小的图片,提高了基准检测的性能。

## 3 仿真设置及结果分析

采用基准数据集进行实验,以展示方法的有效性。GTSDB 数据集包含 900 幅高分辨率的交通标志自然场景图像,每幅图像尺寸为  $1360 \text{ pixel} \times 800 \text{ pixel}$ ,交通标志尺寸从  $16 \text{ pixel} \times 16 \text{ pixel}$  到  $128 \text{ pixel} \times 128 \text{ pixel}$ 。使用 600 幅图像进行训练、300 幅图像进行测试。GTSRB 数据集包含 39209 幅训练图像和 12630 幅测试图像。数据集中的德国交通标志共有 43 个小类,尺寸从  $15 \text{ pixel} \times 15 \text{ pixel}$  到  $250 \text{ pixel} \times 250 \text{ pixel}$ ,如图 3 所示。

采用 TensorFlow 框架来训练本文模型。VGG16 卷积神经网络结构由视觉几何组开发,并在 2014 年度大规模视觉识别挑战赛中赢得了第一。如今,



图 3 交通标志的 3 个示例。(a)指示;(b)禁止;(c)警告

Fig. 3 Three examples of traffic signs.

(a) Mandatory; (b) prohibitory; (c) danger

VGG16 仍然被认为是优秀的视觉模型。该模型具有 16 层的深度,足以进行交通标志图片检测,并且该网络在训练和测试时使用了多尺度进行数据增强。因此将 VGG16 选定为本文的预训练模型,并将改进后的模型与 VGG16 的原始模型进行仿真结果对比。

1) 在 GTSDB 数据集上预训练 VGG16 模型,进行 80000 次迭代,学习率设置为 0.0001。在此培训过程中,图像首先被重新缩放,且始终保持原始宽高比。短边重新缩放为 600,长边为 1000。并使用水平翻转策略,以达到数据增强的目的。在训练过程中,12 个锚点用于 RPN 部分,覆盖总尺寸为  $64 \times 64$ 、 $128 \times 128$ 、 $256 \times 256$ 、 $512 \times 512$  和 3 种纵横比,包括 1:1、1:2 和 2:1。经过非极大值抑制(NMS)算法之后,保留了 1800 个区域提议。对于 Faster R-CNN 分类部分,如果 IoU(intersection over union) 阈值大于 0.5,则将 RoI 视为前景,否则视为背景。为了平衡前景和背景的数量,对这些 RoI 进行采样以保持前景和背景之间的比例为 1:3。

2) 将上述数据集输入网络。当输出区域的信心评分高于 0.8,而 IoU 值与任何地面实况注释均小于 0.5 时,则视为 HNS。然后采用固定学习率为 0.0001 的回归程序进行 10000 次迭代,确保这些 HNS 与其他样本的 RoI 一起被选择。最后,通过 GTSDB 数据集进一步对结果模型进行微调,以产生最终检测模型。

为了在 GTSDB 基准测试中检测本文提出的交通标志检测模型的性能,对于每幅图像,除了执行水平翻转外,还将其较短的一面反馈到网络之前随机调整大小。与步骤 1) 中采取的策略类似,本文确保较长的一面像素不超过 1250。在训练过程中,连接了从 convolution layer 3\_3、4\_3 和 5\_3 图层汇集的功能,在连接后使用的标度可以为细化的或固定的。在训练和测试阶段对二进制对象使用 1:4700 的固定比例。使用 0.001 的固定学习率对模型进行 5000 次迭代微调,以获得最终模型。由于道路环境复杂,且图片分辨率低,ZF

模型和 VGG16 模型均会存在漏检和误检的情况,检测结果示例如图 4 所示。

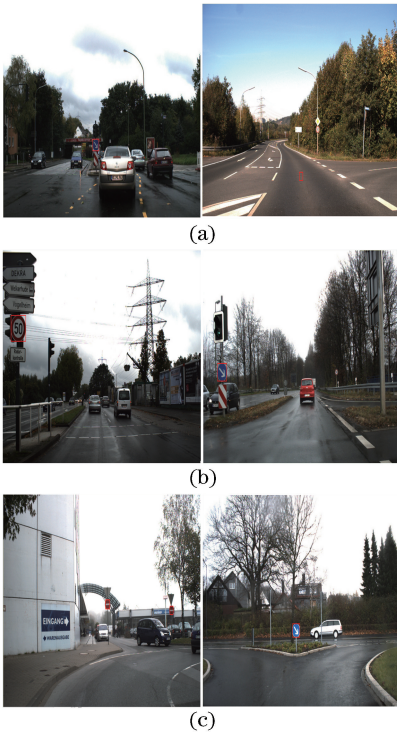


图 4 3 种模型检测结果示例。

(a) ZF; (b) VGG16; (c) NF R-CNN

Fig. 4 Examples of test results by three models.

(a) ZF; (b) VGG16; (c) NF R-CNN

3)对于测试阶段,按照与第 1 阶段相同的原则对图像进行重新缩放。对于每幅图像,RPN 网络在区域提案生成步骤中共生成 80 个区域提案。如果分类置信度得分大于 0.8,则选定的区域提案将被视为交通标志。对于本文方法,将 NMS 阈值设置为 0.3。为了进行分析,还在实验中输出了置信度得

分大于 0.001 的所有区域提案。

以精确度、召回率,以及受试者工作特征曲线 (ROC)下的面积 (AUC)作为评价标准。精确度表示预测为正类样本中真正为正类样本的比例。召回率表示预测为正类样本中被正确分类的比例。表 1 为本文方法在不同时间阈值下的精确度和召回率。随着阈值的增加,精确度升高,召回率降低。当阈值  $\Theta=0.4$  或  $\Theta=0.5$  时,模型的整体性能更高效且更稳健。

为了分析测试结果,设置  $\Theta=0.5$  以确保高精度的检测,当交通标志的置信度得分低于阈值  $\Theta$  时,将过滤交通标志的候选区域。表 2 所示为 3 种模型在数据集上的分类结果。图 5 所示为 3 个超类别精确度-召回率 (P-R) 曲线。

由表 2 可知,本文模型虽然在精确度上略低于 VGG16 模型,但明显高于 ZF 模型,并且可以看出,本文模型相较于其他模型耗时更少,每幅图片仅需 0.016 s。由图 5 可知:1) 相比于其他模型,ZF 模型的卷积深度较浅,导致 ZF 模型在指示标志和禁止标志中的 AUC 值均低于 VGG16 模型和本文模型;2) VGG16模型虽简化了卷积神经网络的结构,但需要较大的训练特征数量,并且禁止标志分布密集、尺寸太小,导致特征不明显,所以 VGG16 模型在警告标志中的 AUC 值明显偏低,仅为 91.58%;3) 多种优化方法的使用均能提升模型的精确度和召回率,故而本文模型在 3 种示例标志中均表现出较高的 AUC 值和最优的稳定性,达到指示标志 94.47%,禁止标志 98.56%和警告标志 95.21%。实验结果表明,HNS 的采集方法和锚的调整可以提高检测率,减少时间消耗。

表 1 不同时间阈值下的精确度和召回率

Table 1 Accuracy and recall rates at different time thresholds

Item	Threshold $\Theta$					
	0.1	0.2	0.4	0.5	0.6	0.7
Precision	0.7654	0.8826	0.9542	0.9688	0.9835	1
Recall	0.9542	0.9403	0.9125	0.8623	0.7854	0.6281

表 2 3 种模型分类结果比较

Table 2 Comparison of classification results by three models

Model		Mandatory / %	Prohibitory / %	Danger / %	Time / s
Model-ZF	P	91.89	97.68	98.65	0.133
	R	79.11	70.56	93.15	
Model-VGG16	P	98.85	99.44	99.86	0.130
	R	75.23	69.26	87.84	
Model-NF R-CNN	P	98.10	99.22	99.23	0.016
	R	80.31	72.96	94.78	

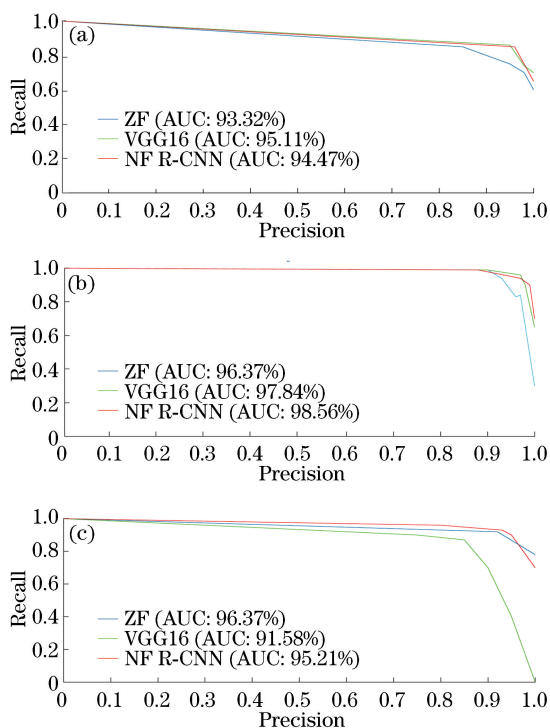


图5 3个示例的P-R曲线图。(a)指示;(b)禁止;(c)警告

Fig. 5 P-R curves of three examples.

(a) Mandatory; (b) prohibitory; (c) danger

3个示例交通标志的AUC值和运行时间如表3所示。文献[8-11]方法是基于传统的检测算法,研究显示,传统方法的检测速率普遍偏低,无法达到实时效果,而本文方法不仅在AUC数值上显示出优越性,而且在处理时间上更是达到了0.016 s,实现了实时性的要求。

表3 AUC值和处理时间

Table 3 AUC values and processing time

Reference	AUC value / %			Time/s
	Mandatory	Prohibitory	Danger	
Ref. [8]	96.54	99.35	96.68	0.15
Ref. [9]	92.56	96.58	91.27	0.31
Ref. [10]	93.47	100	98.64	0.66
Ref. [11]	100	100	99.22	3.47
NF R-CNN	97.02	99.55	99.37	0.016

## 4 结 论

基于交通标志检测的重要性以及现有方法检测速率低的问题,提出了一种基于深度学习的交通标志检测方法。训练了用于通用对象检测的最先进的Faster R-CNN框架,提出了几种有效的策略来改进用于解决交通标志检测任务的Faster R-CNN算法,并在测试集上进行了实验训练与效果比较。实验结果表明,与现有方法相比,所提方法的检测时间

更短,更趋于实时。在未来的工作中,还将有望进一步提高检测速率,并解决所提方法的效率和可扩展性问题。

## 参 考 文 献

- [1] Stewart B T, Yankson I K, Afukaar F, *et al.* Road traffic and other unintentional injuries among travelers to developing countries[J]. *Medical Clinics of North America*, 2016, 100(2): 331-343.
- [2] Zhang S F, Zhu T. A method of traffic sign detection and recognition based on HDR technology[J]. *Laser & Optoelectronics Progress*, 2018, 55(9): 091006. 张淑芳, 朱彤. 一种基于HDR技术的交通标志牌检测和识别方法[J]. *激光与光电子学进展*, 2018, 55(9): 091006.
- [3] Hu C P, Bai X, Qi L, *et al.* Learning discriminative pattern for real-time car brand recognition[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(6): 3170-3181.
- [4] Gao D D, Xu X T, Li B. Research on application of infrared and white light mixed supplemental lighting system in intelligent transportation[J]. *Infrared and Laser Engineering*, 2018, 47(9): 0918006. 高东东, 徐晓婷, 李博. 红外/白光混合补光系统在智能交通中的应用研究[J]. *红外与激光工程*, 2018, 47(9): 0918006.
- [5] Xu Y, Wei Z Y. An improved traffic sign image recognition algorithm[J]. *Laser & Optoelectronics Progress*, 2017, 54(2): 021001. 徐岩, 韦镇余. 一种改进的交通标志图像识别算法[J]. *激光与光电子学进展*, 2017, 54(2): 021001.
- [6] Fleyeh H, Davami E. Eigen-based traffic sign recognition[J]. *IET Intelligent Transport Systems*, 2011, 5(3): 190-196.
- [7] Khan J F, Bhuiyan S M A, Adhami R R. Image segmentation and shape analysis for road-sign detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12(1): 83-96.
- [8] Yang Y, Luo H L, Xu H R, *et al.* Towards real-time traffic sign detection and classification [C] // *Proceedings of 17<sup>th</sup> International IEEE Conference on Intelligent Transportation Systems*, 2014: 87-92.
- [9] Yang Y, Wu F C. Real-time traffic sign detection via color probability model and integral channel features [C] // *Proceedings of Chinese Conference on Pattern Recognition*, 2014: 545-554.
- [10] Liang M, Yuan M Y, Hu X L, *et al.* Traffic sign detection by ROI extraction and histogram features-

- based recognition [C] // Proceedings of the 2013 International Joint Conference on Neural Networks, 2013: 1-8.
- [11] Wang G Y, Ren G H, Wu Z L, *et al.* A hierarchical method for traffic sign classification with support vector machines [C] // Proceedings of the 2013 International Joint Conference on Neural Networks, 2013: 1-6.
- [12] Wang G Y, Jin Y S, Ren G H, *et al.* High-performance VLSI architecture for traffic sign detection[J]. Infrared and Laser Engineering, 2018, 47(9): 0926001.  
王刚毅, 金炎胜, 任广辉, 等. 高性能交通标志检测模块的 VLSI 结构设计[J]. 红外与激光工程, 2018, 47(9): 0926001.
- [13] Belaroussi R, Tarel J P. Angle vertex and bisector geometric model for triangular road sign detection[C]// Proceedings of 2009 Workshop on Applications of Computer Vision, 2009: 1-7.
- [14] Belaroussi R, Tarel J P. A real-time road sign detection using bilateral Chinese transform [C] // Proceedings of International Symposium on Visual Computing, 2009: 1161-1170.
- [15] Houben S. A single targetvoting scheme for traffic sign detection [C] // Proceedings of 2011 IEEE Intelligent Vehicles Symposium, 2011: 124-129.
- [16] Mammeri A, Khiari E H, Boukerche A. Road-sign text recognition architecture for intelligent transportation systems [C] // Proceedings of 2014 IEEE 80<sup>th</sup> Vehicular Technology Conference, 2014: 1-5.
- [17] Yin S Y, Ouyang P, Liu L B, *et al.* Fast traffic sign recognition with a rotation invariant binary pattern based feature[J]. Sensors, 2015, 15(1): 2161-2180.
- [18] Mathias M, Timofte R, Benenson R, *et al.* Traffic sign recognition-how far are we from the solution[C]// Proceedings of the 2013 International Joint Conference on Neural Networks, 2013: 1-8.
- [19] Ruta A, Porikli F, Watanabe S, *et al.* In-vehicle camera traffic sign detection and recognition [J]. Machine Vision and Applications, 2011, 22(2): 359-375.
- [20] He K M, Zhang X Y, Ren S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [21] Girshick R. Fast R-CNN[C] // Proceedings of 2015 IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [22] Uijlings J R R, van de Sande K E A, Gevers T, *et al.* Selective search for object recognition [J]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [23] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [24] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]// Proceedings of European Conference on Computer Vision, 2014: 818-833.
- [25] Redmon J, Divvala S, Girshick R, *et al.* You only look once: unified, real-time object detection[C] // Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [26] Sun X D, Wu P C, Hoi S C H. Face detection using deep learning: an improved faster RCNN approach [J]. Neurocomputing, 2018, 299: 42-50.