人工智能技术与应用

基于嵌入式 GPU 平台的列车运行环境检测算法

熊敏君，李　晨，张慧源，彭联贴，苏　震
（株洲中车时代电气股份有限公司，湖南 株洲　412001）

摘　要：列车运行环境的实时检测是实现列车自动驾驶的重要前提。针对传统列车运行环境检测算法存在的效率低、精度差、鲁棒性弱等问题，文章提出了一种基于图像实例分割的列车运行环境实时检测算法。其通过车载摄像头获取列车运行环境的图像数据，进行样本标注、图像增强等预处理工作；采用改进的实例分割网络 MaskRCNN 进行模型训练，并基于 TensorRT 实现模型优化加速。最后，在嵌入式 GPU 平台 NVIDIA-Xavier 上进行模型的性能验证，优化后模型在验证集下的目标检测精度达到 94.75%，模型推理速度约为原来的 6 倍，满足列车自动驾驶运行环境的实时检测需求。

关键词：实例分割；列车运行环境；空洞卷积；模型加速；嵌入式平台；列车自动驾驶

中图分类号：TP399　　**文献标识码：**A　　**文章编号：**2096–5427(2021)04–0072–06

doi:10.13889/j.issn.2096–5427.2021.04.012

An Algorithm of Train Operation Environment Recognition Based on Embedded GPU Platform

XIONG Minjun, LI Chen, ZHANG Huiyuan, PENG Liantie, SU Zhen
(Zhuzhou CRRC Times Electric Co., Ltd., Zhuzhou, Hunan 412001, China)

Abstract: Real-time detection of train operation environment is an important prerequisite for realizing automatic train operation. Aiming at the problems of low efficiency, poor accuracy and weak robustness of traditional train operation environment sensing algorithms, a real-time detection algorithm of train operation environment based on image instance segmentation is proposed. The image data of train operation environment is acquired by onboard cameras, and the corresponding preprocessing work is carried out, such as sample labeling and image enhancement. The instance segmentation network MaskRCNN based on deep learning is improved and trained, and the model is accelerated based on TensorRT. Finally, the model is validated on the embedded development platform NVIDIA-Xavier. Detection accuracy of the optimized model in the verification set is 94.75%, and the inference speed of the accelerated model is about 6-time of the original one, which meets the real-time detection requirements of train operation environment.

Keywords: instance segmentation; train operation environment; dilated convolution; model acceleration; embedded platform; automatic train operation

0　引言

随着汽车自动驾驶的快速发展，轨道交通领域也在不断开展列车辅助及自动驾驶的研究。列车，尤其是货运列车，有着载重大、运输线路长等特点，辅助及自动驾驶对保障其驾驶安全、提高驾驶效率具有重要的作用。列车运行环境的实时检测是实现列车辅助及自动驾驶的基础。

列车运行环境检测涉及列车运行过程中轨道区域检测、障碍物识别、交通信号灯识别及轨旁标识牌识别等。其中，列车前方轨道区域检测是环境检测的基础，传统方法一般是通过检测两条钢轨线位置从而

收稿日期：2021–03–18
作者简介：熊敏君（1993—），女，硕士，主要从事图像识别技术研究。
基金项目：国家重点研发计划（2016YFB1200401）

确定轨道区域的。文献 [1] 通过改进边缘检测算子进行钢轨识别，但其识别的前提需要两条钢轨和背景存在比较明显的灰度差异，且对光照、遮挡等环境因素敏感。文献 [2] 针对光照因素引起的钢轨特征提取不完整的问题，提出了基于相位一致性的钢轨特征提取方法，但该方法在复杂场景下的轨道检测效果较差。文献 [3] 基于曲率映射图实现了近距离轨道的识别，并基于局部梯度信息实现远距离轨道识别，能够识别百米以内的轨道。此外，有研究者提出采用模板匹配的方法 [4-5] 提取轨道线，和通过高次曲线拟合的方法 [6] 实现对弯轨的识别。在轨道环境较复杂或弯道曲率较大时，基于传统算法的轨道识别漏检率及误检率较大，并受光照、部分遮挡等外界因素的影响较大，导致算法的鲁棒性较差。

基于深度学习的语义 / 实例分割算法近些年发展较快，在公路交通环境检测方面取得了较多应用成果。文献 [7] 提出了全卷积网络（fully convolutional networks, FCN），对传统的卷积神经网络进行了改进，实现了端到端的图像语义分割算法。文献 [8] 采用类似 FCN 结构实现车道线的检测，算法分割效果较好，但实时性较差。为了提高算法的执行效率，文献 [9] 提出了一种编码 – 解码结构 SegNet 网络，编码部分采用传统的卷积网络的池化层，解码部分进行上采样恢复原图像的特征信息。研究者们为了进一步提高算法的精度与效率，又相继提出了 ENet[10]，ERFNet[11]，MaskRCNN[12] 等分割网络结构。其中，MaskRCNN 实例分割网络检测精度较高，主要用来进行目标检测和物体轮廓分割，适用于车辆自动驾驶的交通环境感知。

列车运行的轨道环境较公路交通环境来说相对简单，相应的基于视觉的环境感知算法更易落地实现。本文基于图像实例分割 MaskRCNN 网络，提出一种列车运行环境实时检测算法，通过样本数据图像增强、网络参数与结构的改进、深度学习加速引擎 TensorRT 的模型推理加速等处理，提高算法的检测精度及运行效率；并在嵌入式开发平台 NVIDIA-Xavier 上开展实验，验证了算法的检测精度与运行效率。

1　列车运行环境检测算法原理

基于图像实例分割的列车运行环境检测算法基于嵌入式 GPU 平台 Xavier 实现，由数据处理、模型生成及实时检测 3 个部分组成，本节主要介绍其框架及实例分割模型 MaskRCNN 的基本原理。

1.1　算法框架

图 1 为基于图像实例分割的列车运行环境实时检测算法框架。其中，数据处理环节通过车载摄像头对列车运行环境的视频数据进行采集与存储，并进行筛选、抽帧、标注、图像增强等处理，以获得样本数据；模型生成环节通过改进 MaskRCNN 网络训练、基于 TensorRT 的模型加速获得可推理的模型文件；实时检测环节完成模型文件在 Xavier 上的部署，同时对图像分割结果进行分类，以得到列车运行环境信息。

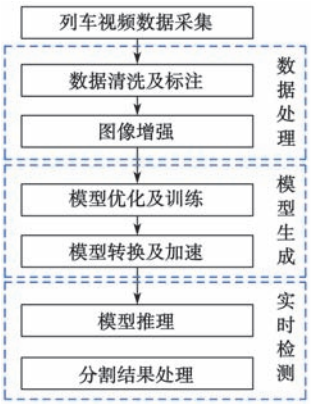

图 1　列车运行环境检测算法框架

Fig. 1　Framework of the train operation environment recognition algorithm

1.2　MaskRCNN 网络模型

MaskRCNN 实例分割网络由主干网络、区域生成网络以及 3 个分支任务组成，如图 2 所示。其中，主干网络由标准卷积神经网络 CNN（ResNet 系列）及其形成的特征金字塔网络 FPN 组成，主要负责特征提取；区域生成网络 RPN 负责从不同尺寸的特征图中获得候选区域并进行特征对齐；最后，通过目标分类、框回归及掩码分支得到目标检测及分割结果。

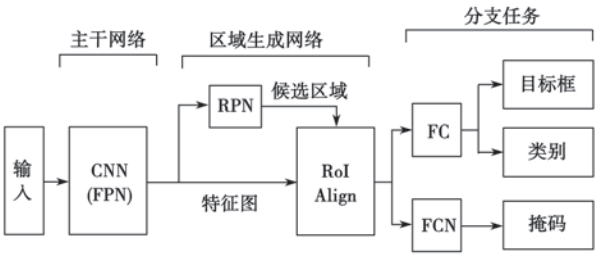

图 2　MaskRCNN 网络示意图

Fig. 2　Structure diagram of MaskRCNN

2　列车运行环境检测算法的实现

针对原始 MaskRCNN 模型在轨道环境目标检

测中存在的目标检测精度低、分割边缘粗糙、算法运行速度慢等问题，本文采用图像增强、主干网络优化、特征金字塔结构优化等措施提高检测精度，并通过基于 TensorRT 的模型加速来满足模型的实时推理需求。

2.1 数据处理与图像增强

视频数据采集装置由多个车载摄像头组成，根据机车车型、运行环境和客户需求的不同，摄像头模组可安装于列车司机室内部或外部，但需保证能够最大视野地获取列车运行前方轨道、轨旁信号灯及标识牌等区域视频数据，且获取的视频数据覆盖列车不同运行时间段、不同运行场景的环境信息。

通过视频数据预处理，获取训练样本数据集，其主要步骤如下：

（1）视频场景分类。对采集的视频数据进行人工清洗，均衡不同场景的视频数量。

（2）关键帧抽取。采用每秒 1 帧的方式对视频抽取关键帧，通过图像质量检测算法筛除重复度较高的图像数据。

（3）图像标注。采用标注工具对图像数据进行像素级的标注。

（4）标注格式转换。将标注完成的标签统一转换成 coco 数据格式。

文献 [13] 验证了图像增强方法在深度学习图像识别中应用的可行性，即采用正常样本 1.43% 数量的样本进行图像增强后可达到与正常样本训练同样的精度。由于轨道环境训练样本数据有限，采用图像处理方法进行图像增强，可进一步扩充样本数据，主要扩充方式有随机裁剪、平移与缩放、高斯噪声、颜色及亮度变换等。图 3 示出图像增强（通道处理）前后图片效果对比。

(a) 原图 (b) 处理后图片

图 3 图像增强处理效果对比

Fig. 3 Comparison of image channel processing effects

2.2 MaskRCNN 模型改进

针对原始 MaskRCNN 实例分割模型存在的大目标分割边界粗糙、小目标检测精度低等问题，本文主要对模型进行以下 3 个方面的改进优化。

（1）采用混合空洞卷积优化主干网络。主干网络提取的特征图信息直接影响后续网络的检测精度。与普通卷积相比，空洞卷积有利于在保证分辨率的前提下扩大感受野，从而获得多尺度的上下文信息，非常适用于图像语义 / 实例分割任务。为了消除空洞卷积带来的网格效应，本文采用不同空洞率组合的卷积核——混合空洞卷积 [14] 对 ResNet 主干网络进行特征提取优化。

（2）特征金字塔的不同尺寸特征融合。特征金字塔结构可以解决不同尺度目标的检测问题，原始 MaskRCNN 模型采用的金字塔结构包含左侧“自底向上”和右侧“自顶向下”两条路径，通过二者相邻层的横向连接实现高层语义信息与低层位置信息的融合（图 4）。为了进一步将低层位置信息融进高层特征，得到更优的目标检测精度，如图 4 所示，在模型的右侧增加一条“自底向上”的路径，以优化传统特征金字塔结构。

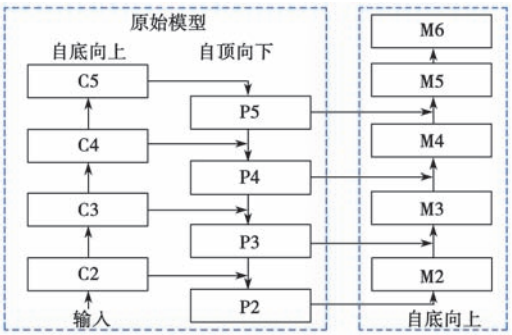

图 4 金字塔结构及其优化示意图

Fig. 4 Pyramid structure and its optimization diagram

（3）提高全卷积掩码输出尺寸。由于原始模型掩码分支默认单个目标的输出大小为 28*28，最终的分割结果图采用上采样得到，导致大目标的分割边缘不清晰，锯齿状边缘现象较严重。为了解决这一问题，考虑提升掩码分支的输出尺寸来优化分割边缘。然而，较大的掩码尺寸输出会造成算法运行效率的下降。因此，根据实验结果对精度与效率进行综合评估，提高全卷积掩码输出尺寸为 56*56，在满足算法运行效率的基础上得到更加平滑的分割轮廓。

2.3 基于 TensorRT 的模型加速

由于车载嵌入式开发平台的算力有限，而实例分割模型资源消耗较大，因此需要对训练好的模型进行加速，从而达到车载实时运行的效果。本文采用高性能深度学习支持引擎 TensorRT 对模型进行加速处理，并在嵌入式开发平台 NVIDIA-Xavier 上实现。

TensorRT 是 NVIDIA 公司针对深度学习网络推出的神经网络推断加速引擎，可极大提高深度学习模型在边缘设备上的推理速度。其工作原理是通过网络模型的层间合并，优化内核选择，指定模型精度（浮点型 32 位 FP32、半精度 FP16 或整型 INT8），执行归一化和转换，优化矩阵计算，从而减少延时、提高吞吐量和效率。

基于 TensorRT 的模型加速算法的主要流程如图 5 所示：

（1）初始化模型推理相关参数，包括待加速的模型文件、engine 文件、图像帧的处理批次、网络输入大小、模型的输入及输出节点等。

（2）判断是否存在本地 engine 文件可读取，若是，则直接进行模型的反序列化；否则，进行模型解析与 engine 生成。

（3）判断模型是否需要进行低精度推理参数设置，并序列化 engine 后保存至本地。

（4）执行模型推理，输入数据为视频流处理模块的输出数据，获得推理结果。

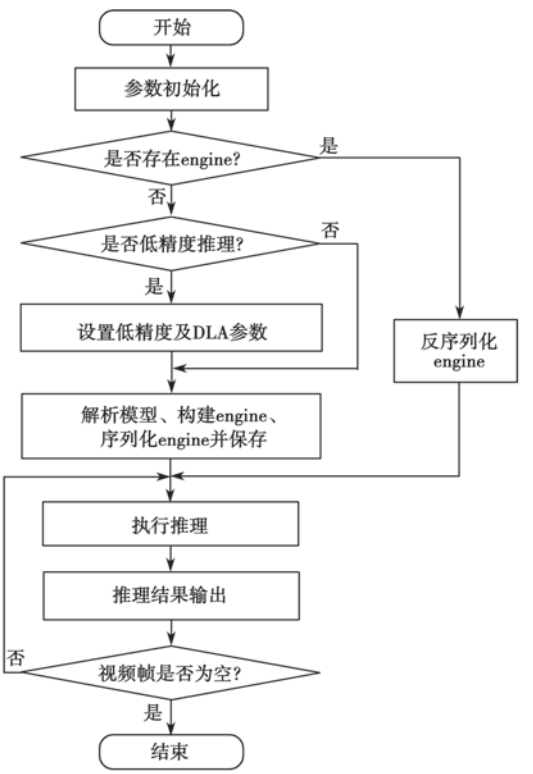

图 5　TensorRT 模型加速算法流程图

Fig. 5　Algorithm flow chart of model acceleration based on TensorRT

3　实验结果分析

为了评估模型优化及加速的效果，本文设计了模型优化实验及推理加速实验，对比了原始模型及优化后模型的检测精度与推理速度。

3.1　环境配置及评价标准

实例分割模型训练的实验环境为两张 NVIDIA-Tesla-V100 显卡（显存为 32 GB），实验基于 TensorFlow 深度学习框架进行。列车运行环境检测的样本数据约 5 000 张，目标类别包含行人、机车、轨道、汽车、信号灯及标识牌。从中随机选取约 800 张作为验证集，样本数据图像大小均为 1920*1080，模型输入大小为 1024*1024。为了提高训练模型的准确率，采用在 ImageNet 上训练好的模型进行迁移学习。文中所有实验训练步数设置均为 epoch 值 180，每个 epoch 步长为 1 000，batch_size 值为 2。模型推理加速实验环境为 NVIDIA-Jetson-Xavier 嵌入式处理平台。

模型优化实验分别计算目标检出的准确率 P 及召回率 R，然后通过式 (1) 计算得到精度值 F 进行检测效果评估。

$$F=\frac{2PR}{P+R}\times 100\% \tag{1}$$

其中轨道项点只进行区域判断，不作为评估依据；其他项点检出值与真实值的交并比 IoU 大于 0.5 即视为正确检出。

推理加速试验采用模型推理时间 T_{infer} 进行加速效果评估。

3.2　结果分析

3.2.1　模型优化实验分析

在进行模型优化实验前，对主干网络 ResNet50 和 ResNet101 进行了对比。虽然采用 ResNet50 训练及推理速度有所提升，但目标检测精度下降明显，因此后续实验均采用 ResNet101 主干网络进行训练及验证。

为了验证模型优化效果，分别对原始模型（MR）、改进空洞卷积（MR_D）模型及改进金字塔结构（MR_{FPN}）模型进行准确率验证，其中采用模型 MR_{FPN} 的检测效果如图 6 所示。验证集共包含 5 类目标物体，约 2 000 个。为了更好地验证小目标检测准确率，将所有目标物体按照像素大小分为大、中、小 3 个类别。不同模型优化后得到的检测结果如表 1 所示，所有结果均取 5 类项点的平均值进行评价，其中下标大、中、小分别代表不同大小的目标的检测结果。

从表 1 检测结果可以看出，采用空洞卷积优化有

助于提高小目标检测的准确率，但对大、中目标的检测精度仅有轻微影响；而改进金字塔结构对不同大小的目标检测精度均有明显提升，平均检测精度提升约 2 个百分点。因此，基于 TensorRT 的模型推理加速实验采用改进金字塔结构的模型进行验证。

表 1　不同模型的检测结果对比
Tab. 1　Comparison of detection results　（%）

指标	MR	MR_D	MR_{FPN}
$P_{大}$	96.10	95.87	97.41
$P_{中}$	95.02	93.74	98.05
$P_{小}$	86.96	93.39	93.24
$R_{大}$	92.51	89.77	98.46
$R_{中}$	93.87	92.13	94.89
$R_{小}$	90.16	92.99	87.36
$F_{大}$	94.27	92.72	97.93
$F_{中}$	94.44	92.93	96.44
$F_{小}$	88.53	93.19	90.20
F_{mean}	92.41	92.94	94.86

场景一　场景二

(a) 原始图片

场景一　场景二

(b) 采用 MR_{FPN} 检测结果

图 6　列车运行环境检测结果图（MR_{FPN}）
Fig. 6　Segmentation results of the train operation environment by MR_{FPN}

为了验证不同掩码尺寸输出对分割区域的影响，采用轨道项点分割结果的平均交并比 mIoU 作为评估指标。M_{28*28} 与 M_{56*56} 分别表示掩码输出尺寸为 28*28 与 56*56。由表 2 结果分析可知，提高掩码输出尺寸后，mIoU 提升 1.05%，有助于分割精度的提升。图 7 示出提高全卷积掩码输出尺寸后目标边缘锯齿化现象改进效果。从轨道边缘可以看出，边缘的锯齿状现象得到了较好的改善，输出边缘平滑，有利于轨道区域限界判定。

表 2　掩码尺寸对轨道区域检测的影响
Tab. 2　Influence of mask size on track area detection

掩码输出尺寸代码	mIoU/%
M_{28*28}	95.08
M_{56*56}	96.13

场景一　场景二

(a) 原始图片

场景一　场景二

(b) 由原始模型得到的分割效果图片

场景一　场景二

(c) 提高掩码尺寸后的输出图片

图 7　分割边缘改进效果对比图
Fig. 7　Comparision of segmentation edge results

3.2.2　推理加速实验分析

模型推理加速实验在 Xavier 平台上实现，主要测试使用 TensorRT 加速前后模型在 Xavier 上的推理速度，该实验采用精度 (FP16) 的模型加速推理进行推理速度验证。模型加速前后的推理时间对比如表 3 所示，T_{before} 和 T_{after} 分别为模型加速前后的推理时间。可以看出，通过 TensorRT 对模型加速，可以显著提升模型的推理速度。模型输入大小为 1024*1024 时，模型的推理速度约为 7 fps，推理加速比超过 6，基本能够实现算法在车载嵌入式平台上的实时运行。

表 3　模型加速前后推理时间对比
Tab. 3　Comparison of inference time before and after model acceleration

模型类别	T_{before}/ms	T_{after}/ms	加速比
MR	871	133	6.55
MR_{FPN}	897	138	6.50

为了确保基于 TensorRT 的模型推理加速不影响检测精度，进行了加速后的精度对比实验。如表 4 所示，F_{before} 和 F_{after} 分别为模型加速前后的平均检测准确率，精度差仅 0.11%。由此可知模型推理加速对于检测精度无明显影响。

表 4　模型加速前后精度对比

Tab. 4　Comparison of model accuracy before and after acceleration　（%）

模型类别	F_{before}	F_{after}	精度差
MR_{FPN}	94.86	94.75	0.11

综合分析实验结果可知，通过改进金字塔结构、提高掩码尺寸及模型推理加速处理，提升了原始模型的检测精度及运行速度，获得了更好的轨道分割效果；同时，在嵌入式 Xavier 平台上模型推理速度达到 7 fps，具备在机车上进行实时环境检测应用的潜力。

4　结语

本文提出一种基于图像实例分割实现的列车运行环境实时检测算法。首先，针对轨道检测边界不清晰及目标检测精度较低的问题，采用空洞卷积、金字塔结构优化等方法改进原有网络模型，提高模型的检测精度及目标边缘分割精度，在验证集下目标检测精度达到 94.75%；其次，通过 TensorRT 进行模型推理加速优化，推理速度达到 7 fps，推理加速比超过 6，并通过嵌入式平台 NVIDIA-Xavier 实现算法的车载实时运行。该系统主要应用于机车和城市轨道交通列车的自动驾驶的环境感知与障碍物检测，辅助机车实现对轨道区域、行人、其他机车、信号灯及标识牌等目标的实时检测。通过模型结构优化及推理加速，提高了系统的检测精度及运行效率，使其满足列车运行环境的实时检测需求。

目前该推理加速算法仅实现了一部分列车运行环境的检测，后续可通过扩充实例分割的目标类别，实现对图像中所有区域的分割；并可通过整型推理进一步提升模型加速的速度。

参考文献：

[1] 王燕芩 , 李沛奇 . 基于改进 Canny 算法的铁轨边缘检测 [J]. 铁道通信信号 , 2015, 51(2):75-78.
WANG Y Q, LI P Q. Rail Edge Detection Based on Improved Canny Algorithm[J]. Railway Signalling & Communication, 2015, 51(2):75-78.

[2] 史红亮 , 荣剑 , 周新民 . 基于相位一致性的铁轨特征提取研究 [J]. 自动化与仪器仪表 , 2015(2):92-94.
SHI H L, RONG J, ZHOU X M. Study on Extraction of the tracks feature based on phase consistency[J]. Automation & Instrumentation, 2015(2):92-94.

[3] 王云泽 . 列车前方轨道识别算法的设计与实现 [D]. 杭州 : 浙江大学 , 2017.
WANG Y Z. Design and Implementation of An Algorithm to Recognize the Rail Ahead of the Train[D]. Hangzhou:Zhejiang University, 2017.

[4] NASSU B T, UKAI M . Rail extraction for driver support in railways[C]// 2011 IEEE Intelligent Vehicles Symposium (IV). Baden-Baden, Germany: IEEE, 2011: 83-88.

[5] 超木日力格 . 机车司机视野扩展系统及路轨障碍物检测的研究 [D]. 北京 : 北京交通大学 , 2012.
CHAO M. Locomotive Driver’s Vision Expansion System and Roadblock Detection Algorithm[D]. Beijing: Beijing Jiaotong University, 2012.

[6] 王前选 , 梁习锋 , 刘应龙 , 等 . 铁路钢轨视觉识别检测方法 [J]. 中南大学学报 (自然科学版), 2014,45(7):2496-2502.
WANG QX, LIANG X F, LIU Y L, et al. Railway rail identification detection method using machine vision[J]. Journal of Central South University(Science and Technology), 2014,45(7):2496-2502.

[7] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(4):640-651.

[8] 李松泽 . 基于深度学习的车道线检测系统的设计与实现 [D]. 哈尔滨：哈尔滨工业大学 , 2016.
LI S Z. The Design and Implementation of the Lane Detection System Based on Deep Learning[D]. Harbin: Harbin Institute of Technology, 2016.

[9] BADRINARAYANAN V , KENDALL A , CIPOLLA R . SegNet: A Deep Convolutional Encoder-Decoder Architecture for Scene Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39: 2481-2495.

[10] PASZKE A, CHAURASIA A, KIM S, et al. ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation[EB/OL].[2020-02-28]. https://arxiv.org/pdf/1606.02147.pdf.

[11] ROMERA E, ÁLVAREZ J M, BERGASA L M, et al. ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation[J]. IEEE Transactions on Intelligent Transportation Systems, 2017(1):1-10.

[12] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]// Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2980-2988.

[13] 孙晓凯 , 倪卿元 , 陈文强 . 图像增强方法在深度学习图像识别场景应用中的可行性研究 [J]. 电信科学 , 2020, 36(增刊 1):172-179.
SUN X K, NI Q Y, CHEN W Q. Feasibility study of image enhancement method in deep learning image recognition scene[J]. Telecommunications Science ,2020, 36(S1):172-179.

[14] WANG P, CHEN P, YUAN Y, et al. Understanding Convolution for Semantic Segmentation[C]// 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.