

车道线检测的 PSPNet 改进算法^{*}

霍爱清 冯若水 李 易

(西安石油大学电子工程学院 西安 710065)

摘 要: 车道线检测已成为智能驾驶领域研究的一项重要课题,而实际应用时,常出现车道线分割不准确、实时检测能力不佳的问题。为此本文提出一种金字塔场景分析网络的改进算法。在编码结构的基础上搭建主体网络 PSPNet,选用 MobileNet v2 轻量级网络作为编码器的主干网络,减少了整体网络的计算复杂度及参数量;网络中添加了空洞卷积,并在不同层间实现特征融合,扩充了模型感受野,同时丰富了局部特征;最后用自适应直线拟合算法对各类型车道线拟合。本文使用 Caltech 车道线数据集进行测试,实验结果显示,改进后的 PSPNet 算法对不同类型的车道线均有较好的分割结果,与 PSPNet 算法相比精度和交并比分别提升了 3.91%、4.14%,且 FPS 达 28 帧/s,本文算法的分割精度和推理速度均优于其他对比算法。

关键词: PSPNet;语义分割;MobileNet v2 网络;空洞卷积;自适应拟合

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Improved PSPNet algorithm for lane detection

Huo Ai Qing Feng Ruoshui Li Yi

(College of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China)

Abstract: Lane detection is a significant research subject in the field of intelligent driving. However, there will always be inaccurate lane segmentation and insufficient real-time processing capabilities in practical applications. Accordingly, an improved algorithm based on the Pyramid Scene Parsing Network is proposed. A main network PSPNet is built on a basis of the encoding structure, and the encoder backbone network is replaced by the lightweight MobileNet v2 network, which effectively cut down the parameter amount and computational complexity of the whole network. Hole convolution is added into the network and feature fusion is realized between different layers, which expands the model receptive field and enriches local feature information. Finally, an adaptive line fitting algorithm is used to fit different lane lines in order to obtain the final prediction result. The Caltech lane data set is come into use for testing. The experimental results show that the improved algorithm has better segmentation for different types of lane lines. Compared with the original algorithm, the Pixel Accuracy and the Intersection over Union is improved by 3.91%, 4.14%, and FPS up to 28 frames per second. The segmentation accuracy and inference speed of the proposed algorithm are superior to other comparison algorithms.

Keywords: PSPNet;semantic segmentation;MobileNet v2 network;atrous convolution;adaptive fitting

0 引 言

车道线检测一直是先进驾驶辅助系统^[1-3]的关键组成部分。带语义分割的车道线检测方法根据各像素的标签结果将车道线与背景进行分割,并获取车道线的轮廓信息及定位信息,以确保车辆能够安全行驶。

2015 年 Huval 等^[4]首次将 CNN 应用到以车道线为研

究对象的检测任务中,为后续车道线检测的发展奠定基础。同年 Long 等^[5]提出的全卷积神经网络更是将图像分割研究推向新的高度。当下许多主流的语义分割算法仍是在 FCN 的研究基础上实现的,如 U-Net^[6]、PSPNet^[7]、SegNet^[8]、Deeplab^[9-11]等。其中编码器部分主要使用的分类网络有 ResNet^[12]以及轻量级的网络 MobileNet^[13]。

近年来,基于深度学习的分割检测算法大致可分为基

收稿日期:2022-08-10

^{*} 基金项目:陕西省教育厅基金(17JS108)、西安石油大学研究生创新与实践能力的培养项目(YCS21213204)、陕西省科技厅一般工业项目(2020GY-152)资助

于编解码器的算法、集合上下文信息的算法及实例分割检测三大类^[14]。Li^[13]提出基于递归神经网络对编码器部分进行改进。庞彦伟等^[16]和 Shao 等^[17]基于道路的特征提取部分进行优化,实验结果得到改善但对边缘提取的结果依赖性太强整体不够稳定。

基于上下文信息集合的算法需要在多粒度特征的组合下平衡局部和全局信息,如 Piao 等^[18]以 Deeplabv3^[10]作为编码器提取并混合大尺度信息和多层次特征,提高了分割的鲁棒性。Chen 等^[19]提出在网络中添加空洞卷积^[20]来增大特征映射的区域范围,但该方法对于存在交叉路口的路边会出现误检现象。Chan 等^[21]将全卷积网络与方向透视映射相结合,获取精确的轮廓和全覆盖的拟合结果,其检测精度仍有提升的空间。

Neven 等^[22]以新的思路提出把车道线的语义分割检测转为实例分割检测,设计了一种语义分支和嵌入分支结合的检测网络 LaneNet。该算法被后续研究中广泛引用,如周苏等^[23]在 LaneNet 双分支基础新增一条对车道线数量预测分支,并将原算法中复杂的聚类算法改为自适应聚类,提高了对环境的适应性。但由于 LaneNet 基础参数量较大,文献[23]仍存在运行速度慢的问题。

通过上述分析,将卷积神经网络应用在车道线检测确实能大幅提高准确性和鲁棒性,但其庞大的参数量和复杂的运算过程也成为不可忽视的问题。因此,本文针对当前算法精确度和速度不平衡的问题提出一种 PSPNet 的改进

检测算法。首先,对网络的编码部分进行改进,选用轻量的 MobileNet v2 网络代替原主干网络 ResNet;其次,增加空洞卷积模块并加强不同层间的特征融合;最后用自适应拟合方法完成后续图像处理。

1 改进的 PSPNet 网络模型

本文采用轻量级网络对 PSPNet 模型编码器进行优化,并结合空洞卷积和集合上下文信息的层间特征融合对网络结构进行改进,提出了一种 PSPNet 改进检测算法。

1.1 PSPNet 网络

金字塔场景分析网络(pyramid scene parsing network, PSPNet)是 Zhao 等^[7]在 2016 年 ImageNet 挑战中提出的一种像素级图像分割网络,其算法总体结构如图 1 所示。其中虚线框内的是 PSPNet 网络的核心,即金字塔池化模块,该模块对 4 个维度不同的特征层分别进行池化,采用 1×1 卷积降维处理后,再经过上采样操作,将其还原的特征与初始特征叠加,形成具有更丰富的全局信息及子区域信息的特征表达。图 1 中,最上层的小块代表最浅层的特征,直接将池化后的上层特征再次平均池化;下面 3 个部分则是将池化后的特征划分为 2×2 、 3×3 、 6×6 的 3 个不同维度子区域,随后进行子区域内平均池化。相较于直接全局池化会造成丢失部分位置信息情况,利用金字塔池化兼顾了全局信息和各子区域间关系,因此使得场景分析中的准确率得到了很大提高。

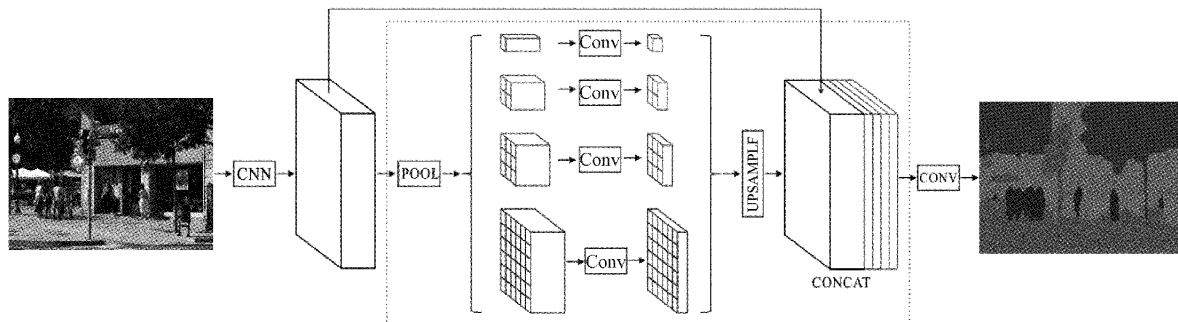


图 1 PSPNet 算法结构

1.2 改进 PSPNet 网络

为了提升检测精度,并降低网络运行时间,本文在 PSPNet 基础上提出了一种改进算法,其用 PSPNet 作为主体网络,以 MobileNet v2 作为编码器的主干特征提取网络,将空洞卷积及不同层间特征融合而成。改进的算法编码器结构如图 2 所示。

由图 2 可知,以 MobileNet v2 替换原主干特征提取网络,同时在主干网络提取的 Conv5 特征层中加入空洞卷积,这样可以在不增加计算量的前提下,扩大感受野,减少提取特征时短小型车道线的信息丢失;而后把 Conv4 和 Conv5 两个特征层混合叠加,用以加强层与层之间的特征信息融合,其结果作为金字塔池化部分的输入和总输出的

特征层分支。

1) MobileNet v2 网络

考虑到降低网络参数的需要,引入轻量级神经网络 MobileNet v2^[13]。为了防止梯度消失,该网络充分利用了残差模块,并参考 ResNet^[12]中瓶颈结构的思路设计了倒残差结构,如图 3 所示的是当步长取 1 时的倒残差卷积示意图。

通过在瓶颈内部采用多通道的方式增大映射空间。其工作过程是先经过一个 1×1 的卷积升维,获取更多特征,然后输入到 3×3 的空间卷积,最后再做一次 1×1 的逐点卷积,进行降维压缩。正是通过这种先升维后降维的思想有效降低了参数量。

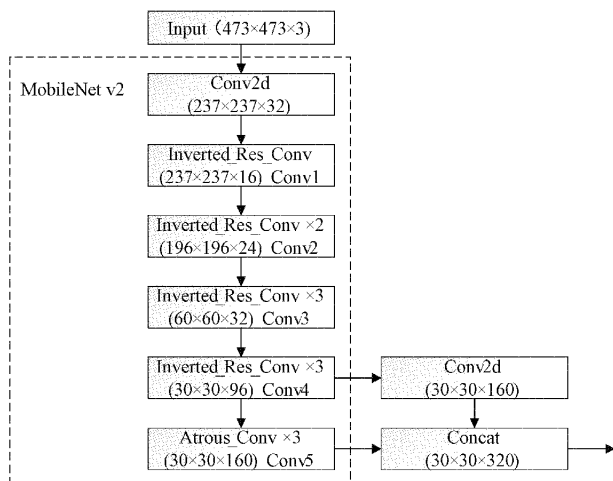


图 2 改进 PSPNet 算法的编码器结构

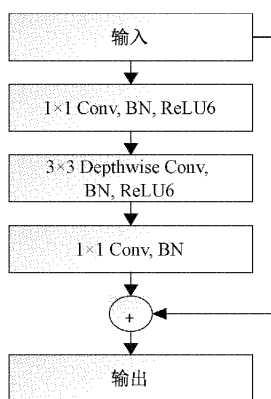


图 3 倒残差卷积结构(步长为 1)

另外,使用线性的激活函数来代替最后一层的 ReLU6 激活层,避免了 ReLU 层在低维度时造成信息的丢失,因此只针对最后一层的激活函数进行替换,可以减少信息损耗,避免有效信息丢失。

2) 空洞卷积

本文在 MobileNet v2 网络中还加入了空洞卷积操作,增大上层卷积感受野,获取更丰富的全局信息。空洞卷积通过给卷积核部分插入空洞使卷积核膨胀,其中空洞的概念就是跨像素的提取特征点。卷积层中引入空洞会产生新的参数空洞率 r , r 所表示的是卷积核做卷积计算时各像素点的间隔。令卷积核的大小取 3×3 , 当 $r=1$ 时,可将其视为普通卷积,即感知域大小不变仍是 3×3 ; 当 $r=2$ 时,可经计算得感知域扩大成 5×5 。理论上越大的空洞可以获得更大的感受野,但过大的空洞率容易造成车道线局部细节丢失,因此本文选择用 3 个空洞卷积的串行,其结构如图 4 所示,将空洞率分别设为 2、3、5,在提高信息利用率的同时避免了网络效应。

1.3 自适应拟合

完成对车道线的分割聚类后,需要对不同的车道线进行拟合,从而实现后续图像的处理。针对实际路面车道

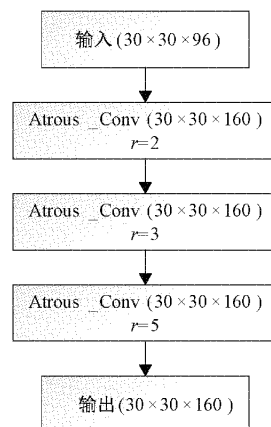


图 4 串行空洞卷积结构

线线型具有多样性,本文提出自适应拟合车道线算法,先用最小二乘法获得直线拟合的样本偏差 δ ,以此来判定是否需要再进行高次项的曲线拟合。具体步骤如下:

- 1) 输入聚类后得到的车道线样本点集合 $M(x, y)$ 。
- 2) 划分不同车道线的样本点集合 $M'(x', y')$, 并对集合内样本点分别拟合。
- 3) 首先对车道线进行直线拟合,由偏差 δ 与设定阈值 h_1 判定是否进行曲线拟合: 若 $\delta \leq h_1$, 则拟合控制另一条车道线的样本点集合 $M''(x'', y'')$; 若 $h_2 \geq \delta \geq h_1$, 则进行二阶曲线拟合; 若 $h_3 \geq \delta \geq h_2$, 则直接进行三阶曲线拟合; 若 $\delta \geq h_3$, 则不拟合。以 δ 最小为原则, 遍历全部车道线样本点集合得到拟合结果。

采用单一的直线拟合算法与本文自适应拟合算法的对比拟合效果如图 5 所示。当检测直线时,两种拟合效果均可满足要求;而当车道线为曲线型时,单一直线的结果不理想,自适应拟合算法则可以将曲线轮廓准确拟合出来,与车道线标签图达到高度重合,展示了良好的适应性。

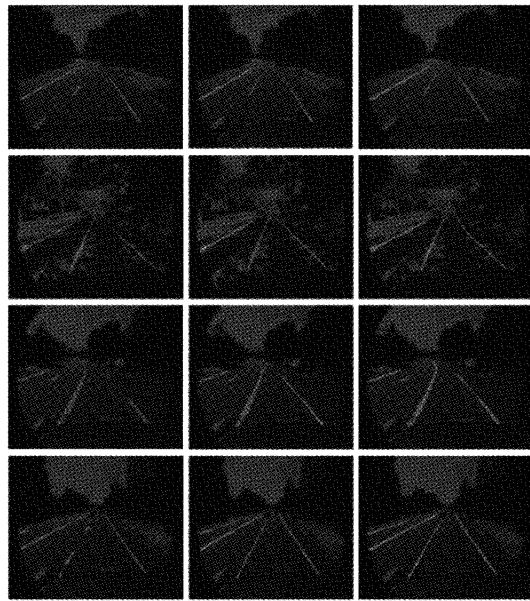
2 实验及结果分析

本文实验阶段,训练模型、测试和验证使用的是在 Ubuntu16.04 操作系统上的 Keras 深度学习框架,网络训练基于单个显卡 NVIDIA GTX1060 6 GB GPU 进行,GPU 的加速基于 CUDA9.0 以及 CUDNN7.05 完成。使用 Caltech 车道线数据集完成模型训练及实现检测效果可视化。

训练中,用梯度下降法进行 100 个 epochs 的训练,并将学习率设为不同大小,每隔 25 个 epochs 分别设置成 0.000 1、0.000 5、0.000 01 和 0.000 05,为防止显存不足,将前 50 和后 50 个周期的批处理大小分别设为 8 和 4。

2.1 实验数据集

本文使用 Caltech 车道线数据集进行研究和实验。该数据集的图片尺寸为 640×480 , 共计 1 225 张图片分为 4 个独立集合。其中, cordoval 有 250 张, cordova2 有 406 张, washington1 有 337 张, washington2 有 232 张。把数据集



(a) 车道线标签 (b) 直线拟合 (c) 自适应拟合

图 5 不同算法拟合效果图

中的图片以 VOC 格式存储,并以 9 : 1 划分出训练集和测试集。标注时将实线和虚线统一视为长短不同的车道线,标作 lane,双实线标作 double-lane。

2.2 实验结果

本文完成了消融实验、可视化实验以及不同算法的对比实验,并由各实验结果进行验证和分析。

首先,为验证本文算法的有效性,设计了相应的消融实验,在 PSPNet 网络中引入了 MobileNet v2、空洞卷积及特征融合模块,分析各模块对参数量、平均交并比(mean intersection over union, MIoU)以及每秒帧率(frames per second, FPS)的 3 个参数影响,实验结果如表 1 所示。

由表 1 可知,本方法用轻量网络 MobileNet v2 作为 PSPNet 算法的主干网络,使模型的参数量下降为原来的 36%,且 FPS 提升到 32 帧/s,而 MIoU 下降了 0.12%;为进一步提升车道线分割的准确率,在 MobileNet v2 网络中增添了空洞卷积和特征融合模块,使 MIoU 达到了 65.51%,与原算法相比,本文改进算法的 MIoU 提高了 4.14%,且视频每帧频率可以达到 28 帧/s,满足实时性。

其次,进一步将本文算法与 PSPNet 算法进行了可视

表 1 模块消融实验

模块	参数量/ M	MIoU/ %	FPS/ (帧·s ⁻¹)
PSPNet	5.29	61.37	21
MobileNet v2	1.93	61.25	32
MobileNet v2+空洞卷积	2.19	63.72	30
MobileNet v2+空洞卷积+ 特征融合	2.21	65.51	28

化对比。分别对弯道、车辆部分遮挡、树荫大面积遮挡、标识干扰等 4 种场景进行车道线检测,可视化结果对比效果如图 6 所示。



(a) 弯道 (b) 车辆部分遮挡 (c) 树荫大面积遮挡 (d) 标识干扰

图 6 可视化结果对比图

图 6(a)中 PSPNet 算法对弯道尽头处的曲线型车道线分割效果不佳,而本文改进算法依然持有良好分割效果,能够展现曲线的线形趋势;PSPNet 算法在图 6(b)、(c)中存在被遮挡部分的分割不连续问题,在图 6(d)中有边缘车道线的缺失问题,本文算法的可视化结果将这两个问题都加以解决。由此可见,相较于 PSPNet 算法在分割时存在局限性,改进算法在复杂场景下能够更好适应路面环境,其分割和拟合效果更好,对于复杂路面具有一定的鲁棒性。

最后,在 Caltech 测试数据集上将本文改进 PSPNet 算法与其他主流车道线分割检测算法作对比实验,实验结果如表 2 所示。其中,将数据集上的车道线分为 lane 和 double-lane 进行测试,并且采用交并比(intersection over union, IoU)、像素精度(Pixel Accuracy, PA)以及两种线型的 MIoU 和 MPA 作为评价指标。

表 2 不同算法对比实验

算法	IoU _{lane}	IoU _{double-lane}	PA _{lane}	PA _{double-lane}	MIoU/%	MPA/%
SegNet ^[8]	56.75	60.63	80.80	83.06	58.69	81.93
U-Net ^[6]	61.13	65.45	85.25	89.59	63.29	87.42
DeeplabV3+ ^[11]	63.73	66.15	85.92	89.94	64.94	87.93
PSPNet ^[7]	60.01	62.73	82.89	88.81	61.37	85.85
改进 PSPNet	64.08	66.94	87.59	91.93	65.51	89.76

从实验结果可得出,改进 PSPNet 算法有良好的检测结果,其 MIoU 和 MPA 均为最高值。UNet 和 PSPNet 两种算法都采用了特征融合的方法提高特征的有效信息。DeeplabV3+ 算法对特征提取部分进行了优化,相较于其他算法具有更好的检测效果,使 MIoU 达到了 64.94%。本文改进算法对于 PSPNet 的编码部分进行优化,引入空洞卷积的同时融合上下文信息获取更丰富的全局信息,获得最好结果,并使 MIoU、MPA 均获得提升。

3 结 论

针对车道线检测过程中存在分割准确度低以及处理速度慢的问题,本文提出了一种 PSPNet 的改进检测算法。首先,选用 MobileNet v2 作为主干网络,有效减少了网络的参数量,提高特征提取的运行速度;引入空洞卷积模块同时利用不同层融合加强局部特征信息,输入金字塔池化层后与全局特征进行拼接,使上下文信息有效结合;通过自适应拟合算法,有效提高了车道线拟合效果。最后与原 PSPNet 算法相比,本文改进算法在测试集上的 MIoU 提高了 4.14%,且视频速度达到 28 帧/s,合理验证了本文改进算法的有效性。根据各实验结果和可视化结果展现,与现有其他主流网络结构相比,本算法能充分结合上下文信息,对于 lane 和 double-lane 两种车道线的刻画更为准确,而且在弯道、遮挡、干扰等场景下均具有良好的分割结果。

参考文献

- [1] ZOU Q, JIANG H, DAI Q, et al. Robust lane detection from continuous driving scenes using deep neural networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1):41-54.
- [2] SAKHARE K V, TEWARI T, VYAS V. Review of vehicle detection systems in advanced driver assistant systems[J]. Archives of Computational Methods in Engineering, 2020, 27(2): 591-610.
- [3] 李成勇,王莎,陈成瑞.基于 CMOS 图像采集的车道偏移识别系统设计与实现[J].中国测试,2022,48(6): 106-110.
- [4] HUVAL B, WANG T, TANDON S, et al. An empirical evaluation of deep learning on highway driving[J]. ArXiv Preprint,2015, ArXiv:1504.01716.
- [5] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. Proc of International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [7] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]. Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.
- [8] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [9] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [10] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. ArXiv Preprint,2017, ArXiv:1706.05587.
- [11] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. Proc of the European Conference on Computer Vision (ECCV), 2018: 801-818.
- [12] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]. Proc of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [13] SINGH B, TOSHNIWAL D, ALLUR S K. Shunt connection: An intelligent skipping of contiguous blocks for optimizing MobileNet-V2 [J]. Neural Networks, 2019, 118: 192-203.
- [14] 吴一全,刘莉.基于视觉的车道线检测方法研究进展[J].仪器仪表学报,2019,40(12): 92-109.
- [15] LI Y. A progressive image semantic segmentation method using recurrent neural network[C]. Proc of 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), 2021: 765-768.
- [16] 庞彦伟,修宇璇.基于边缘特征融合和跨连接的车道线语义分割神经网络[J].天津大学学报:自然科学与工程技术版,2019,52(8):779-787.
- [17] SHAO R, QIAN B, GUO J. A general lane detection algorithm based on semantic segmentation[C]. Proc of the 2nd International Conference on Vision, Image and Signal Processing, ICVISIP, 2018: 1-5.
- [18] PIAO C H, DENG X T, LIU M J. Improved lane detection network based on instance segmentation[C]. Proc of the 2nd International Conference on Computing

- and Data Science, 2021: 1-5.
- [19] CHEN P R, LO S Y, HANG H M, et al. Efficient road lane marking detection with deep learning [C]. Proc of 2018 IEEE 23rd International Conference on Digital Signal Processing(DSP), 2018: 1-5.
- [20] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [J]. ArXiv Preprint, 2016, ArXiv:1511.07122.
- [21] CHAN Y C, LIN Y C, CHEN P C. Lane mark and drivable area detection using a novel instance segmentation scheme [C]. Proc of 2019 IEEE/SICE International Symposium on System Integration(SII), 2019: 502-506.
- [22] NEVEN D, BRABANDERE B D, GEORGOULIS S, et al. Towards end-to-end lane detection: An instance segmentation approach [C]. Proc of 2018 IEEE Intelligent Vehicles Symposium(IV), 2018: 286-291.
- [23] 周苏, 吴迪, 金杰. 基于卷积神经网络的车道线实例分割算法 [J]. 激光与光电子学进展, 2021, 58(8): 381-388.

作者简介

霍爱清, 博士, 教授, 主要研究方向为导向钻井控制、人工智能与深度学习等。

E-mail: aqhuo@xsyu.edu.cn

冯若水, 硕士研究生, 主要研究方向为人工智能, 目标检测等。

E-mail: 410907643@qq.com

李易, 硕士, 主要研究方向为深度学习与计算机图像处理等。

E-mail: 1163897986@qq.com