

杜家宇, 王书强. 基于语义分割网络的智能驾驶车道检测算法研究[J]. 智能计算机与应用, 2026, 16(2): 197-200. DOI: 10.20169/j. issn. 2095-2163. 24041603

基于语义分割网络的智能驾驶车道检测算法研究

杜家宇, 王书强

(河北工程大学 信息与电气工程学院, 河北 邯郸 056038)

摘要: 本文针对当前很多语义分割网络在车道检测任务中因受外界环境因素干扰造成检测精度低的问题, 提出了一种基于 ESPNet 网络优化改进方法。在特征提取网络末端以高效通道注意力和位置注意力模块为基础构建融合注意力机制, 在加强了局部跨通道信息交互的同时增强了空间上下文信息聚合能力。改进后的模型在 BDD100k 数据集上进行验证, 实验结果表明, 改进后语义分割模型检测精度在车道线以及可行驶区域的检测任务中均有提升。

关键词: 智能驾驶; 语义分割; 高效空间金字塔; 融合注意力机制

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2026)02-0197-05

Research on intelligent driving lane detection algorithm based on semantic segmentation network

DU Jiayu, WANG Shuqiang

(School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, Hebei, China)

Abstract: This paper proposes an optimization and improvement method based on ESPNet network to address the problem of low detection accuracy in lane detection tasks caused by external environmental factors in many current semantic segmentation networks. Based on the efficient channel attention and position attention modules at the end of the feature extraction network, a fusion attention mechanism is constructed to enhance the spatial context information aggregation ability while enhancing local cross channel information interaction. The improved model is validated on the BDD100k dataset, and the experimental results show that the detection accuracy of the improved semantic segmentation model is improved in both lane line and drivable area detection tasks.

Key words: intelligent driving; semantic segmentation; efficient spatial pyramid; fusion attention mechanism

0 引言

随着智能驾驶技术的快速发展和智能交通系统的普及, 车道检测在计算机视觉领域中发挥着至关重要的作用。车道线提供了道路结构和交通规则的关键信息, 对于实现自动驾驶、车辆行为分析和驾驶辅助^[1-2]等应用具有重要意义。因此, 高效准确的车道检测算法成为了研究的热点和挑战。

传统检测方法主要依据颜色、形状、纹理^[3]等特征对车道进行划分, 如边缘检测、曲线拟合^[4]等。传统车道检测方法易受到天气、光照等外界因素影响, 抗干扰能力较差, 难以符合当前高精度的检测要求。随着计算机算力的提升, 深度学习在自动驾驶领域展现出了非凡的潜力, 依靠卷积神经网络^[5]强

大的特征提取能力对车道进行检测成为当前主流的方法。图像语义分割^[6-8]作为计算机视觉领域重要的研究方向, 在车道检测领域中取得了众多优秀成果。与传统方法相比, 基于深度学习的语义分割网络能够自动从图像中学习车道线的特征表示, 从而提高了检测的准确性和鲁棒性。

Neven 等学者^[9]在 ECCV 会议上首次提出了 LaneNet 分割网络, 以逆透视变换的方法将车道视角转换为鸟瞰图进行检测, 最后恢复原视角。Zhang 等学者^[10]提出了一种双卷积门控单元 ConvGRU 来处理车道标记的低级特征信息并以连续帧为输入处理时空驾驶信息来应对复杂环境下的车道线检测。Lee 等学者^[11]提出了一种针对车道检测优化的简单但强大的自注意力机制, 称为扩展自注意力(ESA)

作者简介: 杜家宇(2000—), 男, 硕士研究生, 主要研究方向: 图像语义分割。Email: 1952192736@qq.com; 王书强(1975—), 男, 副教授, 主要研究方向: 智能控制, 数据库技术, 人工智能。

收稿日期: 2024-04-16

哈尔滨工业大学主办 ◆ 科技创新与应用

模块,提取全局上下文信息,对遮挡部分加以估计,有效提高了检测精度。鲁维佳等学者^[12]提出了一种基于循环特征融合 Resa-CC 的弯道增强车道线检测算法,利用车道线的形状先验性,捕获图像像素中行与列的空间关系,融合信息生成特征图来提高大曲率车道线检测精度。Wang 等学者^[13]使用基于 Retinex 理论的无监督学习网络来增强夜间图像,提高图片的对比度,然后在 ERFNet 编码层中引入自注意力机制,在车道线整体较暗、照度不均匀的情况下,提高车道检测的准确度。Ati 等学者^[14]提出分布式多尺度金字塔池 (DMPP) 策略,以一种高效的多尺度上下文聚合 (EMCA) 模块,分别使用具有大和小膨胀率的小和大内核的组合来提取多尺度上下文在多个层次的功能层次结构,优化了道路场景的分割精度。

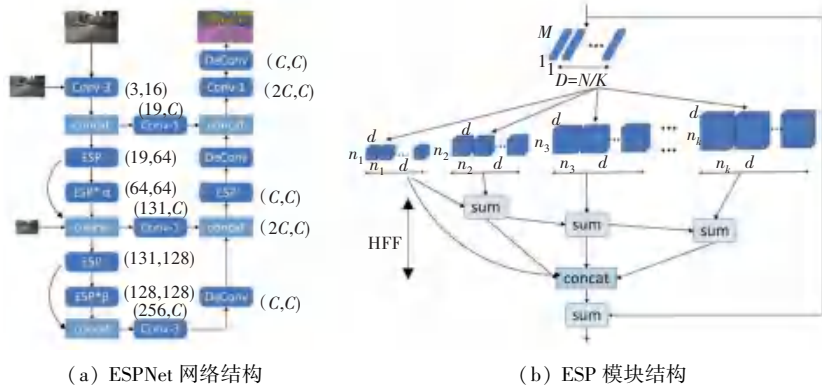


图 1 ESPNet 网络结构及 ESP 模块结构

Fig. 1 ESPNet network architecture and ESP module architecture

ESPNet 基于高效卷积金字塔 (ESP) 模块构建如图 1(b) 所示。ESP 模块将卷积层分解为逐点卷积和扩张卷积金字塔。逐点卷积将高维特征映射到低维空间以减少计算量,扩张卷积金字塔以不同空洞率的扩张卷积同时对特征图进行采样,在不增加参数的条件下扩大了感受野,同时对不同空间级别的特征进行提取。为解决扩张卷积金字塔引入的特征图网格化伪影问题,ESP 模块使用了 HFF 策略,将不同扩张率扩张卷积采样得到的特征图分层融合以消除影响。

1.2 注意力机制

1.2.1 高效通道注意力机制^[17]

当前大多数注意力机制为获得更好的性能效果而增加自身复杂性,这不可避免地增大了模型的计算开销。高效通道注意力模块 (ECA) 在增加少量参数的情况下带来了显著的增益,利用全局上下文信息计算通道权重,不对特征图位置特征进行池化

本文为进一步提高车道检测精度和检测算法在复杂环境中的抗干扰能力,提出一种非对称的编码器-解码器^[15]网络结构,对车道线和可行驶区域进行检测。该模型以 ESPNet 为骨干,搭配融合注意力机制引导的解码器,引入了高效通道注意力以及位置注意力机制。经实验验证,该模型在保障检测速度的同时有效提高了检测精度。

1 方法

1.1 ESPNet 基本原理

Mehta 等学者^[16]提出了一种快速高效的卷积神经网络结构 ESPNet,如图 1(a) 所示。应用于资源限制下的语义分割任务,在参数量较低的同时仍然保持较高的分割精度。

和卷积,减少了计算开销。相比于 SENet^[18],ECA 模块避免了通道降维带来的副作用,并以更为有效的方式捕获局部跨通道交互信息,如图 2 所示。

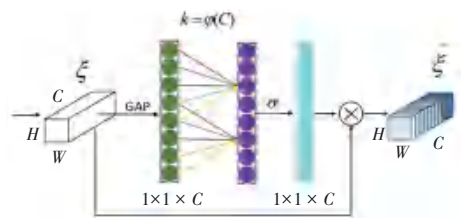


图 2 ECA 注意力模块

Fig. 2 ECA attention module

ECA 模块通过大小为 k 的一维卷积计算通道权重, k 值由通道数 C 映射自适应确定, C 与 k 的映射关系如下:

$$k = f(C) = \left\lfloor \frac{\log_2(C)}{\varepsilon} + \frac{\mu}{\varepsilon} \right\rfloor_{\text{odd}} \quad (1)$$

其中, ε 和 μ 分别为 2 和 1。

经过 f 变换,高维通道呈现更大范围的相互作用

用,低维通道呈现小范围的相互作用,更高效地捕获了跨信道交互信息。

1.2.2 位置注意力机制^[19]

为了获得更加丰富的上下文信息,引入位置注意力机制(PAM)如图 3 所示。经过 Softmax 层得到的注意力图 S 与 D 矩阵相乘,乘参数 α 后再与原特征图 A 逐元素相加得到输出 E 。 α 初始为 0 并通过学习逐渐增加权重,输出特征图 E 中每个元素由各个位置特征和原始特征加权求和得到,能通过注意力图聚合全局信息。

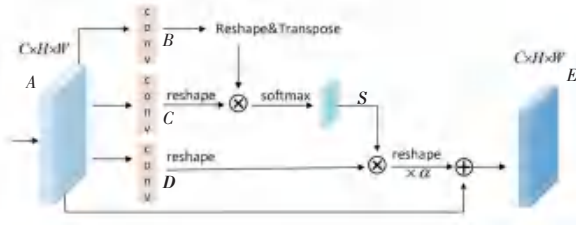


图 3 位置注意力模块

Fig. 3 Position Attention Module

1.2.3 基于融合注意力机制^[20-22]改进的 ESPNet

在车道检测场景中,车道线样本在整体图像中占比较小,为提高模型对细节信息的学习能力,增强细节特征表示,引入 ECA 模块,同时利用 PAM 模块加强上下文信息的聚合能力,在增加少量参数的条件下提高车道检测精度。编码器部分沿用 ESPNet-C,在特征图拼接过程中将图像映射到相同的尺寸空间以增强信息流。融合注意力机制中先以 ECA 模块形成注意力图再与原特征图融合,接着以相同策略进行 PAM 注意力加权。最后,以轻量级解码器恢复特征图尺寸并进行像素分类。基于融合注意力机制改进的模型结构如图 4 所示。



图 4 改进 ESPNet

Fig. 4 Improved ESPNet

1.3 损失函数

结合车道线数据样本不平衡而导致难以分类的特点,本研究采用 Focal 损失函数^[23]缓解该问题,使模型在反向传播过程中更好地更新参数。Focal 损失函数公式如下:

$$\text{Loss}_{\text{Focal}} = -\frac{1}{N} \sum_{j=0}^{J-1} \sum_{i=1}^N p_i(j) (1 - p_i(j))^{\alpha} \log(\hat{p}_i(j)) \quad (2)$$

其中, N 表示输入像素数量; J 表示分割对象类

别; $p_i(j)$ 表示真值; $\hat{p}_i(j)$ 表示预测值。

Focal loss 在数据样本类别不平衡的条件下,对于像素类别的划分表现良好。

2 实验及结果分析

2.1 数据集介绍

BDD100k^[24]数据集是伯克利大学 AI 实验室发布的针对自动驾驶场景的大型数据集,共 10 万张高清图像,图像分辨率为 1280×720 ,其中 7 万张训练集,2 万张测试集和 1 万张验证集。图像包含多种天气、场景以及不同的时间,可应用于车道检测、可行驶区域检测、交通对象检测等多种任务场景。本文采用 BDD100k 中车道线和可行驶区域两种语义分割任务数据进行研究。

2.2 模型评价标准

语义分割任务普遍采用平均准确率、平均交并比等指标作为衡量精度的标准。这些标准基于混淆矩阵计算,设定预测值为正且真实值为正的像素类别为真正 T_p ; 预测值为负且真实值为负的像素类别为真负 T_n ; 预测值为正但真实值为负的像素类别为假正 F_p ; 预测值为负但真实值为正的像素类别为假负 F_n 。混淆矩阵见表 1。

表 1 混淆矩阵

Table 1 Confusion matrix

真实值	预测值	
	正	负
正	T_p	F_n
负	F_p	T_n

准确率 (P_A) 以及平均准确率 (MP_A) 计算公式如下:

$$P_A = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (3)$$

$$MP_A = \frac{1}{K} \sum_{i=1}^K \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (4)$$

交并比 (IoU) 以及平均交并比 (MIoU) 计算公式如下:

$$\text{IoU} = \frac{T_p}{T_p + F_p + F_n} \quad (5)$$

$$\text{MIoU} = \frac{1}{K} \sum_{i=1}^K \frac{T_p}{T_p + F_p + F_n} \quad (6)$$

2.3 实验环境参数

本研究基于 ESPNet 网络,针对车道检测的难点对网络结构进行优化。实验采用的配置为显卡

NVIDIA GeForce RTX 4090, 处理器 Xeon (R) Platinum 8352V, CUDA 11.3.1, Python 3.8.5, Pytorch 1.12.0, 部分训练参数见表 2。

表 2 训练参数

Table 2 Training parameters

优化器	初始学习率	权重衰减	Max epoch	Batch size
Adam	0.000 5	0.000 5	300	16

2.4 对比试验

在 BDD100k 数据集上对改进后模型进行训练, 由于 BDD100k 测试集标签未公开, 模型仅在验证集中进行验证。本文所提出模型与 ESPNet 以及其他同数据集验证过的网络实验结果对比见表 3、表 4。

表 3 可行驶区域分割对比结果

Table 3 Comparison results of driving area segmentation

网络结构	MIoU/%	FPS
MultiNet	71.6	8.6
DLT-Net	72.1	9.3
ESPNet	88.4	96.0
PSPNet	89.6	11.1
本文结构	89.7	85.0

表 4 车道线分割结果对比

Table 4 Comparison of lane line segmentation results

网络结构	IoU/%	MP_A /%	FPS
ENet	14.64	34.12	100
SCNN	15.84	35.79	19
ENet-SAD	16.02	35.56	50
ESPNet	25.26	58.63	96
本文结构	27.12	78.44	85

表 3、表 4 展示本研究改进后的模型与常见模型在可行驶区域分割以及车道线检测任务中的结果对比, 改进 ESPNet 在交并比、准确率、参数量以及推理速度方面均表现优异。相比于 ESPNet 引入了少量参数, 在推理速度不明显降低同时提高了检测精度。

2.5 可视化效果

图 5 为改进 ESPNet 对道路中车道线以及可行驶区域的检测结果。图 5(a)~(d) 中, 左侧图片为白天, 右侧图片为夜间。可以看出该模型在白天以及光线不充分的夜间均有良好的分割效果, 有较强的抗干扰能力。

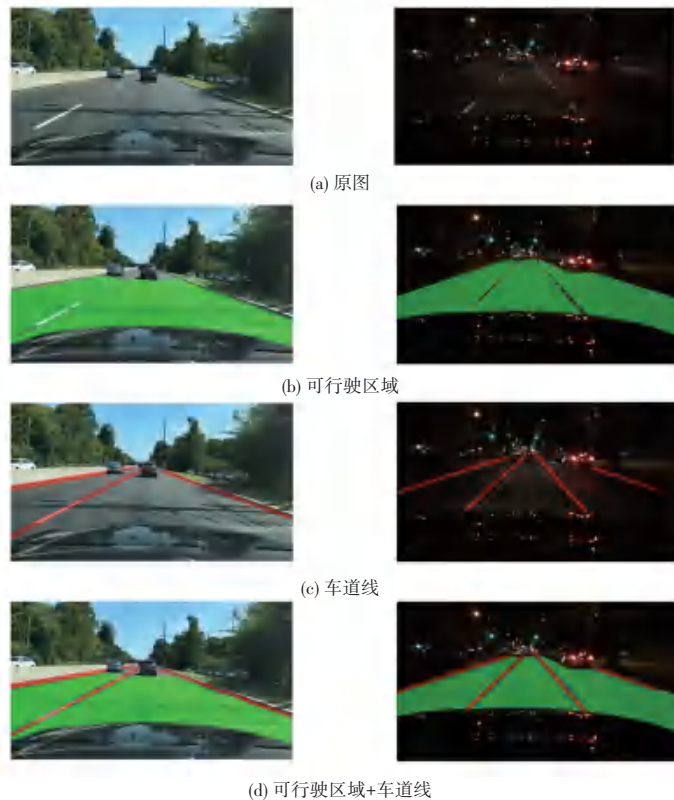


图 5 车道线及可行驶区域检测效果

Fig. 5 Detection effect of lane markings and drivable areas

3 结束语

本研究基于轻量级语义分割网络设计车道检测算法。本研究的重点是优化 ESPNet 语义分割网络,以提高车道检测的性能和效率。针对车道检测任务中由于天气、光线等外界因素干扰而造成车道线检测精度低的问题,本文提出了一种优化策略,以进一步提高车道检测算法的性能。以 ESPNet 为主干,利用扩张卷积金字塔提取多尺度车道特征,引入高效通道注意力模块捕获跨通道信息,增强车道线细节特征,融合位置注意力模块聚合上下文语义信息,最后通过轻量级解码器对可行驶区域和车道线进行分割。通过大量的实验和结果分析,相比于 ESPNet,本算法在 BDD100k 数据集上可行驶区域 mIoU 提升了 1.3%,车道线 IoU 提升了 1.86%。经实验验证,本文提出的改进策略对模型检测精度有提升。

参考文献

[1] 冯霞,陈爽,卢敏,等. 融合多源时空信息鸟瞰图的未来实例分割预测[J]. 吉林大学学报(工学版), 2025, 55(10): 3372-3383.

[2] 王越,曹家乐,孙学斌,等. 融合空间语义的自动驾驶视觉联合感知算法[J]. 太原理工大学学报, 2025, 56(2): 338-347.

[3] 王碧瑶,韩毅,崔航滨,等. 基于图像的道路语义分割检测方法[J]. 山东大学学报(工学版), 2023, 53(5): 37-47.

[4] 蔡汶良,黄俊. 基于 RepVGG 网络的实时车道线检测方法[J]. 计算机科学, 2024, 51(7): 236-243.

[5] 黄雯珂,滕飞,王子丹,等. 基于深度学习的图像分割综述[J]. 计算机科学, 2024, 51(2): 107-116.

[6] 张婧,张策,张茹,等. 图像分割述评:基本概貌、典型算法及比较分析[J]. 计算机技术与发展, 2024, 34(1): 1-8.

[7] SHI J, ZHAO L. A review of lane detection based on semantic segmentation[J]. International Journal of Advanced Network, Monitoring and Controls, 2021, 6(3): 1-8.

[8] 景庄伟,管海燕,彭代峰,等. 基于深度神经网络的图像语义分割研究综述[J]. 计算工程, 2020, 46(10): 1-17.

[9] NEVEN D, BRABANDERE B D, GEORGIOULIS S, et al. Towards end-to-end lane detection: An instance segmentation approach[C]//Proceedings of the IEEE Conference on Intelligent Vehicles Symposium. Piscataway, NJ; IEEE, 2018: 286-291.

[10] ZHANG Jiyong, DENG Tao, YAN Fei, et al. Lane detection model based on spatio-temporal network with double convolutional gated recurrent units[J]. IEEE Transactions on Intelligent Trans-

portation Systems, 2021, 23(7): 6666-6678.

[11] LEE M, LEE J, LEE D, et al. Robust lane detection via expanded self attention[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway, NJ; IEEE, 2022: 533-542.

[12] 鲁维佳,刘泽帅,潘玉恒,等. 基于循环特征融合的弯道增强车道线检测算法[J]. 测绘通报, 2023(12): 25-30.

[13] WANG J, GE Y, HAN C, et al. Nighttime lane detection based on Retinex theory and self-attention distillation[C]// Proceedings of the 35th Chinese Control and Decision Conference. Shenyang, China; Northeastern University, 2023: 309-314.

[14] ATI F, MAZHAR S, AHAMED S R, et al. DMPNet: Distributed multi-scale pyramid network for real-time semantic segmentation[J]. IEEE Access, 2024, 12: 16573-16585.

[15] SHI M, LIN S, YI Q, et al. Lightweight context-aware network using partial-channel transformation for real-time semantic segmentation[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7401-7406.

[16] MEHTA S, RASTEGARI M, CASPI A, et al. ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham; Springer, 2018: 561-580.

[17] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2020: 11534-11542.

[18] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2018: 7132-7141.

[19] FU Jun, LIU Jing, TIAN Haijie, et al. Dual attention network for scene segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2019: 3146-3154.

[20] 李涛,高志刚,管晟媛,等. 结合全局注意力机制的实时语义分割网络[J]. 智能系统学报, 2023, 18(2): 282-292.

[21] 宗豪迈. 基于深度学习的车道线检测方法研究与应用[D]. 西安; 西安理工大学, 2023.

[22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham; Springer, 2018: 3-19.

[23] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ; IEEE, 2017: 2980-2988.

[24] YU F, CHEN Haofeng, WANG Xin, et al. Bdd100k: A diverse driving dataset for heterogeneous multitask learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2020: 2636-2645.