

DOI:10.19651/j.cnki.emt.2417090

基于多尺度空洞融合注意力的车道线检测算法^{*}

李沐原 张兰春 张博源

(江苏理工学院汽车与交通工程学院 常州 213001)

摘要: UFSA-LD算法在提取车道线的细长结构特征时,面临信息丢失、长距离上下文捕获困难以及边界细节识别不敏感等挑战。本文提出一种基于多尺度空洞特征融合注意力的车道线检测算法:在UFSA-LD辅助分割分支加入MDFA模块,借助空洞空间金字塔池化(ASPP)扩展了网络的感受野,多尺度捕捉车道特征;利用融合注意力机制(FCBAM)从通道与空间多维度过滤干扰信息,增强关键特征表示。引入Dice Loss损失函数,更多关注车道线的边缘与局部结构信息。实验结果表明:改进后模型在TuSimple数据集上检测精度由95.81%提升至96.03%;在CULane数据集上F1指标较原文提升1.8,验证了模型改进的有效性。

关键词: 车道线检测;ASPP;融合注意力;Dice Loss

中图分类号: TN209 **文献标识码:** A **国家标准学科分类代码:** 510.99

Lane detection method based on multi-scale dilated fusion attention

Li Muyuan Zhang Lanchun Zhang Boyuan

(School of Automotive and Traffic Engineering, Jiangsu Institute of Technology, Changzhou 213001, China)

Abstract: The UFSA-LD algorithm faces challenges in extracting the thin and long structural features of lane lines, such as information loss, difficulty in capturing long-distance context, and insensitivity to boundary detail recognition. This paper proposes a lane line detection algorithm based on multi-scale atrous feature fusion attention; an MDFA module is added to the UFSA-LD auxiliary segmentation branch, and the receptive field of the network is expanded through atrous spatial pyramid pooling (ASPP) to capture lane features at multiple scales; a fusion channel and spatial attention mechanism (FCBAM) is used to filter out interfering information from channel and spatial dimensions, enhancing the representation of key features. The introduction of the Dice Loss loss function focuses more on the edges and local structural information of the lane lines. Experimental results show that the detection accuracy of the improved model on the TuSimple dataset has been increased from 95.81% to 96.03%; the F1 metric on the CULane dataset has improved by 1.8 compared to the original, validating the effectiveness of the model improvement.

Keywords: lane detection; ASPP; fusion attention; Dice Loss

0 引言

车道线检测作为自动驾驶感知系统的基石,在车道保持系统^[1]、自适应巡航控制^[2]等高级驾驶辅助功能中发挥着至关重要的作用。鉴于其在提升道路交通安全、促进自动驾驶技术及智慧交通系统发展中的关键性,车道线检测技术的研究具有深远的学术价值与应用前景。

在传统车道线检测方法中,研究者通常依赖于手工设计的特征提取技术和基于规则的检测策略。例如,Xu^[3]在HSL颜色空间中探索车道线检测,而姜立标等^[4]则利用改进的Hough变换算法进行实时道路引导线检测。然而,在多变的道路环境中,这些传统方法往往需要复杂的参数调

整和规则设计,其泛化能力和可扩展性受限。近年来,随着人工智能技术的蓬勃发展,基于深度学习的车道线检测方法取得了显著进展,相较于传统算法,这些方法能够自动提取图像中的车道线特征^[5],展现出卓越的建模和特征学习能力。

当前大部分学者都将车道线检测视为语义分割问题。Zheng等^[6]提出了一种创新的递归特征移位聚合器模块,采用水平和垂直的周期性切片特征图移动,以便更有效地收集像素间的跨行和列空间信息。这一方法显著提高了全局信息的聚合能力,从而有效缓解了在长距离信息传播过程中可能出现的丢失问题。Hou等^[7]提出一种突破性的知

收稿日期:2024-10-12

* 基金项目:江苏省高校自然科学基金重大项目(22KJA580002)、江苏理工学院研究生实践创新计划项目(XSJCX23_34)资助

识蒸馏方法:自注意力蒸馏(self attention distillation, SAD),并集成至ENet网络的编码器中,利用深层特征的丰富信息有效增强了浅层特征表现。受此启发,多数学者开始探索注意力机制在本领域的应用。崔建东等^[8]引入一种双维度的自注意力机制,缓解了特征提取时的依赖减弱问题。胡序洋等^[9]设计了一种混合注意力机制与多尺度特征提取网络并行策略,使模型扩大感受野的同时,更加专注于当前任务的特征信息。艾青林等^[10]借鉴前人的思想,将小目标类别注意力机制与特征融合的策略引入到非结构化场景的语义分割领域。贾远鹏等^[11]则另辟蹊径,设计了一种双注意力机制并行策略:采用通道注意力与行列注意力机制分别提取高、低阶特征图的语义信息与空间信息,极大丰富了整体特征。Liu等^[12]引入了基于条件卷积和逐行表述的条件车道检测策略,并设计了一种递归实例模块(recurrent instance module, RIM)来处理具有复杂拓扑结构的车道线。Han等^[13]则在Transformer架构的基础上设计了一种专为车道线优化过的自注意力机制,有效捕获了车道线的形状特征和语义信息。然而,尽管这些模型拥有较强的特征提取能力,但其模型庞大,推理速度慢,计算复杂度高,并且依赖繁琐的后处理步骤,无法满足实时检测的要求。

针对采用像素分割方法的缺点,Qin等^[14]将车道线检测看作一个以行线为单位的全局性特征选择问题,提出了一种基于结构感知的快速车道线检测方法^[15](ultra fast structure-aware deep lane detection, UFSA-LD)。该算法利用具有较大感受野的全局特征来处理复杂的道路场景,并且提出用于建模车道的结构性损失,使模型的计算成本显著降低,并在速度方面取得相当优异的表现。但同时因为在提取车道细长形状特征时,模型存在重要信息丢失、长

距离上下文信息难以捕捉以及对边界信息敏感度不高等问题,模型检测精度不佳。

深入分析UFSA-LD算法在车道特征提取方面存在的问题后,本文提出了一种基于多尺度空洞特征融合注意力机制的车道线检测算法。针对原模型在固定感受野下难以捕获长距离上下文信息的局限,本文在辅助分割层中,创新性地引入多尺度空洞特征融合注意力模块(multi-scale dilated fusion attention, MDFA)。MDFA模块巧妙地结合了空洞空间金字塔池化^[16](atrous spatial pyramid pooling, ASPP)和融合注意力机制(fusion convolutional block attention module, FCBAM),不仅显著扩展了网络的感受野,还有效保留了关键的车道特征信息,并对这些特征进行了精细的选择性增强。此外,考虑到车道线检测任务对边界信息的敏感性,采用Dice Loss^[17]函数替代传统的交叉熵损失函数,以更好地关注车道边界的细节信息。通过消融实验,本文验证了MDFA模块和Dice Loss函数的有效性。在精度上与其他具备代表性的先进方法对比,本文算法虽存在一定差距,但在推理速度上具备一定优势,在精度与速度间达到较好的平衡,满足实时性的需求。

1 UFSA-LD 算法介绍

1.1 UFSA-LD 算法思想

UFSA-LD作为一种轻量化的车道线检测模型,用基于全局特征预定义行锚的方法代替用局部感受野来做逐像素分割的方法,满足了车道检测实时性的要求。具体分析如图1所示。图像被划分为若干行锚后又被划分为若干个细胞单元格,在每个单元格中定位车道线的位置。因为无需进行逐像素分割,UFSA-LD相比其他模型推理速度大幅提升。

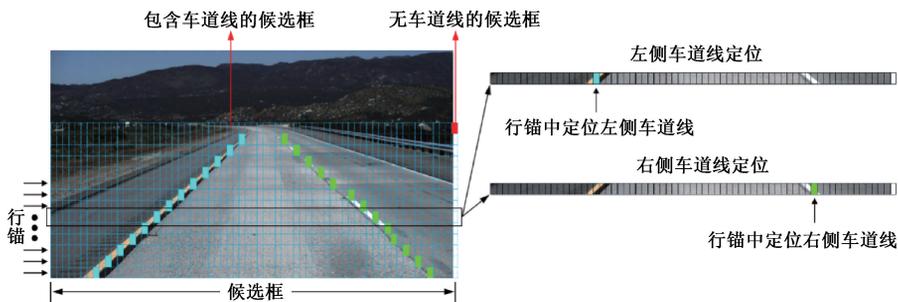


图1 UFSA-LD算法分析

Fig.1 Analysis of UFSA-LD Algorithm

1.2 UFSA-LD 算法模型

UFSA-LD模型结构如图2所示,模型由3部分组成:蓝色方框内是由Res-Net构成的主干特征提取网络;红色方框内为辅助分割模块,用于在训练过程中增强视觉特征,但不参与推理;绿色方框内是分类预测模块,在全局特征上进行行索引来计算候选点,将车道线检测问题转化为分类问题。

2 改进算法介绍

2.1 改进的空洞空间金字塔池化层

UFSA-LD算法在提取特征时无法有效捕捉输入图像多尺度长距离上下文信息,是因为其辅助分割模块由传统卷积层构成,感受野的广度与卷积核的尺寸成正比,且通常为固定值。扩展感受野可以增强模型对车道线长距离

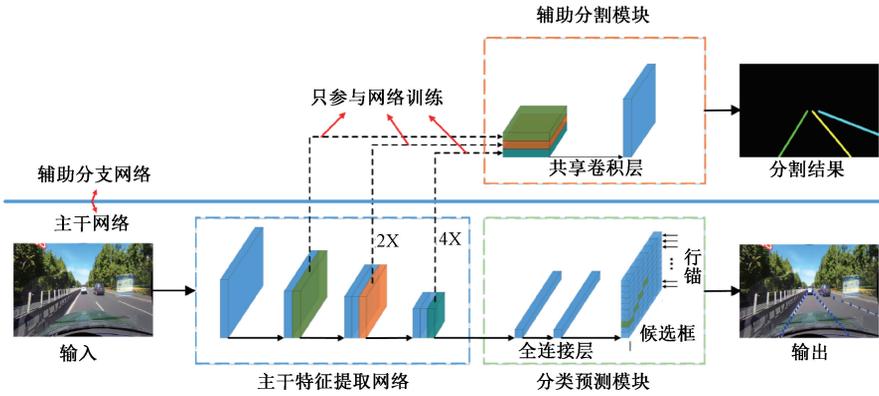


图 2 UFSA-LD 模型结构

Fig. 2 UFSA-LD model structure

上下文信息的捕获能力。尽管增加卷积核尺寸是扩大感受野的一种直接策略,但这不可避免地会导致模型参数数量的增加,显著提升计算成本。

空洞卷积^[18]作为一种扩展普通卷积核感受野的方法,常受分割领域的学者青睐。其显著优势在于:能够在不增加模型参数数量和计算复杂度的情况下,显著扩大卷积核的感受野。如图 3 所示,空洞卷积在卷积核的元素之间引入了“空洞”(见图中深蓝色区域),即在元素之间插入间隔,使得卷积核可以覆盖更广泛的输入区域,而无需增加卷积核的大小或层数。空洞卷积的卷积核范围和其空洞率之间的关系如式(1)。

$$k_{eff} = k + (k - 1) \times (r - 1) \quad (1)$$

式中: r 是空洞率; k 是初始卷积核的范围; k_{eff} 是空洞卷积的实际感受野大小。通过增加空洞率,卷积核可以在更大的输入区域内捕捉特征,这对于图像分割和目标检测等需要大范围上下文信息的任务,尤为重要。

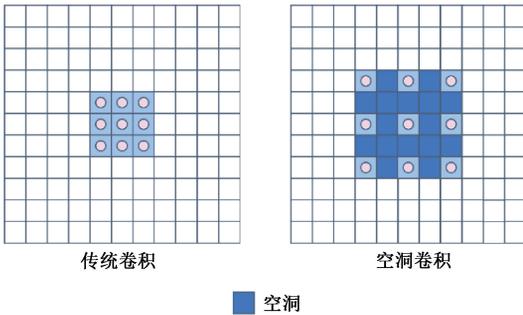


图 3 空洞卷积

Fig. 3 Atrous Convolution

受此启发,本文计划在辅助分割分支引入 ASPP 模块。该模块原本包含 4 个不同空洞率的并行 3×3 卷积,通过降维处理输入特征,在多个尺度上捕获图像的上下文信息。其结构如图 4 所示。

然而,Chen 等^[19]发现当空洞率变大时, 3×3 卷积滤波器退化为 1×1 滤波器,只有中心权重有效。因此,本文使用 1×1 卷积替换图中卷积第 4 分支(即 $rate=24$),并在原

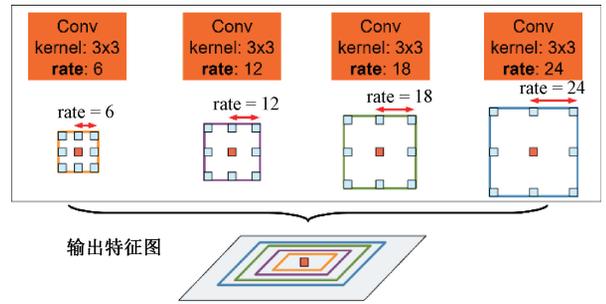


图 4 ASPP 结构

Fig. 4 ASPP structure

模块基础上增加一个全局平均池化分支,在局部特征中融入全局上下文信息,增强模型对整体场景结构的理解力。改进后的结构如图 5 所示。

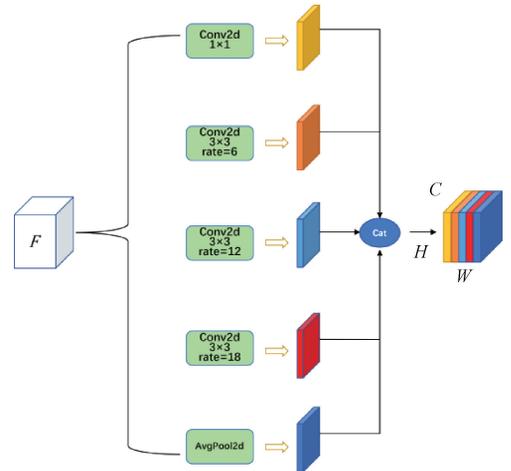


图 5 改进后 ASPP 结构

Fig. 5 Improved ASPP structure

2.2 FCBAM 注意力机制

卷积块注意力(convolutional block attention module, CBAM)^[20]是一种设计精巧且高效的注意力机制,专门用于增强卷积神经网络的表示能力。具体而言,CBAM 首先通过全局平均池化和最大池化生成通道注意力图,然后通

过卷积操作生成空间注意力图,并依次应用这两个注意力图以增强特征图的表达。其整体结构如图6所示。

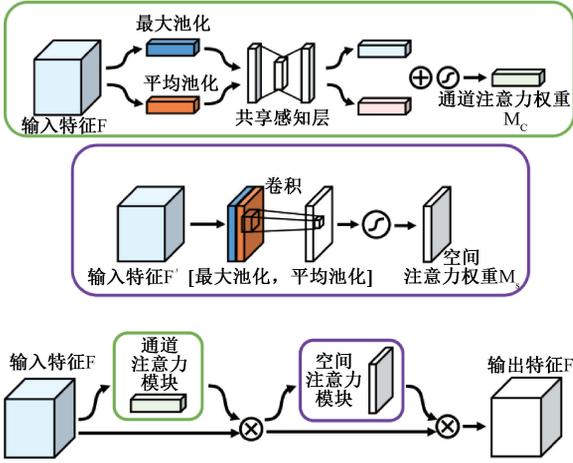


图6 CBAM结构

Fig. 6 CBAM structure

CBAM 包含两个子模块——通道注意力模块(channel attention module, CAM)和空间注意力模块(spatial attention module, SAM)。CAM 模块的目的是关注特征图中哪些类型的特征有意义,应用最大池化和平均池化压缩特征图的空间维度,在通道维度上选择性地强调对当前任务最有用的特征,降低噪声影响。其数学解释如式(2)和式(3)。

$$M_c(F) = \sigma(W_1(W_0(F_{avg})) + W_1(W_0(F_{max}))) \quad (2)$$

$$F' = M_c(F) \otimes F \quad (3)$$

式中: $F \in \mathbb{R}^{C \times H \times W}$ 是输入特征图, $F_{avg}, F_{max} \in \mathbb{R}^{C \times 1 \times 1}$ 是池化后的特征图; $W_0 \in \mathbb{R}^{\frac{C}{r} \times C}$ 和 $W_1 \in \mathbb{R}^{C \times \frac{C}{r}}$ 是可学习的权重矩阵, r 是缩减比例因子, σ 是 Sigmoid 激活函数。 \otimes 表示逐元素相乘操作, $F' \in \mathbb{R}^{C \times H \times W}$ 是经过通道注意力增强后的特征图。

同理, SAM 模块重点关注特征图中哪些位置的特征有意义,沿通道维度应用类似操作提高空间定位能力,增强区域信息。其数学解释如式(4)和(5)。

$$M_s(F) = \sigma(f^{7 \times 7}([\mathbf{F}_{avg}^{spatial}; \mathbf{F}_{max}^{spatial}])) \quad (4)$$

$$F'' = M_s(F') \otimes F' \quad (5)$$

式中: $\mathbf{F}_{avg}^{spatial}, \mathbf{F}_{max}^{spatial} \in \mathbb{R}^{1 \times H \times W}$ 表示池化后的特征, $[\cdot; \cdot]$ 表示通道上的拼接, $f^{7 \times 7}$ 是 7×7 的卷积, $M_s(F) \in \mathbb{R}^{1 \times H \times W}$ 是空间注意力图, F'' 是经过空间注意力增强后的最终特征图。两个子模块结构简单,易于集成,共同作用,能够在图像分类,目标检测等任务中显著提升性能。

针对解决 UFSA-LD 算法在提取车道线特征时存在丢失重要信息,模型精度不佳等问题,本文借鉴 CBAM 的核心思想,提出了 FCBAM。其结构如图7所示。

车道线检测任务要求模型能够同时处理来自不同信息源的多种特征,如纹理、边缘和空间布局,并在复杂场景

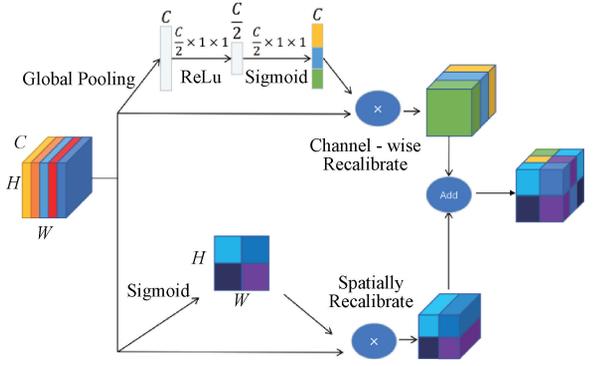


图7 FCBAM结构

Fig. 7 FCBAM structure

中准确识别和分割车道线。相较 CBAM 模块, FCBAM 在设计上进行了若干关键改进,以适应车道线检测任务的特定需求。

区别于 CBAM 中通道和空间注意力顺序执行的方式, FCBAM 采用了并行化的注意力机制,分别在通道和空间维度上对特征图进行选择性强,消除了顺序依赖性,从而使模型能够更精确地捕捉特征图中的关键信息。同时, FCBAM 去除了最大池化操作,避免在通道注意力处理中可能引入的噪声和冗余信息,从而提高了模型的鲁棒性。此外,为提高计算效率, FCBAM 简化了注意力图的生成过程。通过在通道和空间注意力模块中使用 1×1 卷积,降低了计算复杂度并减少了参数量。最后, FCBAM 将通道和空间注意力图通过逐元素最大值操作进行融合,形成最终的特征增强,生成输出特征图。数学推导过程如下:

$$M_c(F) = \sigma(W_1(\text{ReLU}(W_0(F_{avg})))) \quad (6)$$

$$F' = M_c(F) \otimes F \quad (7)$$

$$M_s(F) = \sigma(f^{1 \times 1}(F)) \quad (8)$$

$$F'' = M_s(F) \otimes F \quad (9)$$

$$F_{out} = \max(F'', F') \quad (10)$$

式中: $F \in \mathbb{R}^{C \times H \times W}$ 是输入特征图, $F_{avg} \in \mathbb{R}^{C \times 1 \times 1}$ 是池化后的特征图; $W_0 \in \mathbb{R}^{\frac{C}{r} \times C}$ 和 $W_1 \in \mathbb{R}^{C \times \frac{C}{r}}$ 是可学习的权重矩阵, r 是缩减比例因子, σ 是 Sigmoid 激活函数。 \otimes 表示逐元素相乘操作, $F' \in \mathbb{R}^{C \times H \times W}$ 是经过通道注意力增强后的特征图, $f^{1 \times 1}$ 是 1×1 的卷积, $M_s(F) \in \mathbb{R}^{1 \times H \times W}$ 是空间注意力图, F'' 是经过空间注意力增强后的特征图。

2.3 Dice Loss 损失函数

在语义分割领域,二值交叉熵损失函数因其直观性和计算效率而被广泛采用。然而,对于车道线检测这一特定任务而言,其对边界信息的敏感性要求模型不仅要捕捉对象的几何形态,更需细致地识别边缘特征,包括颜色、纹理和光照变化等细节。传统的二值交叉熵损失函数在处理类别不平衡问题时存在局限,往往难以充分捕捉少数类别的特征表达。

为了应对这一挑战,本研究引入了 Dice Loss 损失函

数,以期提升模型对于车道线这类少数类别的识别能力。Dice Loss 通过计算预测结果与真实标签之间的交集与并集的比值,优化了模型对少数类别的预测性能。该损失函数特别强调了对重叠区域的最大化,从而在边缘信息的捕捉以及结构一致性方面展现出显著优势。通过这种方式,Dice Loss 有助于提升模型对于车道线边界细节的敏感度,进而增强车道线检测的整体性能。

其数学表达式如下:

$$L_{dice} = 1 - \frac{2 \sum_{t=1}^N p_t y_t + \epsilon}{\sum_{t=1}^N p_t + \sum_{t=1}^N y_t + \epsilon} \quad (11)$$

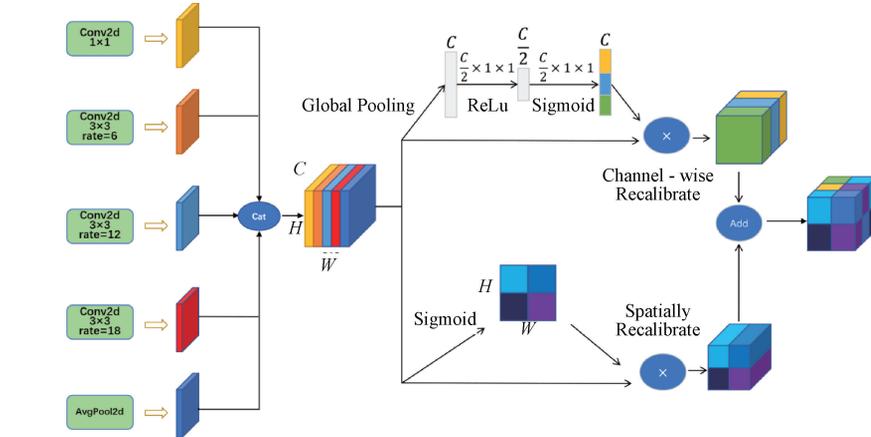


图 8 MDFA 整体结构

Fig. 8 Overall structure of MDFA

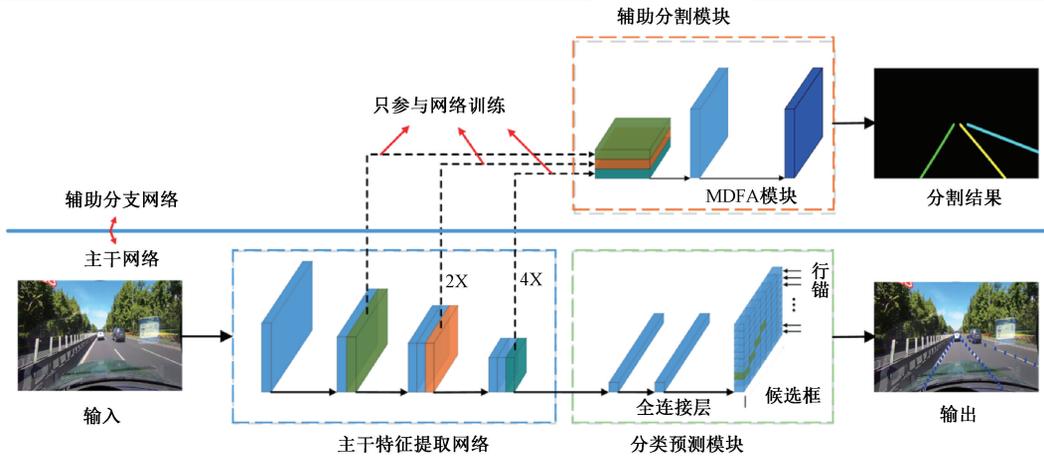


图 9 UFSA-LD 改进后模型结构

Fig. 9 UFSA-LD improved model structure

式中: N 是像素的个数; t 是一个迭代的索引变量; p_t 是预测的概率; y_t 是真实的标签; ϵ 是一个平滑系数。

UFSA-LD 模型损失包括分类损失、结构损失和分割损失。结合 Dice Loss 后,模型整体的损失函数定义为:

$$L_{total} = L_{cls} + \alpha L_{str} + \beta (\mu L_{seg} + \gamma L_{dice}) \quad (12)$$

式中: 分割损失 L_{seg} 的系数 μ 设置为 0.75, L_{dice} 的系数 γ 设置为 0.25。

2.4 模型搭建

依据 2.1 和 2.2 节所述,本文在辅助分割分支中将 ASPP 和 FCBAM 相连接,定义为 MDFA。其结构如图 8 所示。改进后网络的整体结构如图 9 所示。

3 实验

本实验使用的 CPU 为 Intel (R) Core (TM) i5-12490F,内存为 32 GB,GPU 为 NVIDIA GeForce RTX 3060 Ti,系统为 Ubuntu 20.04,开发环境为 PyTorch 2.2.1,CUDA 为 12.1。

3.1 数据集

本文选用 TuSimple 数据集和 CULane 数据集进行相关实验和研究。TuSimple 数据集由图森未来发布,包含 72 520 张训练图像和 55 640 张测试图像,采用 JSON 格式的中心点标注法,通过标注视频的最后一帧来构建。一张图片的标注为一个元组,车道线结构清晰,弯曲弧度不明显。

CULane 数据集由香港中文大学发布,包含 88 880 张训练图像、9 675 张验证图像和 34 680 张测试图像。车道线以粗实线标注,除正常工况外还提供了 8 种挑战性场景。两个数据集的具体介绍如表 1 所示。

表 1 数据集介绍

Table 1 Dataset introduction

数据集	训练集	验证集	测试集	车道线	场景	环境
TuSimple	72 520	—	55 640	≤5	1	高速
CULane	88 880	9 675	34 680	≤4	9	乡镇、高速

3.2 评价方法

在目标检测研究领域,查准率(Precision)与查全率(Recall)构成了评价模型性能的两个基本维度。查准率是指在所有被模型判定为正类的样本中,实际为正类样本所占的比例,它量化了模型预测的精确性。查全率是指在所有实际为正类的样本中,被模型正确判定为正类的样本所占的比例,表征了模型预测的完整性。

CULane 数据集采用 F_1 分数作为评价指标, F_1 是准确率和召回率的加权平均调和结果。三者具体的计算公式为:

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (15)$$

其中,TP(true positive)代表正确识别的正例,即模型准确地将正类别样本预测为正类别;FP(false positive)代表误报的正例,即模型错误地将负类别样本预测为正类别;FN(false negative)代表漏报的负例,即模型错误地将正类别样本预测为负类别。TuSimple 数据集的评价方法是模型精度(Accuracy),其定义为:

$$Accuracy = \frac{\sum_{clip} C_{clip}}{\sum_{clip} S_{clip}} \quad (16)$$

式中: C_{clip} 为正确预测车道点数; S_{clip} 为每个图片中的真实点总数。

3.3 消融实验

UFSA-LD 原文中采用了 ResNet-18 和 ResNet-34 两个主干网络进行测试;本文所改进模型均以 ResNet-18 为主干网络进行测试。训练参数如表 2 所示。在进行对比实验之前,先采用 TuSimple 数据集进行消融实验。实验结果如表 3 所示。

表 2 训练参数

Table 2 Training parameters

数据集	lr	epoch	batch_size	optimizer
TuSimple	4×10^{-4}	50	32	Adam
CULane	0.1	100	32	SGD

表 3 不同改进对模型精度的影响

Table 3 Effects of different improvements on model accuracy

实验编号	MDFA	Dice Loss	Accuracy/%
1	—	—	95.81
2	—	√	95.88
3	√	—	95.96
4	√	√	96.03

注:√表示引入该模块,—表示未引入该模块

由表 3 可知,在 UFSA-LD 原模型基础上只添加 Dice Loss 损失函数时,精度达到 95.88%,添加 MDFA 模块时,精度提升至 95.96%。同时添加 MDFA 模块和 Dice Loss 函数时,模型精度提升有较大提高。

3.4 对比实验

为评估模型性能,本文在两个数据集上选用多种具有代表性先进方法与本文方法进行对比实验。

在 TuSimple 数据集上的测试结果如表 4 所示。以 ResNet-18 为骨干网络的改进模型对比原模型在精度上从 95.81%提高到了 96.03%,误检率(FP)从 19.05%降低到了 18.69%,漏检率(FN)从 3.92%降低到了 3.58%。在精度上较 LaneATT、LkLaneNet^[21] 模型有一定提升;而与 Laneformer^[13]、DAMLAP-Net^[11]、LaneNet 和 EL-GAN 等模型相比,本文模型在精度上存在一定差距,但在推理速度上具有明显优势,满足车道线检测实时性与轻量化的需求。

在 CULane 数据集上的复杂场景测试结果如表 5 所示。改进后的模型相较原模型,在不同场景下的表现均有所提升。具体来说,在正常(Normal)场景下精度从 87.7%提高到了 89.1%,拥挤(Crowded)场景下精度从 66.0%提高到了 69.6%,夜间(Night)场景下精度从 65.1%提高到了 65.3%,阴影(Shadow)场景下精度从 62.8%提高到了 68.9%,箭头(Arrow)场景下精度从 81.0%提高到了 84.4%,曲线(Curve)场景下精度从 57.9%提高到了 59.1%。然而,在无车道线(No-line)场景下精度从 40.2%下降到了 38.4%,在眩光(Dazzlelight)场景下精度从 58.4%下降到了 57.3%。总体而言,模型的精度从 68.4%提高到了 70.2%。

3.5 实例可视化分析

为了充分验证本文方法在不同交通场景中的适用性,从 CULane 数据集中随机挑选了一部分强光照、遮挡、阴影、弯道、夜晚等复杂场景的车道线图像进行可视化检测,检测结果如图 10 所示。选用同为分割类方法的 ResNet-Seg、SCNN 以及自注意力蒸馏办法 SAD 进行对比分析。ResNet-Seg 方法在多个场景中出现漏检和线性不光滑问题;SAD 在大部分场景表现了良好的检测效果,但在无车道场景出现了误检情况;SCNN 方法在正常和阴影场景表现了良好的检测效果,但在其他场景多次出现部分漏检情

表 4 不同网络在 TuSimple 上测试结果和运行时间对比

Table 4 Comparison of test results and running time of different networks on TuSimple

方法	Accuracy/%	FP/%	FN/%	Runtime/ms	Multiple
SCNN	96.53	6.17	1.80	133.5	1.0x
SAD ^[7]	96.64	6.02	2.05	13.4	10.0x
Res18-UFSA-LD ^[14]	95.81	19.05	3.92	3.2	41.7x
LaneATT	95.57	3.56	2.92	4.0	33.4x
LkLaneNet ^[21]	95.94	2.06	3.28	7.9	16.96x
Laneformer ^[13]	96.85	2.01	2.03	20	6.8x
DAMLAP-Net ^[11]	96.46	4.76	5.75	12.7	10.5x
LSCoMer ^{[[22]}	96.68	2.76	3.20	5.6	24.0x
LaneNet	96.38	7.80	2.44	19.0	7.0x
EL-GAN	96.39	4.12	3.36	>100	<1.3x
Res18-UFSA-LD(本文)	96.03	18.69	3.58	7.3	18.69x

注:部分论文结果以本文复现结果为准。

表 5 不同网络在 CULane 上测试结果和运行时间对比

Table 5 Comparison of test results and running time of different networks on CULane

类别	Res50-seg	Res34-SAD	SAD	SCNN	FD-50	Res18-UFSA-LD	Res18-UFSA-LD(本文)
正常	87.4	89.9	90.1	90.6	85.9	87.7	89.1
拥挤	64.1	68.5	68.8	69.7	63.6	66.0	69.6
夜晚	60.6	64.6	66.6	66.1	57.8	62.1	65.3
无车道线	38.1	42.2	41.6	43.4	40.6	40.2	38.4
阴影	60.7	67.7	65.9	66.9	59.9	62.8	68.9
箭头	79.0	83.8	84.0	84.1	79.4	81.0	84.4
眩光	54.1	59.9	60.2	58.5	57.0	58.4	57.3
曲线	59.8	66.0	65.7	64.4	65.2	57.9	59.1
交叉路口	2 505.0	1 960.0	1 998.0	1 990.0	7 013.0	1 743.0	2 218.0
总体	66.7	70.7	70.8	71.6	—	68.4	70.2
Runtime/ms	—	50.5	13.4	133.5	—	3.1	6.90
Multiple	—	2.6x	10.0x	1.0x	—	43.0x	40.6x
FPS	—	19.8	74.6	7.5	—	322.5	143

注:交叉路口一行中的数据为 FP,数值越小检测效果越好。

况;而本文方法中通过空洞空间金字塔多尺度地捕捉车道特征,减少了车道线漏检与末端检测不到位情况,并依靠通道-空间注意力融合机制,多维度地过滤了冗余和干扰信息,注重关键语义信息,缓解了路沿、栏杆、阴影等相似条状物的误检情况,同时借助 Dice Loss 损失函数,更多关注车道线的边缘与局部结构信息,在大部分场景均展现了良好的检测效果,但在无车道和眩光场景存在线性不光滑问题。

3.6 本文工作局限性与未来改进方向

经 3.4 与 3.5 节对比分析可知,本文方法虽在准确性与实时性之间达到了较好的平衡,满足了在实际应用中的需求,但相较于表 4 中 LSCoMer^{[[22]},本文算法在综合性能上

仍存在一定差距,不具竞争力。具体而言,LSCoMer^{[[22]}在精度上比本文方法高出 0.65%,在运行时间上每帧减少了 1.7 ms。LSCoMer^{[[22]}采用 CNN-Transformer 架构,将 Vit-CoMer 密集检测骨干运用于车道线检测任务,并深度融合局部与全局信息,促进 CNN 与 Transformer 间的信息交互,使模型具有优秀的准确性与实时性。未来本文将借鉴这种跨模态融合策略,以缓解本文方法在单一模态下信息不充分的局限性。

同时针对在无车道和眩光场景的线性不光滑问题,将提取视频连续帧之间的光流运动信息,在时间维度将模型提取的车道特征传播复用,进一步增强模型在部分特殊场景的检测连续性,提高模型整体的鲁棒性。

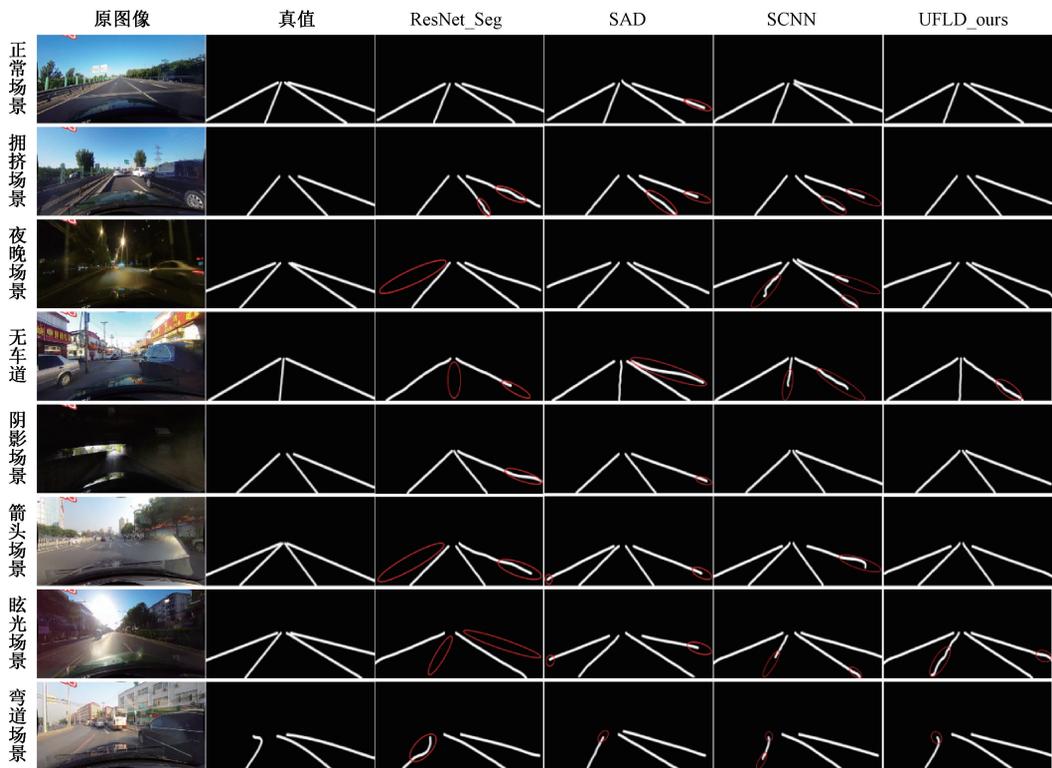


图10 部分场景实例可视化分析

Fig. 10 Visualization analysis of some scene instances

4 结 论

针对 UFSA-LD 在提取车道特征时,存在重要信息丢失、长距离上下文捕获困难以及边界细节识别不敏感等问题,本文提出一种基于多尺度空洞特征融合注意力机制的车道线检测算法。通过在 UFSA-LD 算法内创新性地引入 MDFA 模块和 Dice Loss 函数,多尺度,多维度地对模型特征提取进行增强,有效提升了模型在复杂场景下的检测精度和鲁棒性尽管本算法还存在一定不足,但本文在准确性与实时性之间取得较好的平衡,为自动驾驶车道线检测技术的发展提供了新的思路和方法。

参考文献

- [1] 李凯,韩增文,陈金建,等. 车道保持辅助系统的模糊输出反馈协同控制[J]. 传感器与微系统,2023,42(12): 95-98,102.
LI K, HAN Z W, CHEN J J, et al. Fuzzy output feedback cooperative control for lane-keeping assistance systems[J]. Transducer and Microsystem Technologies, 2023, 42(12): 95-98,102.
- [2] 李旭,谢宁,王建春,等. 面向切入场景的变权重自适应巡航控制策略[J]. 重庆理工大学学报(自然科学), 2023,37(4):10-18.
LI X, XIE N, WANG J CH, et al. Variable-weight

adaptive cruise control strategy for cut-in scenarios[J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(4): 10-18.

- [3] XU D. Color-enhanced lane line detection algorithm based on HSL color space[C]. Third International Conference on Computer Graphics, Image, and Virtualization (ICCGIV 2023). SPIE, 2023, 12934: 483-490.
- [4] 姜立标,李静轩. 基于改进 Hough 变换与双点去除 R-最小二乘法的车道线检测优化算法[J]. 科学技术与工程,2020,20(5):2070-2076.
JIANG L B, LI J X. Optimization algorithm for lane detection based on improved Hough transform and double-point removal R-least squares [J]. Science Technology and Engineering, 2020, 20 (5): 2070-2076.
- [5] DHAIF Z S, EL ABBADI N K. A review of machine learning techniques utilised in self-driving cars [J]. Iraqi Journal For Computer Science and Mathematics, 2024, 5(1): 205-219.
- [6] ZHENG T, FANG H, ZHANG Y, et al. Resa: Recurrent feature-shift aggregator for lane detection[C]. AAAI Conference on Artificial Intelligence, 2021, 35(4): 3547-3554.

- [7] HOU Y, MA Z, LIU C, et al. Learning lightweight lane detection CNNs by self attention distillation[C]. IEEE/CVF International Conference on Computer Vision, 2019: 1013-1021.
- [8] 崔建东, 崔岩. 一种基于多维度自注意力机制的轻量级车道线检测算法[J]. 计算机科学与应用, 2022, 12: 108.
CUI J D, CUI Y. A lightweight lane detection algorithm based on multi-dimensional self-attention mechanism[J]. Computer Science and Applications, 2022, 12: 108.
- [9] 胡序洋, 高尚兵, 汪长春, 等. LaneSegNet: 一种高效的车道线检测方法[J]. 南京信息工程大学学报(自然科学版), 2022(5):14.
HU X Y, GAO SH B, WANG CH CH, et al. LaneSegNet: An efficient lane detection method[J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2022(5):14.
- [10] 艾青林, 张俊瑞, 吴飞青. 基于小目标类别注意力机制与特征融合的 AF-ICNet 非结构化场景语义分割方法[J]. 光子学报, 2023, 52(1): 0110001.
AI Q L, ZHANG J R, WU F Q. AF-ICNet: Unstructured scene semantic segmentation method based on small target class attention mechanism and feature fusion [J]. Acta Photonica Sinica, 2023, 52(1): 0110001.
- [11] 贾远鹏, 陈学文, 哈瑞峰. 双注意力机制下自动驾驶汽车车道线深度感知研究[J]. 重庆理工大学学报(自然科学), 2023, 37(7): 44-50.
JIA Y P, CHEN X W, HA R F. Lane depth perception research for autonomous vehicles under dual attention mechanisms [J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(7): 44-50.
- [12] LIU L, CHEN X, ZHU S, et al. CondLaneNet: A top-to-down lane detection framework based on conditional convolution[C]. IEEE/CVF International Conference on Computer Vision, 2021: 3773-3782.
- [13] HAN J, DENG X, CAI X, et al. LaneFormer: Object-aware row-column transformers for lane detection [C]. AAAI Conference on Artificial Intelligence, 2022, 36(1): 799-807.
- [14] QIN Z, WANG H, LI X. Ultra fast structure-aware deep lane detection[C]. Computer Vision-ECCV 2020: 16th European Conference, 2020: 276-291.
- [15] 朱娟, 朱国吕, 岳晓峰. 融合 CAM 和 ASPP 的车道线检测算法研究[J]. 重庆理工大学学报(自然科学), 2024, 38(5):86-94.
ZHU J, ZHU G L, YUE X F. Research on lane detection algorithm based on CAM and ASPP fusion[J]. Journal of Chongqing University of Technology (Natural Science), 2024, 38(5): 86-94.
- [16] HE H, YANG D, WANG S, et al. Road extraction by using atrous spatial pyramid pooling integrated encoder-decoder network and structural similarity loss [J]. Remote Sensing, 2019, 11(9): 1015.
- [17] ZHAO R, QIAN B, ZHANG X, et al. Rethinking dice loss for medical image segmentation [C]. 2020 IEEE International Conference on Data Mining (ICDM). IEEE, 2020: 851-860.
- [18] 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.
- [19] CHEN L C. Rethinking atrous convolution for semantic image segmentation [J]. ArXiv preprint arXiv:1706.05587, 2017.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [21] LI S, WU X, WU Z. Efficient multi-lane detection based on large-kernel convolution and location [J]. IEEE Access, 2023, 11: 58125-58135.
- [22] 郑剑, 郭亦畅. 基于 LSTR 和 Vit-CoMer 骨干的车道线检测方法[J/OL]. 电子测量技术, 1-9[2024-10-12]. <http://kns.cnki.net/kcms/detail/11.2175.tn.20241021.1659.016.html>.
ZHENG J, GUO Y CH. Lane detection method based on LSTR and Vit-CoMer backbone[J/OL]. Electronic Measurement Technology, 1-9[2024-10-12]. <http://kns.cnki.net/kcms/detail/11.2175.tn.20241021.1659.016.html>.

作者简介

李沐原, 硕士研究生, 主要研究方向为车道线检测。

E-mail: 1404341673@qq.com

张兰春(通信作者), 博士研究生, 教授, 主要研究方向为新能源汽车传动系统控制与优化。

E-mail: zlc@jsut.edu.cn

张博源, 硕士研究生, 主要研究方向为车道线检测。

E-mail: zhangliuliu1022@outlook.com