

融合蛇形卷积与卷积注意力的车道线分割方法研究

黄忠盛, 魏 敏

(成都信息工程大学 计算机学院, 四川 成都 610225)

摘要: 针对车道线分割任务中的精度和鲁棒性问题, 采用一种融合蛇形卷积与卷积注意力的车道线分割方法, 以提高车道线识别准确性和稳定性。在模型设计方面, 首先引入蛇形卷积模块, 将其嵌入到主干网络中, 以增强特征提取能力, 并更好地捕捉车道线的形状信息; 随后, 引入卷积注意力模块, 用于动态调整特征图的权重, 使模型能够更加聚焦于车道线区域, 提升车道线分割精度。为了验证该方法的效果, 在BDD100K数据集上进行了实验评估。实验结果表明, 与主流模型Hybridenet、YOLOP、YOLOM相比, 准确率分别提高了0.8%、1.4%、1.3%, IoU分别提高了1.4%、5.1%、2.8%, 证明了蛇形卷积和卷积注意力在车道线分割中的有效性和互补性。

关键词: 蛇形卷积; 卷积注意力; 车道线分割; 实例分割

DOI: 10.11907/rjdk.241540

中图分类号: TP391.41

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2025)008-0082-06



Research on Lane Line Segmentation Method Integrating Snake Convolution and Convolutional Attention

HUANG Zhongsheng, WEI Min

(School of Computer Science, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: To solve the problem of accuracy and robustness in lane segmentation task, a deep learning model is proposed, which integrates snake convolution and convolution attention mechanism to improve the accuracy and stability of lane recognition. In terms of model design, the snake convolutional module is first introduced and embedded in the backbone network to enhance feature extraction capability and better capture the shape information of the lane lines. Then, the convolutional attention module is introduced to dynamically adjust the weight of the feature map, so that the model can focus more on the lane line area and improve the segmentation accuracy of the lane line. To verify the performance of the proposed fusion model, we performed an experimental evaluation on the BDD100k dataset. The experimental results show that compared with the mainstream models Hybridenet, YOLOP and YOLOM, the accuracy rate increases by 0.8%, 1.4% and 1.3%, and the IoU increases by 1.4%, 5.1% and 2.8% respectively. This result proves the effectiveness and complementarity of snake convolution and convolution attention in lane line segmentation.

Key Words: snake convolution; convolution attention; lane line segmentation; instance segmentation

0 引言

车道线检测是计算机视觉领域的一项重要任务, 在自动驾驶、交通安全和驾驶辅助系统等方面具有广泛应用。车道线是道路上的重要标志之一, 它们提供了关键的空间参考, 能够帮助机器正确地控制车辆的行驶方向和位置。然而, 由于道路状况多样、光照条件变化及其他环境干扰

等, 车道线检测一直是一个具有挑战性的问题^[1]。

随着计算机视觉和深度学习技术的快速发展, 车道线检测取得了显著进展。传统基于图像处理的方法使用各种滤波器、边缘检测和几何变换等技术以提取车道线特征, 但在复杂的场景中往往效果有限^[2]。深度学习方法则通过构建端到端的神经网络模型, 能够从原始图像中学习更丰富的特征表示, 从而实现更准确和鲁棒的车道线检测^[3]。

收稿日期: 2024-08-21

扫描二维码阅读全文:



作者简介: 黄忠盛(1999-), 男, 成都信息工程大学计算机学院硕士研究生, 研究方向为图像处理; 魏敏(1978-), 男, 博士, 成都信息工程大学计算机学院教授, 研究方向为图像处理与目标检测技术、3D仿真技术和虚拟现实应用技术。本文通讯作者: 黄忠盛。

本文主要介绍了车道线检测的基本概念和挑战,探讨了传统方法和深度学习方法在车道线检测中的应用,并对当前研究进展进行了综述。此外,还讨论了车道线检测在自动驾驶和驾驶辅助系统中的重要性,并展望未来发展趋势。通过深入研究和分析,帮助人们更好地理解车道线检测问题及方法,为实现更安全、高效的道路行驶提供有力支持。

1 相关研究

近年来,传统车道线检测算法的主要流程包括图像预处理、特征提取、车道线模型拟合和直线跟踪。首先,从图像中去除一些噪声;其次,利用车道线的特征以提取车道线状区域;再次,使用各种方法对车道线模型进行拟合和跟踪。特征检测是车道线检测中的关键环节,其性能受场景复杂度、硬件约束等多方面因素影响^[2]。图像预处理包括感兴趣区域(ROI)的创建、用于提取车道线信息的图像增强以及非车道线细节的去除。ROI提取方法是通过选择图像的下半部分以有效地减少图像预处理部分中的冗余信息^[1]。此外,ROI的创建最大限度地减少了图像噪声,尽管它不能抵抗阴影或汽车对车道线的影响^[1]。在特征提取过程中,通过提取特定特征以检测车道线,如颜色、边缘、几何等^[2]。具体涉及几种技术,如逆透视图(IPM)/透视变换、滤波技术、基于边缘检测的技术、图像区域提取、形态学算子、基于邻域搜索的特征点、灰度、阈值和聚类。此外,还使用了异构运算符和滑动窗口以减少噪声影响,实现了车道线的快速提取。并且,使用线段检测器(LSD)和基于拟合的方法对车道线模型进行拟合,包括B样条、二次、多项式、抛物线、双曲线和最小二乘。Bresenham线投票空间(BLVS)^[4]、消失点、波形、几何建模、和谐搜索(HS)算法、对比度有限自适应直方图均衡(CLAHE)、随机样本一致性(RANSAC)、基于图的种子填充算法、直方图分析、模型预测控制(MPC)、一种基于区域的迭代种子方法、蚁群优化、嵌套融合和线性回归。Lucas-Kanade方法、Kanade Lucas Tomasi(KLT)^[5]和Lucas-Kanade光流法与车道线模型相匹配。同时,广泛用于跟踪道路车道线检测的算法是卡尔曼滤波器车道线分类和抛物线方程。跟踪通常被用作后处理步骤,以补偿照明波动^[6]。

基于深度学习的方法可分为3类:基于锚点检测^[7-9]、基于参数拟合^[10-12]和基于分割^[13-16]的方法。在基于锚点检测的方法中,UFLD(Unsupervised Feature Learning and Deep Learning)采用了一种简单快速的方法,通过预测车道的位置信息,并利用轻量级主干网络进行高速推理^[7-8]。尽管该方法速度快,但整体性能并不突出。相比之下,CondLaneNet提出了一种更为复杂的策略,它利用条件卷积和行锚点以定位车道线的起始点,再执行基于行锚点的车道检测^[9]。然而,在某些复杂场景下,这种方法可能会

遇到起始点识别困难的问题,从而导致性能相对较差。在基于参数拟合的方法中,PolyLaneNet利用多项式回归拟合车道线,提高了车道线检测效率^[10]。相较之下,LSTR考虑了道路结构和摄像机姿态,通过Transformer获取全局特征,并对车道形状进行建模^[11]。而BezierLane引入了Bezier曲线,进一步提升了检测精度^[12]。尽管这类方法的参数数量较少,但它们对预测参数的敏感性较高,即使微小误差也可能显著影响车道形状。虽然这些方法推理速度快,但要实现更高水平的性能仍然具有挑战性。在基于分割的方法中,LaneNet将车道线检测视为实例分割任务,通过端到端的网络架构,实现了对各种数量车道线的有效检测^[13]。SAD允许模型通过自蒸馏提高性能,无需额外的监督或标签,并且具有较少的参数数量和较高的推理速度,但往往难以达到更高的性能水平^[14]。SCNN提出了一种消息传递机制,以捕获车道的空间关系,从而显著提高了检测性能^[15]。然而,由于其速度较慢,因而并不适合实时应用。相较之下,RESA提出了实时特征聚合模块,有助于网络获取全局特征并提高性能^[16]。尽管这些方法在解决车道线检测任务中取得了一定进展,但它们仍然存在速度、性能和实时应用方面的限制,特别是在夜间场景下,这些限制更加明显。

在复杂的情况下,先进的车道线检测算法已超过了传统方法,但是它们也有局限性,比如只能发现一定数量的车道线,且检测时间成本非常高。随着注意力机制在自然语言处理领域取得良好成效,其很快被广泛应用于各领域,并在车道线检测中逐渐得到了较好应用^[17]。Zhang等^[17]提出了一个基于注意力策略的实时车道线识别系统以解决此问题,该网络包括1个提取车道线特征的编码器模块、2个预测车道线实例特征图的解码器模块、1个二进制解码器和1个可嵌入解码器。在编码器中引入生物学启发的注意力,提取了包含目标区域丰富信息的特征。为了学习语境信息,提出了卷积特征与注意特征之间的相关关系。该方法将上下文信息与解码器中的上采样特征相结合,以补偿所丢失的细节信息。二进制解码器将每个像素分配到两个类别:车道线或背景。采用可嵌入式解码器以获得不同的车道线,再将二进制解码器的输出作为可嵌入解码器的输入之一,可嵌入解码器主要指导车道线上精确像素点的产生。总体而言,注意力机制的引入能够有效地补偿传统方法中存在的不足,如只能检测有限数量的车道线、检测时间成本高等。同时,这也使得算法能够更好地捕捉目标区域的关键特征,并融合上下文信息,从而提高了车道线检测准确性和效率。

2 模型设计

2.1 YOLOv8网络架构

YOLOv8作为目标检测算法YOLO系列的新版本,在

保持YOLO标志性的实时性和准确性特质的同时,针对实例分割任务进行了专门优化。它沿用了端到端的单阶段网络前向传播设计理念,能够在一次前向传播过程中同时预测物体的精确分割掩模和类别概率。在YOLOv8的设计中,针对基于锚框的目标区域预测方式进行改进,选择了Anchor-Free策略,从而减少了锚框对模型性能的限制,增强了模型对不同尺寸和形态实例的泛化能力和分割精度。通过这些优化设计,YOLOv8在实例分割任务中表现优异,尤其适用于车道线分割任务。

YOLOv8整体结构大致分为4个部分,即输入端、骨干网络、颈部网络、预测头网络。其整体架构如图1所示。在输入端,YOLOv8采用自适应图像缩放技术,根据输入图像的尺寸和内容,智能地调整到最优分辨率,以平衡速度和精度,调整后的图像被传递到骨干网络以进行特征提取。

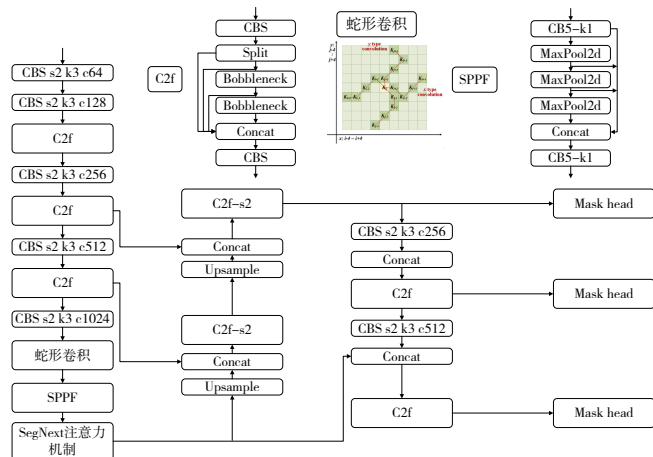


Fig. 1 Overall algorithm steps

图1 算法总体步骤

骨干网络由5个关键模块组成: CBS(卷积—批归一化—SiLU激活)、C2f(改进的CSP结构)、蛇形卷积、SegNext注意力机制和SPPF(空间金字塔池化和融合)。CBS模块通过巧妙的组合,增强了模型稳定性,加速了训练收敛,可有效防止梯度消失问题。C2f模块在经典的CSP结构基础上,引入额外的跳跃连接和Split操作,让梯度流更加丰富,促进了特征的充分融合。为了引入蛇形卷积,需要将CBS后面的卷积层替换成蛇形卷积,以提取到更多的车道线信息。SPPF模块利用不同尺度的池化和卷积,自适应地聚合多尺度特征信息,全面提升模型的特征表达能力。最后,在SPPF模块后面添加卷积注意力机制,利用多尺度卷积特征通过简单元素乘法以唤起空间注意力,可以帮助模型更好地利用全局上下文信息。

在颈部网络中,YOLOv8创新性地结合了PAN(金字塔注意力网络)和FPN(特征金字塔网络)两种结构。通过自上而下和自下而上的跨层连接,实现了不同语义级别特征的充分交互和融合,使得最终的特征图更加丰富。

在YOLOv8的分割任务中,预测头通常包含多个卷积

层,这些卷积层负责从融合的特征图中提取空间信息,以预测每个像素的类别。预测头的输出是一个与输入图像尺寸相同的分割图,其中每个像素的值代表其属于某一特定类别的概率。

2.2 蛇形卷积

对于车道线这种管状类结构的特征精确提取,目前仍然具有挑战性。传统的卷积神经网络通过使用多种卷积核直接对输入图片进行卷积处理,从而能够提取图像中低级到高级的特征。这种方法通过增加卷积核的数量以增强特征提取能力,但其针对性相对较弱。特别是对于那些稀疏且具有细长外观的对象,如车道线,传统卷积神经网络在特征提取上的效果往往不尽如人意。究其原因在于,这种方法未能充分考虑特征图中行和列的空间关系以及通道之间的复杂特征关联。

针对车道线这样的细长和弯曲特征,蛇形卷积提供了一个有效的解决方案。蛇形卷积是一种特殊的卷积技术,其设计是用来更好地适应图像中的线形结构,如图2所示。这种卷积方法通过模拟蛇形路径在行和列之间交错前进,从而更精确地捕捉和处理车道线的几何特性。

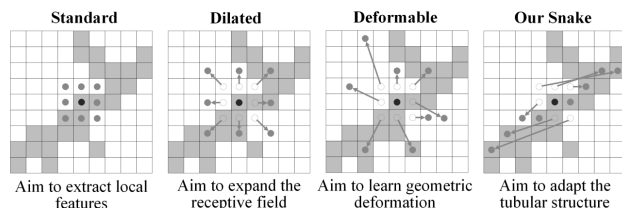


Fig. 2 Snake convolution

图2 蛇形卷积

蛇形动态卷积通过在卷积核设计中加入连续性约束,提高了处理过程的自适应性和连续性。在这种设计中,卷积操作的每一步都以上一步的位置作为参照,从而自由但连续地选择接下来的移动方向。这样既保留了结构的自适应灵活性,又确保了特征感知的顺畅连贯。

为了让卷积核能更有效地适应目标的复杂几何结构,动态蛇形卷积引入了一个可变形的偏移量 Δ 。但是,如果完全让模型自主学习这些偏移量,感知场可能会脱离实际目标,尤其是在处理细长的结构时。因此,动态蛇形卷积采用了一种迭代策略,这种策略可以逐步确定每个目标的下一个观察点,从而保持注意力的连续性,并避免因较大的变形偏移而过度扩展感知区域。为了提取管状结构的局部特征,给定标准的2D卷积坐标 K ,中心坐标为 $K_i = (x_i, y_i)$ 。一个 3×3 的卷积核 K 表示为如式(1)所示。

$$k = \{(x-1, y-1), (x-1, y), \dots, (x+1, y+1)\} \quad (1)$$

在蛇形卷积中,其将标准卷积算子在 x 轴和 y 轴方向上都拉直。蛇形卷积考虑大小为9的卷积核,并以 x 轴方向为例,每个网格在 K 中的具体位置表示为: $k_i \pm c = (x_i \pm c, y_i \pm c)$,其中 $c = \{0, 1, 2, 3, 4\}$ 表示到中心网格的水平距离。卷积核 K 中每个网格位置 $k_i \pm c$ 的选择是一个累

加过程。从中心位置 k_i 开始, 远离中心网格的位置取决于前一网格的位置: 与 k_i 相比, 用偏移量 $\Delta = \{\in [-1, 1]\}$ 以增加 $\delta\delta + 1$ 。因此, 偏移量为 Σ , 从而确保卷积核符合线性形态结构。式(1)在 x 轴方向变为如式(2)所示。

$$K_{i \pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \sum_{i=c}^{i+c} \Delta y), \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \sum_{i=c}^{i-c} \Delta y), \end{cases} \quad (2)$$

并且, y 轴方向上的式(2)变为如式(3)所示。

$$K_{j \pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = (x_j + \sum_{j=c}^{j+c} \Delta x, y_j + c), \\ (x_{j-c}, y_{j-c}) = (x_j + \sum_{j=c}^{j-c} \Delta x, y_j - c), \end{cases} \quad (3)$$

其中, K 表示式(2)和式(3)的分数位置, K' 枚举所有整数空间位置, B 是双线性内插核, 它被分成两个一维核, 如式(4)所示。

$$B(K, K') = b(K_x, K'_x) \cdot b(K_y, K'_y) \quad (4)$$

如图 3 所示, 由于二维 (x 轴、 y 轴) 的变化蛇形卷积在变形过程中覆盖了 9×9 的范围。蛇形卷积的设计是为了更好地适应细长筒体结构的动态结构, 从而更好地感知关键特征。

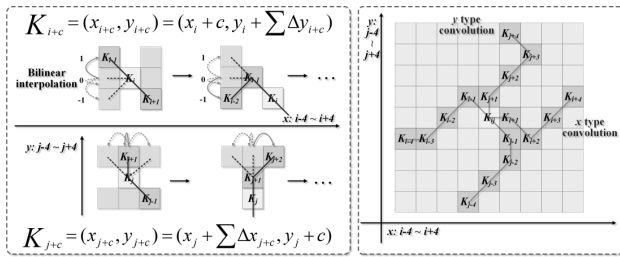


Fig. 3 The working principle of dynamic snake convolution

图 3 蛇形动态卷积工作原理

2.3 segNext 注意力模块

SegNext 注意力模块如图 4 所示。在 SegNext_Attention 中, 通过在编码器和解码器之间引入注意力机制, 模型得以更有效地处理全局上下文信息。这种注意力机制主要通过学习每个像素的注意力权重加以实现, 从而使得模型能够集中关注目标区域, 同时排除背景干扰。SegNext_Attention 的结构主要包括以下核心部分:

(1) 编码器。此部分利用卷积神经网络 (CNN) 以捕捉图像的特征。编码器包括若干卷积层和池化层, 不断缩小特征图的尺寸, 同时增加其通道数, 以深入提取图像特征。

(2) 注意力机制。在编码器生成的特征图上施加注意力机制, 生成与原图像尺寸相匹配的注意力权重图, 该权重图突出显示了图像中各像素的重要性。

(3) 解码器。解码器部分通过上采样和卷积操作将编码器输出的特征图转换为像素级的分割图。在此过程中, 逐步恢复图像尺寸并减少通道数, 以形成最终分割结果。

(4) 损失函数。模型使用损失函数以评估预测分割图与实际标签之间的差异, 以训练优化模型。常见的损失函数有交叉熵损失和 Dice 损失等。

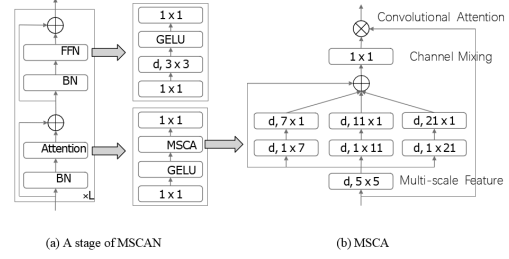


Fig. 4 segNext attention module

图 4 segNext 注意力模块

总体而言, SegNext_Attention 的主要优势在于其能够有效利用全局上下文信息, 并精准关注关键区域, 从而显著提升图像分割准确性和效率。

3 损失函数

本文的损失函数包含分割损失、IoU 损失和 DFL 损失, 通过计算分割损失以衡量预测结果与真实标签之间的差异和相似度, 以指导模型优化。使用 IoU 损失以衡量预测框与真实框之间的重叠程度, 用于评估目标检测模型的性能和优化过程中的损失计算。并且, 使用 DFL 损失以处理难易样本的不平衡问题, 它通过调整正负样本的权重和难易样本的焦点以改善模型对于难例的学习能力。

3.1 分割损失

在车道线分割任务中, 数据集的正负样本差距通常很大, 因为大部分像素都不属于车道线。因此, 使用 Focal Loss 作为损失函数是一个很好的选择, 因为它可以帮助模型更加关注那些难以分类的正样本, 同时减少对大量易分类负样本的关注。Focal Loss 是基于二分类交叉熵 CE, 它是一个动态缩放的交叉熵损失, 通过一个动态缩放因子, 可以动态降低训练过程中易区分样本的权重, 从而将重心快速聚焦在那些难区分的样本上。其运算如式(5)所示。

$$CE(p, y) = \begin{cases} -\log(p) & \text{if } y = 1 \\ -\log(1-p) & \text{otherwise} \end{cases} \quad (5)$$

其中, y 的取值为 1 和 -1, 分别代表前景和背景。 p 的取值范围为 0~1, 是模型预测属于前景的概率。定义一个关于 P 的函数, 如式(6)所示。

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1-p & \text{otherwise} \end{cases} \quad (6)$$

通过上述正负样本和难易样本平衡, 可得到最终的 Focal Loss 形式, 如式(7)所示。

$$FL(p_t) = -\alpha_t (1-p_t)^{\gamma} \log(p_t) \quad (7)$$

3.2 IoU 损失

在图像分割中, 交并比 (Intersection over Union, IoU) 是一个常用的度量标准, 用于评估预测区域与真实区域之间的重叠程度。IoU 也常被用作损失函数, 尤其是在需要精确边界对齐的任务中, 如图像分割和目标定位。

随着检测器技术的快速发展,边框回归技术也取得了显著进步。但现有基于 IoU 的边框回归方法主要聚焦于通过增加新的损失项以加速收敛,却忽略了 IoU 损失本身的局限性。这些方法在实际应用中无法自适应不同的检测器和任务,缺乏泛化性。本文分析了 BBR 模式,并提出了通过使用不同尺度的辅助边框以计算损失,从而区分不同的回归样本,有效加速边框回归过程。对高 IoU 样本使用小尺寸辅助边框,对低 IoU 样本使用大尺寸辅助边框。此外,本文引入了 Inner-ShapeLoss,该方法可以通过关注边界框本身的形状和尺度以计算损失,从而使边界框回归更加准确,其计算公式如式(8)所示。

$$L_{\text{Shape-IoU}} = 1 - \text{IoU} + \text{distance}^{\text{shape}} + 0.5 \times \Omega^{\text{shape}} \quad (8)$$

4 实验与结果分析

4.1 数据集

本文实验采用 BDD100K 数据集,它是一个大规模自动驾驶数据集,包括 100 000 个样本和多任务注释。除庞大的数量外,该数据集的重要性还体现在它的多面性,包括不同的地理位置、环境背景和天气条件,这些优点保证了训练后模型的鲁棒性和通用性。实验中,车道线标记共 100 000 张,其中包括 70 000 张训练集、10 000 张验证集和 20 000 张测试集。

4.2 实验环境

实验操作系统为 Windows10, Intel(R) Core(TM) i7-10700 CPU@2.90 GHz 处理器,显卡型号为 Nvidia RTX A4000(16 GB),使用 PyTorch 深度学习框架搭建模型。

4.3 评价指标

为实现更加全面和准确的预测效果评价,采用 Accuracy 和 IoU 作为评估指标。然而,由于车道线分割中背景和前景像素数量不平衡问题,故采用更有意义的平衡精度进行度量评估。传统的 Accuracy 会使评估结果偏向于样本数量庞大的类别,而平衡精度通过分别考量每个类别的准确性表现,能够为各类别的分类效果提供一个更为公平、客观的度量。具体公式如式(9)一式(11)所示。

$$\text{LineAccuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad (9)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (11)$$

在分割任务中, IoU 是一个非常重要的指标,因为它直观地反映了模型预测精度的高低。高 IoU 值表示模型能够非常精确地覆盖目标区域,而忽略背景或其他非目标区域,它衡量的是预测分割掩码与真实分割掩码的交集与并集之比。

IoU 的计算公式如式(12)所示。

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (12)$$

4.4 实验比较

将本文方法与 YOLOv8-seg、YOLOM、HybridNet 和 YOLOP 等其他模型在 BDD100K 数据集上的表现进行比较。BDD100K 数据集含有丰富的车道线标注,适合用来评估各种车道线检测和分割算法的性能。所有模型都使用固定的 70 000 张训练集、10 000 张验证集进行训练。训练时的 batch_size 设置为 16,使用 SGD 优化器以 0.01 的学习率训练 100 轮,训练时间为 23 h。



Fig. 5 Comparison of prediction performance of different models on the Moving MNIST dataset

图5 不同模型在 Moving MNIST 数据集上的预测效果比较

在对 YOLOv8-seg、YOLOM、HybridNet 和 YOLOP 4 种模型进行主观视觉比较的基础上,通过实验结果发现, YOLOv8-seg 在车道线分割任务上表现最佳,提供了极高的清晰度和连续性,即使在复杂的光照条件或变化的路况下也能维持优异的分割效果。HybridNet 表现良好,尤其适合常见道路环境,但在处理复杂场景时稍显不足。而 YOLOM 和 YOLOP 作为一款多任务模型,虽然具备较好的应用潜力,但在专门的车道线识别任务中稍逊一筹,特别是在光照条件恶劣的场景中。

在主观预测效果上, YOLOv8 和 HybridNet 差距并不明显。单独引入注意力模块后,各项评价指标数值都有所提升,单独引入蛇形卷积模块后也有一定效果,在结合使用 Inner-ShapeLoss 后,得到的本文模型达到更优效果,如表 1 所示。在本文模型中,结合蛇形卷积与分割注意力机制,并使用 Inner-ShapeLoss 进行训练后, IoU 提升至 31.6%,模型预测准确率提升至 86.2%,有效缩小了预测图像与真实图像之间的差异。

4.5 消融实验

为了验证本文方法的有效性,在 BDD 100K 数据集上进行消融实验,设计了 4 种实验以评估 SegNext 注意力机制

Table 1 Prediction metrics of different models on the BDD 100K dataset

表 1 不同模型在 BDD 100K 数据集上的预测指标		
模型名称	Accuracy/%	IoU/%
HybrideNet	85.4	30.2
YOLOP	84.8	26.5
YOLOM	84.9	28.8
Yolov8+Dysnake+SegNext	86.2	31.6

和蛇形卷积模块对模型的影响。

4.5.1 实验配置

实验配置如下:①完整模型,包括蛇形卷积和SegNext注意力机制的完整架构;②无蛇形卷积,在此配置中,蛇形卷积结构被常规卷积层替换,以探索蛇形卷积对性能的贡献;③无SegNext注意力机制,移除SegNext注意力模块,用以评估其对整体模型性能的影响;④无蛇形卷积和SegNext注意力机制,两种技术模块均被移除,作为基线模型进行比较。

4.5.2 实验结果

消融实验结果如表 2 所示。在 IoU 指标上,完整模型表现出色,达到了 30.6%,明显高于其他模型配置。移除蛇形卷积、SegNext 注意力模型 IoU 分别降低了约 2.0%、3.2% 外,基线模型的 IoU 最低,为 26.3%。

Table 2 Ablation experiment results

表 2 消融实验结果		
模型名称	Accuracy/%	IoU/%
Yolov8	80.5	26.3
Yolov8+Dysnake	83.7	28.6
Yolov8+SegNext	84.6	27.4
Yolov8+Dysnake+SegNext	86.2	30.6

实验结果表明,蛇形卷积和SegNext注意力机制对YOLOv8-seg模型的整体性能有着显著的正面影响。蛇形卷积通过其独特的结构增强了模型对局部特征的捕捉能力,而SegNext注意力机制通过关注重要的特征区域进一步提升了模型判别力。完整模型的优越性能验证了这些组件在实现高精度分割中的关键作用。

5 结语

本文针对车道线分割精度和稳定性问题,采用了一种融合蛇形卷积与卷积注意力的车道线分割方法,并在BDD100K数据集上进行了实验评估。通过引入蛇形卷积和SegNext注意力模块,增强了模型对车道线特征的捕捉能力,使其在复杂场景中的识别更为精准和连续。此外,采用的Inner-ShapeLoss损失函数进一步优化了模型对车道线形状的理解和重建能力。实验结果表明,这些改进显著提高了YOLOv8-seg在车道线分割任务中的性能,尤其是在低光照条件下的表现,展示了其在实际应用中的潜力和可行性。未来,将继续探索更多创新技术和方法,以进一步提升车道线分割准确性和鲁棒性,为自动驾驶技术发展提供支持。

展提供支持。

参考文献:

[1] MUNIR F, AZAM S, JEON M, et al. LDNet: end-to-end lane marking detection approach using a dynamic vision sensor[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 9318-9334.

[2] ZHANG R H, PENG J T, GOU W T, et al. A robust and real-time lane detection method in low-light scenarios to advanced driver assistance systems[J]. Expert Systems with Applications, 2024, 256: 124923.

[3] WANG Z, REN W, QIU Q. LaneNet: real-time lane detection networks for autonomous drivin[DB/OL]. <https://arxiv.org/abs/1807.01726>, 2018.

[4] YANG B. Exploration of a new linear algorithm[J]. Computers and Modernization, 2003(5): 45-48.

阳波. 一种新直线算法的探索[J]. 计算机与现代化, 2003(5): 45-48.

[5] HAN Y, KIM C, JANG Y, et al. Parametric analysis of KLT algorithm in autonomous driving[C]//Proceedinds of the 20th International Conference on Control, Automation and Systems, 2020: 184-189.

[6] LEE C, MOON J H. Robust lane detection and tracking for real-time applications[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(12): 4043-4048.

[7] HU H D, LIU G R, WANG L L, et al. A lane detection algorithm based on vanishing point and color filter[J]. Journal of Chongqing Technology and Business University(Natural Science Edition), 2023, 40(5): 25-33.

胡昊东, 刘贵如, 王陆林, 等. 一种基于消失点和颜色过滤器的车道线检测算法[J]. 重庆工商大学学报(自然科学版), 2023, 40(5): 25-33.

[8] QIN Z, WANG H, LI X. Ultra fast structure-aware deep lane detection[C]//Proceedinds of the 16th European Conference on Computer Vision, 2020: 276-291.

[9] QIN Z, ZHANG P, LI X. Ultra fast deep lane detection with hybrid anchor driven ordinal classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(5): 2555-2568.

[10] LIU L, CHEN X, ZHU S, et al. CondLaneNet: a top-to-down lane detection framework based on conditional convolution[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 3773-3782.

[11] TABELINI L, BERRIEL R, PAIXAO T M, et al. PolyLaneNet: lane estimation via deep polynomial regression[C]//Proceedings of the 25th International Conference on Pattern Recognition, 2021: 6150-6156.

[12] LIU R, YUAN Z, LIU T, et al. End-to-end lane shape prediction with transformers[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2021: 3694-3702.

[13] FENG Z, GUO S, TAN X, et al. Rethinking efficient lane detection via curve modeling[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 17062-17070.

[14] NEVEN D, DE BRABANDERE B, GEORGOULIS S, et al. Towards end-to-end lane detection: an instance segmentation approach[C]//Proceedings of the IEEE Intelligent Vehicles Symposium, 2018: 286-291.

[15] HOU Y, MA Z, LIU C, et al. Learning lightweight lane detection CNNs by self attention distillation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1013-1021.

[16] PAN X, SHI J, LUO P, et al. Spatial as deep: spatial CNN for traffic scene understanding[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2018: 7276-7283.

[17] ZHANG L, JIANG F, KONG B, et al. Real-time lane detection by using biologically inspired attention mechanism to learn contextual information[J]. Cognitive Computation, 2021, 13(9): 1333-1344.

(责任编辑:孙 娟)