

# 基于起始点引导多维特征细化的复杂场景 车道线检测

朱松豪, 谭少涵

(南京邮电大学自动化学院、人工智能学院, 江苏南京 210023)

**摘要:** 车道线检测作为智能驾驶系统的基石,在车道保持、自适应巡航等辅助驾驶过程中发挥着至关重要的作用。鉴于车道线检测在提升道路安全、促进智能驾驶及智慧交通发展中的关键作用,车道线检测技术的研究具有深远的学术价值与应用前景。然而,由于车道线的多样性、路况的复杂性及天气的多样性等因素,车道线检测面临诸多挑战。研究提出一种基于起始点引导的多维特征细化的复杂场景车道线检测方法,利用全局特征优化模块和车道线感知聚合模块对特征进行优化,捕捉足够的上下文信息,利用起始点坐标预测模块预测车道线的起始点坐标,生成高质量的锚点。为更好适应二维车道线检测,研究提出更为通用的惩罚车道线交并比损失函数评估车道线的预测结果。与目前车道线检测算法中精度最高的 CLRNet-DLA34 方法相比,所提方法在 CULane 数据集和 CurveLanes 数据集的  $F_1@50$  分别高出了 0.62% 和 0.73%, 达到 81.39% 和 86.83%。实验结果表明,所提方法在复杂场景车道线检测任务中取得了良好的检测效果,且在现有方法中具有很强的竞争力。

**关键词:** 车道线检测;起始点坐标预测;全局特征细化;车道线感知聚合;惩罚车道线交并比

**基金项目:** 国家自然科学基金(No.62001247)

**中图分类号:** TP391.41

**文献标识码:** A

**文章编号:** 0372-2112(XXXX)XX-0001-10

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20240816

## Multi dimension feature refinement based on starting point guidance for lane line detection in complex scene.

ZHU Song-hao, TAN Shao-han

(College of Automation and Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China)

**Abstract:** Lane detection, as the cornerstone of intelligent driving systems, plays a crucial role in assisting driving processes such as lane keeping and adaptive cruise control. Given the crucial role of lane detection in improving road safety, promoting intelligent driving and intelligent transportation development, the research on lane detection technology has profound academic value and application prospects. However, due to the diversity of lane categories, road conditions and weather environments, as well as the different aspect ratios of lane lines, lane detection algorithms face many challenges. This paper proposes a multi-dimensional feature refinement method for complex scene lane detection based on start point guidance. Firstly, a global feature optimization module is utilized to enhance the global feature representation capability and a lane line perception aggregation module is utilized to enhance the correlation of local features, which helps to improve the semantic understanding ability and prediction accuracy of the model. Secondly, a starting point coordinate prediction module is utilized to predict the starting point coordinates of lane line to generate flexible anchors under various complex scenarios. Finally, a more general penalty lane intersection to union ratio is selected as the loss function to represent lane lines with variable virtual widths, which helps to improve the detection accuracy of the model. Compared with the CLRNet-DLA34 method, which has the highest accuracy in current lane detection algorithms, the method proposed in this paper improves the detection accuracy in terms of  $F_1@50$  on the CULane and CurveLanes datasets by 0.62% and 0.73% respectively, reaching 81.39% and 86.83%. The experimental results demonstrate that the proposed method achieves good detection performance in complex scene lane detection tasks and has strong competitiveness among existing methods. Experimental results demonstrate that the proposed method performs well in lane detection tasks in complex scenes and has strong competitiveness among exist-

ing methods.

**Key words:** lane line detection; start point coordinate prediction; global feature refinement; lane aware gather; penalty laneIoU

**Foundation Item(s):** National Natural Science Foundation of China(No.62001247)

## 1 引言

为确保行驶过程中的安全性,智能驾驶系统须准确沿车道线导航车辆,这意味着准确感知车道线至关重要.车道线检测旨在分析车载摄像头捕获的二维图像,并准确提取每条车道线信息.由于细长的形状及复杂的路况,车道线检测面临重大困难.同时,由于车道线检测的快速性,这对算法的轻量化提出了很高要求.

传统的车道线检测方法首先利用人为设计的运算符提取特征,然后使用霍夫变换和随机采样一致性等策略对车道线进行建模<sup>[1-3]</sup>.卷积神经网络的出现加速了车道线检测的快速发展,提高算法的准确性和鲁棒性.然而,尽管现有的深度学习方法<sup>[4-6]</sup>在性能和速度方面取得了显著进展,但在复杂场景下仍有优化空间.

如图1(a)所示,某些车道线完全被遮挡,需要检测网络对上下文信息的深刻理解.此外,由于车道线的细长性和连续性,现有深度学习方法很难捕捉车道线完整的几何特征.首先,由于卷积网络性能无法满足Transformer编码器对大感受野的要求<sup>[7]</sup>,因此,为优化全局特征表示,并扩大感受野,本文提出一个基于卷积注意力的多尺度特征优化方法.其次,车道线局部特征表征会影响车道线检测性能.如图1(b)所示,使用现有深度学习方法获得的车道线并非最优,因为这些方法在固定网格内进行检测,可能无法与不规则形状的车道线对齐.因此,本文提出一个具有局部信息聚合功能的车道线感知聚合模块,利用可变形卷积灵活提取特征表征,增强局部特征的相关性.最后,将经过优化的全局特征与增强后的局部特征进行聚合,捕获更多的全局上下文信息.

基于锚的车道线检测方法检测精度高度依赖锚生成的质量.现有方法采用从图像3个边缘预设锚的策略,无法处理所有场景,这是因为车道线很可能由于车身遮挡、模糊不清、拍摄角度等原因,起始点并不处于图像边缘.如果能对这些特殊的车道线起始点进行捕获、并以此为依据灵活地生成锚,对于提高锚的质量将会很有帮助.因此,本文借鉴目标检测思路,利用起始点坐标预测模块,预测车道线的起始点坐标,生成更灵活可靠的锚.

由于LineIoU使用固定的虚拟宽度表示不同斜率的车道线不完全合理,因为随着车道线越来越水平,其

在二维图像上的水平宽度也越来越大,此时使用恒定的虚拟宽度必然导致精度损失.此外,该方法设置的车道线交并比的值,在预测车道线与实际车道线重合与分离时的分布差异极大,不符合实际情况.因此,本文提出Penalty LaneIoU,使得虚拟宽度随车道线倾斜角变化而变化,同时加入可变虚拟宽度的惩罚项,使交并比分布更均匀.



现有深度学习方法

本文方法

(a) 遮挡场景



现有深度学习方法

本文方法

(b) 弯道场景

图1 复杂场景下的车道线检测示意图

其中:橙线表示真实的车道线,绿线表示预测的车道线.这2组对比实验得到以下2个结论:(1)上下文信息有助于检测完全遮挡的车道线;(2)局部特征的细化有助于获得更为精确的检测结果.

本文的主要贡献可以总结如下:

(1)提出了全局特征优化方法,以最小的计算复杂度增强全局特征表示能力;同时,提出车道线感知聚合模块,增强局部特征的相关性,有助于提高模型的语义能力和预测性能.

(2)提出了起始点坐标预测方法,用于预测车道线的起始点坐标,生成更灵活的锚,有助于处理各种复杂场景.

(3)提出了惩罚车道线交并比损失,使用可变虚拟宽度表示车道线,且设置更为均匀的交并比,有助于提高模型的检测精度.

(4)3个公开数据集上的实验结果表明,所提方法

取得了非常有竞争力的结果,且在CULane数据集的特定场景下取得了最佳性能。

## 2 相关工作

车道线检测旨在准确捕捉车道线边界的形状,并对其进行区分。依据不同的车道线表示方法,可将车道线检测方法分为不同类别。

### 2.1 基于分割的方法

基于分割的方法将车道线检测任务视为像素级预测任务。空间卷积神经网络<sup>[8]</sup>(Spatial Convolutional Neural Network, SCNN)将每条车道线视为一个类别,将车道线检测转化为多类别分割问题,并引入一种在行像素和列像素间传输空间信息的机制。循环特征移位聚合器<sup>[9]</sup>(Recurrent Feature-Shift Aggregator, RESA)利用具有车道线形状先验信息的循环聚合器,提高车道线检测性能。TwinLiteNet<sup>[10]</sup>双精简网(TwinLiteNet)提出一种具有深度可分离卷积和双注意力网络的轻量级模型,用于分割车道线,且在Jetson Xavier NX上实现了60每s帧数(Frame Per Second, FPS)。但是基于分割的方法的缺点在于需要大量的后处理,且泛化能力往往不佳。

### 2.2 基于关键点的方法

基于关键点的方法将车道线检测任务视为关键点检测任务。这些方法首先要预测所有可能的车道线潜在关键点,然后将这些关键点分配至各自车道上。

点实例网络<sup>[11]</sup>(Point Instance Network, PINet)首先利用堆叠沙漏网络预测关键点的位置及相应的特征嵌入,然后利用这些特征嵌入间的相似性对车道线进行分组。聚焦局部车道线方法<sup>[12]</sup>(Focus on Local Lane, FOLOLane)通过识别关键点,实现车道线检测。全局相关性网络<sup>[13]</sup>(Global Association Network, GANet)通过回归确定检测的关键点与相应车道线起始点间的偏移量。基于关键点方法的后处理非常耗时,准确性严重依赖于输入图像分辨率,难以同时实现低延迟和高准确性。

### 2.3 基于曲线的方法

基于曲线方法将车道线表示为一组曲线参数,并通过回归这些参数实现车道线检测。PolyLaneNet<sup>[14]</sup>是第1个使用深度网络回归车道线曲线方程的方法,具有很高的准确性。作为一种端到端的车道线检测方法,BezierLaneNet<sup>[15]</sup>利用参数化的贝塞尔曲线表示车道线。虽然曲线方法易于实现车道线连续性的建模和优化,但其对曲线参数预测具有一定的敏感性,因为高阶系数直接影响车道线形状的预测。为解决这个问题,密集混合建议调制网络<sup>[16]</sup>(Dense Hybrid Proposal Modulation, DHPM)引入密集混合建议调制策略,旨在生成高质量车道线检测结果。虽然基于曲线的方法需要较少的回归参数,导致其推理速度较快,但其对预测参数敏感,难以实现高性能。

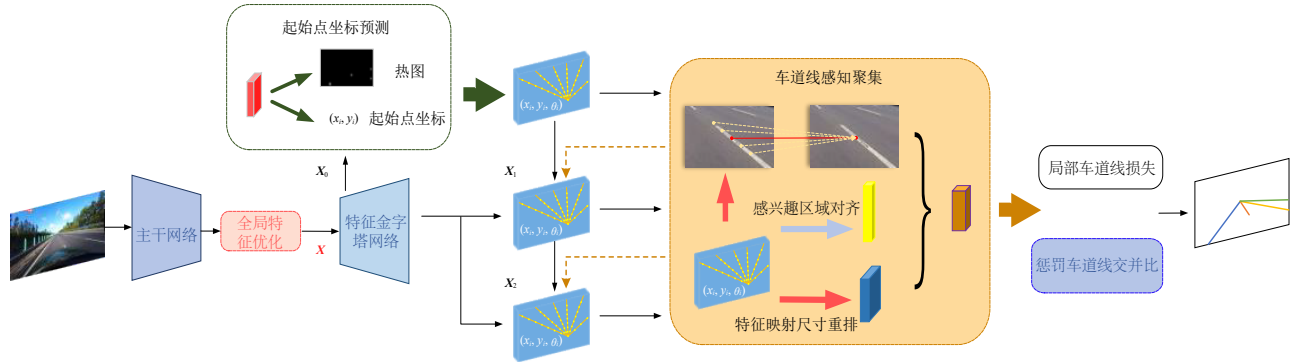


图2 所提方法SMFRNet框架

### 2.4 基于锚的方法

基于锚的方法将图像划分为规则切片,并通过在每个切片进行回归完成车道检测任务。超快速车道检测第2版<sup>[17]</sup>(Ultra-Fast Lane Detection Version 2, UFLDV2)利用车道线的行/列信息,实现车道线检测。改进的跨层细化网络<sup>[18]</sup>(Improved Cross Layer Refinement Network, CLRerNet)提出一种五折交叉验证策略,并将LaneIoU集成至行锚车道线检测框架。锚点分解网络<sup>[19]</sup>(Anchor Decomposition Network, ADNet)将锚分解为起始点和其对应方向,获得较好性能。基于

行锚点的方法能更有效利用形状先验知识,在保持高速推理的同时实现更精确的预测。

## 3 方法

由于行锚方法的高效性能,本文将作为基线方法。图2给出所提方法的整体框架,首先利用提出的全局特征优化模块(Global Feature Refinement, GFR),扩展主干网络的感受野;接着利用提出的起始点坐标预测模块(Start Point Coordinate Prediction, SPCP),预测起始点坐标;然后利用提出的车道线感知聚合模块(Lane

Aware Gather, LAG), 在整合全局特征的同时, 细化局部特征; 最后利用提出的惩罚车道线交并比损失函数 (Penalty LaneIoU, PLIoU), 优化车道线特征.

### 3.1 全局特征优化

由于天气、光照和遮挡等原因, 车道线并不像理想情况下那样清晰. 因此, 全局上下文信息的提取能力至关重要. 受 SegNext<sup>[20]</sup> 下一代语义分割网络 (Next Semantic Segmentation Network, SegNext) 和 ConvNext V2<sup>[21]</sup> 第二代卷积神经网络 (Next Convnet, ConvNext V2) 的启发, 本文设计了如图 3(b) 所示的全局特征优化模块解决这一问题, 并将该模块放置于骨干网络的最后一层, 以扩展感受野并获得有区分性的全局特征.

全局特征优化过程如下所述. 首先, 利用如图 3(a) 所示的多尺度特征注意力子模块, 实现基于卷积的注意力机制: 第 1 步, 使用  $3 \times 3$  卷积核进行特征提取; 第 2 步, 使用大小分别为  $7 \times 7$ 、 $11 \times 11$  和  $21 \times 21$  的卷积核提取不同维度的特征; 第 3 步, 将得到的多维特征与原始特征进行逐元素相乘, 实现卷积注意力. 然后, 使用全局响应规范化子模块对线性化特征进行归一化处理, 有助于特征多样化.

全局特征优化过程可用下式表示

$$\begin{cases} L_{ATT} = L \otimes \sum_{i=0}^3 \text{Conv}_{1 \times 1}(\text{Bra}_i(\text{Conv}_{3 \times 3}(L))) \\ X = L + \text{Conv}_{1 \times 1}(\text{GG}(\text{Conv}_{1 \times 1}(\text{BN}(L_{ATT})))) \end{cases} \quad (1)$$

其中:  $L$  表示从骨干网络最后 1 层提取的特征; BN 表示批归一化层; GG 表示 GELU 高斯误差线性单元 (Gaussian Error Linear Unit, GELU) 激活函数和全局响应规范化的组合. 此外,  $\text{Bra}_i$  表示多尺度特征融合, 其中  $\text{Bra}_0$  不执行任何操作,  $\text{Bra}_1$ 、 $\text{Bra}_2$  和  $\text{Bra}_3$  分别表示大小为  $7 \times 7$ 、 $11 \times 11$  和  $21 \times 21$  的卷积核.

### 3.2 起始点坐标预测

包括 LaneATT 车道线注意力 (Lane Attention, Lane ATT) 和 CLNet 跨层细化网络 (Cross Layer Refinement Network, CLNet) 等在内的许多锚方法, 将锚的起始点设置在图像的底部和侧边. 尽管这类方法检测性能尚可, 但它们严重依赖于锚质量. 真实场景中, 锚的起始点并不总是位于图像边缘, 这可能会影响模型性能. 因此, 受文献[22]启发, 本文为每条车道线预测起始点坐标, 并将它们作为构建新起始点的依据.

此外,  $\oplus$  和  $\otimes$  分别表示矩阵相加和元素相乘. 由于在某些复杂情况下识别车道线起始点可能存在困难, 本文将提出的起始点坐标预测方法与 CLNet 中的锚生成方法相结合, 改善生成锚的质量并增强模型的鲁棒性.

起始点坐标预测的具体过程如下所述. 首先, 利用

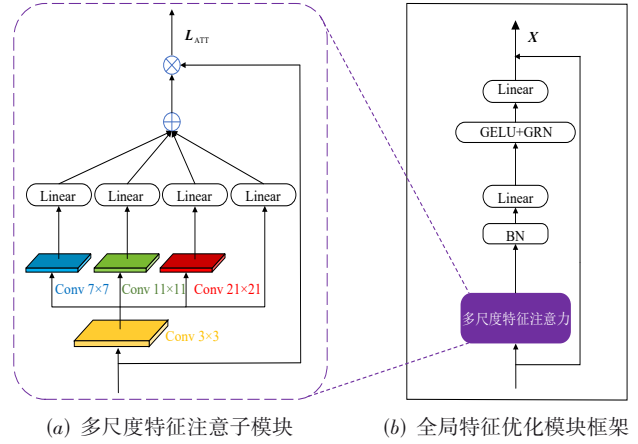


图 3 全局特征优化模块框架

重映射热图获取下采样特征图中可能的起始点, 并利用聚焦损失限制建议热图:

$$\text{IL}_{pr} = \frac{-1}{\text{num}_s} \sum_{x=1}^w \sum_{y=1}^h \begin{cases} (1-p_{\text{start}})^\alpha \log(p_{\text{start}}), & G_{\text{start}} = 1 \\ (1-G_{\text{start}})^\beta (p_{\text{start}})^\alpha \log(1-p_{\text{start}}), & \text{otherwise} \end{cases} \quad (2)$$

其中:  $\text{num}_s$  表示输入图像中建议起始点的数量;  $p_{\text{start}}$  表示建议热图中每个起始点坐标的预测值;  $G_{\text{start}}$  表示预测起始点坐标处的真值.

由于热图是通过下采样获得的, 因此, 可能会降低从热图获得起始点坐标的准确性. 为解决这一问题, 有必要预测起始点位置的偏移量并微调起始点位置, 这可通过下式实现:

$$\begin{cases} \mathbf{O}_{xy} = \left( \frac{x_{gt}^i}{n} - \left\lfloor \frac{x_{gt}^i}{n} \right\rfloor, \frac{y_{gt}^i}{n} - \left\lfloor \frac{y_{gt}^i}{n} \right\rfloor \right) \\ \text{IL}_{\text{offset}} = \frac{\sum_{i=1}^{\text{num}_s} \text{Smooth IL}_1(\mathbf{O}_{xy}, \mathbf{O}_{pre})}{\text{num}_s} \end{cases} \quad (3)$$

其中:  $x_{gt}^i$  和  $y_{gt}^i$  表示真实起始点的坐标;  $\mathbf{O}_{xy}$  表示真实偏移量;  $\mathbf{O}_{pre}$  表示预测偏移量. 此外, 预测偏移量使用平滑的  $\text{IL}_1$  损失进行训练, 并添加至热图坐标, 以获得起始点坐标.

最后, 获得的起始点坐标用于生成车道线先验, 可表示为

$$\left\{ (x_s, y_s) \right\} \cup \left\{ (x_i, y_i) \right\} \Big| i = \left\lfloor \frac{y_s(T-1)}{H} \right\rfloor + 1, \dots, \left\lfloor \frac{y_s(T-1)}{H} \right\rfloor + T \quad (4)$$

其中:  $(x_s, y_s)$  表示起始点坐标;  $H$  表示图像高度;  $T$  表示车道线先验点的最大数量. 沿着图像垂直方向对每个车道线先验点的  $y$  坐标  $y_i$  等距采样, 生成车道线先验. 具体生成过程详见文献[11].

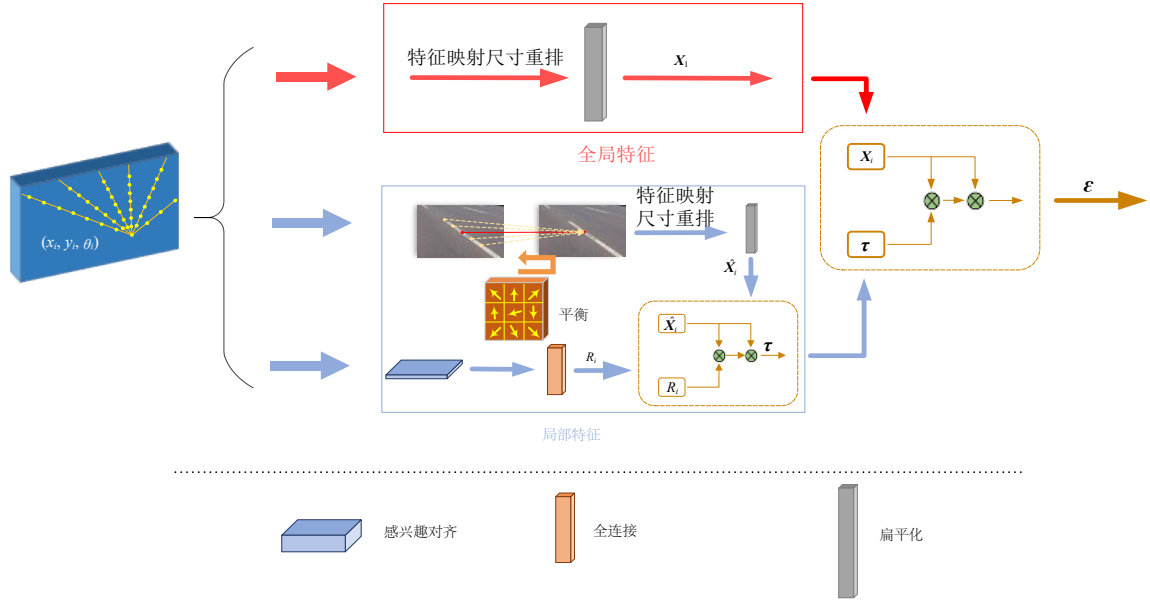


图4 车道线感知聚合模块示意图

### 3.3 车道线感知聚合

在某些情况下,车道线可能被遮挡或无法辨别,使得模型难以正确识别车道线,需获取全局语义信息才能解决这一问题.此外,由于车道线细长的特性,细化局部特征以适应车道线检测至关重要.为解决上述问题,本文使用如图4所示的车道线感知聚合模块实现以下3个功能:保留轻量级的优势,细化车道线的局部特征,并将这些细化的局部特征与全局特征进行聚合,进一步提高模型性能.由于传统卷积获取车道线局部特征的代表能力并不理想,因此,本文使用可变形卷积提取更有区分性的局部特征.首先,利用下式预测位于同一车道线上的某个车道线先验点及其 $N$ 个相邻点间的偏移量:

$$\begin{cases} \Delta o_k = \phi(F(p_i)) \\ \Delta O_k = \{\Delta o_k^n | n=1, 2, \dots, N\} \in \mathbf{R}^{2N} \end{cases} \quad (5)$$

其中: $F(p_i)$ 表示第 $i$ 个车道线先验点的特征表示; $\phi$ 表示一个非线性函数; $\Delta O_k$ 用以聚合第 $i$ 个车道线先验点 $p_i$ 处的相邻点特征.

$$\hat{F}(p_i) = \sum_{n=1}^N w_n \cdot F(p_i + \Delta o_k^n) \quad (6)$$

其中, $w_n$ 表示卷积权重.综上所述,给定输入特征 $X_j \in \mathbf{R}^{C \times H \times W}$ ( $j=0, 1, 2$ ),可以得到相应的细化特征 $\hat{X}_j$ .

将局部特征与全局特征聚合的具体过程描述如下.首先,均匀选择 $M_p$ 个车道线先验点;然后,进行感兴趣区域对齐,得到感兴趣区域特征 $R_i \in \mathbf{R}^{C \times M_p}$ ;其次,利用全连接的感兴趣区域特征 $R_i \in \mathbf{R}^{C \times 1}$ 与细调的特征 $\hat{X}_i \in \mathbf{R}^{C \times H \times W}$ 间的注意力,得到聚合的局部特征 $\tau \in \mathbf{R}^{1 \times C}$ ,具体公式如下

$$\tau = \text{Softmax} \left( \frac{R_i^T \hat{X}_i}{\sqrt{C}} \right) \hat{X}_i^T \quad (7)$$

最后,将 $\tau$ 与输入特征 $X_j \in \mathbf{R}^{C \times H \times W}$ 聚合,得到最终的输出:

$$\Gamma = \text{Softmax} \left( \frac{\tau X_j}{\sqrt{C}} \right) X_j^T \quad (8)$$

### 3.4 惩罚车道线交并比

车道线越水平,其在 $x$ 轴上的跨度越大.因此,恒定的虚拟跨度可能导致精度不足.此外,如图5所示,在LineIoU的绝对值相同时 $|D_i| = |D_{i+1}|$ ,预测车道线与真实车道线间的距离在分离和重合时存在显著差异.这是因为LineIoU的值域介于 $[-1, 1]$ 之间,当LineIoU $> 0$ 时,即使很小的距离变化也会导致LineIoU发生显著变化.相比之下,LineIoU $< 0$ 时,即使很大的距离变化也不对其产生明显影响.

受到文献[18, 19]的启发,本文利用如图5所示的可变车道跨度策略表示不同倾斜角度下的车道线:

$$\text{LIoU} = \frac{\sum_{i=j}^M \min(x_i^p + s_i^p, x_i^q + s_i^q) - \max(x_i^p - s_i^p, x_i^q - s_i^q)}{\sum_{i=j}^M \max(x_i^p + s_i^p, x_i^q + s_i^q) - \min(x_i^p - s_i^p, x_i^q - s_i^q)} \quad (9)$$

其中, $s_i^k$ ( $k \in \{p, q\}$ )是由实际车道线跨度 $s$ 和对应坐标处的倾斜角度 $\theta_i^k$ 推导而来:

$$s_i^k = \frac{s}{\sin \theta_i^k} \quad (10)$$

此外,利用提出的如下式所示的惩罚项构建非重叠车道线间的空间相关性:

$$IL_{\text{pen}} = \frac{\sum_{i=j}^M \text{ReLU}(D_i)}{\sum_{i=j}^M U_i} = \frac{\sum_{i=j}^M \text{ReLU}(U_i - 2(s_i^p + s_i^q))}{\sum_{i=j}^M U_i} \quad (11)$$

当车道线彼此重叠时,  $D_i < 0$  且  $\text{ReLU}(D_i) = 0$ , 表示惩罚项不适用. 当车道线间的距离足够大时,  $D_i > 0$ , 此时的惩罚项对应车道线间的  $IL_i$  距离. 其中:  $s_i^p$  代表预测

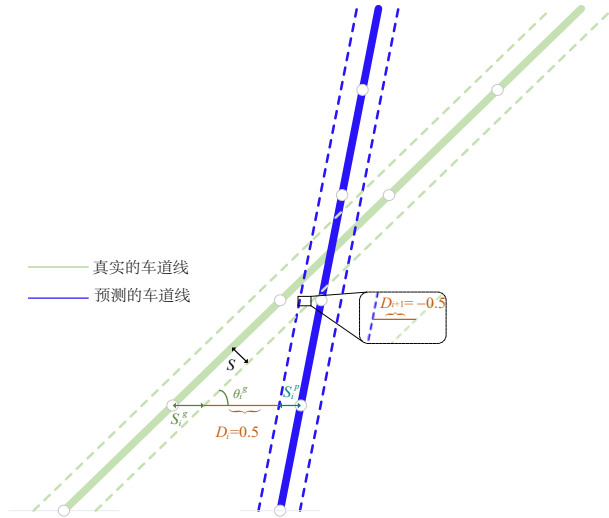


图5 惩罚车道线交并比的示意图

的车道线跨度;  $s_i^q$  代表地面真实的车道线跨度;  $D_i$  代表车道之间的距离.

本文提出的惩罚车道线交并比如下所示:

$$IL_{\text{PLIoU}} = 1 - \text{LIoU} + IL_{\text{pen}} \quad (12)$$

惩罚车道线交并比不仅解决了与车道线跨度变化和空间分布不均匀的问题, 且保留了 LIoU 计算的简单性, 从而提高了模型的整体性能.

整体损失函数可以写成:

$$IL_{\text{total}} = \lambda_{\text{PLIoU}} IL_{\text{PLIoU}} + \lambda_{\text{pro}} IL_{\text{pro}} + \lambda_{\text{offset}} IL_{\text{offset}} + \lambda_{\text{cls}} IL_{\text{cls}} \quad (13)$$

其中,  $IL_{\text{cls}}$  表示预测值与实际值间的焦点损失, 而  $\lambda_{\text{PLIoU}}$ 、 $\lambda_{\text{pro}}$ 、 $\lambda_{\text{offset}}$  和  $\lambda_{\text{cls}}$  分别设置为 1.0、0.3、0.3 和 1.0. 此外, 本文仅在训练阶段引入分割损失辅助模型训练.

## 4 实验

### 4.1 数据集

通过在 3 个公开数据集上进行实验, 验证本文提出方法的检测性能: CULane: City-Urban Lane<sup>[8]</sup>、TuSimple<sup>[23]</sup> 和 CurveLanes<sup>[24]</sup> 数据集. CULane 采集包括夜间、堵车、强光等 9 种交通场景, 每个带标注图像的像素为  $1\,640 \times 590$ . TuSimple 采集了晴朗天气下高速公路的交通场景, 每个带标注图像的像素为  $1\,280 \times 720$ . CurveLanes 采集了曲线车道和密集车道的交通场景, 每个

带标注图像的像素为  $2\,650 \times 1\,440$ .

### 4.2 实施细节

本文在 CULane、TuSimple 和 CurveLanes 上分别进行了 100、60 和 36 轮的训练. 所有输入图像像素均调整为  $320 \times 800$ . 采用运动、水平翻转、随机亮度、随机对比度、随机 HSV 色调、饱和度、数值 (Hue, Saturation, Value, HSV) 调节、中值模糊和随机仿射调节实现数据增强, 并在 CULane 最后 4 轮、TuSimple 最后 10 轮、CurveLanes 最后 6 轮关闭数据增强, 使收敛结果更接近于真实情况. 优化过程采用具有余弦衰减学习率策略的 AdamW 优化器. 将预训练的残差网络<sup>[25]</sup> (Residual Network, ResNet) 和深度层聚合<sup>[26]</sup> (Deep layer aggregation, DLA) 用作骨干网络. 车道线感知聚合中的相邻车道先验点数  $N$  设置为 9, 每个车道先验上点的最大数量  $T$  设置为 80, 通道数  $C$  设置为 64, 实际车道宽度  $s$  设置为 15, 特征聚合过程中车道线先验点数  $M_p$  设置为 40.

### 4.3 评价指标

评估指标因数据集而异. 对于 TuSimple 数据集, 主要评估指标准确率率的计算过程如下:

$$\text{Accuracy} = \frac{C_{\text{clip}}}{S_{\text{clip}}} \quad (14)$$

其中:  $C_{\text{clip}}$  表示每个图像中正确预测的车道线点的数量;  $S_{\text{clip}}$  表示车道线点的真实数量. 如果超过 85% 的预测车道线点在对应的真值范围内 (20pixel 以内), 则认为预测结果准确.

此外, 在 CULane 和 CurveLanes 数据集中采用真值和预测值间的交并比作为评估指标, 计算公式如下:

$$\begin{cases} F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \\ \text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \\ \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \end{cases} \quad (15)$$

其中, TP、FP 和 FN 分别表示真阳性、假阳性和假阴性.

### 4.4 实验结果

CULane 数据集. 表 1 给出了本文方法 SMFRNet 和其他主流车道线检测方法在 CULane 数据集上的实验结果, 可以看到本文方法在各方面都取得了最优的结果. 与以 ResNet34 作为骨干网络的 ADNet (ADNet-ResNet3) 相比, 采用 ResNet34 作为骨干网络的 SMFRNet (SMFRNet-ResNet34) 的 FPS 值达到 101, 且在保持更快处理速度的同时,  $F_1@50$  值提高了 1.59%. 可以看出, 在正常 (Normal) 路况下, SMFRNet-ResNet34 的  $F_1$  值为 93.57%, 比 ADNet-R34 高出 0.67%. 此外, 在一些复杂场景, 如堵车 (Crowded)、强光 (Dazzle)、阴影

(Shadow)等, SMFRNet-ResNet34 的性能具有明显优势, 其  $F_1$  值分别比 ADNet-ResNet34 提高了 1.76%、2.51% 和 3.33%, 在堵车 (Crowded)、强光 (Dazzle)、无车道线 (No line) 等复杂条件下获得的优异检测结果表明, 与现有方法相比, SMFRNet 在全局语义理解和车道线回归能力方面表现出更优越的特性.

与以 DLA34 作为骨干网络的基线方法 CLRNet 相比, SMFRNet-DLA34 在  $F_1@50$  和  $F_1@75$  上分别提高了 0.92% 和 1.84%. 即使选择 ResNet34 作为骨干, SMFRNet 每 s 处理的帧率仍然可以达到 101,  $F_1@50$  和  $F_1@75$  分别可以达到 80.53% 和 63.87%, 超过了基线方法.

从实验结果可以看出, SMFRNet 对于复杂场景检测的性能提升尤为明显: (1) 在遮挡和无车道线场景下, SMFRNet 利用全局特征优化模块优化了全局特征

表示, 并在车道线感知聚合模块中与局部特征进行了聚合, 因此模型可以更好地捕获全局上下文信息, 更好地对被遮挡或缺失的车道线进行预测; (2) SMFRNet 在车道线感知聚合模块中对局部特征进行了细化, 因此在光照变化和阴影等场景下, 可以得到更加精细的局部特征, 提升了此类模糊场景的检测性能; (3) 起始点坐标预测模块优化了车道线检测的锚生成方法, 生成了更灵活的锚; 惩罚车道线交并比损失函数兼顾距离度量及倾斜角度, 增强了损失函数的一般性. 上述优化策略对于所有场景的检测性能均有提升. 由于本文并未对弯曲及交叉车道线进行针对性处理, 因此, 相较于聚焦于曲线车道线检测方法<sup>[7]</sup>, 仍有提升空间. 同时, 每 s 处理的帧率也仍有优化空间. 部分可视化结果可见于图 6.

表 1 不同方法在 CULane 数据集上的实验结果

方法	主干网络	$F_1@50/\%$	$F_1@75/\%$	Normal/ $\%$	Crowded/ $\%$	Dazzle/ $\%$	Shadow/ $\%$	No line/ $\%$	Arrow/ $\%$	Curve/ $\%$	Cross/ $\%$	Night/ $\%$	FPS/帧
SCNN <sup>[8]</sup>	VGG16	71.60	39.84	90.60	69.70	58.50	66.90	43.40	84.10	64.40	1 990	66.10	25.0
RESA <sup>[9]</sup>	ResNet50	75.30	53.39	92.10	73.10	69.20	72.80	47.70	88.30	70.30	1 503	69.90	39.7
FastDraw <sup>[27]</sup>	ResNet50	—	—	85.90	63.60	57.00	69.90	40.60	79.40	65.20	7 013	57.80	—
UFLDv2 <sup>[28]</sup>	ResNet34	76.00	—	92.50	74.80	65.50	75.50	49.20	88.80	70.10	1 910	70.80	114.0
PINet <sup>[11]</sup>	Hourglass	74.40	51.33	90.30	72.30	66.30	68.40	49.80	83.70	65.20	1 427	67.70	25.0
LaneATT <sup>[5]</sup>	ResNet122	77.02	57.50	91.74	76.16	69.47	76.31	50.46	86.29	64.05	1 264	70.81	38.0
SGNet <sup>[29]</sup>	ResNet34	77.27	—	92.07	75.41	67.75	74.31	50.90	87.97	69.65	1 373	72.69	98.0
CondLane <sup>[4]</sup>	ResNet34	78.74	59.39	93.38	77.14	71.17	79.93	51.85	89.89	73.88	1 387	73.92	137.0
CondLane <sup>[4]</sup>	ResNet101	79.48	61.23	93.47	77.44	70.93	80.91	54.13	90.16	75.21	1 201	74.80	58.0
CLRNet <sup>[7]</sup>	ResNet101	80.13	62.96	93.85	78.78	72.49	82.33	54.50	89.79	75.57	1 262	75.51	61.0
CLRNet <sup>[7]</sup>	DLA34	80.47	62.78	93.73	79.59	75.30	82.51	54.58	90.62	74.13	1 155	75.37	110.0
ADNet <sup>[19]</sup>	ResNet18	77.56	—	91.92	75.81	69.39	76.21	51.75	87.71	68.84	1 133	72.33	75.0
ADNet <sup>[19]</sup>	ResNet34	78.94	—	92.90	77.45	71.71	79.11	52.89	89.90	70.64	1 499	74.78	62.0
CLRmatchNet <sup>[30]</sup>	DLA34	79.97	62.10	93.74	79.50	73.57	82.27	54.61	90.96	<b>77.09</b>	2 064	75.19	110.0
SMFRNet(ours)	ResNet18	80.08	63.25	92.38	78.95	73.41	81.64	55.31	89.76	73.47	1 189	75.88	111.0
SMFRNet(ours)	ResNet34	80.53	63.87	93.57	79.21	74.22	82.44	55.71	90.61	75.23	<b>1 093</b>	76.24	101.0
SMFRNet(ours)	ResNet101	81.01	64.53	<b>94.23</b>	79.76	74.88	82.71	55.83	90.79	76.32	1 167	76.34	43.0
SMFRNet(ours)	DLA34	<b>81.39</b>	<b>64.80</b>	94.12	<b>80.21</b>	<b>76.01</b>	<b>83.67</b>	<b>56.53</b>	<b>91.09</b>	75.87	1 254	<b>76.85</b>	84.0

注:加粗数据表示此评价指标下的最佳性能, 其中  $F_1@50$  和  $F_1@75$  分别表示阈值为 0.5 和 0.75 时的评估指标.

CurveLanes 数据集. 表 2 给出了不同方法在 CurveLanes 数据集上的实验结果. 由表 2 所示结果可知: SMFRNet-DLA34 超过了所有其他方法, 在保持高处理速度的同时, 取得了最高的  $F_1$  值 86.83%, 比 CLRNet-DLA34 高出 0.73%. 此外, 与 CondLane-ResNet101 相比, SMFRNet 的 GFlops 显著从 44.9 减少到 22.5, 仅比 CLRNet-DLA34 高出 4.1.

TuSimple 数据集. 由于 TuSimple 数据集的场景相对单一, 不同方法在其上的性能差异在 1% 以内, 已达到饱和. 因此, 复杂的网络结构对于性能的提升并不明

显, 且具有发生过拟合从而导致性能降低的风险. 表 3 显示了所提出的方法在 TuSimple 数据集上达到了与最先进的方法相当的性能.

#### 4.5 消融实验

在 CULane 数据集上进行消融实验, 验证每个子模块的有效性. 所有的消融实验都是在本文提出的 SMFRNet-DLA34 版本上进行, 相应的实验结果如表 4 所示. 第 1 行显示了基线方法的结果, 从第 2 行到第 5 行, 逐步添加了全局特征优化、起始点坐标预测、车道线感知聚合和惩罚车道线交并比, 以观察它们对整体

图6 本文方法和基准方法CLRNet在CULane数据集上检测结果的可视化

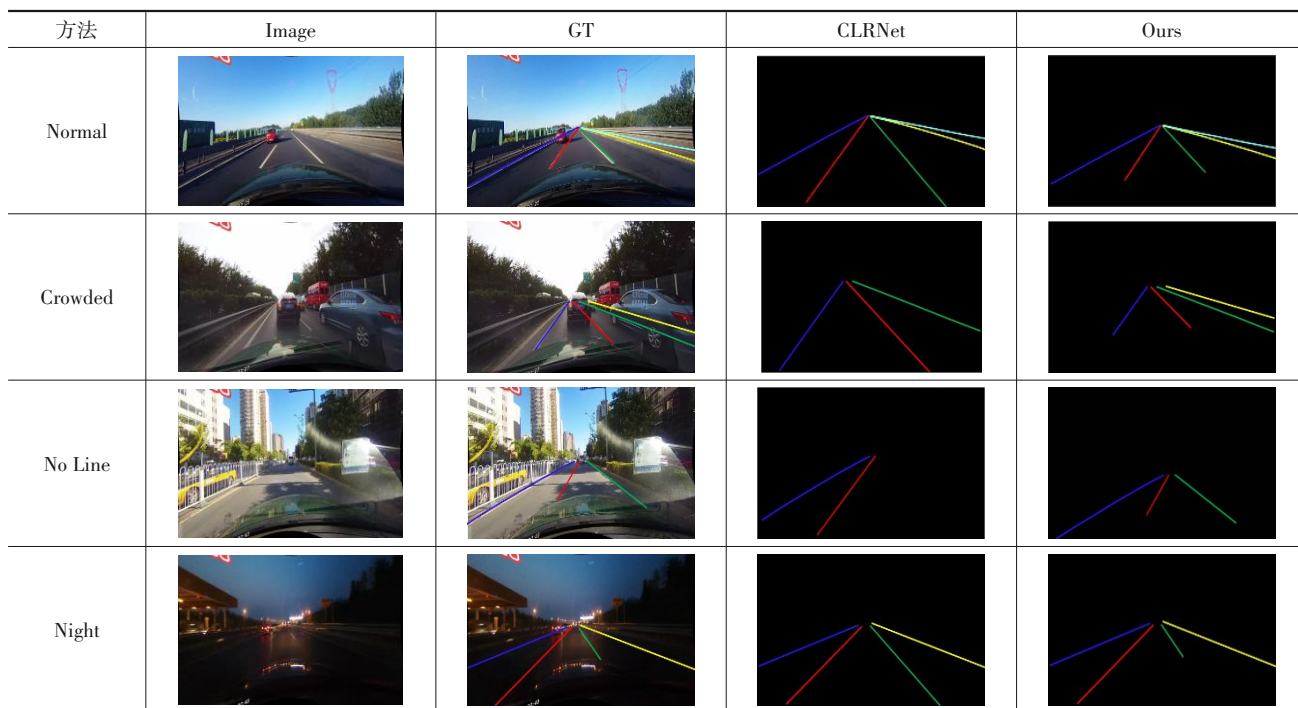


表2 不同方法在CurveLanes数据集上的实验结果

方法	主干网络	$F_1@50/\%$	Precision/%	Recall/%	GFlops/GB
SCNN <sup>[8]</sup>	VGG16	65.02	76.13	56.74	328.4
CondLane <sup>[4]</sup>	ResNet18	85.09	87.75	82.58	10.3
CondLane <sup>[4]</sup>	ResNet34	85.92	88.29	83.68	19.7
CondLane <sup>[4]</sup>	ResNet101	86.10	88.98	83.41	44.9
CLRNet <sup>[6]</sup>	DLA34	86.10	91.40	81.39	18.4
BézierLaneNet <sup>[16]</sup>	ResNet-18	75.03	82.75	68.62	—
SMFRNet(ours)	DLA34	86.83	91.56	82.57	22.5

性能的影响. 本文遵循第4.2小节中概述的实施细节,

针对每种情况训练和验证了10次模型.

第2行的实验结果显示,将全局特征优化模块添加到骨干网络中可将 $F_1$ 值提高0.12%,表明更大的感受野有利于网络更好地理解上下文信息. 第3行的实验结果显示,将起始点坐标预测模块添加到骨干网络中可将 $F_1$ 值提高0.27%,表明预定义锚的准确性和灵活性显著影响了行锚方法的性能. 第4行的实验结果显示,将ROIgather<sup>[12]</sup>替换为车道线感知聚合可将 $F_1$ 值提高0.16%,表明捕获车道线特定的局部模式有助于提取更健壮的特征表示,从而提高回归精度. 最后1行的实验结果显示,将LineIoU损失替换为惩罚车道线交并比可

表3 不同方法在TuSimple数据集上的实验结果

方法	主干网络	$F_1@50$	Accuracy	False Positive	False Negative
SCNN <sup>[8]</sup>	VGG16	95.97	96.53	6.17	1.80
LaneATT <sup>[5]</sup>	ResNet18	96.71	95.57	3.56	3.01
CondLane <sup>[4]</sup>	ResNet101	97.24	96.54	2.01	3.50
CLRNet <sup>[6]</sup>	ResNet18	97.89	96.84	2.28	1.92
ADNet <sup>[19]</sup>	ResNet34	97.31	96.60	2.83	2.53
BézierLaneNet <sup>[16]</sup>	ResNet-18	—	95.61	5.30	3.50
BézierLaneNet <sup>[16]</sup>	ResNet34	—	95.87	5.00	3.40
ADNet <sup>[19]</sup>	ResNet18	96.90	96.23	2.91	3.29
ADNet <sup>[19]</sup>	ResNet34	97.31	96.60	2.83	2.53
PGA-Net <sup>[31]</sup>	ResNet18	97.66	95.43	4.53	4.67
PGA-Net <sup>[31]</sup>	ResNet34	97.62	95.34	3.89	4.13
AR-NET <sup>[32]</sup>	ResNet18	96.11	96.73	2.45	4.06
SMFRNet(ours)	DLA34	97.93	96.91	2.14	2.25

将  $F_1$  值提高 0.37%, 表明与 LineIoU 相比, 惩罚车道线交并比在距离上具有更均匀的分布, 且能更好地处理倾斜的车道线。

表 4 CULane 数据集上的消融实验 %

GFR	SPCP	LAG	PLaneIoU	$F_1@50$
—	—	ROIgather <sup>[6]</sup>	LineIoU <sup>[6]</sup>	80.47
√	—	ROIgather <sup>[6]</sup>	LineIoU <sup>[6]</sup>	80.59
√	√	ROIgather <sup>[6]</sup>	LineIoU <sup>[6]</sup>	80.86
√	√	√	LineIoU <sup>[6]</sup>	81.02
√	√	√	√	<b>81.39</b>

注:加粗数据表示此评价指标下的最佳性能。

## 5 结论

本文提出了一种更灵活的基于行锚的车道线检测方法。在骨干网络中引入全局特征优化模块, 扩展了全局感受野, 优化了特征表示。起始点坐标预测预测每个车道实例的起始坐标, 并以此为依据生成了更适合复杂场景的锚。利用车道线感知聚合, 处理含有车道先验信息的特征图, 增强了局部特征之间的相关性。提出了兼顾距离度量及车道线倾斜角度的惩罚车道线交并比, 增强了损失函数的一般性。通过实验验证, 本文提出的模型在 3 个公开数据集上取得了良好的检测效果, 证明了改进的有效性。

## 参考文献

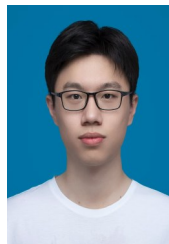
- [1] LIU G L, WÖRGÖTTER F, MARKELIĆ I. Combining statistical hough transform and particle filter for robust lane detection and tracking[C]//2010 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE, 2010: 993-997.
- [2] ZHOU S Y, JIANG Y H, XI J Q, et al. A novel lane detection based on geometrical model and Gabor filter[C]//2010 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE, 2010: 59-64.
- [3] 徐频捷, 陈逸杰, 李之南, 等. 基于事件驱动的车道线识别算法研究[J]. 电子学报, 2021, 49(7): 1379-1385.  
XU P J, CHEN Y J, LI Z N, et al. Research on event-driven lane recognition algorithms[J]. Acta Electronica Sinica, 2021, 49(7): 1379-1385. (in Chinese)
- [4] LIU L Z, CHEN X H, ZHU S Y, et al. CondLaneNet: A top-to-down lane detection framework based on conditional convolution[EB/OL]. (2021-05-11)[2023-02-10]. <https://arxiv.org/abs/2105.05003v3>.
- [5] TABELINI L, BERRIEL R, PAIXAO T M, et al. Keep your eyes on the lane: Real-time attention-guided lane detection[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 294-302.
- [6] ZHENG T, HUANG Y F, LIU Y, et al. CLRNet: Cross layer refinement network for lane detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 888-897.
- [7] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to  $31 \times 31$ : Revisiting large kernel design in CNNs [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 11953-11965.
- [8] PAN X G, SHI J P, LUO P, et al. Spatial as deep: Spatial CNN for traffic scene understanding[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 7276-7283.
- [9] ZHENG T, FANG H, ZHANG Y, et al. RESA: Recurrent feature-shift aggregator for lane detection[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3547-3554.
- [10] CHE Q H, NGUYEN D P, PHAM M Q, et al. TwinLiteNet: An efficient and lightweight model for driveable area and lane segmentation in self-driving cars[C]//2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). Piscataway: IEEE, 2023: 1-6.
- [11] KO Y, LEE Y, AZAM S, et al. Key points estimation and point instance segmentation approach for lane detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(7): 8949-8958.
- [12] QU Z, JIN H, ZHOU Y, et al. Focus on local: Detecting lane marker from bottom up via key point[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 14122-14130.
- [13] WANG J S, MA Y C, HUANG S F, et al. A keypoint-based global association network for lane detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 1382-1391.
- [14] TABELINI L, BERRIEL R, PAIXAO T M, et al. PolyLaneNet: Lane estimation via deep polynomial regression[C]//2020 25th International Conference on Pattern Recognition (ICPR). Piscataway: IEEE, 2021: 6150-6156.
- [15] FENG Z Y, GUO S H, TAN X, et al. Rethinking efficient lane detection via curve modeling[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 17041-17049.
- [16] WU Y J, ZHAO L Q, LU J W, et al. Dense hybrid propos-

- al modulation for lane detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(11): 6845-6859.
- [17] QIN Z Q, ZHANG P Y, 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.
- [18] HONDA H, UCHIDA Y. CLRerNet: Improving confidence of lane detection with LaneIoU[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2024: 1165-1174.
- [19] XIAO L Y, LI X, YANG S, et al. ADNet: Lane shape prediction via anchor decomposition[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2023: 6381-6390.
- [20] GUO M H, LU C Z, HOU Q B, et al. Segnext: Rethinking convolutional attention design for semantic segmentation [C]//Proceedings of the IEEE Conference on Neural Information Processing Systems. Piscataway: IEEE, 2022: 1-17.
- [21] WOO S, DEBNATH S, HU R H, et al. ConvNeXt V2: Co-designing and scaling ConvNets with masked autoencoders[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 16133-16142.
- [22] DUAN K W, BAI S, XIE L X, et al. Centernet: Key point triplets for object detection[C]//Proceedings of the IEEE Conference on Computer Vision. Piscataway: IEEE, 2019: 6569-6578.
- [23] TuSimple. Tusimple benchmark[EB/OL]. (2020-10-7)[2024-01-03]. <https://github.com/TuSimple/tusimple-enchmark/>.
- [24] XU H, WANG S J, CAI X Y, et al. CurveLane-NAS: Unifying lane-sensitive architecture search and adaptive point blending[C]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 689-704.
- [25] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [26] YU F, WANG D Q, SHELHAMER E, et al. Deep layer aggregation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2403-2412.
- [27] PHILION J. FastDraw: Addressing the long tail of lane detection by adapting a sequential prediction network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 11582-11591.
- [28] QIN Z Q, ZHANG P Y, 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.
- [29] SU J M, CHEN C, ZHANG K, et al. Structure guided lane detection[C]//Proceedings of the International Joint Conference on Artificial Intelligence. Sacramento: IJCAI, 2021: 997-1003.
- [30] KONTENTE S, ORFAIG R, BOBROVSKY B Z. CLR-matchNet: Enhancing curved lane detection with deep matching process[EB/OL]. (2023-09-26)[2024-03-31]. <https://arxiv.org/abs/2309.15204v2>.
- [31] LI Q K, YU X W, CHEN J X, et al. PGA-net: Polynomial global attention network with mean curvature loss for lane detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(1): 417-429.
- [32] TONG G X, ZU C Y. AR-NET: Lane detection model with feature balance concerns for autonomous driving[J]. Neural Computing and Applications, 2024, 36(8): 3997-4012.

#### 作者简介



朱松豪 男, 1973年2月生, 江苏如皋人。南京邮电大学自动化学院、人工智能学院副教授。主要研究方向为图像处理、人工智能。  
E-mail: zhush@njupt.edu.cn.



谭少涵 男, 1995年8月生, 江苏南京人。南京邮电大学自动化学院、人工智能学院硕士研究生。主要研究方向为模式识别、深度学习。  
E-mail: 394600079@qq.com.