

融合ResNeXt与坐标注意力的车道线检测方法

杨鹏成¹, 钟琪²

¹西华大学汽车与交通学院, 四川 成都

²庆铃汽车(集团)有限公司, 重庆

收稿日期: 2025年12月9日; 录用日期: 2025年12月31日; 发布日期: 2026年1月13日

摘要

随着智能驾驶与ADAS的发展, 车道线检测对路径规划和车道保持等模块至关重要, 亟需兼顾高精度与实时性的轻量化方法。现有传统方法对遮挡和光照变化鲁棒性不足, 深度学习分割和BEV/3D等方法又存在计算开销大或部署成本高的问题, 基于行锚的UFLD方法虽具速度优势, 但在长距离连续性建模和复杂场景鲁棒性方面仍有提升空间。本文在UFLD框架上引入ResNeXt-50主干网络与坐标注意力机制, 并结合分类、结构和分割三种损失共同约束车道线的平滑性与刚性。在CULane数据集上, 所提方法取得72.51% F1-score和211.42 FPS, 相比其它车道线检测方法达到了精度与速度上平衡。在遮挡、高光、阴影、弯道及夜间等典型场景中, 可视化结果显示该方法的漏检与误检数量少于对比方法。实验结果表明, 本文所提方法在大幅提升推理速度的同时保持了具有竞争力的检测精度, 可用于复杂交通环境下的实时车道线检测。

关键词

车道线检测, 行锚机制, ResNeXt, 坐标注意力

Lane Line Detection Method Integrating ResNeXt and Coordinate Attention

Pengcheng Yang¹, Qi Zhong²

¹School of Automobile and Transportation, Xihua University, Chengdu Sichuan

²Qingling Motors (Group) Co., LTD., Chongqing

Received: December 9, 2025; accepted: December 31, 2025; published: January 13, 2026

Abstract

With the development of intelligent driving and ADAS, lane line detection is crucial for modules such as path planning and lane keeping, and there is an urgent need for lightweight methods that

文章引用: 杨鹏成, 钟琪. 融合 ResNeXt 与坐标注意力的车道线检测方法[J]. 人工智能与机器人研究, 2026, 15(1): 232-241. DOI: 10.12677/airr.2026.151023

balance high precision and real-time performance. The existing traditional methods have insufficient robustness against occlusion and illumination changes. Methods such as deep learning segmentation and BEV/3D have problems of high computational overhead or high deployment cost. Although the UFLD method based on row anchors has a speed advantage, there is still room for improvement in long-distance continuous modeling and robustness in complex scenes. In this paper, the ResNeXt-50 backbone network and coordinate attention mechanism are introduced on the UFLD framework, and the smoothness and rigidity of lane lines are jointly constrained by the three losses of classification, structure and segmentation. On the CULane dataset, the proposed method achieved a score of 72.51% F1-score and 211.42 FPS, achieving a balance between accuracy and speed compared with other lane line detection methods. In typical scenarios such as occlusion, highlights, shadows, curves and at night, the visualization results show that the number of missed detections and false detections of this method is less than that of the comparison method. The experimental results show that the method proposed in this paper not only significantly improves the reasoning speed but also maintains competitive detection accuracy, and can be used for real-time lane line detection in complex traffic environments.

Keywords

Lane Line Detection, Lane Anchor Mechanism, ResNeXt, Coordinate Attention

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着智能驾驶技术的发展，环境感知已成为自动驾驶与高级驾驶辅助系统(ADAS)的关键能力。在感知任务中，车道线检测不仅提供道路边界和行驶引导信息，还为路径规划、车道保持与变道决策等下游模块提供直接约束，因此，实现高精度、低延迟、轻量化的车道线检测，对车辆在复杂交通场景中稳定行驶至关重要。

早期车道线方法依赖人工设计的视觉特征与几何模型，可分为基于特征与基于模型两类。基于特征的方法通过颜色、纹理和边缘等浅层线索分割车道，结合逆透视变换与霍夫变换定位车道，但对遮挡、破损和光照变化敏感且鲁棒性差[1]。基于模型的方法则以直线、多项式或样条等几何先验拟合车道，通过 RANSAC、最小二乘等方法增强对缺失信息的容错性，但这些方法难以覆盖多样化路况，且计算开销大、实时性不足[2]。

近年来，深度学习推动了车道线检测的多样化发展，主要分为语义与实例分割、检测(线锚/行锚)、参数曲线回归、关键点法以及 BEV/3D 方法。分割类的方法在细粒度上有优势，但计算量大且难以推断被遮挡区域[3][4]。参数化曲线和关键点方法能减少后处理，但在表示复杂局部形变时受限且训练不易收敛[5][6]。BEV/3D 方法在几何一致性上表现良好，但对标定、深度信息和算力要求高，限制了其在嵌入式平台的应用[7]。

检测类方法借鉴目标检测思想，通过适配的锚表示(如线锚、行锚)对车道进行建模，兼顾速度与精度。线锚方法能较好恢复车道形状，行锚(row-anchor)则以图像行作为采样单元，通过回归每行的水平偏移或存在概率重建车道，极大降低了计算复杂度并提升实时性，如 UFLD [8]。尽管行锚在速度上具有明显优势，但在长距离连续性、多车道分离、交叉口拓扑以及遮挡下的全局推断方面仍有提升空间。

综上，现有方法在精度、实时性与泛化性之间仍需折中。基于行锚的方案以其简单高效的结构为工

程化部署提供了可行路径。本文提出了基于改进 UFLD 的车道线检测算法, 通过引入 ResNeXt 主干网络 [9]与坐标注意力机制[10], 增强模型的特征表达能力和长距离依赖关系, 提升在复杂场景下的车道线检测精度。

2. 方法

2.1. UFLD 概述

UFLD 算法首次将行锚思想引入车道线检测, 将任务从逐像素分割转为行级分类。该方法把图像在纵向划分为若干行锚, 每行再划分为若干水平单元, 并增加“无车道”类别, 从而用行内分类替代全图像素级预测。与传统逐像素方法相比, 行锚显著压缩了预测维度与计算量, 且行数与单元格数可按需调整以在精度与效率间权衡。

在网络结构上, UFLD 由逐行检测分支与辅助分割分支构成, 如图 1。其中, 逐行检测分支通过 ResNet 主干网络提取全局特征, 随后以轻量回归和分类头映射到预定义行锚, 实现车道位置预测。而辅助分割分支在训练阶段融合多尺度信息以强化特征表达, 但推理时可被裁剪以节省算力。

总的来说, 行锚方法兼顾了局部精度与全局上下文: 每个输出基于全图感受野, 因而可以在遮挡或复杂光线下进行合理推断; 同时其结构简单、并行性强, 便于轻量化与实时部署。但是, 行锚方法虽大幅降低计算量, 却以精度为代价, 整体表现尚未达到最优。将图像离散为单元格也使得对车道线的长距离连续性建模困难, 模型需跨越较远区域才能恢复车道整体走势, 从而影响准确识别。因此, 本文引入 ResNeXt 作为主干以增强特征表达, 并使用坐标注意力模块捕捉长距离依赖, 从而提升复杂场景下的车道线检测精度。

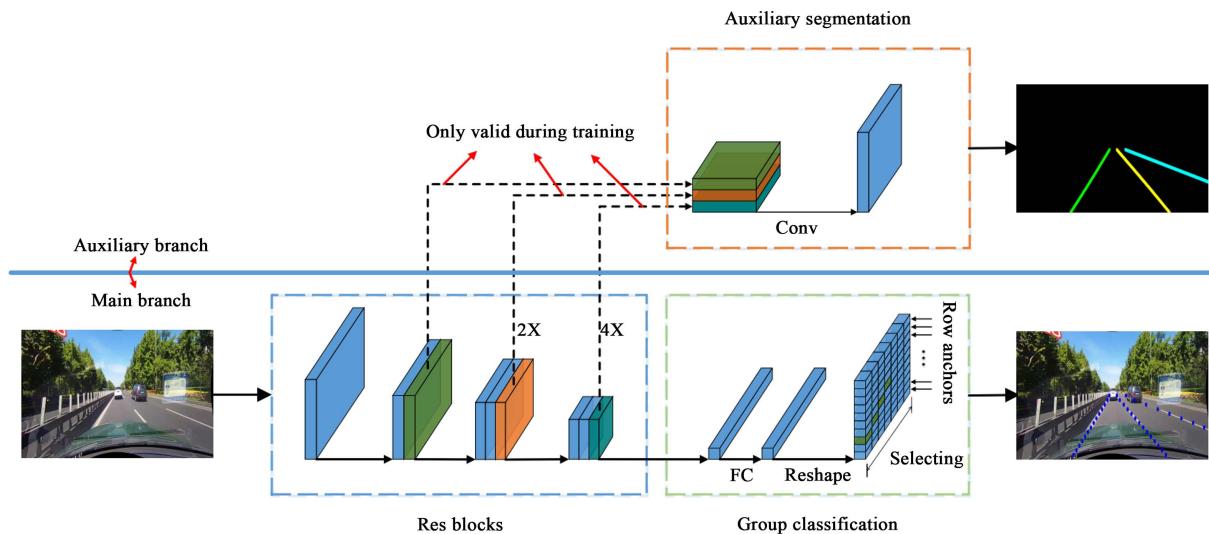


Figure 1. UFLD network architecture [9]

图 1. UFLD 网络结构[9]

2.2. ResNeXt 主干网络

在 UFLD 的网络框架中, 一般以浅层的网络 ResNet-18 和 ResNet-34 作为主干网络, 但在不增加网络深度的前提下仍需强化特征表征以提升检测精度。为此常见的两种改进路线是增宽通道(Wide ResNet)与改进网络结构(如 MobileNetV2、ResNeXt)。Wide ResNet 的思路是加宽残差块的输出通道, 而非盲目加深

网络。通过按倍数放大卷积通道, Wide ResNet 能在相同计算预算下增强模型的表达能力, 缓解深层网络的梯度问题, 从而在一定算力限制下提高特征提取效果, 但通道扩展带来大量可训练参数, 若数据不足或正则化不到位, 模型更容易过拟合。MobileNetV2 面向移动与嵌入式场景, 采用深度可分离卷积与逆残差(扩展 - 深度卷积 - 投影)结构, 并辅以线性瓶颈与 ReLU6 激活。该结构极大降低了参数与计算量, 便于低算力平台实现高效推理, 但在极端精度需求下其表征能力有限。

ResNeXt 则通过增加“基数”(Cardinality)引入多路径并行的分组卷积, 把单一路径的卷积操作拆分为多个子路径并聚合输出。图 2 展示了 ResNet 与 ResNeXt ($C = 32$) 的结构差异, 每一层可表示为“输入通道数, 卷积核大小, 输出通道数”。整个 ResNeXc 的运算流程公式如下:

$$y = x + \sum_{i=1}^C \tau_i(x) \quad (1)$$

其中, x 为输入的特征, C 为并行的路径数量, $\tau_i(x)$ 为第 i 条路径的变换, y 为输出的特征。以图 2 为例, 若输入的通道数为 256, 瓶颈宽度 $d = 4$, 则总参数量为:

$$C \times (256 \times 1 \times 1 \times d + d \times 3 \times 3 \times d + d \times 1 \times 1 \times 256) \quad (2)$$

该总参数量与同样数据输入的 ResNet 结构的参数量一致。这种以分支数量换取表达力的策略, 在几乎不改变整体复杂度的情况下, 显著提升了特征细节捕捉与泛化能力。因此, ResNeXt 在相同计算代价下通常优于同规模的 ResNet。所以本文选用 ResNeXt-50 ($32 \times 4d$) 作为主干网络, 以兼顾精度与实时性。

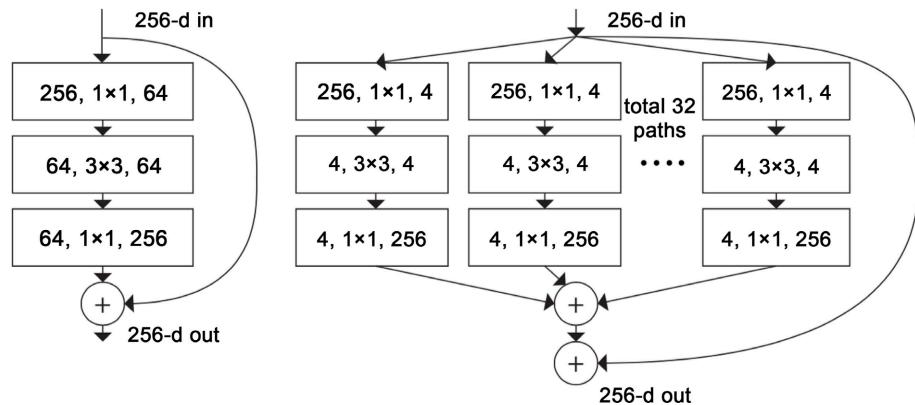


Figure 2. ResNet and ResNeXt ($C = 32$) structures [9]
图 2. ResNet 与 ResNeXt ($C = 32$) 结构[9]

2.3. 坐标注意力机制

在车道线检测任务中, 车道线之间通常存在显著的长距离依赖关系, 即不同位置之间相互关联、彼此影响, 这对于模型准确判断车道线的方向和空间位置十分关键。

然而, 传统注意力机制在建模此类长距离依赖时仍存在明显不足。通道注意力机制 SENet 通过对通道进行加权来强化关键信息, 但完全丢弃了空间位置信息; 在此基础上, 混合注意力机制 CBAM 进一步加入空间注意力模块, 借助大卷积核扩展感受野, 提高了模型对空间信息的关注程度, 但依旧难以充分表征车道线的全局连续特征, 对长距离依赖关系的建模能力仍然有限。

相比之下, 坐标注意力在上述问题上给出了改进方案。坐标注意力分别沿垂直和水平方向聚合空间特征, 并将精确的位置信息嵌入通道注意力的计算过程, 使得网络在关注通道间关系的同时, 还能够显式编码空间位置以及长距离依赖信息。一方面, 坐标注意力沿空间轴进行特征编码, 强化模型对车道线

整体形状和连续性的表征能力。另一方面, 在通道注意力计算中保留空间维度, 使通道重要性与空间位置之间的关联可以被同步建模。坐标注意力机制的结构示意如图 3 所示。

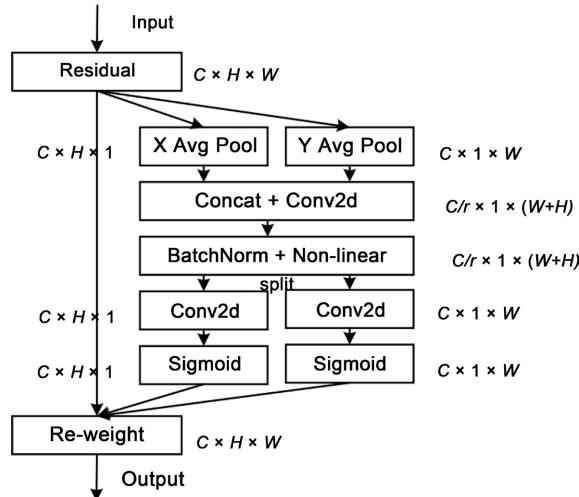


Figure 3. Coordinate attention mechanism structure [10]

图 3. 坐标注意力机制结构[10]

网络采用 ResNeXt-50 ($32 \times 4d$)作为主干, 对车道线特征进行提取。与原始的 ResNet-50 相比, 在计算复杂度相同的条件下, 该网络具有更好的泛化能力。在此基础上引入坐标注意力机制, 使网络更加聚焦于关键车道线位置的检测, 削弱次要区域的影响, 从而提升检测精度, 并缓解由于行锚单元格离散划分带来的、对连贯车道线结构长距离依赖关系难以建模的问题。

坐标注意力机制被插入在 ResNeXt-50 ($32 \times 4d$)网络的浅层以及深层特征之后。将坐标注意力加入浅层特征之后, 有利于捕捉车道线的边缘方向和弱纹理特征, 同时抑制背景噪声, 减少无关背景的干扰。而在深层特征之后加入坐标注意力, 此时特征已经具有较强的高层语义信息和较大的感受野, 能够表征车道整体走向及拓扑结构, 进而有效建立长距离依赖。

2.4. 损失函数

损失函数(Loss Function)是在机器学习和深度学习模型中, 用于度量模型预测结果与真实目标(标签)之间差异的函数。它的主要作用是为模型优化提供依据, 引导模型通过不断调整参数来减小预测误差, 从而提升整体性能。

本文方法采用 UFLD 网络中提出的损失设计, 由分类损失、结构损失和分割损失三部分组成, 分别对应逐行检测、车道线结构关系约束以及特征融合下的辅助分割任务。分类损失的计算公式如下:

$$L_{cls} = \sum_{i=1}^C \sum_{j=1}^h L_{CE}(P_{i,j}, T_{i,j}) \quad (3)$$

其中, $P_{i,j}$ 为 $(w+1)$ 维的向量, 表示第 j 行选择 $w+1$ 个栅格的概率。 $T_{i,j}$ 为目标标签的 one-hot 编码, L_{CE} 为交叉熵损失函数。在已有分类损失的基础上, 进一步引入结构损失函数, 用于约束车道线之间的位置关系, 并显式刻画其结构化信息。车道线主要具有平滑性和刚性两种特征。平滑性强调车道线的连续特性: 同一条车道线上的像素点, 在相近的图像高度处, 其横坐标应当互相接近, 并呈现整体平滑的变化趋势。基于这一性质, 结构损失通过在相邻行锚之间建立约束, 使模型在分类损失之上进一步捕捉车道

线的几何特征, 从而保证预测结果更符合真实车道线的自然形态。相似性损失的计算公式如下:

$$L_{sim} = \sum_{i=1}^C \sum_{j=1}^{h-1} \|P_{i,j} - P_{i,j+1}\|_1 \quad (4)$$

式中用 $\|\cdot\|_1$ 表示 L_1 范数, 用于度量相邻行之间分类分布的差异。

此外, 车道线还体现出刚性特征: 在大多数情形下整体形状近似直线, 即便处于弯道路段, 由于透视效应的存在, 其大部分区域仍接近直线。基于这一特征, 通过引入直线的二阶差分趋近于零的约束来限制车道线形状, 从而体现刚性约束, 对应的形状损失定义为:

$$L_{shp} = \sum_{i=1}^C \sum_{j=1}^{h-2} \left\| \left(Loc_{i,j} - Loc_{i,j+1} \right) - \left(Loc_{i,j+1} - Loc_{i,j+2} \right) \right\|_1 \quad (5)$$

其中, $Loc_{i,j}$ 表示第 i 条车道在第 j 行锚点上的位置, 计算方式为:

$$Loc_{i,j} = \sum_{k=1}^w k \cdot Prob_{i,j,k} \quad (6)$$

$Prob_{i,j,k}$ 为经过 softmax 后的在第 i 条车道的第 j 个行锚中, 第 k 个单元存在车道线的概率。 w 表示每个行锚所包含的单元格数量。最终的结构损失函数定义为:

$$L_{str} = L_{sim} + \lambda L_{shp} \quad (7)$$

其中 λ 为超参数, 用于调节相似性损失与形状损失两部分在整体结构损失中的相对权重。训练阶段启用的辅助分支本质上是一个分割任务, 其损失函数采用交叉熵损失。设 \hat{Y} 为预测得到的分割结果, Y 为对应的真实分割标签, 则有:

$$L_{seg} = L_{CE}(\hat{Y}, Y) \quad (8)$$

综合分类损失、结构损失以及分割损失后, 整体损失函数为:

$$L_{total} = L_{cls} + \alpha L_{str} + \beta L_{seg} \quad (9)$$

其中, α 和 β 分别为结构损失和分割损失在总损失中的权重系数。

3. 实验验证

3.1. CULane 数据集概述

在当前车道线检测领域的研究工作中, 常用的公开数据集主要包括 CULane、Tusimple 和 CurveLanes。为了评估算法在复杂场景中的表现, 重点选用了图像数量更大、并且对多种复杂场景进行了精细划分的 CULane 数据集。

CULane 数据集[3]由香港中文大学的研究人员在 SCNN 相关工作中构建, 数据采集地点位于北京市。研究团队在六辆不同车型上安装车载摄像设备, 连续采集超过 55 小时的行驶视频, 并从中抽取 133,235 帧图像作为数据样本。这些图像覆盖了多种道路类型, 包括城市道路、高速公路以及乡村道路等, 同时还涉及不同时段、各类天气条件和多样的交通流状态, 从而为模型在复杂环境下的鲁棒性训练提供了良好的数据基础。

全部图像的分辨率统一为 590×1640 。每张图像最多标注四条车道线, 一般为当前行驶车道的左右边界线及其两侧相邻车道线。为了降低将对向车道误识别为本侧车道的风险, 数据集中未对对向车道线进行标注。整个数据集被划分为三部分: 训练集 88,880 张、验证集 9675 张以及测试集 34,680 张。测试集在此基础上根据不同路况与光照条件进一步划分为九类典型场景, 用于评估模型在各类复杂环境下的

性能表现, 不同场景的示例图见图 4。



Figure 4. CULane example pictures of different scenarios
图 4. CULane 不同场景示例图

3.2. 实验细节

本文的实验环境为: Ubuntu 20.04 操作系统、Pytorch 2.0.0 框架、Intel Core i5-13600KF 以及 GeForce RTX 4070(12GB)。

为检验本文方法的有效性, 设计了与 SCNN、RESA、ERFNet 以及 UFLD 等代表性车道线检测算法的对比试验。SCNN 选用 VGG16 作为主干网络, RESA 以 ResNet-50 为主干网络, UFLD 则基于 ResNet-152 搭建。表 1 展示了各方法在 CULane 数据集上的整体 F1-score 与帧率(FPS)表现。

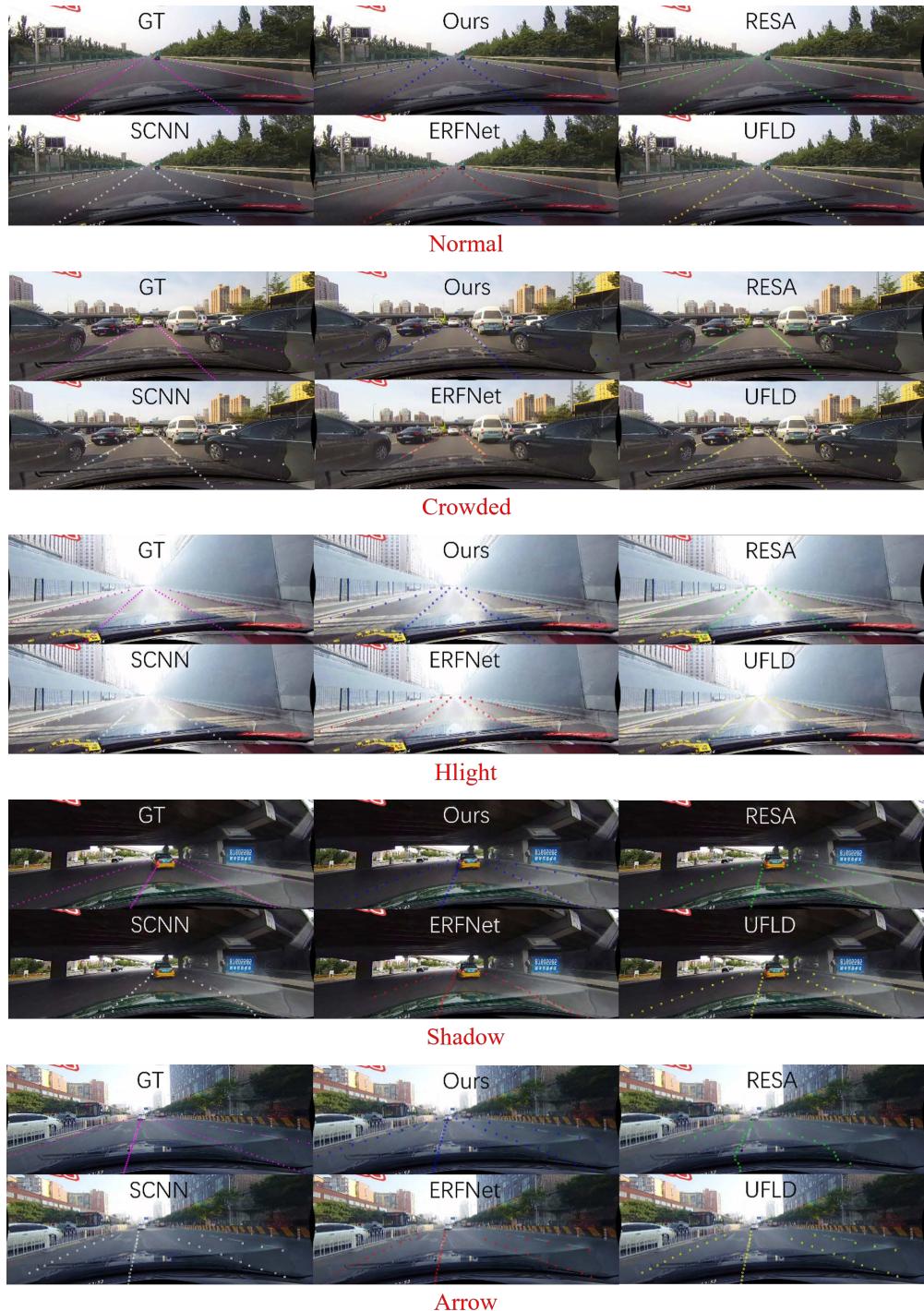
Table 1. Comparison of experimental evaluation results for different network models
表 1. 不同网络模型对比实验评估结果

模型	F1-score (%)	FPS
SCNN-VGG16	68.48	11.25
RESA-ResNet-50	72.88	67.64
ERFNet	71.52	73.86
UFLD-ResNet-152	71.63	104.13
Ours	72.51	211.42

由实验结果可知, 本文方法在 CULane 上的 F1-score 达到 72.51%, 推理速度为 211.42 FPS。与 SCNN-VGG16 相比, 无论精度还是速度均有明显提升。在 F1-score 指标上虽然比 RESA 略低 0.37 个百分点, 但帧率约为 RESA 的三倍, FPS 提高了 143.78。相对于 ERFNet, F1-score 提高约 1 个百分点, 检测速度快出 137.56 帧/秒; 与 UFLD-ResNet152 对比, F1-score 增加 0.88 个百分点, 帧率则提升了 107.29。综合上述结果可以看出, 在保持较高检测准确率的同时, 本文提出的方法在运行效率上也能够很好地满足车道线检测任务的实时性需求。

在普通直道或光照良好的场景中, 五种车道线检测方法整体表现接近, 难以区分明显优劣。然而, 对于基于语义分割的 ERFNet、SCNN 和 RESA, 由于缺乏显式的行锚约束且全卷积结构在感受野方面存在局限, 在多种复杂工况下都出现了不同程度的漏检和误检, 见图 5。具体来说: 在车辆遮挡车道线的情

况下, ERFNet 无法恢复被遮挡部分的车道线;在强光眩光场景中, SCNN 会出现车道线未被检测出的情况;在阴影较重的环境下, RESA 的车道位置预测出现一定偏移, ERFNet 与 SCNN 也都存在漏检问题, 其中 SCNN 更为明显。当地面存在箭头标识等干扰时, RESA 会产生较为严重的错误检测, 而 UFLD 和 ERFNet 对于远处车道线的检出也不够完整。在弯道路段, UFLD 的整体表现落后于其他算法;到了夜间场景, 除本文提出的方法外, 其余模型普遍在车道线定位上出现较大偏差。



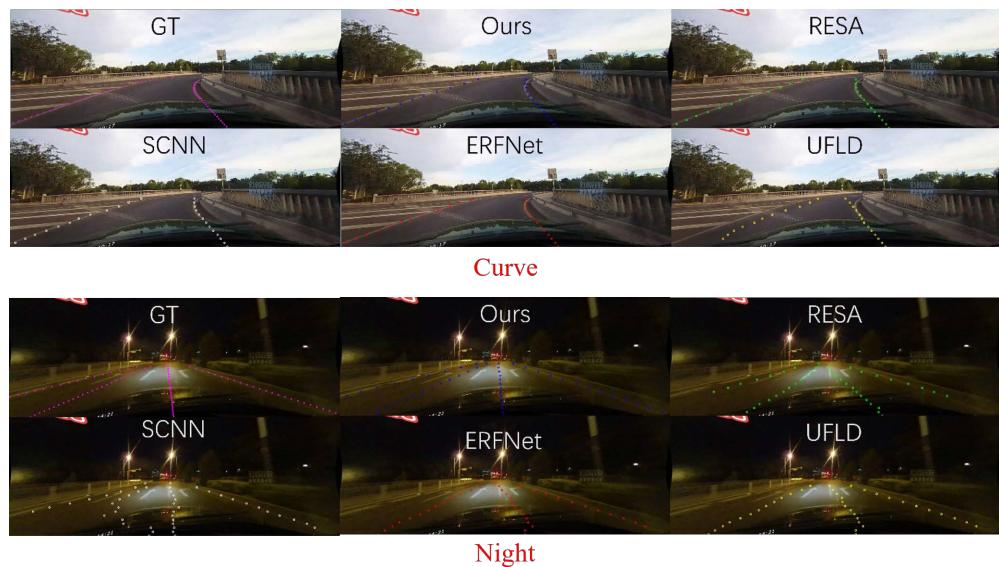


Figure 5. Visual comparison of different network models in various scenarios
图 5. 不同网络模型在不同场景的可视化对比

综合不同场景下的实验结果可以看出, 本文构建的基于行锚的车道线检测方法在复杂道路环境中仍能较为精准地识别出各条车道线, 且漏检与误检现象相对较少, 鲁棒性更强。

4. 总结

本文面向自动驾驶场景下对高精度与高实时车道线检测的需求, 在 UFLD 方法框架基础上提出了一种改进的轻量化车道线检测方法。方法以 ResNeXt-50 作为网络主干, 在保持与原始 ResNet 主干相近计算复杂度的前提下, 通过增加卷积分组基数增强特征表达能力, 并在浅层和深层特征图后分别插入坐标注意力模块, 使模型在保留全局感受野的同时, 更充分利用车道边缘与纹理等局部信息, 从而改善长距离连续性建模效果。在损失设计上, 联合分类损失、结构损失与辅助分割损失, 从平滑性与刚性两方面约束车道形态。基于 CULane 数据集的实验结果表明, 所提方法在保持 211.42 FPS 高帧率的同时, F1-score 达到 72.51%, 在遮挡、高光、阴影、弯道和夜间等复杂场景中, 定性对比显示该方法的漏检与误检情况少于其他对比算法。由此可见, 所提方法在保证实时性的同时提供了具有竞争力的检测性能, 适合在复杂道路环境的车道线感知模块中部署。

参考文献

- [1] Kim, K.B. and Song, D.H. (2017) Real Time Road Lane Detection with RANSAC and HSV Color Transformation. *Journal of Information & Communication Convergence Engineering*, **15**, 187-192.
- [2] 王文奇. 基于数字图像处理的车道线检测算法研究[D]: [硕士学位论文]. 烟台: 烟台大学, 2018.
- [3] Pan, X., Shi, J., Luo, P., Wang, X. and Tang, X. (2018) Spatial as Deep: Spatial CNN for Traffic Scene Understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**, 7276-7283. <https://doi.org/10.1609/aaai.v32i1.12301>
- [4] Neven, D., Brabandere, B.D., Georgoulis, S., Proesmans, M. and Gool, L.V. (2018) Towards End-to-End Lane Detection: An Instance Segmentation Approach. 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, 26-30 June 2018, 286-291. <https://doi.org/10.1109/ivs.2018.8500547>
- [5] Tabelini, L., Berriel, R., Paixao, T.M., Badue, C., De Souza, A.F. and Oliveira-Santos, T. (2021) PolyLaneNet: Lane Estimation via Deep Polynomial Regression. 2020 25th International Conference on Pattern Recognition (ICPR), Milan, 10-15 January 2021, 6150-6156. <https://doi.org/10.1109/icpr48806.2021.9412265>
- [6] Qu, Z., Jin, H., Zhou, Y., Yang, Z. and Zhang, W. (2021) Focus on Local: Detecting Lane Marker from Bottom up via

- Key Point. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, 20-25 June 2021, 14117-14125. <https://doi.org/10.1109/cvpr46437.2021.01390>
- [7] Garnett, N., Cohen, R., Pe'er, T., Lahav, R. and Levi, D. (2019) 3D-LaneNet: End-to-End 3D Multiple Lane Detection. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, 27 October-2 November 2019, 2921-2930. <https://doi.org/10.1109/iccv.2019.00301>
- [8] Qin, Z., Wang, H. and Li, X. (2020) Ultra Fast Structure-Aware Deep Lane Detection. In: *Lecture Notes in Computer Science*, Springer, 276-291. https://doi.org/10.1007/978-3-030-58586-0_17
- [9] Xie, S., Girshick, R., Dollar, P., Tu, Z. and He, K. (2017) Aggregated Residual Transformations for Deep Neural Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 21-26 July 2017, 5987-5995. <https://doi.org/10.1109/cvpr.2017.634>
- [10] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, 19-25 June 2021, 13713-13722.