

DOI: 10.19650/j.cnki.cjsi.J2514147

基于掩码孪生网络与幂调节损失的 恶劣场景车道线检测方法^{*}

邹亮¹, 郭泽沛¹, 李颖娜², 鞠进军³, 雷萌¹

(1. 中国矿业大学信息与控制工程学院 徐州 221116; 2. 唐山学院新材料与化学工程学院 唐山 063000;
3. 陆军工程大学训练基地工程装备系 徐州 221000)

摘要:车道线检测是自动驾驶感知系统的核心任务,在复杂交通环境下具有重要的应用价值。尽管现有方法在常规条件下已取得较好检测效果,但在光照不足、逆光、大雾、雨雪等恶劣场景中车道线检测仍面临诸如模糊、断裂与遮挡等挑战。为提升恶劣场景下的车道线检测性能,基于ADNet框架提出了 α -SimADNet检测网络。该模型基于ADNet网络实现锚点提取与参数回归,通过引入无负样本对比学习与具有交替优化策略的掩码孪生网络,增强主干网络的特征判别力与环境适应性。这些改进使模型在不增加推理计算开销的前提下,显著提升了恶劣条件下的特征表征能力。此外,针对传统IoU损失在回归困难样本时梯度响应不足的问题,设计了具有幂调节机制的 α -GLIoU损失函数,以增强模型对断裂与遮挡车道线的拟合能力。为全面评估所提方法的性能,构建了一个面向恶劣环境的高质量车道线检测数据集HardLane-F100,涵盖106段视频、10 600帧图像,有效缓解了当前公开数据集中极端环境样本覆盖不足的问题。实验结果表明, α -SimADNet在HardLane-F100数据集上的F1@0.5得分达到83.2%,分别较主流方法ADNet与RVLD提升2.7%和1.2%。在更严格的F1@0.7指标下得分60.9%,较ADNet与RVLD分别提升3.8%和3.2%,该方法在多种挑战性场景下均表现出更优性能,充分说明了其在恶劣场景下的有效性。

关键词: 车道线检测;掩码孪生网络;对比学习;恶劣场景; α -SimADNet

中图分类号: TH89 TP183 **文献标识码:** A **国家标准学科分类代码:** 510.40 580.99

Robust lane detection in challenging scenarios using a masked siamese network with power-modulated loss

Zou Liang¹, Guo Zepei¹, Li Yingna², Ju Jinjun³, Lei Meng¹

(1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China;
2. School of New Materials and Chemical Engineering, Tangshan University, Tangshan 063000, China;
3. Department of Engineering Equipment, Train Base of Army Engineering University of PLA, Xuzhou 221000, China)

Abstract: Lane detection is a core task in autonomous driving perception systems, holding significant application value in complex traffic environments. While existing methods perform well under normal conditions, lane detection still faces challenges such as blurriness, disconnection, and occlusion in adverse scenarios like low light, backlighting, heavy fog, rain, and snow. To improve lane detection performance in these harsh conditions, this paper proposes the α -SimADNet detection network, built upon the ADNet framework. This model performs anchor point extraction and parameter regression using ADNet, while enhancing the backbone network's feature discrimination and environmental adaptability by introducing negative sample contrastive learning and a mask twin network with an alternating optimization strategy. These enhancements significantly improve the model's feature representation capabilities in challenging environments, without increasing computational overhead during inference. Additionally, to address the insufficient gradient response from traditional IoU loss in the regression of difficult samples, we introduce the power-adjusted α -GLIoU loss function to improve the model's ability to fit broken and occluded lane lines. To thoroughly assess the proposed method's performance, we constructed a high-quality lane detection dataset, HardLane-F100, focused on harsh environments, which includes 106 video segments and 10 600 image

frames. This dataset effectively mitigates the current public datasets' lack of extreme environmental samples. Experimental results show that α -SimADNet achieves an F1@0.5 score of 83.2% on the HardLane-F100 dataset, outperforming mainstream methods ADNet and RVLD by 2.7% and 1.2%, respectively. Under the more stringent F1@0.7 metric, it scores 60.9%, improving by 3.8% and 3.2% compared to ADNet and RVLD, respectively. This method demonstrates superior performance across various challenging scenarios, fully proving its effectiveness in harsh environments.

Keywords: lane line detection; mask twin networks; contrastive learning; challenging scenarios; α -SimADNet

0 引言

随着自动驾驶技术的快速发展,车辆对环境感知能力的要求日益提高。作为环境感知系统的重要组成部分,车道线检测在路径规划、车道保持以及决策控制等任务中发挥着关键作用。因此,车道线检测的精度与鲁棒性已成为保障自动驾驶系统安全与稳定运行的核心因素之一^[1-2]。

为满足自动驾驶技术对高精度、高可靠感知能力的需求,近年来研究者们提出了大量基于深度学习的车道线检测策略,并在多个公开数据集上展示出良好的性能与应用前景^[3-5]。现有主流车道线检测方法可分为4类:基于关键点的方法^[6]、基于参数化建模的方法^[7]、基于分割的方法^[8]和基于锚点的方法^[9]。其中,基于关键点的方法虽具备一定的几何表达能力,但在建模车道线的长距离拓扑关系方面存在局限,而参数化建模方法通常依赖于预设的车道数量与形态,缺乏灵活性,对复杂道路环境的适应性不足。相比之下,基于语义分割和锚点回归的检测方法能够同时建模全局拓扑关系与局部细节特征,摆脱了对几何先验的刚性依赖,具备更强的环境适应能力,现已逐渐成为车道线检测的研究热点及重点发展方向^[10]。

基于语义分割的车道线检测方法将任务转化为像素级语义识别,通过提取车道区域的分割掩码实现对车道的检测。Yang等^[11]在AtrousFormer(atrous transformer)中设计了双分支结构,分别通过交替行列采样捕获全局拓扑信息,并在局部路径中引入矩形区域约束以增强空间表达能力。Zhao等^[12]提出基于水平与垂直多头注意力的结构,实现行列信息交叉融合,并设计自注意力蒸馏机制,引导低层感知模块学习高层语义,显著提升复杂环境下的检测表现。Hosseini等^[13]结合亲和场机制与二值分割策略,有效增强对车道边界的表征能力,在保持结构完整性的同时具备较强的泛化性能。

为提升车道线检测的效率与结构一致性建模能力,基于锚点回归的方法通常在特征图中预设规则锚点,通过学习其与真实车道线之间的相对偏移实现车道结构的快速拟合,从而避免逐像素分割的高计算成本。Zheng等^[14]提出的CLRNet(cross layer refinement network)采用

跨层特征融合策略,利用高阶语义引导实现车道线的粗定位,并引入面向结构对齐的线交并比(line intersection over union, LIoU)损失函数以提升几何拟合精度。Liu等^[15]提出超锚点技术,结合可变形锚点结构提供车道点的粗估计,并引入多级感知特征聚合模块,有效提升了锚点表达能力与预测精度。Xiao等^[16]将传统锚点解构为起始点及角度热力图形式,解决了锚点灵活性受限的问题,并配合提出的通用车道线交并比(general lane intersection over union, GLIoU)损失以加强对偏差较大的预测曲线的约束,在恶劣场景下展现出优越性能。

尽管上述研究取得了显著进展,但多数方法仍主要依赖于结构清晰、环境理想的公开数据集进行训练与评估,如CULane^[17]、TuSimple、VIL-100^[18](video instance lane detection-100)等。这些方法在特征提取阶段普遍依赖于局部纹理与边缘线索,容易受到背景干扰或图像质量退化的影响,导致性能下降^[19]。此外,传统交并比(intersection over union, IoU)类损失函数在处理断裂、遮挡等困难样本时,其梯度反馈主要来源于重叠区域,对定位误差较大的样本梯度反馈不足,限制了模型的收敛效率和对困难样本的优化能力。因此,如何增强模型的特征表征能力,成为恶劣场景车道线检测的关键问题之一。

与此相关,近年来对比学习(contrastive learning, CL)与孪生网络在图像识别与度量学习任务中展现了出色的特征表征能力,尤其在无监督学习领域引起了广泛关注^[20]。然而,这类方法大多面向分类或匹配任务设计,缺乏针对结构化回归任务(如车道线检测)的适配机制,在训练稳定性与推理效率方面仍存在诸多挑战。

为解决上述问题,本文结合对比学习思想提出了一种面向恶劣场景的鲁棒车道线检测网络幂相似性锚点分解网络(alpha-similarity anchor decomposition network, α -SimADNet)。该网络融合锚点回归主干与掩码引导的孪生特征增强模块,利用无负样本对比学习强化车道结构感知能力,同时引入幂通用车道线交并比(alpha-general lane intersection over union, α -GLIoU)损失以提升模型对困难样本的优化效率和梯度响应强度。本文的主要贡献为:

- 1) 提出了一种仅依赖正样本的掩码孪生特征增强网络,融合锚点回归主干构建双分支检测框架 α -SimADNet。该模块通过掩码引导机制生成语义一致的正样本对,并结合交替优化策略,有效引导特征提取网络学

习车道线的空间位置与上下文关系,显著提升了模型在模糊、遮挡等复杂场景下的鲁棒性,同时保持较低的计算开销。

2) 针对 GLIoU 损失在低 IoU 样本上的优化瓶颈,设计了具备幂调节机制的 α -GLIoU 损失函数。该方法引入幂调节机制动态调整梯度响应强度,增强了模型对困难样本的优化能力。

3) 构建了一个面向恶劣场景的高质量车道线检测数据集 HardLane-F100,包含 106 组视频与 10 600 帧图像,涵盖低照、雨雪、大雾、逆光等典型复杂条件,并提供实例级标注。在 HardLane-F100 数据集上的实验表明, α -SimADNet 在 F1 分数、鲁棒性与精度等关键指标上全面优于当前主流方法,验证了本文方法在恶劣环境下的检测能力与工程适用性。

1 车道线检测方法

1.1 网络整体架构

本文提出的 α -SimADNet 检测框架主要由掩码孪生网络和锚点分解网络 (anchor decomposition network, ADNet) 两部分组成。如图 1 所示,输入图像经数据增强后,保留原图并生成叠加随机网格掩码的掩码图,两者分别输入至两个结构对称的孪生分支进行特征提取。两个孪生分支共享残差神经网络 (residual neural network, ResNet-101) 主干,并输出各自的深层语义特征表示。随后,各分支提取到的特征经由映射器变换后,输入至对应的预测器进行特征预测,其中⑤用于度量映射特征与预测特征之间的相似性。

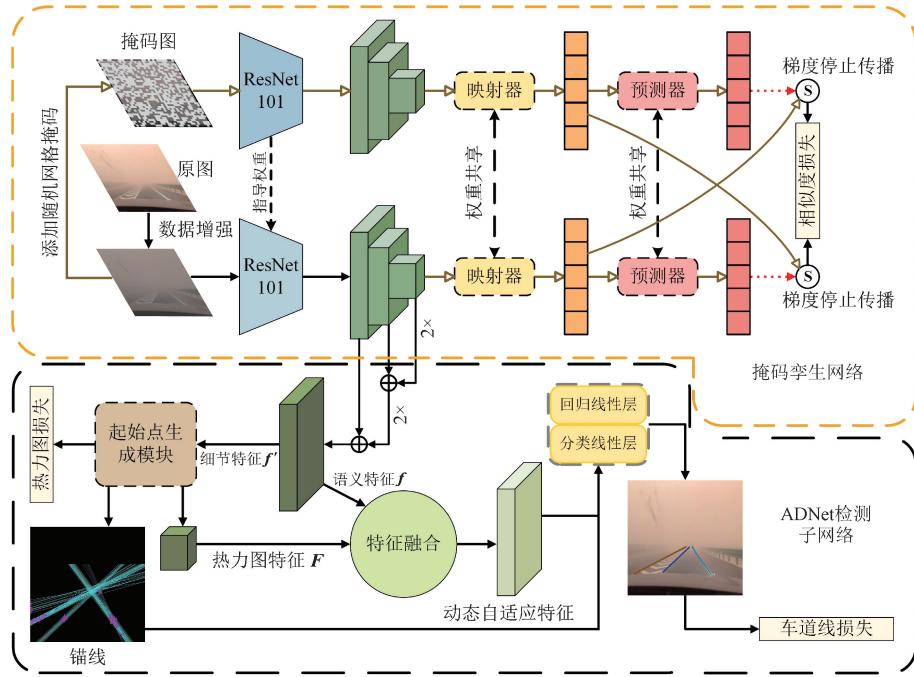


图 1 α -SimADNet 网络整体框架

Fig. 1 Overall framework of the α -SimADNet network

在训练阶段,孪生分支采用交替优化策略:将一个分支的映射特征与对向分支的预测特征计算余弦相似度,而该分支自身预测器的梯度传播被阻断,使对应的映射器无法从自身通路中获得梯度反馈,只能依赖对向通路的预测器进行优化。该非对称式训练机制打破了参数更新对称性,可有效缓解表征退化问题,提升了模型的判别能力,其中预测器之后虚线代表反向传播梯度的停止。在推理阶段,仅保留原始图的输入支路,即将图 1 中下方支路实心箭头部分的网络结构用于前向传播,无额外推理成本。

经孪生网络优化后的特征图被输入至 ADNet 检测网络,以完成车道线的几何建模与检测任务。该网络首先使用特征金字塔 (feature pyramid network, FPN) 融合多尺度语义特征,并利用起始点生成模块提取车道线的结构先验与几何分布信息。随后,网络通过并行的分类分支和回归分支,分别预测车道线的类别置信度与锚点坐标信息。为进一步提升车道线回归的精度,在回归过程中引入改进的 α -GLIoU 损失函数。该损失在 GLIoU 损失的基础上引入幂运算调节项,增加了高 IoU 样本和空间偏离样本的梯度,从而增强模型在复杂场景下对车道线空

间分布的拟合能力。整体网络架构在保证检测精度的同时,显著增强了面对恶劣环境下模糊、遮挡和断裂车道线的识别鲁棒性。

1.2 掩码孪生网络

为提升模型在恶劣场景下的特征表征能力,本文引入对比学习范式,引导模型更有效地区分车道线与背景干扰。然而,现有对比学习方法通常依赖大量正负样本对,不仅训练效率较低,且对批量大小高度敏感,尤其是处理高分辨率图像时,难以在小批量训练条件下实现稳定优化,限制了其在车道检测任务中的实际应用^[21]。

针对上述问题,设计了一种无负样本的掩码孪生网络。借鉴 Chen 等^[22]提出的孪生网络架构设计,该网络由两个共享主干参数的对称分支组成,摒弃了传统方法中使用动量编码器更新分支参数的策略,从而简化了模型结构并提升了训练稳定性。

网络以同一图像的两种增强视图为输入:其中一个视图采用常规的数据增强方法,另一个视图则通过引入随机掩码遮蔽部分区域,用于模拟局部结构缺失。通过在部分区域施加遮蔽扰动,网络被迫在信息不完整的条件下学习保持全局语义一致性的表征,引导模型关注车道线中具有代表性的结构特征,抑制非关键区域的干扰。

如图 2 所示,掩码孪生网络中的映射器与预测器均采用多层前馈神经网络(feedforward neural network,FFN)构建。每个模块由若干全连接层、归一化层及非线性激活函数叠加组成,实现逐层特征变换与语义提取。映射器将其输入特征 $v_{\text{flatten}} \in \mathbb{R}^{B \times (C \times H \times W)}$ 投影至高维的潜在表示空间;预测器则基于该高维嵌入,对另一视角下的数据增强样本进行特征预测。

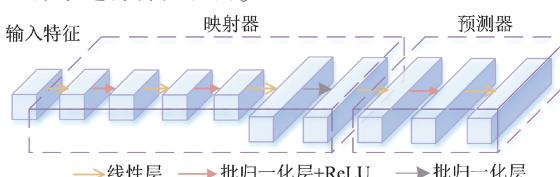


图 2 映射器与预测器结构

Fig. 2 Mapping and prediction structure diagram

由于孪生网络在训练过程中通常仅依赖正样本,即通过最小化同一样本不同视角下的特征表示差异来更新梯度,因此模型极易陷入坍缩解(collapsed solution)问题^[23],即所有样本在潜在空间中的表示趋于一致,这种表示退化现象会显著削弱主干网络对判别性特征的建模能力。为缓解这一问题,引入交替优化策略,在图 1 的双分支掩码孪生网络中,通过在其中一条通路的预测器之后施加梯度停止操作,使得该通路的映射器无法从自身的预测器获得梯度反馈,只能依赖对向通路的预测器进

行梯度优化。这种非对称性的设计打破了双分支的参数共享与更新对称性,提升了网络的训练稳定性与判别能力。此外,本文所设计的掩码机制在每个训练轮次中均以随机方式生成不同结构的网格遮罩图,用以模拟遮挡、模糊或低光照等破坏性因素所造成的视觉缺陷,该设计避免了传统对比学习对负样本构造与大规模批次的依赖,从而在降低训练资源消耗的同时实现对环境扰动的鲁棒模拟^[24]。掩码孪生网络的训练流程为:

- 1)一个批次中每组恶劣条件数据集的图像 m_i ,其中 $i = 1, 2, \dots$
- 2)对图像 m_i 随机使用数据增强 T' ,得到增强后的图像 m'_i
- 3)对图像 m'_i 应用随机掩码 T'' ,得到添加掩码的图像 m''_i
- 4)通过主干网络和映射器从 m'_i 和 m''_i 中提取特征,得到 z'_i 和 z''_i
- 5)使用预测器,通过特征 z'_i 预测数据增强 T'' 的分布期望 $p'_{i''}$ 。同样地,通过特征 z''_i 获得数据增强 T' 的分布期望 $p''_{i''}$
- 6)计算每条通路的分布期望和对向通路映射特征之间的余弦相似度损失:

$$\mathcal{L}_{\text{sim}} = 1 - \frac{1}{2} \frac{\mathbf{p}'_{i''} \cdot \text{detach}(\mathbf{z}''_i)}{\|\mathbf{p}'_{i''}\| \|\text{detach}(\mathbf{z}''_i)\|} - \frac{1}{2} \frac{\mathbf{p}''_{i''} \cdot \text{detach}(\mathbf{z}'_i)}{\|\mathbf{p}''_{i''}\| \|\text{detach}(\mathbf{z}'_i)\|} \quad (1)$$

其中, $\text{detach}(\mathbf{z})$ 表示停止对特征 \mathbf{z}_i 的梯度计算。

7)反向传播并更新梯度。

1.3 起始点生成模块

图 3 展示了起始点生成模块的基本结构。该模块以来自浅层的细节特征为输入,通过两组并行的卷积-线性网络层,分别输出起始点热力图和角度热力图。

图 3 中,起始点热力图表示每个像素作为起始点的概率,而角度热力图用于估计每个像素作为起始点时车道线的方向。起始点热力图和角度热力图分别被输入到两个不同的分支中。在热力图分支中,起始点热力图和角度热力图作为热力图特征进行特征融合,并与热力图真值共同参与损失计算。起始点热力图的损失函数如式(2)、(3)所示。

$$\mathcal{L}_{pts} = \frac{-1}{H'_f \times W'_f} \times \sum_{xy} \begin{cases} (1 - A_{pred}^{pts})^\gamma \log(A_{gt}^{pts}), & A_{gt}^{pts} = 1 \\ (1 - A_{gt}^{pts})^\beta (A_{pred}^{pts})^\gamma \log(1 - A_{pred}^{pts}), & \text{其他} \end{cases} \quad (2)$$

$$A_{gt}^{pts}(x, y) = \exp\left(-\frac{(x - x_{gt}^{start})^2 + (y - y_{gt}^{start})^2}{2\sigma^2}\right), \quad (3)$$

$$A_{gt} \in R^{H'_f \times W'_f}$$

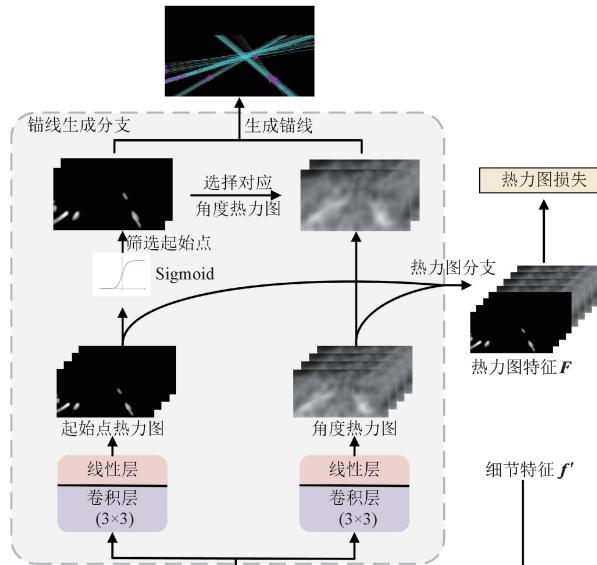


图 3 起始点生成模块示意图

Fig. 3 Schematic diagram of the starting point generation module

式中: $H_{f'}$ 、 $W_{f'}$ 为细节特征 f' 的高宽; A_{gt}^{pts} 为起始点监督热力图; x_{gt}^{start} 、 y_{gt}^{start} 为起始点的真值坐标; x 、 y 是真值起始点热力图上像素点的坐标; γ 、 β 为焦点损失 (focal loss) 的超参数; σ 为高斯分布的标准差参数。角度热力图的损失函数如式(4)、(5)所示。

$$\mathcal{L}_\theta = \frac{1}{H_{f'} \times W_{f'}} \sum_{xy} |A_{pred}^\theta - A_{gt}^\theta| \quad (4)$$

$$A_{gt}^\theta(x, y) = l(A_{gt}^{pts}(x, y) > t_\theta) \cdot \theta(x_{gt}^{start}, y_{gt}^{start}) \quad (5)$$

式中: $l(\cdot)$ 表示指示函数; $\theta(x, y)$ 表示真值中坐标为 (x, y) 的点对应的角度; t_θ 为阈值超参数。通过热力图分支引导网络学习不同场景下的车道线的起始位置与方向信息。

在锚线生成分支中, 起始点热力图通过 Sigmoid 函数归一化后, 产生全图每个像素点作为车道线起点的概率。之后, 按概率值降序选择前 n 个坐标点作为可能的起始点, 并从角度热力图中检索对应位置的角度预测值。最后, 依据这 n 个起始点与角度生成初始锚点, 用于车道线的预测, 作为后续车道线回归的几何先验输入。该模块打破了传统方法对图像边缘起始点的依赖, 使网络具备在任意图像区域灵活定位车道线起点的能力, 特别是在图像边缘存在遮挡、模糊或视角畸变等情况下, 仍能保持锚点生成的稳定性与结构准确性, 为整条车道线的建模提供了关键几何支持。

1.4 α -GLIoU 损失

GLIoU^[16] 损失通过引入惩罚项使偏离较大的预测结果获得更高的损失值, 从而增强了正负样本之间的判别能力。然而, 在恶劣场景中, 由于训练阶段难以获得完整

的监督信号, 网络往往会输出大量离群负样本。尽管 GLIoU 在损失设计上提升了对负样本的惩罚, 但其优化机制未能针对梯度反馈进行有效调控, 导致网络对困难样本的过拟合。为此, 本文在 GLIoU 基础上引入幂调节机制^[25], 提出 α -GLIoU 损失函数, 通过幂参数动态调整梯度响应强度, 以适应不同难度样本的优化需求并提升整体回归精度, 计算公式如式(6)~(8)所示。

$$\mathcal{L}_{\alpha-GLIoU} = 1 - \left(\sum_{i=1}^K d_i^o / \sum_{i=1}^K d_i^u \right) \cdot \left(\left| \sum_{i=1}^K d_i^o / \sum_{i=1}^K d_i^u \right| \right)^{\alpha-1} \quad (6)$$

$$\mathcal{L}_{\alpha-GLIoU} = \mathcal{L}_{\alpha-GLIoU} + \frac{\sum_{i=1}^K \max(d_i^u - 4e, 0)}{\sum_{i=1}^K d_i^u} \quad (7)$$

$$\begin{aligned} d_i^o &= \min(x_i^p + e, x_i^g + e) - \max(x_i^p - e, x_i^g - e) \\ d_i^u &= \max(x_i^p + e, x_i^g + e) - \min(x_i^p - e, x_i^g - e) \end{aligned} \quad (8)$$

其中各符号含义如图 4 所示, 具体定义为: K 表示与 y 轴平行的采样线数量, 每条车道线由 K 个与 y 轴平行线相交的坐标点拟合; e 为计算交集或并集时, 每个锚点左右扩展的长度; d_i^u 为车道线检测中两个线段间的并集; d_i^o 为车道线检测中两个线段间的交集; d_i^u 可能为负值。

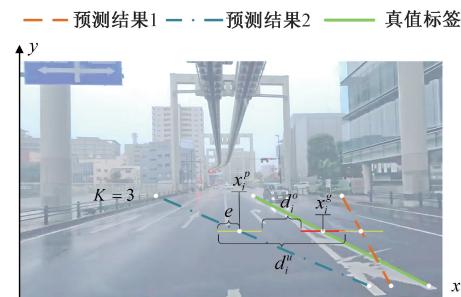


图 4 车道线检测 IoU 计算示意图

Fig. 4 Diagram of lane line detection IoU calculation

由式(6)可得, 当 $\alpha > 1$ 时, 高 IoU 样本对应的梯度响应显著增强, 有助于避免模型在训练初期对简单样本的过早收敛, 从而完善模型在恶劣条件下对模糊边缘的微调, 使模型具备高精度预测能力^[26]。此外, 在式(7)中引入 GLIoU 的原始惩罚项, 与幂运算调节项协同作用于梯度更新过程, 对偏差较大的负样本提供额外的梯度反馈, 以提升模型对低 IoU 样本的识别与边界优化能力。

图 5 展示了幂参数 α 的不同取值下, α -GLIoU 损失函数及其对应梯度随 IoU 变化的趋势。尽管 $\mathcal{L}_{\alpha-GLIoU}$ 在表达形式上与 \mathcal{L}_{GLIoU} 存在差异, 但当 $\alpha = 1$ 时 $\mathcal{L}_{\alpha-GLIoU}$ 可退化为 \mathcal{L}_{GLIoU} , 两者具有相似的数学性质, 包括单调递减性、极值一致性与符号敏感性。在车道线检测中, 对正负方向的响应差异尤为关键。考虑到恶劣场景车道线

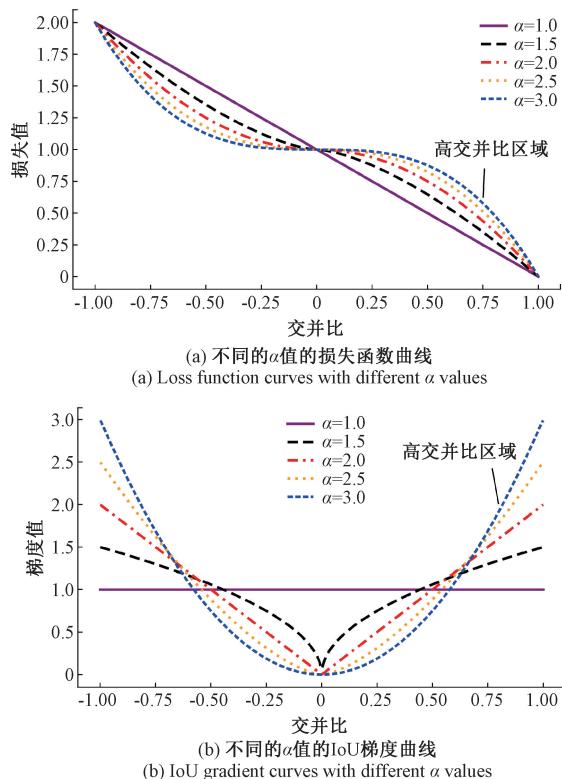


图 5 α -GLIoU 损失在不同 α 值下的损失函数与梯度
Fig. 5 Loss function and gradient plot of α -GLIoU
loss at different α values

检测中需要关注高质量样本的差异, $\mathcal{L}_{\alpha-GLIoU}$ 与 \mathcal{L}_{GLIoU} 相比引入了梯度动态缩放、非线性权重分配以及损失函数的凸性调整。具体体现在3个方面,即:

1) 梯度动态缩放: $\mathcal{L}_{\alpha-GLIoU}$ 在 IoU 接近-1 时,梯度放大为原来的 α 倍,从而加速模型对严重错误预测的修正。在实际的网络输出中,低 IoU 范围内的输出占比通常较大;而在 IoU 接近于 1 时,梯度同样被适度增强,有助于缓解模型陷入局部最优的风险,从而更好地对模糊边缘区域进行微调。

2) 非线性权重分配:通过幂函数与惩罚项联合构建非线性损失加权策略,在不同 IoU 区间实现差异化关注,显著提升对离群样本的识别能力。具体而言,通过 \mathcal{L}_{GLIoU} 的惩罚项放大 $\text{IoU} < 0$ 区间内的损失权重,提升模型对离群负样本的敏感性;而在 $\text{IoU} > 0$ 的区间,高次幂的引入有效增强了反向传播的权重。实现了权重的非线性重加权,提升模型的判别能力。

3) 损失函数的凸性调整:当 $\text{IoU} > 0$,损失函数的二阶导为负,函数呈凹性,梯度下降在接近正确预测时加速;当 $\text{IoU} < 0$ 时,二阶导为正,函数呈凸性,在极端错误的区域使模型快速优化离群负样本。

2 实验结果与分析

2.1 数据集介绍与构建

现有的车道线检测数据集主要分为两类:一类为合成数据集,如 Apollo Synthetic 和 Unity-based Lane Dataset,通过仿真方式快速生成多样化道路与光照环境,适用于模型预训练与评估;另一类为真实采集数据集,如 TuSimple、CULane^[17]、LLAMAS^[27] (unsupervised labeled lane markers dataset) 和 BDD100K^[28] (berkeley deepdrive 100 k),相较于合成数据集,其在天气、光照和交通场景上展现出更强的多样性。尽管现有车道线检测数据集在算法研究中发挥了重要作用,但其标注精度、恶劣场景占比等方面不足,限制了模型在恶劣场景下的性能评估。为弥补上述不足,本文构建了一个更具挑战性的车道线检测数据集——HardLane-F100,以真实视频为基础,覆盖多种复杂气象与光照条件,包括低照度、雨雪、大雾、逆光、反光、阴影、遮挡、弯道、十字路口等场景。数据集包含 106 段视频,每段视频通过 30 倍帧率下采样方式均匀抽取连续的 100 帧图像,总计 10 600 帧。为避免数据泄露,训练、验证与测试集以视频为单位进行划分,确保同一视频中的图像不会同时出现在不同子集中。按照 8:1:1 的比例将数据集划分为训练集、验证集与测试集,涵盖典型极端环境样本,如图 6 所示。



图 6 HardLane-F100 数据集

Fig. 6 HardLane-F100 dataset

表 1 对比了 HardLane-F100 与当前主流车道线检测数据集在多个维度上的差异。从车道线数量来看,HardLane-F100 单帧图像包含 3~7 条车道线,显著高于其他数据集,更加贴近实际城市道路中多车道、交叉路口等复杂交通场景。在恶劣场景占比上,该数据集包含超过 60% 的恶劣场景样本,远超当前主流数据集,显著提升了数据集的场景丰富度。此外,与多数仅提供静态帧的车道线检测数据集不同,HardLane-F100 具备时序连贯性,为引入时序建模机制提供了数据基础,从而进一步提升车道线检测在动态驾驶环境下的连续性与稳定性。综上所述,HardLane-F100 在车道线数量、标注质量与恶劣场景占比等方面具有明显优势,为恶劣场景下的车道线检测研究提供更高质量的评估基准。

表 1 TuSimple、CULane、LLAMAS、BDD100K 与 HardLane-F100 的对比概览

Table 1 Overview comparison of TuSimple, CULane, LLAMAS, BDD100K and HardLane-F100

数据集名称	TuSimple	CuLane ^[17]	LLAMAS ^[27]	BDD100K ^[28]	HardLane-F100
车道线数量	2~4 条	1~4 条	1~3 条	2~5 条	3~7 条
恶劣场景占比	10% 以下	35%~45%	-	15%~25%	60%~75%
图像分辨率	1 280×720	1 640×590	1 276×717	1 280×720	1 366×720
场景丰富度	丰富度低,以白天、高速公路为主	丰富度中等,存在遮挡、夜间等恶劣场景	丰富度低,少量雨夜场景	丰富度中等,以城市与高速道路为主	丰富度高,包含十字路口、沙尘暴、雨雾雪天、夜间等恶劣场景
标注类型	离散点	mask	非连续 mask	粗略折线	折线+实例级 mask
是否时序连贯	×	×	×	×	√

2.2 实验设置与评价指标

本文实验在一台配置为 Intel Xeon Platinum 8336C CPU 和 NVIDIA GeForce RTX 4090 GPU 的工作站上进行,操作系统为 Ubuntu 20.04, CUDA 版本为 11.3, 主要使用 Python3.8 和 PyTorch1.12.1 作为开发环境。网络结构采用 ResNet-101 作为主干, 输入图像尺寸统一为 800 pixels×320 pixels, 单张图像锚点候选数为 100。计算交集或并集时每个车道线与 y 轴平行线的交点扩展的长度 e 设置为 15, y 轴平行线的数量 K 设置为 72, 掩码孪生网络中掩码比设置为 40%, α -GLIoU 损失中幂参数 α 设置为 3。另外, 批量大小(batch size, BS) 设置为 24, 训练轮次(Epoch) 为 80, 优化器为 Adam 优化器, 初始学习率为 0.000 64。在孪生掩码网络中, 映射器的输出通道维度为 1 024 维, 预测器的输出通道维度设置为 512 维。

本文使用精确率(Precision)、召回率(Recall)和 F1 分数作为车道线检测的评价指标, 在预测锚点与真值标签之间计算了 IoU, 其中将 IoU 大于阈值的车道线结果定义为正样本^[29]。实验过程中 IoU 阈值分别设置为 0.5 和 0.7, 以检测在不同标准下的不同模型的车道线检测性能。召回率衡量真实车道线中有多少被模型成功检测; 精确率衡量模型预测为正的车道线中有多少为真实车道线; F1 分数是召回率和精确率的调和平均数, 能够综合评估模型在精确率和召回率之间的平衡情况, 计算公式如式(9)~(11)所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (9)$$

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

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

其中包含真正例(true positive, TP), 假正例(false positive, FP), 假负例(false negative, FN)。

2.3 消融实验

1) 模块有效性验证

为了充分验证本文所提各改进模块的有效性, 以 ADNet 为基线模型, 基于 HardLane-F100 数据集开展了模块消融实验, 结果如表 2 所示。

表 2 本文所提车道线检测方法消融实验结果

Table 2 Ablation experimental results of the lane line detection method proposed in this paper (%)

基线	掩码孪生网络	α -GLIoU 损失	Precision	Recall	F1@0.5
ADNet ^[16]	×	×	86.7	75.2	80.5
	√	×	87.6	77.2	82.1
	×	√	87.4	76.0	81.3
	√	√	89.2	78.1	83.2

引入掩码孪生网络后, 模型在精准率和召回率均有不同程度的提升, F1 分数由 80.5% 上升至 83.2%。此外, 本文设计的 α -GLIoU 损失, 通过提高计算损失时困难样本与车道线细节的比重, 同样使模型在恶劣条件下的检测性能实现了提升, F1 分数达到了 83.2%, 较基线提高了 0.8%。融合掩码孪生网络与 α -GLIoU 损失后, 本文方法达到了 89.2% 的精确率、78.1% 的召回率和 83.2% 的 F1 分数, 分别较基线提升 2.5%、2.9% 和 2.7%, 验证了该方法在恶劣场景下的性能优势。

2) 掩码比的影响

为验证掩码比设置的合理性, 对不同掩码比(即随机掩码在图像中所占比例)进行了实验对比, 结果如表 3 所示。当掩码比为 40% 时, 精确度为 87.6%, 召回率为 77.2%, F1 分数为 82.1%, 综合性能最优。过小的掩码比(例如 30%)难以充分模拟遮挡带来的信息缺失, 导致误检率升高; 而过大的掩码比(例如 50%)则会隐蔽重要信息、甚至可能覆盖小物体, 进而导致检测性能的下降。因此, 合理设定掩码比可在保持原始结构信息的同时增强模型对潜在特征的学习能力, 提升其在复杂环境下的适应性。

表3 不同掩码比下所提车道线检测方法的结果

Table 3 Results of the proposed lane line detection method under different mask ratios (%)

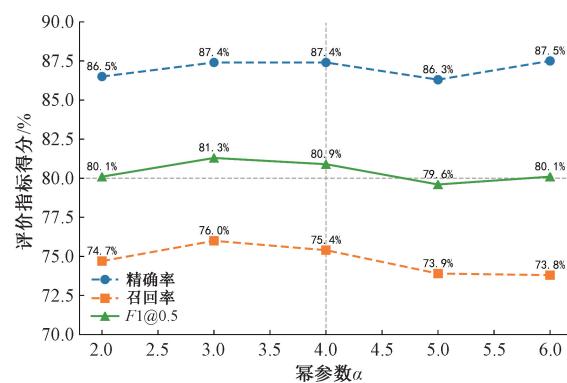
掩码比	Precision	Recall	F1@ 0.5
30	86.8	77.3	81.8
40	87.6	77.2	82.1
50	85.5	77.0	81.0

3) 幂参数 α 的影响

为分析 α -GLIoU 损失函数中幂参数取值对车道线回归性能的影响,本文将 α 设置为 2~6 进行对比实验,结果如图 7 所示。实验结果表明, $\alpha=3$ 时 F1 分数为 81.3%, 综合性能最好; $\alpha=4$ 时 F1 分数为 80.9%, 综合性能次之。尽管 $\alpha=6$ 时, 精确率较 $\alpha=3$ 时稍有提高, 但其召回率显著下降, 漏检率增加, 导致 F1 分数仍然较低。整体来看, $\alpha=3$ 能在提升困难样本学习能力的同时, 避免对易检样本梯度的过度抑制, 实现了精度与召回率之间的权衡。

2.4 对比实验与分析

基于 HardLane-F100 数据集, 本文与 5 种具有代表性的车道线检测方法进行了对比, 包括 HWLane^[12] (hw-transformer for lane detection)、MFIALane^[30] (multiscale feature information aggregator network for lane detection)、LaneTCA^[31] (lane detection with temporal context aggregation)、

图 7 不同 α 时所提车道线检测方法的结果Fig. 7 Results of lane line detection methods at different α

ADNet^[16] 与 RVLD^[32] (recursive video lane detection)。各方法在不同 IoU 阈值下的检测性能指标如表 4 所示。实验结果表明, 本文所提出的 α -SimADNet 模型在 IoU 阈值为 0.5 和 0.7 情况下, F1 得分相较于当前性能最优的车道线检测方法 RVLD 分别有 1.2% 和 3.2% 的提升。此外, 召回率在两个阈值下分别提高了 2.5% 和 3.4%。由此可见, 随着 IoU 阈值的提高, 本文方法相较于其他算法的性能优势进一步扩大, 从而说明本方法对车道细节的拟合能力更强。尽管在 IoU 阈值为 0.5 的情况下, RVLD 在精确率上略占优势, 但 α -SimADNet 在精确率与召回率之间实现了更优的平衡, 从而获得了更高的综合检测性能。

表4 各类车道线检测方法在不同评价指标下的性能比较

Table 4 Performance comparison of various lane line detection methods on different evaluation metrics

模型	Param/M	FLOPs/G	F1@ 0.5/%			F1@ 0.7/%		
			Precision	Recall	F1	Precision	Recall	F1
HWLane ^[12]	22.1	87.99	70.9	61.7	66.0	55.9	48.7	52.0
MFIALane ^[30]	21.4	85.22	73.2	60.8	66.5	56.9	47.2	51.6
LaneTCA ^[31]	25.2	130.51	89.3	71.3	79.3	61.2	49.4	54.7
ADNet ^[16]	43.3	41.20	86.7	75.2	80.5	61.4	53.3	57.1
RVLD ^[32]	25.3	134.00	89.7	75.6	82.0	62.5	53.7	57.7
α -SimADNet	43.3	41.20	89.2	78.1	83.2	65.2	57.1	60.9

表 5 展示了各方法在相同输入尺寸与硬件条件下的平均推理时间。尽管引入孪生掩码网络在训练阶段增加了一定的计算开销, 但在推理阶段, 本文提出的方法仍保持了较高的推理效率, 推理速度达到了 71.23 fps (14.04 ms/帧), 明显优于其他对比方法。

图 8 展示了各模型在 F1 得分、参数规模与推理速度三者之间的综合性能表现, 图形面积代表参数量。结果表明, 现有方法普遍存在模型复杂、推理缓慢的问题, 且在高 IoU 阈值下表现不佳。而 α -SimADNet 在保证效

表5 不同车道线检测方法的推理速度对比

Table 5 Comparison of inference speed of different lane detection methods

模型	FPS	推理速度/ms
HWLane ^[12]	34.52	28.97
MFIALane ^[30]	40.58	24.64
LaneTCA ^[31]	20.43	48.95
ADNet ^[16]	78.96	12.66
RVLD ^[32]	26.72	37.43
α -SimADNet	71.23	14.04

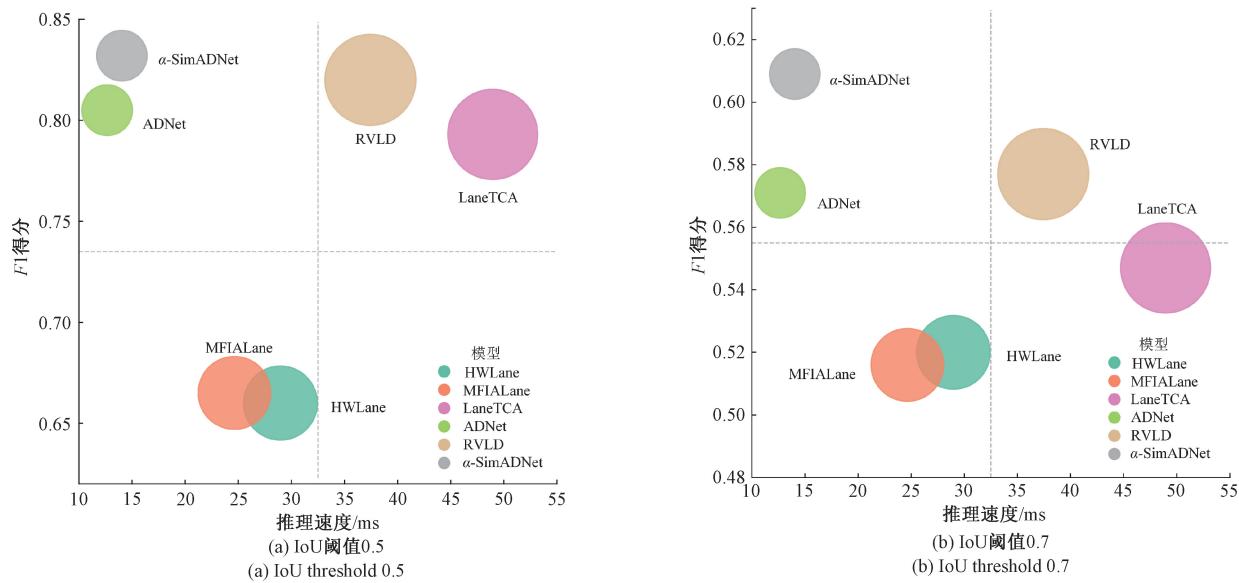


图 8 不同方法在恶劣场景下车道检测性能

Fig. 8 Lane detection performance of different methods under challenging scenarios

率的同时,在检测精度方面展现出更强的优势,尤其在IoU为0.7条件下,其对模糊边缘区域的精细建模能力得以凸显。

图9展示了各方法在典型恶劣场景下的检测可视化结果,具体包括弯道、低照度、多车道、遮挡、大雾和雨天6种恶劣场景。从可视化结果中可以观察到,现有方法

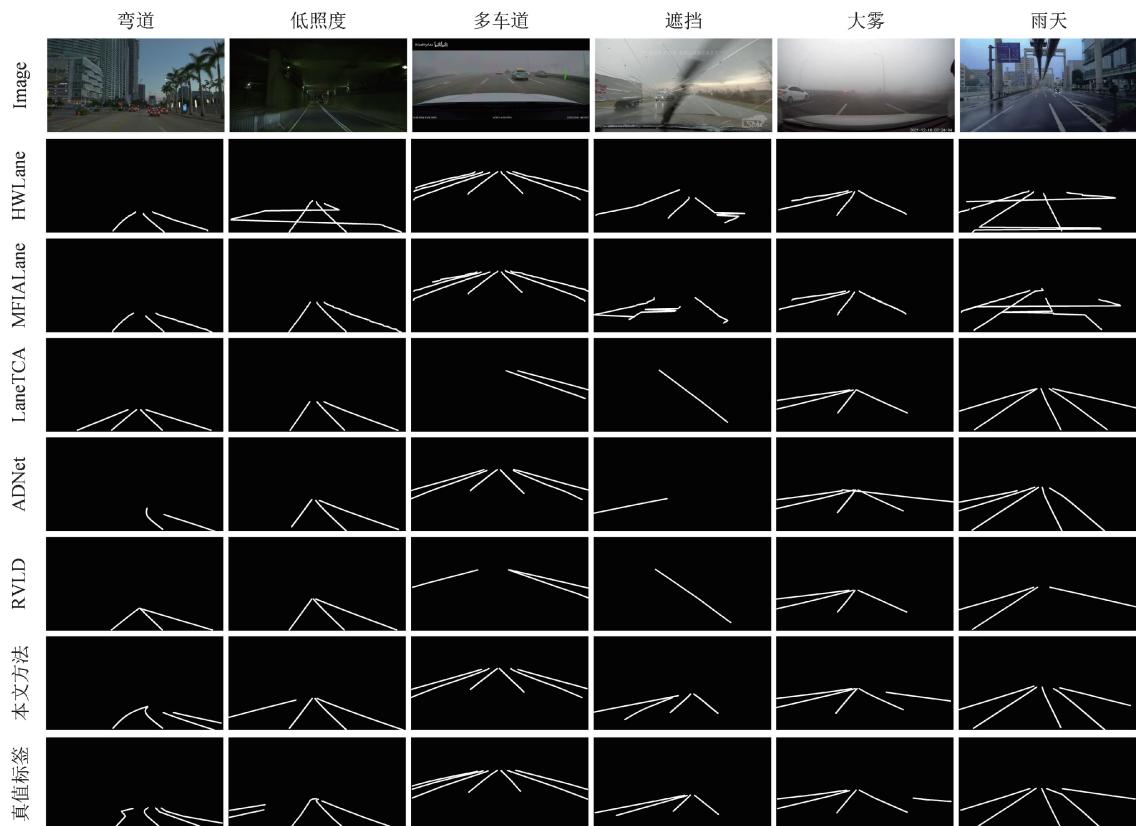


图 9 不同方法在恶劣场景下车道线检测结果的可视化示例

Fig. 9 Visualization of lane detection results using different methods in adverse conditions

在上述场景下普遍存在漏检或误检的问题,难以准确回归车道线结构。相比之下,本文提出的 α -SimADNet方法在各类恶劣场景下依然能够准确、完整地识别出图中所有车道线,验证了其在复杂场景中的鲁棒性与检测精度优势。

3 结 论

针对恶劣场景下车道线检测易受干扰、精度下降的问题,提出一种基于掩码孪生网络与幂调节IoU损失的检测框架 α -SimADNet。该方法以ADNet锚点检测网络为基础,通过结合无负样本的对比学习与交替优化策略构建掩码孪生模块,有效提升了特征提取网络在复杂环境下的表征能力与鲁棒性,同时保持了较高的推理效率。此外,针对传统IoU损失在回归困难样本时梯度响应不足的问题,设计了具备幂调节机制的 α -GLIoU损失函数,通过幂参数动态调节梯度反馈强度,增强了模型对断裂车道和模糊边缘的拟合能力,实现了更精细的边界建模。

针对当前公共数据集中极端环境样本覆盖不足的问题,构建了一个面向恶劣场景的高质量车道线检测数据集HardLane-F100,并在该数据集上与多种典型的车道线检测方法进行了对比。实验结果表明, α -SimADNet模型在IoU阈值为0.5与0.7的情况下,F1分数分别达到83.2%和60.9%,相较于ADNet和RVLD分别取得了2.7%与1.2%,3.8%与3.2%的提升,验证了所提方法在多种恶劣场景下的有效性。未来将进一步探索扩展该方法在视频车道线检测任务中的应用,结合时序信息提升其在动态场景中的连续性与稳定性,为自动驾驶感知系统提供更可靠的车道线检测方案。

参考文献

- [1] LIU F, LU Z H, LIN X K. Vision-based environmental perception for autonomous driving[J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2025, 239(1): 39-69.
- [2] 吴一全, 刘莉. 基于视觉的车道线检测方法研究进展[J]. 仪器仪表学报, 2019, 40(12): 92-109.
- [3] WU Y Q, LIU L. Research and development of the vision-based lane detection methods[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 92-109.
- [4] BEHRENDT K, SOUSSAN R. Unsupervised labeled lane markers using maps[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019: 832-839.
- [5] CHEN Z Y, LIU Y, GONG M M, et al. Generating dynamic kernels via transformers for lane detection[C]. 2023 IEEE/CVF International Conference on Computer Vision, 2023: 6835-6844.
- [6] QIU Q B, GAO H M, HUA W, et al. Priorlane: A prior knowledge enhanced lane detection approach based on transformer[C]. 2023 IEEE International Conference on Robotics and Automation, 2023: 5618-5624.
- [7] WANG J SH, MA Y CH, HUANG SH F, et al. A keypoint-based global association network for lane detection[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 1392-1401.
- [8] FENG ZH Y, GUO SH H, TAN X, et al. Rethinking efficient lane detection via curve modeling[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 17062-17070.
- [9] ZHENG T, FANG H, ZHANG Y, et al. RESA: Recurrent feature-shift aggregator for lane detection[C]. 35th AAAI Conference on Artificial Intelligence, 2021, 35(4): 3547-3554.
- [10] CHEN CH, LIU J, ZHOU CH, et al. Sketch and refine: Towards fast and accurate lane detection[C]. 38th AAAI Conference on Artificial Intelligence, 2024, 38(2): 1001-1009.
- [11] TANG J G, LI S B, LIU P. A review of lane detection methods based on deep learning[J]. Pattern Recognition, 2021, 111: 107623.
- [12] YANG J X, ZHANG L H, LU H CH. Lane detection with versatile atrousformer and local semantic guidance[J]. Pattern Recognition, 2023, 133: 109053.
- [13] ZHAO J, QIU Z Y, HU H Q, et al. HWLane: HW-transformer for lane detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(8): 9321-9331.
- [14] HOSSEINI S R, TAHERI H, TESHNEHLAB M. ENet-21: An optimized light CNN structure for lane detection[J]. ArXiv preprint arXiv: 2403.19782, 2024.
- [15] 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, 2022: 898-907.
- [16] LIU B H, LING Q. Hyper-Anchor based lane detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(10): 13240-13252.
- [17] 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, 2023: 6404-6413.
- [17] PAN X G, SHI J P, LUO P, et al. Spatial as deep: Spatial CNN for traffic scene understanding [C]. Thirty-Second AAAI Conference on Artificial Intelligence, 2018, 32(1): 7276-7283.
- [18] ZHANG Y J, ZHU L, FENG W, et al. Vil-100: A new dataset and a baseline model for video instance lane detection [C]. 2021 IEEE/CVF International Conference on Computer Vision, 2021: 15681-15690.
- [19] 李思纯, 王建军, 宋伟润, 等. 无人驾驶扫地机道路可行驶区域的融合提取研究 [J]. 仪器仪表学报, 2024, 45(12): 190-200.
- LI S CH, WANG J J, SONG W R, et al. Fusion extraction of road drivable areas for unmanned sweepers [J]. Chinese Journal of Scientific Instrument, 2024, 45(12): 190-200.
- [20] 张玥, 陈锡伟, 陈梦丹, 等. 基于对比学习生成对抗网络的无监督工业品表面异常检测 [J]. 电子测量与仪器学报, 2023, 37(10): 193-201.
- ZHANG Y, CHEN X W, CHEN M D, et al. Unsupervised surface anomaly detection of industrial products based on contrastive learning generative adversarial networks [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 193-201.
- [21] HU H G, WANG X Y, ZHANG Y, et al. A comprehensive survey on contrastive learning [J]. Neurocomputing, 2024, 610: 128645.
- [22] CHEN X L, HE K M. Exploring simple siamese representation learning [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 15750-15758.
- [23] LI Y K, CHEN C L P, ZHANG T. A survey on siamese network: Methodologies, applications, and opportunities [J]. IEEE Transactions on Artificial Intelligence, 2022, 3(6): 994-1014.
- [24] CHEN CH P, HSIEH J W, CHEN P Y, et al. SARAS-net: Scale and relation aware siamese network for change detection [C]. 37th AAAI Conference on Artificial Intelligence, 2023, 37(12): 14187-14195.
- [25] HE J B, ERFANI S, MA X J, et al. α -IoU: A family of power intersection over union losses for bounding box regression [J]. Advances in Neural Information Processing Systems, 2021, 34: 20230-20242.
- [26] HAMZENEJADI M H, MOHSENI H. Fine-tuned YOLOv5 for real-time vehicle detection in UAV imagery: Architectural improvements and performance boost [J]. Expert Systems with Applications, 2023, 231: 120845.
- [27] BEHRENDT K, SOUSSAN R. Unsupervised labeled lane markers using maps [C]. 2019 IEEE/CVF International Conference on Computer Vision Workshops, 2019: 832-839.
- [28] YU F, CHEN H F, WANG X, et al. Bdd100K: A diverse driving dataset for heterogeneous multitask learning [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 2636-2645.
- [29] YAN K Y, WANG J, WANG J, et al. An efficient improved YOLOv5-based method for detecting iron waste in ores [J]. Instrumentation, 2025, 12(2): 36-46.
- [30] QIU Z Y, ZHAO J, SUN SH L. MFIALane: Multiscale feature information aggregator network for lane detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24263-24275.
- [31] ZHOU K Y, LI L, ZHOU W G, et al. LaneTCA: Enhancing video lane detection with temporal context aggregation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025, 35(9): 8575-8585.
- [32] JIN D, KIM D, KIM C S. Recursive video lane detection [C]. 2023 IEEE/CVF International Conference on Computer Vision, 2023: 8473-8482.

作者简介



邹亮, 2010 年于安徽大学获得学士学位, 2013 年于中国科学技术大学获得硕士学位, 2017 年于不列颠哥伦比亚大学获得博士学位, 现为中国矿业大学副教授, 博士生导师, 主要研究方向为信号处理、智能语音、计算机视觉。

E-mail: liangzhou@cumt.edu.cn

Zou Liang received his B. Sc. degree from Anhui University in 2010, his M. Sc. degree from the University of Science and Technology of China in 2013, and his Ph. D. degree from the University of British Columbia (UBC) in 2017. He is currently a Ph. D. supervisor and Associate Professor at China University of Mining and Technology. His research interests include signal

processing, intelligent speech, and computer vision.



郭泽沛,2023年于南京工业大学获得学士学位,现为中国矿业大学信息与控制科学学院研究生,主要研究方向为计算机视觉。

E-mail:zpguo@cumt.edu.cn

Guo Zepei received his B. Sc. degree from Nanjing Tech University in 2023. He is currently a graduate student at the School of Information and Control Engineering, China University of Mining and Technology. His research interest is computer vision.



李颖娜,2003年于燕山大学获得学士学位,2006年于天津工业大学获得硕士学位,2018年于天津工业大学获得博士学位。现为唐山学院正高级实验师,主要研究方向智能检测。

E-mail:liyingna_2003@163.com

Li Yingna received her B. Sc. degree from Yanshan University in 2003, her M. Sc. degree from Tianjin Polytechnic University in 2006, and her Ph. D. degree from Tianjin Polytechnic University in 2018. She is currently a senior experimentalist at Tangshan University. Her research interest is intelligent detection.



鞠进军,2002年于山东大学获得学士学位,2009年于工程兵指挥学院获硕士学位,现为陆军工程大学训练基地副教授,主要研究方向为智能检测。

Email:jushi5666@sina.com

Ju Jinjun received his B. Sc. degree from Shandong University in 2002 and his M. Sc. degree from the Engineer Corps Command College of the PLA in 2009. He is currently an associate professor at the Training Base of the Army Engineering University. His research interest is equipment testing.



雷萌(通信作者),2008年、2013年分别于中国矿业大学获得学士、博士学位,现为中国矿业大学信息工程学院副教授,主要研究方向:智能检测。

E-mail:lmsiee@cumt.edu.cn

Lei Meng (Corresponding author) received her B. Sc. and Ph. D. degrees both from China University of Mining and Technology in 2008 and 2013, respectively. She is currently an associate professor with the School of Information and Electrical Engineering, China University of Mining and Technology. Her research interests include intelligent detection and computer vision.