

基于风格迁移的夜间车道线检测方法

李晓旋¹

LI Xiaoxuan

摘要

针对现有车道线数据集类别分布不均衡和采集标注困难等问题,文章提出了一种基于风格迁移的夜间车道线检测方法。首先,将无监督域适应的思想引入到车道线检测任务中,通过对抗学习,让网络学习域不变特征,提高对夜间目标域的适应性,来丰富训练数据的分布,提高网络的泛化能力。然后,在ERFNet中引入自注意力机制,通过预测图像中的遮挡位置提高网络提取全局特征的能力。实验结果表明,相较于ERFNet,所提方法总体精度 F_1 -measure提高了1.81%,夜间场景下的精度 F_1 -measure提高了2.38%。

关键词

车道线检测; 风格迁移; CycleGAN; ERFNet; 数据增强

doi: 10.3969/j.issn.1672-9528.2025.02.030

0 引言

车道线检测能够为自动驾驶车辆的变道辅助决策提供支持,在智能交通和自动驾驶领域发挥着关键作用^[1],对于提高自动驾驶车辆的安全性和推动发展智能网联、车路协同有着重要意义。

基于传统图像处理的车道线检测方法通常分为3个步骤,即图像预处理、特征提取和特征拟合。图像预处理和特征拟合方法已经相对成熟,而现有的特征提取方法往往依赖过强的先验假设,在获取车道线的颜色^[2]、边缘^[3]、纹理^[4]等视觉或高阶特征时存在困难,且算法可移植性弱,缺乏普适性,因此在实际夜间环境中容易产生错判和漏判。

基于深度学习的车道线检测方法可以根据不同的场景自适应调整参数以适应不同环境和车道线形状,为车道线检测问题提供了更可行的解决方案。Pan等人^[5]提出了一种特殊的卷积方式(spatial CNN, SCNN),将特征图在行、列方向上切片,逐片递进卷积,将上、下、左、右不同维度的信息采用element-wise add聚合,实现跨行、跨列的像素级消息传递,提高了车道线检测的精度和鲁棒性。Lee等人^[6]提出了一种多任务分割网络VPGNet,将车道线图像划分为若干个栅格,分别预测栅格级检测框、目标掩码和多类别标签,并通过预测消失点位置完成车道线检测,在复杂恶劣天气(阴雨)场景下获得较高的准确

率。Hou等人^[7]提出了一种基于自注意力蒸馏(self-attention distillation, SAD)的车道线检测模型,在编码器端提取多个注意力图,采用分层自顶向下的注意力蒸馏网络进行层间的表征学习,用深层注意力图来指导浅层注意力图的学习,从而提升整体的分割精度。Qin等人^[8]提出了一种基于逐行搜索的超快速车道线检测方法(ultra fast structure-aware deep lane detection, UFLD),将逐像素分类任务简化为行分类问题,提升车道线检测的速度。但基于深度学习的方法仍存在一些不足,庞大的神经网络模型依赖大量数据进行训练,而现有车道线数据集存在严重的类别分布不平衡问题,以CULane数据集为例,夜间样本仅占20.3%,如图1所示,直接影响了训练出的检测模型在预测夜间车道线时的精度。

受数据增强技术的启发,本文将无监督域适应的思想引入车道线检测,提出了一种基于风格迁移的夜间车道线检测方法,在车道线检测网络SA-ERFNet前端加入了一个风格迁移网络Ad-CycleGAN用于数据增强,在不额外人工标注的同时,提升夜间车道线的检测性能。

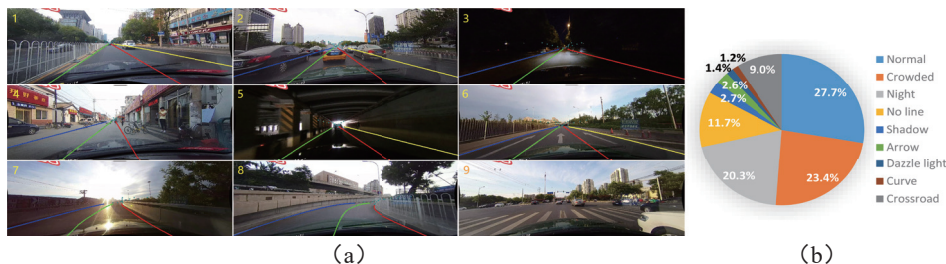


图1 CULane各场景示例图

1. 长安大学 陕西西安 710064

1 风格迁移网络 Ad-CycleGAN

1.1 网络结构

Ad-CycleGAN 网络结构如图 2 所示。两个生成器具有相同的网络结构，如表 1 所示。第 1 层采用卷积核大小为 7×7 、步长为 1 的卷积进行特征提取；第 2~3 层采用卷积核大小为 3×3 、步长为 2 的卷积进行下采样，通过缩小提取特征图的尺寸来减小运算量；第 4~13 层采用 9 个 ResNetV2 残差块^[9]进行特征提取；第 14 层采用卷积核大小为 3×3 、步长为 0.5 的卷积进行上采样；第 15~16 层采用卷积核大小为 7×7 、步长为 1 的卷积进行上采样，弥补下采样造成的细节损失；第 17 层采用 tanh 激活函数将输出图像的像素范围调整至 $[-1,1]$ 。

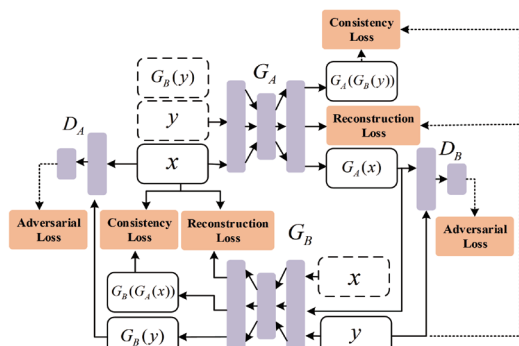


图 2 Ad-CycleGAN 网络结构

表 1 生成器的具体结构表

层数	类型	输入通道数	输出通道数	激活函数
1	初始卷积层	3	64	ReLU
2	下采样卷积层	64	128	ReLU
3	下采样卷积层	128	256	ReLU
4~13	残差块	256	256	ReLU
14	转置卷积层	256	256	
15	转置卷积层	256	128	ReLU
16	转置卷积层	128	64	ReLU
17	输出层	64	3	Tanh

判别器采用 PatchGAN 结构，如表 2 所示。第 1~4 层采用卷积核大小为 4×4 、步长为 2 的卷积进行上采样，得到高分辨率的特征图；第 5 层采用全连接层输出结果，并对结果分类。

表 2 判别器的具体结构表

层数	类型	输入通道数	输出通道数	激活函数
1	卷积层	3	64	LeakyReLU
2	卷积层	64	128	LeakyReLU
3	卷积层	128	256	LeakyReLU
4	卷积层	256	512	LeakyReLU
5	全连接层	512	1	

1.2 损失函数

将白天场景的车道线图像定义为域 X ，夜晚场景的车道线图像定义为域 Y 。生成器 G_A 伪造虚假的夜晚图像，生成器 G_B 伪造虚假的白天图像，判别器 D_B 对虚假和真实的夜间图像进行真伪判断，判别器 D_A 对虚假和真实的白天图像进行真伪判断。从域 X 到域 Y 的对抗性损失（Adversarial loss）公式为：

$$L_{\text{GAN}}(G_A, D_B, X, Y) = E_{y \sim P_{\text{data}}(y)} [\log D_B(y)] + E_{x \sim P_{\text{data}}(x)} [\log(1 - D_B(G_A(x)))] \quad (1)$$

同理，从域 Y 到域 X 的对抗性损失为 $L_{\text{GAN}}(G_B, D_A, Y, X)$ 。

为确保两个生成器网络之间的映射一致，引入循环一致性损失（Consistency loss）。循环一致性损失要求生成器 G_A 前向循环一致和生成器 G_B 反向循环一致，即 $x \rightarrow G_A(x) \rightarrow G_B(G_A(x)) \approx x$ 和 $y \rightarrow G_B(y) \rightarrow G_A(G_B(y)) \approx y$ ，公式为：

$$L_{\text{cyc}}(G_A, G_B) = E_{x \sim P_{\text{data}}(x)} [\|G_B(G_A(x)) - x\|_1] + E_{y \sim P_{\text{data}}(y)} [\|G_A(G_B(y)) - y\|_1] \quad (2)$$

式中： $\|\cdot\|_1$ 表示求 L_1 范数。

由于 HSV 图像的三通道相互独立，且光照对图像的亮度通道 V 最敏感^[10]，因此，本文采用 V 通道信息来计算输入图像与重建图像的 L_1 距离生成重建性损失（Reconstruction loss）。对训练好的生成器 G_A ，若输入图像为夜间图像 y ，那么输出应为伪造夜间图像 $G_A(y)$ ，即 $y \rightarrow G_A(y) \approx y$ ；对训练好的生成器 G_B ，若输入图像为白天图像 x ，那么输出应为伪造白天图像 $G_B(x)$ ，即 $x \rightarrow G_B(x) \approx x$ 。重建性损失公式为：

$$L_{\text{rec}}(G_A, G_B) = E_{x \sim P_{\text{data}}(x)} [\|V(G_B(x)) - V(x)\|_1] + E_{y \sim P_{\text{data}}(y)} [\|V(G_A(y)) - V(y)\|_1] \quad (3)$$

因此，总损失公式为：

$$L(G_A, G_B, D_A, D_B) = L_{\text{GAN}}(G_A, D_B, X, Y) + L_{\text{GAN}}(G_B, D_A, Y, X) + L_{\text{cyc}}(G_A, G_B) + \alpha L_{\text{rec}}(G_A, G_B) \quad (4)$$

2 车道线检测网络 SA-ERFNet

车道线具有细长、连续的结构，可以通过利用车道线像素间的长距离依赖关系，提高车道线检测网络全局特征提取能力。因此，本文提出了一种基于注意力机制的车道线检测网络 SA-ERFNet，其整体网络结构如图 3 所示，详细网络结构如表 3 所示，包括编码 - 解码器模块、自注意力 SA 模块和车道线预测分支模块。

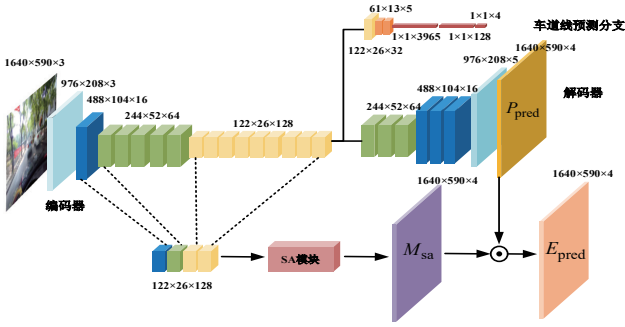


图3 SA-ERFNet 的整体网络结构

表3 SA-ERFNet 网络结构表

层数	类型	输出通道数	输出尺寸
1	下采样块	16	488×104
2	下采样块	64	244×52
3-7	5×Non-bt-1D	64	244×52
8	下采样块	128	122×26
9	Non-bt-1D (dilated 2)	128	122×26
10	Non-bt-1D (dilated 4)	128	122×26
11	Non-bt-1D (dilated 6)	128	122×26
12	Non-bt-1D (dilated 8)	128	122×26
13	Non-bt-1D (dilated 2)	128	122×26
14	Non-bt-1D (dilated 4)	128	122×26
15	Non-bt-1D (dilated 6)	128	122×26
16	Non-bt-1D (dilated 8)	128	122×26
17	上采样块	64	244×52
18-19	2×Non-bt-1D	64	244×52
20	上采样块	16	488×104
21-22	2×Non-bt-1D	16	488×104
23	上采样块	C	976×208

2.1 编码 - 解码器模块

编码器（Encoder）由双分支下采样块和非对称的Non-bt-1D块组成，如图4和图5所示。网络的第1、2和8层采用一种双分支下采样块，将分别经过卷积和最大池化操作得到的特征图按通道拼接。网络的3~7

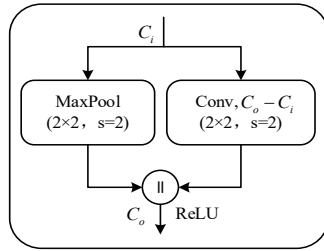


图4 下采样块

层、9~16层为Non-bt-1D块，将3×3的二维卷积分解为一个一维水平方向的卷积（1×3）和一个一维垂直方向的卷积（3×1），并通过残差连接进行特征融合。

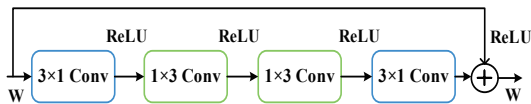


图5 Non-bt-1D 块

解码器（Decoder）由Non-bt-1D块和上采样块组成，其中，网络中第18~19层、21~22层为Non-bt-1D块，第17、20、23层采用卷积核为3×3、步长为2的转置卷积进行上采样操作，将特征图恢复到原图尺寸，并进行像素级分类，输出车道线概率图。

2.2 自注意力模块

如图6所示，自注意力SA模块首先使用双线性插值法将编码 - 解码器模块第2、3、8和16层的输出特征图下采样到相同尺寸，并进行加权平均，以同样的重要性融合不同层的不同分辨率特征图。接着，多分辨率的聚合特征依次通过卷积、最大池化、全连接、激活操作分别得到水平特征图 $f_{sa}^h \in \mathbb{R}^{C \times H \times 1}$ 和垂直特征图 $f_{sa}^v \in \mathbb{R}^{C \times 1 \times W}$ ，增强细长特征的上下文信息，抑制噪声特征。然后，分别将 f_{sa}^h 和 f_{sa}^v 在水平方向和垂直方向上以尺寸 W 、 H 扩展为原始图像大小，得到水平扩展特征图 $M_{sa}^h \in \mathbb{R}^{C \times H \times W}$ 和垂直扩展特征图 $M_{sa}^v \in \mathbb{R}^{C \times H \times W}$ 。最后将二者逐像素相加融合输出自注意力权重图 M_{sa} ，并与特征解码分支输出的车道线概率图 P_{pred} 按像素相乘，输出车道线自注意力加权概率图，计算公式为：

$$E_{pred} = P_{pred} \odot M_{sa} \quad (5)$$

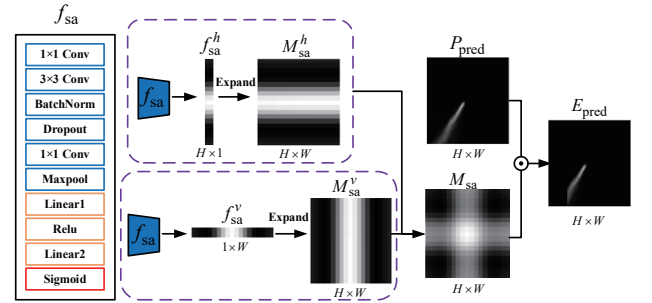


图6 自注意力 SA 模块网络结构

2.3 车道线检测分支模块

为提高车道线的定位准确度和辅助预测车道线，利用车道线预测分支判断车道线是否存在，若概率小于等于0.5，表示车道线不存在，用0表示，否则为1。

3 结果与分析

3.1 数据集

CULane数据集覆盖了各种光照条件、驾驶场景，共包含133 235张1640×590的车道线图像。

3.2 实验设置及评价指标

输入图像大小为976 px×208 px，优化器为Adam，批处理大小为2，初始学习率为0.000 2，迭代次数为100次，重建性损失权重参数为0.5。

为了评估车道线检测网络的性能，使用综合查准率（Precision）和召回率（Recall）的指标 F_1 -measure为最终的评估标准，其计算方法为：

$$F_1\text{-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Precision 计算方法为:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

Recall 计算方法为:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

式中: TP 为实际存在且被检测到的车道线的数量; FP 为不存在但被检测到的车道线的数量; FN 为实际存在但未被检测到的车道线的数量。

3.3 实验结果与分析

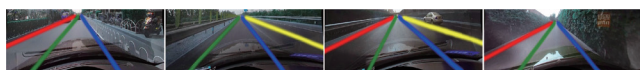
3.3.1 SA-ERFNet 网络有效性验证

为了验证 SA-ERFNet 方法的有效性,选择经典的车道线检测网络 ERFNet^[11]、SCNN、UFLD、SAD 和 SA-ERFNet 进行对比,结果如表 4 所示。从实验结果可得,SA-ERFNet 性能最优,在各种场景下的表现都很均衡,总体精度 F_1 -measure 为 74.39%。而在 Night 场景下,SA-ERFNet、ERFNet、SCNN、UFLD 和 SAD 的 F_1 -measure 分别为 68.45%、67.58%、66.47%、61.40% 和 66.29%,这表明在夜间场景下,SA-ERFNet 的检测能力更强。从运行速度来看,SA-ERFNet 可以满足实时性的要求,数据处理速度达到了 77 帧/s。总的来说,SA-ERFNet 是相对最优的车道线检测方法,达到了较好的准确率和实时性之间的平衡。

表 4 不同车道线检测方法的对比

单位: %

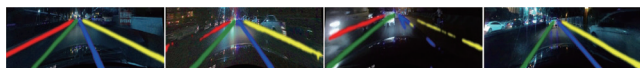
场景	ERFNet	SCNN	UFLD	SAD	SA-ERFNet
Normal	91.31	90.81	88.18	90.70	91.43
Crowded	71.08	71.13	65.52	70.13	73.20
Night	67.58	66.47	61.40	66.29	68.45
No line	45.30	45.67	38.40	43.46	45.92
Shadow	67.83	70.59	58.91	66.98	71.12
Arrow	86.93	85.43	81.38	84.39	86.97
Hlight	67.53	64.80	54.88	59.87	68.15
Curve	69.11	66.30	56.00	65.74	69.40
Crossroad	2065	2236	2291	2052	2463
Total	73.08	72.69	68.12	72.20	74.39
FPS	78	22	213	75	77



(a) Normal 类拟合图



(b) Crowded 类拟合图



(c) Night 类拟合图



(d) No line 类拟合图



(e) Shadow 类拟合图



(f) Arrow 类拟合图



(g) Hlight 类拟合图



(h) Curve 类拟合图



(i) Crossroad 类拟合图

图 7 SA-ERFNet 车道线检测方法结果图

从 CULane 数据集中随机抽取图像,采用 SA-ERFNet 方法对车道线进行拟合,可视化结果如图 7 所示。可以看出,在夜间场景中,SA-ERFNet 的车道线拟合效果有待提升,需要进一步优化。

3.3.2 Ad-CycleGAN + SA-ERFNet 方法的有效性验证

为了验证本文方法(Ad-CycleGAN + SA-ERFNet)的有效性,将本文方法与方法 1(SA-ERFNet)、方法 2(StarGAN^[12] + SA-ERFNet)、方法 3(CycleGAN^[13] + SA-ERFNet)、方法 4(EnlightenGAN^[14] + SA-ERFNet)进行对比实验,结果如表 5 所示。

表 5 不同风格迁移方法 + SA-ERFNet 的对比

单位: %

类别	方法 1	方法 2	方法 3	方法 4	本文方法
Normal	91.43	91.98	92.11	92.08	92.25
Crowd	73.20	71.24	73.22	73.55	73.31
Night	68.45	68.36	69.56	69.04	69.96
No line	45.92	45.77	45.30	46.66	46.49
Shadow	71.12	68.95	71.72	76.17	72.41
Arrow	86.97	87.24	86.93	87.75	87.76
Dazzle light	68.15	66.73	68.14	68.38	70.66
Curve	69.40	64.68	69.11	67.08	70.34
Crossroad	2463	2291	2065	2346	2177
Total	74.39	73.41	74.55	74.65	74.89

从实验结果可得,本文方法性能最优,在各种场景下的表现都很均衡,总体精度 F_1 -measure 为 74.89%。在夜晚场景下,本文方法表现最佳, F_1 -measure 达到了

69.96%，比方法 1 的结果提高了 1.51%，比方法 2 的结果高出 1.6%，比方法 3 的结果高出了 0.4%，比方法 4 的结果高出了 0.92%。

采用 SA-ERFNet 方法训练扩充后的 CULane 数据集，并随机抽取图像，对车道线进行拟合，可视化结果如图 8 所示。实验结果表明，本文方法拟合的车道线精度更高，较少出现漏检或错检，且在遮挡情况下，也能保持车道线的连续性和完整性。



(a) 方法 1 (SA-ERFNet)



(b) 方法 2 (StarGAN + SA-ERFNet)



(c) 方法 3 (CycleGAN + SA-ERFNet)



(d) 方法 4 (EnlightenGAN + SA-ERFNet)



(e) 本文方法 (Ad-CycleGAN + SA-ERFNet)

图 8 风格迁移方法对比结果图

4 结语

本文提出了一种基于风格迁移的夜晚车道线检测方法。首先，利用风格迁移网络 Ad-CycleGAN 生成夜间场景下的图像，并将生成的图像与原始图像混合作为新的数据集。然后，使用结合语义分割和自注意力机制的车道线检测方法 SA-ERFNet 预测、拟合车道线。实验结果表明，本文方法在夜晚场景下具有良好的性能表现。

参考文献：

- [1] 中华人民共和国工业和信息化部. 道路车辆先进驾驶辅助系统 (ADAS) 术语及定义: GB/T 39263—2020[S]. 北京: 中国标准出版社, 2020:1.
- [2] 李楠, 顾苏菁, 叶飞, 等. 一种基于 HSI 颜色模型的车道线检测算法[J]. 信息技术与信息化, 2022(4): 58-61.
- [3] CANNY J. A computational approach to edge detection[J]. IEEE transactions on pattern analysis and machine intelligence, 1986, PAMI-8 (6): 679-698.
- [4] GONG J L, CHEN T, ZHANG Y F. Complex lane detection based on dynamic constraint of the double threshold[J]. Multimedia tools and applications, 2021, 80(5): 27095-27113.

- [5] PAN X G, SHI J P, LUO P, et al. Spatial as deep: spatial CNN for traffic scene understanding[DB/OL].(2017-12-17)[2024-06-19].<https://doi.org/10.48550/arXiv.1712.06080>.
- [6] LEE S, KIM J, YOON J S, et al. VPGNet: vanishing point guided network for lane and road marking detection and recognition[C/OL]// 2017 IEEE International Conference on Computer Vision (ICCV).Piscataway: IEEE, 2017[2024-04-03].<https://ieeexplore.ieee.org/document/8237477>.
- [7] HOU Y N, MA Z, LIU C X, et al. Learning lightweight lane detection CNNs by self attention distillation[C/OL]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV).Piscataway:IEEE, 2019[2024-05-01].<https://ieeexplore.ieee.org/document/9009567>.
- [8] QIN Z Q, WANG H Y, LI X. Ultra fast structure-aware deep lane detection[J].Computer vision-ECCV 2020, 2020, 12369:276-291.
- [9] HE K M, ZHANG X Y, REN S Q, et al. Identity mappings in deep residual networks[DB/OL].(2016-07-25)[2024-06-11].<https://doi.org/10.48550/arXiv.1603.05027>.
- [10] LEZAMA J, QIU Q, SAPIRO G. Not afraid of the dark: NIR-VIS face recognition via crossspectral hallucination and low-rank embedding[DB/OL].(2016-11-21)[2024-07-13].<https://doi.org/10.48550/arXiv.1611.06638>.
- [11] ROMERA E, ÁLVAREZ J M, BERGASA L M, et al. ER-FNet: efficient residual factorized ConvNet for real-time semantic segmentation[J]. IEEE transactions on intelligent transportation systems, 2017, 19(1): 263-272.
- [12] CHOI Y, CHOI M, KIM M, et al. StarGAN: unified generative adversarial networks for multi-domain image-to-image translation[DB/OL].(2018-09-21)[2024-06-19].<https://doi.org/10.48550/arXiv.1711.09020>.
- [13] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C/OL]//2017 IEEE International Conference on Computer Vision (ICCV).Piscataway:IEEE, 2017[2024-04-19].<https://ieeexplore.ieee.org/document/8237506>.
- [14] JIANG Y F, GONG X Y, LIU D, et al. EnlightenGAN: deep light enhancement without paired supervision[J]. IEEE transactions on image processing, 2021, 30: 2340-2349.

【作者简介】

李晓旋 (1998—)，女，安徽宿州人，硕士研究生，研究方向：计算机视觉、图像处理。

(收稿日期：2024-10-21)