http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0601

融合 CBAM 注意力机制与可变形卷积的车道线检测

胡丹丹[1]，张忠婷[1]，牛国臣[1, 2, *]

(1. 中国民航大学 机器人研究所，天津 300300；2. 中国民航大学 民航智慧机场理论与系统重点实验室，天津 300300)

摘 要：为满足自动驾驶及汽车高级驾驶辅助系统（ADAS）对车道线检测准确性和实时性的要求，提出一种融合卷积块注意力机制 (CBAM) 与可变形卷积网络 (DCN) 的车道线检测方法 CADCN。在特征提取模块中嵌入 CBAM 注意力机制，增强有用特征并抑制无用特征响应；引入可变形卷积替换常规卷积，用带偏移的采样学习车道线的几何形变，提高卷积核的建模能力；基于行锚分类思想，对行方向上的位置进行选择和分类分析，预测车道线的位置信息，提高车道线检测模型的实时性。在车道线公开数据集上对所提 CADCN 方法进行训练及验证，在满足实时性的情况下，CADCN 方法在 TuSimple 数据集上准确率达到 96.63%，在 CULane 数据集上综合评估指标 F_1 平均值达到 74.4%，验证了所提方法的有效性。

关 键 词：车道线检测；特征提取；注意力机制；可变形卷积网络；行锚分类
中图分类号：U471.11; TP391.4
文献标志码：A **文章编号**：1001-5965(2024)07-2150-11

车道线作为路面的主要交通标志，具有分隔道路区域、规定行驶方向和为行人提供引导信息的功能，是保证车辆安全行驶的重要信息。车道线检测技术是智能车领域的关键技术之一，被广泛应用于自动驾驶及车道保持、车道偏离预警、前向碰撞预警等汽车高级驾驶辅助系统 (advanced driver assistance system, ADAS)[1] 中。

实际路况中，车道线的形状、颜色具有多样性；行人、车辆等目标遮挡会影响车道线的连续性；在不同的驾驶场景中，车道线的表现形式不同，如栅栏、台阶等；还受到天气因素、光线变化、道路阴影和车道线破损等影响。此外，车道线检测任务还需具有较高的实时性以保障车辆响应速度及安全性。因此，车道线检测研究仍面临着严峻的挑战。

目前，基于计算机视觉的车道线检测研究主要包括传统方法和深度学习方法。传统方法主要通过手工提取不同环境中车道线的颜色、边缘或梯度变化等特征[2-3]，利用霍夫变换[4]、最小二乘法[5] 等后处理来拟合车道线形状。当实际环境中车道线较为清晰时，传统方法表现出良好的准确性和检测效率，但易受道路环境的干扰，在光线变化、路面积水、障碍物严重遮挡和阴影干扰等复杂场景下难以取得理想效果。

为解决传统车道线检测方法存在的检测精度低、对环境要求高、适应性差等问题，国内外学者开始研究基于深度学习的车道线检测方法。Kim 和 Lee[6] 提出一种基于卷积神经网络（convolutional neural network, CNN）的车道线检测算法。Li 等[7] 提出一种将 CNN 和递归神经网络（recurrent neutral network, RNN）相结合的车道线检测方法，利用 CNN 提取车道线每帧图像中的特征信息，输入 RNN 中进行特征学习和车道线预测。Pan 等[8] 提出空间 CNN（spatial CNN, SCNN）模型，用片连片卷

收稿日期：2022-07-09；**录用日期**：2022-11-19；**网络出版时间**：2022-12-15 17:33
网络出版地址：link.cnki.net/urlid/11.2625.V.20221214.1720.008
基金项目：天津市科技计划 (17ZXHLGX00120)；中央高校基本科研业务费专项资金 (3122022PY17,3122017003)
* **通信作者**. E-mail：niu_guochen@139.com

引用格式：胡丹丹，张忠婷，牛国臣. 融合 CBAM 注意力机制与可变形卷积的车道线检测 [J]. 北京航空航天大学学报，2024，50（7）：2150-2160. HU D D，ZHANG Z T，NIU G C. Lane line detection incorporating CBAM mechanism and deformable convolutional network [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（7）：2150-2160（in Chinese）.

积取代传统层接层卷积的连接形式, 使信息可以跨行和跨列传递, 改善无视觉线索情况下的车道线检测问题。但是, 逐片卷积信息传递方式耗时较长, 难以进行实时的车道线检测。

为提高车道线检测速度, Hou 等[9] 提出一种采用自注意蒸馏 (self-attention distillation, SAD) 模块的轻量级网络模型, 推理时间较少, 但会丢失一些详细信息, 对车道线缺失或无视觉线索等复杂环境下的车道线检测效果欠佳。Neven 等[10] 提出 LaneNet 模型, 将车道线检测过程视为实例分割问题, LaneNet 模型由 2 个分支构成, 其中, 二值分割分支通过逐像素的语义分割区分车道线和背景, 像素嵌入分支将车道线像素分解为不同的车道线实例, 结合 2 个分支的结果得到实例分割的效果, 从而完成车道线提取。此外, Tabelini 等[11] 提出 PolyLaneNet 模型, 利用深度学习网络回归车道线的曲线方程, 为每条车道线输出一个多项式及置信度得分, 实时性较高, 但当车道线受到严重遮挡时, 检测性能会显著降低。

为解决实际应用中车道线检测准确性和实时性不平衡的问题, 本文提出一种融合卷积块注意力机制 (convolutional block attention module, CBAM) 与可变形卷积网络 (deformable convolutional networks, DCN) 的车道线检测方法 CADCN, 在满足车道线检测实时性的前提下提高检测精度。通过融合 CBAM 注意力机制[12] 和可变形卷积[13] 重构残差网络 (residual network, ResNet)[14], 端到端地学习全局特征并细化局部特征, 获取更丰富的特征表示, 提升模型的整体车道线识别能力。基于行锚分类思想, 网格化车道线位置代替像素级分类, 提高车道线模型检测的实时性。

1　车道线检测网络框架

本文提出一种融合 CBAM 注意力机制和可变形卷积的 CADCN 车道线检测方法, 其网络框架如图 1 所示, 主要由车道线特征提取模块和车道线预测模块 2 部分组成, 能够实现端到端、实时的车道线检测。

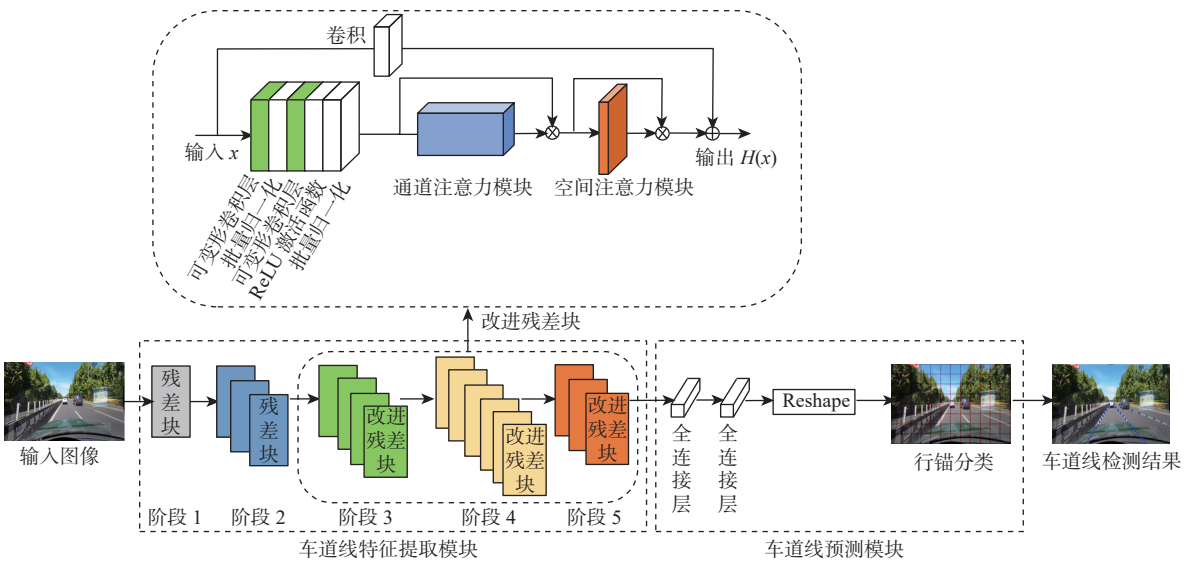


图 1　网络框架

Fig. 1　Network framework

车道线特征提取模块以 ResNet34 作为骨干网络, 其包含 5 个阶段, 即 5 种参数不同的卷积阶段, 经过多层的卷积层进行特征提取。利用 ResNet34 残差网络中特殊的跳连结构缓解因网络深度增加导致的梯度爆炸或梯度弥散等网络性能退化现象, 加快模型收敛, 提升模型训练效果。在阶段 3～阶段 5 的残差块中引入 CBAM 注意力机制和可变形卷积, 提高模型特征提取能力, 提取到更多关于车道线的细节信息, 尤其是低对比度区域的信息, 提升车道线模型检测准确性。

在车道线预测模块中, 使用全局的融合特征, 经过 2 个全连接层 FC 的分类, 解决感受野小的问题。利用上下文的环境信息将不连续但属于同一条车道线的各部分识别为一个整体, 无须复杂的信息传递机制就可以实现很好的车道线识别效果, 改善车道线被遮挡、磨损等复杂场景下的检测性能。为提高车道线检测速度, 改进传统语义分割方法, 借鉴行锚分类思想, 将像素点分类改为对像素块进行分类, 减少需分类的像素点的数量, 在不影响正常检测的前提下, 放大需检测的区域。方法复杂度缩减, 极大提高了车道线检测方法的实时性。通过预测预定义的行锚上所有位置的概率分布选择车

道线的位置，判断车道线是否存在，端到端地输出车道线的位置信息。

2 融合 CBAM 注意力机制与可变形卷积的特征提取模块

ResNet34 的阶段 1 由 1 个卷积层和 1 个最大池化层 MaxPool 组成，阶段 2～阶段 5 由残差块堆叠而成，见图 2。如图 2(a)所示，残差块的输入为 x，输出为 $H(x)$，输入和输出之间通过一个旁路卷积将浅层特征与深层抽象特征直接连接在一起，增加特征的重复利用率，在一定程度上能够避免信息传递时的特征丢失。

本文提出的 CADCN 车道线检测方法，在 ResNet34 的阶段 3～阶段 5 的每个残差块中串联加入 CBAM 注意力机制，分别考虑车道线图像不同通道上像素的重要性及同一通道不同空间位置上像素的重要性，强化有效特征信息，弱化无效特征信息。此外，为更好地提取输入特征，对不同位置对应的尺度或感受野进行自适应调整，将阶段 3～阶段 5 的残差块中的 3×3 卷积层调整为可变形卷积，其他阶段保留为标准卷积层，以减小引入可变形卷积带来的网络参数量的增加对网络负荷的影响。改进后的残差块如图 2(b)所示。

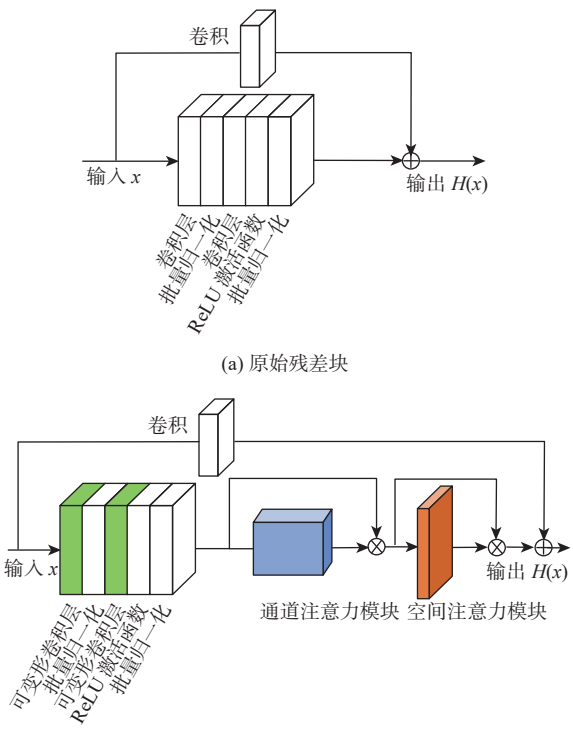


图 2 残差结构改进

Fig. 2 Improvement of residual structure

2.1 基于 CBAM 注意力机制的残差块重构

车道线通常只占图像的小部分区域，大部分内容是自然背景，为剔除与车道线无关的冗余特征信息，提升车道线检测模型的特征提取能力，在 ResNet34 中对残差块进行重构，设计并添加 CBAM 注意力机制，对提取到的车道线特征进行特征重标定，模仿人类视觉感知，自动过滤不重要的信息，把更多的注意资源用于需要重点关注的目标区域，提高视觉信息处理效率与准确性。

CBAM 注意力机制整体流程如图 3 所示。CBAM 由通道注意力机制与空间注意力机制组成，协同学习图像中的关键局部细节信息，为神经网络特征图中车道线区域分配较高的权重，背景分配较低的权重，提高神经网络对图像中车道线的关注度，进而提升网络的特征学习和表达能力。

2.1.1 通道注意力机制

在判断某一像素是否属于车道线时，特征图的各个通道往往具有不同的重要程度，平等地对各个通道的特征信息进行利用，会造成计算资源的浪费。引入通道注意力机制，建模特征图通道间的相互关系，有区别地对待不同通道的特征信息。

如图 3 所示，输入特征图记为 $F \in \mathbf{R}^{C\times H\times W}$，$H$ 和 W 分别为输入特征图的高和宽，C 为输入特征图的通道数。通过全局最大池化和全局平均池化 AvgPool 对特征图 F 的全局空间信息进行压缩，生成 2 个尺寸为 $1\times1\times C$ 的特征图 S_1、S_2。将每个通道特征图在整张图中的位置信息进行融合，可避免小尺寸卷

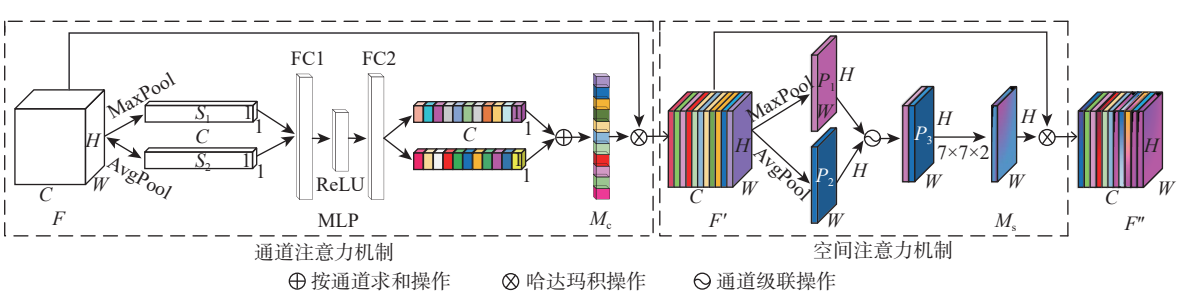


图 3 CBAM 结构

Fig. 3 CBAM structure

积核在提取特征时因感受野范围太小造成的网络评估偏差。

为充分利用压缩操作提取的特征信息, 通过评估操作获得通道间的相关性。S_1和S_2通过共享由全连接层FC1、FC2和ReLU非线性激活函数组成的多层感知机 (multi-layer perception, MLP) 得到2个一维特征图。对其按通道进行求和操作后使用Sigmoid函数进行归一化, 得到各通道的尺寸为$1\times1\times C$的权重统计值M_c, 该过程可表示为

$$M_c(F)=\sigma\left(\mathrm{MLP}\left(\mathrm{AvgPool}(F)\right)+\mathrm{MLP}\left(\mathrm{MaxPool}(F)\right)\right)=\sigma\left(\mathrm{MLP}\left(\frac{1}{H\times W}\sum_{i_0=1}^{H}\sum_{j_0=1}^{W}f_{x_0}(i_0,j_0)\right)+\mathrm{MLP}\left(\max_{i\in H,j\in W}f_{x_0}(i_0,j_0)\right)\right)\tag{1}$$

式中: σ为Sigmoid函数; $f_{x_0}(i_0,j_0)$表示输入特征图F的第x_0通道中坐标为(i_0,j_0)点的像素值。

2.1.2　空间注意力机制

空间注意力机制以通道特征重标定后得到的特征图F'作为输入, 在通道维度上分别做全局最大池化和平均池化操作, 得到特征图$P_1\in\mathbf{R}^{1\times H\times W}$、$P_2\in\mathbf{R}^{1\times H\times W}$。通过通道级联的方式得到特征描述$P_3\in\mathbf{R}^{2\times H\times W}$, 利用卷积层对$P_3$中不同位置的信息进行编码和融合, 得到空间加权信息M_s, 可用于区分图像不同空间位置的重要程度, 过程如下:

$$M_s(F)=\sigma\left(f^{7\times7}\left(\mathrm{AvgPool}(F');\mathrm{MaxPool}(F')\right)\right)\tag{2}$$

式中: $f^{7\times7}$表示卷积核为7×7大小的卷积层。

2.1.3　CBAM 注意力机制

将输入特征图F与各通道权重值M_c的2个矩阵的对应元素相乘, 对F进行特征重标定处理, 得到能够有效体现特征关键通道信息的特征映射F'。在通道加权的基础上, 利用串联的空间注意力机制对空间特征信息进行自适应加权, 将F'作为空间注意力模块的输入, 与空间权重系数M_s这2个矩阵的对应元素相乘, 得到包含通道位置信息、空间位置信息的显著特征图F'', 使网络能够更多关注车道线输入特征强的部分, 提升其空间特征选择能力, 其过程可表示为

$$\begin{cases}F'=M_c\otimes F\\F''=M_s\otimes F'\end{cases}\tag{3}$$

式中: "⊗"表示2个矩阵对应元素相乘。

2.2　可变形卷积替换常规卷积

常规卷积受限于固定的几何结构, 感受野难以感知目标的几何变形, 可提取的有效信息有限, 容易导致目标关键特征信息丢失严重, 网络整体性能下降。为提高CNN对几何变换的建模能力, 引入可变形卷积替换普通卷积, 构造可变形残差块。用带偏移的采样代替固定位置采样, 采样位置自动向目标区域位置偏移, 使卷积核形状自适应车道线形状, 重构ResNet34的卷积层, 以捕获各种形状和尺寸的车道线微小结构细节。

传统的二维卷积, 在局部采样位置p_0的输出特征映射y可定义为

$$y(p_0)=\sum_{p_n\in r}w(p_n)\cdot x(p_0+p_n)\tag{4}$$

式中: $w(p_n)$为采样位置的卷积核权重; $x(p_0+p_n)$为输入特征图的采样位置; p_n为感受野区域中的所有采样位置; r为感受野区域。

在可变形卷积中, 其输出特征映射y定义如下:

$$y(p_0)=\sum_{p_n\in r}w(p_n)\cdot x(p_0+p_n+\Delta p_n)\tag{5}$$

式中: $\{\Delta p_n|n=1,\cdots,N\}, N=|r|$表示在标准卷积的每个采样位置上增加的可学习的偏移量, 通常是一个小数。

因此, 偏移后采样位置的像素$x(p_0+p_n+\Delta p_n)$通常利用双线性插值方法来实现:

$$x(p)=\sum_{q}G(q,p)\cdot x(q)\tag{6}$$

式中: p为区域中的任意位置, $p=p_0+p_n+\Delta p_n$; q为输入特征映射中的所有空间位置, 即$p_0+p_n+\Delta p_n$周围的4个整数点; $x(q)$为特征图上所有整数点的取值; G为双线性插值核函数, 是1个二维核, 可以分解为2个一维核:

$$G(q,p)=\mathrm{g}(q_x,p_x)\mathrm{g}(q_y,p_y)\tag{7}$$

其中: $g(a,b)=\max\{0,1-|a-b|\}$。

引入调节因子[15]$\Delta m_n(0\leqslant\Delta m_n\leqslant1)$对采样范围加以约束, 减少特征学习过程中不相关内容的影响:

$$y(p_0)=\sum_{p_n\in r}w(p_n)\cdot x(p_0+p_n+\Delta p_n)\Delta m_n\tag{8}$$

可变形卷积过程如图4所示。卷积核在偏移量Δp_n的约束下, 车道线特征图的感受野变成一个多边形, 并且多边形的每个位置都受到调节因子Δm_n的影响, 使采样点可以提取车道线不规则的特征, 关注车道线的局部形态变换。偏移量和调节因子的大小通过卷积层进行学习确定。

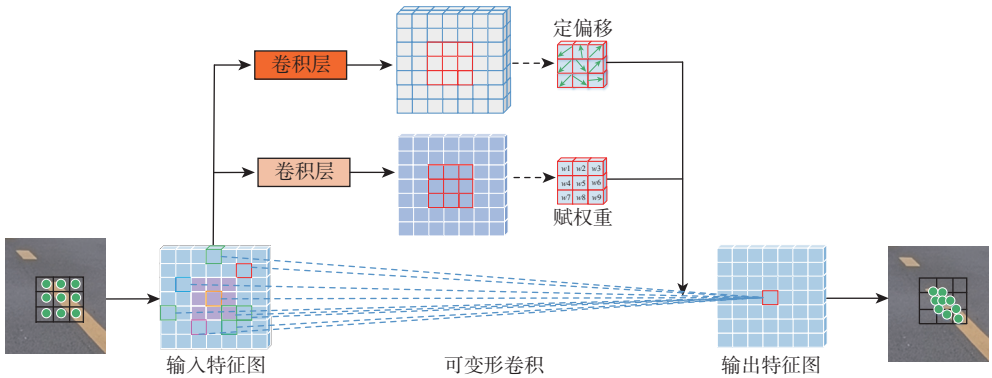


图4 可变形卷积过程

Fig. 4 Process of deformable convolutional network

3 基于行锚分类的车道线预测

考虑车道线长且窄的结构特征，其像素量远小于背景像素量，借鉴超快速结构感知深度车道检测(ultra fast structure-aware deep lane detection, UFLD)方法[16]提出的行锚分类思想，将车道线检测定义为在行方向上进行水平位置的选择和分类问题，由网络模型预测最可能包含车道标记的单元，用网格化车道线位置选取代替像素级分类，旨在复杂道路场景下对车道线进行快速、实时的检测。

将输入通道数为C_0的车道线图像缩放至$H_0\times W_0$大小，本文中H_0取288，W_0取800。车道线特征提取模块处理后的特征图经过全连接层后输出分类张量，利用行锚分类思想，将其重构为C'通道的特征图，每个通道的特征图被分为h行，这些预定义的行称为一个行锚，每一行在垂直方向上再进行切割，形成w个网格，即将$H_0\times W_0$大小的图像聚合成$h\times w$个单元进行处理。在特征图的最右侧增加一个网格单元(即第w+1个)来表示该行中没有车道的情况。利用行锚分类思想，在行方向上对w+1个网格单元进行选择与分类，选择含有车道线元素的网格单元。

以TuSimple数据集为例进行分析，先将输入RGB三通道的车道线图像缩放至288像素×800像素大小，在车道线特征提取模块经过一个卷积层后得到64像素×144像素×400像素的特征图，再经过最大池化将特征维度缩减为64像素×72像素×200像素。经过ResNet34堆叠的残差模块输出512像素×9像素×25像素的特征图，平均池化后得到8像素×9像素×25像素的压缩图，对其进行扁平化处理，展开成1×1 800的一维向量，经过2个全连接层后，其特征维度映射为22 624=(100+1)×56×4，用于最终的SoftMax分类，其中的101为w+1，56表示h，4为设定的图像最多可检测的车道线数量，即C'。SoftMax分类的输出为：$P_{\mathrm{rob}_{i,j,k}}(1\leqslant i\leqslant C', 1\leqslant j\leqslant h, 1\leqslant k\leqslant w+1)$，其中，$P_{\mathrm{rob}_{i,j,k+1}}$表示第$i$个车道出现在第$j$行、第$k$列的网格单元上的可能性，当$k=w+1$时，即$P_{\mathrm{rob}_{i,j,k+1}}$表示第$i$个车道不出现在第$j$行任一个网格单元的可能性。对行锚上的每个网格单元求期望，即可预测出车道线的位置信息。

以LaneNet为例的传统分割方法需处理$H_0\times W_0\times C_0$个分类问题，十分消耗计算资源，难以达到实时性的要求。本文提出的基于行锚分类的CADCN车道线检测方法只需处理$h\times(w+1)\times C'$个分类问题，其中，$h\ll H_0$，$w\ll W_0$，网络模型可以直接提供车道线标志的位置且计算参数量降低，车道线检测实时性提高。

4 实验结果与分析

4.1 数据集及评价标准

为验证CADCN车道线检测模型的性能，分别使用TuSimple数据集和CULane数据集对模型进行训练及评估。数据集的详细信息如表1所示。

4.1.1 TuSimple数据集及评价标准

TuSimple数据集的示例如图5所示。在TuSimple数据集中，使用准确率A对车道线的预测结果进行评估，表示预测正确的像素点在样本中所占的比例，通常$A=\sum_{\mathrm{clip}}C_{\mathrm{clip}}\Big/\sum_{\mathrm{clip}}S_{\mathrm{clip}}$越高表示检测效果

表1 车道线数据集的详细信息

Table 1 Details of lane line dataset

数据集	总帧数	训练集	验证集	测试集	分辨率/(像素×像素)	标注方式	多天气	多时段	多线型	道路类型	车道数	场景种类
TuSimple	6 408	3 268	358	2 782	1 280×720	点坐标	是	是	否	公路	≤5	1
CULane	133 235	88 880	9 675	34 680	1 640×590	点坐标	是	是	否	公路、城市、农村	≤4	9



图 5　TuSimple 数据集示例

Fig. 5　Example of TuSimple dataset

越好, 计算公式如下:

$$A=\sum_{\text{clip}}C_{\text{clip}}\Big/\sum_{\text{clip}}S_{\text{clip}} \tag{9}$$

式中: C_{clip}为预测正确的车道线点数; S_{clip}为标注的车道线真值点数。

4.1.2　CULane 数据集及评价标准

CULane 数据集的测试集分为 9 个类别, 分别为正常、拥挤、夜间、无车道线、阴影、箭头、亮光、弯道和十字路口, 9 种场景所占比例如表 2 所示, CULane 数据集示例如图 6 所示。

表 2　CULane 数据集场景分类及占比

Table 2　CULane dataset scene classification and proportion

场景	场景描述	占比/%
正常	驾驶员视野良好, 车道线标记清晰可见	27.74
拥挤	道路上大量车辆缓行, 车道线被车辆遮挡	23.39
夜间	晚上能见度低, 车道线模糊, 各类灯光干扰	20.27
无车道线	无车道线标记, 通常道路狭窄且路边停车较多	11.73
阴影	天桥或高架桥下光线较暗, 树荫或建筑阴影	2.68
箭头	车道线之间的直行、转弯、调头等标记	2.57
亮光	各类强光造成路边反光, 道路上车道线不清晰	1.40
弯道	前方车道线弯曲	1.22
十字路口	十字路口中斑马线标记干扰, 路口中间无车道线	9.00

在 CULane 数据集中, 使用F_1分数作为综合评估标准, 表示精确率P和召回率R的加权调和平均, F_1分数越高表明模型性能越好, 其计算过程如下:

$$F_1=2\frac{PR}{P+R} \tag{10}$$

$$\begin{cases}P=\dfrac{T_P}{T_P+F_P}\\[2ex] R=\dfrac{T_P}{T_P+F_N}\end{cases} \tag{11}$$



图 6　CULane 数据集示例

Fig. 6　Example of CULane dataset

其中: F_N为假阴性的数目; T_P、F_P分别为真阳性、假阳性的数目。

本文选用准确率A、F_1分数、帧率综合评价 CADCN 车道线检测模型, 验证其性能。

4.2　实验设置

实验使用的硬件环境为: 处理器为 Intel(R) Core (TM) i7-10750H CPU@2.60 GHz 2.59 GHz, GPU 为 GTX-1650Ti。软件环境为: Ubuntu18.04, CUDA-11.1, OpenCV-3.4, Pytorch-1.7.1, Python3.8。使用标准的随机梯度下降法训练模型, 学习动量为 0.9, 权重衰减为 10^{-4}, TuSimple 数据集和 CULane 数据集学习率分别为 2.5×10^{-2} 和 1.6×10^{-2}, 批处理量大小分别为 4 和 2。

4.3　实验对比分析

4.3.1　定量对比实验

基于上述公开车道线数据集进行训练获得车道线检测模型, 将 CADCN 模型分别与 Segnet-Res34[17]、PolyLaneNet[11]、FastDraw[18]、PINet[19]、SAD[9]、LaneNet[10]、EL-GAN[20]、SCNN[8] 等模型进行对比, 用于验证该模型的检测效果。在 TuSimple 数据集上的评估结果如表 3 所示, 在 CULane 数据集上的评估结果如表 4 所示。

从表 3 的实验结果可以看出, 与基于语义分割和聚类实例的 Segnet-Res34[17]、EL-GAN[20]、SCNN[8] 车道线模型相比, 由于 CADCN 模型采用行锚分类的方法对像素块进行分类, 降低了计算量, 在检测速度方面有明显的优势, 可达到 50 帧/s, 是 SCNN[8] 的 7 倍左右; 与 PolyLaneNet[11]、FastDraw[18] 模型相比, CBAM 注意力机制与可变形卷积的加入增强了网络的特征提取能力, 提取到的车道线特征细节信息更加丰富, 使 CADCN 模型在检测准确率方面优势较大; 与 PINet[11]、SAD[9]、LaneNet[10] 模型相比,

表3 TuSimple数据集上不同模型检测结果比较

Table 3 Comparison of detection results of different models on TuSimple dataset

方法	准确率/%	帧率/(帧·s^{-1})
Segnet-Res34[11]	92.84	19.5
PolyLaneNet	93.36	115
FastDraw[18]	95.20	90
PINet[19]	96.51	35
SAD	96.02	42
LaneNet	96.38	52.63
EL-GAN[20]	96.39	9.8
SCNN	96.53	7.5
CADCN	96.63	50

CADCN模型在保持高实时性的前提下，车道线检测准确率最高，可达96.63%，达到检测速度和精度的良好平衡，表现出优越的性能。

表4以F_1分数和帧率比较了CULane数据集中不同场景下的车道线检测效果。可以看出，相较于其他模型，CADCN模型在除亮光及十字路口场景外，F_1指标均表现较优，在夜间、阴影、无车道线等复杂驾驶场景下，表现出良好的抗扰能力。在检测速度方面，CADCN模型基于像素块在行方向上进行分类，其检测速度高于DeepLabV2和FastDraw[18]，可达92.4帧/s，比SCNN[6]提高了10倍以上，相较于LaneNet[10]、SAD[9]、PINet[19]提高了2～3倍，表现出明显的优势。

表4 CULane数据集上不同模型检测结果比较

Table 4 Comparison of detection results of different models on CULane dataset

方法	F_1/%								F_P(十字路口)	平均值/%	帧率/(帧·s^{-1})
	正常	拥挤	夜间	无车道线	阴影	箭头	亮光	弯道			
LaneNet[10]	82.9	61.1	53.4	37.7	56.2	72.2	54.5	59.3	5 928	61.8	44
DeepLabV2	87.4	64.1	60.6	38.1	60.7	79	54.1	59.8	2 505	66.7	77.31
FastDraw[18]	85.9	63.6	57.8	40.6	59.9	79.4	57.0	65.2	7 013	69.7	90.3
SAD[9]	89.8	68.1	64.2	42.5	67.5	83.9	59.8	65.5	1 995	70.5	39.53
SCNN[8]	90.6	69.7	66.1	43.4	66.9	84.1	58.5	64.4	1 990	71.6	7.5
PINet[19]	89.6	71.9	67.0	49.3	67.0	84.2	**65.2**	66.2	**1 505**	73.8	35
CADCN	**90.9**	**72.4**	**67.7**	**46.1**	**69.3**	**85.9**	62.3	**69.5**	2 016	**74.4**	**92.4**

注：加粗字体为每行最优结果。

综上所述，本文提出的融合CBAM注意力机制和可变形卷积的CADCN车道线检测方法能够在一定程度上使模型更加关注与车道线相关的特征信息，提升网络的检测性能，具有较强的鲁棒性和实时性。

4.3.2 消融实验

为验证CADCN模型采用模块的有效性，在设置相同实验参数的情况下做如下对比实验，实验结果如表5所示。可以看出，在可变形卷积替换普通卷积后，分别添加通道注意力机制和空间注意力机制均能提高车道线检测的准确率，但添加CBAM注意力机制后，车道线检测效果提升明显，说明CBAM注意力机制可以对图像中的车道线予以聚焦和关注，并且对干扰信息进行了有效抑制，提高了模型对目标通道和空间信息的聚焦能力。

表5 消融实验结果对比

Table 5 Comparison of ablation test results

组别	可变形卷积	通道注意力机制	空间注意力机制	CBAM注意力机制	准确率/%
1					95.87
2	√				96.07
3	√	√			96.23
4	√		√		96.30
5	√			√	96.63

4.3.3 实验可视化分析

运用本文提出的CADCN模型与LaneNet[10]模型在Tusimple数据集上进行实验验证，部分测试结果如图7所示，白色框标注的区域对比了2种模型对车道线的检测效果。

从图7可以看出，LaneNet[10]模型对车道线的拟合效果欠佳，在场景1(弯道)、场景2(车道线缺失)、场景3(车道线磨损)、场景4(道路阴影)中预测的车道线均出现了偏移，且在场景2和场景3中，因车道线缺失及磨损造成了车道线漏检。而CADCN模型通过输入图像的全局特征有效利用上下文信息，较好处理车道线细节部分，即使车道线



图 7　LaneNet 与 CADCN 检测结果对比

Fig. 7　Comparison of LaneNet and CADCN detection results

出现破损、阴影遮挡等情况, 也能够对其进行鲁棒检测, 且检测结果与真实标签有较高的交并比(intersection over union, IoU)。采用行锚分类预测车道线位置, 多个网格同时负责检测一条车道线, 能够有效避免漏检。然而, 对远端车道线的预测效果不及 LaneNet 模型。

将 CADCN 模型与 PINet[19] 模型在 CULane 数据集及校园环境下采集的图像进行对比, 部分检测效果如图 8 所示。

在场景 1 中, 车道线被车辆遮挡, CADCN 模型可以感受图像的全局信息, 依靠周围车流的走向对车道线进行定位, 借助 CBAM 注意力机制恢复被遮挡的车道线, 与 PINet[19] 模型相比, 减少了车道线漏检情况。

在无车道线视觉线索的场景 2 中, 模型在训练阶段经过多次重复的训练, 即使没有车道线信息, 也可根据道路边缘信息预测出合理的车道线位置。

在场景 3 中, 夜晚环境下道路光照较弱, 且路面有箭头等信息干扰, PINet[19] 模型出现了漏检现象, 由于 CBAM 注意力机制的加入, 使模型能够更加关注车道线的局部细节信息, 可变形卷积替换普通卷积后增强了特征提取能力, 使得提取的特征信息更加丰富。因此, CADCN 模型能够有效检测出车道线, 表现出优越的性能。

在场景 4 和场景 5 中, 车道线被道路两旁的树木阴影干扰, 图像阴影交界的明暗差异容易对模型预测车道线造成干扰。PINet[19] 模型在该情况下对车道线的特征捕捉不充分, 车道线识别不全, 融合 CBAM 注意力机制和可变形卷积的 CADCN 模型能够充分提取全局的空间语义信息, 将不连续但是属于同一条车道线的各个部分识别为一个整体。

场景 6 为实际校园环境下采集的图像, 车道线出现破损且路面存在积水, 车道线颜色与地面接近。从检测结果可以看出, PINet[19] 模型只检测到 2 条车道线, 而 CADCN 模型能够根据图像的全局特征及车道线局部细节特征, 准确定位 3 条车道线的位置, 具有良好的实际应用效果。

由上述测试对比结果可以看出, CADCN 模型在各种复杂场景下都能够较为准确地预测出车道线的位置信息, 显著提高了车道线检测的精度, 同

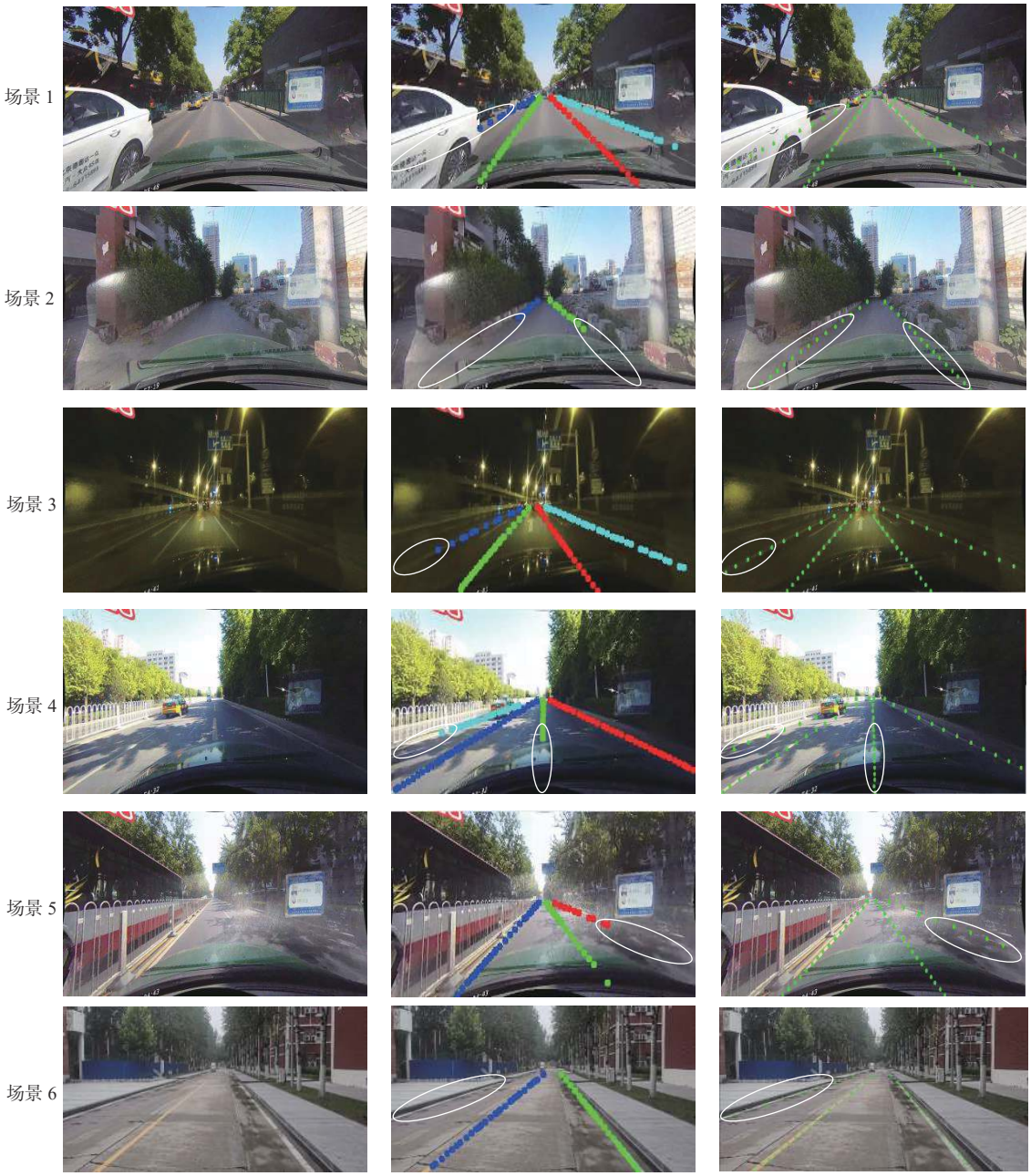


图 8 PINet 与 CADCN 检测结果对比

Fig. 8 Comparison of PINet and CADCN detection results

时具有很好的泛化能力和抗干扰能力。

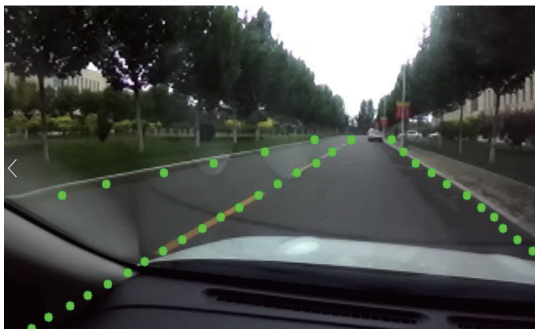
4.3.4 实际道路测试

为验证 CADCN 模型的有效性，在无人驾驶汽车平台上进行了实际道路的车道线检测。车辆的整体架构如图 9 所示，在长城欧拉黑猫 R1 新能源汽车的基础上，搭载了工控机、AGX 计算平台、ZED 2i 双目立体视觉相机和其他外设设备，如电源、鼠标、键盘和显示器等。车道线检测结果如图 10 所示。可以看出，CADCN 模型可以有效地对实际道路中的车道线进行检测，能够应对多种复杂场景，且实时性能够满足无人驾驶车辆对车道线检测算法的要求 (帧率>30 帧/s)，具有良好的实际应用效果。



图 9 无人驾驶汽车平台架构

Fig. 9 Driverless vehicle platform architecture



(a) 白天、马路边沿道路车道线检测



(b) 黑夜、积水道路车道线检测

图 10　实际道路车道线检测

Fig. 10　Lane line detection on actual road

5　结　论

针对复杂的道路驾驶环境, 为提升车道线检测方法性能, 提出 CADCN 车道线检测方法, 经实验验证可得到以下结论:

1) 在提高方法准确性方面, 采用 CBAM 注意力机制使模型能够专注于学习车道线特征, 提高了模型对车道线检测的灵敏度及辨别能力; 将部分 3×3 卷积层调整为可变形卷积, 提高了模型对车道线几何形变的建模能力和学习效率; 通过强化骨干网络的深层语义特征提取能力, 方法的鲁棒性及环境适应性均得到提高。在 TuSimple 数据集上准确率达到 96.63%, 在 CULane 数据集上其综合评估指标 F_1 平均值达 74.4%。

2) 在提高方法实时性方面, 基于行锚分类思想, 将车道线检测定义为在预定义的行锚上进行水平位置的选择和分类问题, 用网格化车道线位置选取代替像素级分类, 计算量减少。采用端到端的方法, 无后处理操作, 实时性较高, 能够满足无人驾驶车辆实时的道路分析检测需求。

本文方法虽然能快速、实时地对车道线进行检测, 但对远端车道线的检测效果欠佳, 下一步的研究工作是考虑对远端和近端车道线使用不同大小的网格分割以提高车道线检测准确率。

参考文献 (References)

[1] WAYKOLE S, SHIWAKOTI N, STASINOPOULOS P. Review on lane detection and tracking algorithms of advanced driver assistance system[J]. Sustainability, 2021, 13(20): 11417.

[2] 杜恩宇, 张宁, 李艳荻. 基于 Gabor 滤波器的车道线快速检测方法[J]. 红外与激光工程, 2018, 47(8): 314-321.
DU E Y, ZHANG N, LI Y D. Lane line quick detection method based on Gabor filter[J]. Infrared and Laser Engineering, 2018, 47(8): 314-321(in Chinese).

[3] ANDRADE D C, BUENO F, FRANCO F R, et al. A novel strategy for road lane detection and tracking based on a vehicle's forward monocular camera[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(4): 1497-1507.

[4] LUO S, ZHANG X Q, HU J, et al. Multiple lane detection via combining complementary structural constraints[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(12): 7597-7606.

[5] 蔡英凤, 张田田, 王海, 等. 基于实例分割和自适应透视变换算法的多车道线检测[J]. 东南大学学报 (自然科学版), 2020, 50(4): 775-781.
CAI Y F, ZHANG T T, WANG H, et al. Multi-lane detection based on instance segmentation and adaptive perspective transformation [J]. Journal of Southeast University (Natural Science Edition), 2020, 50(4): 775-781(in Chinese).

[6] KIM J, LEE M. Robust lane detection based on convolutional neural network and random sample consensus[C]//Proceeding of the International Conference on Neural Information Processing. Berlin: Springer, 2014: 454-461.

[7] LI J,MEI X,PROKHOROV D,et al. Deep neural network for structural prediction and lane detection in traffic scene[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(3): 690-703.

[8] PAN X G, SHI J P, LUO P, et al. Spatial as deep: Spatial CNN for traffic scene understanding[C]//Proceeding of the Thirty-Second AAAI Conference on Artificial Intelligence. Palo Alco: AAAI, 2018: 7276-7283.

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

[10] NEVEN D, DE BRABANDERE B, GEORGOULIS S, et al. Towards end-to-end lane detection: An instance segmentation approach[C]//Proceeding of the IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2018: 286-291.

[11] TABELINI L, BERRIEL R, PAIXÃO T M, et al. PolyLaneNet: Lane estimation via deep polynomial regression[C]//Proceeding of the 25th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2021: 6150-6156.

[12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceeding of the European Conference on Computer Vision. Berlin: Springer, 2018: 3-19.

[13] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]/Proceeding of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 764-773.

[14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.

[15] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: More deformable, better results[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 9300-9308.

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

[17] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.

[18] PHILION J. FastDraw: Addressing the long tail of lane detection by adapting a sequential prediction network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 11574-11583.

[19] KO Y, LEE Y, AZAM S, et al. Key points estimation and point instance segmentation approach for lane detection[C]//Proceedings of the IEEE Transactions on Intelligent Transportation Systems. Piscataway: IEEE Press, 2022: 8949-8958.

[20] GHAFOORIAN M, NUGTEREN C, BAKA N, et al. EL-GAN: Embedding loss driven generative adversarial networks for lane detection[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2019: 256-272.

Lane line detection incorporating CBAM mechanism and deformable convolutional network

HU Dandan[1], ZHANG Zhongting[1], NIU Guochen[1, 2, *]

(1. Robotics Institute, Civil Aviation University of China, Tianjin 300300, China;
2. Key Laboratory of Smart Airport Theory and System, Civil Aviation University of China, Tianjin 300300, China)

Abstract: In order to meet the accuracy and real-time requirements of autonomous driving and advanced driver assistance systems (ADAS) for lane line detection, a CADCN lane line detection method incorporating convolutional block attention module (CBAM) mechanism and deformable convolutional network (DCN) was proposed. Firstly, the CBAM mechanism was embedded in the feature extraction module to enhance the useful features and suppress the useless feature responses. Secondly, DCN was used to replace the conventional convolutional network, and the geometric deformation of lane lines was learned by sampling with offset to improve the modeling capability of the convolution kernel. Finally, based on the idea of row anchor classification, the location point along the row was selected and classified, so as to predict the lane line location information and thus improve the real-time performance of the lane line detection model. The CADCN model was trained and validated on the public lane line dataset. While ensuring real-time performance, the accuracy rate of the model on the TuSimple dataset reaches 96.63%, and the comprehensive evaluation index F_1 on the CULane dataset reaches 74.4%, which verifies the effectiveness of the algorithm.

Keywords: lane line detection; feature extraction; attention mechanism; deformable convolutional network; row anchor classification

Received: 2022-07-09; **Accepted**: 2022-11-19; **Published Online**: 2022-12-15 17:33

URL: link.cnki.net/urlid/11.2625.V.20221214.1720.008

Foundation items: Tianjin Science and Technology Plan (17ZXHLGX00120); The Fundamental Research Funds for the Central Universities (3122022PY17,3122017003)

* **Corresponding author**. E-mail: niu_guochen@139.com