

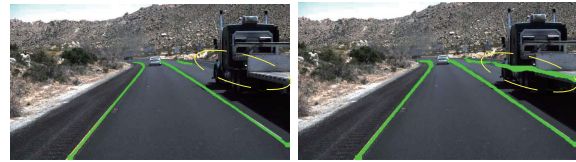


DOI: 10.12086/oe.2022.210378

基于语义分割的实时车道线检测方法

张冲, 黄影平*, 郭志阳, 杨静怡

上海理工大学光电信息与计算机工程学院, 上海 200093



摘要: 车道线识别是自动驾驶环境感知的一项重要任务。近年来,基于卷积神经网络的深度学习方法在目标检测和场景分割中取得了很好的效果。本文借鉴语义分割的思想,设计了一个基于编码解码结构的轻量级车道线分割网络。针对卷积神经网络计算量大的问题,引入深度可分离卷积来替代普通卷积以减少卷积运算量。此外,提出了一种更高效的卷积结构 LaneConv 和 LaneDeconv 来进一步提高计算效率。为了获取更好的车道线特征表示能力,在编码阶段本文引入了一种将空间注意力和通道注意力串联的双注意力机制模块 (CBAM) 来提高车道线分割精度。在 Tusimple 车道线数据集上进行了大量实验,结果表明,本文方法能够显著提升车道线的分割速度,且在各种条件下都具有良好的分割效果和鲁棒性。与现有的车道线分割模型相比,本文方法在分割精度方面相似甚至更优,而在速度方面则有明显提升。

关键词: 车道线检测; 语义分割; 卷积神经网络; 自动驾驶

中图分类号: TP391.4

文献标志码: A

张冲, 黄影平, 郭志阳, 等. 基于语义分割的实时车道线检测方法 [J]. 光电工程, 2022, 49(5): 210378

Zhang C, Huang Y P, Guo Z Y, et al. Real-time lane detection method based on semantic segmentation[J]. *Opto-Electron Eng.* 2022, 49(5): 210378

Real-time lane detection method based on semantic segmentation

Zhang Chong, Huang Yingping*, Guo Zhiyang, Yang Jingyi

School of Optical-Electronic and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Abstract: Lane line recognition is an important task of automatic driving environment perception. In recent years, the deep learning method based on convolutional neural network has achieved good results in target detection and scene segmentation. Based on the idea of semantic segmentation, this paper designs a lightweight Lane segmentation network based on encoding and decoding structure. Aiming at the problem of large amount of computation of convolution neural network, the deep separable convolution is introduced to replace the ordinary convolution to reduce the amount of convolution computation. Moreover, a more efficient convolution structure of laneconv and lanedeconv is proposed to further improve the computational efficiency. Secondly, in order to obtain better lane line feature representation ability, in the coding stage, a dual attention mechanism module (CBAM) connecting spatial attention and channel attention in series is introduced to improve the accuracy of lane line

收稿日期: 2021-11-24; 收到修改稿日期: 2022-01-31

基金项目: 上海自然科学基金资助项目 (20ZR14379007); 国家自然科学基金面上项目 (61374197)

*通信作者: 黄影平, huangyingping@usst.edu.cn。

版权所有©2022 中国科学院光电技术研究所

segmentation. A large number of experiments are carried out on tusimple lane line data set. The results show that this method can significantly improve the lane line segmentation speed, and has a good segmentation effect and robustness under various conditions. Compared with the existing lane line segmentation models, the proposed method is similar or even better in segmentation accuracy, but significantly improved in speed.

Keywords: lane detection; semantic segmentation; convolutional neural networks; automatic driving

1 引言

车道线识别是自动驾驶环境感知的一项重要任务。准确且快速地检测出车道线, 是无人车安全可靠行驶的前提。车道线检测可分为传统视觉车道线检测以及基于深度学习的车道线检测。传统车道线检测大多是基于几何方法、使用一些曲线模型来拟合车道线^[1-4]。最近几年, 卷积神经网络快速发展, 在目标检测和场景语义分割中取得了很好的效果。由此, 研究人员开始使用卷积神经网络实现车道线检测并且取得了一定的进展^[5-12], 但在模型的实时性及检测精度上仍有提升的空间。本文借鉴语义分割的思想, 设计了一个基于编码解码结构的轻量级车道线分割网络。

主要贡献有: 1) 提出了一种高效的车道线识别方法, 采用图像语义分割的思想, 通过编码器-解码器结构的深度学习模型将车道线识别问题转化为像素分类问题。2) 引入深度可分离卷积, 并且设计了有效的轻量级卷积结构, 包括一个用于特征编码的下采样卷积模块 LaneConv 和一个用于特征解码的上采样卷积模块 LaneDeconv, 提高计算效率以解决常规语义分割方法中因复杂的卷积计算而导致实时性差的问题。3) 相关的研究表明采用注意力机制的 CBAM^[13] 模型可以充分利用局部信息来聚焦场景的显著特征。本文在编码阶段加入 CBAM 模块获得更好的车道线特征提取能力, 以提高车道线分割精度。4) 使用 Tusimple 车道线数据集进行实验评估, 与多种基于深度学习的车道线检测算法的对比分析表明, 本文方法在分割精度和分割速度上具有综合优势。

2 相关工作

2.1 传统方法

大多数经典的车道检测方法都是基于手工特征, 并采用预定义的曲线模型如线模型或抛物线模型来匹配线特征。Yue 等人^[1] 提出了一种基于样条的车道模型, 该模型描述了平行线对通用车道边界的透视效果。

与其他车道模型如直线模型和抛物线模型相比, 它可以通过不同的控制点形成任意形状, 以描述更大范围的车道结构。Jiang 等人^[2] 使用双抛物线模型来描述更精确和可靠的车道线。朱鸿宇等人^[3] 提出一种基于级联霍夫变换的车道线检测方法, 使用感兴趣区域 (region of interest, ROI) 选取, 滤波, 边缘检测, 非极大值抑制等预处理, 克服了传统霍夫变换在映射过程时需对每一个点进行极坐标转换的缺点, 使计算更简单。Tian 等人^[4] 提出了一种基于直线段检测器 (line segment detector, LSD)、自适应角度滤波器和双卡尔曼滤波器的 ADAS 车道检测与跟踪方法。在车道检测过程中, 将输入图像中的感兴趣区域 (ROI) 转换为灰度图像, 然后对灰度图像进行中值滤波、直方图均衡化、图像阈值化和透视映射等预处理。然后, 在 ROI 上应用快速鲁棒的 LSD 算法, 并使用所提出的自适应角度滤波器, 更有效地消除不正确的车道线。

2.2 深度学习方法

目前基于深度学习的车道检测方法可分为两类: 基于检测的方法^[5-8] 和基于分割的方法^[9-12]。基于检测的方法会将场景图片划分为一个个大小相同的小网格区域, 然后利用 YOLO 这样的目标检测器对每一行的网格区域进行检测判断是否属于车道线, 最后将属于车道线的网格拟合为车道线。基于分割的方法是将场景图片的每一个车道线像素都进行分类, 判断该像素是否属于车道线或者背景。基于检测方法的优点是速度较快, 处理直线车道能力较强。但是当环境复杂且弯道较多的情况下, 检测效果要明显比基于分割的方法更差。基于分割的方法一般检测效果都要比基于检测的方法更好, 但是速度上要稍逊。

在基于检测的方法中, Qin 等人^[5] 将车道检测过程视为基于行的选择问题, 它为无视觉线索的车道线检测问题提供了很好的解决方案。但是, 其结构损失函数导致了在弯曲处车道线检测效果不佳。Chen 等人^[6] 通过将线的表示构造为可学习的点来提出了 PointLaneNet, 然后使用目标检测网络来检测这些车

道线点进而得到车道线。Tabelini 等人^[7]对输入图像进行特征提取, 并使用多项式对车道线候选标记进而直接输出车道线曲线。Ji 等人^[8]通过改进的 YOLO 网络, 并根据车道线分布来确定不同的检测尺度, 从而提升检测效果和速度。通常, 基于检测的方法可以实现快速的运行速度, 而准确性却不是很突出。在基于分割的方法中, Neven 等人^[9]和刘彬等人^[10]是通过结合分割方法和聚类算法, 将车道检测过程视为一个实例分割问题。但是, 这种方法在车道线不清晰或车道被遮挡的情况下不是很有效, 并且无法达到实时性。田锦等人^[11]基于实例分割方法 Mask-RCNN 对车道线进行实例分割, 并且提出了一种自适应拟合的方法, 通过多项式拟合对不同视野内的车道线特征点进行拟合。Pan 等人^[12]提出了 SCNN, 通过用逐片段卷积代替普通卷积来改善无视觉提示的问题。但是, 逐片段卷积需要更多卷积运算以融合局部上下文和层中的通道特征且 SCNN 不适用于包含任意数量车道的场景。

3 方法

本文方法的网络结构如图 1 所示, 整体网络结构是基于编码器-解码器结构。其中特征编码阶段采用轻量级卷积 LaneConv (LC) 与深度可分离卷积 (DWConv) 进行快速特征下采样。接着利用 CBAM

模块获得更多的全局语义信息以及学习更多的通道特征, 保留更多的且重要的车道线细节。在特征解码阶段采用轻量级的 LaneDeConv (LD) 与深度可分离卷积 (DWConv) 进行快速特征上采样, 还原特征图尺寸, 最后将分割的二值车道线聚类产生最终的车道线语义分割结果。

网络的详细参数设置如图 2 所示, 图 2 中的每个分支包含多个卷积层。每一层的运算 (opr) 可以是 DWConv、LC 或 LD, 其中 DWConv 表示深度可分离卷积层, LC 表示 LaneConv 层, LD 表示 LaneDeconv 层。每个操作包含一个核大小 k , 步幅 s , 输出通道 c , 重复次数 r 。

3.1 深度可分离卷积

类似于膨胀卷积^[15], 深度可分离卷积 (DWConv) 也是一种可以有效减少卷积计算量的轻量化卷积。深度可分离卷积是在 MobileNet^[14] 网络中最先提出, 随后 Xception^[16] 也使用了它, 在参数量不增加的情况下提升了模型效果。深度可分离卷积过程分为两步, 分别是逐通道卷积和逐点卷积, 如图 3(a) 和图 3(b) 所示。逐通道卷积是指一个 3×3 的卷积核只会与一个通道进行卷积运算, 避免了普通卷积每个卷积核需要和每一个通道进行卷积运算所带来的高昂运算, 此过程

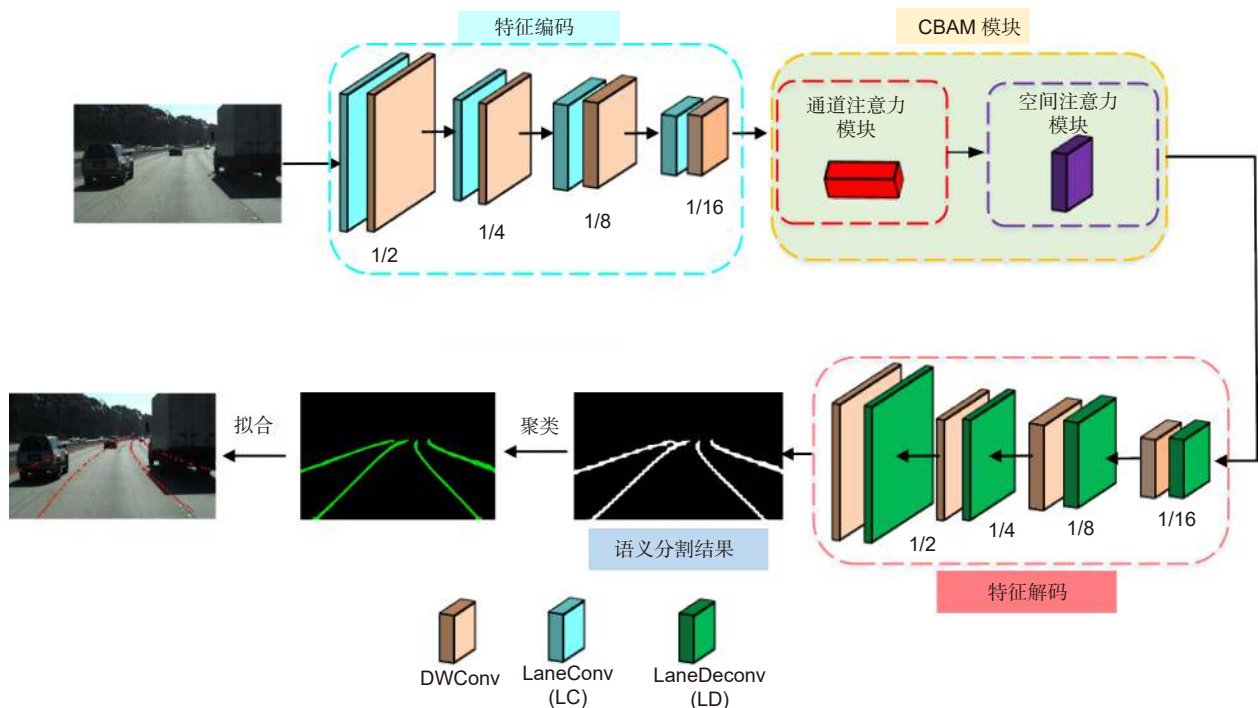


图 1 方法框图

Fig. 1 Framework of the method

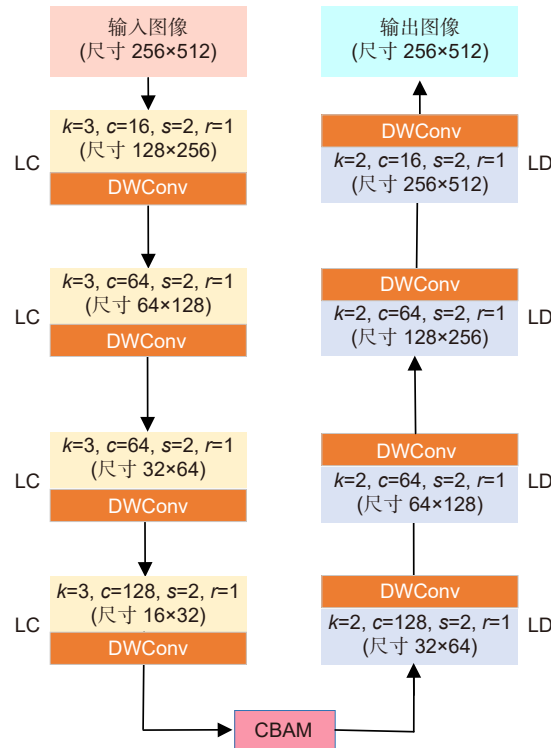


图 2 网络参数与轻量化卷积结构图

Fig. 2 The parameters of the network and illustration of LaneConv and LaneDeConv

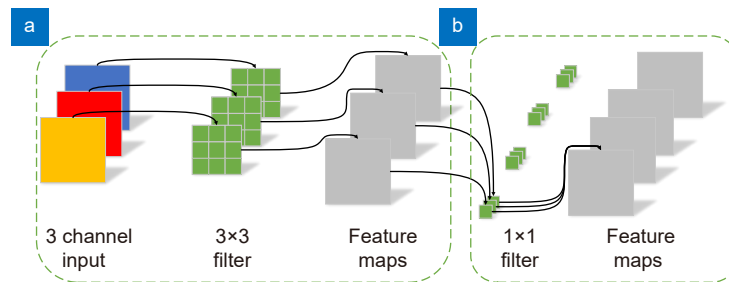


图 3 深度可分离卷积。

(a) 逐通道卷积; (b) 逐点卷积

Fig. 3 Depth separable convolution.

(a) Channel by channel convolution; (b) Pointwise convolution

会产生与输入的通道数相同的特征图。但是可以看出，逐通道卷积并没有将输入特征的不同通道进行关联，因此还需要逐点卷积来加强不同的通道间的联系。逐点卷积的卷积核大小为 1×1 ，这里的卷积运算会将上一步的特征图的不同通道进行一个加权组合，生成新的特征图。可以看出，不同于普通卷积，深度可分离卷积是将卷积过程分成两步，从而大量减少卷积所带来的运算并且具有和普通卷积相同的感受野和相同的输出特征通道数。

3.2 轻量化卷积与反卷积网络

本文由 Inception-V3^[17] 得到启发，设计了一种全

新的快速下采模块 LaneConv 以获得更快的检测速度。如图 4 (a) 所示，首先将大小为 $H \times W \times C$ 的输入张量输入到一个 1×1 卷积中，使通道数减小到 $C/2$ ，然后用一个 3×3 卷积来进行下采样，再加上一个并行的 3×3 最大池化来恢复 C 的通道大小。给定相同的输入和输出大小， 3×3 卷积层需要 $9C^2$ 参数和 $9HWC^2$ 计算量，而 LaneConv 只需要 $3C^2$ 参数量和 $9HWC^2/8$ 计算量。由 SqueezeSeg^[18] 得到启发，本文也设计了一个快速上采样模块 LaneDeConv。如图 4 (b) 所示，先使用 1×1 卷积进行降维，再进行反卷积操作，之后将不同卷积核的特征输出进行拼接操作。其中：Conv

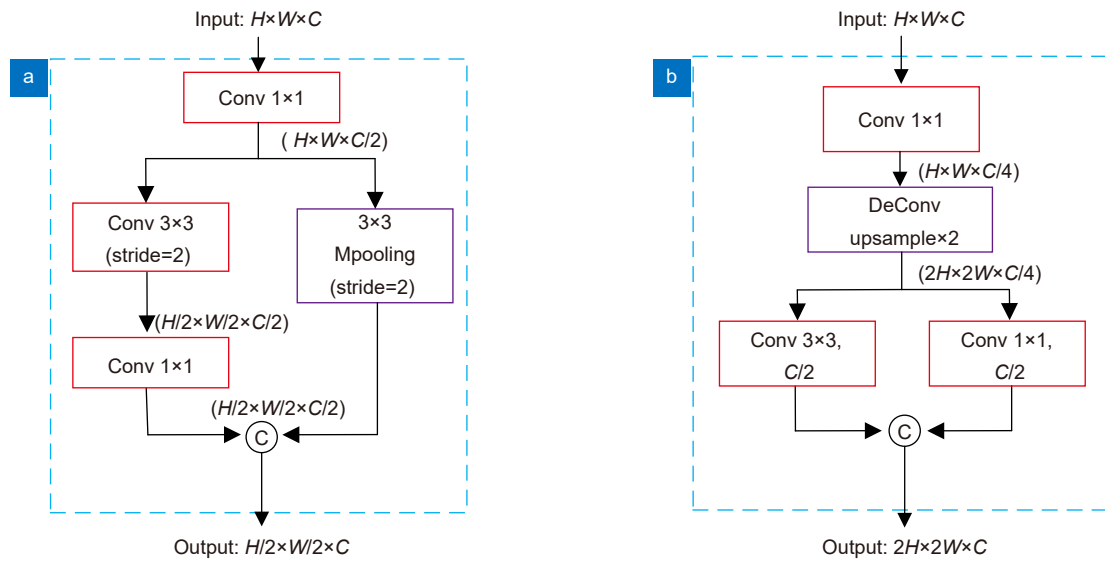


图 4 (a) LaneConv 结构图; (b) LaneDeconv 结构图

Fig. 4 (a) Laneconv structure; (b) Lanedeconv structure

是卷积运算。Mpooling 是最大池化，⊕表示拼接操作。同时，1×1 和 3×3 表示卷积核大小， $H \times W \times C$ 表示张量形状 (高度、宽度、深度)。如表 1 所示，LaneConv 和 LaneDeConv 显然比普通卷积和反卷积具有更高的计算效率。

3.3 引入注意力机制的特征提取模块 (CBAM)

CBAM 模块结合了通道注意力与空间注意力来提高特征的表达能力。在 ResNet, Vgg 等分类网络中加入 CBAM 模块，分类效果会有明显的提升^[13]。而在 Faster-RCNN^[19] 等目标检测网络中加入 CBAM 时，其目标检测效果也会提升^[13]。因此本文尝试在车道线分割网络中加入 CBAM 模块，以增强网络的车道线特征提取能力。

在判断像素是否属于车道线时，需要考虑不同通道特征图上的特征信息。每个通道对于不同任务的重要性是不同的。通道注意力通过赋予每个通道不同的权重，以增强判断重要特征通道的能力。空间注意力负责捕获全局上下文信息，可以将所有车道线像素的

位置信息与特征图上的像素相结合，有助于推断出模糊像素。下面给出空间注意力模块的作用原理。

如图 5(a) 所示，特征 $F \in R^{H \times W \times C}$ 作为输入，接着使用平均池化和最大池化操作将 F 转换为平均池化特征 F_{avg}^c 和最大池化特征 F_{max}^c 。然后，将这两个输出特征送到可以融合所有通道特征的多层感知器 (multiple perceptron, MLP)。最后，使用逐元素求和将两个输出特征向量进行合并，以生成通道注意力，该通道注意力机制计算方式可以表示为

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(w_1(w_0(F_{avg}^c)) + w_1(w_0(F_{max}^c))), \quad (1)$$

其中： σ 表示 sigmoid 函数， w_0 和 w_1 表示 MLP 权重， $w_0 \in R^{C/r \times C}$ 和 $w_1 \in R^{C \times C/r}$ ，其中减速比 $r = 1/8$ 。

如图 5(b) 所示，假设将特征编码后的最后一层特征 $F \in R^{H \times W \times C}$ 作为输入，分别转换为三个特征图，其中 $\{F_1, F_2, F_3\} \in R^{C \times H \times W}$ 。 F_1 和 F_2 用于计算远程相关性； F_3 用于计算远程依赖关系。然后，将它们重塑为 $R^{C \times N}$ ，其中 N 是像素数。 C_{ij} 表示特征图中第 i 位置和第 j 位置之间的相关性，如下所示：

表 1 参数量和计算量对比

Table 1 Comparison of parameters and computations

Name	Parameters	Computations
3*3 Conv	$9C^2$	$9HWC^2$
LaneConv	$3C^2$	$9HWC^2/8$
2*2 DeConv	$4C^2$	$4HWC^2$
LaneDeConv	$7C^2/4$	$7HWC^2/4$

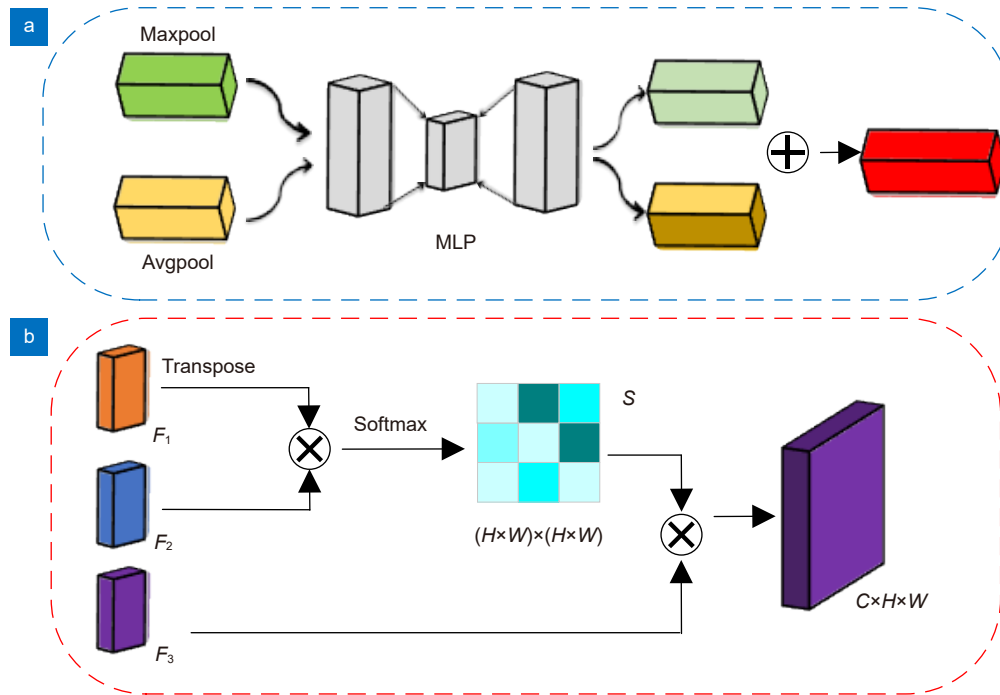


图 5 (a) 通道注意力; (b) 空间注意力结构图
Fig. 5 (a) Channel attention; (b) Spatial attention

$$C_{ij} = F_1(x_i)^T F_2(x_j). \quad (2)$$

应用 Softmax 函数将相关 C_{ij} 归一化为 $S_{ij} \in [0, 1]$ 并计算空间注意力图 $S \in R^{N \times N}$, 如下所示:

$$S_{ij} = \frac{\exp(C_{ij})}{\sum_{i=1}^N \exp(C_{ij})}, \quad (3)$$

其中: S_{ij} 衡量第 i 位置对第 j 位置的影响。两个位置的特征表示越相似, 它们之间的相关性就越高。最后, 本文在 S 和 $F_3(x_i)$ 的转置之间执行矩阵乘法, 并将最终输出重塑为 $R^{C \times H \times W}$, 表示为 $X = (x_1, x_2, \dots, x_j, \dots, x_{H \times W})$ 。每个最终特征图可以计算为

$$X_j = \sum_{i=1}^N s_{ij} F_3(x_i). \quad (4)$$

3.4 二值分割

训练二值分割是将车道线像素二值化, 从而指明哪些像素属于车道线, 哪些像素属于背景。为了处理遮挡或阴影情况, 本文在构造标签时将真实车道线上的点连接起来以形成一条车道, 其中包括诸如汽车遮挡或车道线褪色之类的对象。另外, 使用标准交叉熵损失函数用于计算该分支的损失, 其输出的结果是一维的二值图。对于像素 x 的特定类别, 其损失函数计算如下:

$$L_{\text{bin}}(x, \text{class}) = -\log \left(\frac{\exp(x[\text{class}])}{\sum_j \exp[x[j]]} \right). \quad (5)$$

3.5 像素聚类及曲线拟合

像素聚类及曲线拟合是对网络输出二值分割结果的后处理。为了后续的车道线拟合以及使车道线形状更加平整, 本文使用 DBSCAN 聚类对分割得到的像素进行聚类。DBSCAN 聚类是一种基于密度的聚类, 它将同一类别的像素聚集在一起形成一条车道线。DBSCAN 聚类需要设置两个参数 ϵ 和 $MinPts$, 其中参数 ϵ 是指邻域的距离阈值, 某个像素点 ϵ 半径范围以内区域称之为其邻域, 即如图 6 所示的红色虚线圆所示, $MinPts$ 是指某一邻域中像素点的数量阈值, 某像素点邻域如果至少含有 $MinPts$ 个车道像素点, 则称它为核对象, 如图 6 中红色实线小圆所示。DBSCAN 聚类的过程就是通过将核对象和其邻域内的像素点聚类成一条车道线, 而不在红圈范围内的像素点则会被当作噪点。其聚类过程如图 6 所示, 需要注意的是, 两个核对象要在同一个邻域内, 才会被聚成一个类簇。

网络的输出是二值图, 如图 7(a) 所示, 对二值图的像素进行聚类后, 得到如图 7(b) 所示的聚类输出图。因为 Tusimple 数据集评估需要得到车道线的具体坐

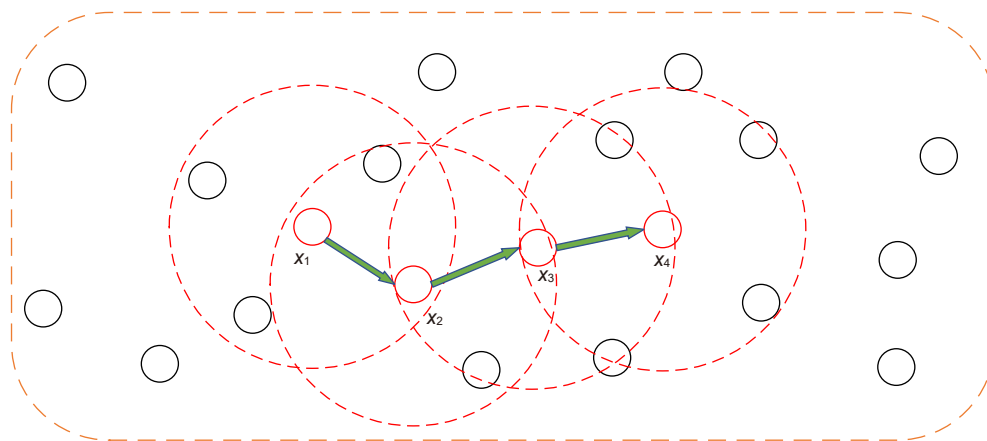


图 6 DBSCAN 聚类过程

Fig. 6 DBSCAN cluster

标, 所以还需对聚类后的车道线进行拟合, 拟合效果输出如图 7(c) 所示。为了提升网络的检测速度以及进行端到端的输出, 本文使用了简单的三次多项式对车道线进行拟合。

4 实验结果与分析

4.1 Tusimple 数据集

本文采用 Tusimple 数据集来验证模型的网络性能。Tusimple 数据集是目前最常用的车道线检测数据集, 它含有 3626 张训练图片和 2782 张测试图片, 包含多种天气状况, 交通状况。车道线多样性方面包含了 2 车道、3 车道、4 车道, 以及少量的 5 车道。图像的注释是 json 格式的文件, 文件中存储着车道线上离散点的位置信息, 即初始标注是以点的连续车道线

点的形式。在实际使用时, 本文将这些车道线点连接起来转换成车道线的二值图。

4.2 评价指标

实验结果采用与真实标记相同的格式进行评价, 即每个车道线用一组固定 y 坐标间隔的点表示。因此, 需要对车道分割结果进行后处理后拟合车道线点, 本文采用三次多项式拟合作为后处理。主要评价指标由 Tusimple 数据集官方提供的准确度 (Acc), 假阳性率 (FP) 与假阴性率 (FN) 指标, 计算方式如下所示:

$$Acc = \sum \frac{C_{pred}}{C_{gt}}, \quad (6)$$

其中: C_{pred} 是预测正确点, 即预测点和真实点在一定阈值内, C_{gt} 是真实点。除了准确度, 还提供了假阳性率 (FP) 与假阴性率 (FN) 指标:

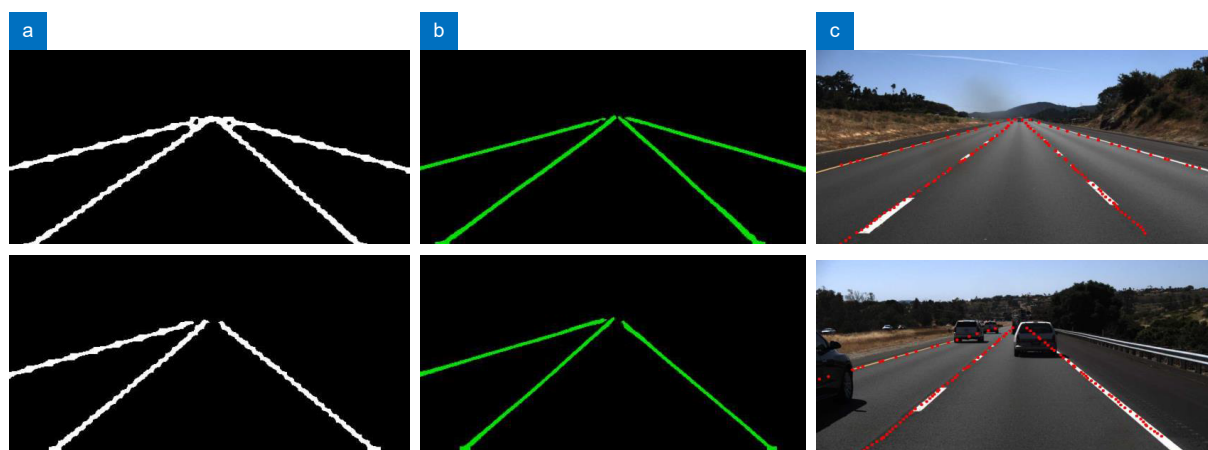


图 7 不同阶段的输出图。

(a) 二值输出图; (b) 聚类输出图; (c) 拟合输出图

Fig. 7 The output in different stages.

(a) Binary output; (b) Clustering output; (c) Fitting output

$$FP = \frac{F_{\text{pred}}}{N_{\text{pred}}}, \quad (7)$$

$$FN = \frac{M_{\text{pred}}}{N_{\text{gt}}}, \quad (8)$$

其中: F_{pred} 是预测错误的车道线数量, N_{pred} 是预测正确的车道线数量, M_{pred} 是没有正确预测真实车道线数量, N_{gt} 是真实的车道线数量。

另外, 由于本文的模型是语义分割网络, 那么均交并比 (mIoU) 也将作为一个重要的评估指标。在本文中车道线的宽度设为 10 pixels, mIoU 计算方式可表示为

$$mIoU = \frac{1}{n_c} \sum_i \frac{n_{ii}}{\left(t_i + \sum_j n_{ji} - n_{ii} \right)}, \quad (9)$$

其中: n_c 是类别数, t_i 是真实情况下类别 i 的像素总数, n_{ij} 表示预测属于类别 j 的 i 类像素的数量。

4.3 结果对比

4.3.1 可视化评估

本文方法与 LaneNet^[11] 都是一个端到端的车道线分割网络。所以本文主要与 LaneNet 进行可视化结果对比。数据均采用 Tusimple 数据集中的图片, 对比结果如图 8 所示。

图 8 中红色虚线框是两个网络的效果对比。可以看出, LaneNet 对车道线细节部分处理的并不好, 而本文的检测效果与真值相比, 在各种情况下都能提供完美的检测结果。红色虚线框代表了因为遮挡等原因造成的车道线漏检, 而本文的网络可以借助 CBAM 恢复被遮挡的车道线。可以看出, LaneNet 在红色框中生成错误的检测结果, 具体地说, LaneNet 没有很好地检测到图 8 中遮挡的车道线, 而本文的网络可以更清楚地检测到这些被遮挡的车道线。在第 3 行中, LaneNet 仅仅识别出小部分未被遮挡的车道线, 而本

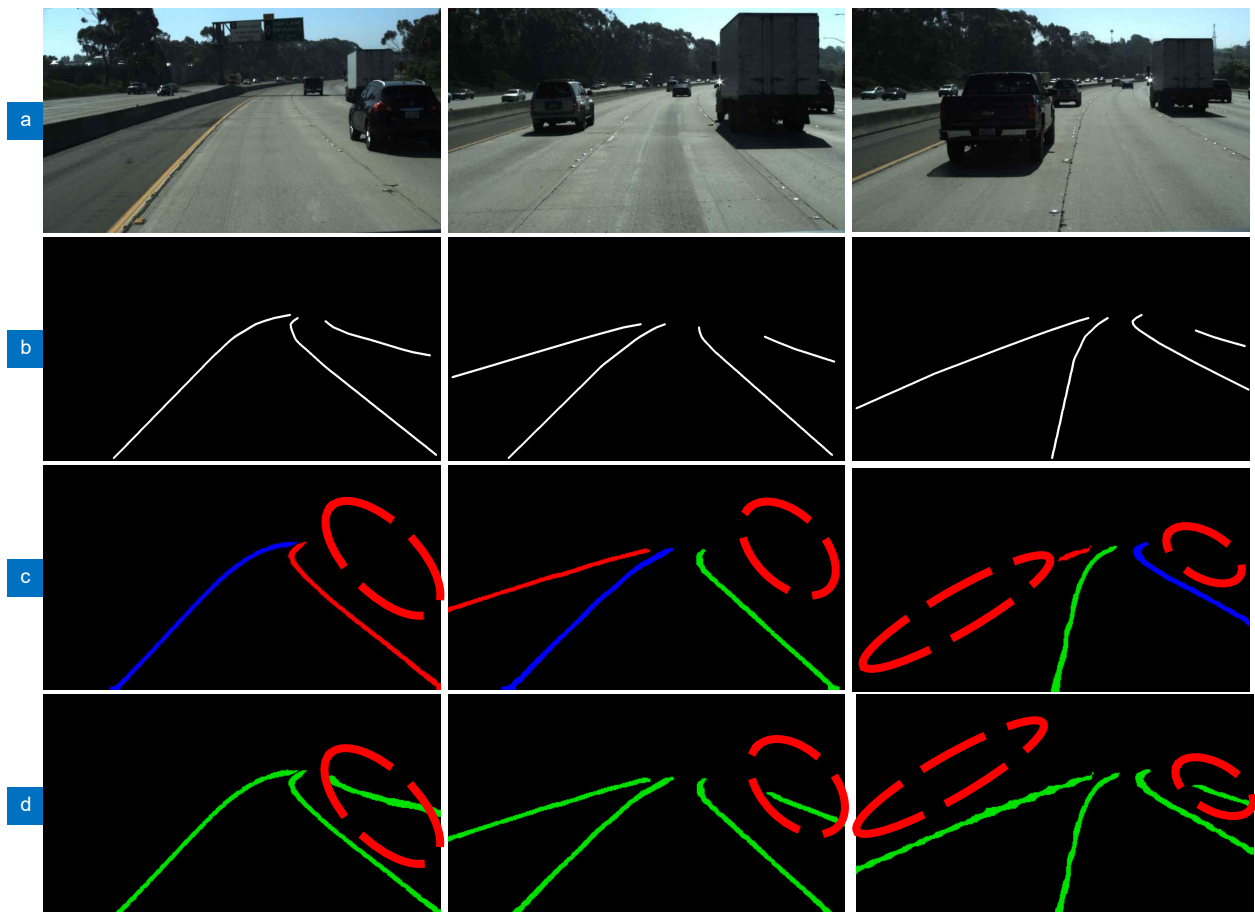


图 8 在 TuSimple 数据集下本文方法和基准的可视化结果比较。

(a) 原始场景; (b) 真值; (c) LaneNet 结果; (d) 本文结果

Fig. 8 Comparison between visualization results of baseline and our method on TuSimple.

(a) Original scene; (b) True value; (c) Lanenet results; (d) Results of our method

文网络可以完整准确地识别出这些车道线。此外, 为了探究 CBAM 对网络的影响, 本文还进行了消融实验, 效果对比如图 9 黄色虚线部分所示, 可以看出加入了 CBAM 之后, 能够提升网络的车道线分割精度。

图 10 显示了本文的网络在 Tusimple 数据集的测试集上的一些经典场景下生成的视觉结果。这些场景包括遮挡、阴影、多车道、曲线等情况。可以看出, 本文的方法在各种情况下均表现良好。

4.3.2 定量评估

在定量评估中, 本文选取了几个最近几年提出的车道线检测方法来与本文方法的数据进行定量比较。这些方法包括了基于检测的方法和基于分割的方法。使用的数据集为 Tusimple 数据集, 对比结果如表 2 所示。

可以看出本文的网络的 mIoU 达到了 64.46%, 超过了表中大部分基于分割的车道线识别方法, 这表明

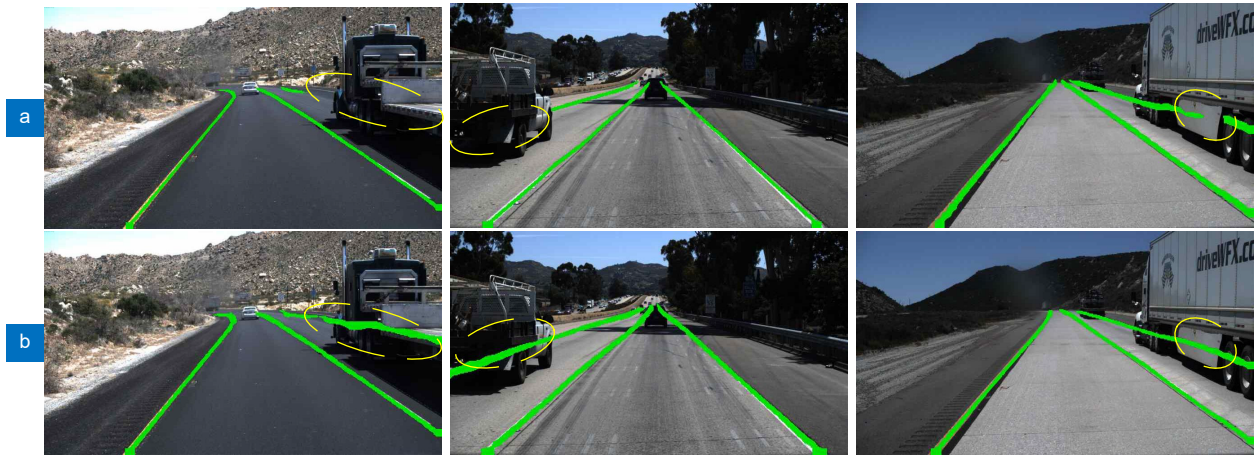


图 9 加入 CBAM 前后效果对比。

(a) 未加入 CBAM; (b) 加入了 CBAM

Fig. 9 Comparison of effects before and after adding CBAM.

(a) Not joined CBAM; (b) Joined CBAM

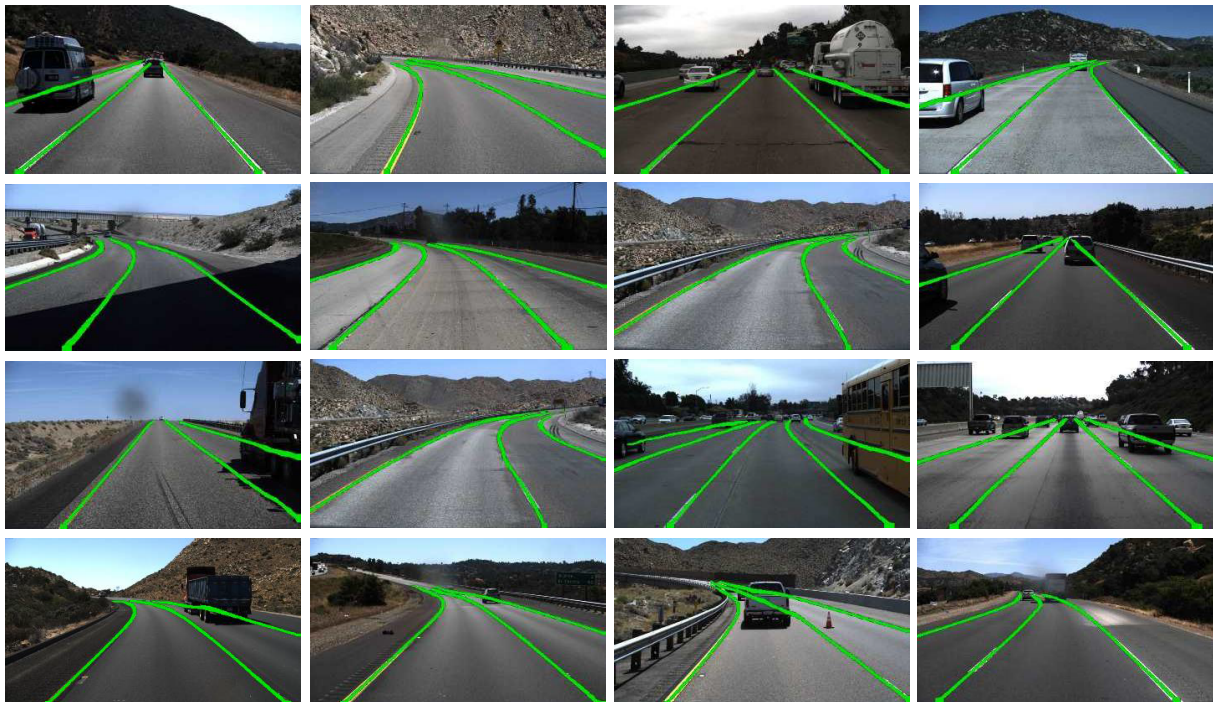


图 10 本文方法在某些典型情况下生成的视觉结果

Fig. 10 Visual results generated by our method on some of typical scenarios

表 2 与其他方法在 Tusimple 数据集上的比较结果
Table 2 Comparison results with other methods on tusimple dataset

方法		Acc/(%)	FP/(%)	FN/(%)	Speed/(f/s)	mIoU/(%)
基于检测的方法	PointLaneNet	96.34	4.67	5.18	71.0	N/A
	PolyLaneNet	93.36	9.42	9.33	115.0	N/A
基于分割的方法	SCNN	96.53	6.17	1.80	7.5	57.37
	VGG-LaneNet	94.03	10.2	11.0	1.7	41.34
	LaneNet	94.42	9.0	9.0	62.5	56.59
	Method 1	94.34	9.1	8.4	102.4	56.08
	Method 2	95.70	8.3	4.31	58.6	65.22
	Method 3	95.64	8.5	4.45	98.7	64.46

注意: 表中的N/A表示相关论文未提及或无法复制该项目, Method 1加入了新引入的卷积结构但未加入CBAM, Method 2使用普通卷积结构但加入了CBAM, Method 3同时加入了CBAM和新引入的卷积结构

了本文能够很好的将车道线从场景中分割出来。需要注意的是, 基于检测的车道线识别方法无法使用 mIoU 这一指标来评估。另外, 由于 *Acc*, *FP* 和 *FN* 这三个评估指标十分依赖于后处理, 而本文仅使用了简单的三项式拟合, 这样的拟合在弯道情况下可能会出现拟合偏差。而 PointLaneNet 可以直接得到车道线点的坐标, 更加契合 Tusimple 数据集的评估方式, 因此本文方法在评估指标上略差于 PointLaneNet。而 SCNN 使用了更加复杂的后处理, 因此评估指标上更加优秀但效率上要稍差。然而在分割网络对比中, mIoU 能够更加直观反应模型是否能够精确地将车道线分割出来。从表 2 的 mIoU 对比中, 本文方法体现出明显的优势, 也反应了 CBAM 能够更好地提高车道线像素识别的精度。此外, 为了了解新引入的卷积结构除提高检测速度外, 对车道线检测精度的影响, 本文加入了消融实验, 结果如表 2 中本文方法 2 所示。和本文方法 3 的对比结果显示, 新卷积结构的加入会对分割精度有一定程度的降低, 但是分割速度会有明显的提升。

在速度方面, 本文的网络可以达到 98.7 f/s 的速度, 能够完成实时的车道线识别任务, 明显优于其他基于分割的车道线识别网络, 但要稍逊于基于检测的车道线识别网络。因为分割网络计算量要比检测网络更大, 而且本文的网络还需要对像素进行聚类, 增加了计算量。总的来说, 本文的模型在检测精度和速度方面展现了更综合的优势。

5 总结

本文提出了一种高效的车道线语义分割方法, 通过编码器和解码器结构以及加入轻量化卷积, 并结合 CBAM 改善了车道线分割网络的速度与精度。实验表明, CBAM 能够提升分割的细节和增强网络的全局信息捕获能力, 从而提高了车道线像素分割能力。此外, 本文设计了高效的卷积结构 LaneConv 和 LaneDeconv 并引入了深度可分离卷积来降低模型计算复杂性, 提高运行速度。在 Tusimple 数据集上进行的定量和定性实验, 结果表明, 与其他方法相比, 本文的模型具有综合优势。

参考文献

- [1] Wang Y, Shen D G, Teoh E K. Lane detection using spline model[J]. *Pattern Recognit Lett*, 2000, 21(8): 677-689
- [2] Jiang R Y, Reinhard K, Tobi V, et al. Lane detection and tracking using a new lane model and distance transform[J]. *Mach Vis Appl*, 2011, 22(4): 721-737
- [3] Zhu H Y, Yang F, Gao X Q, et al. A fast lane detection algorithm based on cascade Hough transform[J]. *Comput Technol Dev*, 2021, 31(1): 88-93
朱鸿宇, 杨帆, 高晓倩, 等. 基于级联霍夫变换的车道线快速检测算法[J]. *计算机技术与发展*, 2021, 31(1): 88-93
- [4] Tian J, Liu S W, Zhong X Y, et al. LSD-based adaptive lane detection and tracking for ADAS in structured road environment[J]. *Soft Comput*, 2021, 25(7): 5709-5722
- [5] Qin Z Q, Wang H Y, Li X. Ultra fast structure-aware deep lane detection[C]//16th European Conference on Computer Vision, 2020: 276-291.
- [6] Chen Z P, Liu Q F, Lian C F. PointLaneNet: efficient end-to-end CNNs for accurate real-time lane detection[C]//2019 IEEE Intelligent Vehicles Symposium (IV), 2019: 2563-2568.

- [7] Tabelini L, Berriel R, Paixão T M, et al. PolyLaneNet: lane estimation via deep polynomial regression[C]//2020 25th International Conference on Pattern Recognition (ICPR), 2021: 6150–6156.
- [8] Ji G Q, Zheng Y C. Lane line detection system based on improved Yolo V3 algorithm[Z]. Research Square: 2021. <https://doi.org/10.21203/rs.3.rs-961172/v1>.
- [9] Neven D, De Brabandere B, Georgoulis S, et al. Towards end-to-end lane detection: an instance segmentation approach[C]//2018 IEEE Intelligent Vehicles Symposium (IV), 2018: 286–291.
- [10] Liu B, Liu H Z. Lane detection algorithm based on improved Enet network[J]. *Comput Sci*, 2020, 47(4): 142–149
刘彬, 刘宏哲. 基于改进Enet网络的车道线检测算法[J]. *计算机科学*, 2020, 47(4): 142–149
- [11] Tian J, Yuan J Z, Liu H Z. Instance segmentation based lane line detection and adaptive fitting algorithm[J]. *J Comput Appl*, 2020, 40(7): 1932–1937
田锦, 袁家政, 刘宏哲. 基于实例分割的车道线检测及自适应拟合算法[J]. *计算机应用*, 2020, 40(7): 1932–1937
- [12] 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: 7276–7283.
- [13] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the 15th European Conference on Computer Vision (ECCV), 2018: 3–19.
- [14] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[Z]. arXiv:1704.04861, 2017. <http://www.arxiv.org/abs/1704.04861>.
- [15] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[C]//4th International Conference on Learning Representations, 2016.
- [16] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1800–1807.
- [17] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 2818–2826.
- [18] Wu B C, Wan A, Yue X Y, et al. SqueezeSeg: convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud[C]//2018 IEEE International Conference on Robotics and Automation (ICRA), 2018: 1887–1893.
- [19] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137–1149

作者简介



张冲 (1994-), 男, 硕士研究生, 主要从事计算机视觉、机器学习的研究。

E-mail: 1368033001@qq.com

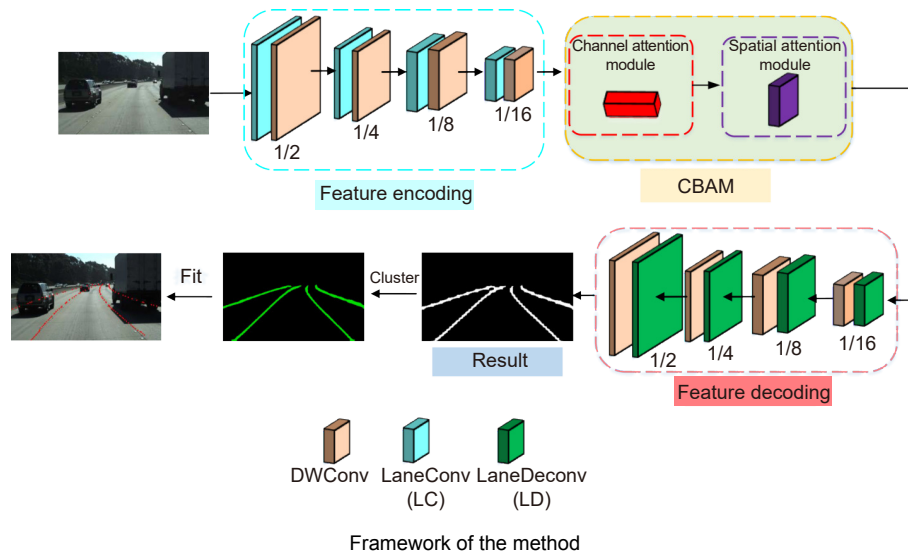


【通信作者】黄影平 (1966-), 男, 教授, 主要从事汽车电子、计算机视觉的研究。

E-mail: huangyingping@usst.edu.cn

Real-time lane detection method based on semantic segmentation

Zhang Chong, Huang Yingping*, Guo Zhiyang, Yang Jingyi



Overview: In recent years, the rapid development of neural network has greatly improved the efficiency of lane detection. However, convolutional neural network has become a new problem restricting the development of lane detection because of its large amount of calculation and high hardware requirements. Lane detection methods based on deep learning can be divided into two categories: detection based methods and segmentation based methods. The method based on detection has the advantages of high speed and strong ability to deal with straight lane. However, when the environment is complex and there are many curves, the detection effect is obviously not as good as the segmentation based method. This paper adopts the segmentation based method, and considers that the performance of lane detection can be improved by establishing global context correlation and enhancing the effective expression of important Lane feature channels. Attention mechanism is a model that can significantly improve network performance. It imitates the human visual processing mechanism, strengthens the attention to important information, so as to reasonably allocate network resources and improve the detection efficiency and accuracy of the network. Therefore, this paper uses the CBAM model. In this model, channel attention and spatial attention are serial to obtain better feature representation ability. Spatial attention learns the positional relationship between lane line pixels, and channel attention learns the importance of different channel features. In addition, in order to solve the problem of complex convolution calculation and slow running speed based on segmentation model, a more efficient convolution structure is proposed to improve the computational efficiency. A new fast down sampling module laneconv and a new fast up sampling module laneconv are introduced, and the depth separable convolution is introduced to further reduce the amount of calculation. They are located in the coding part of the network. The decoding part outputs the binary segmentation result. Then, the results are clustered by DBSCAN to obtain the lane line. After clustering, compared with the complex post-processing in other literature, this paper only uses simple cubic fitting to fit the lane line, which further improves the speed. Therefore, the running speed of the model proposed in this paper is better than most segmentation based methods. Finally, a large number of experiments are carried out on tusimple Lane database. The results show that the method has good robustness under various road conditions, especially in the case of occlusion. Compared with the existing models, it has comprehensive advantages in detection accuracy and speed.

Zhang C, Huang Y P, Guo Z Y, et al. Real-time lane detection method based on semantic segmentation[J]. *Opto-Electron Eng*, 2022, 49(5): 210378; DOI: [10.12086/oe.2022.210378](https://doi.org/10.12086/oe.2022.210378)

Foundation item: The Shanghai Natural Science Foundation of Shanghai Science and Technology Commission, China (20ZR14379007), and National Natural Science Foundation of China (61374197)

School of Optical-Electronic and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

* E-mail: huangyingping@usst.edu.cn