[文章编号] 1003－4684(2025)05-0056-07

基于语义分割的实时车道线检测方法

张富杰，柯敏毅

（湖北工业大学计算机学院，湖北 武汉 430068）

[摘 要] 为提升自动驾驶中车道线的检测效果，提出一种基于多尺度语义特征和注意力机制的车道检测网络。网络采用编解码结构，编码器基于轻量化网络 MobileNetV2，能实时且准确提取车道线特征；针对车道检测中语义特征没有充分利用的问题，在编码器和解码器间增加多尺度特征融合模块，以提高模型的分割精度，同时引入混合注意力机制，结合全局和局部的多尺度特征信息，实现对车道线的准确定位。在 Tusimple 数据集上的测试表明，与现有车道检测模型相比，所提方法具有更好的分割精度和较好的速度提升。

[关键词] 自动驾驶；语义分割；多尺度特征融合；注意力机制

[中图分类号] TP391.4 [文献标识码] A

车道检测是车辆环境感知的核心技术[1]，必须快速、准确识别出车道线，为后续模块提供保障。真实道路复杂多样，车道检测时容易受环境影响而错检、漏检。复杂场景车道检测是极具挑战性的任务，也是当前研究的难点。

车道线检测分为基于特征、基于模型和基于深度学习的检测[2]，通过颜色、结构等特征与霍夫变换结合，对结果后处理，获得车道线信息，是基于特征检测的思路[3]。王鑫[4]等基于感兴趣区域和扫描线，结合霍夫变换拟合备选点集，获得结果。刘源[5]等基于边缘特征点聚类，结合边缘的梯度、灰度分布提取最优点，拟合获得结果。这类方法计算量小，实时性好，但易受环境影响，鲁棒性不强。

基于模型方法使用数学建模，将问题转化为模型求参，通过最小二乘法、RANSAC 等方法拟合车道线[6]。Shin[7]等基于顶视图设计算法，构建曲线模型来扩展 ROI(感兴趣区域)底部的车道线。这类方法对特定场景的识别精度高，但模型复杂度会受环境影响，增加计算成本，影响实时性。

Lee 等设计了端到端多任务处理网络 VPG-Net[8]，通过灭点引导检测车道线。Pan 等设计了空间卷积网络 SCNN[9]，逐片卷积提升模型对细、长分割对象的学习能力，能处理路面闭塞、磨损等情况，但速度不足。Neven 等提出 LaneNet[10]，通过语义分割和聚类，将问题转化为实例分割，但对空间信息利用不足，车道不清晰或被遮挡后效果会下降。Hou 等提出自注意力蒸馏 SAD[11]，编码器间通过相互学习提高特征识别能力，推理时间少，但同样对空间信息利用不足。与传统方法比，深度方法速度快、效率高、鲁棒性好，但效果受网络性能影响，合理搭建网络，让模型在实时、精度上保持平衡，一直是业界关注的重点。

本文基于语义分割的思想，使用编解码结构设计了一种轻量实时车道分割网路。采用轻量级网络 MobileNetV2 作为主干，改进常规分割网络实时性不足的问题。针对空间信息利用不足的问题，主干网络中采用空洞卷积提升网络感受野，引入混合注意力的 MFAM 模块，对车道线精准定位，简化空洞空间金字塔池化模块 TASPP，加强模型的特征提取能力，以提高分割精度。

1 基于多尺度语义特征与注意力机制的车道线检测网络

车道检测对车道线空间位置的检测精度要求高，若使用目标检测，将输出目标所在的矩形框，需要在后处理时矫正精度，会增加后处理难度；而语义分割是逐像素点分类，针对车道检测，精度相对目标检测高，不会增加后处理难度[12]。本文算法将车道检测转化为逐像素点分类问题，先获得车道线的分割二值图，通过聚类得到车道线掩码，再对所得车道线进行后处理，最终得到检测结果。

很多优秀的语义分割网络都是重编码器轻、解

[收稿日期] 2023－01－04

[第一作者] 张富杰(1996－)，男，湖北监利人，湖北工业大学硕士研究生，研究方向为图像处理。

[通信作者] 柯敏毅(1967－)，男，湖北黄石人，湖北工业大学教授，研究方向为计算机控制系统，智能交通。

码器结构,减少了计算量,也取得了很好的分割效果,本文同样采用此结构。编码器通过连续的卷积和池化获得长范围的上下文信息,以获得更好的分类结果。但是,下采样会降低特征图的分辨率,导致图像细节的丢失,需要在解码器上采样以恢复图像分辨率。解码器将编码器学习到特征信息的语义投射到像素空间中,得到密集像素点的分类。本文网络结构如图 1 所示。网络由三部分组成:编码器,解码器和中间的特征强化模块,其中编码器采用改进的 MobileNetV2 网络,保证了特征提取的实时性和效率,倒残差结构是 MobileNetV2 的特殊结构,在最后一层倒残差中使用空洞卷积扩大网络感受野,以提高网络性能。编码器与解码器之间的简化空洞空间金字塔池化 TASPP 模块和多尺度混合注意力特征提取 MFAM 模块均属于特征强化模块。TASPP 有助于获得更丰富的多尺度语义信息,保留更多的车道线细节特征,MFAM 有助于结合上下文信息,使网络更加关注于车道线及其相关信息,强化车道线的定位效果,残差结构能对被误检为车道线的像素点进行特征修正,三者综合提升模型的特征提取能力。解码器使用特殊上采样 laneDeconv 模块适当提升上采样精度,上采样还原图像尺寸,最后对分割结果进行聚类,得到最终检测结果。

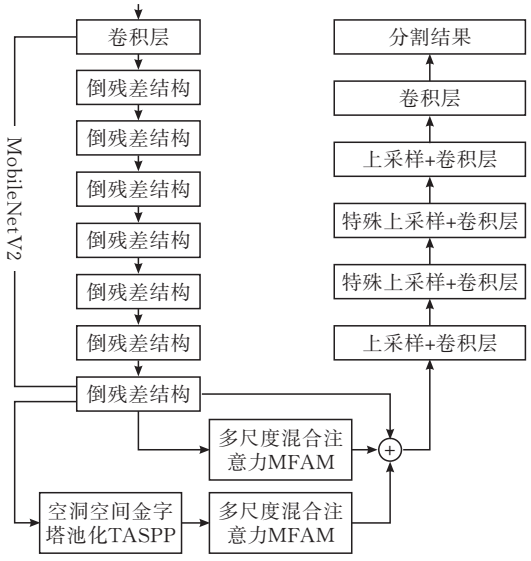


图 1　网络结构

2　车道检测网络模块设计

2.1　编码器

语义分割是逐像素分类,对计算资源要求大,而精度高的网络对网络深度也有一定要求,这些都会导致运行时间长,影响模型速度,车道检测如果不能实时运行,将影响实用性,因而本文挑选轻量实时的 MobileNetV2[13] 作为主干网络。MobileNet2017 年由谷歌提出,旨在高效利用移动设备和嵌入式设备的计算资源,保证模型的准确性。MobileNetV1 采用深度可分离卷积[14]减少计算量和参数,大大降低了计算开销,加快了模型的运行速度;MobileNetV2 基于 V1 提出了倒残差结构,先用 1×1 的卷积扩展通道数(升维),然后用 3×3 的 Dw 卷积提取特征,最后用 1×1 卷积压缩通道(降维);V2 进一步改善了网络性能,在实时性、精度方面取得了不错的平衡。

MobilNetV2 特征提取在前 8 层,在第 9 层会将通道数变为 1280,这会极大增加网络运算量,为节省内存和计算资源,本文选取其前 8 层作为编码器。原先的 MobileNet 不适合做语义分割,因为编码器特征提取所得图片大小为原始图片的 1/32,语义分割会对图像不停卷积以提取特征,受计算资源的限制,输入图像的大小不可能一直保持原有大小,需要对图像进行压缩,在解码器阶段将图像上采样恢复到原来大小,但下采样会降低图像的分辨率,丢失更多的细节,网络没法完全恢复丢失的信息。因此将网络第七层的第一次升维步长 s 修改为 1,使最终编码器输出的图片为原始图片的 1/16,且不损失更多的细节信息。空洞卷积[14]可以在不丢失信息的情况下,增大网络感受野,本文将第 7 层、第 8 层的普通卷积替换成空洞卷积,空洞系数 r 分别设为 2 和 4。

改进后的网络各层结构如表 1 所示,通道扩展倍数 t 即倒残差结构中 1×1 卷积升维倍数,例网络第 3 层倒残差中的升维会将 16 通道变为 96(×6,t=6);输出通道数 c-out 即本层输出特征图的通道数;倒残差重复次数 n 是本层倒残差结构的使用次数,例第 3 层会经过 2 次倒残差结构,每次倒残差会用 1×1 卷积将 16 通道升维为 96(×6),DW 卷积提取特征后,用 1×1 卷积将通道数降回 16;第一次升维步长 s 是第一次倒残差中升维使用 1×1 卷积的步长,为 1 不改变图片大小,为 2 则将输入特征图尺寸减半;空洞系数 r 为 1 表示普通卷积,为 k 表示使用空洞卷积。

表 1　改进后的 MobileNetV2 网络各层结构空洞系数

层号	$c\times h\times w$	操作	t	c-out	n	s	r
1	3×256×512	卷积	—	32	1	2	1
2	32×128×256	倒残差结构	1	16	1	1	1
3	16×128×256	倒残差结构	6	24	2	2	1
4	24×64×128	倒残差结构	6	32	3	2	1
5	32×32×64	倒残差结构	6	64	4	2	1
6	64×32×64	倒残差结构	6	96	3	1	1
7	96×16×32	倒残差结构	6	160	3	1	2
8	160×16×32	倒残差结构	6	320	1	1	4

2.2 多尺度混合注意力特征提取模块(MFAM)

注意力机制源于对人类视觉的研究,人类在处理信息时,会选择性关注重要信息,“忽略”不重要信息,因而在处理复杂信息时能更加快速、高效。注意力机制能决定输入信息里需要关注哪部分和分配有限的处理资源给重要部分,这种机制后来被研究人员引入计算机视觉来提升网络性能。

卷积通过混合通道和空间两个维度的信息来提取特征,因此在通道和空间中加入注意机制有助于从这两个维度提取信息。CBAM[16]是一种同时关注通道和空间的卷积模块,能提升特征图的特征表达能力,并且模块较轻量,可以无缝集成到网络中。

本文提出了一种多尺度混合注意模块(MFAM),利用不同尺度的分支提取注意权重。其中一个分支是通道关注模块,一个分支提取全局特征关注,另一个分支提取局部特征的通道关注,通过整合三分支注意力模块,形成新的通道注意力模块MCAM,再将其输入空间注意力模块进行处理。将该模块加入到车道线分割网络中,能抑制无关信息,加强网络的车道线特征提取能力。MFAM 具体结构如图 2a 所示。判断图中每个像素点是否属于车道线时,要考虑不同通道的信息,每个通道的重要性不同,表现为权重不同,通道注意力(CAM)[15]是赋予通道不同权重的操作,其结构如图 2b,分支一的通道关注结构如图 2c 所示,对于输入特征 F,分别进行平均池化(avgpool)和最大池化(maxpool)获得两个大小为通道数 1×1 的特征图 F^{c}_{avg}、F^{c}_{max},将它们送到可以融合所有通道特征的多层感知器(MLP)中,把得到的两个新的特征图按元素加起来,通过 sigmoid 得到各层的加权系数。计算公式:

$$M_c(F)=\sigma\{\mathrm{fc2}[\mathrm{relu}(\mathrm{fc1}(\mathrm{AvgPool}(F))+\mathrm{fc1}(\mathrm{MaxPool}(F)))]\} \quad (1)$$

σ 是 sigmoid 激活函数,MLP 的具体实现是 fc1、relu 和 fc2,fc1 和 fc2 是卷积层,fc1 将通道数缩小为原来的 1/16,fc2 将通道数恢复到原有大小,relu 是激活函数,将最终所得的各层权重系数和最初输入特征图点乘,获得新特征图。分支二是通过逐点卷积(1×1 卷积)和 3×3 的空洞卷积来提取局部特征的通道关注,分支三通过全局平均池化来提取全局特征的注意力;通过将三个注意力特征融合,网络能更好整合语义和尺度不一致的特征。

空间注意力聚焦于特征图上的有效信息 ,在车道检测中加入有助于结合车道线位置信息来推测模糊像素点的类别。图 2d 是空间注意力的具体结构,对于输入的特征图 F,分别沿 channel 方向平均池化、最大池化,获得两个大小为 1×高×宽的特征图 F^{s}_{max}、F^{s}_{avg},将它们在 channel 维度进行拼接,再通过 3×3 空洞卷积将通道变为 1,最后通过 sigmoid,得到空间注意力权重,计算公式:

$$M_s(F)=\sigma(f^{3\times3}([\mathrm{AvgPool}(F);\mathrm{MaxPool}(F)])) \quad (2)$$

σ 是 sigmoid 激活函数。将通道注意力的输出和空间注意力权重进行逐元素相乘,获得最终特征图。

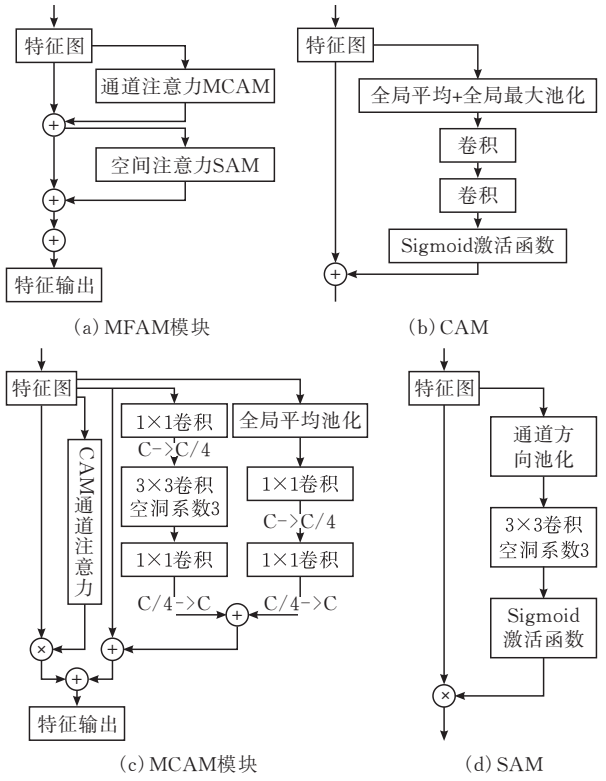


图 2 多尺度混合注意模块说明

2.3 空洞空间金字塔池化 TASPP 模块

ASPP 最早在 Deeplabv2[17]模型中提出,它是通过融合不同尺度的特征、提高网络特征提取能力的并行多尺度卷积模块。考虑车载设备计算力有限,对算法的实时性要求高,而 ASPP 计算消耗大,执行时间长,本文对 ASPP 进行轻量处理即 TASPP,具体结构如图 3 所示。与 ASPP 相比,TASPP 移除了特征池化分支,通过一个 3×3 普通卷积和两个 3×3 的空洞卷积(空洞系数分别为 4 和 8)提取多尺度的特征信息,将这些特征信息在 channel 维度拼接后,送入 1×1 卷积层中获取车道线信息更为丰富的上下文语义特征。ASPP 表明多尺度卷积核可以从低分辨图中较好得到更多高频图像的细节信息,本文为节省算力,统一使用大小为 3 的卷积核,不同空洞系数的空洞卷积可以提供不同的感受野,获得不同尺度的车道线信息。考虑到多尺度信息虽能提升特征提取能力,但可能将道路两侧边缘误检为车道线,因而参考 ResNet 引入跳跃连接的残差结构进行特征修正,防止丢失必要信息。轻量化后的 TASPP 在减少了计算参数、加速算法运行的同

时，有效减少了上下文车道线信息的丢失，添加的捷径分支提高了算法的收敛速度，减少了车道线误检的概率。

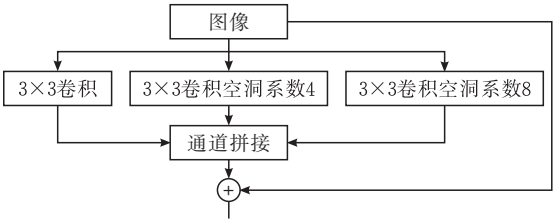


图 3　TASPP 模块

2.4　解码器模块

语义分割实质是对每个像素点分类，区分出哪些像素点属于背景，哪些像素点属于车道，是二分类问题。将得到的特征图在解码器中逐级上采样，恢复图像分辨率，最终得到车道分割结果。

很多 CNN 模型的准确率已经很高了，学界、业界的重点都聚焦于速度和准确率的平衡，希望模型又快又准。普通卷积中每个卷积核都要和每个通道进行计算，运算量大；优化模型的重点是在保证模型精度的情况下，尽量减少网络的计算量和参数，以加快模型的运算速度。VGG[18] 网络提出用多个小卷积替代大卷积，在保证感受野不变的情况下减少运算量；Inception[19] 结构提出用 1×1 卷积核改变图像通道数减少参数量。参考以上想法，本文采用一个特殊上采样模块，结构如图 4 所示，使用 1×1 卷积对特征进行降维，减少之后的运算量，通过转置卷积进行上采样，以各自不同大小的卷积输出特征，在 channel 维进行拼接。

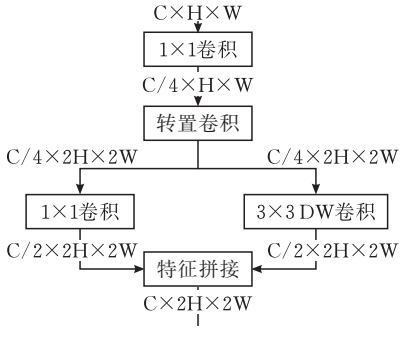


图 4　特殊上采样模块

二分类常用损失函数有交叉熵函数，但车道检测图像中车道线与背景样本高度不均衡，因而本文采用的损失函数是 Focal loss[20]，计算公式：

$$\begin{cases}\text{FocalLoss}(p,y,\alpha)= \\ \quad -\log(p)*\alpha*(1-p)^{\gamma} \qquad \text{if } y=1 \\ \text{Focalloss}(p,y,\alpha)= \\ \quad -\log(1-p)*(1-\alpha)*p^{\gamma} \quad \text{其他}\end{cases} \tag{3}$$

其中，p 为模型预测为正样本的概率，y 为标注值，式中第一项是标准交叉熵，第二项为正负样本平衡参数，第三项为样本难易分类权值。

仅区分车道线和背景还不够，还需获得不同车道的实例，以区分每条车道线。设计像素嵌入分支，将输入图像的每个像素映射到高维空间，得到像素向量嵌入映射，便于计算像素间的距离。该分支将每个像素区分开，使同一车道像素之间的向量距离最小，并使不同车道之间的像素距离最大。参考 Neven D[21] 提出的基于距离度量学习进行实例分割的思路，使用聚类损失函数。损失函数：

$$\begin{cases}L_{\text{var}}=\dfrac{1}{C}\displaystyle\sum_{c=1}^{C}\dfrac{1}{N_c}\sum_{i=1}^{Ne}[\,\|\,\mu_c-x_i\,\|\,-\delta_v]_+^2 \\ L_{\text{dist}}=\dfrac{1}{C(C-1)}\displaystyle\sum_{cB=1,cA\neq cB}^{C}[\delta_d-\|\,\mu_{cA}-\mu_{cB}\,\|\,]_+^2\end{cases} \tag{4}$$

其中，L_{var} 为方差项，它让每个像素都能接近所属车道的平均向量，减少类内距离，从而使同一车道线的像素逐渐聚类；L_{dist} 为距离项，它保持每条车道线的平均向量之间的距离，使不同车道线的像素逐渐远离。C 是车道线数量，N_c 是属于同一车道线的像素点数量，$\boldsymbol{\mu}_c$ 是车道线的均值向量，$\boldsymbol{x}_i$ 是像素向量。通过该分支，成功区分每条车道线。

语义分割分支输出的是二值图，像素嵌入分支输出的是实例分割结果，拟合聚类后的车道掩码，得到各车道线的曲线方程，得到连续光滑的车道线，方便后续相关系统使用。

3　实验结果与分析

3.1　Tusimple 数据集

本文使用图森公司提供的 Tusimple 数据集来验证性能。这是现在常用的车道线检测数据集之一。主要包括直线、曲线、光照、道路破损、外部设施干涉、阴影遮挡等情况的 3626 张训练集图像和 2782 张测试集图像。尺寸为 1280×720。标注文件为 Json 文件，文件有 3 个字段，h_sample 是车道线纵坐标，lanes 对应车道线标注的横坐标，是个二维数组，−2 表示这个点是无效点，raw_file 是每一个数据段第 20 帧图像的路径。通过解读标注文件，获得车道线的二值图像和实例分割图像作为网络的输入数据。

3.2　评估指标

深度学习需要根据模型的预测目标选择合适的评估指标，正确评估模型的效果。Tusimple 数据集官方提供的评估指标是准确度（Acc）、假阳性率（FP）、假阴性率（FN）。Acc 的计算公式：

$$\text{Acc}=\sum\frac{C_{\text{pred}}}{C_{gt}} \tag{5}$$

分子 C_{pred} 是模型预测正确的像素点，分母 C_{gt} 是标

签中所有点。FP 计算公式：

$$FP = \frac{F_{\text{pred}}}{N_{\text{pred}}} \tag{6}$$

分子 F_{pred} 是预测错误的车道线数量，分母 N_{pred} 是总共预测的车道线数量。FN 计算公式：

$$FN = \frac{M_{\text{pred}}}{N_{gt}} \tag{7}$$

分子 M_{pred} 是未能预测出的车道线数量，分母 N_{gt} 是标签中所有车道线数量。

由于进行了语义分割，可选语义分割常用的 mIoU 作为评价指标，其计算公式：

$$\text{mIoU} = \frac{1}{n_c}\sum_i \frac{n_{ii}}{(t_i + \sum_j n_{ji} - n_{ii})} \tag{8}$$

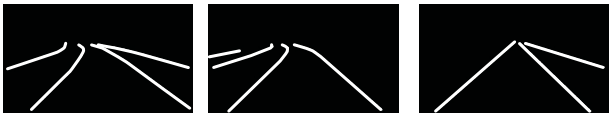
n_c 是类别数，t_i 为真实情况下目标为类别 i 的总像素个数，n_{ji} 表示类别 j 被预测分成类别 i 的像素个数。

3.3 结果对比

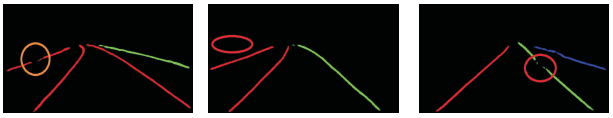
3.3.1 可视化评估 本文方法主要和 LaneNet 进行对比，数据集使用 Tusimple，检测结果如图 5 所示。



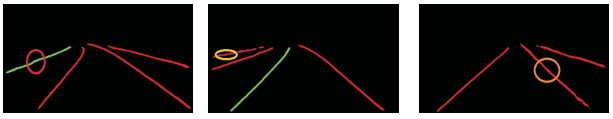
(a)原图



(b)标注图



(c)LaneNet 结果



(d)本文结果

图 5 本文方法和 LaneNet 对比

对比图 5 第 3 行 LaneNet 效果图和第 2 行标注图，可以看出 LaneNet 在车道线的细节处理上有待提升，处理一些被遮挡的车道线时会出现漏检现象。而本文网络使用 TASPP，结合多尺度信息以强化车道线特征，配合 MFAM 结合上下文语义信息，能更为准确地定位被遮挡的车道线位置，因而检测效果更好。图 5 第 4 行清楚地检测出了这部分被车辆或阴影遮挡的像素点，体现了本文方法的优势。

同时，本文设计了消融实验，以探究 TASPP 和 MFAM 对模型的影响。图 6 是本文网络加入 TASPP 模块前后检测效果的对比图。



(a)加入 TASPP 前



(b)加入 TASPP 后

图 6 加入 TASPP 模块前后检测效果的对比

对比图 6 最左边第一张图，可以看出，加入 TASPP 前，网络将最右边的道路边缘误检为车道线，加入 TASPP 后纠正了检测结果；从对比中能够看出，TASPP 能将被遮挡导致漏检的特征点检测出来，起到特征强化的效果。

图 7 是本文网络加入 MFAM 模块前后检测效果的对比图，从第一张图的上下对比可以看出，加入 MFAM 前，网络将道路边缘误检为车道线，加入 MFAM 结合上下文信息，极大修正了检测效果；从对比中能够看出，MFAM 使用的混合注意力能强化车道线定位能力，检测出被遮挡的像素点，从而提升检测精度。



(a)加入 MFAM 前



(b)加入 MFAM 后

图 7 加入 MFAM 模块前后检测效果的对比

图 8 是本文方法在 Tusimple 测试集上一些特定场景的检测效果，可以看出本文方法在曲线、遮挡等条件下都有良好的检测效果。



图 8 本文方法在典型场景的检测效果

3.3.2 定量评估 选取近年来常用的车道检测方法与本文方法进行定量比较。采用 Tusimple 数据集，比较结果见表 2。其中，Method1 使用普通上采样且未使用 TASPP 和 MFAM，Method2 使用 laneDeconv 结构未使用 TASPP 和 MFAM，Method3 使用 laneDecov、TASPP 未使用 MFAM，Method4 使用 laneDecov、MFAM 未使用 TASPP，Method5 同时使用 laneDecov、TASPP、MFAM。

表 2 不同方法结果对比

方法	Acc/%	FP/%	FN/%	Speed/FPS	*mIoU* /%
SCNN	96.53	6.17	1.8	7.5	57.6
LaneNet	94.4	9.03	8.97	62.5	56.6
VGG-LaneNet	94.02	10.3	10.9	2.9	42.4
PolyLaneNet	93.36	9.42	9.33	115	—
Method1	94.12	9.32	8.92	76	56.37
Method2	94.31	9.15	8.46	73	56.53
Method3	95.58	8.41	4.47	71.3	64.7
Method4	95.63	8.33	4.35	69.8	65.2
Method5	95.69	8.31	4.33	66	65.6

Acc、FP、FN 这些指标极其依赖后处理，SCNN 使用了相对复杂的后处理，因而其评估数值更精细，但由于进行逐切片卷积，极大影响了检测速度，无法达到实时水平；PolyLaneNet 使用的是多项式回归，实时性高，但 Acc 相对不足，并由于不是基于分割，故未采用 mIoU 评估；LaneNet 能够达到实时标准，更换更合适的主干网络能提升网络性能。

mIoU 指标能精确地反映出模型的车道线分割能力，从表中 mIoU 中能够看出本文方法相对其他方法具有不错的优势，方法二、方法三、方法四之间的对比也说明加入 TASPP、MFAM 模块确实可以起到提高识别精度的效果；方法二和方法一对比，可以看出 laneDeconv 以牺牲少量速度为代价，换取模型准确性和 mIoU 上的提升；本文的网络速度达 66 fps，可以完成车道线的实时检测任务，并在分割网络里效果优秀；但分割方法针对像素，需要进行聚类，计算量更大，因而检测速度会逊于基于检测的方法，故速度上不及 PolyLaneNet。综上所述，本文方法在检测速度与精度之间获得了良好平衡，在保证实时的同时，能够准确检测出车道线。

4 结论

本文基于语义分割经典编解码器结构，提出了一种高效实时的车道线检测网络。该网络通过选取轻量 backbone 加快网络的特征提取速度，采用 TASPP 和 MFAM 模块增强车道特征提取能力，提高分割精度。系列实验表明，TASPP 能有效结合多尺度特征信息提升模型的特征提取能力，MFAM 能结合上下文语义信息强化车道线的定位效果，提高网络的车道线分割能力。在 Tusimple 数据集上开展定量、定性实验后发现，可以看到本文网络的 mIoU 达到了 65.6%，性能超过参与对比的大部分分割方法，表明模型具有良好的车道线分割能力，能很好地将车道线从场景中分割出来，同时，模型运行速度达 66 fps 保证了模型能够实时运行。

[参 考 文 献]

[1] 范延军.基于机器视觉的先进辅助驾驶系统关键技术研究[D].南京：东南大学，2016.

[2] 赵紫霞.基于卷积神经网络的车道线检测技术研究[D].北京：北京交通大学，2020.

[3] 陈戈珩，潘晓旭，侯作辉.基于车道线识别和多特征的前车检测算法[J].科学技术与工程，2016，16(15)：1671-1815.

[4] 王鑫，刘玉超，海丹.基于双重 ROI 和变间距扫描的车道线检测[J].指挥与控制学报，2017，3(02)：154-159.

[5] 刘源，周聪玲，刘永才，等.基于边缘特征点聚类的车道线检测[J].科学技术与工程，2019，19(27)：247-252.

[6] 叶飞，刘子龙.一种采用区域知识蒸馏网络的车道线检测算法[J/OL].(2021-10-25).[2022-06-08].https://kns.cnki.net/kcms/detail/21.1106.tp.20211019.1525.004.html.

[7] SHIN J，LEE E，KWON K K，et al.Lane detection algorithm based on top-view image using random sample consensus algorithm and curve road model[C].2014 Sixth International Conference on Ubiquitous and Future Networks(ICUFN). IEEE，2014：1-2.

[8] LEE S U，KIM J，YOON J，et al. VPGNet：vanishing point guided network for lane and road marking detectionand recognition[C].IEEE International Conference on Computer Vision，2017：1965-1973.

[9] PAN X G，SHI J P，LUO P，et al. Spatial as deep：spatial CNN fortraffic scene understanding[C]. Thirty-Second AAAIConference on Artificial Intelligence，2018：7276-7283.

[10] 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.

[11] HOU Y, MA Z, LIU C, et al. Learning lightweight lane detection cnns by self attention distillation[C]. Proceedings of the IEEE International Conference on Computer Vision. 2019: 1013-1021.

[12] 赵泽威,杨雪银,员富强.基于自注意力机制的车道线检测算法[J].长江信息通信,2021,34(09):24-27.

[13] SANDLER M, HOWARD A, ZHU Menglong, et al. Mobilenetv2: inverted residuals and linear bottlenecks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt lake, Utah,2018:4510-4520.

[14] AGHOWARD, ZHU M, CHEN B,et al.MobileNets: efficient convolutional neural networks for mobile vision applications [DB/OL]. (2017-04-17). [2022-08-09].https: // doc.taixueshu.com/foreign /arXiv170404861.html.

[15] 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 & Machine Intelligence, 2016, 40 (04): 834-848.

[16] WOO S,PARK J,LEE J Y,et al. CBAM:convolutional block attention module[C] // Proceedings of the 15th EuropeanConference on Computer Vision (ECCV), 2018:3-19.

[17] CHEN L C, PAPANDREOU G, SCHROFFF, et al. Rethinking atrous convolution for semantic image segmentation[DB/OL]. (2017-12-05). [2022-09-12]. https: // doc. taixueshu. com/foreign/arXiv170605587.html.

[18] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[DB/OL].(2014-04-10).[2022-09-14].https: // doc.taixueshu.com/foreign/arXiv14091556.html.

[19] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking theinception architecture for computer vision [C]. 016 IEEEConference on Computer Vision and Pattern Recognition(CVPR), 2016: 2818-2826.

[20] LINTY,GOYAL P,GIRSHICK R,et al.Focal loss for dense objectdetection[DB/OL].(2018-02-07).[2022-09-18]. https: // doc. taixueshu. com/foreign/arXiv170802002.html.

[21] DE BRABANDERE B, NEVEN D,et al.Semantic instancesegmentation with a discriminative loss function [DB/OL].(2017-08-08).[2022-09-22].https: // doc.taixueshu.com/foreign/arXiv170802551.html.

Real-time Lane Detection Method Based on Semantic Segmentation

ZHANG Fujie, KE Minyi

(*School of Computer Science, Hubei Univ. of Tech., Wuhan* 430068, *Chine*)

Abstract: In order to improve the detection effect of lane lines in automatic driving, a lane line detection network based on multi-scale semantic features and attention mechanism is proposed. The network is based on the encoder decoder structure, in which the encoder is based on the lightweight network MobileNetV2 to extract the lane line features in real time accurately; To solve the problem that semantic features are not fully utilized in lane detection, multi-scale feature fusion module is added between the encoder and decoder to improve the segmentation accuracy of lane detection. By introducing the attention mechanism of mixed channel domain and spatial domain, combined with global and local multi-scale feature information, the positioning effect of lane lines is strengthened. Through training and testing on Tusimple dataset, the results show that compared with the existing lane line segmentation model, this method is better on segmentation accuracy and speed.

Keywords: automatic driving; semantic segmentation; multi-scale feature fusion; attention mechanism

[责任编校: 赵 瞳]