光电技术及应用

DOI: 10.16818/j.issn1001－5868.20240817001

基于双分支特征提取的车道线检测算法研究

李亚琳，宋小军*

（上海电力大学 电子与信息工程学院，上海 200090）

摘　要： 针对车道线检测算法结构复杂、参数量大等问题，将车道线检测视为语义分割任务，构建了一种基于双分支特征提取的车道线检测模型（DP-RESA）。首先，该模型采用并行设计的语义分支和细节分支进行特征提取。语义分支使用轻量化的MobileNetV2提取高级特征；细节分支具有更多通道数，用于提取包含更多空间细节的低级特征。其次，利用高级特征的权重对低级特征进行筛选这一策略，快速且高效地融合了两路分支的特征信息。在Tusimple数据集上的实验结果表明，相较于基准模型，所提出的DP-RESA算法的准确率提高至96.58%，参数量降至5.12 M，单图推理时间降低11.76 ms，能够满足车道线检测任务在资源有限的嵌入式平台上部署的需求。

关键词： 车道线检测；语义分割；MobileNetV2

中图分类号： TP389　**文章编号：** 1001－5868(2025)03－0522－06

Lane-Detection Algorithm Based on Dual-branch Feature Extraction

LI Yalin, SONG Xiaojun

(**College of Electronic Information Engineering, Shanghai University of Electric Power, Shanghai 200090, CHN**)

Abstract: To address the problems of complex algorithm structure and large number of parameters existing in road lane detection, a lane-detection method based on dual-branch feature extraction (DP-RESA) is proposed. This model treats road-lane detection as a semantic segmentation task. First, lane features were extracted using two parallel branches: a semantic branch and a detail branch. The semantic branch uses a lightweight MobileNetV2 model to extract high-level features, while the detail branch uses wide channels and shallow layers to capture low-level details with more spatial details. Furthermore, the model quickly and efficiently fuses features by using high-level features as weights to filter essential semantic information embedded within the low-level features. Finally, the experimental results for the Tusimple dataset show that compared with the baseline, DP-RESA network lane detection accuracy can reach 96.58%. In addition, the number of model parameters is reduced to 5.12 MB, and the single-image inference time is reduced by 11.76 ms, making the model well suited for lane-detection tasks deployed on resource-constrained embedded platforms.

Key words: lane-detection; semantic segmentation; MobileNetV2

0　引言

随着智能化技术在汽车行业的不断发展，自动驾驶已经成为汽车工程领域的前沿研究方向[1]。车道线作为关键的交通标志，在自动驾驶技术中扮演着至关重要的角色，其检测效果直接影响车道偏离

收稿日期：2024－08－17.

* **通信作者：**宋小军

预警、车道保持以及车道中心线追踪等下游任务的执行效果。高效且实时的车道识别可以提升自动驾驶汽车的行驶安全性。目前主流的车道线检测主要分为传统方法和深度学习方法。

传统方法通常基于边缘检测、滤波等方式，从图像中提取车道线的边缘、纹理和颜色等低级特征信息，然后与霍夫变换、粒子或卡尔曼滤波器相结合，从而实现对车道线的检测。Chiu 等[2]利用颜色信息选择图像中的感兴趣区域，然后通过边缘检测实现对车道线的检测。然而，该算法仅适合直线车道线环境，当车道线弯曲或被遮挡时效果欠佳。Huang 等[3]利用逆透视变换法生成车道线俯视图，并结合卡尔曼滤波器实现车道线检测。该算法运行速度较慢，且仍然不能准确检测复杂环境下的车道线。Sheng 等[4]通过在霍夫变换中引入长度、平行、分布、对角和均匀宽度约束实现车道线的检测，但这种方法计算复杂度高且不能很好处理曲线车道线。

随着深度学习在计算机视觉领域的不断发展，现如今大部分研究人员已将研究重点放在利用深度学习方法检测车道线上。基于深度学习的车道线检测网络可以自主学习不同环境下的车道线特征，具有较强的鲁棒性。目前基于卷积神经网络（Convolutional Neural Network，CNN）的车道线检测方法主要分为两大类：基于检测的方法和基于分割的方法。

在基于检测的方法中，Qin 等[5]将车道检测过程视为基于全局特征的行选择问题并使用结构损失法对车道结构进行建模，具有较高的推理速度，但参数量过大导致耗费过多内存资源，不易于部署在嵌入式平台。Tabelini 等[6]根据车道线细长的特点，将具有细长形状的锚点应用在注意力机制中聚合局部和全局特征，使模型可以学习来自其他车道的信息。该算法实现了精度和速度的平衡，且具有较少的参数量，但不适用于弯道的检测。总体而言，基于检测的方法可以实现快速的运行速度，但准确性欠佳。

在基于分割的方法中，Neven 等[7]通过结合实例分割和曲线拟合实现车道线的检测，但在复杂路况下检测效果不佳。Pan 等[8]提出一种空间卷积网络，通过在水平和垂直方向上传递信息，使模型学习到特征的空间关系，以适应复杂路况下的车道线检测。然而，该算法的逐片传递信息机制导致其实时性效差。为解决该问题，Zheng 等[9]提出利用循环特征移位聚合器（REcurrent Feature-Shift Aggregator，RESA）聚合全局信息，通过并行的方式在垂直和水平方向上循环移动切片特征图传递信息，实现了推理速度的提升。

本文综合了上述算法的优缺点，受 RESA 车道线检测算法的启发，提出一种基于双分支特征提取的车道线检测算法（DP-RESA）。该算法旨在降低网络参数量，提高车道线检测精度，使其易于部署在嵌入式平台。

1　基于双分支特征提取的车道线检测模型

针对现有的车道线检测算法在计算资源有限的车载嵌入式设备中存在实时性差的问题，在像素级语义分割网络 RESA 的基础上，提出一种基于双分支特征提取的车道线检测算法。该算法整体模型由双分支特征提取编码器、循环特征移位聚合器和解码器组成。算法整体网络架构如图 1 所示。

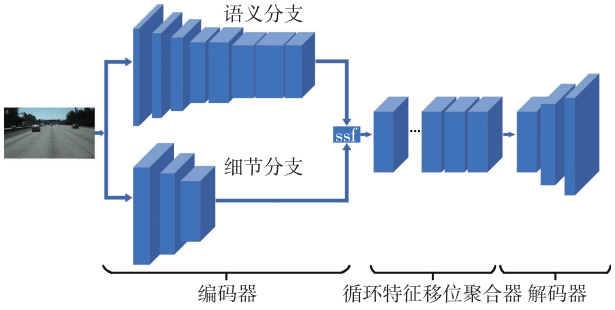


图 1　DP-RESA 结构

1.1　编码器

RESA 算法以 ResNet[10]作为特征提取器。ResNet 通过引入残差学习机制，使网络在深层时依然能够有效地传播梯度，避免梯度消失或爆炸问题，在许多应用中表现出色。但其复杂的模型结构占用了较多内存资源，不适合部署到资源有限的车载嵌入式平台。为此，设计一种轻量化的双分支特征提取网络，将空间信息和语义信息分开处理，以实现实时语义分割的高精度和高效率。该网络包括语义分支、细节分支和特征融合。

1.1.1　语义分支

语义分支旨在捕获高级语义信息，该分支具有更深层次的特征和较少的通道数。为达到轻量化的目的，选用参数量较小的网络 MobileNetV2[11]。

MobileNetV2通过堆叠多个BottleNeck模块实现特征提取，具体结构如表1所示，其中t为扩展因子，c为输出通道数，n为BottleNeck的重复次数，s为步长。BottleNeck为倒残差结构，主要通过使用深度可分离卷积（Depthwise Separable Convolution，DWConv）达到轻量化的目的，其结构如图2所示，两个1×1 Conv通过调整通道数进行升维和降维，3×3 DWConv实现车道线特征的提取。DWConv通过将标准的卷积分为两个步骤以实现轻量化的目的：先进行深度卷积，即对每个通道独立卷积；然后通过1×1卷积（称为逐点卷积）来组合特征图的通道，结构如图3所示。为更直观展示DWConv轻量化的特点，以3×3的卷积核为例，Conv和DWConv的参数量和计算量如表2所示，其中H，W和C分别为高、宽和通道数。由表2可以看出，相比普通卷积，DWConv的参数量和计算量都有大幅度的降低。通过将ResNet替换为MobileNetV2可以有效降低计算复杂度，减少计算资源的消耗，使车道线检测任务更易于部署在车载嵌入式平台上。

表1 MobileNetV2结构

输入	操作方法	t	c	n	s
368×640×3	Conv	—	32	1	2
184×320×32	BottleNeck	1	16	1	1
184×320×16	BottleNeck	6	24	2	2
92×160×24	BottleNeck	6	32	3	2
46×80×32	BottleNeck	6	64	4	2
23×40×64	BottleNeck	6	96	3	1
23×40×96	BottleNeck	6	160	3	2
11×20×160	BottleNeck	6	320	1	1
11×20×320	Conv	—	128	1	1
11×20×128	—	—	—	—	—

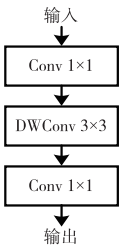


图2 BottleNeck结构

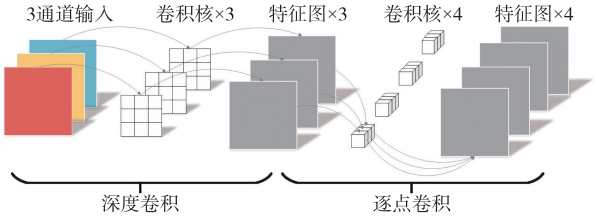


图3 深度可分离卷积流程图

表2 Conv和DWConv参数量和计算量对比

不同卷积方法	参数量	计算量
3×3 Conv	$9C^2$	$9HWC^2$
3×3 DWConv	$9C+C^2$	$9HWC+HWC^2$

1.1.2 细节分支

轻量化的特征提取网络往往会导致特征提取能力不足，难以捕捉复杂数据集中的抽象特征，如车辆遮挡、阴影、车道线掉漆等。因此，与语义分支并行，增加用于捕获低级细节并生成高分辨率特征表示的细节分支[12]。细节分支只关注低级细节，因此堆叠三个卷积层，通过保留更多的通道数提取更多的空间细节，即低级信息。细节分支的具体网络结构如图4所示，主要包含三个阶段，其中每一层都是由卷积层，批量归一化（Batch Normalization，BN）和激活函数（Rectified Linear Unit，Relu）组成。每个阶段第一层的步长s=2，因此，细节分支提取的输出特征图是原始输入的1/8。

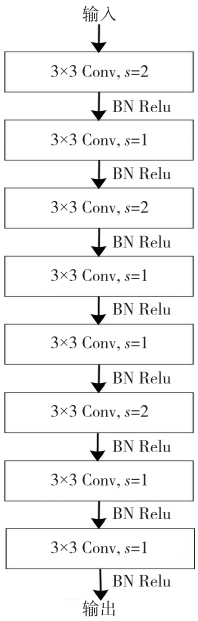


图4 细节分支

1.1.3 特征融合

语义分支具有丰富语义信息但目标定位粗略的高级特征。相反，细节分支包含有限的语义信息但低级特征可以提供精确的目标位置，二者的特征表示是互补的。基于轻量化的要求，简单高效地合并两种类型的特征表示至关重要。由于语义分支输出的特征空间维度小于细节分支，需要对语义分支的输出特征进行上采样，以匹配细节分支的输出。考虑到准确性和效率，使用SFF[13]模块对细节分支和语义分支的特征进行融合。SFF的具体网

络结构如图 5 所示。给定一个语义分支输出的高级特征 $f_{high}\in R^{C\times H\times W}$ 和一个细节分支输出的低级特征 $f_{low}\in R^{C\times H\times W}$，高级特征首先经过步长为 2 的转置卷积（Transposed Convolution，T-Conv）上采样 2 倍，得到 $f_{\widehat{high}}\in R^{C\times 2H\times 2W}$。然后，为统一高级特征和低级特征的维度，使用双线性插值（Bilinear Interpolation，BL）将高级特征上采样到和低级特征相同的维度，即 $f_{att}\in R^{C\times H_1\times W_1}$。最后，通过 CA[14] 注意力机制得到高级特征的权重信息，并利用高级特征的权重信息对低级特征进行筛选，将二者相加得到融合后的特征表示 $f_{out}\in R^{C\times H_1\times W_1}$。使用到的公式：

$$f_{att}=\mathrm{BL}\left(\mathrm{T}-\mathrm{Conv}\left(f_{high}\right)\right) \tag{1}$$

$$f_{out}=f_{low}\times \mathrm{CA}\left(f_{att}\right)+f_{att} \tag{2}$$

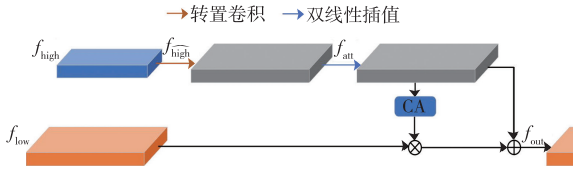


图 5　SFF 的网络结构

1.2　循环特征移位聚合器

车道线由远及近是连续的，且由左向右延伸，线与线之间空间依赖性较强。基于这个特点，循环特征移位聚合器[9]通过水平和垂直移动切片特征图来收集空间信息，如图 6 所示。在每次迭代中，信息传递有四个方向，分别为从下到上、从上到下、从左到右和由右到左，总迭代次数为 4。切片的特征图通过四个方向循环移动传递信息，使具有相同偏移步长的卷积层权重在同一方向所有切片上共享。

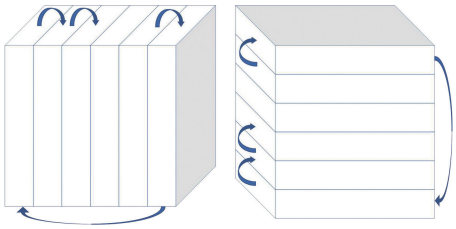


图 6　水平和垂直方向上信息传递

以“从右到左”的信息传递为例，如图 7 所示。当第一次迭代时，步长 $s=1$ 且每列中的 X_i 可以接收 X_{i+1} 移位特征。由于采用循环移位的方式，X_{w-1} 可以接收 X_0 移位特征。当第二次次迭代时，步长 $s=2$ 且每列中的 X_i 可以接收 X_{i+2} 移位的特征。以 X_0 为例，X_0 在第二次迭代中接收 X_2 的信息，由于到 X_0 已经从 X_1 接收到信息，而 X_2 在第一次迭代中已经从 X_3 接收到信息。因此在两次迭代中，X_0 总共从 X_0，X_1，X_2 和 X_3 接收到信息。经过 4 次迭代后，每个位置都能感知并聚合同一特征图中的所有空间信息。

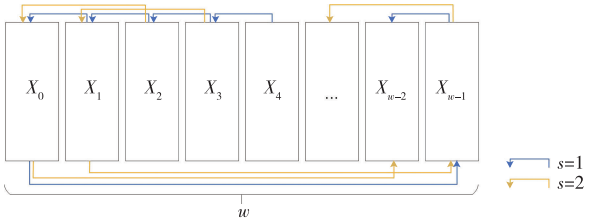


图 7　从右到左特征传递

1.3　解码器

解码器通过堆叠三个双边上采样模块[9]使特征图恢复输入大小。双边上采样模块由两个分支组成，一个分支恢复粗粒度特征，另一个分支修复细粒度丢失，结构如图 8 所示。粗粒度分支使用 1×1 卷积调整通道数后直接使用双线性插值法将特征图上采样 2 倍；细粒度分支通过转置卷积将特征图上采样 2 倍后，堆叠两个 Non-BottleNeck 块[15] 恢复粗粒度分支的信息丢失。Non-BottleNeck 块通过交替使用 1×3 和 3×1 的卷积以保持在特征图大小不变的情况下分别捕捉水平和垂直方向的特征，提升模型对车道线目标的检测和分割能力。

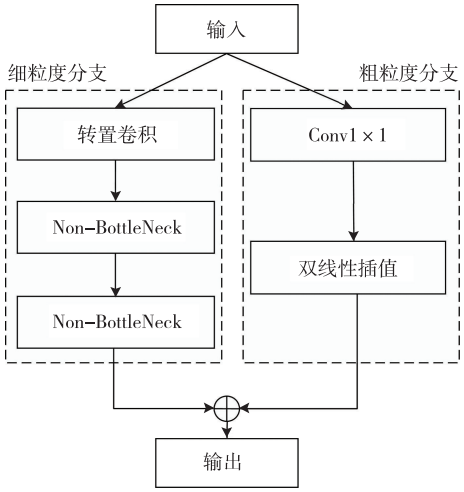


图 8　双边上采样模块

2　实验结果分析

2.1　实验环境

实验环境为 ubuntu18.04，12 vCPU Intel（R）Xeon（R）Platinum 8255C CPU @ 2.50 GHz，GPU 为 11 GB 显存的 NVIDIA GeForce GTX 2080Ti。在算法实现方面，使用 PyTorch 深度学习框架进行

神经网络模型搭建。算法内各超参数设置见表3。

表3 超参数设置

参数	参数值
迭代次数/次	150
批处理大小/次	8
学习率	0.025
优化器	SGD
损失函数	二值交叉熵损失

2.2 数据集和评价标准

使用TuSimple数据集对网络性能进行评估。Tusimple数据集是图森公司在2006年公开的以摄像头为传感器采集的数据集，也是目前车道线检测研究中使用最广泛的数据集。该数据集的采集是通过视频分帧转换的方式形成多张图片，其中已标注的训练集图像有3 626张，已标注的测试集图像有2 782张。

Tusimple数据集将每个车道线用一组固定y坐标间隔的点表示，因此将准确率定义为预测图像中正确的车道线点数在标签图像中车道线点数的占比。其使用的评价指标包括：正确率（Acc）、误检率（FP）和漏检率（FN）。计算公式分别为：

$$Acc=\frac{\sum_{\text{clip}} C_{\text{clip}}}{\sum_{\text{clip}} S_{\text{clip}}}\times 100\% \tag{3}$$

$$FP=T_{\text{fp}}/N_{\text{t}} \tag{4}$$

$$FN=T_{\text{fn}}/N_{\text{T}} \tag{5}$$

式中，C_{clip}是准确预测到的车道线点数，S_{clip}是标签中的车道线点数，T_{fp}是预测错误的车道线点数，N_{t}是所有预测出来的正确车道线点数，T_{fn}是没有正确预测的车道线点数，N_{T}是所有真实的车道线点数。

2.3 消融实验

为验证所提出的双分支特征提取模型的有效性，设计如表4所示的消融实验。

表4 超参数设置

方案	*Acc*/(%)	*FP*/(%)	*FN*/(%)	参数量/M	运行时间/ms
RESA	96.46	3.64	3.39	23.94	37.58
方案1	96.19	4.83	3.51	4.44	24.32
方案2	96.46	5.69	3.26	4.96	24.46
方案3	96.58	3.79	2.84	5.12	25.69

其中RESA为主干使用ResNet34的基础模型；方案1为主干只使用语义分支MobileNetV2；方案2为主干使用语义分支和细节分支且仅使用双线性插值法调整语义分支特征图大小后合并；方案3为本文设计的双分支特征提取的车道线检测算法DP-RESA。

由方案1和基础模型的对比可知，轻量化的特征提取主干MobileNetV2尽管可以带来参数量的显著降低和单图推理速度的提升，但在一定程度上牺牲了精度；方案2在不引入过多参数量的前提下增加了细节分支以弥补语义分支的不足，准确率由96.19%提高至96.46%，与ResNet34-RESA持平；方案3利用高级特征的权重信息对细节分支的低级特征进行筛选这一思想，对语义分支和细节分支进行融合，最终实现准确率提高至96.58%，参数量仅为原模型的21.4%，单图推理时间由37.58 ms降低至11.76 ms。以上实验证明了改进的有效性。

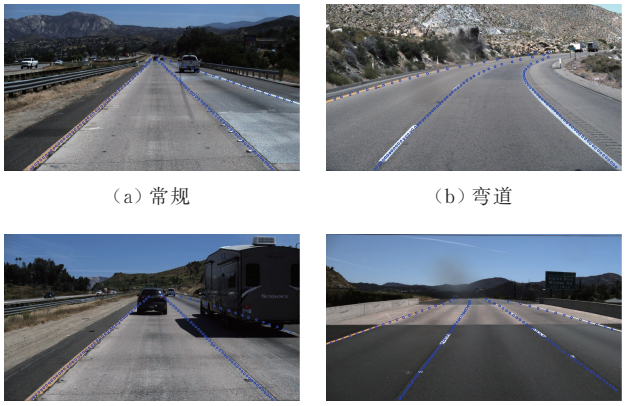
2.4 实验结果与分析

为对DP-RESA的有效性进行验证，选取近年来报道的车道检测算法进行定量分析，对DP-RESA进行横向比较。对比的算法主要分为基于检测和基于分割的算法：前者包括LaneAtt[6]和UFLD[5]；后者包括LaneNet[7]，SCNN[8]和RESA[9]。表5所示为所提方法与其他算法在TuSimple数据集上的对比结果。

表5 对比试验

方法	算法	*Acc*/(%)	*FP*/(%)	*FN*/(%)	参数量/M	运行时间/ms
基于检测的方法	LaneAtt	95.63	3.53	2.92	8.55	5.84
	UFLD	95.86	18.91	3.71	94.55	5.9
基于分割的方法	LaneNet	93.38	7.80	2.24	3.67	19.01
	SCNN	96.53	6.17	1.80	21.94	133.5
	RESA	96.46	3.64	3.39	23.94	37.58
	DP-RESA	96.58	3.79	2.84	5.12	25.69

由表5可知，本文改进的模型DP-RESA在Tusimple数据集上达到了96.58%的准确率，优于其他模型，且保证了在准确率提高的前提下，参数量大幅度降低，使其更易于部署在嵌入式平台。与基于检测的方法相比，虽然推理速度略有逊色，但参数量和精度都具有显著优势。与基于分割的方法相比，DP-RESA的准确率高于其他网络。尽管参数量与单图推理时间略逊色于LaneNet，但准确率远高于LaneNet，展现了一定的竞争优势。除此之外，还进行了可视化实验，各场景下车道线检测效果如图9所示，可以看到所提算法在各种路况下都具有较好的检测效果。



（a）常规　（b）弯道　（c）遮挡　（d）阴影

图9　车道线检测效果

3　结论

车道线检测是智能辅助驾驶系统的重要组成部分，为满足车道线检测需部署在算力有限的车载嵌入式平台的需求，本文在分割模型RESA的基础上，提出一种基于双分支特征提取的车道线检测模型DP-RESA。语义分支使用MobileNetV2，利用深度可分离卷积实现参数量的大幅度降低；细节分支在不引入较多参数量的前提下，通过保留更多通道数以实现空间细节信息的提取；两分支的融合策略上，使用语义分支的权重信息对细节分支提取的特征进行筛选后融合。最后在TuSimple数据集上进行实验，改进后算法的准确率为96.58%，参数量仅为原模型的21.4%，单图推理时间降低11.76 ms，能够满足嵌入式设备实时检测的需求。

参考文献：

[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] Chiu K Y, Lin S F. Lane detection using color-based segmentation [C]// Proceedings of the IEEE Intelligent Vehicles Symposium, 2005: 706-711.

[3] Huang Y, Li Y, Hu X, et al. Lane detection based on inverse perspective transformation and Kalman filter [J]. KSII Transactions on Internet and Information Systems, 2018, 12(2): 643-661.

[4] Luo S, Zhang X, 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] Qin Z, Wang H, Li X. Ultra fast structure-aware deep lane detection[C]// Computer Vision-ECCV 2020: 16th European Conference, 2020: 276-291.

[6] Tabelini L, Berriel R, Paixão T M, et al. Keep your eyes on the lane: Real-time attention-guided lane detection [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 294-302.

[7] Neven D, De Brabandere B, Georgoulis S, et al. Towards end-to-end lane detection: An instance segmentation approach [C]// IEEE Intelligent Vehicles Symposium, 2018: 286-291.

[8] Pan X, Shi J, Luo P, et al. Spatial as deep: Spatial CNN for traffic scene understanding [C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2018: 7276-7283.

[9] Zheng T, Fang H, Zhang Y, et al. RESA: Recurrent feature-shift aggregator for lane detection [C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 3547-3554.

[10] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.

[11] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.

[12] Yu C, Gao C, Wang J, et al. BiSeNet V2: Bilateral network with guided aggregation for real-time semantic segmentation [J]. International Journal of Computer Vision, 2021, 129(11): 3051-3068.

[13] Chen Y, Zhang C, Chen B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. Computers in Biology and Medicine, 2024, 170: 107917.

[14] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 13708-13717.

[15] Romera E, Alvarez J M, Bergasa L M, et al. ERFNet: Efficient residual factorized ConvNet for real-time semantic segmentation [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(1): 263-272.

作者简介：

李亚琳（1998—），女，河南省人，硕士研究生，主要研究方向为自动驾驶；

宋小军（1985—），女，北京市人，博士，副教授，主要研究方向为自动驾驶。