

# 基于事件驱动的车道线识别算法研究

徐频捷<sup>1,2</sup>, 陈逸杰<sup>3</sup>, 李之南<sup>4</sup>, 赵 地<sup>1</sup>

(1. 中国科学院计算技术研究所, 北京 100190; 2. 中国科学院大学, 北京 100049;  
3. 北京邮电大学信息与通信工程学院, 北京 100876; 4. 北京邮电大学国际学院, 北京 100876)

**摘 要:** 动态视觉传感器(Dynamic Vision Sensor, DVS)相比于传统彩色相机有更高的时间分辨率、动态范围,且功耗更低、带宽更低,在自动驾驶领域有很好的应用前景,因此吸引了越来越多研究者的注意.然而由于事件驱动的数据是异步的且缺少一种统一的表示形式,在复杂的交通场景下,以车道检测为代表的基于事件驱动数据的交通场景分割任务难以应用传统的语义分割算法.针对以上问题,本文提出了一种三通道的事件数据编码方式,综合考虑事件数据的时空特征,将其作为卷积神经网络的输入;提出了一种基于编解码模型的事件数据车道检测算法,在基于事件驱动的车道线检测数据集DET上,本文方法的mIoU(mean Intersection over Union)达到了58.76%,比基准方法提高了4.4%.

**关键词:** 事件驱动; 卷积神经网络; 车道线检测; 编解码模型; 语义分割; 动态视觉传感器; 事件表示

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 0372-2112 (2021) 07-1379-07

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20201375

## Research on Event-Driven Lane Recognition Algorithms

XU Pin-jie<sup>1,2</sup>, CHEN Yi-jie<sup>3</sup>, LI Zhi-nan<sup>4</sup>, ZHAO Di<sup>1</sup>

(1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;

4. International School, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** Compared with the traditional color cameras, the dynamic vision sensor, a type of event-based sensor, has higher time resolution, dynamic range, lower power consumption and lower bandwidth requirements. It has good application prospects in the field of automatic driving, which attracts more and more researchers' attention. However, event-driven data is asynchronous and lacks a unified representation. At the same time, in the complex traffic scenario, the traditional semantic segmentation model is difficult to be applied to the event-driven data-based traffic scene segmentation task, for instance, the lane detection task. In view of the above problems, our study proposes a three-channel encoding method for event data, which is successfully used as the input of convolution neural network by considering the spatio-temporal characteristics of event data comprehensively. This paper also proposes a lane segmentation algorithm based on encoding-decoding model, which is superior to the traditional event-based lane line segmentation. On the DET data set, with mIoU(mean Intersection over Union) as the evaluation index, this paper reaches 58.76%, which is 4.4% higher than the benchmark.

**Key words:** event-based; convolution neural network; lane detection; encoder-decoder model; semantic segmentation; dynamic vision sensor; event representation

## 1 引言

车道检测任务对于自动驾驶来说十分重要,在过去的几年里,深度学习技术在该任务上取得较大突破<sup>[1]</sup>.但是这些方法普遍都是为彩色相机拍摄的逐帧

图片专门设计的,一定程度上限制了车道线检测算法在更复杂的城市环境中的表现.为了解决这些问题,动态视觉传感器受到了越来越多的关注.动态视觉传感器是受人类视网膜原理的启发而设计的一种神经拟态

视觉传感器,这种相机通过事件驱动的方式来捕捉场景中的动态变化.与传统相机相比,动态视觉传感器具有很多优势:拥有更高的时间分辨率,可以在微秒级时间捕捉事件;拥有非常高的动态范围,可以在复杂多变的光照环境下稳定的采集信息;拥有更低的功耗,同时对带宽要求更低.

事件驱动的数据表达方式相比于传统的图像有很大的优势,基于事件驱动的光流估计<sup>[2]</sup>、目标检测<sup>[3]</sup>技术已经受到了较多的关注;越来越多的研究者也将目光放在事件驱动的自动驾驶技术上,如DDD17<sup>[4]</sup>等事件驱动的交通场景分割数据集,DET<sup>[5]</sup>事件驱动车道线检测数据集,也有如evsegnet<sup>[6]</sup>等对事件驱动的交通场景做分割的研究.综合来看这些工作主要把目光放在数据集的收集、事件数据的表达以及事件驱动技术与传统图像处理算法的结合上.

在车道线检测任务上,常用的算法可以分为基于图像处理的方法和基于深度学习的方法两大类.基于图像处理的技术,往往将道路图像校正成鸟瞰视图的形式<sup>[7]</sup>,再针对如夜间<sup>[8]</sup>、复杂光照<sup>[9]</sup>等不同的场景,通过滤波、边缘检测等方法预处理图片,再用霍夫变换等方法拟合车道线.但这些方法难以适应复杂的交通场景,容易受到环境噪声的影响.目前被更为广泛使用的是基于卷积神经网络的深度学习.考虑到通用语义分割算法在车道线检测任务上的不足,针对车道线畸变和无限延伸的特点,Neven等使用神经网络自适应学习相机的视角变换<sup>[10]</sup>;Pan等提出的网络结构在设计上考虑了车道线的形状特点<sup>[11]</sup>,达到了目前在自然光图像上车道线检测任务的最高水平,同时该方法也是基于事件驱动的车道线数据集DET<sup>[5]</sup>公开的基准算法之一.但上述深度学习方法都缺少对事件数据的优化,难以适应事件数据的稀疏性和不连续性.文献[12]对于基于机器视觉的车道线检测算法做了详细的综述.

## 2 事件驱动的车道线数据表示

### 2.1 基于事件驱动的车道线数据集

本文使用基于事件驱动的车道线检测数据集DET<sup>[5]</sup>.该数据集将动态视觉传感器安装在汽车上,在武汉市收集了5h的原始事件数据.以时间30ms的时间间隔记录原始事件流,生成分辨率为1280×800的传感器图像,这些图像包括复杂的交通场景和各种类型的车道线.DET共标注了五种不同类别车道线,分别以{0,1,2,3,4}为标号,标号0指背景,其他标号指代不同类型的车道线.该数据集收录了多种车道类型的图像,包括单点划线、单实线、平行虚线、平行实线和虚线等.

### 2.2 基于事件驱动的车道线数据集

#### 2.2.1 事件数据

动态视觉传感器与传统的彩色相机不同,并不采用三个颜色通道编码场景,而是采用捕获每个像素位置上光的强度(即像素的灰度值)的变化编码场景.动态视觉传感器的原始输出不是三维图像(高度、宽度和通道),而是事件流,事件由信号强度对数的变化表示:

$$\log(I_{t+1}) - \log(I_t) \geq \sigma \quad (1)$$

$I_{t+1}$ 和 $I_t$ 是在两个连续的时间戳上感知的事件强度,当强度对数的变化超过阈值 $\sigma$ 时,传感器记录一个事件 $e_i$ ,而每个 $e_i$ 由一个三元组定义:测量点像素坐标 $(x_i, y_i)$ ,以及时间戳 $t_i$ :

$$e_i = \{x_i, y_i, t_i\} \quad (2)$$

#### 2.2.2 事件表示

事件数据具有较大的动态范围,可以在很短的时间间隔存储大量的信息,但是一直难以找到一种有效的编码方式使其作为卷积神经网络的输入.目前常用的事件编码解决方案<sup>[13]</sup>为基于事件位置的密度编码.该方案将事件数据信息编码成如式(3)所示的单通道的图像.该编码方案将通过累加 $W$ 时间间隔内的事件数将事件数据编码成一张事件图像,由式(3)定义:

$$H(x, y) = \sum_{i=1, t_i \in W}^N \delta(x_i, x, y_i, y) \quad (3)$$

式(3)中,根据采集设备的采集频率将时间间隔细分成 $N$ 个时间戳,统计位置 $(x, y)$ 发生的事件总数,用狄拉克函数 $\delta(x_i, x, y_i, y)$ 表示在时间戳 $t_i$ 时位置 $(x_i, y_i)$ 是否存在事件,如果存在则记为1,否则记为0.

上述编码方式只包含了事件在时间间隔 $W$ 内的空间信息,并没有考虑到时间戳之间包括时间信息在内的其他可用于识别任务的信息.为了解决上述问题,本文改进了基于事件位置的密度编码方法,提出了一种融合空间和时间信息的事件数据编码方法,在式(3)的基础上再增加两个编码时间信息的通道,这两个通道可以简单有效存储时间间隔 $W$ 期间内时间戳之间的时间信息.两个通道分别存储每个像素位置上发生的事件的时间戳的均值( $M$ )和标准差( $S$ ),由式(4)和式(5)表示:

$$M(x, y) = \frac{1}{H(x, y)} \sum_{i=1, t_i \in W}^N t_i \delta(x_i, x, y_i, y) \quad (4)$$

$$S(x, y) = \sqrt{\frac{\sum_{i=1, t_i \in W}^N (t_i \delta(x_i, x, y_i, y) - M(x, y))^2}{H(x, y) - 1}} \quad (5)$$

### 3 用于车道线分割的编解码模型

本文提出的算法综合考虑事件的时空信息,将异步的事件数据编码成了三通道的图像格式.接下来将对如何从事件图像中提取车道线进行论述.受到卷积神经网络在图像分割<sup>[14]</sup>等任务上取得较大成功的启发,本文提出了一种基于编解码模型的事件数据车道线分割算法.本文使用Xception网络<sup>[15]</sup>作为编码器,通过组合双线性插值上采样和卷积层构建解码器.编解码模型的整体结构如图1所示.

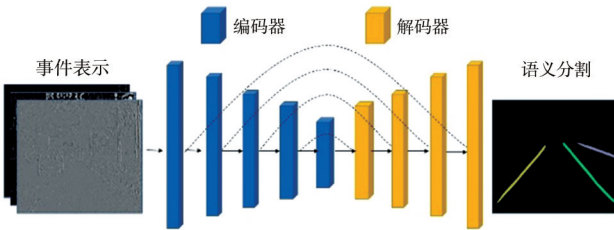


图1 编解码模型的整体结构

**编码器** Xception模型<sup>[15]</sup>在图像分类和分割领域取得了很好的效果,且该模型的参数量较小,可以部署在车载设备上.为了适应事件数据在事件驱动的车道线检测任务的需求,需要对作为编码器的Xception模型做一些修改.由于编码器仅需从事件数据中提取特征,而我们关注的是Xception模型学习到的特征图,因此可以删除原模型最后的两个全连接层和全局平均池化层以缩小模型的规模.从多层次特征提取的角度考虑,可以将Xception模型按层的深度划分成开头、中间、结尾三个部分.不同深度的层可以学习到不同层次的特征信息,开头部分的卷积层可以学习到事件数据在局部时间、空间上的低层次信息,而结尾部分可以学习到综合全局时空特征的抽象语义信息.为了让解码器能充分利用到编码器学习到的不同层次的特征信息,本文使用残差结构连接不同层次的编解码器,将模型如图2灰色部分所示的第2、3、4、13、14个Seption块的输出提取出来,与对称位置的解码块以特征图求和的方式直接向连,如图1连接编解码块的虚线所示.其中,第2、3、4个Seption块来自Xception模型的开始部分,关注浅层特征;第13、14个Seption块来自结尾部分,关注语义特征.

**解码器** 事件数据与传统彩色图像相比,表达的信息更丰富也更抽象,因此本文希望设计的模型可以在解码器部分学习尽量少的参数,而让精心设计的编码器承担尽可能多的学习任务.因此我们在解码器部分仅使用少量的卷积操作,这样既可以利用Xception的特征提取能力,又能保持整个模型轻量化的特点,既可以适应事件数据的特征,又有利于将来模型在车载嵌

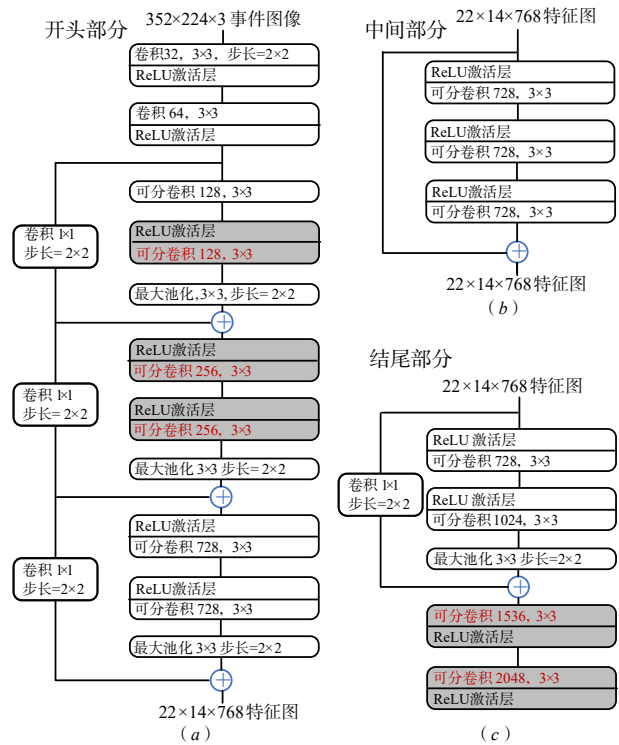


图2 修改后的Xception网络,图中灰色部分连接编码块

入式设备中的部署.如图1所示,为了能融合来自编码器不同深度的特征图,设计了与编码块对应大小的四个解码块.与常用的反卷积增大特征图方法不同,本文直接使用基于双线性插值的上采样,将输入的特征图映射成更大尺度的特征图,从而逐步恢复原始图像的尺度.上采样无需通过反向传播更新参数,而是将参数学习的重心主要安排在精心设计且参数密集的解码器部分,使得模型更加轻量、有效.解码块的结构如图3所示,在每次上采样之后使用少量的可分卷积块学习非线性特征,且可分卷积块中的卷积、可分离卷积都不改变输入特征图的尺度和通道数.在解码模块使用的可分卷积将保持输入特征图的尺度和通道数不变的前提下,通过组合不同大小的卷积核,减少需要学习的参数,从而实现获得更佳的特征学习能力的目的<sup>[16]</sup>.

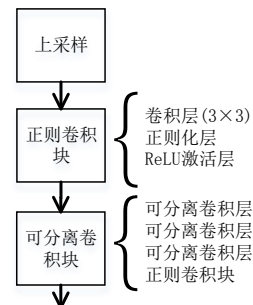


图3 解码块结构

### 4 实验与结果

#### 4.1 训练数据和参数

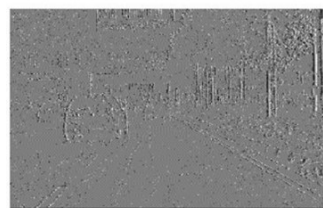
以 30ms 为间隔,按照 2.2.2 节提出的方法编码异步的事件数据,得到包含事件数据时空特性的三通道表示,如图 4 所示. DET 的原始数据集有标注样本 5424 个,将其中 3589 个样本作为训练集训练模型,1835 个样本作为测试集测试算法性能. 为了防止模型在训练集上过拟合,本文对训练集仅需适量的扩充,具体采用了随机裁剪、平移的方法<sup>[17]</sup>. 扩充后的训练集包含 10767 个事件数据样本. 本实验在深度学习框架 tensorflow<sup>[18]</sup> 的基础上,使用 NVIDIA Tesla P100 显卡加速训练,初始学习率为 0.001;模型使用软最大交叉熵(soft-max cross entropy)损失函数;使用 RAdam<sup>[19]</sup>算法优化模型参数,训练 500 个 epoch 后模型收敛,得到与标注结果强相关的模型.

#### 4.2 直观分析

图 5 展示了本模型的部分分割结果. 可以看出本文提出的分割算法在直道分割任务上表现较好,对于有一定距离间隔的虚线也可以直接连续分割出来,传感器感受范围边缘的模糊虚线也可以被准确分割,如图 5 的第一行所示. 当道路前方出现车辆时,本算法也可以将车辆遮挡住的车道线分割出来;对于远离镜头的单实线,传感器采集到的道路数据清晰度不如靠镜



(a) 直方图通道H



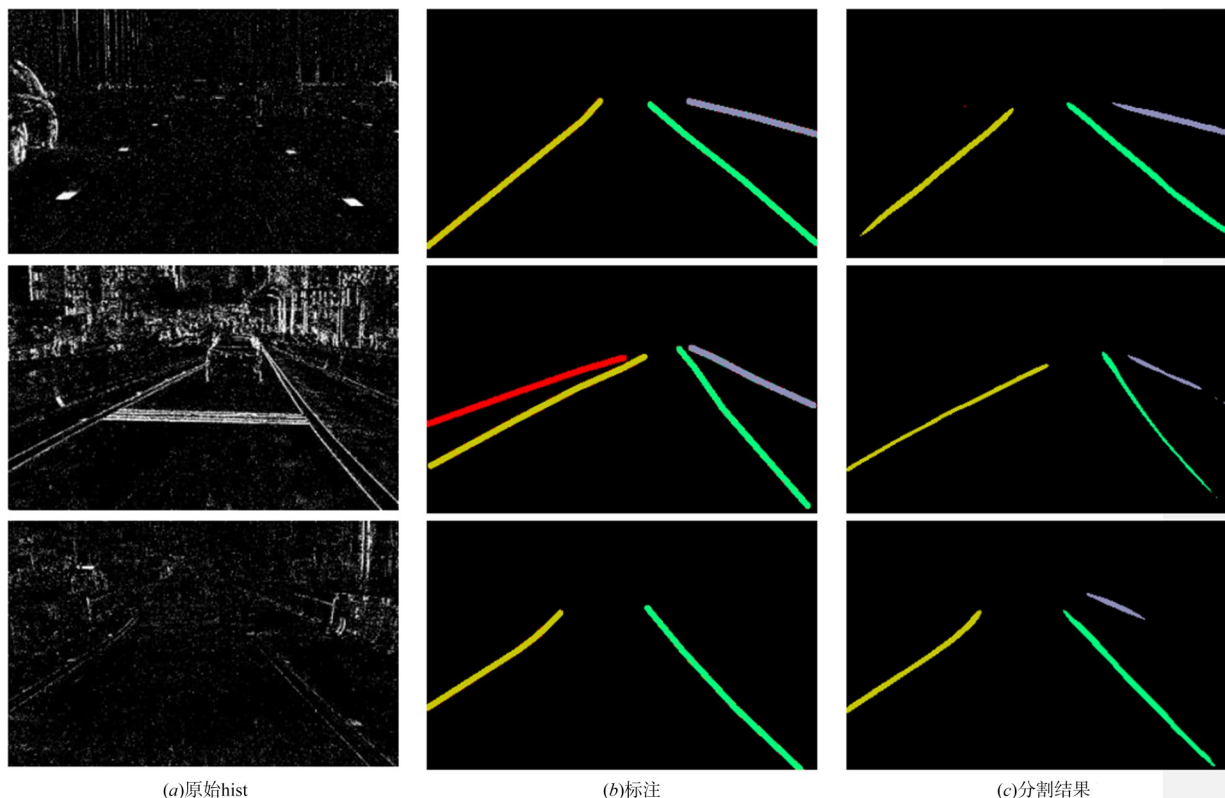
(b) 均值通道M



(c) 方差通道S

图 4 三通道数据编码结果

头的车道线,因此本算法会出现漏检和误检现象,如图 5 第二行所示.



(a) 原始hist

(b) 标注

(c) 分割结果

图 5 道路分割结果展示

### 4.3 比较分析

上述定性分析虽然能直观看出本文提出的事件数据车道检测算法的效果,但是不足以评价本算法的性能.与传统图像分割任务类似,本文选用 mIoU 来衡量算法性能.与当前在车道线分割领域上取得较好效果的方法 FCN (Fully Convolutional Networks)<sup>[14]</sup>、DeepLabv3<sup>[20]</sup>、RefineNet<sup>[21]</sup>、LaneNet<sup>[10]</sup>、SCNN (Spatial CNN)<sup>[11]</sup>进行对比,本方法以 58.76% 的 mIoU 优于这些方法. FCN、RefineNet 和 RefineNet 是通用语义分割算法,并没有针对车道线检测任务做专门的优化,因此性能是较差的, mIoU 均在 50% 或以下;而 LaneNet 和 SCNN 是针对车道线检测任务精心优化过的算法,这些算法采用的网络结构都针对车道线的形状、持续延伸等特点进行了优化,所以可以获得较高的得分.但是事件驱动数据表达的车道线无论是线条的连续性还是和

周边物体的区分度都和传统彩色图像有较大差异,使得这两种算法的性能均不如本方法.本方法在利用传统图像分割算法先进网络结构的基础上,考虑事件数据的时空信息,增强网络学习事件间关系的能力.考虑到事件数据分割的难点在于目标与背景差异不够,本文方法将网络学习的重心放在编码器部分,而采用轻量级的解码器,最终可以发现本方法相比于效果较好的 SCNN 和 LaneNet 有 2.47% 和 4.17% 的提升.

图 6 展示了在一些较难样本上本方法与其他方法的分割效果比较.对于如图 6 第一行所示的有复杂城市背景的弯道场景,本算法检测出了远处视野尽头的车道线弯曲,而其他算法均未检测出此处的弯道变化.在光线昏暗的隧道环境下,彩色图像车道线检测难度较大,但是使用基于事件驱动的数据表示就可以很好的解决这个问题,如图 6 第三行所示,所有的算法都可以较好的检测昏暗隧道里的车道线.

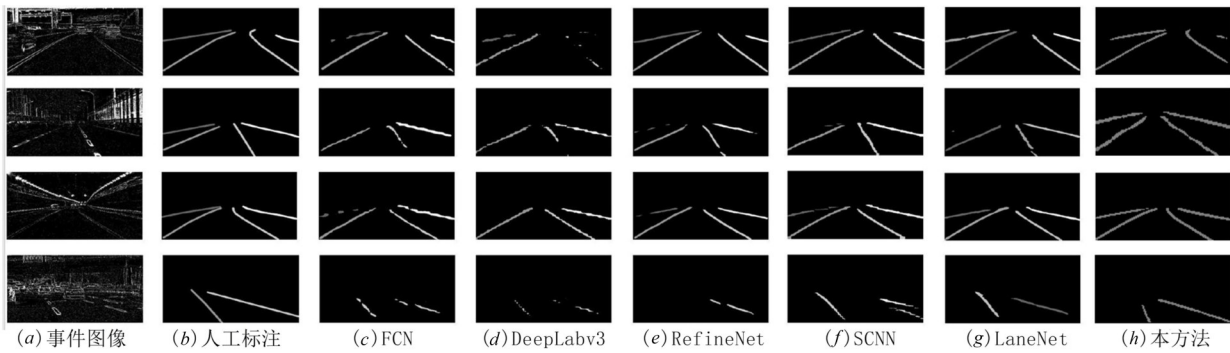


图 6 本方法与其他方法的分割效果比较

表 1 目前主流算法在 DET 数据集上的车道分割表现,以 mIoU(%) 为评价指标

方法	mIoU
FCN	47.36
DeepLabv3	47.30
RefineNet	50.29
LaneNet	53.59
SCNN	56.29
本方法	58.76

### 4.4 数据编码

本文在标准的基于直方图的事件数据编码方法的基础上,提出了融合事件的时间、空间的三通道编码方法.为了验证本文提出的编码方法的有效性,对数据的编码方式做了消融实验,分别测试了相同网络结构下,基于直方图编码(H)、仅使用时间信息编码(M+S)、融合时空信息编码(H+M+S)三种编码方式的性能,如表 2 所示.可以看出,仅基于直方图编码的算法性能低于本

文提出的融合时空信息的编码方式;仅使用时间信息的编码方式结果最不理想,说明仅使用时间信息是不足以表达事件数据包含的主要信息的;但是当时间信息和空间信息被整合成三通道表示的时候,算法性能相比于直方图编码能得到 4.35% 的提高.

表 2 事件数据的编码方式对算法性能的影响

编码方式	mIoU
直方图编码(H)	53.41
时间信息编码(M+S)	21.37
融合时空信息编码(H+M+S)	58.76

## 5 总结

为了解决动态视觉传感器采集到的数据是异步的、缺少一种统一的表示形式的问题,本文提出了一种三通道的事件数据编码方式,综合考虑了事件的时空关系,并为当前主流深度学习算法在事件数据上的应用提供了一种思路.综合考虑到事件数据图像的特点,

本文提出了一种基于编解码模型的事件车道检测算法. 但本文方法仍有不足, 目前本文提出的事件数据编码整合的时间信息还比较简单, 在信息的表达能力上仍有不足. 在卷积神经网络结构的设计上, 虽然本文考虑了事件数据与传统彩色图像的差异, 但网络结构对于车道线本身特征的优化不足. 综合来看, 本文设计的实验结果表明, 本文提出的三通道事件数据编码方式和车道线检测网络是有效的, 且以 mIoU 作为评价指标, 本文方法达到了 58.76%, 优于目前主流的车道分割算法.

#### 参考文献

- [1] HILLEL A B, LERNER R, et al. Recent progress in road and lane detection: a survey[J]. *Machine Vision and Applications*, 2014, 25(3): 727 – 745.
- [2] GALLEGO G, LUND J E A, et al. Event-based, 6-DOF camera tracking from photometric depth maps[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(10): 2402 – 2412.
- [3] CHEN N F Y. Pseudo-labels for supervised learning on dynamic vision sensor data, applied to object detection under ego-motion[A]. *Proceedings 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* [C]. Salt Lake City: IEEE, 2018. 757 – 766.
- [4] BINAS J, NEIL D, et al. DDD17: End-to-end DAVIS Driving Dataset [EB/OL]. <https://arxiv.org/abs/1711.01458>, 2017.
- [5] CHENG W, LUO H, et al. DET: a high-resolution DVS dataset for lane extraction[A]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*[C]. Long Beach: IEEE, 2019. 1666 – 1675.
- [6] ALONSO I, MURILLO A C. EV-SegNet: semantic segmentation for event-based cameras[A]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*[C]. Long Beach: IEEE, 2019. 1624 – 1633.
- [7] 成春阳, 黄渊博, 等. 基于主动红外滤光环视成像的车道线检测算法[J]. *激光与光电子学进展*, 2018, 55(12): 288 – 296.  
CHENG Chun-yang, HUANG Yuan-bo, et al. Lane detection based on active infrared filter and around-view imaging[J]. *Laser & Optoelectronics Progress*, 2018, 55(12): 288 – 296.(in Chinese)
- [8] 王正家, 王若, 等. 夜间车道线检测的研究[J]. *现代电子技术*, 2020, 43(24): 26 – 29.  
WANG Zheng-jia, WANG Ruo, et al. Research on nighttime lane detection [J]. *Modern Electronics Technique*, 2020, 43(24): 26 – 29.(in Chinese)
- [9] 罗强, 王国胤, 等. 复杂光照下的缩微道路车道线检测方法[J]. *计算机科学*, 2014, 41(03): 46 – 49, 70.  
LUO Qiang, WANG Guo-yin, et al. Lane detection in micro-traffic under complex illumination [J]. *Computer Science*, 2014, 41(03): 46 – 49, 70.(in Chinese)
- [10] NEVEN D, DE B B, et al. Towards end-to-end lane detection: an instance segmentation approach [A]. *2018 IEEE Intelligent Vehicles Symposium* [C]. Changshu: IEEE, 2018. 286 – 291.
- [11] PAN X, SHI J, et al. Spatial as deep: spatial CNN for traffic scene understanding [A]. *32nd AAAI Conference on Artificial Intelligence* [C]. New Orleans: AAAI, 2018. 7276 – 7283.
- [12] 吴一全, 刘莉. 基于视觉的车道线检测方法研究进展[J]. *仪器仪表学报*, 2019, 40(12): 92 – 109.  
WU Yi-quan, LIU Li. Research and development of the vision-based lane detection methods[J]. *Chinese Journal of Scientific Instrument*, 2019, 40(12): 92 – 109.(in Chinese)
- [13] NGUYEN A, DO T T, et al. Real-time 6-DOF pose relocalization for event cameras with stacked spatial LSTM networks[A]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* [C]. Long Beach: IEEE, 2019. 1638 – 1645.
- [14] LONG J, SHELLHAMER E, et al. Fully convolutional networks for semantic segmentation[A]. *2015 IEEE Conference on Computer Vision and Pattern Recognition*[C]. Boston: IEEE, 2015. 3431 – 3440.
- [15] CHOLLET F. Xception: deep learning with depthwise separable convolutions [A]. *30th IEEE/CVF Conference on Computer Vision and Pattern Recognition* [C]. Honolulu: IEEE, 2017. 1251 – 1258.
- [16] HOWARD A G, ZHU M, et al. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications[EB/OL]. <https://arxiv.org/abs/1704.04861>, 2017.
- [17] ESWARAN K P, TARJAN R E. Augmentation problems [J]. *SIAM Journal on Computing*, 1976, 5(4): 653 – 665.
- [18] ABADI M, BARHAM P, et al. Tensorflow: a system for large-scale machine learning[A]. *Proceedings of OSDI'16: 12th USENIX Symposium on Operating Systems Design and Implementation*[C]. Savannah: USENIX Association, 2016. 265 – 283.
- [19] LIU L, JIANG H, et al. On the Variance of the Adaptive Learning Rate and Beyond [EB/OL]. <https://arxiv.org/abs/1908.03265>, 2019.

- [20] CHEN L C, ZHU Y, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [A]. Computer Vision-ECCV 2018, PT VII [C]. Munich: Springer, 2018. 801 – 818.

- [21] LIN G, MILAN A, et al. Refinenet: multi-path refine-

ment networks for high-resolution semantic segmentation [A]. 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition [C]. Honolulu: IEEE, 2017. 1925 – 1934.

#### 作者简介



**徐频捷** 男,1997年生于江西上饶.现为中科院计算技术研究所硕士研究生.  
E-mail:xupinjie19g@ict.ac.cn



**李之南** 男,1998年生于新疆乌鲁木齐.现为北京邮电大学本科生.主要研究方向为深度学习加速器,机器学习.  
E-mail:a710859414@bupt.edu.cn



**陈逸杰** 男,1999年生于云南昆明.现为北京邮电大学信息与通信工程学院本科生.  
E-mail:yj-chen@bupt.edu.cn



**赵地(通信作者)** 男,1978年生于湖南岳阳.现为中科院计算技术研究所副研究员,主要研究方向为类脑计算和深度学习.  
E-mail: zhaodi@ncic.ac.cn