

一种基于多维度自注意力机制的轻量级车道线检测算法

崔建东, 崔岩*

五邑大学, 广东 江门

收稿日期: 2021年12月22日; 录用日期: 2022年1月18日; 发布日期: 2022年1月25日

摘要

车道线检测是自动驾驶领域中一个基础的任务。现在最先进的方法主要将车道线检测视为逐像素分割的问题。作为一个密集预测任务, 车道线检测需要较大的计算量, 因此预测速度较为缓慢。但是作为自动驾驶的子任务, 具备实时性是一个重要的要求。所以我们提出了一个极为轻量化的侧导向检测模型, 来实现实时的车道线检测。而且针对卷积网络在特征提取过程中所造成的特征之间的依赖减弱问题, 我们引入了双维度的自注意力机制来解决这个问题。通过在现有的车道线检测基准数据集上进行的大量实验表明, 我们的方法可以在速度和准确性方面达到具备竞争性的性能。

关键词

车道线检测, 轻量级, 实时性, 自注意力机制

A Lightweight Lane Line Detection Algorithm Based on Multi-Dimensional Self-Attention Mechanism

Jiandong Cui, Yan Cui*

Wuyi University, Jiangmen Guangdong

Received: Dec. 22nd, 2021; accepted: Jan. 18th, 2022; published: Jan. 25th, 2022

Abstract

Lane line detection is a basic task in the field of autonomous driving. Current state-of-the-art me-

*通讯作者。

thods mainly treat lane detection as a problem of pixel-by-pixel segmentation. As an intensive prediction task, lane line detection requires a large amount of calculation. So the speed of prediction is relatively slow. But as a subtask of autonomous driving, the real-time is an important requirement. Therefore, we propose an extremely lightweight side guide detection model to achieve real-time lane line detection. And for the problem of reduced dependence between features caused by the convolutional network in the feature extraction process, we introduce a two-dimensional self-attention mechanism to solve this problem. A large number of experiments on existing lane detection benchmark data sets show that our method can achieve competitive performance in terms of speed and accuracy.

Keywords

Lane Line Detection, Lightweight, Real-Time, Self-Attention Mechanism

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

车道线检测是计算机视觉领域一个重要问题。被广泛应用于自动驾驶和路径规划等领域。主流的车道线检测方法有两种,传统的图像处理方法和深度分割方法最近,由于深度学习的发展得到了广泛关注。SCNN [1]通过提出相邻像素之间的消息传递机制显著提高了深度分割方法的性能。但是像素级通信需要更多的计算成本[2]。引入了生成对抗网络,提升了模型在弱光条件下的车道线检测的鲁棒性[3]。提出了一个轻量级的车道线检测网络,通过引入空洞卷积金字塔提升了模型预测的精度。尽管深度分割的车道线检测方法在精度上有了大幅度的提升,但是基于深度分割的车道线检测仍然存在着巨大的挑战。

车道线检测作为自动驾驶的子任务对实时性的要求非常高。尽管现在图形显卡的计算力有了大幅度的提升可以满足大多数视觉任务的运算需求,但是大量的运算同时也会带来很高的能量消耗。在汽车行驶过程中车道线检测算法需要持续进行运算,所以如果单次运算的能量消耗过高这会严重影响车辆的续航能力。因此在确保精度的同时我们还要降低模型的运算量,以达到低耗能实时运算的目的。为了实现低耗能实时运算的需求,我们提出了一个轻量化的车道线检测,并且为了解决卷积神经所带来的特征间依赖降低和特征离散问题,提出了一个融合自注意模块,通过融合空间注意力和通道注意力实现模型更有导向性地进行迭代。

2. 相关工作

随着深度学习的发展,基于深度神经网络的方法在车道检测方面表现出优越性。基于深度学习的车道线方法一般将车道线任务视为一个语义分割任务。例如,VPGNet [4]提出了一个由消失点引导的多任务网络,用于车道和道路标记检测。除了主流的分割公式外,还提出了其他公式,如序列预测和聚类。Fast-Draw [5]预测每个车道点的车道方向,并依次绘制出来。在[6]中,车道检测的问题被视为聚类二进制段[7]。将自注意力机制引入了车道线检测模型一次提升模型性能。但是[7]他只关注了空间上自注意机制。而忽略了通道自注意力机制,因此在本文中我们在模型中引入了空间自注意力机制和通道自注意力机制来关注这两个维度的信息。

3. 算法实现

本论文主要提出了一个轻量化的实时车道线检测模型, 具体网络模型如图 1 所示。

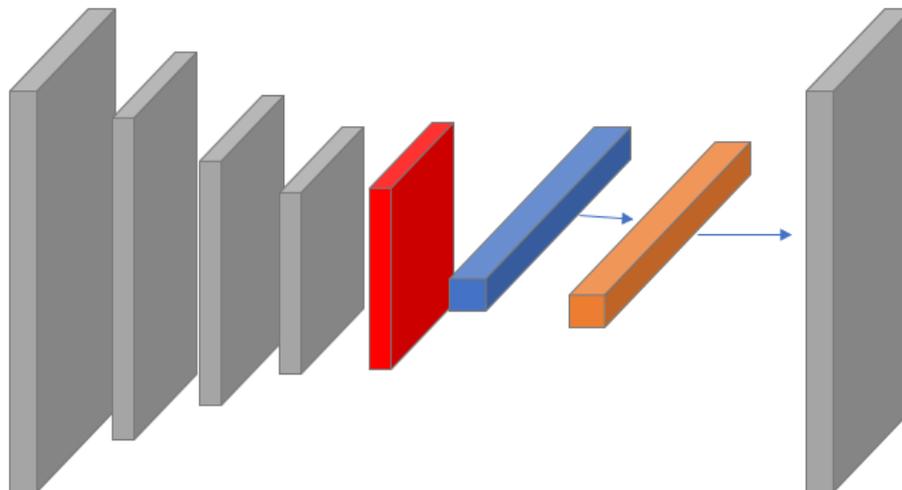


Figure 1. The model framework
图 1. 网络框架图

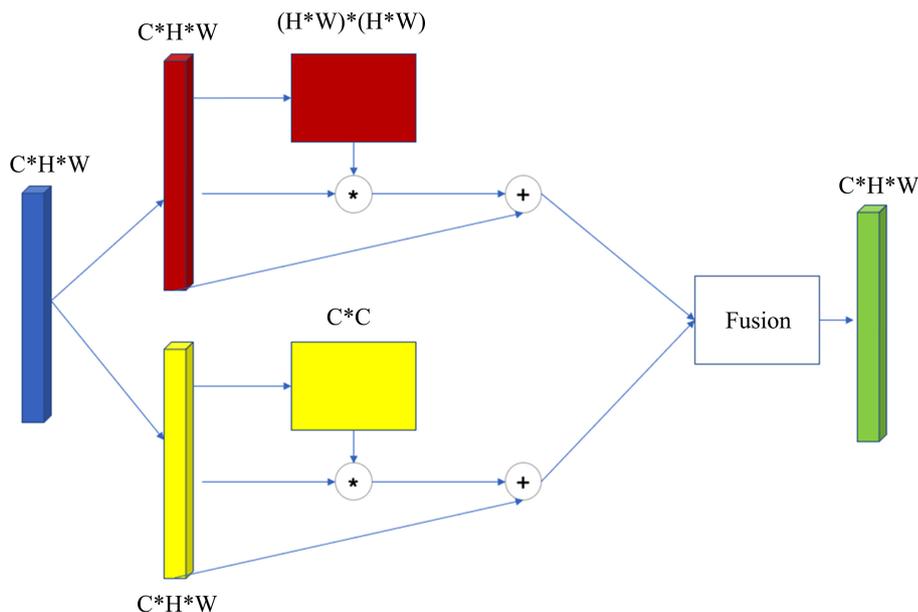


Figure 2. The attention module
图 2. 注意力模块

在本文中我们选用 Resnet34 [8]作为基础网络, Resnet34 有着参数量少、运算速度快等优点。本文使用 Resnet 作为编码器, 在解码阶段会将特征图进行压缩, 并将其拉成一个特征块, 然后分别用两个卷积对特征块进行处理, 最后把处理后的特征块展开得出最终预测结果。

图 1 中的红色模块是我们引入的双头自注意机制模块[9], 其结构如图 2 所示。双头自注意机制主要分成两部分: 1) 空间自注意力模块(图 2 上); 2) 通道自注意模块(图 2 下)。图 2 中, 对于输入的大小

为 $C \times H \times W$ 特征图, 空间注意力模块把空间上的每个点都可看作一个长度为 C 的向量, 并通过 1×1 的卷积对特征图进行压缩得到长度为 $C/8$ 代表 key 的向量。接着用同样的方法得到长度为 $C/8$ 代表 query 的向量。将 key 和 query 相乘可以得到大小为 $(H \times W) \times (H \times W)$ 的相似矩阵。相似矩阵可以理解成空间上每个点与其他点之间所包含信息的相似度。对于空间上每个点得到的大小为 $H \times W$ 的相似向量我们通过 softmax 使得其总和大小为 1, 处理后的 $(H \times W) \times (H \times W)$ 的矩阵可以看作通过注意力机制得到的权重。将这个权重与提取的 value 特征图相乘, 再和原来的特征图相加得到最终的结果。对大小为 $C \times H \times W$ 输入特征图, 通道自注意力模块直接通过重塑输入特征图得到 key, query 向量相乘后得到通道与通道之间的相似度。由于每个通道和自己的相似度是最大的, 所以每一行要减去这个最大的相似值。然后使用 softmax 对每一行进行处理, 使其总和为 1, 得到注意力权重。将这个权重与原特征图相乘, 再整体乘上一个可学习的系数 γ 即可得到最终输出。最后我们将两个从自注意力模块输出的特征图进行融合得到最终的融合特征图。为了使模型更加轻量化, 在解码阶段我们对特征图进行压缩, 并将其拉成一个大小为 $1 \times 1 \times 1800$ 的特征块, 然后分别用两个 1×1 的卷积对特征块进行处理。最后把处理过的特征块展开得到最终预测结果。在损失函数方面我们遵循[10]提出的分类损失函数。

4. 实验

在这一章节我们将详细介绍我们的实验细节。首先我们将介绍相关数据集, 而后我们将介绍实验环境和相关超参数的设置。最后我们将展示我们的实验效果还有我们和一些先进方法的方法的比较。

4.1. 数据集

CULane 是由香港中文大学提出的用于行车道检测学术研究的大规模数据集。它由安装在六辆车上的摄像机收集。收集了超过 55 小时的视频, 并提取了 133,235 帧。其中 88,880 帧被分训练集, 9675 帧被分为验证集, 34,680 帧被分为测试集。

4.2. 实验室环境和超参设置

使用 Adam [11]来训练我们的模型。batch size 设置为 32, 数据集的训练 epoch 总数设置为 50, 学习率为 0.0004。所有模型均使用 pytorch [12]和 Tesla V100 进行训练和测试。

4.3. 评价指标

为了对算法的检测精度进行评价, 本文将每个车道线视作一条宽 30 像素的线, 通过计算真实标签与预测结果之间的交并比(IoU)。在本中我们将阈值设置为 0.5, IoU 大于阈值的预测结果被视作真阳性(TP)。最后我们用 $F1$ 分数来对模型性能进行评价, $F1$ 分数定义为:

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (1)$$

$F1$ 分数($F1$ Score)为统计学中用来权衡二分类模型精确度的指标。它同时顾及了分类模型的精确率和召回率。 $F1$ 分数可以看作是模型精确率和召回率的一种调和平均, $F1$ 分数($F1$ Score)又叫做平衡 F 分数(balanced F Score), 他被定义为精确率(precision)和召回率(Recall)的调和平均数。用精确率和召回率的指标来评价二分类模型的分析结果。

4.4. 实验效果展示和对比

在这部分我们对本文的算法和其他先进算法进行对比以证明本文算法的有效性, 对比结果如表 1 所示。

Table 1. Performance comparison results of the algorithm in this paper and other algorithms

表 1. 本文算法与其他算法的性能对比结果

环境	[1]	[2]	[13]	本文
正常	90.6	91.8	90.1	90.2
拥挤	69.7	71.7	68.8	69.2
夜晚	66.1	69.3	66.0	66.7
无车道线	43.4	45.7	41.6	43.3
阴影	66.9	76.5	65.9	68.2
眩光	58.5	66.5	60.2	65.5
十字路	2505	2346	1998	2015
综合	71.6	73.8	70.8	71.3
FPS	7.5	-	74.6	124.5

从表 1 可知我们的算法在各种环境下进行车道线检测都能取得优秀的效果。并且在相对于其他[1] [2] [13]的算法, 我们的算法在运行效率上取得了最佳的效果。这是由于我们的框架非常的轻量化的网络框架保证了模型的运行效率, 同时我们又引入多维度的自注意机制来提升模型的预测精度。由于自注意力机制会产生较大的运算量, 但由于我们在将特征输入自注意力模块之前就先把特征通道进行了压缩, 从而有效地缓解自注意机制的高计算量的问题。为了更有效地证明我们算法的有效性, 我们对预测效果进行了可视化, 可视化结果如图 3 所示。由可视化结果可知我们的模型的真实预测效果十分好。



Figure 3. Visualize lane line detection results

图 3. 可视化车道线检测结果

5. 结论

本文提出了一种轻量化的实时车道线检测算法。为了提高模型的速度,我们在模型中没有采用任何多尺度特征融合模块,而是使用了极其轻量的模型结构。同时提出了一个多维度的自注意力模块来提升模型预测精度。由于我们在将特征输入自注意力模块前先对特征进行了压缩,因此我们的自注意力模块并不会引入过多的计算量。实验证明,我们的算法很好地实现了实时性,并取得了优秀的检测效果。

参考文献

- [1] Pan, X., Shi, J., Luo, P., Wang, X. and Tang, X. (2018) Spatial as Deep: Spatial CNN for Traffic Scene Understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, 7276-7283.
- [2] 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]. 小型微型计算机系统: 1-7. <http://kns.cnki.net/kcms/detail/21.1106.tp.20211019.1525.004.html>, 2021-11-04.
- [3] 陈正斌, 叶东毅. 带语义分割的轻量化车道线检测算法[J]. 小型微型计算机系统, 2021, 42(9): 1877-1883.
- [4] Lee, S., Kim, J., Shin Yoon, J., Shin, S., Bailo, O., Kim, N., Lee, T.H., Seok Hong, H., Han, S.H. and So Kweon, I. (2017) Vpnet: Vanishing Point Guided Network for Lane and Road Marking Detection and Recognition. *Proceedings of the IEEE International Conference on Computer Vision*, Venice, 22-29 October 2017, 1965-1973. <https://doi.org/10.1109/ICCV.2017.215>
- [5] Phillion, J. (2019) FastDraw: Addressing the Long Tail of Lane Detection by Adapting a Sequential Prediction Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, 15-20 June 2019, 11582-11591. <https://doi.org/10.1109/CVPR.2019.01185>
- [6] Hsu, Y.C., Xu, Z., Kira, Z. and Huang, J. (2018) Learning to Cluster for Proposal-Free Instance Segmentation. *Proceedings of the International Joint Conference on Neural Networks*, Rio de Janeiro, 8-13 July 2018, 1-8. <https://doi.org/10.1109/IJCNN.2018.8489379>
- [7] 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(9): 24-27.
- [8] He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [9] Fu, J., Liu, J., Tian, H.J., Li, Y., Bao, Y.J., Fang, Z.W. and Lu, H.Q. (2019) Dual Attention Network for Scene Segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, 15-20 June 2019, 3146-3154. <https://doi.org/10.1109/CVPR.2019.00326>
- [10] Qin, Z.Q., Wang, H.Y. and Li, X. (2020) Ultra Fast Structure-Aware Deep Lane Detection. *Computer Vision-ECCV 2020: 16th European Conference*, Glasgow, 23-28 August 2020, 276-291. https://doi.org/10.1007/978-3-030-58586-0_17
- [11] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980
- [12] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., et al. (2019) Pytorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, **32**, 8026-8037.
- [13] Hou, Y., Ma, Z., Liu, C. and Loy, C.C. (2019) Learning Lightweight Lane Detection CNNs by Self Attention Distillation. *Proceedings of the IEEE International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 1013-1021. <https://doi.org/10.1109/ICCV.2019.00110>