

复杂城市交通场景下的自动驾驶语义分割方法

周 勇 刘泓滨 侯亚东

(昆明理工大学机电工程学院 昆明 650500)

摘 要:多尺度特征金字塔可以缓解语义分割在复杂交通场景下通常存在漏分割、错误分割、边界分割不清晰等问题,但现有的多尺度特征金字塔在获取丰富的语义信息时,不得不下采样特征图,牺牲空间细节信息,而这导致了最终的分割结果仍然精度受限。针对该问题,本文提出了特征强化模块,使得有利于正确分类像素的相似特征在下采样过程之前得到基于不同矢量间余弦相似度的进一步加强,降低下采样带来的负面影响。另外,结合空洞卷积和条带卷积原理,本文对大卷积核进行了改造,并构建新的多尺度特征金字塔模块,以获取尺度不同且具备更大感受野的语义信息。该分割方法实时高效,能够满足自动驾驶语义分割要求,在 VOC2012 数据集上的实验表明,所提出的方法分割结果 mIoU 达到了 74.36%,FPS 达到了 43,优于目前的主流语义分割方法。

关键词: 语义分割;自动驾驶;特征强化模块;多尺度特征金字塔;相似特征

中图分类号: TP391.41;U121 **文献标识码:** A **国家标准学科分类代码:** 510.40

Automatic driving semantic segmentation method for complex urban traffic scene

Zhou Yong Liu Hongbin Hou Yadong

(Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: The multi-scale feature pyramid can alleviate the problems of semantic segmentation in complex traffic scenes, such as missing segmentation, wrong segmentation and unclear boundary segmentation. However, the existing multi-scale feature pyramid has to downsample the feature maps and sacrifice the spatial detail information for rich semantic information, leading to the limited accuracy of the final segmentation result. Aiming at this problem, a feature enhancement module is proposed to further reinforce similar features based on cosine similarity between different vectors before downsampling, alleviating the negative influence of downsampling. In addition, combined with the principle of dilated convolution and strip convolution, the large convolution kernel is modified to build a new multi-scale feature pyramid module for semantic information with different scales and larger receptive fields. The proposed segmentation method is real-time and efficient, and can meet the requirements of automatic driving. Experiments on the VOC2012 dataset show that the mIoU of the proposed method reaches 74.36%, and the FPS reaches 43, which is superior than the current prevailing semantic segmentation methods.

Keywords: semantic segmentation; automatic driving; feature enhancement module; multi-scale feature pyramid; similar features

0 引 言

自动驾驶能够节约学习驾驶技术的时间和成本,缓解驾驶压力,规范城市交通,大大减少交通事故发生的概率。

近年来,自动驾驶技术发展迅猛,但主要依靠激光雷达、超声波传感器等获取较为精确的环境信息。然而,这

种方式的成本非常高。因此通过语义分割技术实现车辆的环境感知更有利于推进自动驾驶工业化进程。

语义分割技术的发展对自动驾驶发展至关重要。通过对场景的解析,自动驾驶车辆可以更加准确地响应路况变化,做出正确的行驶决策。然而,自动驾驶语义分割技术仍然存在许多挑战。而如何实现高效准确的语义分割,一直是分割领域的难点,需要不断探索和创新。深入

研究和发自动驾驶语义分割技术,推动技术的不断进步和优化,是实现自动驾驶技术普及和提高行驶安全性和效率的重要举措。

1 相关工作

语义分割是自动驾驶车辆感知环境的技术手段之一,在应用于复杂多变的城市交通场景时,通常存在以下精度类型问题:占据像素较少的目标难以被分割、目标边缘轮廓分割不准确、属于目标整体的一部分被错误分类为其他类别等。

对于解决错误分类问题而言,丰富的语义信息是非常重要的,一些经典的工作如: ResNet^[1]、Xception^[2]、VGG16^[3]等以及近年的优秀工作^[4-7]可以通过不断加深的网络逐步提取到高级的语义信息,在一些主流的语义分割工作中还提出或引入了多尺度金字塔模块^[8-10],将不同具备不同尺度感受野的语义信息相结合,得到对分类判断有利的特征表达。这些工作对于解决这种目标部分分割错误的问题非常有效,但它们都依赖下采样来获取高级语义信息,牺牲了空间细节信息。

获取空间细节信息与高级语义信息总是矛盾的,但对于解决边缘轮廓分割不精确问题而言,图像的空间细节信息也同样重要。为了解决该问题,一些双分支网络工作^[11-12]被提出,通过新的分支单独提取加工空间细节信息,并设计特征融合模块来融合高级语义信息和空间细节信息,这对于分割边缘轮廓是很有效的,但同时也带来了更大的计算量。除此以外,以大卷积核的方式提取语义信息来保留更多的空间信息这类型的工作^[13]同样效果显著,但仍然有着相同的问题。

解决分割精度类型问题往往需要消耗更多的计算资源和网络训练时间,而模型的规模大小和分割速度同样是需要考虑的,为此,许多工作提出或引入了注意力机制模块^[14-19]来帮助网络克服卷积需要消耗更多资源来建立长距离依赖的缺陷。一些轻量化的网络例如 MobileNet 系列^[20-21]同样也有助于模型规模的缩减。

目前自动驾驶领域对基于深度学习的主流语义分割所做的改进工作,大多从四个方面入手:添加特征融合、引入注意力机制、改良或引入多尺度金字塔模块,改良骨干网络。

为了兼顾分割速度和精度,孟俊熙等^[22]基于 deeplabv3+模型设计了异感受野拼接 ASPP 模块,以提取更丰富的语义信息,缓解错误分割、漏分割等问题,并引入深度可分离空洞卷积,减小模型的参数量。石敏等^[23]针对常用骨干网络无法兼顾参数计算量与特征表达能力的问题,设计了新的通道注意力瓶颈模块,构建骨干网络用于分割,达到了较好的分割效果。樊博等^[24]结合混合注意力模块对 BiSeNet V2 模型进行改进,优化深层特征图的权重,以增强网络对关键特征的提取,解决边缘分割和小目标分割效果差的问题。

尽管这些工作已取得了很好的成果,但没有对下采样过程丢失相似空间信息的问题进行改进,这对于分割同类型但占据像素更少的小目标、分割目标边缘轮廓都是不利的。

空间信息丢失是限制目前主流分割网络性能上限的重要因素,而参数计算量将影响分割时效性。因此本文提出了一种新的基于编码器解码器结构的分割网络,选择提取特征能力良好,且非常轻量化的骨干网络 MobileNetV2 作为编码器,提出特征强化模块 FEM (feature enhancement module),以减少相似特征丢失的概率,引入条带卷积及空洞卷积原理对大卷积核改良,构建多尺度金字塔模块 ASPSP (atrous spatial pyramid strip pooling),提取丰富的语义信息。在兼顾分割速度的前提下,网络实现了很好的分割效果。

2 网络模型设计

2.1 整体结构

本文提出的分割模型整体结构如图 1 所示,先利用 MobileNetV2 作为骨干网络提取具备更多空间信息的浅层特征以及需要被进一步加强的深层特征。

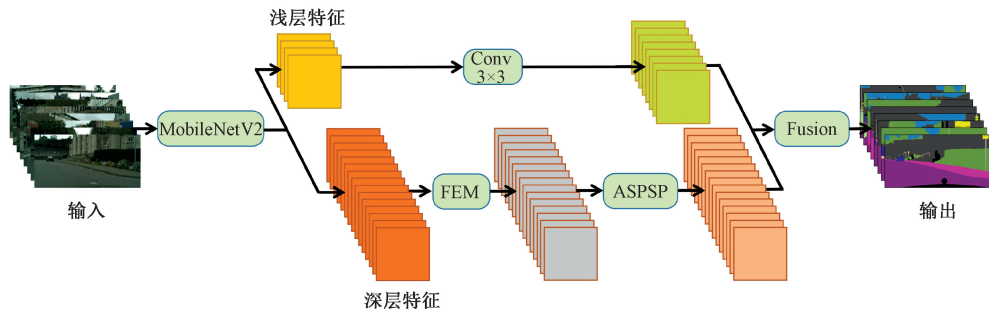


图 1 整体结构

Fig. 1 Overall structure

在提取更丰富的语义信息之前,深层特征先经过 FEM 模块进行加强,而后再经过 ASPSP 模块提取多尺度语义信息,将加强后的深层特征与经过调整通道数后的浅层特征相结合,经过上采样后解码完成分割。

其中输入图像尺寸为 512×512 ,先经过一个 3×3 的卷积下采样后,变成 $32 \times 256 \times 256$ 的张量,经过 MobileNetV2 下采样一次后,得到 $24 \times 128 \times 128$ 的浅层特征,下采样两次后,得到 $320 \times 64 \times 64$ 的深层特征。

浅层特征经过卷积进行升维后,通道数翻倍为 48,深层特征经过 FEM 模块和 ASPSP 模块后,通道数和分辨率仍旧保持不变。

2.2 FEM 模块

余弦相似度常被用于自然语言处理中估计文本之间的相似度,而在图像语义分割中,从属于同类的像素之间也具备相似度。

本文提出的 FEM 模块基于此概念,以张量中每个通道维度的元素组成空间矢量,组成方式如图 2 中红色方块所示,并以空间矢量夹角的余弦值来评估特征图中空间矢量之间的相似度,得到的相似度矩阵将用于放大对应的空间元素,使得相似特征被加强。

其具体原理如图 2 所示。假定输入的深层特征为 $4 \times 3 \times 3$,那么将该张量分别拉伸并拼接成 9×4 的矩阵以及 4×9 的矩阵,其中 9×4 矩阵的行向量与 4×9 矩阵的列向量代表不同像素点的空间矢量。计算矢量余弦相似度,得到输入张量的相似度矩阵图,再左乘 9×4 矩阵并重组,得到每个通道的相似度张量。

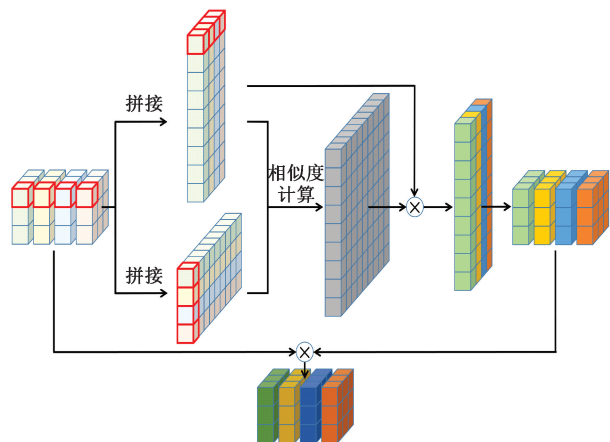


图 2 FEM 原理图

Fig. 2 Schematic diagram of FEM

相似度张量与输入的深层特征对应元素相乘,实现对深层特征基于相似度的特征强化。其中,矢量余弦相似度的计算可表示为:

$$f(x_i, x_j) = \frac{x_i \cdot x_j}{|x_i| |x_j|} \cdot g(x_i, x_j) \quad (1)$$

式中: x 为输入张量, i 与 j 表示空间矢量的编号,最大为 9; $g(x_i, x_j)$ 表示对输入张量的拉伸拼接。那么加强后的深层特征可表示为:

$$F(x_i, x_j) = g^{-1}(f(x_i, x_j)) * x \quad (2)$$

式中: $g^{-1}(x_i, x_j)$ 表示对矩阵重组为高维张量,可看作将张量拉伸拼接的逆操作。

2.3 ASPSP 模块

大卷积核对于语义分割提取信息非常有利,但因过多的参数,很少被用于多尺度信息提取。因此本文结合空洞卷积、条带卷积来降低大卷积核的参数,将其用于提取多尺度语义信息,以实际尺寸为 4×4 大小的卷积核为例,设步长为 1,膨胀系数为 2,则捕获到的感受野可等效为 7×7 的大卷积核,原理如图 3 所示。

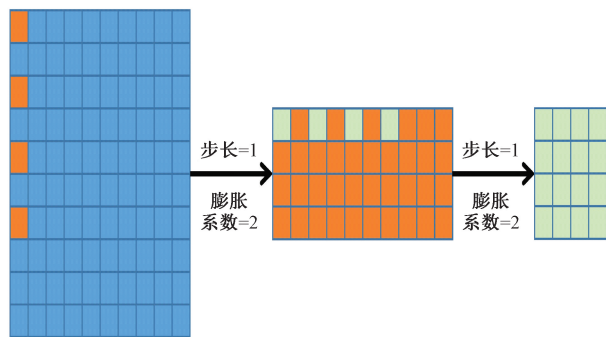


图 3 条带空洞卷积原理图

Fig. 3 Schematic diagram of strip dilated convolution

以条带空洞卷积为基础,构建出的 ASPSP 模块结构如图 4 所示。

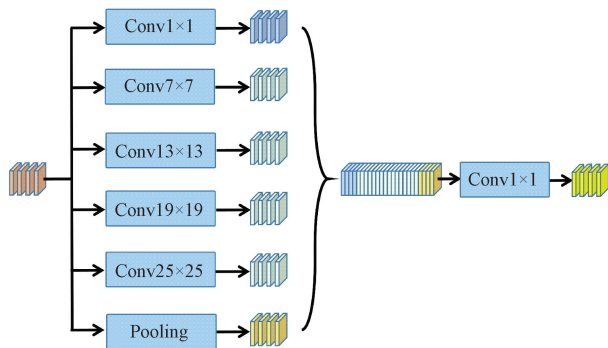


图 4 条带空洞卷积原理图

Fig. 4 Schematic diagram of strip dilated convolution

将加强后的深层特征进行系数缩放、全局池化,同时用不同尺寸的条带空洞卷积提取特征,再将所得的三种特征堆叠,使用 1×1 卷积调整通道,融合特征。通过 ASPSP 模块处理后,所得的深层特征即为加强后的、具备更丰富多尺度语义信息的深层特征。

条带空洞卷积提取特征的过程中,虽然卷积核尺寸

不同,但膨胀系数仍然都为 2,通过 padding 操作使得不同尺寸的卷积输出的特征图尺寸都保持一致。需要说明的是,尽管 ASPSP 模块并不损失通道数和分辨率,这并不代表特征图进入 ASPSP 模块后不会产生空间信息丢失,因为 padding 操作弥补的信息都为 0。所以在 ASPSP 模块之前使用 FEM 模块仍然是必要的。

3 实验验证

3.1 实验环境

本文实验中所需的操作系统、硬件配置、开发环境如表 1 所示。

表 1 配置表
Table 1 Configuration table

配置	参数
处理器	Intel Core i5-11400 2.60 GHz
显卡	NVIDIA GeForce RTX 3080 12 GB
内存	16 GB
Pytorch	1.10.0
Python	3.7.0
CUDA	11.3
CUDNN	8.2.0
操作系统	Windows 10

3.2 数据集介绍

本文使用的数据集为 VOC2012 增强数据集,该数据集在 VOC2012 数据集的原版基础上,结合 SBD 数据集进行了扩充,一共包含了包括背景在内的 21 个类别,其中有行人、汽车、自行车、摩托车等城市交通场景常见事物,分为 10 582 张训练集、1 449 张验证集以及 1 446 张测试集。

为了格式统一,训练前将分割标签全部转化为 8 bit 的 PNG 灰度图,部分图例如图 5 所示。



图 5 VOC2012 增强数据集(左:原图;右:标签)

Fig. 5 VOC2012aug dataset (the left means original photos and the right means labels)

3.3 实验评价指标

本文采用平均交并比 mIoU 作为判断分割准确度的指标,采用每秒传输帧数 FPS 作为度量分割速度的指标。mIoU 的计算公式可表示为:

$$mIoU = \frac{1}{K} \sum_{i=0}^K \frac{T_{TP}}{F_{FN} + F_{FP} - T_{TP}} \tag{3}$$

式中: T_{TP} 表示真正例, F_{FP} 表示假正例, F_{FN} 表示假负例, K 表示为类别数。

FPS 的计算从图像上传至 GPU 开始,计算公式可表示为:

$$FPS = \frac{n}{\sum_{i=0}^n t_i} \tag{4}$$

式中: n 表示预测图像张数, t_i 表示预测第 i 张所用时间。

3.4 实验结果

本文实验参数设置如表 2 所示,由于获取骨干网络的预训练权重参数便利,因此将网络训练过程分为冻结阶段和解冻阶段,冻结阶段加载骨干网络的预训练权重文件,此时的 Batch size 为 16,到第 51 轮迭代时,网络解冻,开始训练整个网络的参数,此时的 Batch size 为 8。

表 2 实验参数
Table 2 Experimental parameters

参数	数值
输入尺寸	512×512
优化器	ADAM
Batch size	16/8
迭代轮数	100
初始学习率	0.000 08
动量参数	0.9

训练完成后,模型分割效果如图 6 所示。该结果为模型分割图与原图的混合,从图 6 中可以看出,本文所提出的网络模型对自动驾驶场景常见的行人、车辆等具有较好的分割效果。

为进一步说明本文网络设计的有效性,本文以同样的迭代轮数对目前的主流分割网络复现,进行对比实验,定量的对比结果如表 3 所示。

表 3 对比实验结果
Table 3 Contrast experiment results

模型	Backbone	mIoU/%	FPS
DANet	ResNet50	70.1	26.72
GCNet	ResNet50	71.46	28.58
ISANet	ResNet50	71.82	32.73
PSPNet	ResNet50	71.57	30.19
DeepLabV3+	ResNet50	70.82	27.14
本文	MobileNetV2	74.36	43.11

可以看出相对于主流的分割模型,本文提出的分割方法在 mIoU 指标上最大提升了 4.26%,在 FPS 指标上最大提升了 16.39。与表现最好的 ISANet 相比,本文提出的网络也分别有 2.54% 的 mIoU 提升和 10.38 FPS 的提升,这说明本文相对于主流的分割网络,有效地提高了

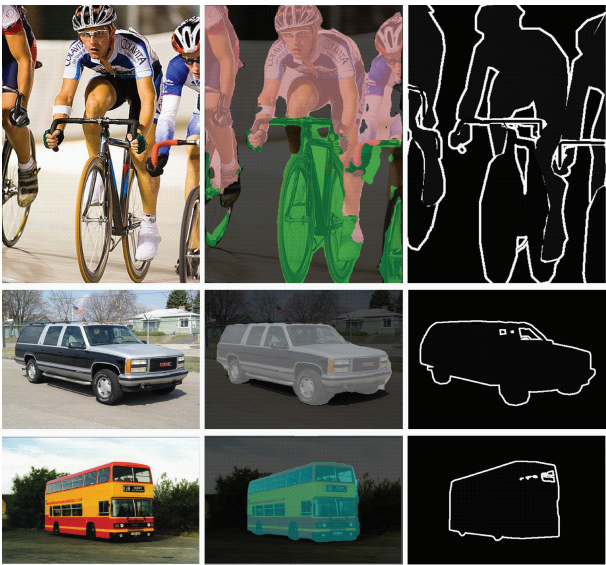


图 6 分割效果展示

Fig. 6 Display of segmentation results

分割精度和分割速度。

以典型常见的城市交通自动驾驶场景为例,定性的对比分割结果如图 7 所示,从分割结果对比可以看到,相对于主流分割网络,本文提出的分割方法在解决小目标分割、目标边缘轮廓分割不准确属于目标整体的一部分被错误分类等常见分割问题的效果更好。

如图 7 所示,FEM 模块对相似特征的加强会使得对属于同一类事物的像素分类结果更不容易出错,因此在对人与摩托车的分割中,容易与摩托车混淆的腿部像素有更高的几率被归类为人的类别。而在对占据像素较少的轿车分割时,尽管目标很小,但网络仍然成功将几辆小轿车都分割了出来。

ASPSP 模块中的卷积核对输入特征图能够获取等效于大卷积核的感受野,提取的上下文信息丰富,这会使得不属于一类的像素,也更不容易混淆。例如在对小型摩托车分割中,部分网络将车身一部分判别为自行车,一部分判别为摩托车。在对自行车与轿车尾部的分割中,部

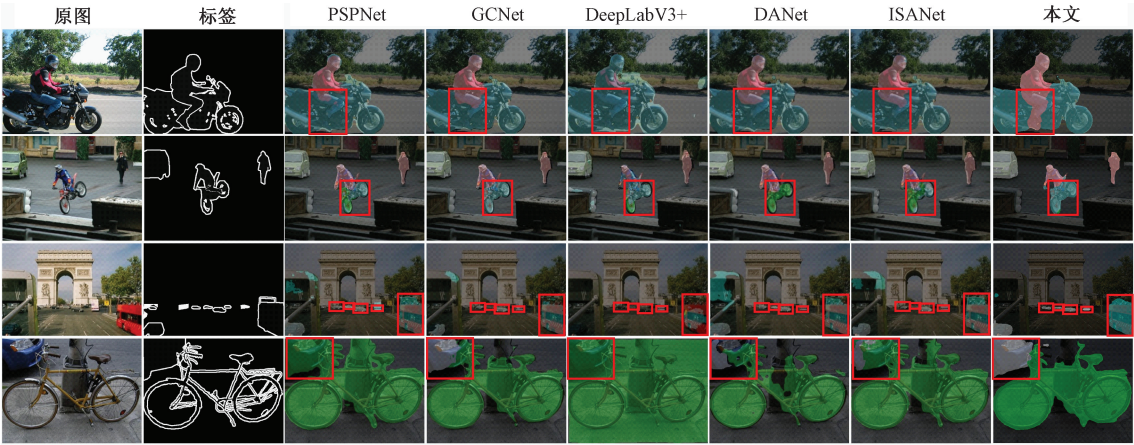


图 7 分割结果对比

Fig. 7 Contrast of segmentation results

分网络将轿车尾部一部分像素归类为轿车类别,一部分像素归类为自行车类别,这都是不符合常理的,但丰富的上下文信息使得本文提出的分割方法避免了这样的错误。

图 7 的分割结果对比说明了模型不仅对不同类别事物的分割有更好的效果,对于同类事物的轮廓分割也更为精准,因此本文提出的网络设计是有效的。

为了进一步说明本文对网络设计的合理性,本文进一步进行了模块间的消融实验。其结果如表 4 所示。

从表 4 的实验结果可以看到,在不添加 FEM 模块以及 ASPSP 模块时,网络模型的分割精度为 65.95%。在此基础上加入 FEM 模块后,分割精度提升了 0.19%,提升不多;不加入 FEM 模块,只加入 ASPSP 模块时,分割精度提高了 7.38%,这说明 ASPSP 对模型分割精度提升

非常有效;当 FEM 和 ASPSP 都在网络中时,模型又提升了 1.03%,这说明针对 ASPSP 模块损失空间信息所设计

的 FEM 模块也是有效的。

表 4 消融实验结果

Table 4 Ablation experiment results		
FEM	ASPSP	mIoU/%
		65.95
✓		66.14
	✓	73.33
✓	✓	74.36

由此可以看出,FEM 模块和 ASPSP 模块都有效地提升了分割性能,因此这说明本文提出的网络模型的设计是合理的。

本文所提出的模型不仅在分割准确性和实时性两方

面与目前的主流分割网络相比更具优势。在输入图片尺寸同等大小的情况下,浮点数运算量和参数量也更少,如表 5 所示。

表 5 模型复杂度对比
Table 5 Model complexity contrast

模型	Input size	#Params/M	FLOPs/GFLOPs
DANet	512×512	49.85	199.45
GCNet	512×512	49.62	197.98
ISANet	512×512	37.69	149.69
PSPNet	512×512	48.98	178.76
DeepLabV3+	512×512	43.58	176.63
Ours	512×512	17.74	98.68

从表 5 可看出,本文提出的模型参数量比最小的模型 ISANet 少 19.95 M,浮点数运算律也少 51.01 GFLOPs,这说明本文提出的模型相比于主流分割网络,对实际的自动驾驶工程应用更加便利,更易于自动驾驶工业化推进。

4 结 论

本文针对主流分割网络在应用于城市交通场景下,下采样过程损失空间信息不足导致的错误分割、漏分割等问题,提出了一种高效准确的实时语义分割方法,通过提出的 FEM 模块加强了下采样前的特征层,避免相似特征的丢失,这有利于分割属于同类别的像素。另外本文结合空洞卷积和条带卷积将大卷积核应用于 ASPSP 模块,提取了丰富的语义信息。

实验结果证明,该分割方法中的 FEM 模块和 ASPSP 模块设计合理有效,模型能够对行人、车辆、摩托车等常见自动驾驶场景中的事物进行准确有效的分割。相比于主流的分割模型,本文提出的分割网络有更好的分割精度,且分割速度也更快,达到了每秒处理 43 帧的速度,能够满足自动驾驶实时分割要求。ASPP 模块和 FEM 模块非常便利,且同样能应用于分类、识别、检测等其他视觉任务,能轻松添加到其他网络中的任意位置,即插即用。

需要说明的是,本文所提出的 FEM 模块在处理特征图时会占用一定量的内存,因此在实际的工程应用中,如果输入图像尺寸过大且经过骨干网络降采样的次数较少,则会对硬件内存有更高的要求;另外在实际的工程应用中,需要使用图片数量更大,类别标注更多的数据集对模型进行训练,避免自动驾驶车辆在行车过程中,对数据集里未曾出现过的类别产生错误判断。

自动驾驶实时语义分割领域中,模型如何以更少的参数提高车辆感知系统解析外界环境速度和精度;如何提高算法对雨雾、强光照等外界环境干扰的鲁棒性;如何

提高对小样本、小目标、高相似度等物体的分割精度,是复杂城市交通场景下,进一步推进基于语义分割的自动驾驶环境感知技术发展值得研究的方向。

参考文献

[1] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[2] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.

[3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv Preprint, 2014, ArXiv:1409.1556.

[4] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [J]. ArXiv Preprint, 2020, ArXiv:2010.11929.

[5] ZHENG S, LU J, ZHAO H, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 6881-6890.

[6] VASWANI A, SHAZEER N, PARMAR, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.

[7] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.

[8] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing net-work [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.

[9] 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 and Machine Intelligence, 2017, 40 (4): 834-848.

[10] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Re thinking atrous convolution for semantic image segmentation [J]. ArXiv Preprint, 2017, ArXiv:1706.05587.

[11] YU C, WANG J, PENG C, et al. Bisenet: Bilateral segmentation network for real-time semantic segmentation [C]. Proceedings of the European Conference on Computer

- Vision (ECCV), 2018: 325-341.
- [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: 3051-3068.
- [13] PENG C, ZHANG X, YU G, et al. Large kernel matters—improve semantic segmentation by global convolutional net-work [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4353-4361.
- [14] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [15] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [17] WANG Q, WU B, ZHU P, et al. Eca-net: Efficient channel attention for deep convolutional neural net-works [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [18] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.
- [19] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7794-7803.
- [20] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv Preprint, 2017, arXiv:1704.04861.
- [21] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [22] 孟俊熙, 张莉, 曹洋, 等. 基于 Deeplabv3+ 的图像语义分割算法优化研究 [J]. 激光与光电子学进展, 2022, 59(16): 1610009.
- MENG J X, ZHANG L, CAO Y. et al, Research on optimization of image semantic segmentation algorithms based on Deeplabv3+ [J]. Laser & Optoelectronics Progress, 2022, 59(16): 1610009.
- [23] 石敏, 沈佳林, 易清明, 等. 快速超轻量城市交通场景语义分割 [J]. 计算机科学与探索, 2022, 16(10): 2377-2386.
- SHI M, SHEN J L, YI Q M, et al. Rapid and ultra-lightweight semantic segmentation in urban traffic scene [J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(10): 2377-2386.
- [24] 樊博, 高玮玮, 单明陶, 等. 融合注意力机制与重影特征映射的无人机交通场景目标轻量级语义分割 [J]. 电子测量与仪器学报, 2023, 37(3): 21-28.
- FAN B, GAO W W, SHAN M T, et al. Lightweight semantic segmentation of UAV traffic scene objects combining attention mechanism and ghost feature mapping [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3): 21-28.

作者简介



周勇, 2019 年于重庆理工大学获得学士学位, 现为昆明理工大学在校硕士研究生, 主要研究方向为深度学习、语义分割。

E-mail: zhou_yog@163.com

Zhou Yong received his B. Sc. degree from Chongqing University of Technology in 2019. Now he is a M. Sc. candidate in Kunming University of Science and Technology. His main research interests include deep learning and semantic segmentation.



刘泓滨 (通信作者), 1983 年于昆明工学院获得学士学位, 1995 年于昆明工学院获得硕士学位, 现为昆明理工大学教授, 主要研究方向为机器人技术及应用。

E-mail: lihhong6696@163.com

Liu Hongbin (Corresponding author) received his B. Sc. degree from Kunming Institute of Technology in 1983, M. Sc. degree from Kunming Institute of Technology in 1983, respectively. Now he is professor in Kunming University of Science and Technology. His main research interests include robot technology and application.



侯亚东, 2020 年于天水师范学院获得学士学位, 现为昆明理工大学在校硕士研究生, 主要研究方向为深度学习、目标检测。

E-mail: yadong_hou@163.com

Hou Yadong received his B. Sc. degree from Tianshui Normal University in 2020. Now he is a M. Sc. candidate in Kunming University of Science and Technology. His main research interests include deep learning and object detection.