



(12) 发明专利

(10) 授权公告号 CN 112634289 B

(45) 授权公告日 2022.05.27

(21) 申请号 202011585360.4

(22) 申请日 2020.12.28

(65) 同一申请的已公布的文献号  
申请公布号 CN 112634289 A

(43) 申请公布日 2021.04.09

(73) 专利权人 华中科技大学  
地址 430074 湖北省武汉市洪山区珞喻路  
1037号

(72) 发明人 周瑜 龚石 白翔 方聪 李益群

(74) 专利代理机构 深圳市六加知识产权代理有  
限公司 44372  
专利代理师 向彬

(51) Int. Cl.

- G06T 7/11 (2017.01)
- G06V 10/46 (2022.01)
- G06V 10/764 (2022.01)
- G06V 10/80 (2022.01)
- G06V 10/82 (2022.01)
- G06K 9/62 (2022.01)
- G06N 3/04 (2006.01)
- G06N 3/08 (2006.01)

(56) 对比文件

- CN 110059768 A, 2019.07.26
- CN 110020606 A, 2019.07.16
- CN 109816100 A, 2019.05.28
- CN 110263833 A, 2019.09.20
- CN 111767922 A, 2020.10.13
- CN 110992379 A, 2020.04.10
- CN 111161279 A, 2020.05.15
- CN 111882620 A, 2020.11.03
- CN 109948524 A, 2019.06.28
- CN 111914654 A, 2020.11.10
- CN 110751111 A, 2020.02.04
- CN 110490884 A, 2019.11.22
- US 10067509 B1, 2018.09.04
- US 2018157904 A1, 2018.06.07

Z. Zhu et al..Asymmetric Non-Local  
Neural Networks for Semantic  
Segmentation.《2019 IEEE/CVF International  
Conference on Computer Vision》.2020,1-10.

Wang, J. et al..ADSCNet: asymmetric  
depthwise separable convolution for  
semantic segmentation in real-time.  
《Applied intelligence》.2019,1-12.

审查员 黎成超

权利要求书2页 说明书4页 附图1页

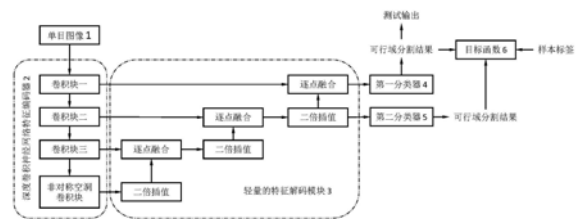
(54) 发明名称

一种基于非对称空洞卷积的快速可行域分割方法

(57) 摘要

本发明公开了一种基于非对称空洞卷积的快速可行域分割方法:对单目相机采集的图像进行多尺度图像特征提取,在最后一个卷积块使用非对称空洞卷积块消除局部噪声,获得具有区分力的表征;特征解码模块对特征编码器提取的多尺度图像特征进行逐点融合,得到高分辨率的高区分力的图像特征图;采用分类器基于融合后的图像特征,预测输出图像场景中的可行域分割结果,将图像中的所有像素划分为可行域和不

可行域区域两类。本发明引入了一种全新的非对称空洞卷积模块来提高特征的区分力,大幅减少对不可行驶道路的误判,并且没有引入额外的计算量。基于轻量化的模块和网络设计,在保证精度的前提下,实现了对可行域的快速分割。



1. 一种基于非对称空洞卷积的快速可行域分割方法,其特征在于,包括以下步骤:

步骤S1,多尺度特征提取,利用深度卷积神经网络特征编码器对单目相机采集的图像进行多尺度图像特征提取;所述步骤S1中的深度卷积神经网络特征编码器,采用ResNet作为骨干网络进行多尺度图像特征提取,将第四个卷积块替换成非对称空洞卷积块;从第一个到第四个卷积块输出的图像特征图尺寸分别为输入图像的1/4,1/8,1/16和1/32,通道数分别为 $C_1, C_2, C_3, C_4$ ,其中 $C_1, C_2, C_3, C_4$ 分别为预设值;

步骤S2,采用特征解码模块对步骤S1提取的多尺度图像特征进行融合,得到分辨率扩大的高层语义的特征图;步骤S2中特征解码模块采用反向逐层融合的策略,在每层进行逐点融合,具体包括:步骤S2-1,将步骤S1中第四个卷积块输出的1/32分辨率的特征图进行2倍双线性插值,并将它与第三个卷积块输出的1/16分辨率特征图在通道维度上堆叠,使用 $1 \times 1$ 卷积进行逐点融合  $R^{(C_4+C_3)} \rightarrow R^C$ ,然后使用ReLU激活函数进行激活;步骤S2-2,将步骤S2-1输出的特征图进行2倍双线性插值,并将其与步骤S1中第二个卷积块输出的1/8分辨率特征图在通道维度上堆叠,使用 $1 \times 1$ 卷积进行逐点融合  $R^{(C+C_2)} \rightarrow R^C$ ,使用ReLU进行激活;步骤S2-3,将步骤S2-2输出的特征图进行2倍双线性插值,并将其与步骤S1中第一个卷积块输出的1/4分辨率特征图在通道维度上堆叠,使用 $1 \times 1$ 卷积做逐点融合  $R^{(C+C_1)} \rightarrow R^C$ ,将并使用ReLU激活;

步骤S3,逐像素预测,基于步骤S2输出的图像特征,采用分类器预测图像场景中的可行域分割结果,将输入单帧图像的所有像素划分为可行驶区域和不可行驶区域两类。

2. 如权利要求1所述的基于非对称空洞卷积的快速可行域分割方法,其特征在于,所述深度卷积神经网络特征编码器中将第四个卷积块替换成非对称空洞卷积块,具体为:

将普通的 $3 \times 3$ 卷积替换成 $3 \times 3$ 非对称空洞卷积,非对称空洞卷积核的水平半径和垂直半径分别为 $r_h$ 和 $r_v$ ,非对称率为 $r = \frac{r_h}{r_v}, r \gg 1$ ,基本空洞率为 $d, d, r_h$ 和 $r_v$ 分别为预设值。

3. 如权利要求1所述的基于非对称空洞卷积的快速可行域分割方法,其特征在于,在训练阶段,步骤S3包含第一分类器和第二分类器两个分类器,各自输出最终的可行域分割概率图;第一分类器的输入为步骤S2-3的输出,它使用 $1 \times 1$ 卷积进行逐像素的预测 $R^C \rightarrow R^2$ ,其中 $R^2$ 表示可行域/非可行域的分类得分,使用softmax函数对得分进行归一化,将结果插值到输入图像尺寸作为最终预测的结果;第二分类器作为辅助分支,它的输入为步骤S2-2的输出,同样使用 $1 \times 1$ 卷积预测分类分数,并经过softmax函数做概率归一化。

4. 如权利要求3所述的基于非对称空洞卷积的快速可行域分割方法,其特征在于,在测试阶段,仅保留第一分类器输出的可行域预测结果。

5. 如权利要求1所述的基于非对称空洞卷积的快速可行域分割方法,其特征在于,采用端到端的方式进行模型的参数学习,优化的目标函数为:

$$L=L_1+\gamma \cdot L_2$$

其中 $L_1$ 和 $L_2$ 分别为第一分类器和第二分类器输出的概率图与真实值之间的逐像素难样本挖掘-交叉熵损失, $\gamma$ 为超参数用来控制两项损失值的权重。

6. 如权利要求5所述的基于非对称空洞卷积的快速可行域分割方法,其特征在于,难样本挖掘-交叉熵损失如下公式所示:

$$L_{hm} = \frac{1}{N_p} \sum_{i,j} I(p^{i,j} < \lambda_b) L_{bce}^{i,j}$$

其中 $\lambda_b$ 为批数据中的置信度阈值, $p$ 是预测的概率, $I(\cdot)$ 是指示函数, $N_p$ 是批数据中的像素数量, $i, j$ 是图像中的像素索引, $L_{bce}$ 是二元交叉熵损失函数。

## 一种基于非对称空洞卷积的快速可行域分割方法

### 技术领域

[0001] 本发明属于计算机视觉技术领域,更具体地,涉及一种基于非对称空洞卷积的快速可行域分割方法。

### 背景技术

[0002] 近年来,可行域分割已经成为自动驾驶技术中的研究热点。由于车载相机相比高精度3D激光雷达具有极大的成本优势,因此以单目图像作为输入的可行域分割成为了自动驾驶中不可或缺的部分。在可行域分割任务中,图像中的像素被预定义为三类:背景,可行驾驶道路,以及不可行驶道路(比如高速公路的逆向车道,人行道等)。这里的可行域分割任务是要将图像中的可行驾驶道路像素分割出来。

[0003] 早期的可行域分割方法使用低层次的特征,比如颜色,边缘和纹理等,进行逐像素或者逐块分类,并引入条件随机场来建模全局依赖关系。近年来,由于高性能并行计算能力的提升和大规模数据集的提出,深度卷积网络在计算机视觉领域中取得了突破性进展,可行域分割领域也随之进展迅速。

[0004] 现有的基于深度卷积网络的方法,大多依赖于复杂的网络设计来提升精度,但是造成了极大的计算复杂度,难以应用于实时性的自动驾驶场景。还有少数方法通过降低模型复杂度来提升模型推理速度,但是它们也无法在移动嵌入式平台达到实时性。

### 发明内容

[0005] 针对现有技术的以上缺陷或改进需求,本发明提供了一种快速、准确的基于非对称空洞卷积的快速可行域分割方法。为达到上述目的,本发明提供如下技术方案:

[0006] 一种基于非对称空洞卷积的快速可行域分割方法,包括以下步骤:

[0007] 步骤S1,多尺度特征提取,利用深度卷积神经网络特征编码器(2)对单目相机采集的图像(1)进行多尺度图像特征提取;

[0008] 步骤S2,采用特征解码模块(3)对步骤S1提取的多尺度图像特征进行融合,得到分辨率扩大的高层语义的特征图;

[0009] 步骤S3,逐像素预测,基于步骤S2输出的图像特征,采用分类器预测图像场景中的可行域分割结果,将输入单帧图像的所有像素划分为可行驾驶区域和不可行驶区域两类。

[0010] 优选地,深度卷积神经网络特征编码器(2)分为四个卷积块:

[0011] 将第四个卷积块替换成非对称空洞卷积块,以消除局部噪声;从第一到第四个卷积块输出的图像特征图尺寸分别为输入图像的 $1/4, 1/8, 1/16$ 和 $1/32$ ,通道数分别为 $C_1, C_2, C_3, C_4$ ,其中 $C_1, C_2, C_3, C_4$ 分别为预设值。

[0012] 优选地,所述深度卷积神经网络特征编码器(2)中将第四个卷积块替换成非对称空洞卷积块,假设非对称空洞卷积核的水平半径和垂直半径分别为 $r_h$ 和 $r_v$ ,非对称率为 $r = \frac{r_h}{r_v}$ ,基本空洞率为 $d$ ,则由于不同的区域(即不可行驶道路,可行驾驶道路和背景)分布在

水平方向上,  $r \gg 1$  的情况下非对称空洞卷积获取水平方向上不同区域之间的上下文关联, 同时尽可能避免引入背景噪声,  $d$ 、 $r_h$  和  $r_v$  分别为预设值。

[0013] 优选地, 特征解码模块 (3) 采用反向逐层融合的策略, 在每层进行逐点融合, 具体包括:

[0014] 步骤S2-1, 将步骤S1中第四个卷积块输出的1/32分辨率的特征图进行2倍双线性插值, 并将它与第三个卷积块输出的1/16分辨率特征图在通道维度上堆叠, 使用  $1 \times 1$  卷积进行逐点融合  $R^{(C_4+C_3)} \rightarrow R^C$ , 然后使用ReLU激活函数进行激活。

[0015] 步骤S2-2, 将步骤S2-1输出的特征图进行2倍双线性插值, 并将其与步骤S1中第二个卷积块输出的1/8分辨率特征图在通道维度上堆叠, 使用  $1 \times 1$  卷积进行逐点融合  $R^{(C+C_2)} \rightarrow R^C$ , 使用ReLU进行激活。

[0016] 步骤S2-3, 将步骤S2-2输出的特征图进行2倍双线性插值, 并将其与步骤S1中第一个卷积块输出的1/4分辨率特征图在通道维度上堆叠, 使用  $1 \times 1$  卷积做逐点融合  $R^{(C+C_1)} \rightarrow R^C$ , 将并使用ReLU激活。

[0017] 优选地, 在训练阶段, 步骤S3包含第一分类器和第二分类器两个分类器, 各自输出最终的可行域分割概率图; 第一分类器 (4) 的输入为S2-3的输出 (1/4的分辨率), 它使用  $1 \times 1$  卷积进行逐像素的预测  $R^C \rightarrow R^2$  ( $C$  被设置为128), 其中  $R^2$  分别表示可行域/非可行域的分类得分, 使用softmax函数对得分进行归一化, 将结果插值到输入图像尺寸作为最终预测的结果。第二分类器 (5) 作为辅助分支, 它的输入为步骤S2-2的输出 (1/8的分辨率), 同样使用  $1 \times 1$  卷积预测分类分数, 并经过softmax函数做概率归一化。在测试阶段, 仅保留第一分类器 (4) 输出的可行域预测结果。

[0018] 优选地, 本方法采用端到端的方式进行模型的参数学习, 优化的目标函数 (6) 为:

$$[0019] \quad L = L_1 + \gamma \cdot L_2$$

[0020] 其中  $L_1$  和  $L_2$  分别为第一分类器和第二分类器输出的概率图与真实值之间的逐像素难样本挖掘-交叉熵损失 (7),  $\gamma$  为超参数用来控制两项损失值的权重。

[0021] 优选地, 难样本挖掘-交叉熵损失 (7) 旨于解决数据集中仅包含少量的难样本 (如可行域边缘, 阴影区域, 过曝区域等), 以及大量的容易样本 (如平坦的路面区域, 天空等), 自动地给难样本以更大的训练强度, 如下公式所示:

$$[0022] \quad L_{hm} = \frac{1}{N_p} \sum_{i,j} I(p^{i,j} < \lambda_b) L_{bce}^{i,j}$$

[0023] 其中  $\lambda_b$  为批数据中的置信度阈值,  $p$  是预测的概率,  $I(\cdot)$  是指示函数,  $N_p$  是批数据中的像素数量,  $i, j$  是图像中的像素索引,  $L_{bce}$  是二元交叉熵损失函数。

[0024] 总体而言, 通过本发明所构思的以上技术方案与现有技术相比, 具有如下有益效果:

[0025] (1) 本发明提出了一种新颖的非对称空洞卷积模块, 该模块采用非对称空洞卷积核, 用来在水平方向上捕捉更多的上下文信息, 增强不同区域之间的特征区分力, 减少对非可行道路路的误判。该模块可以作为特征提取网络的一部分, 并没有带来额外的计算开销, 使得提取的特征更具鉴别力, 使得后续的特征解码器可以设计得更加轻巧而不会带来精度

下降；

[0026] (2) 本发明提出了一种新颖的特征解码模块,该模块摒弃了传统的解码器的复杂设计,完全基于轻量的 $1 \times 1$ 卷积操作,用于逐步恢复特征图的分辨率。与经典的特征解码器U-Net相比,该模块降低了94%的参数和93%的计算开销。

## 附图说明

[0027] 图1是本发明提供的基于非对称空洞卷积的快速可行域分割方法的整体流程图。

## 具体实施方式

[0028] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。此外,下面所描述的本发明各个实施方式中所涉及到的技术特征只要彼此之间未构成冲突就可以相互组合。

[0029] 本发明提供了一种基于非对称空洞卷积的快速可行域分割方法,如图1所示,包括以下步骤:

[0030] 步骤S1,多尺度特征提取,利用深度卷积神经网络特征编码器(2)对单目相机采集的图像(1)进行多尺度图像特征提取;

[0031] 步骤S2,采用特征解码模块(3)对步骤S1提取的多尺度图像特征进行融合,得到高分辨率高语义的图像特征;

[0032] 步骤S3,逐像素预测,基于步骤S2输出图像特征,采用分类器预测图像中的可行域分割结果,将输入单帧图像的所有像素划分为可行驶区域和不可行驶区域两类。

[0033] 主要有三个实现部分:1)多尺度的有鉴别力的特征提取;2)轻量化的特征解码器;3)分类器和目标函数。接下来对本发明中步骤进行详细说明。

[0034] 1. 多尺度的有鉴别力的特征提取

[0035] 本发明实施例采用ResNet18作为骨干网络进行多尺度特征提取,将第四个卷积块替换成非对称空洞卷积块,它使用 $3 \times 3$ 非对称空洞卷积,假设非对称空洞卷积核的水平半径和垂直半径分别为 $r_h$ 和 $r_v$ ,非对称率为 $r = \frac{r_h}{r_v}$ ,基本空洞率为 $d$ ,则由于不同的区域(即可行驶道路,可行驶道路和背景)分布在水平方向上, $r \gg 1$ 的情况下非对称空洞卷积获取水平方向上不同区域之间的上下文关联,同时尽可能避免引入背景噪声。实施例中 $d$ 的数值从前往后依次为3,5,5和12。为满足数据集中道路纵横比特点,实施例中将 $r$ 设置为3。

[0036] 2. 轻量化的特征解码器

[0037] 特征解码模块(3)采用反向逐层融合的策略,在每层进行逐点融合,具体包括:

[0038] 步骤S2-1,将步骤S1中第四个卷积块输出的 $1/32$ 分辨率的特征图进行2倍双线性插值,并将它与第三个卷积块输出的 $1/16$ 分辨率特征图在通道维度上堆叠,使用 $1 \times 1$ 卷积进行逐点融合 $R^{(C_4+C_3)} \rightarrow R^C$ ,然后使用ReLU激活函数进行激活。

[0039] 步骤S2-2,将步骤S2-1输出的特征图进行2倍双线性插值,并将其与步骤S1中第二个卷积块输出的 $1/8$ 分辨率特征图在通道维度上堆叠,使用 $1 \times 1$ 卷积进行逐点融合

$R^{(C+C_2)} \rightarrow R^C$ , 并使用ReLU进行激活。

[0040] 步骤S2-3, 将步骤S2-2输出的特征图进行2倍双线性插值, 并将其与步骤S1中第一个卷积块输出的1/4分辨率特征图在通道维度上堆叠, 使用 $1 \times 1$ 卷积做逐点融合  $R^{(C+C_1)} \rightarrow R^C$ , 将并使用ReLU激活。

[0041] 3. 分类器和目标函数

[0042] 本发明实施例在训练阶段, 步骤S3包含第一分类器(4)和第二分类器(5)两个分类器, 各自输出最终的可行域分割概率图; 第一分类器(4)的输入为S2-3的输出(1/4的分辨率), 它使用 $1 \times 1$ 卷积进行逐像素的预测  $R^C \rightarrow R^2$  (C被设置为128), 其中 $R^2$ 表示可行域/非可行域的分类得分, 使用softmax函数对得分进行归一化, 将结果插值到输入图像尺寸作为最终预测的结果。第二分类器(5)作为辅助分支, 它的输入为步骤S2-2的输出(1/8的分辨率), 同样使用 $1 \times 1$ 卷积预测分类分数, 并经过softmax函数做概率归一化。在测试阶段, 仅保留第一分类器(4)输出的可行域预测结果。本发明实施例采用端到端的方式进行模型的参数学习, 优化的目标函数(6)为:

$$[0043] \quad L = L_1 + \gamma \cdot L_2$$

[0044] 其中 $L_1$ 和 $L_2$ 分别为第一分类器和第二分类器输出的概率图与真实值之间的逐像素难样本挖掘-交叉熵损失,  $\gamma$ 为超参数用来控制两项损失值的权重, 本发明实施例设定 $\gamma = 0.5$ 。

[0045] 优选地, 难样本挖掘-交叉熵损失旨在解决数据集中仅包含少量的难样本(如可行域边缘, 阴影区域, 过曝区域等), 以及大量的容易样本(如平坦的路面区域, 天空等), 自动地给难样本以更大的训练强度, 如下公式所示:

$$[0046] \quad L_{hm} = \frac{1}{N_p} \sum_{i,j} I(p^{i,j} < \lambda_b) L_{bce}^{i,j}$$

[0047] 其中 $\lambda_b$ 为批数据中的置信度阈值,  $p$ 是预测的概率,  $I(\cdot)$ 是指示函数,  $N_p$ 是批数据中的像素数量,  $i, j$ 是图像中的像素索引,  $L_{bce}$ 是二元交叉熵损失函数。

[0048] 本发明提出了一种基于非对称空洞卷积的快速可行域分割方法。采用了一个浅层的特征编码器进行多尺度的特征提取, 并设计了一个轻量级的特征解码器恢复特征图的分辨率。为了避免网络简化带来的精度下降, 本发明引入了一种全新的非对称空洞卷积模块来提高特征的区分力, 大幅减少对不可行驶道路的误判, 并且没有引入额外的计算量。基于轻量化的模块和网络设计, 本发明在保证精度的前提下, 实现了对可行域的快速分割。

[0049] 本领域的技术人员容易理解, 以上所述仅为本发明的较佳实施例而已, 并不用以限制本发明, 凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等, 均应包含在本发明的保护范围之内。

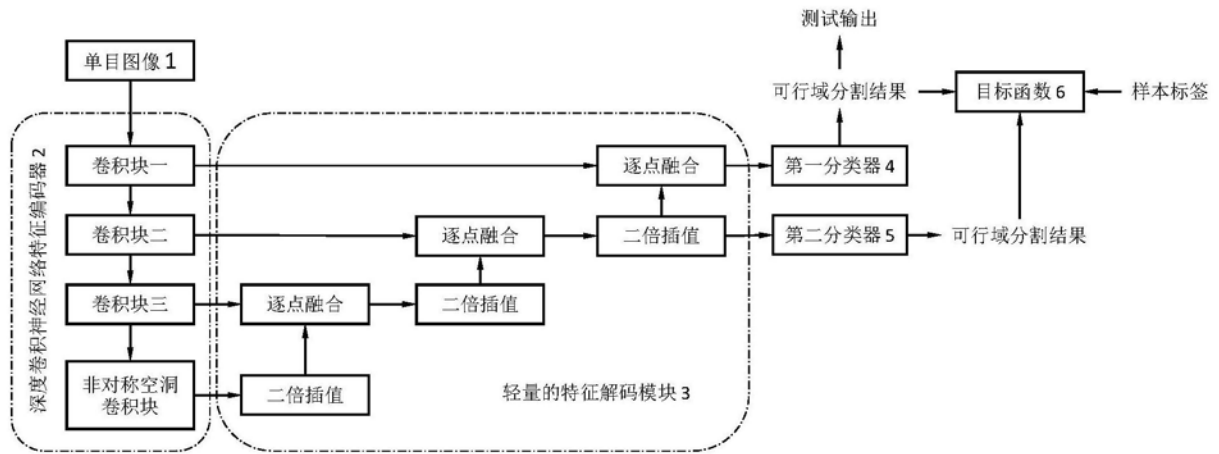


图1