



(12) 发明专利申请

(10) 申请公布号 CN 105654067 A

(43) 申请公布日 2016. 06. 08

(21) 申请号 201610073326. 6

(22) 申请日 2016. 02. 02

(71) 申请人 北京格灵深瞳信息技术有限公司
地址 100091 北京市海淀区玉带路 6 号

(72) 发明人 张德兵

(74) 专利代理机构 北京新知远方知识产权代理
事务所(普通合伙) 11397
代理人 申楠

(51) Int. Cl.

G06K 9/00(2006. 01)

G06K 9/46(2006. 01)

G06K 9/62(2006. 01)

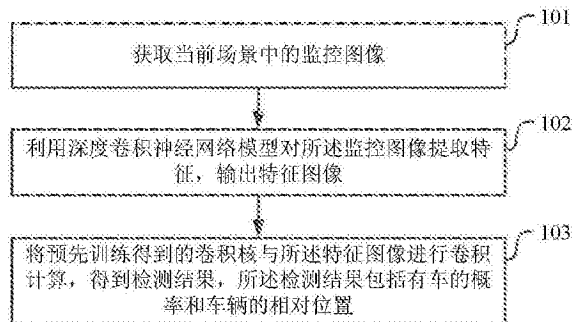
权利要求书2页 说明书11页 附图2页

(54) 发明名称

一种车辆检测方法及装置

(57) 摘要

本申请提供了一种车辆检测方法及装置,包括:获取当前场景中的监控图像;利用深度卷积神经网络模型对所述监控图像提取表达能力更强的视觉特征,输出特征图像;将预先训练得到的卷积核与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。由于本申请实施例中利用深度卷积神经网络模型对所述监控图像提取特征输出特征图像,后续只需要将卷积核与所述特征图像进行计算即可,无需对原始监控图像进行预处理、无需再对整张原始监控图像进行滑动框扫描,且不需要对原始图像进行多尺度特征计算,从而节省了计算时间、大大提高了检测效率和检测准确度。



1. 一种车辆检测方法,其特征在于,包括如下步骤:
获取当前场景中的监控图像;
利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;
将预先训练得到的卷积核与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。
2. 如权利要求1所述的方法,其特征在于,所述利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像,具体为:利用深度卷积神经网络模型对所述监控图像提取不同的特征,输出k个特征图像;
所述将预先训练得到的卷积核与所述特征图像进行卷积计算,具体为:将预先训练得到的 $5*k*scale*ratio$ 个卷积核与所述k个特征图像分别进行卷积计算;其中,所述scale为车辆尺度的种类,所述ratio为车辆长宽比的种类。
3. 如权利要求1所述的方法,其特征在于,所述卷积计算具体为全卷积计算。
4. 如权利要求1所述的方法,其特征在于,所述预先训练得到的卷积核为多个,所述将预先训练得到的卷积核与所述特征图像进行卷积计算,具体为:
将所述多个预先训练得到的卷积核并行与所述特征图像进行卷积计算。
5. 如权利要求4所述的方法,其特征在于,在进行卷积计算时如果同一辆车在不同位置产生响应,采用非最大值抑制方法将输出的所述车辆的多个相对位置进行合并。
6. 如权利要求1所述的方法,其特征在于,进一步包括:
根据所述车辆的相对位置和预先设置的相对参考框,计算所述车辆在所述监控图像中的绝对位置。
7. 如权利要求1所述的方法,其特征在于,进一步包括:
利用校验模型对所述检测结果进行校验,所述校验模型是通过对正例和负例进行训练得到的。
8. 如权利要求1所述的方法,其特征在于,所述卷积核的训练过程具体为:
获取若干车辆样本图像,所述样本图像中每辆车标注有外接矩形框;
利用深度卷积神经网络模型提取所述样本图像的特征,输出特征图像;
根据卷积核与所述特征图像检测所述样本图像中的车辆,根据检测结果以及所述外接矩形框调整所述卷积核内的参数,直至所述检测结果中车辆的位置与所述外接矩形框接近或重合。
9. 一种车辆检测装置,其特征在于,包括:
获取模块,用于获取当前场景中的监控图像;
特征提取模块,用于利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;
训练模块,用于训练卷积核;
检测模块,用于将所述训练得到的卷积核与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。
10. 如权利要求9所述的装置,其特征在于,所述特征提取模块具体用于利用深度卷积神经网络模型对所述监控图像提取不同的特征,输出k个特征图像;所述检测模块具体用于将预先训练得到的 $5*k*scale*ratio$ 个卷积核与所述k个特征图像分别进行卷积计算,得到

检测结果,所述检测结果包括有车的概率和车辆的相对位置;其中,所述scale为车辆尺度的种类,所述ratio为车辆长宽比的种类。

11.如权利要求9所述的装置,其特征在于,所述卷积计算具体为全卷积计算。

12.如权利要求9所述的装置,其特征在于,所述检测模块具体用于将所述训练得到的多个卷积核并行与所述特征图像进行卷积计算。

13.如权利要求12所述的装置,其特征在于,所述检测模块进一步用于在进行卷积计算时如果同一辆车在不同位置产生响应,采用非最大值抑制方法将输出的所述车辆的多个相对位置进行合并。

14.如权利要求9所述的装置,其特征在于,进一步包括:

计算模块,用于根据所述车辆的相对位置和预先设置的相对参考框,计算所述车辆在所述监控图像中的绝对位置。

15.如权利要求9所述的装置,其特征在于,进一步包括:

校验模块,用于利用校验模型对所述检测结果进行校验,所述校验模型是通过对正例和负例进行训练得到的。

16.如权利要求9所述的装置,其特征在于,所述训练模块具体包括:

获取单元,用于预先获取若干车辆样本图像,所述样本图像中每辆车标注有外接矩形框;

特征提取单元,用于利用深度卷积神经网络模型提取所述样本图像的特征,输出特征图像;

训练单元,用于根据卷积核与所述特征图像检测所述样本图像中的车辆,根据检测结果以及所述外接矩形框调整所述卷积核内的参数,直至所述检测结果中车辆的位置与所述外接矩形框接近或重合。

一种车辆检测方法及装置

技术领域

[0001] 本申请涉及计算机视觉技术领域,尤其涉及一种车辆检测方法及装置。

背景技术

[0002] 精确鲁棒的车辆检测对于安防监控(特别是在卡口和电子警察场景下)具有重要的应用意义,路上经过的每一辆车都是安防监控领域中非常关心的目标。

[0003] 现有的车辆检测技术大多采用如下方式:首先对监控画面进行图像预处理,比如直方图均衡、对比度调整等;然后提取预先定义的一系列特征,比如边缘、纹理、各个方向上的梯度特征等;最后利用上述特征,依次对经过预处理的图像中不同大小、长宽比或位置的窗口逐一进行分类判断,如果大于某个阈值则认为车辆。

[0004] 采用上述方式对车辆进行检测,需要首先对图片进行预处理,然后对整张图片进行特征提取,再将经过处理的图片按照不同尺度缩放、分类处理、多尺度计算等,耗费了大量时间,导致检测效率极低。

[0005] 现有技术不足在于:

[0006] 现有车辆检测方案计算量大、检测速度慢、检测场景单一、检测精度低。

发明内容

[0007] 本申请实施例提出了一种车辆检测方法及装置,以解决现有技术中车辆检测方案计算量大、检测速度慢、检测场景单一、检测精度低的技术问题。

[0008] 本申请实施例提供了一种车辆检测方法,包括如下步骤:

[0009] 获取当前场景中的监控图像;

[0010] 利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;

[0011] 将预先训练得到的卷积核与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。

[0012] 本申请实施例提供了一种车辆检测装置,包括:

[0013] 获取模块,用于获取当前场景中的监控图像;

[0014] 特征提取模块,用于利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;

[0015] 训练模块,用于预先训练卷积核;

[0016] 检测模块,用于将预先训练得到的卷积核与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。

[0017] 有益效果如下:

[0018] 本申请实施例所提供的车辆检测方法及装置,在获取到当前场景中的监控图像后,可以利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像,将训练得到的卷积核与所述特征图像进行卷积计算得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。由于本申请实施例中利用深度卷积神经网络模型对所述监控图像提取特征

输出特征图像,后续只需要将卷积核与所述特征图像进行计算即可,可以适用于卡口或电子警察等多个检测场景,无需对原始监控图像进行预处理、无需再对整张原始监控图像进行滑动框扫描,且不需要对原始图像进行多尺度特征计算,从而节省了计算时间、大大提高了检测效率和检测精度。

附图说明

[0019] 下面将参照附图描述本申请的具体实施例,其中:

[0020] 图1示出了本申请实施例中车辆检测方法实施的流程示意图;

[0021] 图2示出了本申请实施例中车辆检测过程示意图;

[0022] 图3示出了本申请实施例中车辆检测装置的结构示意图。

具体实施方式

[0023] 为了使本申请的技术方案及优点更加清楚明白,以下结合附图对本申请的示例性实施例进行进一步详细的说明,显然,所描述的实施例仅是本申请的一部分实施例,而不是所有实施例的穷举。并且在不冲突的情况下,本说明中的实施例及实施例中的特征可以互相结合。

[0024] 针对现有技术的不足,本申请实施例提出了车辆方法及装置,可以同时支持多种复杂场景、支持任意大小/比例的车辆且可以在CPU/GPU快速实现高精度的实时监测,下面进行说明。

[0025] 图1示出了本申请实施例中车辆检测方法实施的流程示意图,如图所示,所述车辆检测方法可以包括如下步骤:

[0026] 步骤101、获取当前场景中的监控图像;

[0027] 步骤102、利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;

[0028] 步骤103、将预先训练得到的卷积核作为滑动窗口与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。

[0029] 具体实施时,可以在需要监控的场景中预先布置监控摄像头,由摄像头回传当前场景中的监控图像。实际使用中,摄像头回传的可能为监控视频,监控视频可以包括若干帧的监控图像。

[0030] 本申请实施例中的监控图像可以为红绿蓝RGB图像,一个图像可以按照颜色的不同分为三张图像,具体实施时可以分别以三张图像进行后续操作。

[0031] 深度学习在处理图像问题中展现了巨大的优势,其深层的网络结构和端到端的优化思想使得让机器自动从数据中进行学习成为可能。卷积神经网络(Convolutional Neural Networks,简称CNN)是深度学习较为热门的方法之一,一般地,卷积神经网络的基本结构可以包括两层,一层是特征提取层,一层是特征映射层。

[0032] 本申请实施例利用深度卷积神经网络模型对所述监控图像进行特征提取,输出特征图像,这个输出就是通过神经网络学习得到的用于检测的特征。其中,卷积神经网络模型可以采用现有的模型,例如现有的Google Net模型,这个模型是由卷积层和感知模块Inception Module堆叠而成。

[0033] 本申请实施例中所述卷积核可以理解为根据车的部分或整体特征(车头、车灯、车

尾、车轱辘、车身等)进行样本训练得到的数字表示,所述卷积核可以为 $1*1$ 、 $3*3$ 、 $5*5$ 等大小,卷积核中包括多个数值,例如: $3*3$ 的卷积核中包括9个数值。

[0034] 假设对监控图像提取特征后得到1个特征图像,采用一定大小(假设为 $3*3$)的卷积核,可以利用所述卷积核与上述特征图像进行卷积计算,最终得到一个输出(大小可以为 $18*31$),所述输出的每个位置可以表示在所述特征图像中该位置存在一个固定大小和比例的车辆的概率。

[0035] 再利用1个卷积核与上述特征图像进行卷积计算,得到左上角x的相对偏移位置;同理,左上角y的相对偏移位置、右下角x的相对偏移位置、右下角y的相对偏移位置等各需要1个卷积核与特征图像进行计算得到。

[0036] 因此,对于一个固定比例和尺度的车辆的检测,当提取出的特征图像为1个时,可以利用5个一定大小(假设为 $3*3$)的卷积核(卷积核中的参数总共大小为 $5*3*3$)进行计算,卷积核与特征图像进行卷积计算之后,可以得到5张图(大小为 $5*18*31$ 的输出,即5张 $18*31$ 的图)。

[0037] 其中,第一张图可以表示每个位置有车的概率,所有的数字都在 $0\sim 1$ 之间,另外四张图分别可以表示车辆的相对位置的左上角x坐标、左上角y坐标、右下角x坐标、右下角y坐标的相对偏移位置。

[0038] 如果要对多尺度、多长宽比的车辆进行检测,假设需要处理scale个尺度、ratio种长宽比的车辆检测,本申请实施例可以利用 $5*scale*ratio$ 个卷积核进行卷积计算,最终输出 $5*scale*ratio$ 个图,每张输出的图可以表示某种尺度和比例的、某个属性(有车的概率或者相对偏移位置的四个坐标)的输出。

[0039] 其中,scale可以为正整数,例如可以为7,面积可以为 $50*50*2^i$,i可以为0、1、2、3、4、5、6等,ratio可以为正整数,例如可以为5,长宽比可以为 $3:1$ 、 $2:1$ 、 $1:1$ 、 $1:2$ 、 $1:3$ 等。

[0040] 本申请实施例通过深度卷积神经网络将原始的监控图像提取特征后缩小为特征图像,在所述特征图像中进行后续扫描、检测等操作与在原始监控图像中操作相比可以极大降低卷积核的滑动或检测时间。

[0041] 以 $3*3$ 大小的滑动窗口为例,本申请实施例只需要扫描 $3*3$ 的滑动窗口,不需要将监控图像变换不同的尺度进行扫描、检测,即,本申请实施例不需要根据物体的尺度变化而变化,降低了计算复杂度、提高了检测效率。

[0042] 具体实施时,卷积核中的参数与所述卷积神经网络模型中的参数可以同时训练得到,也可以先进行卷积核的训练、再进行模型的训练,本申请对卷积核训练和模型训练的顺序不作限制。

[0043] 本申请实施例所提供的车辆检测方法,无需对监控图像进行任何预处理,可以接受任意大小、任意长宽比的图像输入,本申请实施例可以对获取到的图像提取出表达能力更强的视觉特征,浓缩得到特征图像,后续只需要对所述特征图像进行检测即可,可以适用于卡口或电子警察等多个检测场景,无需对原始监控图像进行预处理、无需对整张原始监控图像进行滑动框扫描,且不需要对原始图像进行多尺度特征计算,从而节省了计算时间、大大提高了检测效率和检测精度。

[0044] 本申请实施例所提供的车辆检测方法不仅可以检测车辆,还可以对三轮车、非机动车、行人等目标进行检测,只需要预先训练这些目标的卷积核即可。

[0045] 实施中,所述监控图像可以为灰度图像、RGB图像、RGBD图像、YUV图像。

[0046] 具体实施中,本申请实施例所获取的监控图像可以为灰度图像、RGB图像、RGBD图像、YUV图像等多种图像格式,其中,所述RGBD图像表示包括深度信息的图像。当车辆之间有遮挡的情况下,也能准确的检测到被遮挡的车辆。

[0047] 实施中,所述利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像,具体可以为:利用深度卷积神经网络模型对所述监控图像提取不同的特征,输出k个特征图像;

[0048] 所述将预先训练得到的卷积核与所述特征图像进行卷积计算,具体可以为:将预先训练得到的 $5*k*scale*ratio$ 个卷积核与所述k个特征图像分别进行卷积计算;其中,所述scale为车辆不同尺度的数量,所述ratio为车辆不同长宽比的数量。

[0049] 经卷积神经网络模型提取特征后输出的特征图像可以为k张,每张特征图像则可以为 $18*31$ 等不同大小,特征图像的总的大小可以为 $k*18*31$ 。

[0050] 用于检测的卷积核可以有 $k*scale*ratio$ 个,具体实施时,scale可以有7种、ratio可以有5种,其中,不同的scale可以负责检测不同大小的车辆,例如: $50*50*2^0$ 大小的车辆、 $50*50*2^6$ 大小的车辆,不同的ratio可以负责检测不同长宽比的车辆,例如:1:3、1:2、1:1、2:1、3:1等长宽比的车辆。

[0051] 假设本申请实施例中利用深度卷积神经网络模型对所述监控图像提取出k个特征图像,那么可以利用k个卷积核分别与上述的k张特征图像进行对应的卷积计算,并将结果求和,这样得到的结果是一定大小的输出,每个位置上表示,在特征图像的该位置存在一个关系大小和比例的车辆的概率。在计算相对的左上角x的相对偏移位置时,可以再利用k个卷积核来进行计算。同理,左上角y的相对偏移、右下角x的相对偏移、右下角y的相对偏移分别可以利用k个卷积核计算得到。

[0052] 因此,对于一个固定比例和尺度的车辆进行检测,当提取得到k个特征图像时,本申请实施例可以利用 $k*5$ 个一定大小(假设为 $3*3$)的卷积核(卷积核中的参数总共大小可以为 $k*5*3*3$)进行计算求和,卷积核与特征图像进行对应的卷积计算之后,输出5张图(大小为 $5*18*31$ 的输出,即5张 $18*31$ 的图),分别代表相应当前尺度和比例的车概率,左上角x的相对偏移,左上角y的相对偏移,右下角x的相对偏移,右下角y的相对偏移。

[0053] 如果要对多尺度、多长宽比的车辆进行检测,假设需要处理scale个尺度、ratio种长宽比的车辆检测,本申请实施例可以利用 $k*5*scale*ratio$ 个卷积核进行卷积计算,最终输出 $5*scale*ratio$ 个图,每张输出的图可以代表某种尺度和比例的、某个属性(概率或者相对偏移位置的四个坐标)的输出。

[0054] 传统技术在具体实施时,通常根据原始监控图像提取一些基本特征,比如纹理,颜色,折线,各个方向的梯度等等,但是这些特征与最终的车辆检测任务并不直接相关。而本申请实施例充分借鉴了深度学习模型的高效特征表达能力。深度学习的优势在于,它可以根据任务本身,专门从数据和数据的标注中自动学习得到最紧凑有效的特征。本申请实施例可以利用深度卷积神经网络模型从所述监控图像中自动学习得到更加丰富、更加有表达和区分能力的特征,这些特征里包含了所有能够区分车与非车的信息,比如有没有车牌、车轮、车窗、车的角度或者车的各个部分的位置关系等等。

[0055] 实施中,所述卷积计算具体可以为全卷积计算。

[0056] 现有技术中的进行卷积计算,是将卷积核作为滑动窗口在所述特征图像的某个位置进行卷积计算,在计算结束后再移动到下一个位置进行卷积计算,...,直至完成所述特征图像的全部位置的卷积计算。

[0057] 而本申请实施例采用的可以是全卷积计算,即,所述卷积核作为滑动窗口一次性的与所述特征图像上所有位置进行卷积计算,显然,与现有技术相比,本申请实施例所提供的方案检测效率更高。

[0058] 实施中,所述预先训练得到的卷积核可以为多个,所述将预先训练得到的卷积核作为滑动窗口与所述特征图像进行卷积计算,具体可以为:

[0059] 将所述多个预先训练得到的卷积核作为滑动窗口并行与所述特征图像进行卷积计算。

[0060] 具体实施时,卷积核可以有多个,按照大小不同可以分为几类,例如:1*1、3*3或5*5等,将这些卷积核分别在所述特征图像上进行滑过即可完成卷积计算。

[0061] 本申请实施例可以根据实际车辆的特征预先训练出多个卷积核,根据不同场景、不同环境、不同距离、不同角度、不同长宽比的车辆样本中的车灯、车尾、车轱辘、车身、或者车门把手等,经过不断的训练得到可以用于检测车辆的卷积核。例如:根据不同场景、不同环境、不同距离、不同角度、不同长宽比的车辆样本中的车门把手训练得到卷积核,从而可以在监控图像中出现车门把手时,采用所述卷积核可以快速、准确的检测出图像中存在车门把手且输出车门把手的位置,从而确定图像中有车以及车的位置。

[0062] 本申请实施例采用多个卷积核进行计算,经汇总后可以得到更为准确的有车概率和车的位置,例如,有的卷积核可能是根据车门把手确定监控图像中有辆车、有的卷积核可能是根据车轱辘确定监控图像中有辆车的,当多个卷积核均检测出监控图像中存在一辆车,则检测结果的准确性更高。

[0063] 由于本申请实施例在检测计算过程中,可以多个卷积核同时与所述特征图像进行卷积计算,假设GPU有10个核,那么GPU的10个核可以均匀并行利用,进一步提高了检测效率,而现有技术则只能一个一个串行处理,无法拆分成计算量相同的子任务,因此也无法充分利用GPU多核特性。具体实施时,GPU通常可以有200~4000个核,本申请对GPU的核数不作限制。

[0064] 实施中,在卷积核进行卷积计算时如果同一辆车在不同位置产生响应,可以采用非最大值抑制方法将输出的所述车辆的多个相对位置进行合并。

[0065] 具体实施时,卷积核的大小可能小于所述车辆、也可能大于所述车辆,不同大小的卷积核可能会与同一辆车产生响应,即输出有车的概率大于预设阈值,此时,所述车辆可能被多个框所覆盖,本申请实施例可以采用非最大值抑制方法将这些框进行合并。

[0066] 除此之外,所述车辆也可能与对应于不同长宽比的卷积核产生响应,即输出有车的概率大于预设阈值,此时,所述车辆可能被多个框所覆盖,本申请实施例可以采用非最大值抑制方法将这些框进行合并。

[0067] 本申请实施例中,卷积核在扫描时,多次都对同一辆车产生较强响应时,可以对那些输出的车辆相对位置进行合并;或者说,同一辆车在卷积核扫描时,在不同位置产生了多次响应(所述响应可以为输出的有车概率大于预设阈值),那么可以把输出的车辆相对位置(对应于车辆的框)进行合并。

[0068] 非最大值抑制(Non-maximum suppression)可以看做是局部最大值的搜索问题,可以简单理解为抑制不是极大值的元素(具体实施时可以将非极大值的像素点所对应的灰度值置为0)、搜索局部的极大值。

[0069] 具体实施时,当任意两个框的交并比大于预设第二阈值时,则认为这两个框覆盖的是同一辆车,即可根据上述非最大值抑制方法将这两个框进行合并。

[0070] 本申请实施例中在检测时,如果同一辆车被多个框所覆盖,则可以采用非最大值抑制方法将这些框进行合并。

[0071] 实施中,所述方法可以进一步包括:

[0072] 根据所述车辆的相对位置和预先设置的相对参考框,计算所述车辆在所述监控图像中的绝对位置。

[0073] 本申请实施例中在得到有车的概率和车辆的相对位置后,可以利用多尺度、多长宽比的相对参考框来计算出车辆在原始监控图像中的绝对位置坐标。

[0074] 具体实施时,可以采用7个尺度的相对参考框,每种尺度可以有5种长宽比。例如:相对参考框的面积可以等比例(面积的比例为2,边长的比例为根号2)从50*50到400*400,即,50*50、71*71、100*100、141*141、200*200、282*282、400*400,长宽比可以为3:1,2:1,1:1、1:2、1:3,假设参考框面积为200*200、长宽比为2:1,那么这个参考框的大小则为282*141。

[0075] 假设车辆的相对位置(也就是相对于相对参考框)为 (lx, ly, rx, ry) ,即所述车辆的相对位置的左上角坐标为 (lx, ly) 、右下角坐标为 (rx, ry) ,利用面积为 a 、长宽比为 b 的相对参考框,进行如下计算:

[0076] 假设在图像的 (w, h) 位置,有一个相对参考框的大小为 $x*y$ 的,根据公式 $x = \sqrt{a*b}$, $y = \sqrt{a/b}$ 进行计算,得到相对参考框绝对位置为:

[0077] 左上角: $[w-x/2, h-y/2]$,右下角 $[w+x/2, h+y/2]$;

[0078] 计算得到车辆在原始监控图像中的绝对位置为

[0079] 左上角 $[w-x/2+lx, h-y/2+ly]$,右下角 $[w-x/2+rx, h-y/2+ry]$

[0080] 本申请实施例中所述长、宽并不一定代表车身的长、宽,具体可以是车辆显示在图片上的大小,例如:假设车辆的高度为0.5、长度为1.2、宽度为0.7,假设车辆正对镜头,那么检测到的长则为0.7、宽为0.5,而如果车头在左、车尾在右,那么检测到的长则为1.2、宽为0.5。

[0081] 实施中,所述方法可以进一步包括:

[0082] 利用校验模型对所述检测结果进行校验,所述校验模型是通过对正例和负例进行训练得到的。

[0083] 具体实施时,假设检测结果确定当前场景中有车,还可以进一步用校验模型去验证结果是否正确,以降低误差。所述校验模型是通过对正例和负例进行训练得到的。其中,正例和负例可以采用1:1的方式生成。

[0084] 具体实施时,所述校验模型同样可以采用卷积神经网络进行训练得到,所述校验模型的结构可以为:

[0085] 输入图像(假设为 $3*128*128$);

[0086] 经过卷积层进行卷积计算Conv(7, 2, 32),卷积核大小为 $7*7$,滑动时每次移动2个

像素,总共32个特征图像;

[0087] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;

[0088] 经过卷积层进行卷积计算Conv(3,1,64),卷积核大小为3*3,滑动时每次移动1个像素,总共64个特征图像;

[0089] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;

[0090] 经过卷积层进行卷积计算Conv(3,1,128),卷积核大小为3*3,滑动时每次移动1个像素,总共128个特征图像;

[0091] 经过卷积层进行卷积计算Conv(3,1,128),卷积核大小为3*3,滑动时每次移动1个像素,总共128个特征图像;

[0092] 经过卷积层进行卷积计算Conv(3,1,128),卷积核大小为3*3,滑动时每次移动1个像素,总共128个特征图像;

[0093] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;

[0094] 经过全连接层FC(512),全连接层的节点个数为512个;

[0095] 经过全连接层FC(512),全连接层的节点个数为512个;

[0096] 最终输出两个数字输出(2),一个是有车的概率,一个是没车的概率。

[0097] 其中,Conv(a,b,c)表示卷积核大小为a*a,滑动时每次移动b个像素,总共有c个特征图像;Pooling(a,b)表示采用最大池化(max pooling),范围大小是a*a,每次移动b个像素;FC(a)表示这是一个全连接层,节点个数为a个。

[0098] 校验模型的训练可以与之前训练相对参考框的训练方式相同,可以利用反向传播(BP,backpropagation)算法调整参数,使得在测试集上输出的结果与真实的标注尽可能接近。

[0099] 本申请实施例可以预先采集较多真实车辆样本,正例可以为和某个真实车辆外接矩形框的交并比超过预设第一阈值的所有样本,负例则可以是和图像上所有车辆外接矩形框的交并比均小于所述预设第一阈值、且至少和某个外接矩形框的交并比大于预设第二阈值的所有样本。

[0100] 其中,所述交并比可以指两个框的交集的面积除以两个框的并集的面积所得到的数值,其中一个框可以为真实的标注框,即所述车辆的外接矩形框;另一个框可以是对所述外接矩形框进行一定的扰动所得到的框。具体实施中,当两个框的交并比大于预设第二阈值(比如0.5)时,则可以经过扰动的框认为是正例,否则可以认为是负例。

[0101] 假设预设阈值为0.5,本申请实施例中正例可以是和图像上车辆外接矩形框的交并比大于0.5的所有样本,负例可以是和这张图上所有车辆外接矩形框交并比都在0.5以下、且至少和某个外接矩形框的交并比在0.1以上。

[0102] 具体实施时,由于负例的比例太多,本申请实施例中,为了取得有质量的负例,将负例限定为交并比小于预设第一阈值且至少存在一个交并比大于第二阈值。

[0103] 采用上述方式可能使得纯背景中存在一些剔除不掉的负例,本申请实施例可以加入一些与图像中任意的框都不相交且容易被当成误检测的位置作为负例,从而进一步提高模型的判别能力。

[0104] 实施中,所述卷积核的训练过程具体可以为:

[0105] 获取若干车辆样本图像,所述样本图像中每辆车标注有外接矩形框;

[0106] 利用深度卷积神经网络模型提取所述样本图像的特征,输出特征图像;

[0107] 根据卷积核与所述特征图像检测所述样本图像中的车辆,根据检测结果以及所述外接矩形框调整所述卷积核内的参数,直至所述检测结果中车辆的位置与所述外接矩形框接近或重合。

[0108] 具体实施时,获取若干车辆样本图像后,可以为每张图像上的车辆标注一个外接矩形框,作为后续模型检测的标准。

[0109] 利用深度卷积神经网络模型提取所述样本图像的特征,输出特征图像后,可以利用预先设置好初始参数的卷积核与所述特征图像进行计算来检测所述样本图像中是否有车辆以及车辆的位置等,然后将检测结果与实际标注的外接矩形框进行比较,如果不一致则调整卷积核中的数值、重新计算,直至检测结果不断趋近于实际的车辆位置,最终训练得到所述卷积核,确定所述卷积核可以正确检测出是否有车辆以及车辆的位置。

[0110] 本申请实施例所提出的基于全卷积深度神经网络的端到端的车辆检测方法,具有如下优点:

[0111] 1)无需对图像进行任何预处理;

[0112] 2)可以接受任意大小、任意长宽比的图像输入;

[0113] 3)可以同时处理白天、黑夜、车头、车尾等情形;

[0114] 4)不仅能够检测车辆,还可以很容易的扩展到对三轮车、非机动车、行人等目标的检测;

[0115] 5)可以准确的计算车辆轮廓,精度可以达到1个像素,而传统方法通常只能支持固定比例的检测框;

[0116] 6)可以同时CPU和GPU上运行,检测效率高。

[0117] 为了便于本申请的实施,下面以实例进行说明。

[0118] 本申请实施例中可以包括数据收集和标注、网络模型设计、离线训练检测器、在线检测这四个步骤。

[0119] 一、数据收集和标注

[0120] 为了实现不同场景(不同角度的卡口或者电子警察场景等)、不同环境(白天、黑夜、雨雪天等)、不同距离(远、近)、不同角度(车头、车尾、侧面)、不同长宽比的车辆进行统一检测,首先可以收集覆盖上述情形的大量真实数据,并进行准确标注。

[0121] 标注形式可以为给每一辆车标注一个外接矩形框,最终的标注结果可以为一张监控图像中有很多矩形框,每个框可以和每辆车一一对应。如果车辆之间有遮挡,则对应的矩形框之间可能会重叠。

[0122] 二、网络模型设计

[0123] 假设输入的图片大小为M*N(这里以600*1000为例),本申请实施例的网络模型的前半部分可以采用卷积神经网络经典模型Google Net(该模型由卷积层和Inception Module堆叠而成),该部分的最终输出大小为18x31(这个输出就是通过神经网络学习得到的用于检测的特征),然后通过卷积的方式模拟滑动窗口。

[0124] 三、离线训练检测器

[0125] 1、设置模型训练相关参数

[0126] 学习率:0.001;

- [0127] 批次大小(mini-batch size):100,其中包含50个正例、50个负例;
- [0128] 翻转操作(flip):对每张图进行左右翻转,从而扩大训练集;
- [0129] 动量(momentum):0.9;
- [0130] 权值衰减率(weight decay):0.0005。
- [0131] 2、正负样本选择机制
- [0132] 检测任务的本质可以为2分类的任务,检测结果为有车辆或没有车辆,本申请实施例可以生成训练用的正例和负例,正例和负例采用1:1的方式生成。
- [0133] 其中,正例可以指和某个真实的车辆框的交并比(两个框的交的面积除以两个框的并的面积)超过0.5的所有样本,而负例可以指和这张图上所有车辆交并比都在0.5以下,而且至少和某一个框交并比在0.1以上。
- [0134] 四、在线检测
- [0135] 图2示出了本申请实施例中车辆检测过程示意图,如图所示,所述车辆检测过程可以包括如下步骤:
- [0136] 步骤1、利用RGB摄像头获取监控场景的视频图像(假设为600*1000)或视频(每帧为视频图像);
- [0137] 步骤2、采用深度卷积神经网络模型对所述视频图像提取特征,输出特征图像(18*31);
- [0138] 具体实施可以为,输入所述视频图像(600*1000);
- [0139] 经过卷积层进行卷积计算Conv1(7,2),卷积核大小为7*7,滑动时每次移动2个像素;
- [0140] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;
- [0141] 经过卷积层进行卷积计算Conv2(5,2),卷积核大小为5*5,滑动时每次移动2个像素;
- [0142] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;
- [0143] 利用googlenet中的感知模块Inception3进行相关性聚集;
- [0144] 经过池化层采用最大池化Pooling(3,2),池化范围大小为3*3,每次移动2个像素;
- [0145] 利用googlenet中的感知模块Inception4进行相关性聚集;
- [0146] 最终得到特征图像。
- [0147] 每次经过一个卷积层Conv或者池化层Pooling,图像的大小就大概会缩小一倍。
- [0148] 所述特征图像可以理解为所述视频图像的特征浓缩后的结果。
- [0149] 步骤3、利用预先训练得到的卷积核(假设为3*3,对应到原视频图像中卷积核的大小则为48*48)模拟滑动窗口,与所述特征图像进行卷积计算,得到5个参数;
- [0150] 其中,第一个参数代表有车的概率(假设为80%),第2~5个参数代表车辆的位置(假设为A、B、C、D)。
- [0151] 最终,卷积核的滑动位置会停留在图像中检测有车的位置,第2~5个参数分别为卷积核的左上坐标和右下坐标,经计算得到卷积核的最终位置停留在左上坐标为(A,B)、右下坐标为(C,D)的位置,在左上坐标为(A,B)、右下坐标为(C,D)的卷积核位置则为所述车辆的位置。
- [0152] 具体检测时,可能最终输出很多的候选框,且同一辆车可能会被多个框覆盖,此

时,本申请实施例可以进行框的融合。具体实施时,可以采用非最大值抑制策略,把对应于同一辆车的多个框合并成一个框,得到该车辆最可能出现的位置。

[0153] 步骤4、利用相对参考框计算所述车辆在原始图像中的位置,从近似的卷积核位置回归出准确的车辆位置;

[0154] 本申请实施例所提供的车辆检测方法,可以利用端到端的深度全卷积神经网络进行车辆的识别,只利用单个模型即可广泛适用于各种情形,且能够自动处理多尺度、多角度以及复杂环境等情形。

[0155] 基于同一发明构思,本申请实施例中还提供了一种车辆检测装置,由于这些设备解决问题的原理与一种车辆检测方法相似,因此这些设备的实施可以参见方法的实施,重复之处不再赘述。

[0156] 图3示出了本申请实施例中车辆检测装置的结构示意图,如图所示,所述车辆检测装置可以包括:

[0157] 获取模块,用于获取当前场景中的监控图像; f

[0158] 特征提取模块,用于利用深度卷积神经网络模型对所述监控图像提取特征,输出特征图像;

[0159] 训练模块,用于训练卷积核;

[0160] 检测模块,用于将所述训练得到的卷积核作为滑动窗口与所述特征图像进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置。

[0161] 实施中,所述特征提取模块具体可以用于利用深度卷积神经网络模型对所述监控图像提取不同的特征,输出 k 个特征图像;所述检测模块具体可以用于将预先训练得到的 $5*k*scale*ratio$ 个卷积核与所述 k 个特征图像分别进行卷积计算,得到检测结果,所述检测结果包括有车的概率和车辆的相对位置;其中,所述 $scale$ 为车辆尺度的种类,所述 $ratio$ 为车辆长宽比的种类。

[0162] 实施中,所述卷积计算具体可以为全卷积计算。

[0163] 实施中,所述检测模块具体可以用于将所述训练得到的多个卷积核作为滑动窗口并行与所述特征图像进行卷积计算。

[0164] 实施中,所述检测模块可以进一步用于在进行卷积计算时如果同一辆车在不同位置产生响应,采用非最大值抑制方法将输出的所述车辆的多个相对位置进行合并。

[0165] 实施中,所述装置可以进一步包括:

[0166] 计算模块,用于根据所述车辆的相对位置和预先设置的相对参考框,计算所述车辆在所述监控图像中的绝对位置。

[0167] 实施中,所述装置可以进一步包括:

[0168] 校验模块,用于利用校验模型对所述检测结果进行校验,所述校验模型是通过正例和负例进行训练得到的。

[0169] 实施中,所述训练模块具体包括:

[0170] 获取单元,用于预先获取若干车辆样本图像,所述样本图像中每辆车标注有外接矩形框;

[0171] 特征提取单元,用于利用深度卷积神经网络模型提取所述样本图像的特征,输出特征图像;

[0172] 训练单元,用于根据卷积核与所述特征图像检测所述样本图像中的车辆,根据检测结果以及所述外接矩形框调整所述卷积核内的参数,直至所述检测结果中车辆的位置与所述外接矩形框接近或重合。

[0173] 为了描述的方便,以上所述装置的各部分以功能分为各种模块或单元分别描述。当然,在实施本申请时可以把各模块或单元的功能在同一个或多个软件或硬件中实现。

[0174] 本申请实施例中采用全卷积神经网络,可以通过一次前向计算得到所有滑动框的结果,而且基于选定的不同大小和比例的相对参考框,能够在不带来额外运算的情形下,同时预测多种尺度和长宽比的车辆。深度网络在自动提取任务相关的有效特征方面相比传统方法有很大的优势,能够得到更加精确的检测结果。而且,深度全卷积神经网络内部结合了多尺度卷积感知模块(Inception Module),且由于模型较深,感受野(Receptive Field)较大,预测车辆时,能自然的利用到车辆周围的背景信息,从而进一步提升车辆的检测效果。

[0175] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0176] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0177] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0178] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0179] 尽管已描述了本申请的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例作出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

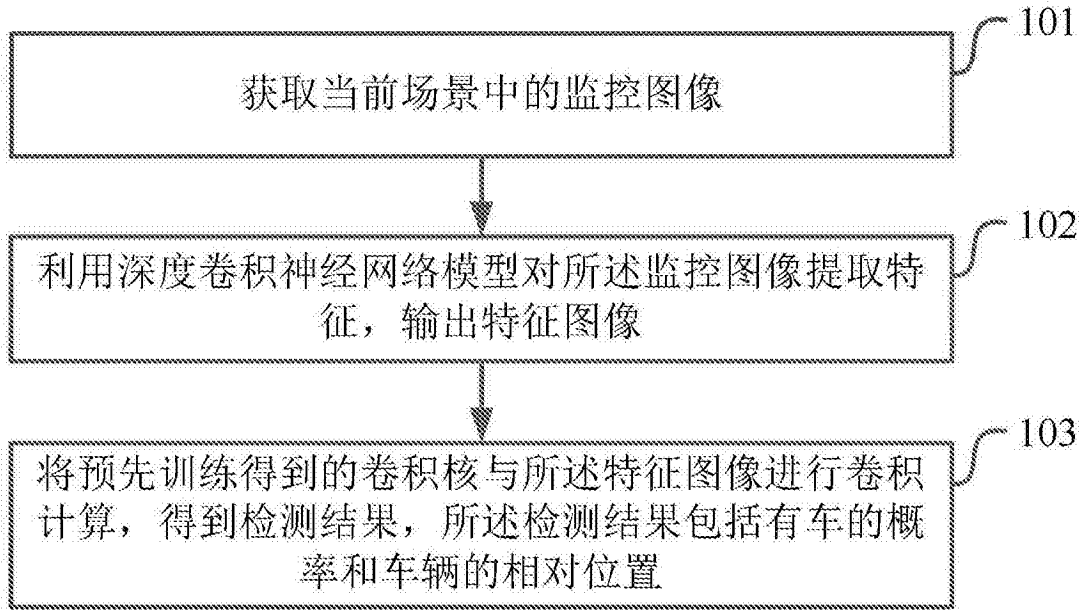


图1

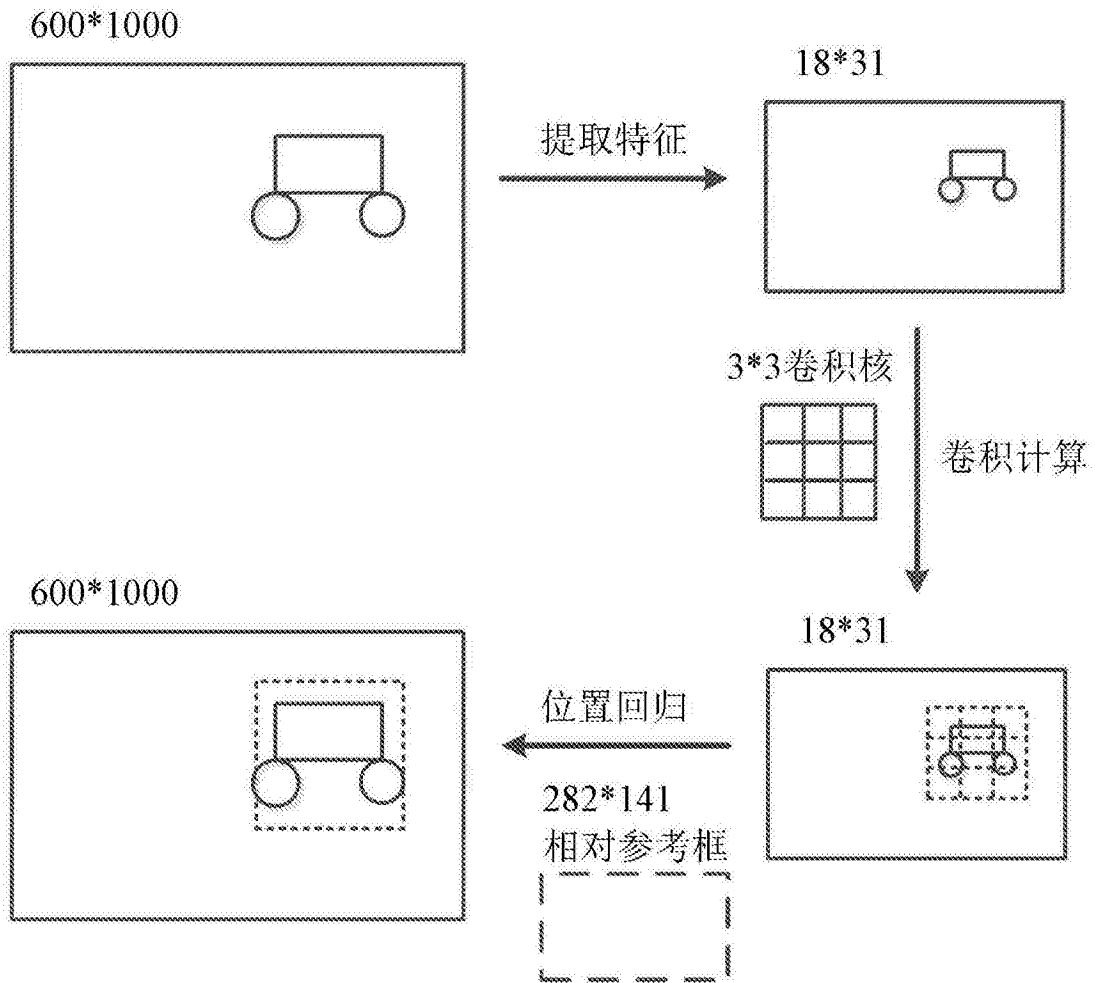


图2

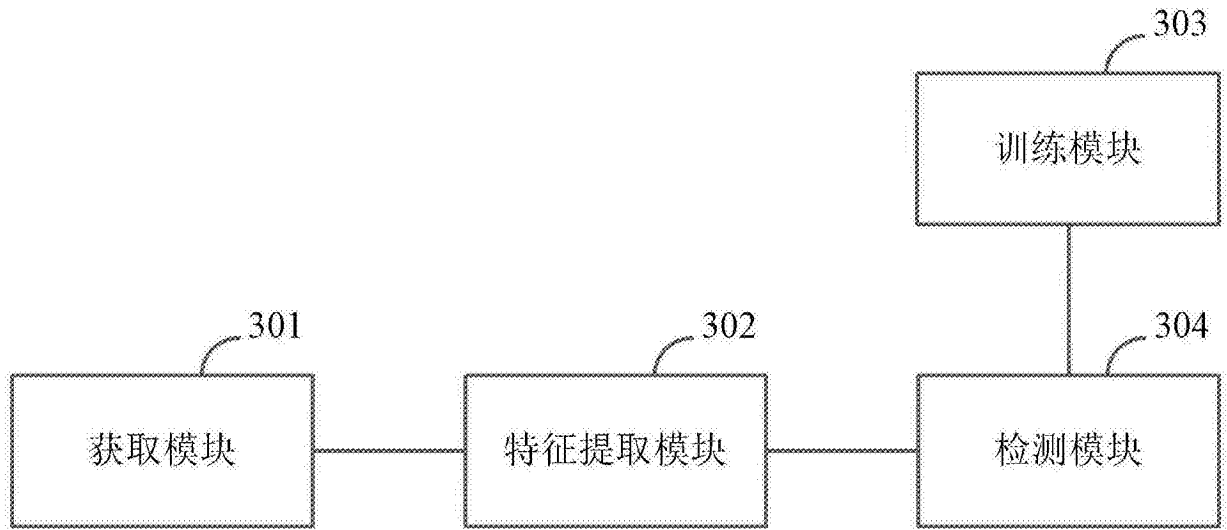


图3