



(12)发明专利申请

(10)申请公布号 CN 108875902 A

(43)申请公布日 2018. 11. 23

(21)申请号 201711262814.2

(22)申请日 2017.12.04

(71)申请人 北京旷视科技有限公司

地址 100190 北京市海淀区科学院南路2号
A座313

申请人 北京迈格威科技有限公司

(72)发明人 王曼晨

(74)专利代理机构 北京市磐华律师事务所

11336

代理人 高伟 张建

(51)Int.Cl.

G06N 3/04(2006.01)

G06K 9/62(2006.01)

G06K 9/32(2006.01)

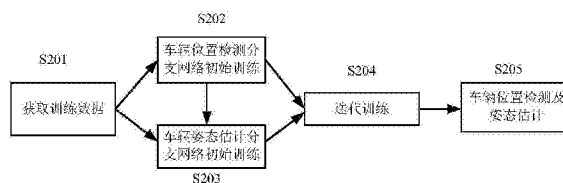
权利要求书2页 说明书15页 附图5页

(54)发明名称

神经网络训练方法及装置、车辆检测估计方法及装置、存储介质

(57)摘要

本发明公开了一种神经网络训练方法及装置、车辆检测估计方法及装置、存储介质,用于车辆位置检测和姿态估计,该神经网络训练方法包括:获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;重复执行以下步骤直到满足设定条件为止:利用所述训练数据训练所述车辆位置检测分支网络,利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。该神经网络训练方法可以实现端到端训练,并且无需较高数据标注精度,并且可以提高网络的准确率。该车神经网络训练装置、车辆检测估计方法及装置、存储介质具有类似的优点。



1. 一种神经网络的训练方法,其特征在于,所述神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,所述训练方法包括:

获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;

重复执行以下步骤直到满足设定条件为止:

利用所述训练数据训练所述车辆位置检测分支网络,

利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

2. 根据权利要求1所述的训练方法,其特征在于,所述神经网络还包括深层神经网络、区域建议网络和感兴趣区域池化层,所述训练所述车辆位置检测分支网络包括:

通过所述深层神经网络对所述训练数据进行处理以得到训练用特征图;

通过所述区域建议网络对所述训练用特征图进行处理以得到训练用候选框;

将所述训练用特征图和所述训练用候选框输入所述感兴趣区域池化层做池化处理,以得到所述训练用候选框对应的训练用第一特征向量;

将所述训练用第一特征向量输入所述车辆位置检测分支网络,以得到所述训练用车辆检测框。

3. 根据权利要求2所述的训练方法,其特征在于,所述训练所述车辆姿态估计分支网络包括:

将所述训练用车辆检测框输入所述感兴趣区域池化层做池化处理,以得到所述训练用车辆检测框对应的训练用第二特征向量;

将所述训练用第二特征向量输入所述车辆姿态估计分支网络,以得到车辆姿态估计信息。

4. 一种神经网络的训练装置,所述神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,其特征在于,所述训练装置包括:

数据获取单元,用于获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;

第一训练单元,用于利用所述训练数据训练所述车辆位置检测分支网络;

第二训练单元,用于利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

5. 一种神经网络的训练装置,其特征在于,所述训练装置包括存储器和处理器,所述存储器上存储有由所述处理器运行的计算机程序,所述计算机程序在被所述处理器运行时执行如权利要求1-3中的任一项所述的神经网络的训练方法。

6. 一种存储介质,其特征在于,所述存储介质上存储有计算机程序,所述计算机程序在运行时执行如权利要求1-3中的任一项所述的神经网络的训练方法。

7. 一种用于车辆位置检测和姿态估计的方法,其特征在于,包括:

获取待检测图像;

利用预先训练的用于车辆位置检测和姿态估计的神经网络对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息;

其中,所述预先训练的用于车辆位置检测和姿态估计的神经网络包括车辆位置检测分

支网络和车辆姿态估计分支网络，

所述车辆位置检测分支网络用于基于待测图像输出车辆检测框；

所述车辆姿态估计分支网络用于基于所述待测图像和所述车辆检测框输出车辆姿态估计信息。

8. 根据权利要求7所述的方法，其特征在于，所述预先训练的用于车辆位置检测和姿态估计的神经网络还包括深层神经网络、区域建议网络和感兴趣区域池化层，所述方法还包括：

通过所述深层神经网络对所述待测图像进行处理以得到特征图；

通过所述区域建议网络对所述特征图进行处理以得到候选框；

将所述候选框和所述特征图输入所述感兴趣区域池化层做池化处理，以得到所述候选框对应的第一特征向量。

9. 根据权利要求8所述的方法，其特征在于，还包括：

通过所述位置检测分支网络基于所述第一特征向量输出所述车辆检测框和置信度；

将所述车辆检测框输入所述感兴趣区域池化层做池化处理，以得到所述车辆检测框对应的第二特征向量；

通过所述车辆姿态估计分支网络基于所述第二特征向量输出所述车辆姿态估计信息。

10. 根据权利要求7-9任意一项所述的方法，其特征在于，所述车辆姿态估计信息包括车辆三维包围框顶点投影坐标或车辆三维包围框顶点投影坐标相对所述车辆检测框中心点坐标的位置。

11. 根据权利要求9所述的方法，其特征在于，还包括：根据所述车辆检测框和置信度基于用户设置或选择获得车辆位置检测结果。

12. 根据权利要求11所述的方法，其特征在于，还包括：根据所述车辆位置检测结果和所述车辆姿态估计信息获得车辆姿态估计结果。

13. 一种用于车辆位置检测和姿态估计的装置，其特征在于，包括：

图像获取模块，用于获取待检测图像；

车辆位置检测和姿态估计模块，用于对所述待检测图像进行处理，以获取车辆检测框和车辆姿态估计信息；

其中，所述车辆位置检测和姿态估计模块包括车辆位置检测分支网络和车辆姿态估计分支网络，

所述车辆位置检测分支网络用于基于待测图像输出车辆检测框；

所述车辆姿态估计分支网络用于基于所述待测图像和所述车辆检测框输出车辆姿态估计信息。

14. 一种用于车辆位置检测和姿态估计的装置，其特征在于，所述装置包括存储器和处理器，所述存储器上存储有由所述处理器运行的计算机程序，所述计算机程序在被所述处理器运行时执行如权利要求7-12中的任一项所述的用于车辆位置检测和姿态估计的方法。

15. 一种存储介质，其特征在于，所述存储介质上存储有计算机程序，所述计算机程序在运行时执行如权利要求7-12中的任一项所述的用于车辆位置检测和姿态估计的方法。

神经网络训练方法及装置、车辆检测估计方法及装置、存储介质

技术领域

[0001] 本发明涉及图像检测技术领域,尤其涉及一种神经网络的训练方法及装置、用于车辆位置检测和姿态估计的方法及装置,以及存储介质。

背景技术

[0002] 在过去几年中,随着科技的发展,借助智能系统实现汽车的自动驾驶在学术界与产业界都变得越来越炙手可热。而在自动驾驶中,最基本的功能之一就是检测车辆、行人、障碍物等等,这就转换为了物体检测问题。图像中的物体检测问题一直是计算机视觉领域的热门话题,深度学习时代的到来给检测精度带来了非常大的提升。深度学习在2D物体检测中已取得一些非常重要的成果,如RCNN (Regions with Convolutional Neural Network Features具有卷积神经网络特征的区域)、FastRCNN、FasterRCNN等工作。在自动驾驶中,尽管物体检测针对场景中内容提供了重要信息,但二维检测框对于三维真实世界场景描述是依然不足的。对于一辆自动驾驶的汽车来说,它必须能够根据从图片场景中提取的信息理解当前的交通状况并预测出下一秒可能的道路情况。为了获取周围车辆的速度、方向等信息,需要对不同时刻的车辆姿态进行准确预测。另外,为了更好的理解道路交通状况,如何有效描述周围的车辆非常重要。

发明内容

[0003] 在发明内容部分中引入了一系列简化形式的概念,这将在具体实施例部分中进一步详细说明。本发明的发明内容部分并不意味着要试图限定出所要求保护的技术方案的关键特征和必要技术特征,更不意味着试图确定所要求保护的技术方案的保护范围。

[0004] 为了解决上述问题,本发明提供了一种神经网络的训练方法,所述神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,所述训练方法包括:

[0005] 获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;

[0006] 重复执行以下步骤直到满足设定条件为止:

[0007] 利用所述训练数据训练所述车辆位置检测分支网络,

[0008] 利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

[0009] 在本发明的一个实施例中,所述神经网络还包括深层神经网络、区域建议网络和感兴趣区域池化层,所述训练所述车辆位置检测分支网络包括:

[0010] 通过所述深层神经网络对所述训练数据进行处理以得到训练用特征图;

[0011] 通过所述区域建议网络对所述训练用特征图进行处理以得到训练用候选框;

[0012] 将所述训练用特征图和所述训练用候选框输入所述感兴趣区域池化层做池化处理,以得到所述训练用候选框对应的训练用第一特征向量;

[0013] 将所述训练用第一特征向量输入所述车辆位置检测分支网络,以得到所述训练用车辆检测框。

[0014] 在本发明的一个实施例中,所述训练所述车辆姿态估计分支网络包括:

[0015] 将所述训练用车辆检测框输入所述感兴趣区域池化层做池化处理,以得到所述训练用车辆检测框对应的训练用第二特征向量;

[0016] 将所述训练用第二特征向量输入所述车辆姿态估计分支网络,以得到车辆姿态估计信息。

[0017] 根据本发明的另一个方面,还提供了一种神经网络的训练装置,所述神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,所述训练装置包括:

[0018] 数据获取单元,用于获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;

[0019] 第一训练单元,用于利用所述训练数据训练所述车辆位置检测分支网络;

[0020] 第二训练单元,用于利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

[0021] 在本发明的一个实施例中,所述神经网络还包括深层神经网络、区域建议网络和感兴趣区域池化层,所述第一训练单元执行以下步骤:

[0022] 通过所述深层神经网络对所述训练数据进行处理以得到训练用特征图;

[0023] 通过所述区域建议网络对所述训练用特征图进行处理以得到训练用候选框;

[0024] 将所述训练用特征图和所述训练用候选框输入所述感兴趣区域池化层做池化处理,以得到所述训练用候选框对应的训练用第一特征向量;

[0025] 将所述训练用第一特征向量输入所述位置检测分支网络,以得到所述训练用车辆检测框。

[0026] 在本发明的一个实施例中,所述第二训练单元执行以下步骤:

[0027] 将所述训练用车辆检测框输入所述感兴趣区域池化层做池化处理,以得到所述训练用车辆检测框对应的训练用第二特征向量;

[0028] 将所述训练用第二特征向量输入所述车辆姿态估计分支网络,以得到车辆姿态估计信息。

[0029] 根据本发明的神经网络的训练方法和训练装置,一方面训练数据只依靠标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片,不需要额外更精细的标注,对训练数据标注的要求低,降低了训练难度;另一方面,同时训练车辆位置检测分支网络和车辆姿态估计分支网络,基于训练数据可以同时输出车辆的二维检测框和车辆的姿态估计信息,不需要人工预处理和后续处理。

[0030] 进一步地,根据本发明的神经网络的训练方法和装置,在训练车辆姿态估计分支网络时通过车辆位置检测分支网络输出的训练用车辆检测框对应的特征向量辅助预测车辆姿态信息,提高了车辆姿态估计的准确率;并且车辆姿态估计任务的加入也提高了车辆位置检测的准确度,两种任务互相辅助,互相促进。

[0031] 根据本发明的另一个方面,还提供了一种用于车辆位置检测和姿态估计的神经网络的方法,包括:

[0032] 获取待检测图像;

[0033] 利用预先训练的用于车辆位置检测和姿态估计的神经网络对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息;

[0034] 其中,所述预先训练的用于车辆位置检测和姿态估计的神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,

[0035] 所述车辆位置检测分支网络用于基于待测图像输出车辆检测框;

[0036] 所述车辆姿态估计分支网络用于基于所述待测图像和所述车辆检测框输出车辆姿态估计信息。

[0037] 在本发明的一个实施例中,所述预先训练的用于车辆位置检测和姿态估计的神经网络还包括深层神经网络、区域建议网络和感兴趣区域池化层,所述方法还包括:

[0038] 通过所述深层神经网络对所述待测图像进行处理以得到特征图;

[0039] 通过所述区域建议网络对所述特征图进行处理以得到候选框;

[0040] 将所述候选框和所述特征图输入所述感兴趣区域池化层做池化处理,以得到所述候选框对应的第一特征向量。

[0041] 在本发明的一个实施例中,还包括:

[0042] 通过所述位置检测分支网络基于所述第一特征向量输出所述车辆检测框和置信度;

[0043] 将所述车辆检测框输入所述感兴趣区域池化层做池化处理,以得到所述车辆检测框对应的第二特征向量;

[0044] 通过所述车辆姿态估计分支网络基于所述第二特征向量输出所述车辆姿态估计信息。

[0045] 在本发明的一个实施例中,所述车辆姿态估计信息包括车辆三维包围框顶点投影坐标或车辆三维包围框顶点投影坐标相对所述车辆检测框中心点坐标的位置。

[0046] 在本发明的一个实施例中,还包括:根据所述车辆检测框和置信度基于用户设置或选择获得所述车辆位置检测结果。

[0047] 在本发明的一个实施例中,还包括:根据所述车辆位置检测结果和所述车辆姿态估计信息获得所述车辆姿态估计结果。

[0048] 根据本发明的另一个方面,还提供了一种用于车辆位置检测和姿态估计的神经网络的装置,包括:

[0049] 图像获取模块,用于获取待检测图像;

[0050] 车辆位置检测和姿态估计模块,用于对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息;

[0051] 其中,所述车辆位置检测和姿态估计模块包括车辆位置检测分支网络和车辆姿态估计分支网络,

[0052] 所述车辆位置检测分支网络用于基于待测图像输出车辆检测框;

[0053] 所述车辆姿态估计分支网络用于基于所述待测图像和所述车辆检测框输出车辆姿态估计信息。

[0054] 在本发明的一个实施例中,所述车辆位置检测和姿态估计模块还包括:

[0055] 深层神经网络,用于对所述待测图像进行处理以得到特征图;

[0056] 区域建议网络,用于对所述特征图进行处理以得到候选框;

[0057] 感兴趣区域池化层,用于对所述候选框和所述特征图做池化处理,以得到所述候选框对应的第一特征向量。

[0058] 在本发明的一个实施例中,所述位置检测分支网络基于所述第一特征向量输出所述车辆检测框和置信度;

[0059] 所述感兴趣区域池化层还用于对所述车辆检测框和所述特征图做池化处理,以得到所述车辆检测框对应的第二特征向量;

[0060] 所述车辆姿态估计分支网络基于所述第二特征向量输出所述车辆姿态估计信息。

[0061] 在本发明的一个实施例中,所述车辆姿态估计信息包括车辆三维包围框顶点投影坐标或车辆三维包围框顶点投影坐标相对所述车辆检测框中心点坐标的位置。

[0062] 在本发明的一个实施例中,所述车辆位置检测和姿态估计模块还包括:车辆位置检测结果获取子模块,用于根据所述车辆检测框和置信度基于用户设置或选择获得所述车辆位置检测结果。

[0063] 在本发明的一个实施例中,所述车辆位置检测和姿态估计模块还包括:车辆姿态估计结果获取子模块,用于根据所述车辆位置检测结果和所述车辆姿态估计信息获得所述车辆姿态估计结果。

[0064] 根据本发明的用于车辆位置检测和姿态估计的方法和装置,可以对输入的待检测图像同时进行车辆位置检测和车辆姿态估计,并且由于采用了经过上述训练方法或装置训练的用于车辆位置检测和姿态估计的神经网络,具有更高的准确度。

[0065] 进一步地,根据本发明的用于车辆位置检测和姿态估计的方法和装置在进行车辆姿态估计时,将车辆姿态估计问题转化为了预测车辆的三维包围框(立方体)的8个三维顶点在图片上的投影(即8个二维坐标)问题,并且优选地在实际任务中预测的是这8个二维坐标相对于车辆位置检测得到的车辆检测框中心点的相对位置,进一步提高了预测的准确率。

[0066] 根据本发明的另一个方面,还提供了一种神经网络的训练装置,所述训练装置包括存储器和处理器,所述存储器上存储有由所述处理器运行的计算机程序,所述计算机程序在被所述处理器运行时执行如上所述的神经网络的训练方法。

[0067] 根据本发明的另一个方面,还提供了一种存储介质,所述存储介质上存储有计算机程序,所述计算机程序在运行时执行如上所述的的神经网络的训练方法。

[0068] 根据本发明的另一个方面,还提供了一种用于车辆位置检测和姿态估计的神经网络的装置,所述装置包括存储器和处理器,所述存储器上存储有由所述处理器运行的计算机程序,所述计算机程序在被所述处理器运行时执行如上所述的用于车辆位置检测和姿态估计的方法。

[0069] 根据本发明的另一个方面,还提供了一种存储介质,所述存储介质上存储有计算机程序,所述计算机程序在运行时执行如上所述的用于车辆位置检测和姿态估计的方法。

附图说明

[0070] 本发明实施例的下列附图在此作为本发明的一部分用于理解本发明。附图中示出了本发明的实施例及其描述,用来解释本发明的原理。在附图中,

[0071] 图1是用于实现根据本发明实施例的神经网络的训练方法及装置以及用于车辆位

置检测和姿态估计的方法及装置的示例电子设备的示意性框图；

[0072] 图2为根据本发明实施例的神经网络的训练方法中初始训练的示意图；

[0073] 图3为根据本发明实施例的神经网络的训练方法中车辆位置检测分支网络训练的示意图；

[0074] 图4为根据本发明实施例的神经网络的训练方法中车辆姿态估计分支网络训练的示意图；

[0075] 图5为根据本发明实施例的神经网络的训练方法中迭代训练的示意图；

[0076] 图6为根据本发明实施例的神经网络的训练装置的示意性结构框图；

[0077] 图7为根据本发明实施例的用于车辆位置检测和姿态估计的方法的示意图；

[0078] 图8为根据本发明实施例的用于车辆位置检测和姿态估计的方法详细示意图。

[0079] 图9为根据本发明实施例的用于车辆位置检测和姿态估计的装置的结构示意图；

[0080] 图10为图9中车辆位置检测和姿态估计模块的详细结构示意图。

具体实施例

[0081] 在下文的描述中,给出了大量具体的细节以便提供对本发明更为彻底的理解。然而,对于本领域技术人员来说显而易见的是,本发明实施例可以无需一个或多个这些细节而得以实施。在其他的例子中,为了避免与本发明实施例发生混淆,对于本领域公知的一些技术特征未进行描述。

[0082] 应当理解的是,本发明能够以不同形式实施,而不应当解释为局限于这里提出的实施例。相反地,提供这些实施例将使公开彻底和完全,并且将本发明的范围完全地传递给本领域技术人员。在附图中,为了清楚,部件、元件等的尺寸以及相对尺寸可能被夸大。自始至终相同附图标记表示相同的元件。

[0083] 为了使得本发明的目的、技术方案和优点更为明显,下面将参照附图详细描述根据本发明的示例实施例。显然,所描述的实施例仅仅是本发明的一部分实施例,而不是本发明的全部实施例,应理解,本发明不受这里描述的示例实施例的限制。基于本发明中描述的本发明实施例,本领域技术人员在没有付出创造性劳动的情况下所得到的所有其它实施例都应落入本发明的保护范围之内。

[0084] 首先,图1是用于实现根据本发明实施例的神经网络的训练方法及装置以及用于车辆位置检测和姿态估计的方法及装置的示例电子设备的示意性框图100。如图1所示,电子设备100包括一个或多个处理器102、一个或多个存储装置104、输入/输出装置106、通信接口108以及一个或多个图像采集装置110,这些组件通过总线系统112和/或其它形式的连接机构(未示出)互连。应当注意,图1所示的电子设备100的组件和结构只是示例性的,而非限制性的,根据需要,所述电子设备也可以具有其他组件和结构,也可以不包括前述的部分组件,例如可以不包括图像采集装置。

[0085] 所述处理器102一般表示任何类型或形式的能够处理数据或解释和执行指令的处理单元。一般而言,处理器可以是中央处理单元(CPU)或者具有数据处理能力和/或指令执行能力的其它形式的处理单元,并且可以控制所述电子设备100中的其它组件以执行期望的功能。此外,处理器也可以是神经网络处理器、图形处理器(GPU)、现场可编程门阵列(FPGA)或数字信号处理器(DSP),或者它们的一种或多种的组合。在特定实施例中,处理器

102可以接收来自软件应用或模块的指令。这些指令可以导致处理器102完成本文描述和/或示出的一个或多个示例实施例的功能。

[0086] 所述存储装置104可以包括一个或多个计算机程序产品,所述计算机程序产品可以包括各种形式的计算机可读存储介质,例如易失性存储器和/或非易失性存储器。所述易失性存储器例如可以包括随机存取存储器(RAM)和/或高速缓冲存储器(cache)等。所述非易失性存储器例如可以包括只读存储器(ROM)、硬盘、闪存等。在所述计算机可读存储介质上可以存储一个或多个计算机程序指令,处理器102可以运行所述程序指令,以实现下文所述的本发明实施例中(由处理器实现)的客户端功能以及/或者其它期望的功能。在所述计算机可读存储介质中还可以存储各种应用程序和各种数据,例如所述应用程序使用和/或产生的各种数据等。

[0087] 所述输入/输出装置106可以是用户用来输入指令和向外部输出各种信息的装置,例如输入装置可以包括按钮、键盘、鼠标、麦克风和触摸屏等中的一个或多个。输出装置可以包括显示器、扬声器、投影仪、网卡等中的一个或多个。

[0088] 通信接口108广泛地表示任何类型或形式的能够促进示例电子设备100和一个或多个附加设备之间的通信的设备或适配器。例如,通信接口108可以促进电子设备100和前端或附件电子设备以及后端服务器或云端的通信。通信接口108的示例包括但不限于有线网络接口(诸如网络接口卡)、无线网络接口(诸如无线网络接口卡)、调制解调器和任何其他合适的接口。在一实施例中,通信接口108通过与诸如因特网的网络的直连提供到远程服务器/远程前端设备的直连。在特定实施例中,通信接口108通过与专用网络,例如视频监控网络、天网系统网络等网络的直连提供到远程服务器/远程前端设备的直连。通信接口108还可以间接提供这种通过任何其它合适连接连接。

[0089] 所述图像采集装置110可以拍摄用户期望的图像(例如照片、视频等),并且将所拍摄的图像存储在所述存储装置104中以供其它组件使用。图像采集装置110可以采用各种合适的图像传感器或摄像头装置。

[0090] 示例性地,用于实现根据本发明实施例的神经网络的训练方法及装置以及用于车辆位置检测和姿态估计的方法及装置的电子设备可以集成在车辆的一个装置中,也可以分布布置在车辆的不同装置中。

[0091] 示例性地,用于实现根据本发明实施例的神经网络的训练方法及装置可以部署在云端或服务器端,用于车辆位置检测和姿态估计的方法及装置可以部署在车辆中。

[0092] 图2为根据本发明实施例的神经网络的训练方法中初始训练的示意图;图3为根据本发明实施例的神经网络的训练方法中迭代训练的示意图;图4为根据本发明实施例的神经网络的训练方法中车辆位置检测分支网络训练的示意图;图5为根据本发明实施例的神经网络的训练方法中车辆姿态估计分支网络训练的示意图。下面将结合图2~图5对根据本发明实施例的训练方法进行描述。

[0093] 本发明实施例公开一种用于神经网络的训练方法,所述神经网络用于车辆位置检测和姿态估计,所述神经网络包括车辆位置检测分支网络和车辆姿态估计分支网络,如图2所示该训练方法包括:

[0094] 首先,在步骤S201中,获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片。

[0095] 在本实施例中,将车辆位置检测问题转化为预测车辆位置坐标,将车辆姿态估计问题转化为车辆三维包围框顶点投影坐标,因此在获取训练数据时仅需标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片即可,无需更精细的标注,降低了数据获取难度。

[0096] 需要说明的是,所述三维包围框指的图片中包围车辆的三维框,其例如为一个立方体,相应地,车辆三维包围框顶点投影坐标例如为该立方体的八个顶点在图片上的投影点坐标。也即,本申请中将车辆姿态估计问题转化为预测车辆的三维包围框(也就是一个立方体)对应的八个顶点在图片上的投影点,这样更容易和车辆位置检测一起训练,训练效果也更好。

[0097] 接着,在步骤S202中,进行车辆位置检测分支网络的初始训练。

[0098] 当在步骤S201中获取训练数据后,则利用该训练数据(即标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片)训练所述车辆位置检测分支网络,也即进行车辆位置检测分支网络的初始训练。

[0099] 车辆位置检测分支网络的详细初始训练过程如图3所示,包括:

[0100] 步骤301,获取训练数据,所述训练数据包括标注了车辆位置坐标的图片和车辆三维包围框顶点投影坐标。

[0101] 步骤302,通过深层神经网络获取训练用特征图。即将标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片输入深层神经网络,经过卷积或池化才做将输入的整张图片转化成为训练用特征图,即提取图片的特征。作为一种举例,深层神经网络可以是卷积层、ReLU层、池化层等中一种或几种的组合,可以根据需要进行合适的构建。

[0102] 步骤S303,通过所述区域建议网络(即RPN网络)对所述训练用特征图进行处理以得到训练用候选框。

[0103] 作为一种举例,区域建议网络(即RPN网络)包括滑动窗口、两个并行的卷积层/全连接层,以及其它所需要的网络层。示例性地,区域建议网络(即RPN网络)获得训练用候选框的过程为:将深层神经网络得到并输入的训练用特征图,用 3×3 (即 $3 \text{scale} \times 3 \text{aspect ratio}$)的滑窗滑动扫描,通过映射为低维的特征向量后采用ReLU,每个滑窗位置对应 k 个anchors(候选区域);然后,将低维特征向量输入两个并行连接的卷积层,分别用于回归region proposals产生bounding-box和对是否为前景或背景打分。在本实施例中,RPN网络具体操作相当于在训练用特征图的每个点生成3个scale和3个ratio的候选框,然后对这些候选框进行调整,输出候选框属于前景或者背景的概率以及坐标修正值。

[0104] 步骤S304,将所述训练用候选框和所述特征图输入感兴趣区域池化层做池化处理,以得到所述训练用候选框对应的训练用第一特征向量。

[0105] 在步骤S303中,区域建议网络(即RPN网络)输出的不止一个候选框(在本实施例中,候选框为矩形框),在本步骤中对多个感兴趣区域(ROI)进行池化(Pooling),ROI Pooling的过程就是将一个个大小不同的box矩形框,都映射成大小为 $W \times H$ 的矩形框,输出batch个vector,其中batch的值等于roi的个数,向量(vector)的大小为 $\text{channel} \times W \times H$,channel表示通道,H和W是层超参数。

[0106] ROI Pooling的过程为先把roi中的坐标映射到训练用特征图上,映射规则例如为把各个坐标除以输入图片与训练用特征图的大小的比值,得到了训练用特征图上的box坐标后,使用池化得到输出;由于输入的图片大小不一,所以这里我们使用的spp pooling,

spp pooling在pooling的过程中需要计算pooling后的结果对应的两个像素点反映社到feature map上所占的范围,然后在那个范围中进行取最大或者取平均。

[0107] 即,对于每个候选框,感兴趣区域(RoI)池化层从特征图中提取固定长度的特征向量。RoI池化层使用最大池化或平均池化方法将任何有效的感兴趣区域内的特征转换成具有 $H \times W$ (例如, 7×7)的固定空间范围的小特征地图,其中 H 和 W 是层超参数,独立于任何特定的RoI层。在本文中,每个RoI是一个矩形框,用一个四元组 (r, c, h, w) 来表示,其中 (r, c) 定义了矩形框的左上角坐标, (h, w) 定义了矩形框的高度和宽度。

[0108] 示例性地,在本实施例中,使用最大池化方法,RoI最大池化就是将 $h \times w$ 的RoI窗口分成 $H \times W$ 的网格,每个子窗口的大小近似为 $h/H \times w/W$,然后最大池化每个子窗口的值到相应的输出网格单元。

[0109] 步骤S305,将所述训练用第一特征向量输入所述车辆位置检测分支网络进行处理,以得到所述训练用车辆检测框。

[0110] 当得到所述训练用第一特征向量后,将其输入车辆检测位置分支网络,在车辆检测位置分支网络中,通过回归法得到训练用车辆检测框。更具体地,车辆位置检测分支网络输出矩形框属于每一种车辆类型的概率,以及相对于区域建议网络(即RPN网络)输出的候选框的修正值,然后判断每个检测框属于车辆的分数和相对于RPN输出的物体检测框(候选框的)修正值,然后以分数最高的检测框作为车辆检测框输出。

[0111] 需要说明的是,车辆检测框是图片上检测出来的包围车辆的矩形框,由矩形的左上角顶点和长宽共4个参数表示(即类似上述的左上角 $(r; c)$ 及其高度和宽度 $(h; w)$)。

[0112] 接着,在步骤S203中,进行车辆姿态估计分支网络的初始训练。

[0113] 当在步骤S201中获取训练数据后,则利用所述训练数据(标注了车辆三维包围框顶点投影坐标和车辆三维包围框顶点投影坐标的图片)和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

[0114] 车辆姿态估计分支网络的详细初始训练过程如图4所示,包括:

[0115] 步骤401,获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片。

[0116] 步骤402,通过深层神经网络获取训练用特征图。即将标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片输入深层神经网络,通过将输入的整张图片转化成为训练用特征图,即提取图片的特征。作为一种举例,深层神经网络可以是卷积层、ReLU层、池化层等中一种或几种的组合,可以根据需要进行合适的构建。

[0117] 步骤403,通过所述感兴趣区域池化层将所述训练用车辆检测框在所述训练用特征图做池化处理,以得到所述训练用车辆检测框对应的训练用第二特征向量。

[0118] 即将车辆位置检测分支网络输出的训练用车辆检测框和深层神经网络得到训练用特征图输入到感兴趣区域池化层,通过感兴趣区域池化层处理得到训练用车辆检测框对应的训练用第二特征向量。

[0119] 由于车辆位置检测分支网络输出的车辆检测框的大小是不一样的(因为有大车有小车,车辆远近导致在图片上车辆大小也不一样),所以特征向量维数也是不一样的。感兴趣区域池化层是将这些不一样维数的特征向量通过一些方式转换成同样维数的特征向量,然后进行后续车辆姿态预测训练。池化过程与前述类似,即将训练用车辆检测框映射到训

练用特征图上,从而提取固定长度的特征向量。

[0120] 步骤404,将所述训练用第二特征向量训练输入所述车辆姿态估计分支网络进行处理,以得到车辆姿态估计信息。

[0121] 当得到所述训练用第二特征向量后,将其输入车辆姿态估计分支网络,在车辆姿态估计分支网络,通过回归法得到车辆姿态估计信息。如前所述,在本实施例中,将车辆姿态预测问题转化为了预测车辆的三维包围框(立方体)的8个三维顶点在图片上的投影(即8个二维坐标)问题,因此这里的车辆姿态估计信息即8个三维顶点在图片上的投影坐标。

[0122] 进一步在本实施例中,为了提高预测准确率,我们预测的是这8个二维坐标 (x_i, y_i) 相对于之前得到的车辆检测框中心点 (x_c, y_c) 的相对位置 $((x_i - x_c) / w, (y_i - y_c) / h)$,这里 w 和 h 是车辆检测框的宽和高,这样预测的准确率比较高。

[0123] 应当理解,上述步骤S202和S203的相关描述,仅仅给出训练过程各网络层的处理过程,得到训练用车辆检测框和车辆姿态估计信息后,所述训练还包括与训练数据上的标准数据进行比对的过程,以及根据比对结果通过反向传播等方法调整网络参数等训练过程,这些过程采用神经网络训练的常用方法,在此不再赘述。

[0124] 接着,在步骤S204中,进行迭代训练。

[0125] 即,重复执行步骤S202和步骤S203,直到满足设定条件为止。所述设定条件例如为所述车辆位置检测分支网络的损失函数和所述车辆姿态估计分支网络的损失函数分别满足各自设定的阈值。

[0126] 迭代训练过程如图5所示,包括:

[0127] 首先,执行步骤501中,利用获取的训练数据,进行车辆位置检测分支网络的迭代训练,例如进行第 N 次训练, N 为大于等于的自然数。

[0128] 接着,执行步骤502,利用获取的训练数据,进行车辆姿态估计分支网络的迭代训练,例如第 N 次训练。并且在进行车辆姿态估计分支网络的迭代训练时,利用车辆位置检测分支网络输出的训练用车辆检测框辅助训练车辆姿态估计分支网络,以提高姿态估计的准确率。至于如何利用车辆位置检测分支网络输出的训练用车辆检测框辅助训练车辆姿态估计分支网络,则可以参见图4以及相关描述。

[0129] 当完成第 N 次训练后,则接着执行步骤S503,利用已获取的训练数据或新获取的训练数据,继续进行车辆位置检测分支网络的迭代训练,即进行第 $N+1$ 次训练。

[0130] 接着,执行步骤S504,利用已获取的训练数据或新获取的训练数据,进行车辆姿态估计分支网络的迭代训练,例如第 $N+1$ 次训练。并且在进行车辆姿态估计分支网络的迭代训练时,利用车辆位置检测分支网络输出的训练用车辆检测框辅助训练车辆姿态估计分支网络,以提高姿态估计的准确率。至于如何利用车辆位置检测分支网络输出的训练用车辆检测框辅助训练车辆姿态估计分支网络,则可以参见图4以及相关描述。

[0131] 当执行完第 $N+1$ 次训练后,则继续执行第 $N+2$ 次迭代训练,训练过程类似,在此不再赘述,如此往复迭代训练,直到满足设定条件为止,以获得训练好的用于车辆位置检测和车辆姿态估计的神经网络。

[0132] 当完成车辆位置分支网络的和车辆姿态估计分支网络迭代训练后,则执行步骤S205,利用训练完成的神经网络对图片进行检测,通过车辆位置检测分支网络的输出获得车辆检测框和置信度,通过车辆姿态估计分支网络的输出获得车辆姿态估计信息。

[0133] 示例性地,根据本实施例的神经网络的训练方法可以在具有存储器和处理器的设备、装置或者系统中实现。

[0134] 根据本实施例的神经网络的训练方法,一方面训练数据只依靠标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片,不需要额外更精细的标注,对训练数据标注的要求低,降低了训练难度;另一方面,同时训练车辆位置检测分支网络和车辆姿态估计分支网络,基于训练数据可以同时输出车辆的二维检测框和车辆的姿态估计信息,不需要人工预处理和后续处理。

[0135] 进一步地,根据本发明的神经网络的训练方法,在训练车辆姿态估计分支网络时通过车辆位置检测分支网络输出的训练用车辆检测框对应的特征向量辅助预测车辆姿态信息,提高了车辆姿态估计的准确率;并且车辆姿态估计任务的加入也提高了车辆位置检测的准确度,两种任务互相辅助,互相促进。

[0136] 图6为根据本发明实施例的神经网络的训练装置的示意性结构框图。下面结合图6对根据本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练装置进行详细描述。

[0137] 如图6所示,本实施例公开一种神经网络的训练装置,所述神经网络包括深层神经网络、区域建议网络、感兴趣区域池化层、车辆位置检测分支网络和车辆姿态估计分支网络,所述训练装置600包括数据获取单元601、第一训练单元602和第二训练单元603。

[0138] 数据获取单元601用于获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片。示例性地,数据获取单元601由图1所示的电子设备中的图像采集装置来实现或者由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练方法中的步骤S201、S301和S401。

[0139] 第一训练单元602用于利用标注了车辆位置坐标的图片训练所述车辆位置检测分支网络。第一训练单元602可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练方法中的步骤S202、S302~S305、S501和S503。

[0140] 第二训练单元603用于利用标注了车辆三维包围框顶点投影坐标的图片和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。第二训练单元603可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练方法中的步骤S203、S402~S404、S502和S504。

[0141] 示例性地,根据本实施例的神经网络的训练装置可以在具有存储器和处理器的设备、装置或者系统中实现。

[0142] 根据本实施例的神经网络的训练装置,一方面训练数据只依靠标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片,不需要额外更精细的标注,对训练数据标注的要求低,降低了训练难度;另一方面,同时训练车辆位置检测分支网络和车辆姿态估计分支网络,基于训练数据可以同时输出车辆的二维检测框和车辆的姿态估计信息,不需要人工预处理和后续处理。

[0143] 进一步地,根据本发明的神经网络的训练装置,在训练车辆姿态估计分支网络时

通过车辆位置检测分支网络输出的训练用车辆检测框对应的特征向量辅助预测车辆姿态信息,提高了车辆姿态估计的准确率;并且车辆姿态估计任务的加入也提高了车辆位置检测的准确度,两种任务互相辅助,互相促进。

[0144] 图7为根据本发明实施例的用于车辆位置检测和姿态估计的方法的示意图;图8为根据本发明实施例的用于车辆位置检测和姿态估计的方法详细示意图。

[0145] 本实施例公开一种用于车辆位置检测和姿态估计的方法,如图7所示,该方法包括:

[0146] 步骤S701,获取待检测图像。

[0147] 示例性地,通过车辆上配置的摄像头或图像采集装置,获取待检测图像,所述待检测图像例如为车辆前方的图像。

[0148] 步骤S702,利用预先训练的用于车辆位置检测和姿态估计的神经网络对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息。

[0149] 其中,所述预先训练的用于车辆位置检测和姿态估计的神经网络包括深层神经网络、区域建议网络、感兴趣区域池化层、车辆位置检测分支网络和车辆姿态估计分支网络。

[0150] 本实施例的用于车辆位置检测和姿态估计的方法的详细过程,如图8所示,包括:

[0151] 步骤S801,获取待检测图像。

[0152] 步骤S802,通过深层神经网络对所述待测图像进行处理以得到特征图。

[0153] 步骤S803,通过区域建议网络对所述特征图进行处理以得到候选框。

[0154] 步骤S804,将所述候选框和所述特征图输入感兴趣区域池化层做池化处理,以得到所述候选框对应的第一特征向量。

[0155] 步骤S805,通过位置检测分支网络基于所述第一特征向量输出所述车辆检测框和置信度。

[0156] 步骤S806,将所述车辆检测框和所述特征图输入感兴趣区域池化层做池化处理,以得到所述车辆检测框对应的第二特征向量。

[0157] 步骤S807,通过所述车辆姿态估计分支网络基于所述第二特征向量输出车辆姿态估计信息。所述车辆姿态估计信息包括车辆三维包围框顶点投影坐标或车辆三维包围框顶点投影坐标相对所述车辆检测框中心点坐标的位置。

[0158] 步骤S808,获取车辆位置检测结果和车辆姿态估计结果。

[0159] 具体地,根据车辆位置检测分支网络输出的车辆检测框和置信度,基于设置或用户选择(例如设定阈值、非最大抑制等)获取车辆位置检测结果,以及根据车辆姿态估计信息以及车辆检测框获取车辆姿态估计结果。

[0160] 需要说明的是,设定阈值表示例如当置信度小于设定值的车辆检测框则被略去,非最大抑制表示将置信度不是最大值的车辆检测框抑制掉。

[0161] 根据本发明的用于车辆位置检测和姿态估计的方法,可以对输入的待检测图像同时进行车辆位置检测和车辆姿态估计,并且由于采用了经过上述训练方法或装置训练的用于车辆位置检测和姿态估计的神经网络,具有更高的准确度。

[0162] 进一步地,根据本发明的用于车辆位置检测和姿态估计的方法在进行车辆姿态估计时,将车辆姿态估计问题转化为了预测车辆的三维包围框(立方体)的8个三维顶点在图片上的投影(即8个二维坐标)问题,并且优选地在实际任务中预测的是这8个二维坐标相对

于车辆位置检测得到的车辆检测框中心点的相对位置,进一步提高了预测的准确率。

[0163] 图9为根据本发明实施例的用于车辆位置检测和姿态估计的装置的结构示意图;图10为图9中车辆位置检测和姿态估计模块的详细结构示意图。下面结合图9和图10对根据本发明实施例的用于车辆位置检测和姿态估计的装置进行详细描述。

[0164] 如图9所示,根据本发明实施例的用于车辆位置检测和姿态估计的装置900包括图像获取模块901以及车辆位置检测和姿态估计模块902。

[0165] 图像获取模块901用于获取待检测图像。图像获取模块901可以由图1所示的电子设备中的图像采集装置来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法中的步骤S701和S801。

[0166] 车辆位置检测和姿态估计模块902用于对所述待检测图像进行处理,以获取车辆位置检测结果和车辆姿态估计结果。车辆位置检测和姿态估计模块902包括车辆位置检测分支网络和车辆姿态估计分支网络,所述车辆位置检测分支网络用于基于待测图像输出车辆检测框和置信度;所述车辆姿态估计分支网络用于基于所述待测图像和所述车辆检测框输出车辆姿态估计信息。车辆位置检测和姿态估计模块902可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法中的步骤S702,以及S802~S808。

[0167] 进一步地,如图10所示,在本实施例中,车辆位置检测和姿态估计模块902包括深层神经网络1000、RPN网络1001、感兴趣区域池化层1002、车辆位置检测分支网络1003、车辆位置检测结果获取子模块1004、车辆姿态估计分支网络1005和车辆姿态估计结果获取子模块1006。

[0168] 深层神经网络1000用于对所述图像获取模块901获取的待测图像进行处理以得到特征图。深层神经网络1000可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S802。

[0169] 区域建议网络(RPN网络)1001用于对深层神经网络1000输出的所述特征图进行处理以得到候选框。区域建议网络(RPN网络)1001可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S803。

[0170] 感兴趣区域池化层1002用于对所述候选框和所述特征图做池化处理,以得到所述候选框对应的第一特征向量,以及对所述车辆检测框和所述特征图做池化处理,以得到所述车辆检测框对应的第二特征向量。感兴趣区域池化层1002可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S804和S806。

[0171] 车辆位置检测分支网络1003用于基于所述第一特征向量输出所述车辆检测框和置信度。车辆位置检测分支网络1003可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S805。

[0172] 车辆位置检测结果获取子模块1004用于根据所述车辆检测框和置信度基于用户设置或选择获得所述车辆位置检测结果。车辆位置检测结果获取子模块1004可以由图1所

示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S808。

[0173] 车辆姿态估计分支网络1005用于基于所述第二特征向量输出车辆姿态估计信息。车辆姿态估计分支网络1005可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S807。

[0174] 车辆姿态估计结果获取子模块1006用于根据所述车辆位置检测结果和所述车辆姿态估计信息获得所述车辆姿态估计结果。车辆姿态估计结果获取子模块1006可以由图1所示的电子设备中的处理器102运行存储装置104中存储的程序指令来实现,并且可以执行根据本发明实施例的用于车辆位置检测和姿态估计的方法的步骤S808。

[0175] 根据本发明的用于车辆位置检测和姿态估计的装置,可以对输入的待检测图像同时进行车辆位置检测和车辆姿态估计,并且由于采用了经过上述训练方法或装置训练的用于车辆位置检测和姿态估计的神经网络,具有更高的准确度。

[0176] 进一步地,根据本发明的用于车辆位置检测和姿态估计的装置在进行车辆姿态估计时,将车辆姿态估计问题转化为了预测车辆的三维包围框(立方体)的8个三维顶点在图片上的投影(即8个二维坐标)问题,并且优选地在实际任务中预测的是这8个二维坐标相对于车辆位置检测得到的车辆检测框中心点的相对位置,进一步提高了预测的准确率。

[0177] 此外,根据本发明的实施例,还提供了一种用于车辆位置检测和姿态估计的神经网络的训练装置,所述训练装置包括存储器和处理器,所述存储器上存储有由所述处理器运行的计算机程序,所述计算机程序在被所述处理器运行时执行如所述的用于车辆位置检测和姿态估计的神经网络的训练方法的相应步骤,并且用于实现根据本发明实施例的训练装置中的数据获取单元、第一训练单元和第二训练单元。

[0178] 在一个实施例中,所述计算机程序在被所述处理器运行时执行以下步骤:获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;重复执行以下步骤直到满足设定条件为止;利用所述训练数据训练所述车辆位置检测分支网络;利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车辆姿态估计分支网络。

[0179] 此外,根据本发明实施例,还提供了一种存储介质,在所述存储介质上存储了计算机程序,在所述计算机程序被计算机或处理器运行时用于执行本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练方法的相应步骤,并且用于实现根据本发明实施例的用于车辆位置检测和姿态估计的神经网络的训练装置中的数据获取单元、第一训练单元和第二训练单元。所述存储介质例如可以包括智能电话的存储卡、平板电脑的存储部件、个人计算机的硬盘、只读存储器(ROM)、可擦除可编程只读存储器(EPROM)、便携式紧致盘只读存储器(CD-ROM)、USB存储器、或者上述存储介质的任意组合。所述计算机可读存储介质可以是一个或多个计算机可读存储介质的任意组合。

[0180] 在一个实施例中,所述计算机程序在被计算机运行时执行以下步骤:获取训练数据,所述训练数据包括标注了车辆位置坐标和车辆三维包围框顶点投影坐标的图片;重复执行以下训练直到满足设定条件为止;利用所述训练数据训练所述车辆位置检测分支网络;利用所述训练数据和所述车辆位置检测分支网络输出的训练用车辆检测框训练所述车

辆姿态估计分支网络。

[0181] 进一步地,根据本发明的实施例,还提供了一种用于车辆位置检测和姿态估计的装置,所述装置包括存储器和处理器,所述存储器上存储有由所述处理器运行的计算机程序,所述计算机程序在被所述处理器运行时执行如所述的用于车辆位置检测和姿态估计的方法的相应步骤,并且用于实现根据本发明实施例的用于车辆位置检测和姿态估计的装置中的图像获取模块以及车辆位置检测和姿态估计模块。

[0182] 在一个实施例中,所述计算机程序在被所述处理器运行时执行以下步骤:获取待检测图像;利用预先训练的用于车辆位置检测和姿态估计的神经网络对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息。

[0183] 此外,根据本发明实施例,还提供了一种存储介质,在所述存储介质上存储了计算机程序,在所述计算机程序被计算机或处理器运行时用于执行本发明实施例的用于车辆位置检测和姿态估计的方法的相应步骤,并且用于实现根据本发明实施例的用于车辆位置检测和姿态估计的装置中的图像获取模块以及车辆位置检测和姿态估计模块。所述存储介质例如可以包括智能电话的存储卡、平板电脑的存储部件、个人计算机的硬盘、只读存储器(ROM)、可擦除可编程只读存储器(EPROM)、便携式紧致盘只读存储器(CD-ROM)、USB存储器、或者上述存储介质的任意组合。所述计算机可读存储介质可以是一个或多个计算机可读存储介质的任意组合,例如一个计算机可读存储介质包含用于人眼是否正常睁开判断的计算机可读的程序代码,另一个计算机可读存储介质包含用于驾驶员是否处于远光灯照射状态判断的计算机可读的程序代码。

[0184] 在一个实施例中,所述计算机程序在被计算机运行时执行以下步骤:获取待检测图像;利用预先训练的用于车辆位置检测和姿态估计的神经网络对所述待检测图像进行处理,以获取车辆检测框和车辆姿态估计信息。

[0185] 尽管这里已经参考附图描述了示例实施例,应理解上述示例实施例仅仅是示例性的,并且不意图将本发明的范围限制于此。本领域普通技术人员可以在其中进行各种改变和修改,而不偏离本发明的范围和精神。所有这些改变和修改意在包括在所附权利要求所要求的本发明的范围之内。

[0186] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0187] 在本申请所提供的几个实施例中,应该理解到,所揭露的设备和方法,可以通过其它的方式实现。例如,以上所描述的设备实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个设备,或一些特征可以忽略,或不执行。

[0188] 在此处所提供的说明书中,说明了大量具体细节。然而,能够理解,本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中,并未详细示出公知的方法、结构和技术,以便不模糊对本说明书的理解。

[0189] 类似地,应当理解,为了精简本发明并帮助理解各个发明方面中的一个或多个,在

对本发明的示例性实施例的描述中,本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而,并不应将该本发明的方法解释成反映如下意图:即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说,如相应的权利要求书所反映的那样,其发明点在于可以用少于某个公开的单个实施例的所有特征的特征来解决相应的技术问题。因此,遵循具体实施方式的权利要求书由此明确地并入该具体实施方式,其中每个权利要求本身都作为本发明的单独实施例。

[0190] 本领域的技术人员可以理解,除了特征之间相互排斥之外,可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述,本说明书(包括伴随的权利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0191] 此外,本领域的技术人员能够理解,尽管在此所述的一些实施例包括其它实施例中包括的某些特征而不是其它特征,但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如,在权利要求书中,所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

[0192] 本发明的各个部件实施例可以以硬件实现,或者以在一个或者多个处理器上运行的软件模块实现,或者以它们的组合实现。本领域的技术人员应当理解,可以在实践中使用微处理器或者数字信号处理器(DSP)来实现根据本发明实施例的物品分析设备中的一些模块的一些或者全部功能。本发明还可以实现为用于执行这里所描述的方法的一部分或者全部的装置程序(例如,计算机程序和计算机程序产品)。这样的实现本发明的程序可以存储在计算机可读介质上,或者可以具有一个或者多个信号的形式。这样的信号可以从因特网网站上下载得到,或者在载体信号上提供,或者以任何其他形式提供。

[0193] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制,并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了若干装置的单元权利要求中,这些装置中的若干个可以是通过同一个硬件项来具体体现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0194] 以上所述,仅为本发明的具体实施方式或对具体实施方式的说明,本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。本发明的保护范围应以权利要求的保护范围为准。

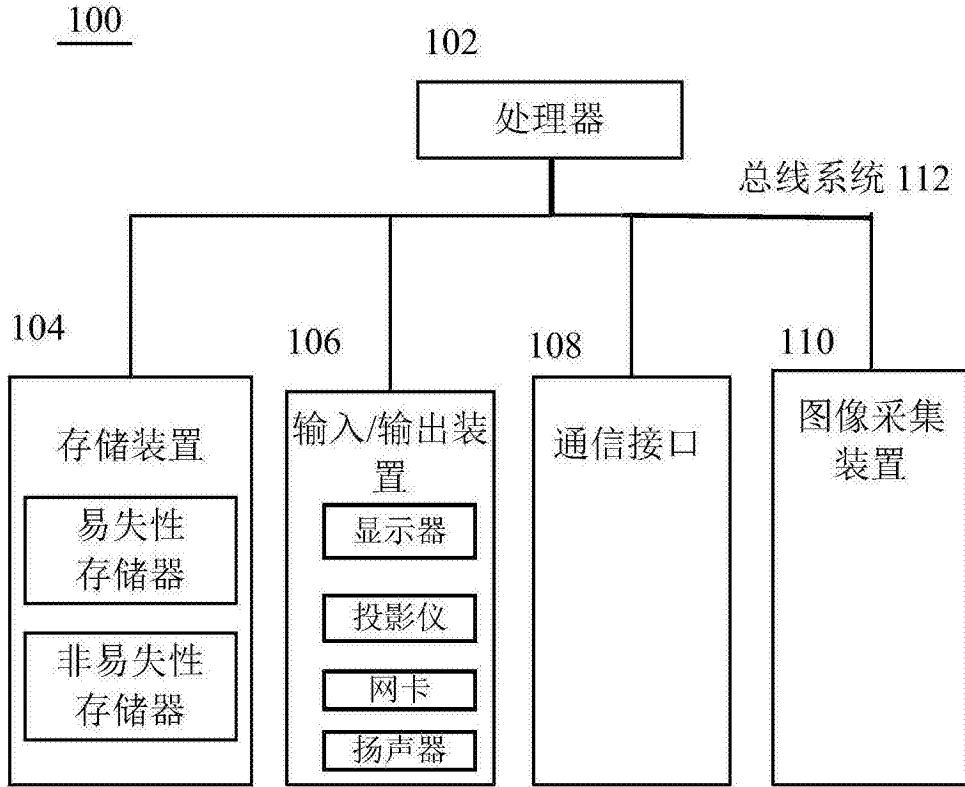


图1

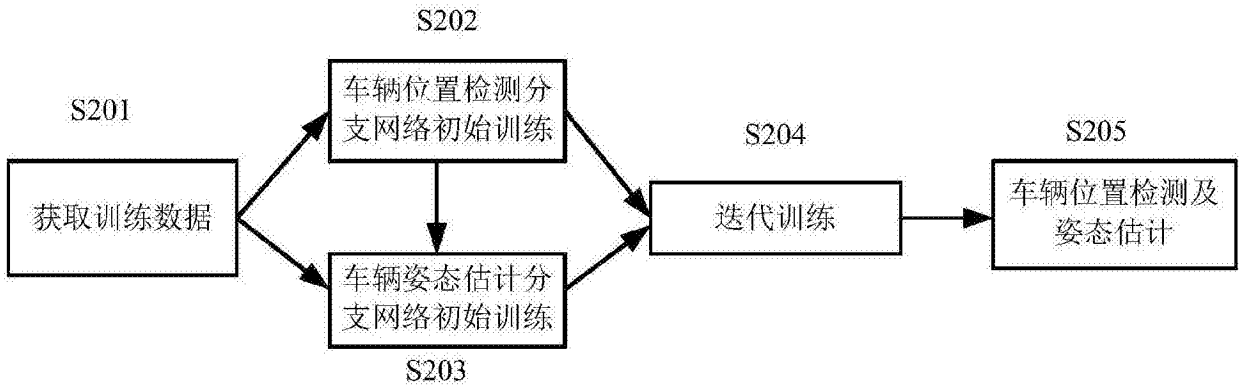


图2

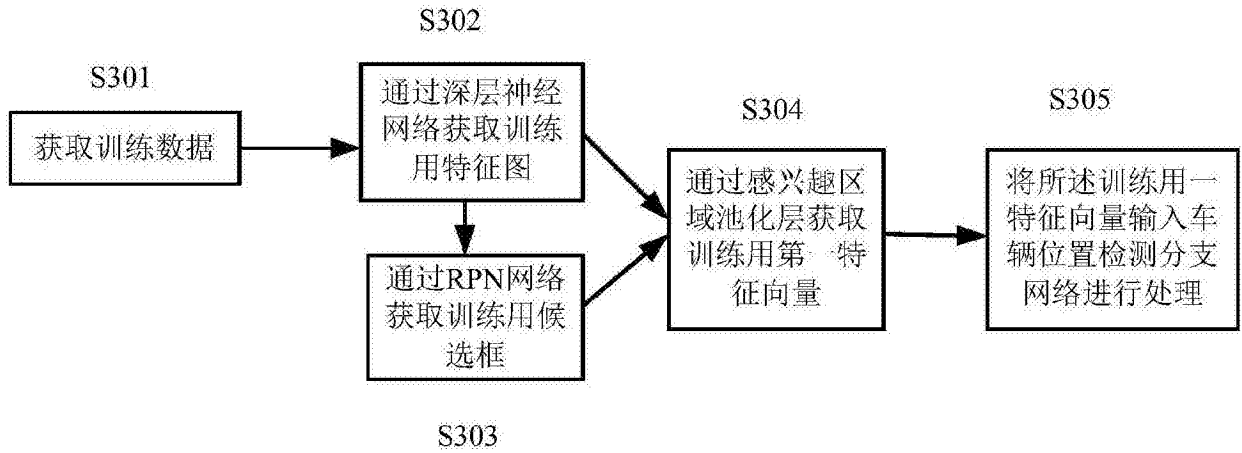


图3

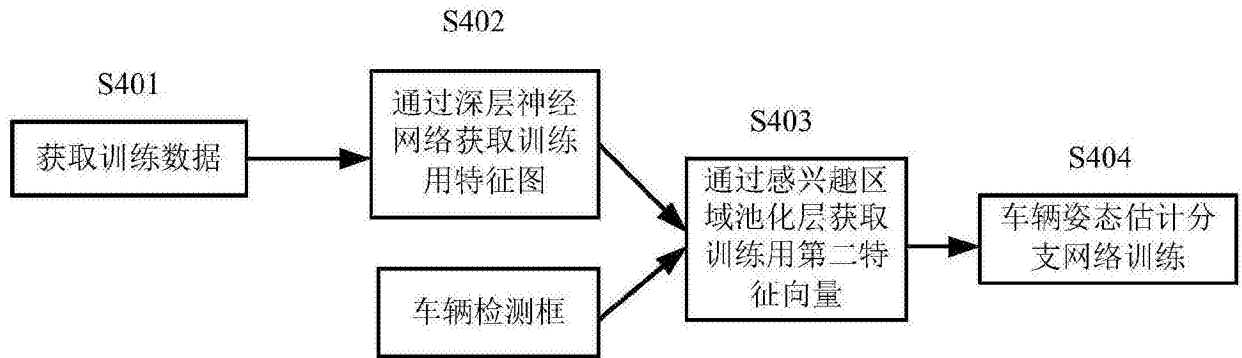


图4

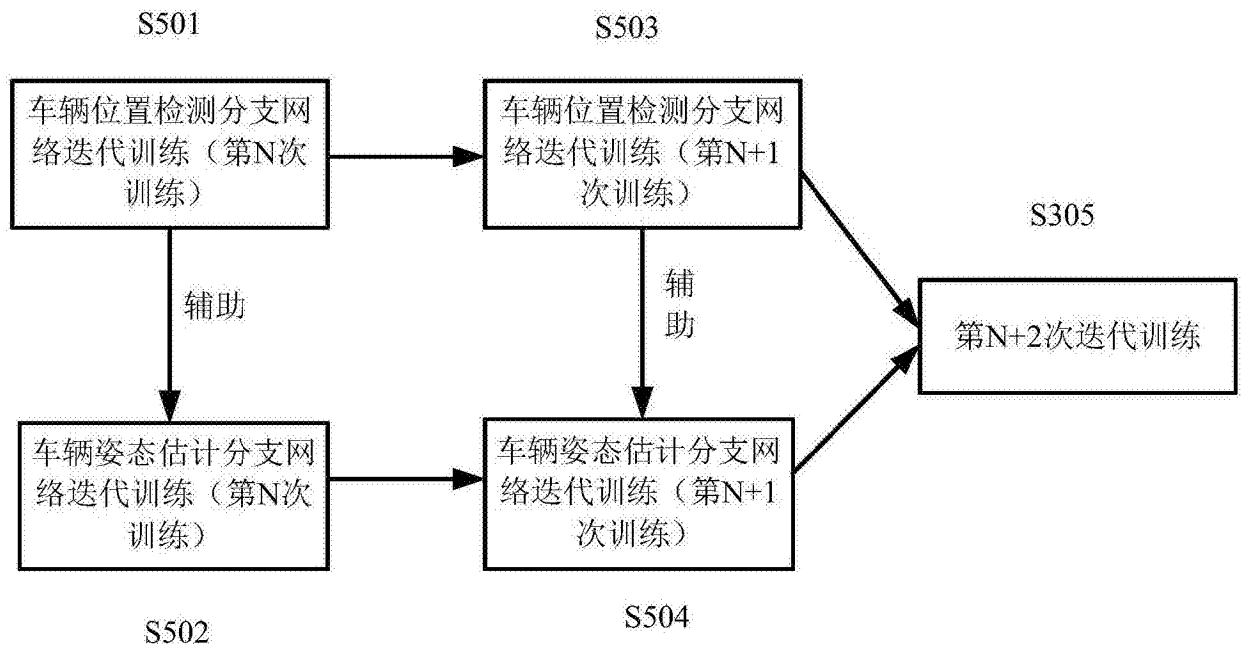


图5

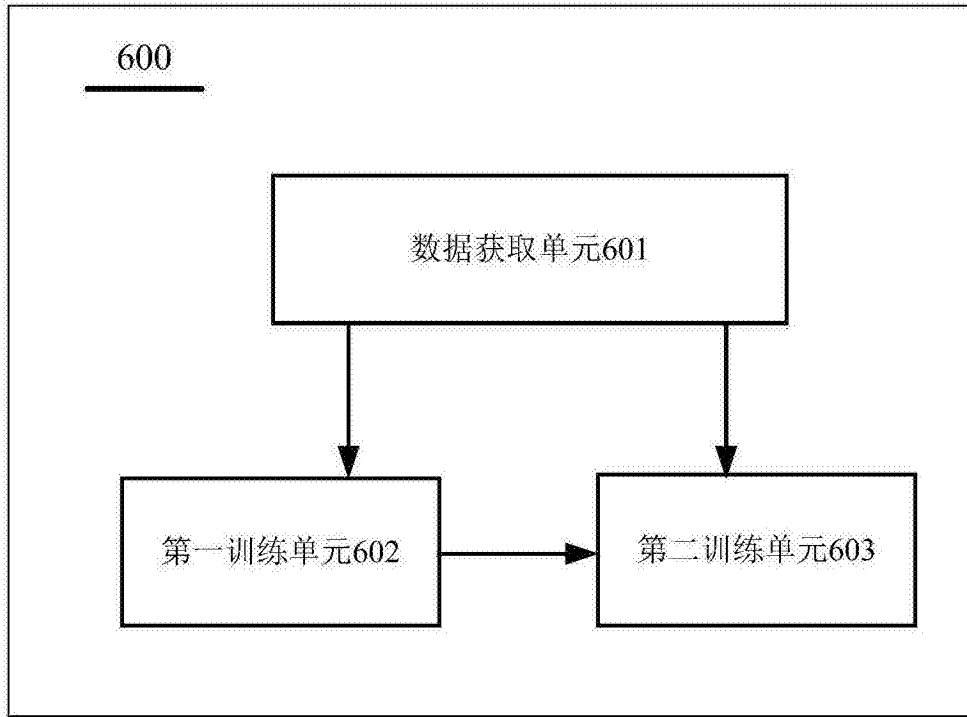


图6

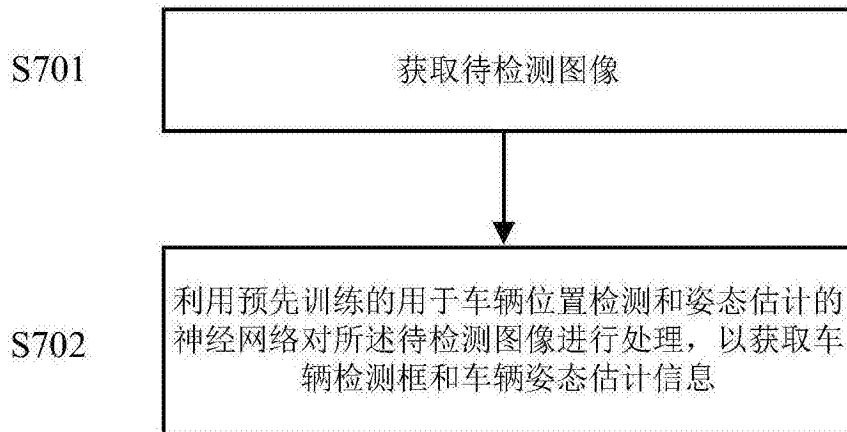


图7

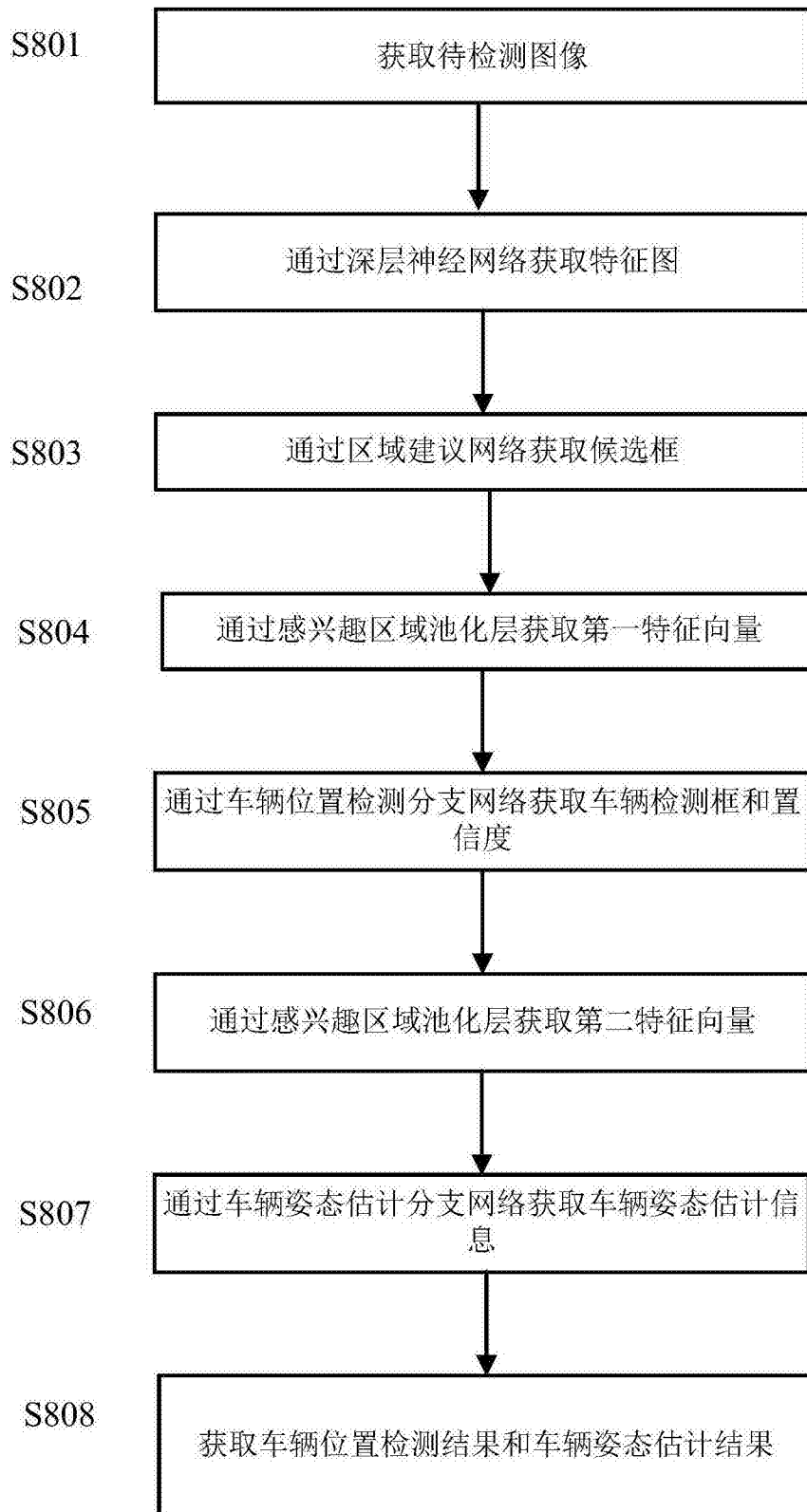


图8

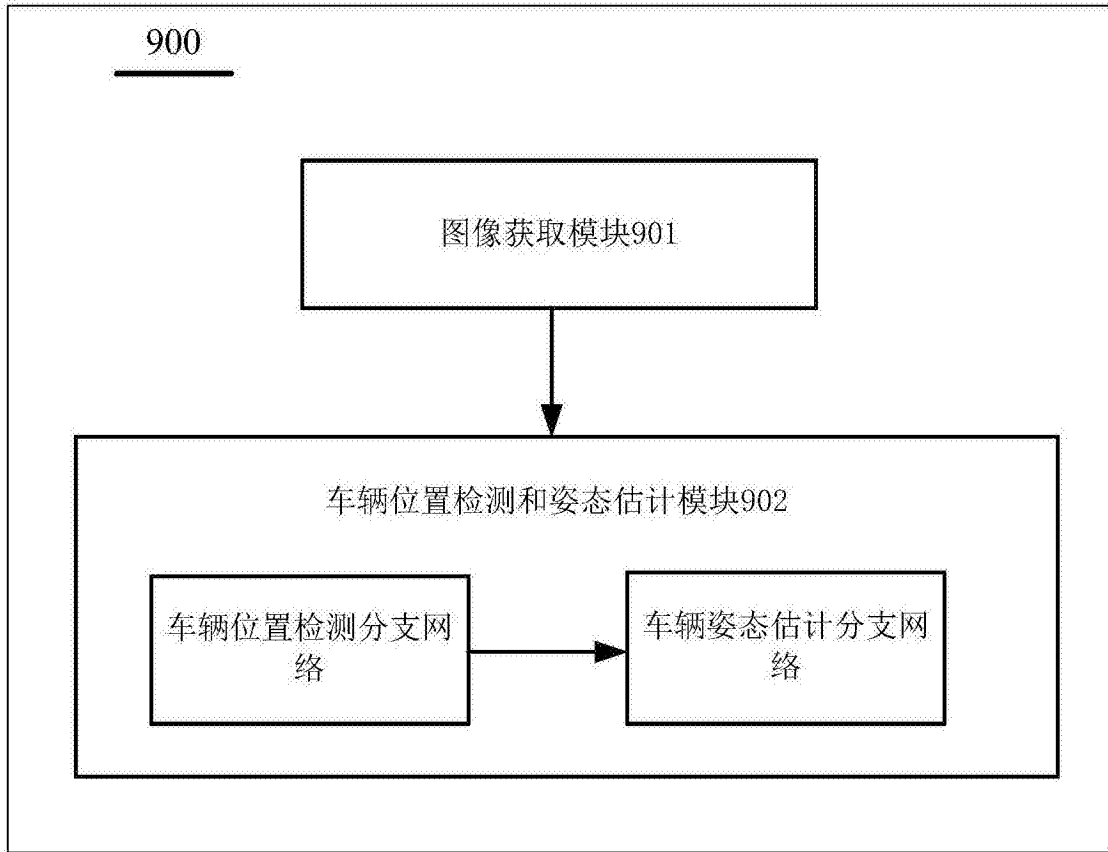


图9

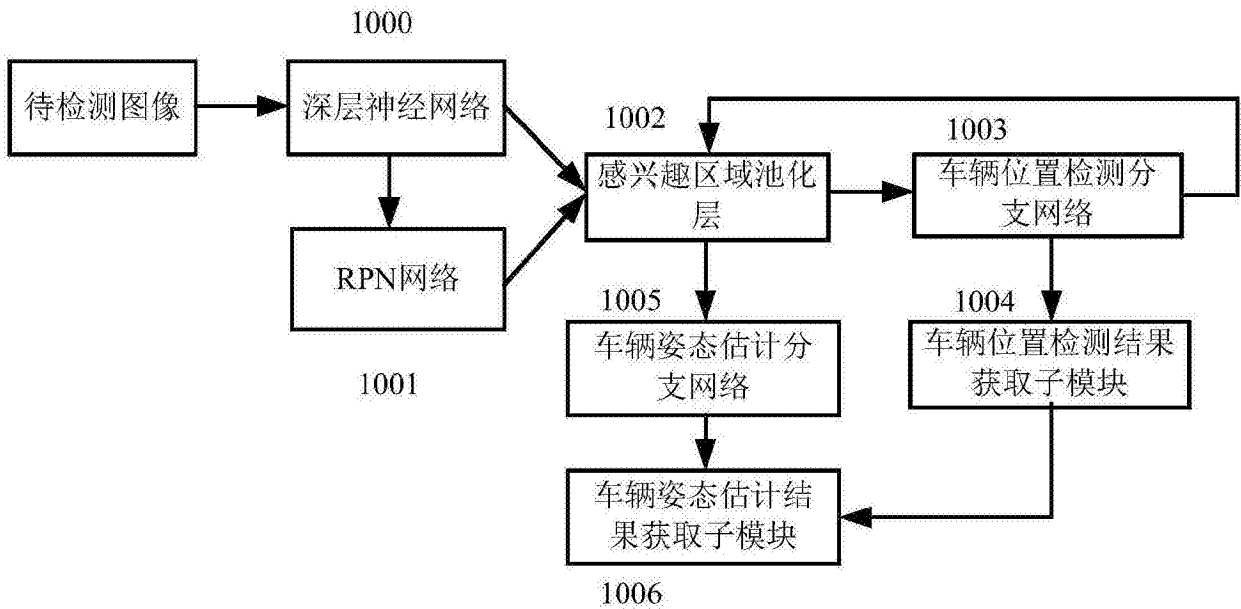


图10