



(12)发明专利申请

(10)申请公布号 CN 108280451 A

(43)申请公布日 2018.07.13

(21)申请号 201810055743.7

(22)申请日 2018.01.19

(71)申请人 北京市商汤科技开发有限公司

地址 100084 北京市海淀区中关村东路1号
院3号楼7层710-712房间

(72)发明人 石建萍

(74)专利代理机构 北京思源智汇知识产权代理
有限公司 11657

代理人 毛丽琴

(51) Int. Cl.

G06K 9/34(2006.01)

G06K 9/62(2006.01)

G06K 9/00(2006.01)

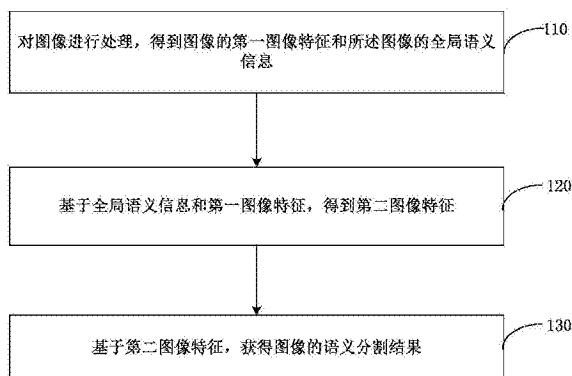
权利要求书1页 说明书22页 附图3页

(54)发明名称

语义分割及网络训练方法和装置、设备、介
质、程序

(57)摘要

本公开实施例公开了一种语义分割及网络
训练方法和装置、设备、介质、程序,其中方法包
括:对图像进行处理,获得所述图像对应的至少
一个第一图像特征,和全局语义信息;基于所述
全局语义信息和所述第一图像特征,得到第二图
像特征;基于所述第二图像特征,获得所述图像
的语义分割结果。本公开上述实施例,获得的图
像语义分割结果,能够基于全局语义信息得到图
像的语义分割结果,从而有利于提高语义分割性
能。



1. 一种语义分割方法,其特征在于,包括:
对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息;
基于所述全局语义信息和所述第一图像特征,得到第二图像特征;
基于所述第二图像特征,获得所述图像的语义分割结果。
2. 根据权利要求1所述的方法,其特征在于,所述全局语义信息包括与至少一个语义类别对应的至少一个缩放因子向量。
3. 根据权利要求1或2所述的方法,其特征在于,所述基于所述全局语义信息和所述第一图像特征,得到第二图像特征,包括:
对所述全局语义信息和所述第一图像特征进行卷积处理,得到第二图像特征。
4. 一种语义分割网络的训练方法,其特征在于,包括:
利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;
利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,其中,所述语义编码损失网络包括编码层;
基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。
5. 一种语义分割装置,其特征在于,包括:
第一处理单元,用于对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息;
第二处理单元,用于基于所述全局语义信息和所述第一图像特征,得到第二图像特征;
分割单元,用于基于所述第二图像特征,获得所述图像的语义分割结果。
6. 一种语义分割网络的训练装置,其特征在于,包括:
第三预测单元,用于利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;
第四预测单元,用于利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,其中,所述语义编码损失网络包括编码层;
第二训练单元,用于基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。
7. 一种电子设备,其特征在于,包括处理器,所述处理器包括权利要求5所述的语义分割装置或权利要求6所述的语义分割网络的训练装置。
8. 一种电子设备,其特征在于,包括:存储器,用于存储可执行指令;
以及处理器,用于与所述存储器通信以执行所述可执行指令从而完成权利要求1至3任意一项所述语义分割方法或权利要求4所述的语义分割网络的训练方法的操作。
9. 一种计算机存储介质,用于存储计算机可读的指令,其特征在于,所述指令被执行时执行权利要求1至3任意一项所述语义分割方法或权利要求4所述的语义分割网络的训练方法的操作。
10. 一种计算机程序,包括计算机可读代码,其特征在于,当所述计算机可读代码在设备上运行时,所述设备中的处理器执行用于实现权利要求1至3任意一项所述语义分割方法或权利要求4所述的语义分割网络的训练方法中各步骤的指令。

语义分割及网络训练方法和装置、设备、介质、程序

技术领域

[0001] 本公开涉及计算机视觉技术,尤其是一种语义分割及网络训练方法和装置、设备、介质、程序。

背景技术

[0002] 场景分析(Scene parsing)是基于语义分割,对描述场景的图像中的每一个像素赋予一个类别,并对不同的类别标注不同的颜色。场景分析是目标识别、图像检索、视频标注等的基础,在复杂场景中如果能够更好的利用场景分析对于场景理解的像素级理解非常关键,近年来已成为计算机视觉领域的研究热点。

发明内容

[0003] 本公开实施例提供一种语义分割技术。

[0004] 根据本公开实施例的一个方面,提供一种语义分割方法,包括:

[0005] 对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息;

[0006] 基于所述全局语义信息和所述第一图像特征,得到第二图像特征;

[0007] 基于所述第二图像特征,获得所述图像的语义分割结果。

[0008] 在基于本发明上述方法的另一个实施例中,所述全局语义信息包括与至少一个语义类别对应的至少一个缩放因子向量。

[0009] 在基于本发明上述方法的另一个实施例中,所述基于所述全局语义信息和所述第一图像特征,得到第二图像特征,包括:

[0010] 对所述全局语义信息和所述第一图像特征进行卷积处理,得到第二图像特征。

[0011] 在基于本发明上述方法的另一个实施例中,所述对图像进行处理,得到所述图像的第一图像特征,包括:

[0012] 对所述图像进行卷积处理,得到所述图像的第一图像特征。

[0013] 在基于本发明上述方法的另一个实施例中,所述对图像进行处理,得到所述图像的全局语义信息,包括:

[0014] 基于所述第一图像特征,得到所述全局语义信息。

[0015] 在基于本发明上述方法的另一个实施例中,所述基于所述第一图像特征,得到所述全局语义信息,包括:

[0016] 对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果;

[0017] 基于所述第一图像特征的编码结果,得到所述全局语义信息。

[0018] 在基于本发明上述方法的另一个实施例中,所述对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果,包括:

[0019] 确定字典和所述字典对应的权重,所述字典包括至少一个字典基;

[0020] 基于所述字典和所述字典对应的权重,对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果。

- [0021] 在基于本发明上述方法的另一个实施例中,基于所述第一图像特征的编码结果,得到所述全局语义信息,包括:
- [0022] 对所述第一图像特征的编码结果进行全连接处理,得到图像向量;
- [0023] 利用激活函数对图像向量进行处理,得到所述全局语义信息。
- [0024] 在基于本发明上述方法的另一个实施例中,所述第一图像特征包括至少一个特征图;
- [0025] 基于所述字典和所述字典对应的权重,对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果,包括:
- [0026] 分别计算至少一个所述特征图中的每个特征图与所述字典的每个字典基的对应元素的差值,并基于所述差值获得编码结果。
- [0027] 在基于本发明上述方法的另一个实施例中,基于所述全局语义信息和所述第一图像特征,得到第二图像特征,包括:
- [0028] 将所述第一图像特征与所述全局语义信息中的缩放因子向量的点积结果确定为所述第二图像特征。
- [0029] 在基于本发明上述方法的另一个实施例中,基于所述第二图像特征,获得所述图像的语义分割结果,包括:
- [0030] 对所述第二图像特征执行卷积处理,得到卷积结果;
- [0031] 对所述卷积结果进行分类处理,得到所述图像的语义分割结果。
- [0032] 在基于本发明上述方法的另一个实施例中,对所述第二图像特征执行卷积处理,得到卷积结果,包括:
- [0033] 对所述第二图像特征执行仅一次卷积处理,得到卷积结果。
- [0034] 在基于本发明上述方法的另一个实施例中,所述基于所述第二图像特征,获得所述图像的语义分割结果之前,还包括:
- [0035] 对所述第二图像特征进行放大处理,以将所述第二图像特征的维度放大至所述第一图像特征的维度;
- [0036] 所述对所述第二图像特征执行卷积处理,得到卷积结果,包括:
- [0037] 对放大处理后的所述第二图像特征执行卷积处理,得到卷积结果。
- [0038] 在基于本发明上述方法的另一个实施例中,所述对所述卷积结果进行分类处理,得到所述图像的语义分割结果之前,还包括:
- [0039] 对所述卷积结果进行放大处理,以将所述卷积结果的维度放大至所述第一图像特征的维度;
- [0040] 所述对所述卷积结果进行分类处理,得到所述图像的语义分割结果,包括:
- [0041] 对放大处理后的所述卷积结果进行分类处理,得到所述图像的语义分割结果。
- [0042] 在基于本发明上述方法的另一个实施例中,所述方法通过语义分割网络实现,所述方法还包括:
- [0043] 利用所述语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;
- [0044] 利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果;
- [0045] 基于所述第一预测语义分割结果、第二语义分割预测结果和所述样本图像的标注

语义分割信息,训练所述语义分割网络。

[0046] 在基于本发明上述方法的另一个实施例中,利用所述语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果,包括:

[0047] 利用所述语义分割网络对样本图像进行处理,获得所述样本图像的样本图像特征;

[0048] 利用所述语义分割网络对所述样本图像特征进行处理,得到所述样本图像的第一语义分割结果;

[0049] 所述利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,包括:

[0050] 利用所述语义编码损失网络对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果;

[0051] 利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0052] 在基于本发明上述方法的另一个实施例中,在所述语义编码损失网络中,不同的物体类别具有相同的权重。

[0053] 在基于本发明上述方法的另一个实施例中,利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果,包括:

[0054] 利用所述语义编码损失网络中的全连接层和响应函数对所述编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0055] 在基于本发明上述方法的另一个实施例中,所述基于所述第一预测语义分割结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络,包括:

[0056] 基于所述第一预测语义分割结果和所述样本图像的标注语义分割信息得到第一差异,基于所述第二预测语义分割结果和所述样本图像的标注语义分割信息得到第二差异;

[0057] 根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数。

[0058] 在基于本发明上述方法的另一个实施例中,根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数,包括:

[0059] 基于所述第一差异与所述第二差异的和,得到第三差异;

[0060] 基于所述第三差异,调整所述语义分割网络中的参数。

[0061] 根据本公开实施例的另一个方面,提供一种语义分割网络的训练方法,包括:

[0062] 利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;

[0063] 利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,其中,所述语义编码损失网络包括编码层;

[0064] 基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。

[0065] 在基于本发明上述方法的另一个实施例中,所述利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果,包括:

[0066] 利用所述语义分割网络对样本图像进行处理,得到所述样本图像的样本图像特

征；

[0067] 利用所述语义分割网络对所述样本图像进行处理,得到所述样本图像的第一语义分割结果；

[0068] 所述利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,包括：

[0069] 利用语义编码损失网络的编码层对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果；

[0070] 利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0071] 在基于本发明上述方法的另一个实施例中,所述利用语义编码损失网络对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果,包括：

[0072] 确定字典和所述字典对应的权重,所述字典包括至少一个字典基；

[0073] 基于所述字典和所述字典对应的权重,对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果。

[0074] 在基于本发明上述方法的另一个实施例中,所述样本图像特征包括至少一个样本特征图；

[0075] 基于所述字典和所述字典对应的权重,对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果,包括：

[0076] 分别计算至少一个所述样本特征图中的每个样本特征图与所述字典的每个字典基的对应元素的差值,并基于所述差值获得编码结果。

[0077] 在基于本发明上述方法的另一个实施例中,所述语义编码损失网络还包括全连接层和响应函数层；

[0078] 利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果,包括：

[0079] 利用所述语义编码损失网络中的全连接层和响应函数层对所述编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0080] 在基于本发明上述方法的另一个实施例中,基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像标注的语义分割信息,训练所述语义分割网络,包括：

[0081] 基于所述第一语义分割预测结果和所述样本图像标注的语义分割信息得到第一差异,基于第二语义分割预测结果和所述样本图像标注的语义分割信息得到第二差异；

[0082] 根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数。

[0083] 在基于本发明上述方法的另一个实施例中,根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数,包括：

[0084] 基于所述第一差异与所述第二差异的和,得到第三差异；

[0085] 基于所述第三差异,调整所述语义分割网络中的参数。

[0086] 根据本公开实施例的另一个方面,提供一种语义分割装置,包括：

[0087] 第一处理单元,用于对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息；

[0088] 第二处理单元,用于基于所述全局语义信息和所述第一图像特征,得到第二图像特征；

- [0089] 分割单元,用于基于所述第二图像特征,获得所述图像的语义分割结果。
- [0090] 在基于本发明上述装置的另一个实施例中,所述全局语义信息包括与至少一个语义类别对应的至少一个缩放因子向量。
- [0091] 在基于本发明上述装置的另一个实施例中,所述第二处理单元,具体用于对所述全局语义信息和所述第一图像特征进行卷积处理,得到第二图像特征。
- [0092] 在基于本发明上述装置的另一个实施例中,所述第一处理单元,包括:
- [0093] 图像处理模块,用于对所述图像进行卷积处理,得到所述图像的第一图像特征。
- [0094] 在基于本发明上述装置的另一个实施例中,所述第一处理单元,还包括:
- [0095] 特征处理模块,用于基于所述第一图像特征,得到所述全局语义信息。
- [0096] 在基于本发明上述装置的另一个实施例中,所述特征处理模块,包括:
- [0097] 编码模块,用于对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果;
- [0098] 语义模块,用于基于所述第一图像特征的编码结果,得到所述全局语义信息。
- [0099] 在基于本发明上述装置的另一个实施例中,所述编码模块,包括:
- [0100] 字典确定模块,用于确定字典和所述字典对应的权重,所述字典包括至少一个字典基;
- [0101] 编码结果模块,用于基于所述字典和所述字典对应的权重,对所述第一图像特征进行编码处理,得到所述第一图像特征的编码结果。
- [0102] 在基于本发明上述装置的另一个实施例中,所述语义模块,具体用于对所述第一图像特征的编码结果进行全连接处理,得到图像向量;利用激活函数对图像向量进行处理,得到所述全局语义信息。
- [0103] 在基于本发明上述装置的另一个实施例中,所述第一图像特征包括至少一个特征图;
- [0104] 所述编码结果模块,具体用于分别计算至少一个所述特征图中的每个特征图与所述字典的字典基的对应元素的差值,并基于所述差值获得编码结果。
- [0105] 在基于本发明上述装置的另一个实施例中,所述第二处理单元,具体用于将所述第一图像特征与所述全局语义信息中的缩放因子向量的点积结果确定为所述第二图像特征。
- [0106] 在基于本发明上述装置的另一个实施例中,所述分割单元,包括:
- [0107] 卷积模块,用于对所述第二图像特征执行卷积处理,得到卷积结果;
- [0108] 分类模块,用于对所述卷积结果进行分类处理,得到所述图像的语义分割结果。
- [0109] 在基于本发明上述装置的另一个实施例中,所述卷积模块,具体用于对所述第二图像特征执行仅一次卷积处理,得到卷积结果。
- [0110] 在基于本发明上述装置的另一个实施例中,所述分割单元,还包括:
- [0111] 特征放大模块,用于对所述第二图像特征进行放大处理,以将所述第二图像特征的维度放大至所述第一图像特征的维度;
- [0112] 所述卷积模块,具体用于对放大处理后的所述第二图像特征执行卷积处理,得到卷积结果。
- [0113] 在基于本发明上述装置的另一个实施例中,所述分割单元,还包括:

[0114] 结果放大模块,用于对所述卷积结果进行放大处理,以将所述卷积结果的维度放大至所述第一图像特征的维度;

[0115] 所述分类模块,具体用于对放大处理后的所述卷积结果进行分类处理,得到所述图像的语义分割结果。

[0116] 在基于本发明上述装置的另一个实施例中,所述装置基于语义分割网络实现,所述装置还包括:

[0117] 第一预测单元,利用所述语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;

[0118] 第二预测单元,用于利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果;

[0119] 第一训练单元,用于基于所述第一预测语义分割结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。

[0120] 在基于本发明上述装置的另一个实施例中,所述第一预测单元,具体用于利用所述语义分割网络对样本图像进行处理,获得所述样本图像的样本图像特征;

[0121] 利用所述语义分割网络对所述样本图像特征进行处理,得到所述样本图像的第一语义分割结果;

[0122] 所述第二预测单元,包括:

[0123] 样本编码模块,用于利用所述语义编码损失网络对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果;

[0124] 样本处理模块,用于利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0125] 在基于本发明上述装置的另一个实施例中,在所述语义编码损失网络中,不同的物体类别具有相同的权重。

[0126] 在基于本发明上述装置的另一个实施例中,所述样本处理模块,具体用于利用所述语义编码损失网络中的全连接层和响应函数对所述编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0127] 在基于本发明上述装置的另一个实施例中,所述第一训练单元,包括:

[0128] 差异模块,用于基于所述第一预测语义分割结果和所述样本图像的标注语义分割信息得到第一差异,基于所述第二预测语义分割结果和所述样本图像的标注语义分割信息得到第二差异;

[0129] 参数调整模块,用于根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数。

[0130] 在基于本发明上述装置的另一个实施例中,所述参数调整模块,具体用于基于所述第一差异与所述第二差异的和,得到第三差异;基于所述第三差异,调整所述语义分割网络中的参数。

[0131] 根据本公开实施例的另一个方面,提供一种语义分割网络的训练装置,包括:

[0132] 第三预测单元,用于利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;

[0133] 第四预测单元,用于利用语义编码损失网络,获得所述样本图像的第二语义分割

预测结果,其中,所述语义编码损失网络包括编码层;

[0134] 第二训练单元,用于基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。

[0135] 在基于本发明上述装置的另一个实施例中,所述第三预测单元,具体用于利用所述语义分割网络对样本图像进行处理,得到所述样本图像的样本图像特征;利用所述语义分割网络对所述样本图像进行处理,得到所述样本图像的第一语义分割结果;

[0136] 所述第四预测单元,包括:

[0137] 特征编码模块,用于利用语义编码损失网络的编码层对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果;

[0138] 语义预测模块,用于利用所述语义编码损失网络对所述样本图像特征的编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0139] 在基于本发明上述装置的另一个实施例中,所述特征编码模块,包括:

[0140] 权重字典模块,用于确定字典和所述字典对应的权重;所述字典包括至少一个字典基;

[0141] 样本结果模块,用于基于所述字典和所述字典对应的权重,对所述样本图像特征进行编码处理,得到所述样本图像特征的编码结果。

[0142] 在基于本发明上述装置的另一个实施例中,所述样本图像特征包括至少一个样本特征图;

[0143] 所述样本结果模块,具体用于分别计算至少一个所述样本特征图中的每个样本特征图与所述字典的每个字典基的各对应元素的差值,并基于所述差值获得编码结果。

[0144] 在基于本发明上述装置的另一个实施例中,所述语义编码损失网络还包括全连接层和响应函数层;

[0145] 所述语义预测模块,具体用于利用所述语义编码损失网络中的全连接层和响应函数层对所述编码结果进行处理,得到所述样本图像的第二语义分割预测结果。

[0146] 在基于本发明上述装置的另一个实施例中,所述第二训练单元,包括:

[0147] 差异获得模块,用于基于所述第一语义分割预测结果和所述样本图像标注的语义分割信息得到第一差异,基于第二语义分割预测结果和所述样本图像标注的语义分割信息得到第二差异;

[0148] 网络参数模块,用于根据所述第一差异和所述第二差异,调整所述语义分割网络中的参数。

[0149] 在基于本发明上述装置的另一个实施例中,所述网络参数模块,具体用于基于所述第一差异与所述第二差异的和,得到第三差异;基于所述第三差异,调整所述语义分割网络中的参数。

[0150] 根据本公开实施例的另一个方面,提供一种电子设备,包括处理器,所述处理器包括如上所述的语义分割装置或语义分割网络的训练装置。

[0151] 根据本公开实施例的另一个方面,提供一种电子设备,包括:存储器,用于存储可执行指令;

[0152] 以及处理器,用于与所述存储器通信以执行所述可执行指令从而完成如上所述语义分割方法或语义分割网络的训练方法的操作。

[0153] 根据本公开实施例的另一个方面,提供的一种计算机存储介质,用于存储计算机可读取的指令,所述指令被执行时执行如上所述语义分割方法或语义分割网络的训练方法的操作。

[0154] 根据本公开实施例的另一个方面,提供的一种计算机程序,包括计算机可读代码,当所述计算机可读代码在设备上运行时,所述设备中的处理器执行用于实现如上所述语义分割方法或语义分割网络的训练方法中各步骤的指令。

[0155] 根据本公开实施例的再一个方面,提供的一种计算机程序产品,用于存储计算机可读指令,所述指令被执行时使得计算机执行上述任一可能的实现方式中所述的语义分割方法或语义分割网络的训练方法。

[0156] 在一个可选实施方式中,所述计算机程序产品具体为计算机存储介质,在另一个可选实施方式中,所述计算机程序产品具体为软件产品,例如SDK等。

[0157] 根据本公开实施例还提供了另一种语义分割方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品,其中,该方法包括:对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息;基于所述全局语义信息和所述第一图像特征,得到第二图像特征;基于所述第二图像特征,获得所述图像的语义分割结果。

[0158] 根据本公开实施例还提供了另一种语义分割网络的训练方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品,其中,该方法包括:利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,其中,所述语义编码损失网络包括编码层;基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。

[0159] 基于本公开上述实施例提供的一种语义分割及网络训练方法和装置、设备、介质、程序,对图像进行处理,得到图像的第一图像特征和图像的全局语义信息,基于全局语义信息和第一图像特征,得到第二图像特征,基于第二图像特征,获得图像的语义分割结果,能够基于全局语义信息得到图像的语义分割结果,从而有利于提高语义分割性能。

附图说明

[0160] 构成说明书的一部分的附图描述了本公开的实施例,并且连同描述一起用于解释本公开的原理。

[0161] 参照附图,根据下面的详细描述,可以更加清楚地理解本公开,其中:

[0162] 图1为本公开语义分割方法一些实施例的流程示意图。

[0163] 图2为本公开实施例的语义分割方法训练的一个具体示例的结构示意图。

[0164] 图3为本公开语义分割装置一些实施例的结构示意图。

[0165] 图4为本公开语义分割网络训练方法一些实施例的流程示意图。

[0166] 图5为本公开语义分割网络的训练装置实施例的结构示意图。

[0167] 图6为适于用来实现本公开实施例的终端设备或服务器的电子设备的结构示意图。

具体实施方式

[0168] 现在将参照附图来详细描述本公开的各种示例性实施例。应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本公开的范围。

[0169] 同时，应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

[0170] 以下对至少一个示例性实施例的描述实际上仅仅是说明性的，决不作为对本公开及其应用或使用的任何限制。

[0171] 对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

[0172] 应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

[0173] 本公开实施例可以应用于计算机系统/服务器，其可与众多其它通用或专用计算机系统环境或配置一起操作。适于与计算机系统/服务器一起使用的众所周知的计算系统、环境和/或配置的例子包括但不限于：个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境，等等。

[0174] 计算机系统/服务器可以在由计算机系统执行的计算机系统可执行指令（诸如程序模块）的一般语境下描述。通常，程序模块可以包括例程、程序、目标程序、组件、逻辑、数据结构等等，它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施，分布式云计算环境中，任务是由通过通信网络链接的远程处理设备执行的。在分布式云计算环境中，程序模块可以位于包括存储设备的本地或远程计算机系统存储介质上。

[0175] 图1为本公开语义分割方法一些实施例的流程示意图，其中，可选地，该语义分割结果可以通过语义分割网络执行，但本公开实施例不限于此。如图1所示，该方法包括：

[0176] 步骤110，对图像进行处理，得到图像的第一图像特征和图像的全局语义信息。

[0177] 在本公开实施例中，可以对图像进行一种或多种处理，得到图像的第一图像特征，其中，在一个或多个可选的实施例中，可以对图像进行卷积处理，例如：利用卷积神经网络（可以是任意预先训练的模型，如：ResNet50, ResNet101等网络）对图像进行卷积处理，得到图像的第一图像特征，其中，该卷积处理的次数可以为一次或多次，例如，利用卷积神经网络的多个卷积层依次对图像进行处理。或者，也可以对图像进行卷积和其他类型的处理，得到图像的第一图像特征，本公开实施例对此不做限定。

[0178] 可选地，第一图像特征是高层语义特征，在具体示例中，高层语义特征可表示为多维特征向量。

[0179] 在一个可选示例中，该操作110可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的第一网络执行。

[0180] 此外，还可以得到图像的全局语义信息，该全局语义信息也可以称为场景全局信息，可以包括该图像的场景相关的全局信息。在一些可选实施例中，全局语义信息包括与至少一个语义类别对应的至少一个缩放因子向量，其中，可选地，不同的语义类别可以对应于相同或不同的缩放因子向量，并且在至少一个语义类别的数量为多个的情况下，存在对

应于不同缩放因子向量的至少两个语义类别。这样,可以通过全局语义信息中的缩放因子向量对场景中的一些类别进行强调,和/或对场景中一些类别进行抑制,使图像中的所有类别都能被准确的识别,其中,该需要强调和/或抑制的类别可以根据应用需求确定。例如,可以通过全局语义信息中的缩放因子向量对小物体对应的类别进行强调,和/或通过全局语义信息中的缩放因子向量对大物体对应的类别进行抑制,但本公开实施例对此不做限定。

[0181] 这样,通过全局语义信息中的缩放因子向量,能够在第二图像特征中有效提升图像中场景相关类别特征的权重,同时抑制与场景信息不匹配的类别特征的权重。

[0182] 在一些可选实施例中,该全局语义信息可以是通过对该图像的图像特征进行编码处理得到的,其中,该图像特征可以是通过对该图像进行卷积处理和其他处理得到的,本公开实施例对此不做限定。

[0183] 步骤120,基于全局语义信息和第一图像特征,得到第二图像特征。

[0184] 可选地,该第一图像特征可以包含类别相关的信息。

[0185] 在一个或多个可选的实施例中,可以对全局语义信息和第一图像特征进行卷积处理,得到第二图像特征。

[0186] 可选地,可以将第一图像特征与全局语义信息中的缩放因子向量的点积结果确定为第二图像特征。

[0187] 在一个可选示例中,该操作120可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的编码网络执行。

[0188] 步骤130,基于第二图像特征,获得图像的语义分割结果。

[0189] 其中,语义分割结果可以包括图像中至少一个像素中每个像素所属的语义类别,其中,该至少一个像素可以为该图像的部分或所有像素,例如,该语义分割结果可以包括图像中所有像素的类别,但本公开实施例对此不做限定。

[0190] 在一个可选示例中,该操作130可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的第二网络执行。

[0191] 基于本公开上述实施例提供的一种语义分割方法,对图像进行处理,得到图像的第一图像特征和图像的全局语义信息,基于全局语义信息和第一图像特征,得到第二图像特征,基于第二图像特征,获得图像的语义分割结果,能够基于全局语义信息得到图像的语义分割结果,从而有利于提高语义分割性能。

[0192] 本公开语义分割方法的一些实施例中,全局语义信息可以是基于第一图像特征得到的。可选地,可以对该第一图像特征进行编码处理,得到该全局语义信息,但该全局语义信息也可以是通过对该第一图像特征进行其他类型的处理得到的,本公开实施例不限于此。

[0193] 在一个或多个可选的实施例中,可以对第一图像特征进行编码处理,得到第一图像特征的编码结果,并基于该第一图像特征的编码结果,得到该全局语义信息。

[0194] 可选地,可以对该第一图像特征进行稀疏编码。在一个具体例子中,获得编码结果的过程包括:确定字典和字典对应的权重,其中,字典包括至少一个字典基;

[0195] 基于字典和字典对应的权重,对第一图像特征进行编码处理,得到第一图像特征的编码结果。

[0196] 该字典及其对应的权重可以通过训练获得的,但本公开实施例不限于此。

[0197] 可选地,第一图像特征可以包括至少一个特征图;相应地,可以分别计算至少一个特征图中的每个特征图与字典的每个字典基的对应元素的差值,并基于差值获得编码结果。

[0198] 该特征图可以具体为三维张量,可以对每个特征图与每个字典基进行逐元素求差,并基于每个特征图与每个字典基的逐元素差值,获得第一图像特征的编码结果。

[0199] 在一个或多个可选的实施例中,当操作110输出的第一图像特征为N个 $H \times W \times C$ 大小的特征图,其中,H、W、C分别为特征向量的长、宽、通道数(channel)。将该第一图像特征表示为矩阵为 $X = \{X_1, \dots, X_N\}$, X_i 表示一个特征图(i的取值范围为1至N),对应的学习到的字典表示为 $D = \{d_1, \dots, d_k\}$, d_j 表示一个字典基(j的取值范围为1至k),字典D包括K组字典基,其中,K可以等于或不等于N,对应字典的权重 $S = \{s_1, \dots, s_k\}$,每个字典基对应一个权重值;

[0200] 基于字典基和特征图计算差值可以通过以下公式(1)实现:

$$r_{ik} = X_i - d_j \quad \text{公式(1)}$$

[0202] 其中,i的取值范围为1至N, r_{ik} 为特征图与字典基的差值,通过上述公式(1)的计算将得到 $N \times K$ 个差值,基于 $N \times K$ 个差值获得编码结果。

[0203] 在一个可选的实施例中,基于上述得到的 $N \times K$ 个差值和对应字典的权重 $S = \{s_1, \dots, s_k\}$ 可通过以下公式(2)获得编码结果:

$$e_k = \frac{\exp(-s_k \|r_{ik}\|^2)}{\sum_{j=1}^K \exp(-s_j \|r_{ij}\|^2)} r_{ik} \quad \text{公式(2)}$$

[0205] 其中, e_k 为编码结果的第k个分量,exp表示以自然常数e为底的指数函数,通过上述公式(2)结合权重S获得编码结果。

[0206] 可选地,可以对该第一图像特征的编码结果进行激活处理和/或全连接处理等处理,得到全局语义信息。在一些可选实施例中,可以对第一图像特征的编码结果进行全连接处理,得到图像向量,并利用激活函数对图像向量进行处理,得到全局语义信息。

[0207] 其中,可选地,激活函数为Sigmoid函数,可以用作神经网络的阈值函数,将变量映射到0到1之间。

[0208] 可选地,在全局语义信息可以包括至少一个缩放因子向量的情况下,缩放因子向量 γ 可以表示为 $\gamma = \delta(We)$,其中,W是层权重, δ 是sigmoid函数层,e表示编码结果;该公式表达的即为编码向量经过全连接层和激活函数层的变换获得缩放因子向量。

[0209] 本公开语义分割方法的一些实施例中,图像的语义分割结果可以是通过对第二图像特征进行卷积处理得到的。在一个具体例子中,可以对第二图像特征执行卷积处理,得到卷积结果,并基于该卷积结果,得到图像的语义分割结果。

[0210] 可选地,该第二图像特征的网络层可以位于语义分割网络的最后一个卷积层之前。此时,可以对第二图像特征执行仅一次卷积处理,得到卷积结果。

[0211] 基于第二图像特征经过卷积得到的分类结果,实现对图像基于语义的像素分类,具体的,可获得图像中哪些像素对应哪些类别,并且,由于第二图像是基于全局语义信息和第一图像特征获得的,实现了对图像中某些类别(例如小物体类别)的强调,使图像中的所有类别都能够被准确的识别。

[0212] 在一个或多个可选的实施例中,操作130之前,还包括:

[0213] 对第二图像特征进行放大处理,以将第二图像特征的维度放大至第一图像特征的

维度；

[0214] 对第二图像特征执行卷积处理,得到卷积结果,包括:

[0215] 对放大处理后的第二图像特征执行卷积处理,得到卷积结果。

[0216] 由于第二图像特征是第一图像特征与缩放因子向量执行点乘获得的,此时第二图像特征的大小与第一图像特征的大小不同,为了实现对场景图像中的各像素进行分类,将第二图像特征放大至第一图像特征的大小,具体可选地可以通过上采样层或其他方式实现特征放大,以实现输出对应各像素的分类结果。

[0217] 具体地,可采用深度神经网络,该深度神经网络至少包括卷积层和分类层,将第二图像特征作为一个特征图输入该深度神经网络,得到对应该第二图像特征的分类结果,及对应图像的语义分割结果。通过深度神经网络中的卷积层对各第二图像特征进行降维度(也称为:降维)处理。在一个可选示例中,该操作可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的深度神经网络或该深度神经网络中的卷积层执行。相应地,后续操作中,可以结合第一图像特征与上述至少一个降维后的第二图像特征,对图像进行场景分析。

[0218] 上述实施例还存在另一种情况,操作130之前,还包括:

[0219] 对卷积结果进行放大处理,以将卷积结果的维度放大至第一图像特征的维度;

[0220] 对卷积结果进行分类处理,得到图像的语义分割结果,包括:

[0221] 对放大处理后的卷积结果进行分类处理,得到图像的语义分割结果。

[0222] 基于第二特征获得分类结果的过程需要应该一次卷积处理和一次放大处理,具体地,先进行放大或先进行卷积都可以,本实施例先进行放大处理,先放大后卷积可以使特征参数增加,参数越多,成本越大,得到的分割效果会越好,但同样的,由于参数增加,会导致显存消耗变大,运行较慢。

[0223] 在一个或多个可选的实施例中,本公开上述实施例中各实施例通过语义分割网络实现,该实施例方法还包括:

[0224] 利用语义分割网络对样本图像进行处理,获得样本图像的第一语义分割预测结果;

[0225] 利用语义编码损失网络,获得样本图像的第二语义分割预测结果;

[0226] 基于第一预测语义分割结果、第二语义分割预测结果和样本图像的标注语义分割信息,训练语义分割网络。

[0227] 在本申请上述实施例的语义分割方法之前,还可以基于样本场景图像,结合语义编码损失网络,采用误差反向传播方法,对初始设置的语义分割网络进行训练,获得本申请上述各实施例中使用的语义分割网络,语义编码损失网络实现约束对语义分割网络的训练。在一个可选示例中,对语义分割网络进行网络训练操作可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的网络训练单元执行。

[0228] 在一个或多个可选实施例中,利用语义分割网络对样本图像进行处理,获得样本图像的第一语义分割预测结果,包括:

[0229] 利用语义分割网络对样本图像进行处理,获得样本图像的样本图像特征;

[0230] 利用语义分割网络对样本图像特征进行处理,得到样本图像的第一语义分割结果;

[0231] 具体地,通过语义分割网络获得的样本图像特征与上述各实施例获得的第二图像特征是基于相同处理过程得到的。

[0232] 利用语义编码损失网络,获得样本图像的第二语义分割预测结果,包括:

[0233] 利用语义编码损失网络对样本图像特征进行编码处理,得到样本图像特征的编码结果;

[0234] 利用语义编码损失网络对样本图像特征的编码结果进行处理,得到样本图像的第二语义分割预测结果。在语义编码损失网络中,不同的物体类别具有相同的权重。

[0235] 具体地,利用语义编码损失网络的全连接层和响应函数,对编码结果进行处理,得到对应样本图像的第二语义分割预测结果。

[0236] 语义编码损失网络(SemanticEncodingLoss)使用很小的计算量约束网络理解全局语义信息。利用语义分割网络得到的编码结果,利用一个小的全连接层和一个响应函数,得到中间特征,学习某个物体类别在该场景中是否存在的一系列二分类问题。因为该损失函数相比较于语义分割网络获得的图像分割损失函数,对图像各个类别权重相等,因此该损失函数会对小物体提供相对较大的学习率,提升了小物体的性能。

[0237] 在一个或多个可选的实施例中,基于第一预测语义分割结果、第二语义分割预测结果和样本图像的标注语义分割信息,训练语义分割网络,包括:

[0238] 基于第一预测语义分割结果和样本图像的标注语义分割信息得到第一差异,基于第二预测语义分割结果和样本图像的标注语义分割信息得到第二差异;

[0239] 根据第一差异和第二差异,调整语义分割网络中的参数。

[0240] 本实施例结合第一差异和第二差异调整语义分割网络中的参数,具体地,调整语义分割网络中的参数,除了各卷积层和全连接层中的参数,还包括获得编码结果的字典和对应字典的权重;通过训练获得的语义分割网络可实现对图像输出全局信息,实现对图像的场景分析。

[0241] 可选地,根据第一差异和所述第二差异,调整语义分割网络中的参数,包括:

[0242] 基于第一差异与第二差异的和,得到第三差异;

[0243] 基于第三差异,调整语义分割网络中的参数。

[0244] 可选地,为了将第二差异应用到误差反向传播方法实现对语义分割网络进行网络训练,可以基于第一差异和第二差异求和得到第三差异,还可以对第一差异和第二差异分别分配不同的权重值,根据不同的权重值分别于第一差异和第二差异求积后再求和获得第三差异,根据反向传播第三差异,实现结合语义编码损失网络对语义分割网络进行训练。

[0245] 图2为本公开实施例提供的语义分割网络以及语义分割网络的训练流程的一个具体示例的示意图。如图2所示,语义分割网络是通过在卷积神经网络中添加语义编码模块得到的,其中,在该例子中,该语义编码模块位于卷积神经网络的最后一个卷积层之前,但本公开实施例不限于此。该语义编码模块包括编码层、全连接层、激活函数层以及卷积层,其中,在图2中,全连接层和激活函数层通过标有FC的箭头表示。此时,该语义编码网络可以包括卷积神经网络的第一部分、语义编码模块和卷积神经网络的第二部分,其中,卷积神经网络的第二部分包括卷积层及上采样层。具体地,卷积神经网络的第一部分可以包括多个卷积层,可以用于对图像进行卷积处理,得到N个 $H \times W \times C$ 大小的特征图,即第一图像特征,语义编码模块基于特征图获得缩放因子向量,并将缩放因子向量与特征图执行点乘操作,得

到第二图像特征；卷积神经网络的第二部分利用卷积层及上采样层对第二图像特征进行处理，得到语义分割结果。

[0246] 如图2所示，语义编码损失网络包括全连接层和响应函数；利用语义编码损失网络的全连接层和响应函数，对语义编码模块中输出的编码结果进行处理，得到对应图像中各类别权重相等的语义分割预测结果。

[0247] 在一些实施例中，可以结合语义编码损失网络训练语义分割网络，该语义分割网络包括第一网络(图中CNN)，经第一网络处理获得N个 $H \times W \times C$ 大小的特征图，其中，H、W、C分别为特征向量的长、宽、通道数(channel)；将第一网络获得的特征图输入语义编码模块(图中context encoding module语义编码模块)，语义编码模块基于特征图获得缩放因子向量，并将缩放因子向量与特征图执行点乘，得到第二图像特征；第二特征图输入第二网络(图中CONV和上采样)，第二网络经过卷积和上采样操作得到场景分析结果；语义编码损失网络从编码网络接收余量编码，经过语义编码损失网络(图中SE-loss)的全连接层(图中fc)和响应函数，输出对应样本场景图像中各类别权重相等的第二场景分类预测结果。

[0248] 应理解，图2所示的例子仅为了帮助本领域技术人员更好地理解本公开的技术方案，而不应理解成对本公开技术方案的限制，本领域普通技术人员在图2的基础上进行的各种变换，均属于本公开的技术方案。

[0249] 本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0250] 图3为本公开语义分割装置一些实施例的结构示意图。该实施例的装置可用于实现本公开上述各方法实施例。如图4所示，该实施例的装置包括：

[0251] 第一处理单元31，用于对图像进行处理，得到图像的第一图像特征和图像的全局语义信息。

[0252] 在本公开实施例中，可以对图像进行一种或多种处理，得到图像的第一图像特征，其中，在一个或多个可选的实施例中，可以对图像进行卷积处理，例如：利用卷积神经网络(可以是任意预先训练的模型，如：ResNet50，ResNet101等基础网络)对图像进行卷积处理，得到图像的第一图像特征，其中，该卷积处理的次数可以为一次或多次，例如，利用卷积神经网络的多个卷积层依次对图像进行处理；。或者，也可以对图像进行卷积和其他类型的处理，得到图像的第一图像特征，本公开实施例对此不做限定。

[0253] 可选地，通过卷积神经网络得到的第一图像特征是高层语义特征，在具体示例中，高层语义特征可表示为多维特征向量。

[0254] 此外，还可以得到图像的全局语义信息，该全局语义信息也可以称为场景全局信息，可以包括该图像的场景相关的全局信息。在一些可选实施例中，全局语义信息包括与至少一个语义类别对应的至少一个缩放因子向量，其中，可选地，不同的语义类别可以对应于相同或不同的缩放因子向量，并且在至少一个语义类别的数量为多个的情况下，存在对应于不同缩放因子向量的至少两个语义类别。

[0255] 第二处理单元32，用于基于全局语义信息和第一图像特征，得到第二图像特征。

[0256] 可选地，该第一图像特征可以包含类别相关的信息。

[0257] 在一个或多个可选的实施例中,可以对全局语义信息和第一图像特征进行卷积处理,得到第二图像特征。

[0258] 可选地,可以将第一图像特征与全局语义信息中的缩放因子向量的点积结果确定为第二图像特征。

[0259] 分割单元33,用于基于第二图像特征,获得图像的语义分割结果。

[0260] 其中,语义分割结果可以包括图像中至少一个像素中每个像素所属的语义类别,其中,该至少一个像素可以为该图像的部分或所有像素,例如,该语义分割结果可以包括图像中所有像素的类别,但本公开实施例对此不做限定。

[0261] 基于本公开上述实施例提供的一种语义分割装置,对图像进行处理,得到图像的第一图像特征和图像的全局语义信息,基于全局语义信息和第一图像特征,得到第二图像特征,基于第二图像特征,获得图像的语义分割结果,能够基于全局语义信息得到图像的语义分割结果,从而有利于提高语义分割性能。

[0262] 在一个或多个可选的实施例中,特征处理模块,包括:

[0263] 编码模块,用于对第一图像特征进行编码处理,得到第一图像特征的编码结果;

[0264] 语义模块,用于基于第一图像特征的编码结果,得到全局语义信息。

[0265] 可选地,编码模块,包括:

[0266] 字典确定模块,用于确定字典和字典对应的权重,字典包括至少一个字典基;

[0267] 编码结果模块,用于基于字典和字典对应的权重,对第一图像特征进行编码处理,得到第一图像特征的编码结果。

[0268] 可选地,语义模块,具体用于对第一图像特征的编码结果进行全连接处理,得到图像向量;利用激活函数对图像向量进行处理,得到全局语义信息。

[0269] 可选地,第一图像特征包括至少一个特征图;

[0270] 编码结果模块,具体用于分别计算至少一个特征图中的每个特征图与字典的字典基的对应元素的差值,并基于差值获得编码结果。

[0271] 在一个或多个可选的实施例中,当第一处理单元31输出的第一图像特征为 N 个 $H \times W \times C$ 大小的特征图,其中, H 、 W 、 C 分别为特征向量的长、宽、通道数(channel)。将该第一图像特征表示为矩阵为 $X = \{X_1, \dots, X_N\}$, X_i 表示一个特征图(i 的取值范围为1至 N),对应的学习到的字典表示为 $D = \{d_1, \dots, d_k\}$, d_j 表示一个字典基(j 的取值范围为1至 k),字典 D 包括 K 组字典基,其中, K 可以等于或不等于 N ,对应字典的权重 $S = \{s_1, \dots, s_k\}$,每个字典基对应一个权重值;

[0272] 基于字典基和特征图计算差值可以通过上述公式(1)实现。

[0273] 本公开语义分割方法的一些实施例中,分割单元33,包括:

[0274] 卷积模块,用于对第二图像特征执行卷积处理,得到卷积结果;

[0275] 分类模块,用于对卷积结果进行分类处理,得到图像的语义分割结果。

[0276] 可选地,该第二图像特征的网络层可以位于语义分割网络的最后一个卷积层之前。此时,可以对第二图像特征执行仅一次卷积处理,得到卷积结果。

[0277] 基于第二图像特征经过卷积得到的分类结果,实现对图像基于语义的像素分类,具体的,可获得图像中哪些像素对应哪些类别,并且,由于第二图像是基于全局语义信息和第一图像特征获得的,实现了对图像中某些类别(例如小物体类别)的强调,使图像中的所

有类别都能够被准确的识别。

[0278] 在一个或多个可选的实施例中,分割单元33,还包括:

[0279] 特征放大模块,用于对第二图像特征进行放大处理,以将第二图像特征的维度放大至第一图像特征的维度;

[0280] 卷积模块,具体用于对放大处理后的第二图像特征执行卷积处理,得到卷积结果。

[0281] 由于第二图像特征是第一图像特征与缩放因子向量执行点乘获得的,此时第二图像特征的大小与第一图像特征的大小不同,为了实现场景图像中的各像素进行分类,将第二图像特征放大至第一图像特征的大小,具体可选地可以通过上采样层或其他方式实现特征放大,以实现输出对应各像素的分类结果。

[0282] 具体地,可采用深度神经网络,该深度神经网络至少包括卷积层和分类层,将第二图像特征作为一个特征图输入该深度神经网络,得到对应该第二图像特征的分类结果,及对应图像的语义分割结果。通过深度神经网络中的卷积层对各第二图像特征进行降维度(也称为:降维)处理。在一个可选示例中,该操作可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的深度神经网络或该深度神经网络中的卷积层执行。相应地,后续操作中,可以结合第一图像特征与上述至少一个降维后的第二图像特征,对图像进行场景分析。

[0283] 上述实施例还存在另一种情况,分割单元33,还包括:

[0284] 结果放大模块,用于对卷积结果进行放大处理,以将卷积结果的维度放大至第一图像特征的维度;

[0285] 分类模块,具体用于对放大处理后的卷积结果进行分类处理,得到图像的语义分割结果。

[0286] 分割单元33中仅包括结果放大模块或特征放大模块中的一个。

[0287] 基于第二特征获得分类结果的过程需要应该一次卷积处理和一次放大处理,具体地,先进行放大或先进行卷积都可以,本实施例先进行放大处理,先放大后卷积可以使特征参数增加,参数越多,成本越大,得到的分割效果会越好,但同样的,由于参数增加,会导致显存消耗变大,运行较慢。

[0288] 在一个或多个可选的实施例中,本公开上述实施例中各实施例装置基于语义分割网络实现,该实施例装置还包括:

[0289] 第一预测单元,利用语义分割网络对样本图像进行处理,获得样本图像的第一语义分割预测结果;

[0290] 第二预测单元,用于利用语义编码损失网络,获得样本图像的第二语义分割预测结果;

[0291] 第一训练单元,用于基于第一预测语义分割结果、第二语义分割预测结果和样本图像的标注语义分割信息,训练语义分割网络。

[0292] 在本申请上述实施例的语义分割之前,还可以基于样本场景图像,结合语义编码损失网络,采用误差反向传播方法,对初始设置的语义分割网络进行训练,获得本申请上述各实施例中使用的语义分割网络,语义编码损失网络实现约束对语义分割网络的训练。在一个可选示例中,对语义分割网络进行网络训练操作可以由处理器调用存储器存储的相应指令执行,也可以由被处理器运行的网络训练单元执行。

[0293] 在一个或多个可选实施例中,第一预测单元,具体用于利用语义分割网络对样本图像进行处理,获得样本图像的样本图像特征;

[0294] 利用语义分割网络对样本图像特征进行处理,得到样本图像的第一语义分割结果;

[0295] 第二预测单元,包括:

[0296] 样本编码模块,用于利用语义编码损失网络对样本图像特征进行编码处理,得到样本图像特征的编码结果;

[0297] 样本处理模块,用于利用语义编码损失网络对样本图像特征的编码结果进行处理,得到样本图像的第二语义分割预测结果。

[0298] 具体地,在语义编码损失网络中,不同的物体类别具有相同的权重。

[0299] 可选地,样本处理模块,具体用于利用语义编码损失网络中的全连接层和响应函数对编码结果进行处理,得到样本图像的第二语义分割预测结果。

[0300] 在一个或多个可选实施例中,第一训练单元,包括:

[0301] 差异模块,用于基于第一预测语义分割结果和样本图像的标注语义分割信息得到第一差异,基于第二预测语义分割结果和样本图像的标注语义分割信息得到第二差异;

[0302] 参数调整模块,用于根据第一差异和所述第二差异,调整语义分割网络中的参数。

[0303] 具体地,参数调整模块,具体用于基于第一差异与第二差异的和,得到第三差异;基于第三差异,调整语义分割网络中的参数。

[0304] 图4为本公开语义分割网络的训练方法一个实施例的流程示意图。如图4所示,该实施例方法包括:

[0305] 步骤410,利用语义分割网络对样本图像进行处理,获得样本图像的第一语义分割预测结果。

[0306] 在一个或多个实施例中,该语义分割网络可以采用现有技术中常用的具有语义分割功能的网络,例如:去除编码层的上述实施例提供的语义分割方法中的网络。

[0307] 步骤420,利用语义编码损失网络,获得样本图像的第二语义分割预测结果。

[0308] 其中,语义编码损失网络包括编码层。

[0309] 步骤430,基于第一语义分割预测结果、第二语义分割预测结果和样本图像的标注语义分割信息,训练语义分割网络。

[0310] 基于本公开上述实施例提供的一种语义分割网络的训练方法,通过结合语义编码损失网络训练语义分割网络,使得到的语义分割网络对图像中的大物体和小物体具有同等权重,实践中能够有效提升小物体预测的性能。

[0311] 在一个或多个可选的实施例中,操作410包括:

[0312] 利用语义分割网络对样本图像进行处理,得到样本图像的样本图像特征;

[0313] 利用语义分割网络对样本图像进行处理,得到样本图像的第一语义分割结果。

[0314] 操作420,包括:

[0315] 利用语义编码损失网络的编码层对样本图像特征进行编码处理,得到样本图像特征的编码结果;

[0316] 该图像特征的编码结果的具体获取过程与本公开语义分割方法上述各实施例中基于第一图像特征获得编码结果的过程相同。

- [0317] 可选地,确定字典和字典对应的权重,其中,字典包括至少一个字典基;
- [0318] 基于字典和字典对应的权重,对样本图像特征进行编码处理,得到样本图像特征的编码结果。
- [0319] 该实施例中字典的定义与上述实施例相同,初始随机生成,经过训练获得字典及字典对应的权重。
- [0320] 可选地,样本图像特征可以包括至少一个样本特征图;
- [0321] 基于字典和字典对应的权重,对样本图像特征进行编码处理,得到样本图像特征的编码结果,可以包括:
- [0322] 分别计算至少一个样本特征图中的每个样本特征图与字典的每个字典基的对应元素的差值,并基于差值获得编码结果。
- [0323] 在一个或多个可选的实施例中,当操作110输出的第一图像特征为 N 个 $H \times W \times C$ 大小的特征图,其中, H 、 W 、 C 分别为特征向量的长、宽、通道数(channel)。将该第一图像特征表示为矩阵为 $X = \{X_1, \dots, X_N\}$, X_i 表示一个特征图(i 的取值范围为1至 N),对应的学习到的字典表示为 $D = \{d_1, \dots, d_k\}$, d_j 表示一个字典基(j 的取值范围为1至 k),字典 D 包括 K 组字典基,其中, K 可以等于或不等于 N ,对应字典的权重 $S = \{s_1, \dots, s_k\}$,每个字典基对应一个权重值;
- [0324] 基于字典基和特征图计算差值可以通过上述公式(1)实现,基于 $N \times K$ 个差值获得编码结果。
- [0325] 利用语义编码损失网络对样本图像特征的编码结果进行处理,得到样本图像的第二语义分割预测结果。
- [0326] 在一个或多个可选的实施例中,语义编码损失网络还包括全连接层和响应函数层;
- [0327] 利用语义编码损失网络对样本图像特征的编码结果进行处理,得到样本图像的第二语义分割预测结果,包括:
- [0328] 利用语义编码损失网络中的全连接层和响应函数层对编码结果进行处理,得到样本图像的第二语义分割预测结果。
- [0329] 语义编码损失网络(SemanticEncodingLoss)使用很小的计算量约束网络理解全局语义信息。利用语义分割网络得到的编码结果,利用一个小的全连接层和一个响应函数层,得到中间特征,学习某个物体类别在该场景中是否存在的一系列二分类问题。因为该损失函数相比较于语义分割网络获得的图像分割损失函数,对图像各个类别权重相等,因此该损失函数会对小物体提供相对较大的学习率,提升了小物体的性能。
- [0330] 在一个或多个可选的实施例中,基于第一语义分割预测结果、第二语义分割预测结果和样本图像标注的语义分割信息,训练语义分割网络,包括:
- [0331] 基于第一语义分割预测结果和样本图像标注的语义分割信息得到第一差异,基于第二语义分割预测结果和样本图像标注的语义分割信息得到第二差异;
- [0332] 根据第一差异和所述第二差异,调整语义分割网络中的参数。
- [0333] 本实施例结合第一差异和第二差异调整语义分割网络中的参数,具体地,调整语义分割网络中的参数,除了各卷积层和全连接层中的参数,还包括获得编码结果的字典和对应字典的权重;通过训练获得的语义分割网络可实现对图像输出全局信息,实现对图像的场景分析。

[0334] 可选地,根据第一差异和所述第二差异,调整语义分割网络中的参数,包括:

[0335] 基于第一差异与第二差异的和,得到第三差异;

[0336] 基于第三差异,调整语义分割网络中的参数。

[0337] 可选地,为了将第二差异应用到误差反向传播方法实现对语义分割网络进行网络训练,可以基于第一差异和第二差异求和得到第三差异,还可以对第一差异和第二差异分别分配不同的权重值,根据不同的权重值分别于第一差异和第二差异求积后再求和获得第三差异,根据反向传播第三差异,实现结合语义编码损失网络对语义分割网络进行训练。

[0338] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0339] 图5为本公开语义分割网络的训练装置实施例的结构示意图。该实施例的装置可用于实现本公开上述各方法实施例。如图5所示,该实施例的装置包括:

[0340] 第三预测单元51,用于利用语义分割网络对样本图像进行处理,获得样本图像的第一语义分割预测结果。

[0341] 在一个或多个实施例中,该语义分割网络可以采用现有技术中常用的具有语义分割功能的网络,例如:去除编码层的上述实施例提供的语义分割方法中的网络。

[0342] 第四预测单元52,用于利用语义编码损失网络,获得样本图像的第二语义分割预测结果,其中,语义编码损失网络包括编码层。

[0343] 第二训练单元53,用于基于第一语义分割预测结果、第二语义分割预测结果和样本图像的标注语义分割信息,训练语义分割网络。

[0344] 基于本公开上述实施例提供的一种语义分割网络的训练方法,通过结合语义编码损失网络训练语义分割网络,使得到的语义分割网络对图像中的大物体和小物体具有同等权重,实践中能够有效提升小物体预测的性能。

[0345] 在一个或多个可选的实施例中,第三预测单元51,具体用于利用语义分割网络对样本图像进行处理,得到样本图像的样本图像特征;利用语义分割网络对样本图像进行处理,得到样本图像的第一语义分割结果;

[0346] 第四预测单元52,包括:

[0347] 特征编码模块,用于利用语义编码损失网络的编码层对样本图像特征进行编码处理,得到样本图像特征的编码结果;

[0348] 语义预测模块,用于利用语义编码损失网络对样本图像特征的编码结果进行处理,得到样本图像的第二语义分割预测结果。

[0349] 可选地,特征编码模块,包括:

[0350] 权重字典模块,用于确定字典和字典对应的权重;字典包括至少一个字典基;

[0351] 样本结果模块,用于基于字典和字典对应的权重,对样本图像特征进行编码处理,得到样本图像特征的编码结果。

[0352] 可选地,样本图像特征包括至少一个样本特征图;

[0353] 样本结果模块,具体用于分别计算至少一个样本特征图中的每个样本特征图与字典的每个字典基的各对应元素的差值,并基于差值获得编码结果。

[0354] 在一个或多个可选的实施例中,语义编码损失网络还包括全连接层和响应函数层;

[0355] 语义预测模块,具体用于利用语义编码损失网络中的全连接层和响应函数层对编码结果进行处理,得到样本图像的第二语义分割预测结果。

[0356] 语义编码损失网络(SemanticEncodingLoss)使用很小的计算量约束网络理解全局语义信息。利用语义分割网络得到的编码结果,利用一个小的全连接层和一个响应函数层,得到中间特征,学习某个物体类别在该场景中是否存在的一系列二分类问题。因为该损失函数相比较于语义分割网络获得的图像分割损失函数,对图像各个类别权重相等,因此该损失函数会对小物体提供相对较大的学习率,提升了小物体的性能。

[0357] 在一个或多个可选的实施例中,第二训练单元52,包括:

[0358] 差异获得模块,用于基于第一语义分割预测结果和样本图像标注的语义分割信息得到第一差异,基于第二语义分割预测结果和样本图像标注的语义分割信息得到第二差异;

[0359] 网络参数模块,用于根据第一差异和第二差异,调整语义分割网络中的参数。

[0360] 本实施例结合第一差异和第二差异调整语义分割网络中的参数,具体地,调整语义分割网络中的参数,除了各卷积层和全连接层中的参数,还包括获得编码结果的字典和对应字典的权重;通过训练获得的语义分割网络可实现对图像输出全局信息,实现对图像的场景分析。

[0361] 可选地,网络参数模块,具体用于基于第一差异与第二差异的和,得到第三差异;基于第三差异,调整语义分割网络中的参数。

[0362] 根据本公开实施例的另一个方面,提供了一种电子设备,包括处理器,处理器包括本公开上述任一实施例的语义分割装置或语义分割网络的训练装置。

[0363] 根据本公开实施例的另一个方面,提供了一种电子设备,包括:存储器,用于存储可执行指令;

[0364] 以及处理器,用于与存储器通信以执行可执行指令从而完成本公开预测网络性能的方法上述任一实施例的操作。

[0365] 根据本公开实施例的另一个方面,提供了一种计算机存储介质,用于存储计算机可读的指令,指令被执行时执行本公开语义分割方法或语义分割网络的训练方法上述任一实施例的操作。

[0366] 根据本公开实施例的另一个方面,提供了一种计算机程序,包括计算机可读代码,当计算机可读代码在设备上运行时,设备中的处理器执行本公开语义分割方法或语义分割网络的训练方法中各步骤的指令。

[0367] 根据本公开实施例的再一个方面,提供了一种计算机程序产品,用于存储计算机可读指令,所述指令被执行时使得计算机执行上述任一可能的实现方式中所述的语义分割方法或语义分割网络的训练方法。

[0368] 在一个或多个可选实施方式中,本公开实施例还提供了一种计算机程序程序产品,用于存储计算机可读指令,所述指令被执行时使得计算机执行上述任一实施例中所述的语义分割方法或语义分割网络的训练方法。

[0369] 该计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选例

子中,所述计算机程序产品具体体现为计算机存储介质,在另一个可选例子中,所述计算机程序产品具体体现为软件产品,例如软件开发包(Software Development Kit,SDK)等等。

[0370] 根据本公开实施例还提供了另一种语义分割方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品,其中,该方法包括:对图像进行处理,得到所述图像的第一图像特征和所述图像的全局语义信息;基于所述全局语义信息和所述第一图像特征,得到第二图像特征;基于所述第二图像特征,获得所述图像的语义分割结果。

[0371] 根据本公开实施例还提供了另一种语义分割网络的训练方法及其对应的装置和电子设备、计算机存储介质、计算机程序以及计算机程序产品,其中,该方法包括:利用语义分割网络对样本图像进行处理,获得所述样本图像的第一语义分割预测结果;利用语义编码损失网络,获得所述样本图像的第二语义分割预测结果,其中,所述语义编码损失网络包括编码层;基于所述第一语义分割预测结果、第二语义分割预测结果和所述样本图像的标注语义分割信息,训练所述语义分割网络。

[0372] 在一些实施例中,该语义分割或语义分割网络的训练指示可以具体为调用指令,第一装置可以通过调用的方式指示第二装置执行语义分割或语义分割网络的训练,相应地,响应于接收到调用指令,第二装置可以执行上述语义分割方法或语义分割网络的训练方法中的任意实施例中的步骤和/或流程。

[0373] 应理解,本公开实施例中的“第一”、“第二”等术语仅仅是为了区分,而不理解成对本公开实施例的限定。

[0374] 还应理解,在本公开中,“多个”可以指两个或两个以上,“至少一个”可以指一个、两个或两个以上。

[0375] 还应理解,对于本公开中提及的任一部件、数据或结构,在没有明确限定或者在前后文给出相反启示的情况下,一般可以理解为一个或多个。

[0376] 还应理解,本公开对各个实施例的描述着重强调各个实施例之间的不同之处,其相同或相似之处可以相互参考,为了简洁,不再一一赘述。

[0377] 本公开实施例还提供了一种电子设备,例如可以是移动终端、个人计算机(PC)、平板电脑、服务器等。下面参考图6,其示出了适于用来实现本申请实施例的终端设备或服务器的电子设备600的结构示意图:如图6所示,计算机系统600包括一个或多个处理器、通信部等,所述一个或多个处理器例如:一个或多个中央处理单元(CPU)601,和/或一个或多个图像处理器(GPU)613等,处理器可以根据存储在只读存储器(ROM)602中的可执行指令或者从存储部分608加载到随机访问存储器(RAM)603中的可执行指令而执行各种适当的动作和处理。通信部612可包括但不限于网卡,所述网卡可包括但不限于IB(Infiniband)网卡。

[0378] 处理器可与只读存储器602和/或随机访问存储器630中通信以执行可执行指令,通过总线604与通信部612相连、并经通信部612与其他目标设备通信,从而完成本申请实施例提供的任一项方法对应的操作,例如,对图像进行处理,得到图像的第一图像特征和图像的全局语义信息;基于全局语义信息和第一图像特征,得到第二图像特征;基于第二图像特征,获得图像的语义分割结果。

[0379] 此外,在RAM 603中,还可存储有装置操作所需的各种程序和数据。CPU601、ROM602以及RAM603通过总线604彼此相连。在有RAM603的情况下,ROM602为可选模块。RAM603存储可执行指令,或在运行时向ROM602中写入可执行指令,可执行指令使处理器601执行上述通

信方法对应的操作。输入/输出(I/O)接口605也连接至总线604。通信部612可以集成设置,也可以设置为具有多个子模块(例如多个IB网卡),并在总线链接上。

[0380] 以下部件连接至I/O接口605:包括键盘、鼠标等的输入部分606;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分607;包括硬盘等的存储部分608;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分609。通信部分609经由诸如因特网的网络执行通信处理。驱动器610也根据需要连接至I/O接口605。可拆卸介质611,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器610上,以便于从其上读出的计算机程序根据需要被安装入存储部分608。

[0381] 需要说明的,如图6所示的架构仅为一种可选实现方式,在具体实践过程中,可根据实际需要对上述图6的部件数量和类型进行选择、删减、增加或替换;在不同功能部件设置上,也可采用分离设置或集成设置等实现方式,例如GPU和CPU可分离设置或者可将GPU集成在CPU上,通信部可分离设置,也可集成设置在CPU或GPU上,等等。这些可替换的实施方式均落入本公开公开的保护范围。

[0382] 特别地,根据本公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括有形地包含在机器可读介质上的计算机程序,计算机程序包含用于执行流程图所示的方法的程序代码,程序代码可包括对应执行本申请实施例提供的方法步骤对应的指令,例如,对图像进行处理,得到图像的第一图像特征和图像的全局语义信息;基于全局语义信息和第一图像特征,得到第二图像特征;基于第二图像特征,获得图像的语义分割结果。在这样的实施例中,该计算机程序可以通过通信部分609从网络上被下载和安装,和/或从可拆卸介质611被安装。在该计算机程序被中央处理单元(CPU)601执行时,执行本申请的方法中限定的上述功能。

[0383] 可能以许多方式来实现本公开的方法和装置、设备。例如,可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法和装置、设备。用于方法的步骤的上述顺序仅是为了进行说明,本公开的方法的步骤不限于以上具体描述的顺序,除非以其它方式特别说明。此外,在一些实施例中,还可将本公开实施为记录在记录介质中的程序,这些程序包括用于实现根据本公开的方法的机器可读指令。因而,本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

[0384] 本公开的描述是为了示例和描述起见而给出的,而并不是无遗漏的或者将本公开限于所公开的形式。很多修改和变化对于本领域的普通技术人员而言是显然的。选择和描述实施例是为了更好说明本公开的原理和实际应用,并且使本领域的普通技术人员能够理解本公开从而设计适于特定用途的带有各种修改的各种实施例。

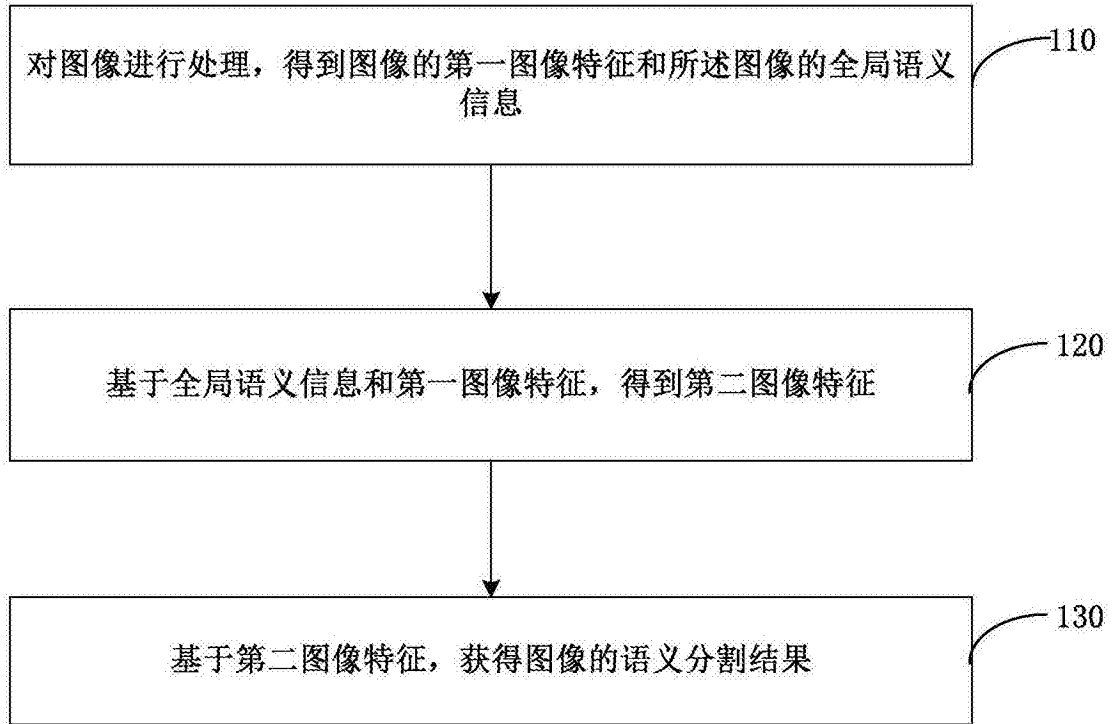


图1

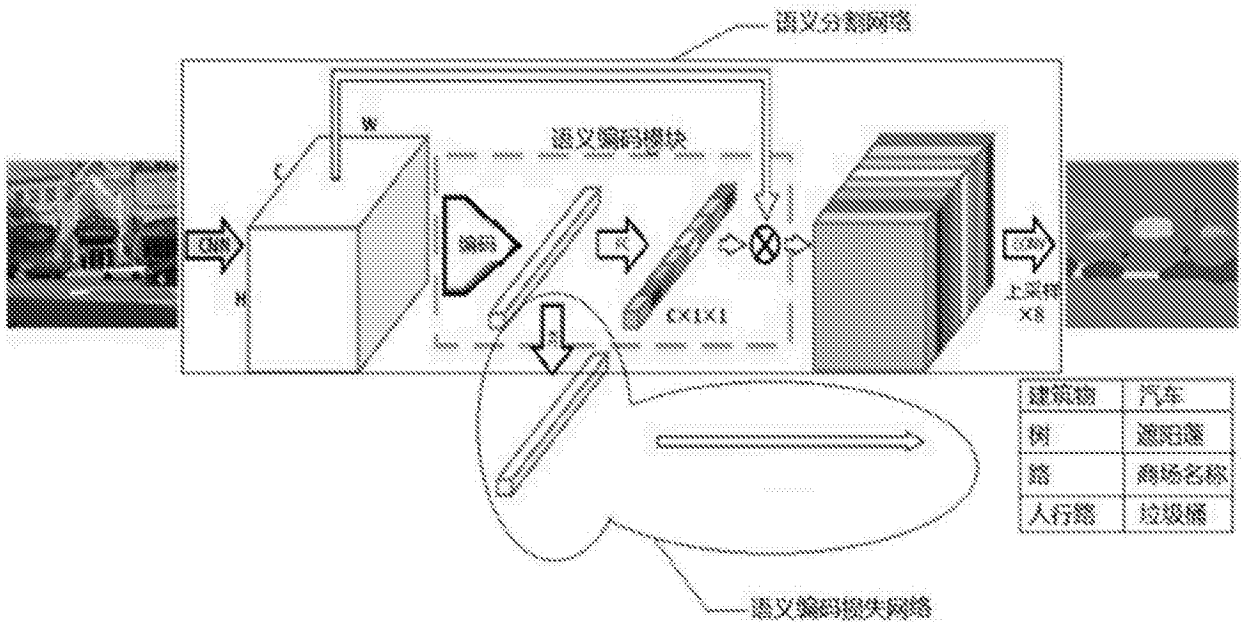


图2



图3

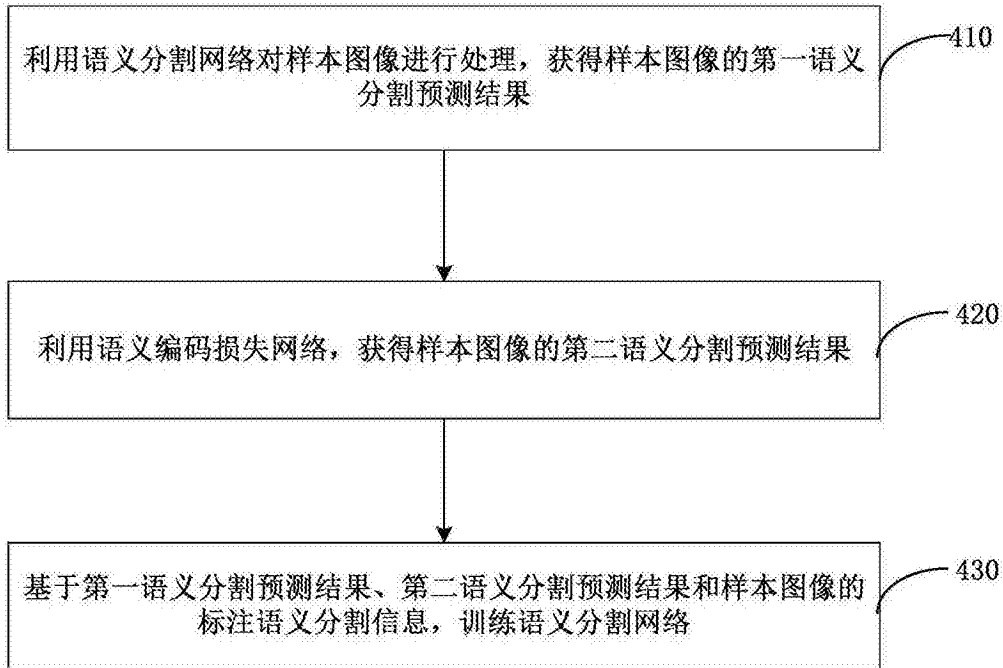


图4



图5

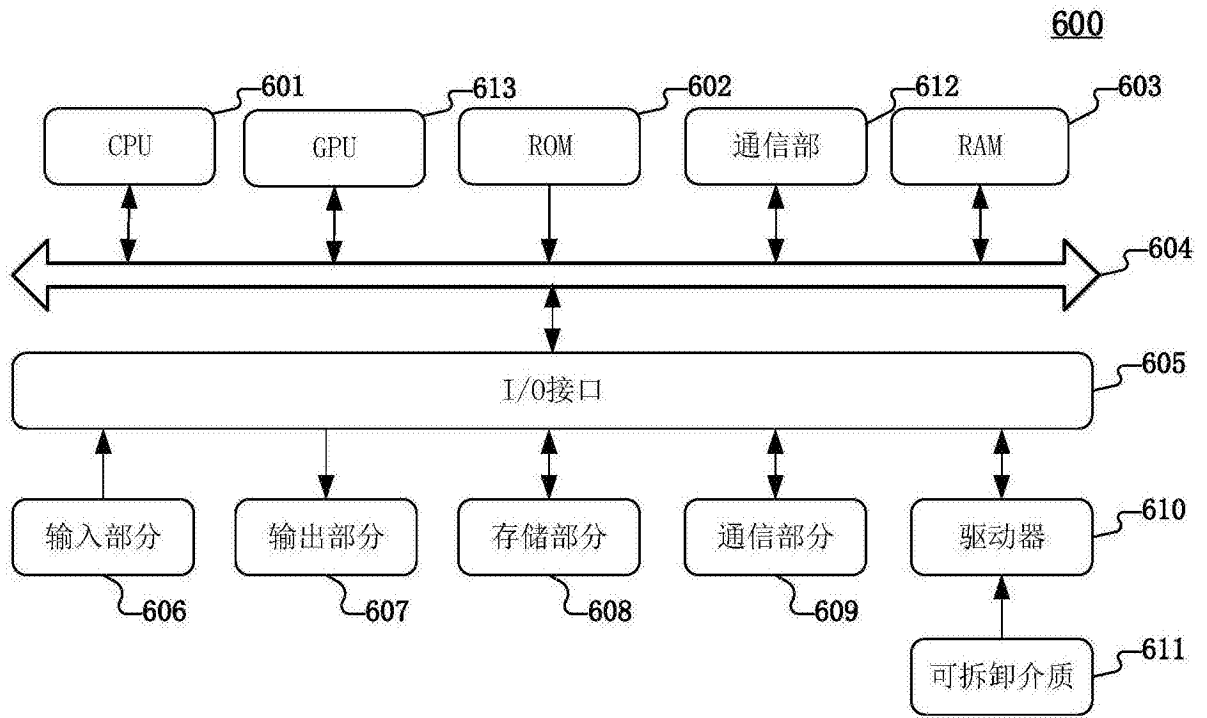


图6