



# (12) 发明专利申请

(10) 申请公布号 CN 115136192 A

(43) 申请公布日 2022. 09. 30

(21) 申请号 202180015890.1

(22) 申请日 2021.02.05

(30) 优先权数据

20158708.6 2020.02.21 EP

(85) PCT国际申请进入国家阶段日

2022.08.19

(86) PCT国际申请的申请数据

PCT/EP2021/052750 2021.02.05

(87) PCT国际申请的公布数据

W02021/165053 EN 2021.08.26

(71) 申请人 皇家飞利浦有限公司

地址 荷兰艾恩德霍芬市

(72) 发明人 N·佩佐蒂 D·马弗罗伊迪斯

(74) 专利代理机构 北京市金杜律师事务所

11256

专利代理师 丁君军

(51) Int.Cl.

G06T 7/00 (2017.01)

G06V 10/26 (2022.01)

G06V 10/44 (2022.01)

G06V 10/764 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

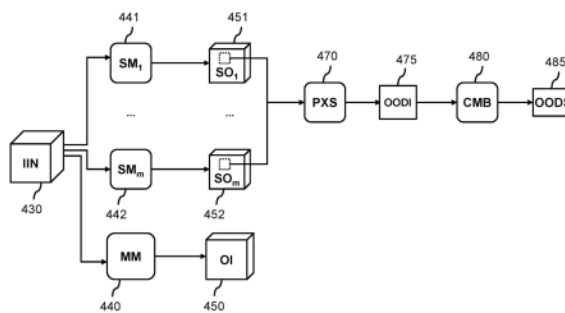
权利要求书3页 说明书13页 附图5页

## (54) 发明名称

对模型的输入实例的分布外检测

## (57) 摘要

本发明涉及一种用于对主模型的输入实例进行分布外(OOD)检测的系统(200)。主模型从输入实例生成输出图像。OOD检测使用在与训练主模型相同的训练数据集上训练的多个辅助模型。为了对输入实例执行OOD检测,为输入实例的辅助模型的输出图像确定每像素OOD得分。像素的像素OOD得分被确定为该像素在相应的辅助模型输出图像中的相应值之间的可变性。ID实例的这种可变性通常低于OOD实例,因此提供了输入实例是否为OOD的度量。将所确定的像素OOD得分组合成总OOD得分,该总OOD得分指示输入实例相对于训练数据集是否为OOD。



1. 一种用于对主模型的输入实例进行分布外OOD检测的系统(200),其中所述主模型在训练数据集上被训练,所述主模型被配置为从所述输入实例生成输出图像,其中分布外表示由高于预定义阈值得分的总像素OOD得分表示的、所述输入实例和所述训练数据集之间的不相似性,所述系统包括:

—数据接口(220),用于访问表示多个辅助模型的数据(040),所述多个辅助模型在对所述主模型的所述OOD检测中被使用,辅助模型在训练所述主模型的相同的所述训练数据集上被训练;

—处理器子系统(240),被配置为:

—获取所述主模型的输入实例;

—将相应的所述多个辅助模型应用于所述主模型的所述输入实例,以获取相应的辅助模型输出图像;

—确定所述相应的辅助模型输出图像的多个像素的多个像素OOD得分,像素的像素OOD得分被确定为所述像素在所述相应的辅助模型输出图像中的相应值之间的可变性;

—将所确定的所述多个像素OOD得分组合成总OOD得分,所述总OOD得分指示所述输入实例相对于所述训练数据集是否为OOD;

—基于所述总OOD得分生成输出信号,所述输出信号指示所述输入实例是否为OOD。

2. 根据权利要求1所述的系统(200),其中所述主模型被配置为从医学成像设备的输入数据确定所述输出图像,所述医学成像设备例如是CT扫描仪或MR扫描仪。

3. 根据权利要求2所述的系统(200),其中所述主模型是被配置为从由所述医学成像设备产生的信号来重建所述输出图像的医学图像重建模型。

4. 根据权利要求2所述的系统(200),其中所述主模型是被配置为确定定位输入图像中的病理的输出图像的医学图像分析模型。

5. 根据前述权利要求中任一项所述的系统(200),其中辅助模型在所述训练数据集的按比例缩减的训练输入实例和/或按比例缩减的训练输出图像上被训练。

6. 根据前述权利要求中任一项所述的系统(200),其中辅助模型包括比所述主模型更少的训练参数。

7. 根据前述权利要求中任一项所述的系统(200),其中所述处理器子系统还被配置为:至少在所述OOD得分没有指示所述输入实例为OOD的情况下:

—将所述主模型应用于所述输入实例以获取主模型输出图像;

—输出所述主模型输出图像。

8. 根据前述权利要求中任一项所述的系统(200),其中所述输出信号还指示所述辅助模型输出图像中对所述输入实例为OOD有贡献的一个或多个像素。

9. 根据任一前述权利要求所述的系统(200),还包括用于将所述输出信号输出到绘制设备的输出接口(260),所述绘制设备用于以感官可感知的方式将所述输出信号绘制给用户。

10. 一种用于使能对主模型的输入进行分布外OOD检测的系统(100),其中所述主模型在训练数据集上被训练,所述主模型被配置为从所述输入实例生成输出图像,其中分布外表示由高于预定义阈值得分的总像素OOD得分表示的、所述输入实例和所述训练数据集之间的不相似性,所述系统包括:

—数据接口(120),用于访问表示所述训练数据集的数据(030),所述主模型在所述训练数据集上被训练;

—处理器子系统(140),被配置为:

—训练多个辅助模型,辅助模型在训练所述主模型的所述训练数据集上被训练,辅助模型用于针对在所述OOD检测中使用的输入实例确定辅助模型输出图像;

—将所述多个辅助模型与所述主模型相关联以启用所述OOD检测。

11.根据权利要求10所述的系统(100),还被配置为:确定用于所述OOD检测的阈值OOD得分,并将所述阈值OOD得分与所述主模型相关联,所述阈值OOD得分基于对所述主模型的多个输入实例的多个OOD得分而被确定,输入实例的OOD得分通过以下方式被确定:

—应用相应的所述多个辅助模型,以获取相应的辅助模型输出图像;

—确定所述相应的辅助模型输出图像的多个像素的多个像素OOD得分,像素的像素OOD得分被确定为所述像素在所述相应的辅助模型输出图像中的相应值之间的可变性;

—将所确定的所述多个像素OOD得分组合成总OOD得分,所述总OOD得分指示所述输入实例相对于所述训练数据集是否为OOD。

12.根据权利要求10或11所述的系统(100),其中辅助模型通过初始化可训练模型的一组参数并基于所述初始化来优化所述一组参数而被训练,相应的辅助模型通过基于相应的随机初始化训练相同的所述可训练模型而被训练。

13.一种对主模型的输入实例进行分布外OOD检测的计算机实现的方法(500),其中所述主模型在训练数据集上被训练,所述主模型被配置为从所述输入实例生成输出图像,其中分布外表示由高于预定义阈值得分的总像素OOD得分表示的、所述输入实例和所述训练数据集之间的不相似性,所述方法包括:

—访问(510)表示多个辅助模型的数据,所述多个辅助模型在对所述主模型的所述OOD检测中使用,其中辅助模型在训练所述主模型的相同的所述训练数据集上被训练;

—获取(520)所述主模型的输入实例;

—将相应的所述多个辅助模型应用(530)于所述主模型的所述输入实例,以获取相应的辅助模型输出图像;

—确定(540)所述相应的辅助模型输出图像的多个像素的多个像素OOD得分,像素的像素OOD得分被确定为所述像素在所述相应的辅助模型输出图像中的相应值之间的可变性;

—将所确定的所述多个像素OOD得分组合(550)成总OOD得分,所述总OOD得分指示所述输入实例相对于所述训练数据集是否为OOD;

—基于所述总OOD得分生成(560)输出信号,所述输出信号指示所述输入实例是否为OOD。

14.一种使能对主模型的输入进行分布外OOD检测的计算机实现的方法(600),其中所述主模型在训练数据集上训练,所述主模型被配置为从输入实例生成输出图像,其中分布外表示所述输入实例和所述训练数据集之间的不相似性,所述方法包括:

—访问(610)表示所述训练数据集的数据,所述主模型在所述训练数据集上被训练;

—训练(620)多个辅助模型,辅助模型在训练所述主模型的所述训练数据集上被训练,辅助模型用于针对在所述OOD检测中使用的输入实例确定辅助模型输出图像;

—将所述多个辅助模型与所述主模型相关联(630)以启用所述OOD检测。

15. 一种计算机可读介质(700),包括表示以下一项或多项的瞬态或非瞬态数据(710):
- 指令,在由处理器系统执行时,所述指令使得所述处理器系统执行根据权利要求13和/或权利要求14所述的计算机实现的方法;
  - 多个辅助模型,所述辅助模型与主模型相关联以使能OOD检测,其中所述主模型在训练数据集上被训练,所述主模型被配置为从输入实例生成输出图像,所述辅助模型在训练所述主模型的所述训练数据集上被训练,辅助模型用于针对在所述OOD检测中使用的输入实例确定辅助模型输出图像。

## 对模型的输入实例的分布外检测

### 发明领域

[0001] 本发明涉及一种用于对模型的输入实例进行分布外检测(out-of-distribution)的系统,以及一种对应的计算机实现的方法。本发明还涉及一种用于使能分布外检测的系统,以及一种对应的计算机实现的方法。本发明还涉及一种计算机可读介质。

### 背景技术

[0002] 来自诸如MR和CT扫描仪之类的医学成像设备的数据处理流水线涉及各种复杂的图像处理操作。例如,来自扫描仪的原始数据可被转换成要供临床医生检查的图像,这是一项被称为图像重建的任务。作为另一示例,可在由医学成像设备产生的图像中识别并突出显示特定对象,这是一项被称为语义分割的任务。图像分析也可用于例如定位病理。通常,这类图像处理模型从输入实例(例如,输入图像)或以所谓的k空间表示(对于MR扫描仪)或以正弦图(对于CT扫描仪)表示的原始扫描仪数据生成输出图像。

[0003] 鉴于所涉及的各种图像处理操作的复杂性,将机器可学习模型用于医学图像处理变得越来越有吸引力。可在训练数据集上自动训练这类机器可学习模型以执行各种任务,并且有希望提供比手动设计的算法更好的结果。例如,在N.Pezzotti等人的“Adaptive-CS-Net:FastMRI with Adaptive Intelligence”(可在<https://arxiv.org/abs/1912.12259>获取并通过引用并入本文)中,提出了一种用于从下采样k空间数据重建MR图像的机器可学习模型,该模型表现出优于传统方法的性能。另一个应用是低剂量CT扫描的去噪。此外,在各种其它应用领域中,例如,自动驾驶和视觉监控,提供输出图像的图像处理模型,诸如分割模型,在质量上不断地提高,并且越来越多地部署在现实生活中。

[0004] 当在现实生活中应用这类机器可学习的图像处理模型时,特别是在诸如医学图像分析和自动驾驶的关键应用领域中,重要的是可依赖由图像处理模型提供的输出图像。然而,当机器可学习的图像处理模型被应用于与训练期间使用的数据太不相似的输入实例时,它们典型地不能保证提供准确的结果。太不相似的输入实例典型地被称为“分布外”(OOD)。此外,从模型的输出图像中可能无法直接观察到有什么不妥。例如,在医学成像中,已经观察到一些基于深度学习的图像处理模型可产生不同于真实解剖结构的逼真的、解剖学上看似合理的图像。因为人工制品看起来像真正的解剖结构,所以放射科医师可能无法标识它们。这可导致影响诊断的误解,降低对产品价值/质量的信心,和/或放射科医师的额外负担。

[0005] 不幸的是,各种已知的OOD检测技术并不适用于以图像(与例如分类相反)为输出的图像处理模型。因此,希望能够对图像处理模型的输入实例执行OOD检测。

### 发明内容

[0006] 各种实施例涉及估计要由产生输出图像的模型(例如,深度神经网络)处理的输入实例是否包含在用于训练的数据集中,并且因此,是否可预期该模型产生可靠的结果。

[0007] 根据本发明的第一方面,提出了一种用于对主模型的输入实例进行分布外(OOD)

检测的系统。根据本发明的另一方面,提出了一种对主模型的输入实例进行OOD检测的计算机实施的方法。主模型可被配置为从输入实例生成输出图像。主模型可为已经或正在训练数据集上训练的机器可学习模型。为了执行OOD检测,可使用多个辅助模型。可在与训练主模型的训练数据集相同的训练数据集上训练辅助模型。在给定输入实例的情况下,可对其应用辅助模型以获取相应的辅助模型输出图像。对于每个像素,像素OOD得分可被确定为该像素在相应的辅助模型输出图像中的相应值之间的可变性。可将多个像素OOD得分组合成总OOD得分,该总OOD得分指示输入实例相对于训练数据集是否为OOD。基于总OOD得分,可生成指示输入实例是否为OOD的输出信号。

[0008] 为了提供OOD检测,各个方面使用在与主模型相同的训练数据集上训练的多个辅助模型。多个辅助模型可一起被称为辅助模型的“集合”。训练数据集典型地包括多个训练输入实例和对应的训练输出图像。在给定训练输入实例的情况下,可训练主模型以产生对应的训练输出图像。可在相同的训练数据集上训练辅助模型,在这种意义上,它们可被训练以在给定相同的训练输入实例的情况下产生相同的训练输出图像,或者至少从输入产生输出(其中这些输出和输入与原始训练输出和输入相关,例如,通过按比例缩减、通道减少或另一图像处理操作)。辅助模型可使用与主模型相同的模型架构(可选地,主模型可为辅助模型中的一个),或其变体,例如,具有较少的可训练参数的简化。例如,可通过基于其一组参数的相应随机初始化训练相同的可训练模型来训练相应的辅助模型。

[0009] 通过在类似的数据上训练和/或通过使用与主模型类似的架构,在将辅助模型应用于类似于训练数据的输入实例时,可预期辅助模型表现出与主模型以及彼此基本相同的行为。因此,至少当类似于训练数据的输入实例,或者换句话说,在分布(ID)中的输入实例,被输入到辅助模型时,辅助模型通常可产生类似的结果。因此,辅助模型的输出图像的每像素方差通常较低。然而,当输入实例是分布外的(OOD)时,辅助模型没有被训练来为该输入实例提供类似的输出,并且由于它们是单独训练的,因此它们的输出图像通常具有较高的每像素方差。因此,输出图像的像素的每个单独的像素OOD得分可被认为是输入实例是否为OOD的度量,并且因此,输入实例为OOD的总体评估可通过将一个或多个像素的像素OOD得分组合成总OOD得分来获取。

[0010] 因此,可按有效和准确的方式对图像生成模型执行OOD检测。相对有限数目的辅助模型足以获取精确的总OOD得分,例如,至多20或甚至至多10或至多5个辅助模型。本发明人已经能够用五个辅助模型获取良好的结果。有效地,因为每个像素OOD得分本身可被认为是输入实例是否为OOD的度量,并且对于有限数目的辅助模型,足够的可用于确定输入实例的OOD状态。此外,通过使用具有较少的可训练参数和/或较小的输入和/或较小的输出的辅助模型,可限制相对于应用主模型计算OOD得分的开销。所提供的OOD检测技术的另一个优势在于它们在很大程度上是模型不可知(model-agnostic)的,例如,不依赖于主模型和/或辅助模型的特定模型架构。

[0011] 根据本发明的另一个方面,提出了一种用于使能对主模型的输入进行分布外(OOD)检测的系统。根据本发明的另一方面,提出了一种对应的计算机实施的方法。在该系统和方法中,可在已经训练主模型的同时训练数据集上训练多个辅助模型。然后,辅助模型可与主模型相关联以启用OOD检测。根据本发明的另一方面,提出了一种计算机可读介质,其包括表示用于执行如本文所描述的计算机实施的方法的指令中的一个或多个的瞬态或

非瞬态数据,或者与主模型相关联以使能如本文所描述的OOD检测的辅助模型。

[0012] 在各种实施例中,本文所描述的OOD检测技术可应用于医学图像处理。用于医学图像处理的各种模型本身在本领域中是已知的,并且可与所呈现的技术相组合。在医学图像处理中,主模型的输出图像可根据医学成像设备(例如CT扫描仪或MR扫描仪)的输入数据来确定。有趣的是,当使用来自MR扫描仪的数据时,能够进行如本文所描述的OOD检测可使得能够安全地使用下采样的k空间数据(例如,通过加速扫描获取的),因为如果输入实例不足以类似于已知的训练实例而被应用于主模型,则输出图像可能被标记为OOD。由于更少的数据必须在扫描仪中被采集,因此可减少检查时间,同时仍然获取可靠的结果。类似地,当使用来自CT扫描仪的数据时,OOD检测技术的使用可使得能够安全地施加具有可靠结果的较低剂量,从而导致对患者的较低辐射。

[0013] 一般而言,当使用来自医学成像设备的输入数据时,该输入数据可为由医学成像设备产生的信号。例如,对于MR扫描仪,可通过k空间表示信号,或者通过对k空间表示应用逆傅立叶变换来在图像空间中表示信号。类似地,对于CT扫描仪,信号可被表示为正弦图,或者通过将逆拉东变换应用于正弦图表示而在图像空间中被表示。在使用原始信号的这类情况下,主模型例如可为被配置为根据信号重建输入图像的医学图像重建模型。这类重建模型也可被称为去噪模型或包括去噪模型。例如,CT图像可从CT信号重建,或者MR图像可从MR信号重建,尽管该模型也可被训练为例如从MR信号重建CT图像或者从CT信号重建MR图像。代替使用信号作为输入,医学图像处理器模型还可对从这类信号重建的输入图像进行操作。

[0014] 主模型可为用于医学图像处理中或其它应用领域中的分割模型,例如语义分割模型。分割模型可被配置为指示输入实例中表示特定特性的一部分,该特性例如可以是位于图像中的对象或位于图像中的其它方面的类型。例如,特性可以是可存在于输入实例中的特定位置处的任何特定类型的对象(例如,囊肿、肿瘤、细胞核、淋巴细胞、坏死组织等),或者是存在于输入实例中的对象可具有的特定特性(例如,黑暗、噪声、尖峰等)。例如,主模型可为每个图像像素提供该像素是否属于该特性的指示。主模型还可被配置为针对输入实例的相应像素指示该像素位置处的特性(例如,细胞或细胞核密度、组织中的灌注等)的存在的相关量或范围。

[0015] 主模型还可为被配置为确定定位输入图像中的病理的输出图像的医学图像分析模型;这类模型可根据需要对重建的图像或对医学成像设备的信号进行操作。例如,主模型可为例如用于前列腺或卵巢肿瘤的肿瘤检测模型,或者被配置为将异常值分配给在输入图像中标识的肿瘤的部分的肿瘤分级模型。

[0016] 如本领域技术人员所理解的,本文所提供的技术不限于医学图像处理,并且还可用于图像处理的各种其它应用领域中,诸如自动驾驶或视频监控。

[0017] 在各个方面,输出信号可基于总OOD得分来生成,该总OOD得分指示输入实例是否为OOD。这类输出信号可按各种方式使用,例如由用户使用或用于进一步的自动化处理中。

[0018] 可选地,总OOD得分可用于决定是否将主模型应用于输入实例。在这类情况下,只有当OOD得分不指示输入实例为OOD时,主模型才可被应用于输入实例以获取主模型输出图像,并且主模型输出图像可例如以感官可感知的方式或数字地输出给用户以用于进一步的自动处理。然而,对于被指示为OOD的输入实例,也可将主模型应用于输入实例,例如,在这

类情况下,输出图像可与输入实例被确定为OOD的警告或错误一起被显示。

[0019] 可选地,输出信号还可指示输出图像中对输入实例为OOD有贡献的一个或多个像素,例如,具有最高像素OOD得分的像素子集。例如,所指示的像素可为其OOD得分超过阈值的像素,或者它们可为具有最高OOD得分的固定百分比或数目的像素。通过指示有贡献的像素,可提供有帮助的反馈以用于理解为什么输入实例被确定为OOD,并且因此理解可以如何对此进行补救。例如,通过在训练数据集中包含在被指示为有贡献的那些方面类似于输入实例的附加训练数据,可以进行补救。

[0020] 一般而言,输入实例为OOD可能有各种原因,因此,在OOD检测时可应用各种补救措施。现在论述可单独或组合使用的若干种原因和补救措施。

[0021] 输入实例为OOD的可能原因可为输入实例的质量太低。例如,如果输入实例非常类似于训练实例,则输入实例可表示“常规案例”,主模型仍然可确信地为其提供模型输出。在这类情况下,一般而言,辅助模型也可一致地提供类似的输出图像。然而,如果输入实例与训练数据集不太相似,则辅助模型可能发散,因此在主模型输出中的置信度较低。因此,输入实例可被确定为OOD。例如,输入实例可为来自CT扫描仪的数据,该CT扫描仪在过低的剂量下操作而不能可靠地为被扫描的特定受试者产生输出图像,或者来自MR扫描仪的数据,该MR扫描仪在过高的加速度下操作。因此,在一些实施例中,在确定输入实例为OOD时,可获取较高质量(例如,较高剂量或较低加速度)的输入实例的新测量结果(例如,新CT或MR扫描),对于该新测量结果,可再次确定总OOD得分。如果新的输入实例不为OOD,则可还使用该输入实例,例如呈现给用户或自动处理。可选地,对于多个质量设置可重复该过程。

[0022] 输入实例为OOD的另一可能原因可能是输入实例中的测量假象。例如,由于在医学成像设备中被扫描的受试者的运动可能存在的运动假象,或者金属假象等。因此,代替或除了以更高的质量执行新的测量结果以外,也可按相同的质量执行新的测量结果。如果新的输入实例不为OOD,则它可被进一步使用,而先前的输入实例作为假象被忽略。

[0023] 输入实例为OOD的另一可能原因在于由输入实例所测量的受试者是分布外的。例如,被扫描的受试者可能具有在数据集中未被充分呈现的病理。这种原因可通过确定表示相同受试者的多个输入实例为OOD,但其它受试者的输入实例不为OOD来识别。例如,当确定受试者为OOD时,可将受试者报告给用户,例如临床医生,用于进一步的分析。可选地,在检测到受试者为OOD时,可使用后备模型来确定输入实例的另一输出图像,并且可使用另一输出图像来代替主模型的输出图像。后备模型例如可为非可训练模型。例如,对于MR扫描,如果受试者为OOD,则可使用本身已知的基于SENSE的重建作为后备模型。

[0024] 输入实例为OOD的另一可能原因可能是持久的测量问题,例如测量设备中的缺陷或用于测量的错误环境条件。这种原因可通过确定表示多个受试者的输入实例为OOD来识别。在这类情况下,持久测量问题可被报告用于修复。

[0025] 本领域技术人员将理解,本发明的上述实施例、实施方式和/或可选方面中的两个或更多个可按任何被认为有用的方式进行组合。

[0026] 基于本说明书,本领域技术人员可对任何计算机实施的方法和/或任何计算机可读介质进行修改和变化,这些修改和变化对应于所描述的对应系统的修改和变化,反之亦然。

## 附图说明

[0027] 本发明的这些和其它方面将从以下描述中以示例的方式描述的实施例并参考附图而变得清楚且还阐明,在附图中:

[0028] 图1示出了用于使能对主模型的输入进行分布外(OOD)检测的系统;

[0029] 图2示出了用于对主模型的输入实例进行分布外(OOD)检测的系统;

[0030] 图3示出了与本文所描述的技术一起使用的模型的详细示例,在这种情况下是U-Net型模型;

[0031] 图4示出了如何确定输入实例的总OOD得分的详细示例;

[0032] 图5示出了对主模型的输入实例进行分布外(OOD)检测的计算机实施的方法;

[0033] 图6示出了使能对主模型的输入进行分布外(OOD)检测的计算机实施的方法;

[0034] 图7示出了包括数据的计算机可读介质。

[0035] 应注意,附图纯粹是示意性的而不是按比例绘制的。在附图中,与已经描述的元件相对应的元件可具有相同的附图标记。

## 具体实施方式

[0036] 图1示出了用于使能对主模型的输入进行分布外(OOD)检测的系统100。主模型可被配置为从输入实例生成输出图像。可在训练数据集上训练主模型。系统100可包括数据接口120和处理器子系统140,它们可经由数据通信121进行内部通信。数据接口120可用于访问表示训练数据集的数据030,在该训练数据集上训练主模型。数据接口120还可用于访问主模型和/或多个辅助模型040,如下面更详细地论述的。辅助模型040可用于根据本文所描述的方法(例如,由图2的系统200)的OOD检测。OOD检测的使能和OOD检测本身可组合在单个系统或方法中,例如,系统100和200可组合在单个系统中。系统,例如其处理器子系统,可还被配置为在训练数据集030上训练主模型。

[0037] 处理器子系统140可被配置为在系统100的操作期间且使用数据接口120来访问数据030、040。例如,如图1所示,数据接口120可提供对外部数据存储器021的访问122,该外部数据存储器021可包括所述数据030、040。另选地,可从作为系统100的一部分的内部数据存储设备访问数据030、040。另选地,可经由网络从另一实体接收数据030、041。一般而言,数据接口120可采取各种形式,诸如到局域网或广域网(例如因特网)的网络接口、到内部或外部数据存储设备的存储接口等。数据存储设备021可采取任何已知且合适的形式。

[0038] 处理器子系统140可被配置为在系统100的操作期间且使用数据接口120来训练多个辅助模型040。可在训练主模型的训练数据集030上训练辅助模型040。辅助模型040可用于确定在OOD检测中使用的输入实例的辅助模型输出图像。

[0039] 处理器子系统140还可被配置为将多个辅助模型040与主模型相关联以使能OOD检测。辅助模型可按任何适当的方式与主模型相关联,或者更具体地与表示主模型的数据相关联,例如通过将辅助模型包含在模型数据本身中,例如作为文件头、XML元件等,或者提供辅助模型作为单独的文件,或者以任何其它方式。例如,辅助模型040可存储在与主模型相同的数据容器中,例如存储在相同的(多个)文件中,但是也可作为与主模型相关联的单独的辅助模型来提供。例如,在一些实施例中,主模型可链接到辅助模型,例如通过含有辅助模型040可访问的URL,或者辅助模型040可链接到主模型。各种其它关联方式同样是可想到

的,并且在本领域技术人员的范围内。

[0040] 作为可选组件,系统100可包括输入接口(未示出),用于从传感器获取传感器数据,例如由诸如CT扫描仪或MR扫描仪的医学成像设备产生的信号。训练数据集的一个或多个训练输入实例可基于这类传感器数据。关于图2更详细地论述了经由输入接口获取输入实例,并且那里描述的选项也可应用于系统100中。被配置为测量传感器数据的传感器本身,例如CT扫描仪或MR扫描仪,可为系统100的一部分。

[0041] 将参考图3和4还阐述系统100的操作的各种细节和方面,包含其可选方面。

[0042] 一般而言,系统100可体现为单个设备或装置,或者体现在单个设备或装置中,诸如工作站,例如基于膝上型电脑或台式电脑,或者服务器。该设备或装置可包括执行适当软件的一个或多个微处理器。例如,处理器子系统可由单个中央处理单元(CPU)体现,也可由这类CPU和/或其它类型的处理单元的组合或系统体现。软件可被下载和/或存储在对应的存储器中,例如诸如RAM的易失性存储器或诸如闪存的非易失性存储器。另选地,系统的功能单元,例如数据接口和处理器子系统,可按可编程逻辑的形式在设备或装置中实施,例如作为现场可编程门阵列(FPGA)和/或图形处理单元(GPU)。一般而言,系统的每个功能单元可按电路的形式实施。注意,系统100还可按分布式方式实施,例如,涉及不同的设备或装置,诸如分布式服务器,例如以云计算的形式。

[0043] 图2示出了用于对主模型的输入实例进行分布外(OOD)检测的系统200。主模型可被配置为从输入实例生成输出图像。可在训练数据集上训练主模型。系统200可包括数据接口220和处理器子系统240,它们可经由数据通信221进行内部通信。数据接口220可用于访问表示在OOD检测中使用的多个辅助模型的数据040。可在与训练主模型的训练数据集相同的训练数据集上训练辅助模型。数据接口220还可用于访问主模型。辅助模型数据可从用于使能OOD检测的系统(例如,图1的系统100)获取,或者通过执行如本文所描述的用于使能OOD检测的方法获取。

[0044] 处理器子系统240可被配置为在系统200的操作期间且使用数据接口220来访问数据040。例如,如图2所示,数据接口220可提供对可包括所述数据040的外部数据存储器022的访问222。另选地,可从作为系统200的一部分的内部数据存储器访问数据040。另选地,可经由网络从另一实体接收数据040。一般而言,数据接口220可采取各种形式,诸如到局域网或广域网(例如因特网)的网络接口、到内部或外部数据存储设备的存储接口等。数据存储设备022可采取任何已知和合适的形式。

[0045] 处理器子系统240可被配置为在系统200的操作期间且使用数据接口220获取输入实例。处理器子系统240可还被配置为应用相应的多个辅助模型以获取相应的辅助模型输出图像。处理器子系统240可还被配置为确定相应的辅助模型输出图像的多个像素的多个像素OOD得分。像素的像素OOD得分可被确定为该像素在相应的辅助模型输出图像中的相应值之间的可变性。处理器子系统240可还将所确定的多个像素OOD得分组合成总OOD得分。总OOD得分可指示输入实例相对于训练数据集是否为OOD。

[0046] 处理器子系统240还可基于总OOD得分生成输出信号225,输出信号225指示输入实例是否为OOD。例如,输出信号可为总OOD得分本身,或者如果输入实例为OOD时发出的警报。输出信号还可指示输出图像中对输入实例为OOD有贡献的一个或多个像素,从而提供输入实例为什么是OOD的解释,该解释可被输出到绘制设备或另一软件组件以用于进一步的自

动化处理。

[0047] 可选地,至少在OOD得分不指示该输入实例为OOD的情况下,处理器子系统240可将主模型应用于输入实例以获取主模型输出图像,并且例如经由如下所述的输出接口将主模型输出图像输出给用户,或者输出给另一软件组件以进行进一步的自动处理。例如,为了确定所确定的OOD得分是否指示OOD状态,可将OOD得分与阈值OOD得分(例如,固定得分或与辅助模型相关联的得分)进行比较。即使输入实例在分布中,也未必将主模型应用于输入实例,例如,在一些实施例中,可确定数据集的多个输入实例的OOD得分以计算数据集和训练数据集之间的相似性,而不必将主模型应用于这些输入实例。

[0048] 作为可选组件,系统200可包括输入接口260,用于从传感器获取传感器数据223,例如,从诸如CT扫描仪或MR扫描仪的医学成像设备产生的信号224获取传感器数据223。图中示出了MR扫描仪072。例如,信号可为来自医学成像设备的原始信号,例如,表示在k空间中(在MR扫描仪的情况下)或表示为正弦图(在CT扫描仪的情况下),或者原始信号可以被变换到图像空间中,例如分别通过逆傅立叶变换或逆拉冬变换。

[0049] 例如,MR扫描仪可被配置为例如以至少为2,至少为4或至少为6的加速因子执行加速扫描。在这类情况下,使用适当的训练模型仍然可能成功地重建或分析MR数据,但是对于训练数据集中未充分表示的输入而言,存在错误结果的较高风险,这使得OOD检测特别重要。类似的考虑适用于在低剂量下操作的CT扫描仪的情况,例如,至多全剂量的一半或至多全剂量的四分之一。

[0050] 然而,各种其它类型的传感器和对应的信号224是可能的,例如,传感器可为产生图像的相机、产生视频的视频相机等。一般而言,输入接口260可被配置用于各种类型的传感器信号,例如,视频信号、雷达/激光雷达信号、超声波信号等。对其执行OOD检测的输入实例可基于传感器数据223,例如,输入实例等于传感器数据223,或者可对其执行可选的预处理步骤。被配置为测量信号224的传感器本身,例如CT扫描仪或MR扫描仪072,可为系统200的一部分。

[0051] 作为可选组件,系统200可包括显示器输出接口280或用于将输出信号225输出到诸如显示器290等绘制设备的任何其它类型的输出接口。例如,显示器输出接口280可生成用于显示器290的显示数据282,这使得显示器290以感官可感知的方式绘制输出信号(例如,作为屏幕上的可视化292)。例如,如果输入实例为OOD,则错误或警告消息可显示在主模型的模型输出旁边。在输入实例为OOD的情况下,对输入实例是OOD有贡献的辅助模型输出图像的特定像素也可被绘制,例如在主模型的输出图像上被突出显示。

[0052] 将参考图3和4还阐述系统200的操作的各种细节和方面,包含其可选方面。

[0053] 一般而言,系统200可体现为单个设备或装置,或者体现在单个设备或装置中,诸如工作站,例如基于膝上型电脑或台式电脑,或者服务器。该设备或装置可包括执行适当软件的一个或多个微处理器。例如,处理器子系统可由单个中央处理单元(CPU)体现,也可由这类CPU和/或其它类型的处理单元的组合或系统体现。软件可被下载和/或存储在对应的存储器中,例如诸如RAM的易失性存储器或诸如闪存的非易失性存储器。另选地,系统的功能单元,例如数据接口和处理器子系统,可按可编程逻辑的形式在设备或装置中实施,例如作为现场可编程门阵列(FPGA)和/或图形处理单元(GPU)。一般而言,系统的每个功能单元可按电路的形式实施。注意,系统200还可按分布式方式实施,例如,涉及不同的设备或装

置,诸如分布式服务器,例如以云计算的形式。

[0054] 图3示出了与本文所描述的技术一起使用的模型的详细但非限制性的示例。例如,用于本文所描述的技术中的主模型和/或一个或多个辅助模型可根据该图中描述的模型架构来构建。在该图中示出的特定示例是完全卷积神经网络,更具体地,是U-Net型模型。这类模型可用于例如各种图像到图像的转换,包含来自医学成像设备的图像的重建。

[0055] 此处所示的模型针对输入实例IIN 330确定输出图像OIM 350。在该示例中,输入实例IIN是例如表示为 $M \times N \times c$ 大小矩阵的图像,其中通道 $c$ 的数目在灰度图像的情况下可为 $c=1$ ,在RGB图像的情况下可为 $c=3$ ,等等。然而,通道的数目不需要对应于不同的颜色,并且在这个意义上,输入实例IIN也可更一般地称为输入层。该示例中的输出图像OIM是与输入实例IIN具有相同空间维度的图像,例如具有相同的宽度和高度。输出图像OIM可具有与输入图像相同数目的通道,但是通道的数目也可不同。

[0056] 此处的模型是卷积网络的示例,也称为卷积神经网络。一般而言,术语卷积网络可用来指包括至少一个卷积层的任何神经网络。如本领域所公知的,卷积层是通过执行卷积或滑动点积运算来操作的层。因此,在卷积层中,可使用每个均在输入层上方卷积的 $c'$ 滤波器将 $m \times n \times c$ 大小的输入层变换为 $m' \times n' \times c'$ 大小的输出层。例如,一层处的滤波器的数目可为至多或至少8个,至多或至少32个,或至多或至少128个。一层的输出的空间维度 $m' \times n'$ 可不同于其输入的空间维度 $m \times n$ ,尽管通常保持与输入IIN的空间对应。例如,一层的输出的空间维度可小于其输入,例如,卷积层可执行下采样。该层的输出的空间维度也可大于输入的空间维度,例如,该层可为通过对输入特征图进行上采样然后应用卷积操作而实施的所谓的“上卷积”层。卷积神经网络可包括除卷积层之外的各种其它类型的层,例如一个或多个ReLU层和/或一个或多个池化层。卷积网络中的卷积层的数目例如可为至少5层,或至少10层。

[0057] 更具体地,该图所示的模型是所谓的全卷积网络。这类模型将输入实例IIN变换为层序列中的输出图像OIM,每个层保留与输入实例的空间对应,例如卷积层、池化层、ReLU层等。

[0058] 该图中的模型是编码器-解码器模型。一般而言,这类模型包括收缩部分CP 310(也称为“编码器路径”)和膨胀部分EP 320(也称为“解码器路径”)。收缩部分CP可包括为输入实例IIN产生后续激活层的一个或多个层。图中示出了激活层AV1 341、AV2 342,直到AV $k-1$  343和AV $k$  344。激活层可由模型的一个或多个层根据先前的激活层来确定,如箭头361、362、363和364所示:典型地,最大池化随后是具有相关联ReLU操作的一个或多个卷积层。如图所示,激活层的空间维度通常在整個收缩部分CP的大小上减小,例如,激活层AV2可具有比激活层AV1更小的空间维度(尽管通道的数目实际上可增加),并且对于所示的其它激活层也是类似的。

[0059] 然后可在膨胀部分EP中处理由U-net模型的收缩部分CP产生的激活层AV $k$ 。膨胀部分还可包括一个或多个层,这些层产生后续激活层,例如激活层AV $k+1$  345,直到AV $2k-2$  346,并且最终输出图像OIM。如上所述,激活层可由模型的一个或多个层根据先前的激活层来确定,如箭头365至366和367所示。在这种情况下,在典型的示例中,可使用上卷积(例如,上采样随后是卷积),随后是具有相关联的卷积操作的一个或多个卷积层。如图所示,在膨胀部分EP中,激活层的空间维度典型地在大小上增加,例如,激活层AV $k+1$ 可具有比激活层

AV<sub>k</sub>更大的空间维度(尽管通道的数目可减少),并且对于所示的其它激活层也是类似的。

[0060] 有趣的是,膨胀部分EP可可选地包括所谓的跳过连接,其中收缩部分CP的激活层AV<sub>i</sub>可与膨胀部分EP的激活层连接。具有一个或多个跳过连接的编码器-解码器模型可被称为U-Net型模型。例如,上卷积的结果可与收缩部分CP的对应激活层连接,之后可应用一个或多个卷积。收缩部分的特征图可被裁剪以考虑边界像素。例如,如图中箭头365'所示,激活层AV<sub>k-1</sub>可与激活层AV<sub>k</sub>的上卷积连接,然后从其结果可确定激活层AV<sub>k+1</sub>。类似地,激活层AV<sub>2k-2</sub>的确定可使用到相应的激活层AV<sub>2</sub>的跳过连接366';激活层OIM的确定可使用到相应激活层AV<sub>1</sub>的跳过连接367',等等。

[0061] 在O.Ronneberger等人的“U-Net:Convolutional Networks for Biomedical Image Segmentation”(可在<https://arxiv.org/abs/1505.04597>获取并通过引用并入本文)中提出了U-Net模型的一个示例。另一个已知示例是所谓的ResNet,或残余网络。在N.Pezzotti等人的“Adaptive-CS-Net:FastMRI with Adaptive Intelligence”中提出了在迭代过程中使用U-net类型架构并可在本文中应用的详细MR数据重建模型示例。

[0062] 图4示出了如何确定主模型的输入实例的总分布外(OOD)得分并由此执行对输入实例的OOD检测的详细但非限制性的示例。

[0063] 如图所示,可按确定总OOD得分OODS<sub>485</sub>的形式执行OOD检测。在一个实施例中,可将输入IIN给予若干辅助模型SM<sub>i</sub><sub>441-442</sub>。然后,输出SO<sub>i</sub><sub>451-452</sub>可用于创建所谓的OOD图像OODI<sub>475</sub>。OOD图像OODI的像素值可表示模型SM<sub>i</sub>的输出图像SO<sub>i</sub>中的对应像素的可变性。然后可通过减小OOD图像OODI,例如通过对OOD图像求平均,取最大值或最小值来获取OOD值OODS。一旦获取,OOD得分OODS就可用于评估模型MM的输入IIN是在训练分布中还是在训练分布之外。这类评估可以是整体和/或主模型MM中的模型的结果质量的良好指标。输入还可由主模型MM<sub>440</sub>处理,主模型MM<sub>440</sub>可产生主输出图像OI<sub>450</sub>(例如,重建的MR图像、去噪的CT图像等、CT中的去噪图像)。如下面更详细地论述的,在一些实施例中,主模型MM例如在参数数目方面大于辅助模型SM<sub>i</sub>,从而允许更高质量的主输出OI,同时保持OOD得分计算OODS的计算时间在控制之下。

[0064] 具体地,图中所示的是要对其执行OOD检测的输入实例IIN<sub>430</sub>。还如关于图3所论述的,输入实例IIN可为输入图像,例如包括一个、三个或其它数目的通道。图像可由相机捕获,但也可使用其它类型的传感器数据,例如,音频数据或表示为图像的多个传感器测量的时间序列数据。输入实例IIN也不必是图像。例如,输入实例IIN可为特征向量,主模型从该特征向量生成输出图像,例如,生成对抗网络(GAN)的生成器部分或自动编码器(例如变分自动编码器(VAE))的解码器部分的潜在表示。在后面的部分中,输入实例IIN可表示合成数据,从该合成数据生成合成输出图像,但是输入实例IIN也可表示真实世界输入,例如,通过被确定为自动编码器的编码器部分的输出或者基于自动编码器的编码器部分的输出。

[0065] 作为另一示例,输入实例IIN可表示由例如用于医学图像重建或医学图像分析任务中的医学成像设备产生的信号。例如,在MR扫描仪的情况下,输入实例可表示所谓的k空间中的扫描仪信号,或者表示通过对k空间数据应用逆傅立叶变换而得到的图像空间中的扫描仪信号。类似地,在CT扫描仪的情况下,输入实例可将扫描仪信号表示为正弦图,或者通过应用逆拉东变换在图像空间中表示。

[0066] 在任何情况下,在各种实施例中,输入实例IIN可包括除传感器/图像数据之外的

额外信息,例如元数据,诸如日志数据或患者特异性信息,其可帮助确定更准确的输出。

[0067] 图中还示出了主模型MM 440。主模型可为图像处理模型。一般而言,此处所使用的术语“图像处理模型”是指具有图像,例如具有宽度、深度和多个通道(其可为一个、三个或任何其它数目)的层作为输出(但不需要具有图像作为输入)的模型。主模型MM的输出图像在图中示为OI 450。输出图像OIM可具有单个通道,如可为各种分割模型、MR或CT重建模型等的情况;三个通道,如可为各种生成模型、图像到图像转换模型等的情况;或者适用于现有应用的任何其它数目的通道。输出图像可为离散的,例如二进制的,例如在掩模的情况下;或者是连续的,例如在生成模型的情况下。在输入实例是图像的情况下,输出图像在许多情况下具有与输入相同的空间维度,可能达到边界效应。

[0068] 在训练数据集(未示出)上训练主模型MM:典型地,标记数据集包括多个训练输入实例和对应的期望训练输出图像(例如,至少1000个或至少1000000个训练实例)。因此,主模型是可训练模型(也称为机器可学习模型或机器学习模型)。这类可训练模型典型地通过学习一组可训练参数的值来训练。例如,主模型的参数数目可为至少1000个、至少100000个或至少1000万个。从训练效率的观点来看,使用适于基于梯度的优化的模型是有益的,例如,该模型在其一组参数中是连续的和/或可微分的。

[0069] 各种类型的模型可用作主模型MM。例如,主模型可为U-net型模型,或更一般地,任何类型的编码器-解码器模型、完全卷积神经网络、卷积神经网络或其它类型的神经网络,如关于图3所论述的。这类神经网络也可用作较大模型的一部分,例如,迭代模型的一次或多次迭代中,如在例如在N.Pezzotti等人的“Adaptive-CS-Net:FastMRI with Adaptive Intelligence”中所做的。如上所述,主模型可为生成模型,例如,GAN的生成部分或VAE的解码器部分。一般而言,神经网络也被称为人工神经网络。当使用神经网络时,该组参数可包括神经网络的节点的权重。例如,模型的层数可为至少5层或至少10层,并且节点和/或权重的数目可为至少1000或至少10000。取决于特定应用,可使用各种已知的神经网络架构和其它类型的机器可学习模型。

[0070] 为了对主模型MM执行输入实例IIN的OOD检测,可使用多个辅助模型SM<sub>i</sub>。图中示出了辅助模型SM<sub>1</sub> 441,直到SM<sub>m</sub> 442。例如,可使用至多或至少五个或至多或至少十个辅助模型。通过应用相应的辅助模型SM<sub>i</sub>,可获取相应的辅助模型输出图像SM<sub>i</sub>,如图中由辅助模型输出图像SO<sub>1</sub> 451直到SO<sub>m</sub> 452所示。

[0071] 一般而言,辅助模型SM<sub>i</sub>可为在与主模型MM相同的训练数据集上训练的模型。因此,当应用于来自训练数据集或类似于来自训练数据集的实例的输入实例IIN时,可预期辅助模型SM<sub>i</sub>表现出与主模型以及彼此类似的行为。因此,可预期辅助模型的输出之间相对低的每像素可变性。然而,在不是来自训练数据集的输入实例IIN上,没有这类保证,因此,可预期更高的每像素可变性。

[0072] 应注意,为了实现上述行为,辅助模型SM<sub>i</sub>既不需要使用与主模型MM或彼此完全相同的模型架构;也不能在完全相同的训练输入和输出上训练辅助模型。通过使用对于训练和/或使用更有效的辅助模型SM<sub>i</sub>,可利用这种观察来提高效率。例如,可在按比例缩减的训练输入实例上训练一个、若干个或全部辅助模型。在使用中,这类辅助模型可通过首先按比例缩减输入实例来应用于输入实例IIN。另选地或附加地,可在按比例缩减的训练输出图像上训练一个、若干个或全部的辅助模型,并且因此,在使用中,产生较小的输出图像,如果需

要的话,随后可该输出图像其放大。例如,输入和/或输出可按比例缩减至少2倍,或至少4倍。

[0073] 一般而言,使用较小的输入和/或输出可使得能够使用较小的辅助模型,例如,具有比主模型更少的可训练参数,从而降低存储和计算要求。还可按其它方式简化辅助模型以减少可训练参数的数目,例如通过使用具有较少层的神经网络,通过使用具有较少迭代的迭代模型等。例如,辅助模型可具有可训练参数数目的至多一半、至多25%或至多10%作为主模型。例如,如果主模型是迭代模型,则辅助模型也可为迭代次数的至多一半、至多25%或至多10%的迭代模型,而不是减少单次迭代中使用的参数的数目或除此之外。例如,通过使用具有3次迭代的辅助模型,发明人对于具有15次迭代的主模型获取了良好的结果。

[0074] 在一些实施例中,辅助模型 $SM_i$ 具有公共模型架构,例如,它们可各自通过初始化公共模型的一组参数并基于初始化优化该组参数来训练,但是使用不同的随机初始化。这种公共架构可与主模型 $MM$ 的架构相同或不同。该技术的优势在于,无论是在训练还是使用辅助模型时,该技术均是普遍适用的,并且非常适于并行化。

[0075] 注意,虽然在该图中主模型 $MM$ 与辅助模型 $SM_i$ 分开示出,但是主模型可用作辅助模型之一,例如,既用于确定主模型输出又用于确定辅助模型输出。主模型 $MM$ 还可由辅助模型中的若干个或全部组成,例如,主模型输出可基于辅助模型 $SM_i$ 中的一个或多个的输出来确定,例如,通过平均或另一种类型的组合操作。

[0076] 在操作PXS 470中,可使用辅助模型输出图像 $SO_i$ 来确定相应的辅助模型输出图像 $SO_i$ 的像素的像素OOD得分。例如,可确定相应的图像 $SO_i$ 的具有坐标(1,1)的左上像素的像素OOD得分,并且对于其它像素(i,j)也是类似的。在一些实施例中,可为每个像素确定像素OOD得分。然而,不需要确定所有像素的像素OOD得分,而是相反获取像素样本,例如,出于效率原因。像素中的若干个或全部的像素OOD得分本身可被视为形成可被称为输入实例 $IIN$ 的“OOD图像”的图像。图中示出的为OOD图像OODI 475。

[0077] 可选地,操作PXS可包括调整大小(例如,将图像按比例缩减到最小大小)和/或通道缩减辅助模型输出图像 $SO_i$ 以使其大小对应(例如,将图像转换为灰度)。

[0078] 像素的像素OOD得分可被确定为相应的辅助模型输出图像中的像素的相应值之间的可变性。例如,可变性可被确定为包括次输出图像 $SO_i$ 的像素值(例如0和1之间的灰度像素值、离散化像素值等)的向量的可变性度量。可采取任何适当的可变性度量,例如方差或标准偏差;香农熵(Shannon entropy);等等。在多通道次输出图像的情况下,输出图像可被转换成单通道图像,例如转换成灰度。还可确定每个通道的像素OOD得分,在这种情况下,可通过组合每个通道的得分(例如,作为最大值、最小值或平均值),或者通过将每个通道的得分的向量视为像素OOD得分来获取像素OOD得分。

[0079] 已经确定了辅助模型输出图像 $SO_i$ 的像素的像素OOD得分OODI,在组合操作CMB 480中,可将这些得分组合成总OOD得分OODS 485。总OOD得分OODS可指示输入实例相对于训练数据集是否为OOD。

[0080] 取决于总OOD得分OODS的各种方式是可能的。例如,可使用最大值、最小值或平均值,或者通常为较高的像素OOD得分OODI提供较高得分的任何其它计算,例如在每个像素OOD得分中不减少或增加的函数。不同的得分类型提供了在不同情况下可能有用的OOD状态的不同指示。该最大值可用作高保证选项,以确保不存在输出图像的辅助模型发散的部分,

并且因此,确保主模型的输出图像的每个部分是可信的。平均值和最小值可用于获取主模型输出的可信度的更全面的度量。由于不同的度量提供了关于输入实例IIN的OOD状态的不同种类的信息,因此也可输出多个OOD得分值,例如,总OOD得分可包括相应的组成OOD得分值,例如,最大值和平均值。

[0081] 为了基于总OOD得分OODS确定输入实例是否为OOD,可应用阈值OOD得分。在一些实施例中,作为训练辅助模型的一部分,可基于针对主模型的多个输入实例确定的OOD得分来自动确定阈值得分。例如,阈值得分可被确定为属于分布内输入实例的OOD得分的统计分布的总OOD得分的统计测试的阈值。可计算阈值得分,例如,作为训练集(含有分布内样本)和保持集(含有分布外样本)的最佳阈值得分。

[0082] 如本领域本身已知的,可使用各种已知方法来训练辅助模型。可选地,主模型由与辅助模型相同的系统训练,尽管也可获取预训练的主模型并训练辅助模型以对其使能OOD检测。典型地,使用诸如随机梯度下降的随机方法来执行训练,例如,使用如Kingma和Ba的“Adam:A Method for Stochastic Optimization”(可在<https://arxiv.org/abs/1412.6980>获取并通过引用并入本文)中所公开的Adam优化器。众所周知,这类优化方法可为启发式的和/或达到局部最优。训练可在逐个实例的基础上或分批地执行,例如至多或至少64个或至多或至少256个实例。

[0083] 图5示出了对到主模型的输入实例进行分布外(OOD)检测的计算机实施的方法500的框图。主模型可被配置为从输入实例生成输出图像。可在训练数据集上训练主模型。方法500可对应于图2的系统200的操作。然而,这不是限制,因为方法500也可使用另一系统,装置或设备来执行。

[0084] 方法500可包括,在题为“访问辅助模型”的操作中,访问510表示在OOD检测中使用的多个辅助模型的数据。可在与训练主模型的训练数据集相同的训练数据集上训练辅助模型。

[0085] 方法500可包括,在题为“获取输入实例”的操作中,获取520输入实例。

[0086] 方法500可包括,在题为“应用辅助模型”的操作中,应用530相应的多个辅助模型以获取相应的辅助模型输出图像。

[0087] 方法500可包括,在题为“确定像素的OOD得分”的操作中,确定540相应的辅助模型输出图像的像素的像素OOD得分。像素的像素OOD得分可被确定为该像素在相应的辅助模型输出图像中的相应值之间的可变性。

[0088] 方法500可包括,在题为“组合成总OOD得分”的操作中,将所确定的像素OOD得分组合550成总OOD得分。总OOD得分可指示输入实例相对于训练数据集是否为OOD。

[0089] 方法500可包括,在题为“生成输出信号”的操作中,基于总OOD得分生成560输出信号,该输出信号指示输入实例是否为OOD。

[0090] 图6示出了对主模型的输入使能分布外(OOD)检测的计算机实施的方法600的框图。主模型可被配置为从输入实例生成输出图像。可在训练数据集上训练主模型。方法600可对应于图1的系统100的操作。然而,这不是限制,因为方法600也可使用另一系统、装置或设备来执行。

[0091] 方法600可包括,在题为“访问训练数据集”的操作中,访问610表示训练数据集的数据,在该训练数据集上训练主模型。

[0092] 方法600可包括,在题为“训练辅助模型”的操作中,训练620多个辅助模型。可在训练主模型的训练数据集上训练辅助模型。辅助模型可用于确定用于例如根据方法500的OOD检测中的输入实例的辅助模型输出图像。

[0093] 方法600可包括,在题为“将模型与主模型相关联”的操作中,将多个辅助模型与主模型相关联630以使能OOD检测。

[0094] 应理解,一般而言,图5的方法500和图6的方法600的操作可按任何合适的顺序来执行,例如,连续地、同时地,或其组合,在适用的情况下,例如,通过输入/输出关系,需要特定的顺序。这些方法中的一些或全部也可组合,例如,使能OOD检测的方法600可连续用于使用方法500执行OOD检测。

[0095] 该方法(一种或多种)可在计算机上实施为计算机实施的方法、专用硬件或两者的组合。同样如图7所示,用于计算机的指令(例如,可执行代码)可存储在计算机可读介质700上,例如,例如,以机器可读物理标记的系列710的形式和/或作为具有不同电(例如,磁)或光属性或值的系列元件。可执行代码可按瞬态或非瞬态方式存储。计算机可读介质的示例包含存储设备、光存储设备、集成电路、服务器、在线软件等。图7示出了光盘700。另选地,计算机可读介质700可包括表示多个辅助模型的瞬态或非瞬态数据710。辅助模型可与主模型相关联以使能OOD检测。主模型可被配置为从输入实例生成输出图像。可在训练数据集上训练主模型。可在训练主模型的训练数据集上训练辅助模型。辅助模型可用于确定用于OOD检测中的输入实例的辅助模型输出图像。可根据计算机实施的600获取辅助模型。辅助模型可用于计算机实施的方法500中。

[0096] 示例、实施例或可选特征,无论是否被指示为非限制性的,均不应被理解为限制所要求保护的本发明。

[0097] 应注意,上述实施例说明而非限制本发明,并且所属领域的技术人员将能够在不脱离所附权利要求书的范围的情况下设计许多替代实施例。在权利要求中,置于括号中的任何附图标记均不应解释为对权利要求的限制。动词“包括(comprise)”及其动词变体的使用不排除权利要求中所陈述的元件或阶段之外的元件或阶段的存在。元件之前的冠词“一(a)”或“一个(an)”不排除多个这类元件的存在。诸如“至少一个(at least one of)”的表述当在一个元件列表或元件组之前时,表示从该列表或组中选择所有元件或任何元件子集。例如,表述“A、B和C中的至少一种(at least one of A,B,and C)”应理解为仅包含A、仅包含B、仅包含C、包含A和B两者、包含A和C两者、包含B和C两者、或包含A、B和C的全部。本发明可借助于包括若干个不同元件的硬件以及借助于适当编程的计算机来实施。在列举了若干种方式的设备权利要求中,这些方式中的若干个可由同一项硬件来体现。在相互不同的从属权利要求中叙述某些措施的事实并不指示不能有利地使用这些措施的组合。

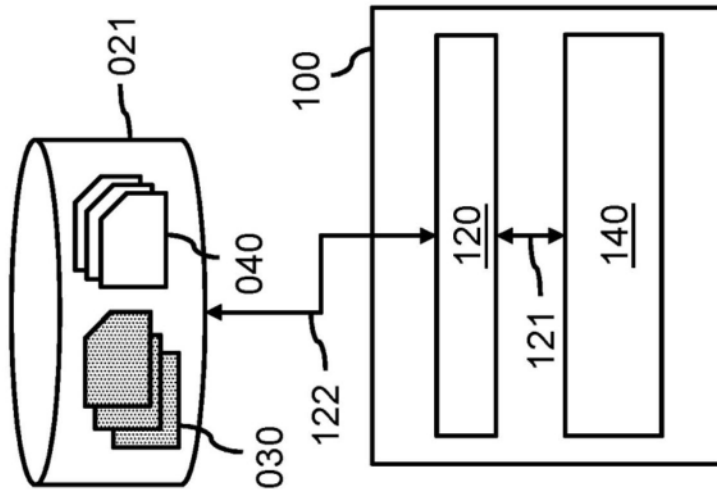


图1

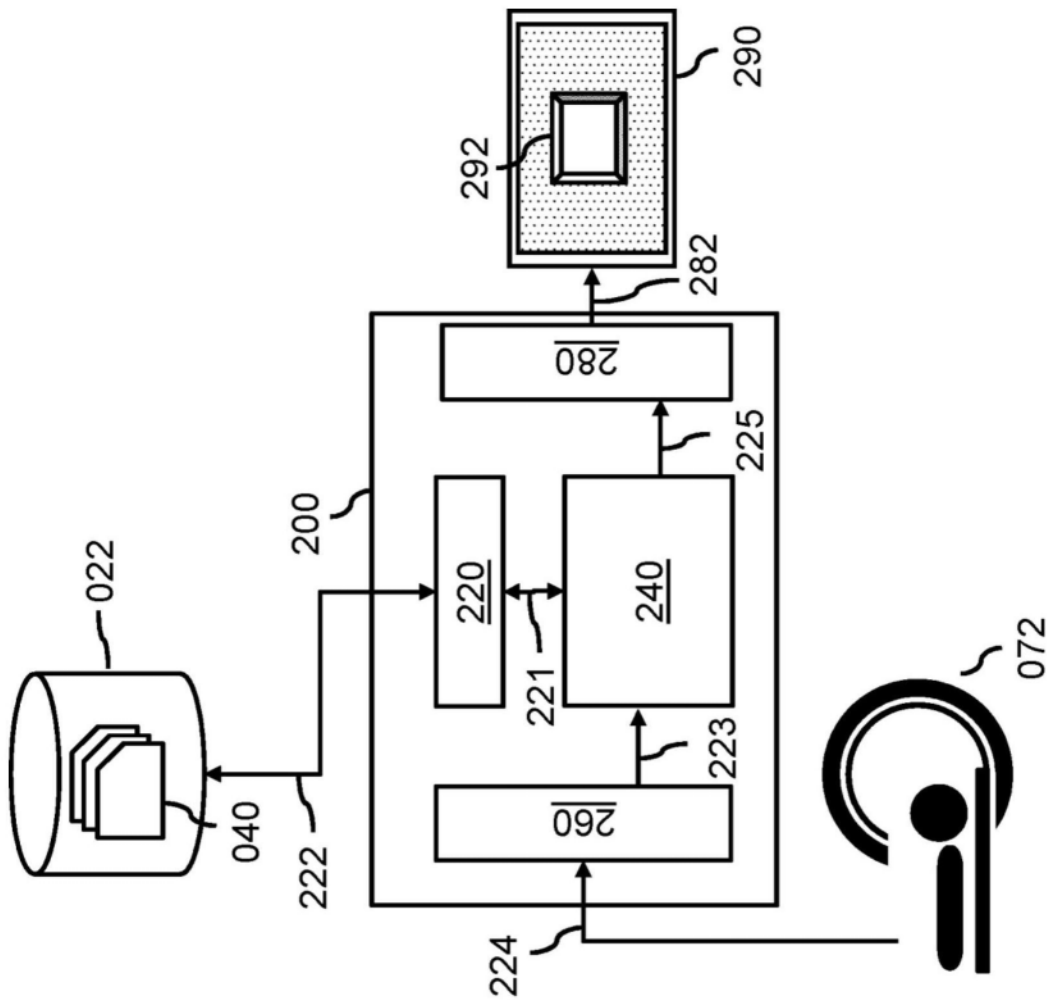


图2

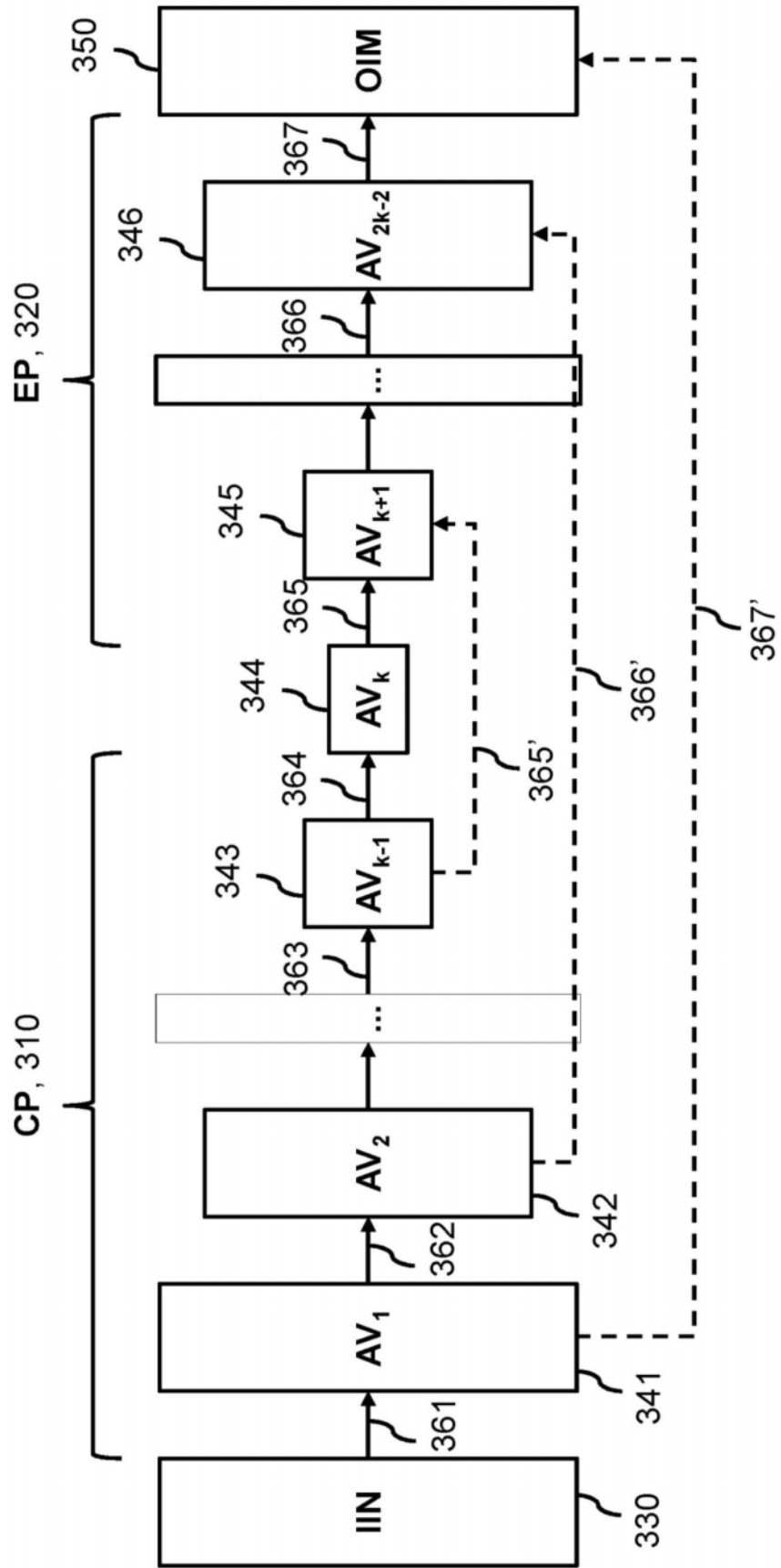


图3

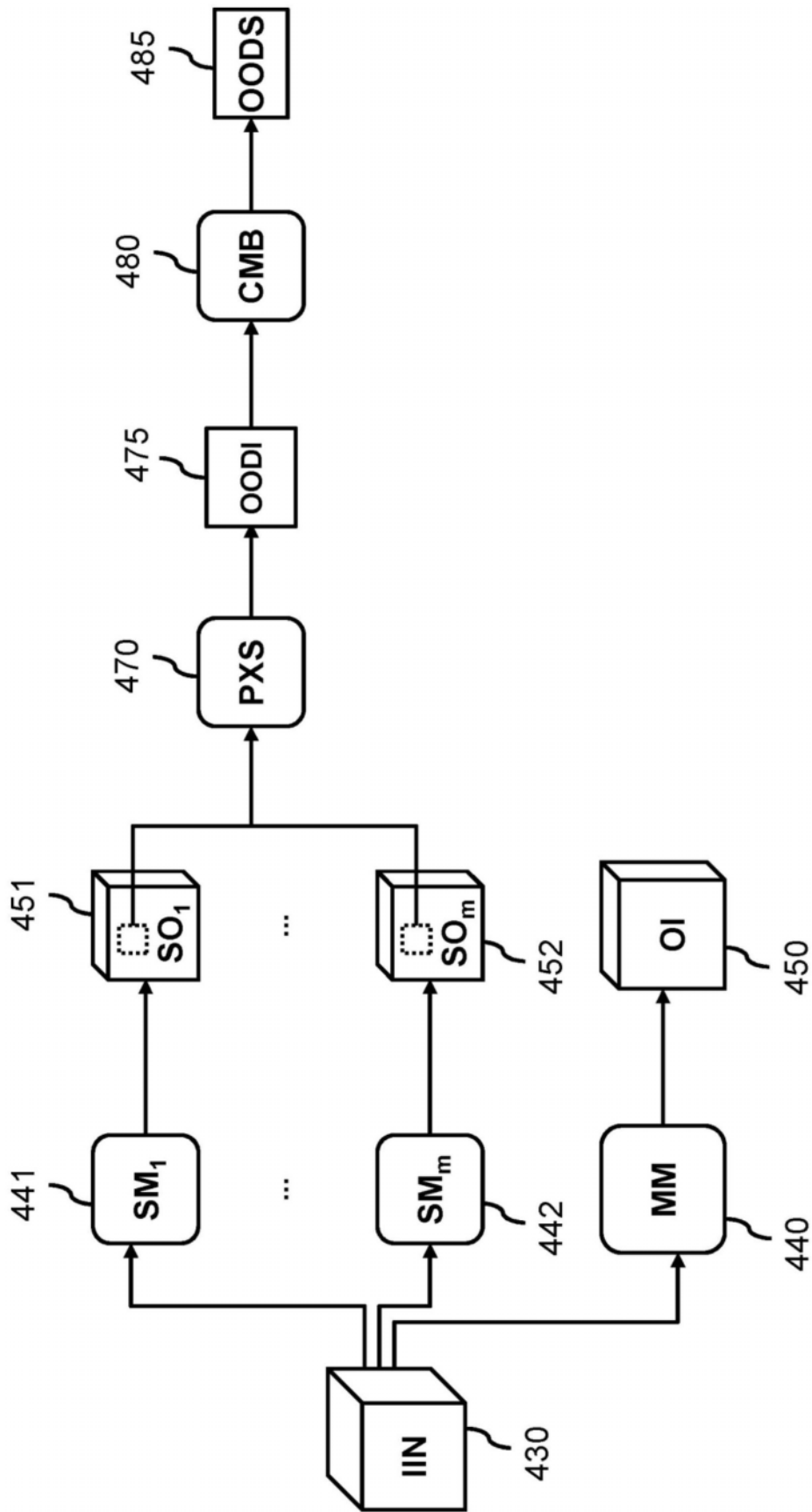


图4

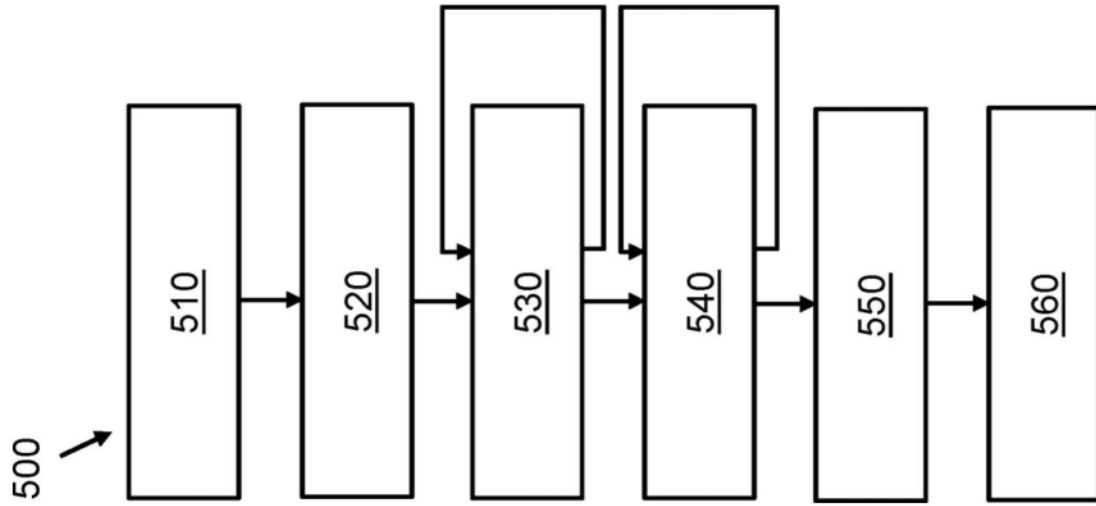


图5

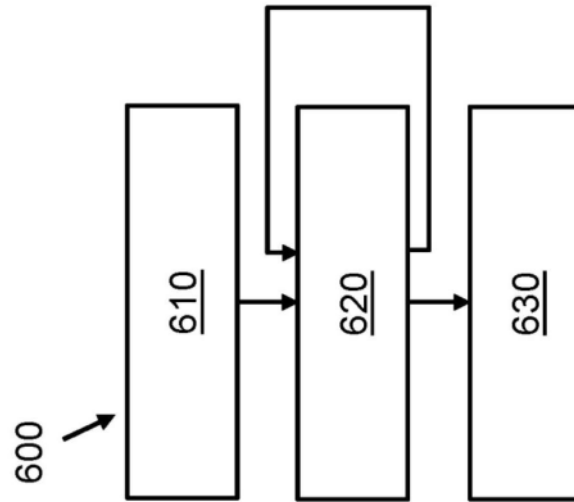


图6

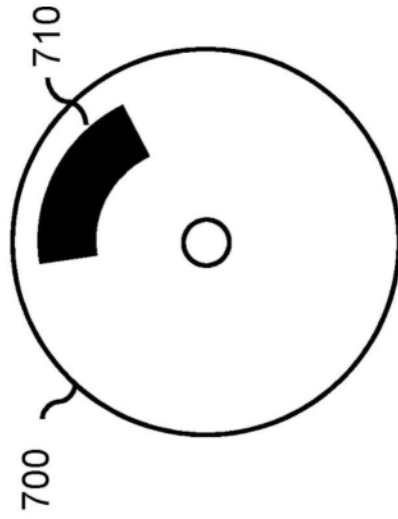


图7