



(12) 发明专利

(10) 授权公告号 CN 108053407 B

(45) 授权公告日 2021.04.13

(21) 申请号 201711415016.9

(22) 申请日 2017.12.22

(65) 同一申请的已公布的文献号
申请公布号 CN 108053407 A

(43) 申请公布日 2018.05.18

(73) 专利权人 联想(北京)有限公司
地址 100085 北京市海淀区上地信息产业
基地创业路6号

(72) 发明人 姜譞

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 周天宇

(51) Int. Cl.
G06T 7/10 (2017.01)

(56) 对比文件

CN 102750549 A, 2012.10.24

CN 106778834 A, 2017.05.31

CN 105930841 A, 2016.09.07

CN 102436583 A, 2012.05.02

CN 104573669 A, 2015.04.29

CN 107330455 A, 2017.11.07

US 7165029 B2, 2007.01.16

CN 106951925 A, 2017.07.14

CN 104318562 A, 2015.01.28

CN 107330263 A, 2017.11.07

CN 106529485 A, 2017.03.22

CN 107273891 A, 2017.10.20

CN 103258214 A, 2013.08.21

CN 103565470 A, 2014.02.12

CN 105069136 A, 2015.11.18

审查员 马金驹

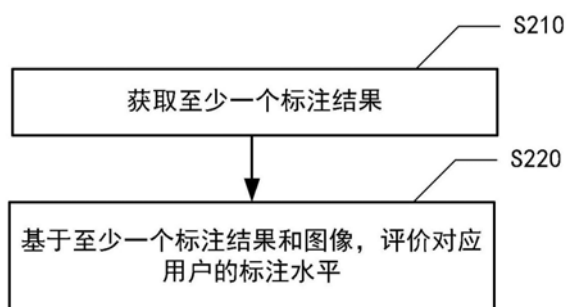
权利要求书2页 说明书8页 附图5页

(54) 发明名称

数据处理方法及数据处理系统

(57) 摘要

本公开提供了一种数据处理方法,包括:获取至少一个标注结果,其中,至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的;以及基于至少一个标注结果和图像,评价对应用户的标注水平。本公开提供的数据处理方法,利用获取的至少一个用户给图像中的对象标注数据形成的至少一个标注结果,评价用户的标注水平,在大数据背景下采用一种基于人工智能的深度学习的方法,至少部分地克服了相关技术中评价标注水平需要耗费较高成本的技术问题,实现对标注水平的自动评价,并实时的输出对标注水平的评价分数的技术效果。此外,本公开还提供了一种数据处理系统。



1. 一种数据处理方法,包括:

获取至少一个标注结果,其中,所述至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的;以及

基于所述至少一个标注结果和所述图像,对所述至少一个标注结果中的一个或多个标注结果进行评价,以评价所述对应用户的标注水平;

其中,基于所述至少一个标注结果和所述图像,评价所述对应用户的标注水平包括:

确定所述至少一个标注结果中各标注结果的初始权重值;

根据所述初始权重值和所述各标注结果,确定与所述至少一个标注结果对应的期望标注数据,期望标注数据A计算公式如下:

$$A = \sum_{i=1}^N w_i A_i$$

其中, w_i 为初始权重值,第i个用户的标注结果用 A_i 表示, $i=1,2,\dots,N$;

基于所述期望标注数据和所述图像,确定用于评价所述对应用户的标注水平的期望分割结果;以及

以所述期望分割结果为参照,评价所述对应用户的标注水平;

其中,以所述期望分割结果为参照,评价所述对应用户的标注水平包括:

依次比较所述期望分割结果和所述至少一个标注结果中各标注结果的差异;

根据差异比较结果,更新所述至少一个标注结果中各标注结果的所述初始权重值,以获得所述各标注结果的目标权重值;以及

根据所述目标权重值,评价与所述各标注结果对应的用户的标注水平。

2. 根据权利要求1所述的方法,其中,基于所述期望标注数据和所述图像,确定用于评价所述对应用户的标注水平的期望分割结果包括:

加载用于图像分割的初始模型;

将所述期望标注数据输入所述初始模型以训练出用于图像分割的目标模型;以及

将所述期望标注数据和所述图像输入所述目标模型,以输出用于评价所述对应用户的标注水平的期望分割结果。

3. 根据权利要求1所述的方法,其中,所述方法还包括:

根据所述目标权重值和所述各标注结果,重新确定与所述至少一个标注结果对应的期望标注数据。

4. 一种数据处理系统,包括:

获取模块,用于获取至少一个标注结果,其中,所述至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的;以及

评价模块,用于基于所述至少一个标注结果和所述图像,对所述至少一个标注结果中的一个或多个标注结果进行评价,以评价所述对应用户的标注水平;

其中,评价模块包括:

第一确定单元,用于确定所述至少一个标注结果中各标注结果的初始权重值;

第二确定单元,用于根据所述初始权重值和所述各标注结果,确定与所述至少一个标注结果对应的期望标注数据,期望标注数据A计算公式如下:

$$A = \sum_{i=1}^N w_i A_i$$

其中, w_i 为初始权重值, 第 i 个用户的标注结果用 A_i 表示, $i=1, 2 \cdots N$;

第三确定单元, 用于基于所述期望标注数据和所述图像, 确定用于评价所述对应用户的标注水平的期望分割结果; 以及

评价单元, 用于以所述期望分割结果为参照, 评价所述对应用户的标注水平;

其中, 评价单元包括:

比较子单元, 用于依次比较所述期望分割结果和所述至少一个标注结果中各标注结果的差异;

更新子单元, 用于根据差异比较结果, 更新所述至少一个标注结果中各标注结果的所述初始权重值, 以获得所述各标注结果的目标权重值; 以及

评价子单元, 用于根据所述目标权重值, 评价与所述各标注结果对应的用户的标注水平。

5. 根据权利要求4所述的系统, 其中, 第三确定单元包括:

加载子单元, 用于加载用于图像分割的初始模型;

训练子单元, 用于将所述期望标注数据输入所述初始模型以训练出用于图像分割的目标模型; 以及

输出子单元, 用于将所述期望标注数据和所述图像输入所述目标模型, 以输出用于评价所述对应用户的标注水平的期望分割结果。

6. 根据权利要求4所述的系统, 其中, 所述系统还包括:

确定模块, 用于根据所述目标权重值和所述各标注结果, 重新确定与所述至少一个标注结果对应的期望标注数据。

数据处理方法及数据处理系统

技术领域

[0001] 本公开涉及一种数据处理方法及数据处理系统。

背景技术

[0002] 利用人工智能和深度学习的方法,可以自动分割出图像中的对象。如在医疗领域,可以自动的分割出计算机断层摄影(Computed Tomography, 简称为CT)图像中的器官,甚至是存在于器官上的病灶组织。而深度学习模型的训练依赖于医生在CT图像中标注好的数据,换言之,医生的标注水平在很大程度上决定了训练出的模型的质量水平。

[0003] 在实现本公开实施例的过程中,发明人发现相关技术中至少存在如下问题:对医生的标注水平进行评价需要耗费较高的成本。

[0004] 针对相关技术中的上述问题,目前还未提出有效的解决方案。

发明内容

[0005] 鉴于此,本公开提供了一种数据处理方法及数据处理系统。

[0006] 本公开的一个方面提供了一种数据处理方法,包括:获取至少一个标注结果,其中,上述至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的;以及基于上述至少一个标注结果和上述图像,评价上述对应用户的标注水平。

[0007] 可选地,基于上述至少一个标注结果和上述图像,评价上述对应用户的标注水平包括:确定上述至少一个标注结果中各标注结果的初始权重值;根据上述初始权重值和上述各标注结果,确定与上述至少一个标注结果对应的期望标注数据;基于上述期望标注数据和上述图像,确定用于评价上述对应用户的标注水平的期望分割结果;以及以上述期望分割结果为参照,评价上述对应用户的标注水平。

[0008] 可选地,基于上述期望标注数据和上述图像,确定用于评价上述对应用户的标注水平的期望分割结果包括:加载用于图像分割的初始模型;将上述期望标注数据输入上述初始模型以训练出用于图像分割的目标模型;以及将上述期望标注数据和上述图像输入上述目标模型,以输出用于评价上述对应用户的标注水平的期望分割结果。

[0009] 可选地,以上述期望分割结果为参照,评价上述对应用户的标注水平包括:依次比较上述期望分割结果和上述至少一个标注结果中各标注结果的差异;根据差异比较结果,更新上述至少一个标注结果中各标注结果的上述初始权重值,以获得上述各标注结果的目标权重值;以及根据上述目标权重值,评价与上述各标注结果对应的用户的标注水平。

[0010] 可选地,上述方法还包括:根据上述目标权重值和上述各标注结果,重新确定与上述至少一个标注结果对应的期望标注数据。

[0011] 本公开的另一个方面提供了一种数据处理系统,包括:获取模块,用于获取至少一个标注结果,其中,上述至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的;以及评价模块,用于基于上述至少一个标注结果和上述图像,评价上述对应用户的标注水平。

[0012] 可选地,评价模块包括:第一确定单元,用于确定上述至少一个标注结果中各标注结果的初始权重值;第二确定单元,用于根据上述初始权重值和上述各标注结果,确定与上述至少一个标注结果对应的期望标注数据;第三确定单元,用于基于上述期望标注数据和上述图像,确定用于评价上述对应用户的标注水平的期望分割结果;以及评价单元,用于以上述期望分割结果为参照,评价上述对应用户的标注水平。

[0013] 可选地,第三确定单元包括:加载子单元,用于加载用于图像分割的初始模型;训练子单元,用于将上述期望标注数据输入上述初始模型以训练出用于图像分割的目标模型;以及输出子单元,用于将上述期望标注数据和上述图像输入上述目标模型,以输出用于评价上述对应用户的标注水平的期望分割结果。

[0014] 可选地,评价单元包括:比较子单元,用于依次比较上述期望分割结果和上述至少一个标注结果中各标注结果的差异;更新子单元,用于根据差异比较结果,更新上述至少一个标注结果中各标注结果的上述初始权重值,以获得上述各标注结果的目标权重值;以及评价子单元,用于根据上述目标权重值,评价与上述各标注结果对应的用户的标注水平。

[0015] 可选地,上述系统还包括:确定模块,用于根据上述目标权重值和上述各标注结果,重新确定与上述至少一个标注结果对应的期望标注数据。

附图说明

[0016] 为了更完整地理解本公开及其优势,现在将参考结合附图的以下描述,其中:

[0017] 图1示意性示出了根据本公开实施例的适用于数据处理方法的应用场景;

[0018] 图2示意性示出了根据本公开实施例的数据处理方法的流程图;

[0019] 图3A示意性示出了根据本公开实施例的评价对应用户的标注水平的流程图;

[0020] 图3B示意性示出了根据本公开实施例的确定用于评价对应用户的标注水平的期望分割结果的流程图;

[0021] 图3C示意性示出了根据本公开实施例的以期望分割结果为参照,评价对应用户的标注水平的流程图;

[0022] 图4示意性示出了根据本公开实施例的数据处理系统的框图;

[0023] 图5A示意性示出了根据本公开实施例的评价模块的框图;

[0024] 图5B示意性示出了根据本公开实施例的第三确定单元的框图;

[0025] 图5C示意性示出了根据本公开实施例的评价单元的框图;以及

[0026] 图6示意性示出了根据本公开实施例的适于实现数据处理方法及其系统的计算机系统的框图。

具体实施方式

[0027] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0028] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用的术语“包括”、“包含”等表明了所述特征、步骤、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、步骤、操作或部件。

[0029] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0030] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。本领域技术人员还应理解,实质上任意表示两个或更多可选项目的转折连词和/或短语,无论是在说明书、权利要求书还是附图中,都应被理解为给出了包括这些项目之一、这些项目任一方、或两个项目的可能性。例如,短语“A或B”应当被理解为包括“A”或“B”、或“A和B”的可能性。

[0031] 附图中示出了一些方框图和/或流程图。应理解,方框图和/或流程图的一些方框或其组合可以由计算机程序指令来实现。这些计算机程序指令可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器,从而这些指令在由该处理器执行时可以创建用于实现这些方框图和/或流程图中所说明的功能/操作的装置。

[0032] 因此,本公开的技术可以硬件和/或软件(包括固件、微代码等)的形式来实现。另外,本公开的技术可以采取存储有指令的计算机可读介质上的计算机程序产品的形式,该计算机程序产品可供指令执行系统使用或者结合指令执行系统使用。在本公开的上下文中,计算机可读介质可以是能够包含、存储、传送、传播或传输指令的任意介质。例如,计算机可读介质可以包括但不限于电、磁、光、电磁、红外或半导体系统、装置、器件或传播介质。计算机可读介质的具体示例包括:磁存储装置,如磁带或硬盘(HDD);光存储装置,如光盘(CD-ROM);存储器,如随机存取存储器(RAM)或闪存;和/或有线/无线通信链路。

[0033] 本公开提供了一种数据处理方法,包括:获取至少一个标注结果,其中,至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的,基于至少一个标注结果和图像,评价对应用户的标注水平。本公开提供的数据处理方法,利用获取的至少一个用户给图像中的对象标注数据形成的至少一个标注结果,评价用户的标注水平,在大数据背景下采用一种基于人工智能的深度学习的方法,至少部分地克服了相关技术中评价标注水平需要耗费较高成本的技术问题,实现对标注水平的自动评价,并实时的输出标注水平的评价分数的技术效果。

[0034] 图1示意性示出了根据本公开实施例的适用于数据处理方法的应用场景。

[0035] 本公开提供的数据处理方法涉及一种人工智能的应用领域,可以实现基于多个输入数据,通过机器学习如深度学习的方法,对多个输入数据中的一个或多个进行评价的任意场景,具体地,可以涉及基于输入的至少一个标注结果对至少一个标注结果中的任意一个或多个标注结果进行评价,并得到对应标注结果的评价结果这一输出结果的各类场景,数据可以是任意试图通过本公开提供的数据处理方法获得该数据评价结果的存在于各行各业中的数据,如可以是在医学领域中,医生对CT图像中的各器官甚至是病灶进行标注,形成的标注数据,也可以是对交通图像进行图像识别形成的数据,包括拍摄地点周围的环境

信息,建筑物,街道,交通工具等,还可以是对包含地下岩石图像中的各岩石进行识别,标注形成的标注数据等。

[0036] 需要注意的是,图1所示的应用场景仅为可以应用本公开实施例的场景的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。

[0037] 以下将以医学领域中,利用人工智能和深度学习方法,自动评价标注结果水平为例,对本公开的实施例的具体实施方式做详细说明,可以理解的是,其他应用场景的数据处理方法可以依据本公开的精神做适当展开,此处不再赘述。

[0038] 图2示意性示出了根据本公开实施例的数据处理方法的流程图。

[0039] 如图2所示,该数据处理方法可以包括操作S210和S220,其中:

[0040] 在操作S210,获取至少一个标注结果,其中,至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的。

[0041] 在操作S220,基于至少一个标注结果和图像,评价对应用户的标注水平。

[0042] 可以理解的是,不同的图像具有不同的属性,在此对图像的种类不做限定。属于同一幅图像中的不同对象,具有不同的性质,可以具有一定的形状和/或构造,不同种类的对象所具有的独特性质根据图像的种类不同而不同。例如人体腹部的CT图像中包含有人体的多个器官,可以包括肝,脾,胆和胰腺等,由于每个器官的大小和形态都不同,其被电子设备扫描并处理后,得到多个形状和大小不同的器官的模糊轮廓信息。

[0043] 针对同一幅人体腹部的CT图像,N个医生对图像中包含的多个器官进行识别标注可以形成不同的N个标注结果,甚至是差异很大的N个标注结果。

[0044] 本公开实施例就是利用这N个标注结果的输入,通过综合各标注结果和分割算法的输出,对这N个标注结果进行标注水平的评价,实时的输出对标注结果的标注水平的评价分数。

[0045] 通过本公开的实施例,利用获取的至少一个用户给图像中的对象标注数据形成的至少一个标注结果,评价用户的标注水平,在大数据背景下采用一种基于人工智能的深度学习方法,至少部分地克服了相关技术中评价标注水平需要耗费较高成本的技术问题,实现对标注水平的自动评价,并实时的输出标注水平的评价分数的技术效果。

[0046] 下面参考图3A~图3C,结合具体实施例对图2所示的数据处理方法做进一步说明。

[0047] 图3A示意性示出了根据本公开实施例的评价对应用户的标注水平的流程图。

[0048] 如图3A所示,前述操作S220可以包括操作S311、S312、S313和S314。其中:

[0049] 在操作S311,确定至少一个标注结果中各标注结果的初始权重值。

[0050] 在操作S312,根据初始权重值和各标注结果,确定与至少一个标注结果对应的期望标注数据。

[0051] 在操作S313,基于期望标注数据和图像,确定用于评价对应用户的标注水平的期望分割结果。

[0052] 在操作S314,以期望分割结果为参照,评价对应用户的标注水平。

[0053] 根据本公开的实施例,各标注结果的初始权重值 W_i 与标注结果的数目N有关,一般可以设置为 $W_i = 1/N$ 。

[0054] 第i个医生的标注结果用 A_i 表示, $i = 1, 2, \dots, N$ 。

[0055] 对N个医生的标注结果,根据初始权重值 W_i 做加权处理,形成信息的标注数据,即期望标注数据A,如下式:

$$[0056] \quad A = \sum_{i=1}^N w_i A_i$$

[0057] 根据期望标注数据A对图像的分割结果,确定用于评价对应用户的标注水平的期望分割结果,以期望分割结果为参照,评价对应用户的标注水平。

[0058] 通过本公开的实施例,将各标注结果赋予相应的权重,生成一个新的标注,即期望标注数据,确定期望分割结果,可以综合多个标注结果,提高标注的准确度。

[0059] 图3B示意性示出了根据本公开实施例的确定用于评价对应用户的标注水平的期望分割结果的流程图。

[0060] 如图3B所示,前述操作S313可以包括操作S321、S322和S323。其中:

[0061] 在操作S321,加载用于图像分割的初始模型。

[0062] 在操作S322,将期望标注数据输入初始模型以训练出用于图像分割的目标模型。

[0063] 在操作S323,将期望标注数据和图像输入目标模型,以输出用于评价对应用户的标注水平的期望分割结果。

[0064] 根据本公开的实施例,根据期望标注数据A对图像的分割结果,确定用于评价对应用户的标注水平的期望分割结果可以利用深度学习模型,需要说明的是,深度学习模型的产生需要用医生标注好的标注结果数据进行训练,不断变化更新的标注结果可以使得训练模型不断更新,越来越趋近于真实的分割结果。换言之,N个医生的标注结果不但可以作为输入数据,还充当着用于对图像进行分割的模型的训练数据,是一个自我学习的过程,由于模型的训练方法并不是本公开的重点,相关技术中任何可以实现模型训练的算法都在本公开的包含范围内,此处不再赘述。

[0065] 将上述期望标注数据A和CT图像,输入分割的深度学习初始模型中加以训练,可以得到训练好的用于图像分割的目标模型,使用训练好的目标模型对CT图像做分割,如U-Net,可以输出分割图,作为用于评价对应用户的标注水平的期望分割结果,可以理解的是,分割图可以是概率图,也可以是基于阈值的二元图,此处不再赘述。

[0066] 通过本公开的实施例,基于模型的期望分割结果确定方法,实现标注结果的自动评价,可以满足大数据的计算需求,且效率高,大大降低了人工成本。

[0067] 图3C示意性示出了根据本公开实施例的以期望分割结果为参照,评价对应用户的标注水平的流程图。

[0068] 如图3C所示,前述操作S314可以包括操作S331、S332和S333。其中:

[0069] 在操作S331,依次比较期望分割结果和至少一个标注结果中各标注结果的差异。

[0070] 在操作S332,根据差异比较结果,更新至少一个标注结果中各标注结果的初始权重值,以获得各标注结果的目标权重值。

[0071] 在操作S333,根据目标权重值,评价与各标注结果对应的用户的标注水平。

[0072] 根据本公开的实施例,在获得用于评价对应用户的标注水平的期望分割结果后,可以将每个医生的标注结果 A_i 与这一期望分割结果做比较,计算差异 d_i ,可以用距离的度量,也可以用其他统计量,此处不做限定。将 d_i 做归一化处理,更新各标注结果的初始权重值 W_i ,得到各标注结果对应的目标权重值,将此目标权重值作为各标注结果的评价结果。

[0073] 通过本公开的实施例,比较各个标注结果与期望标注结果的差异,并更新初始权重值,评价标注水平,可以准确客观的反映标注结果的真实水平。

[0074] 根据本公开的实施例,该数据处理方法还可以包括根据所述目标权重值和所述各标注结果,重新确定与所述至少一个标注结果对应的期望标注数据。

[0075] 需要说明的是,当有新的一批标注结果和图像可以用来训练时,可以目标权重值作为新的初始权重值,进行训练,得到新的期望标注数据,如此,用于对图像进行分割的深度学习模型处于不断的更新和逐渐优化的状态。

[0076] 通过本公开的实施例,重新确定期望标注数据,以新的期望标注数据进行训练,可以提高模型的准确度,进一步提高评价结果的准确度。

[0077] 图4示意性示出了根据本公开实施例的数据处理系统的框图。

[0078] 如图4所示,该数据处理系统400可以包括获取模块410和评价模块420。其中:获取模块410用于获取至少一个标注结果,其中,所述至少一个标注结果中的每个为对应用户给图像中的对象标注数据形成的。评价模块420用于基于所述至少一个标注结果和所述图像,评价所述对应用户的标注水平。

[0079] 可以理解的是,获取模块410和评价模块420可以合并在一个模块中实现,或者其中的任意一个模块可以被拆分成多个模块。或者,这些模块中的一个或多个模块的至少部分功能可以与其他模块的至少部分功能相结合,并在一个模块中实现。根据本发明的实施例,获取模块410和评价模块420中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以以对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式的适当组合来实现。或者,获取模块410和评价模块420中的至少一个可以至少被部分地实现为计算机程序模块,当该程序被计算机运行时,可以执行相应模块的功能。

[0080] 通过本公开的实施例,利用获取的至少一个用户给图像中的对象标注数据形成的至少一个标注结果,评价用户的标注水平,在大数据背景下采用一种基于人工智能的深度学习方法,至少部分地克服了相关技术中评价标注水平需要耗费较高成本的技术问题,实现对标注水平的自动评价,并实时的输出标注水平的评价分数的技术效果。

[0081] 图5A示意性示出了根据本公开实施例的评价模块的框图。

[0082] 如图5A所示,评价模块420可以包括第一确定单元511、第二确定单元512、第三确定单元513和评价单元514。其中:第一确定单元511用于确定所述至少一个标注结果中各标注结果的初始权重值。第二确定单元512用于根据所述初始权重值和所述各标注结果,确定与所述至少一个标注结果对应的期望标注数据。第三确定单元513用于基于所述期望标注数据和所述图像,确定用于评价所述对应用户的标注水平的期望分割结果。评价单元514用于以所述期望分割结果为参照,评价所述对应用户的标注水平。

[0083] 通过本公开的实施例,将各标注结果赋予相应的权重,生成一个新的标注,即期望标注数据,确定期望分割结果,可以综合多个标注结果,提高标注的准确度。

[0084] 图5B示意性示出了根据本公开实施例的第三确定单元的框图。

[0085] 如图5B所示,第三确定单元513可以包括加载子单元521、训练子单元522和输出子单元523。其中:加载子单元521用于加载用于图像分割的初始模型。训练子单元522用于将

所述期望标注数据输入所述初始模型以训练出用于图像分割的目标模型。输出子单元523用于将所述期望标注数据和所述图像输入所述目标模型,以输出用于评价所述对应用户的标注水平的期望分割结果。

[0086] 通过本公开的实施例,基于模型的期望分割结果确定方法,实现标注结果的自动评价,可以满足大数据的计算需求,且效率高,大大降低了人工成本。

[0087] 图5C示意性示出了根据本公开实施例的评价单元的框图。

[0088] 如图5C所示,评价单元514可以包括比较子单元531、更新子单元532和评价子单元533。其中:比较子单元531用于依次比较所述期望分割结果和所述至少一个标注结果中各标注结果的差异。更新子单元532用于根据差异比较结果,更新所述至少一个标注结果中各标注结果的所述初始权重值,以获得所述各标注结果的目标权重值。评价子单元533用于根据所述目标权重值,评价与所述各标注结果对应的用户的标注水平。

[0089] 通过本公开的实施例,比较各个标注结果与期望标注结果的差异,并更新初始权重值,评价标注水平,可以准确客观的反映标注结果的真实水平。

[0090] 根据本公开的实施例,该数据处理系统还可以包括确定模块,用于根据所述目标权重值和所述各标注结果,重新确定与所述至少一个标注结果对应的期望标注数据。

[0091] 通过本公开的实施例,重新确定期望标注数据,以新的期望标注数据进行训练,可以提高模型的准确度,进一步提高评价结果的准确度。

[0092] 图6示意性示出了根据本公开实施例的适于实现数据处理方法及其系统的计算机系统的框图。图6示出的计算机系统仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0093] 如图6所示,计算机系统600可以包括处理器610和计算机可读存储介质620。该计算机系统600可以执行上面参考图2以及图3A~图3C描述的数据处理方法:利用获取的至少一个用户给图像中的对象标注数据形成的至少一个标注结果,评价用户的标注水平,在大数据背景下采用一种基于人工智能的深度学习方法,至少部分地克服了相关技术中评价标注水平需要耗费较高成本的技术问题,实现对标注水平的自动评价,并实时的输出对标注水平的评价分数的技术效果。

[0094] 具体地,处理器610例如可以包括通用微处理器、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC)),等等。处理器610还可以包括用于缓存用途的板载存储器。处理器610可以是用于执行参考图2以及图3A~图3C描述的根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0095] 计算机可读存储介质620,例如可以是能够包含、存储、传送、传播或传输指令的任意介质。例如,可读存储介质可以包括但不限于电、磁、光、电磁、红外或半导体系统、装置、器件或传播介质。可读存储介质的具体示例包括:磁存储装置,如磁带或硬盘(HDD);光存储装置,如光盘(CD-ROM);存储器,如随机存取存储器(RAM)或闪存;和/或有线/无线通信链路。

[0096] 计算机可读存储介质620可以包括计算机程序621,该计算机程序621可以包括代码/计算机可执行指令,其在由处理器610执行时使得处理器610执行例如上面结合图2以及图3A~图3C所描述的方法流程及其任何变形。

[0097] 计算机程序621可被配置为具有例如包括计算机程序模块的计算机程序代码。例

如,在示例实施例中,计算机程序621中的代码可以包括一个或多个程序模块,例如包括621A、模块621B、……。应当注意,模块的划分方式和个数并不是固定的,本领域技术人员可以根据实际情况使用合适的程序模块或程序模块组合,当这些程序模块组合被处理器610执行时,使得处理器610可以执行例如上面结合图2以及图3A~图3C所描述的方法流程及其任何变形。

[0098] 根据本公开的实施例,处理器610可以执行上面结合图2以及图3A~图3C所描述的方法流程及其任何变形。

[0099] 根据本发明的实施例,上文所述的模块中的至少一个可以实现为参考图6描述的计算机程序模块,其在被处理器610执行时,可以实现上面描述的相应操作。

[0100] 本领域技术人员可以理解,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合,即使这样的组合或结合没有明确记载于本公开中。特别地,在不脱离本公开精神和教导的情况下,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0101] 尽管已经参照本公开的特定示例性实施例示出并描述了本公开,但是本领域技术人员应该理解,在不背离所附权利要求及其等同物限定的本公开的精神和范围的情况下,可以对本公开进行形式和细节上的多种改变。因此,本公开的范围不应该限于上述实施例,而是应该不仅由所附权利要求来进行确定,还由所附权利要求的等同物来进行限定。

100

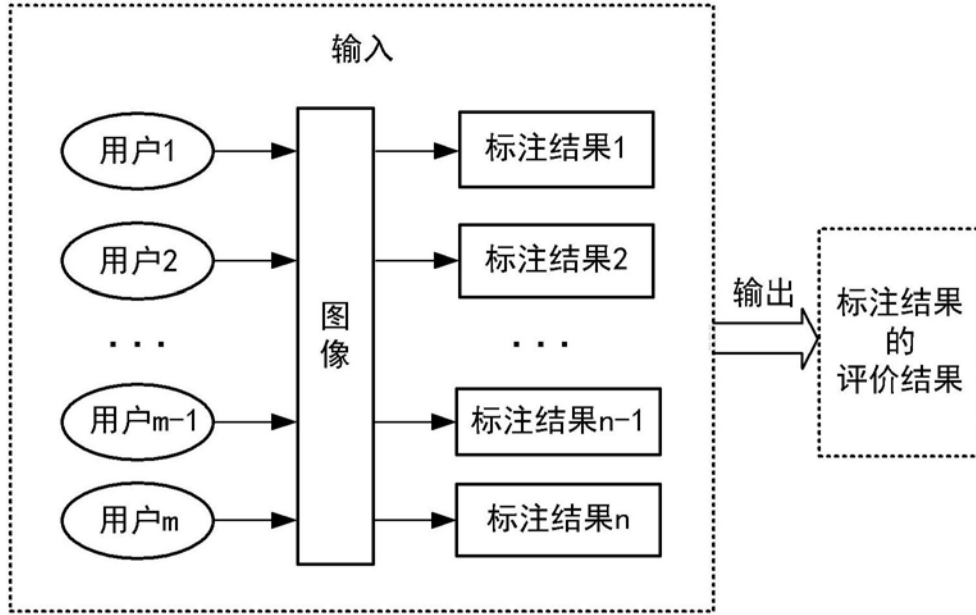


图1

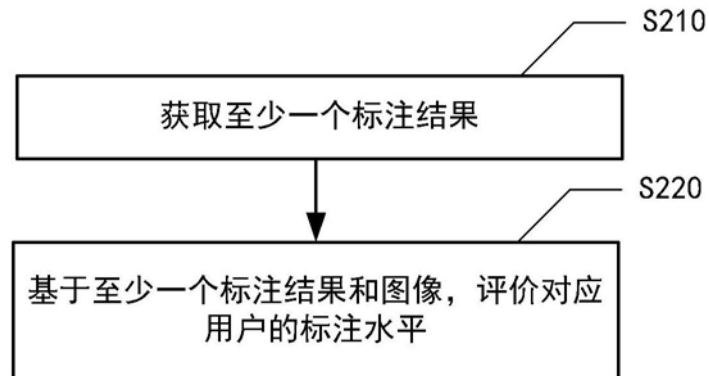


图2

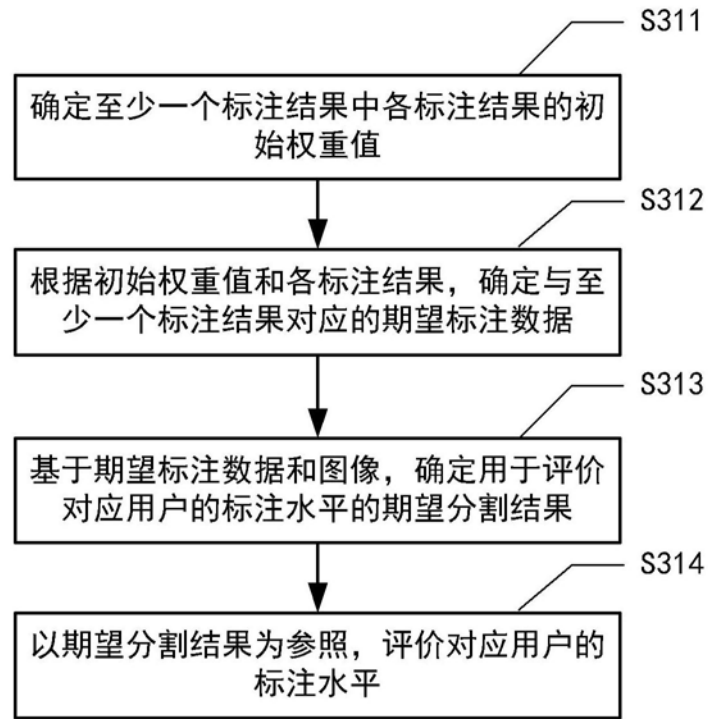


图3A

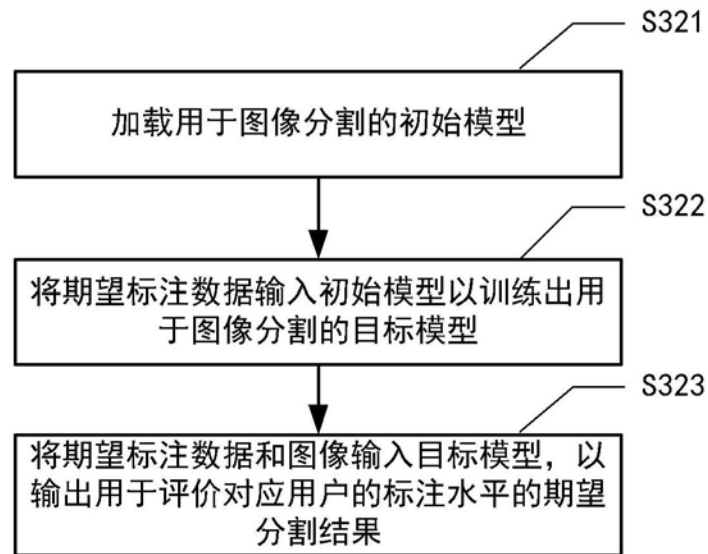


图3B

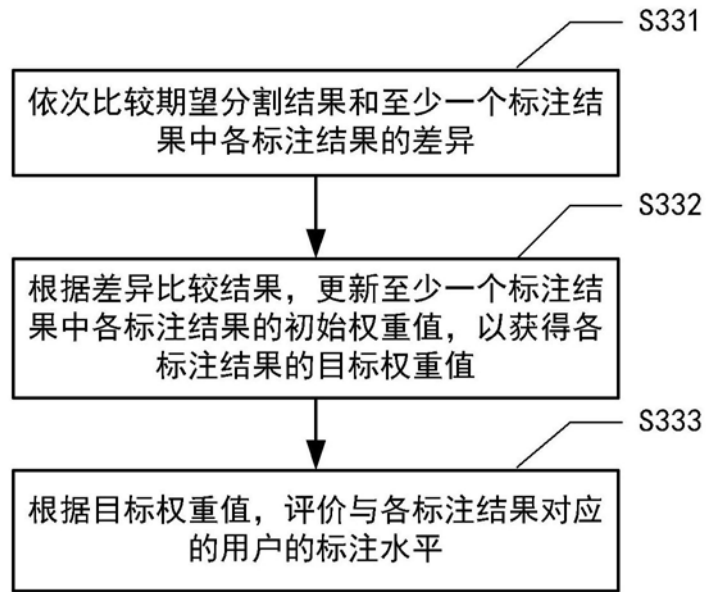


图3C

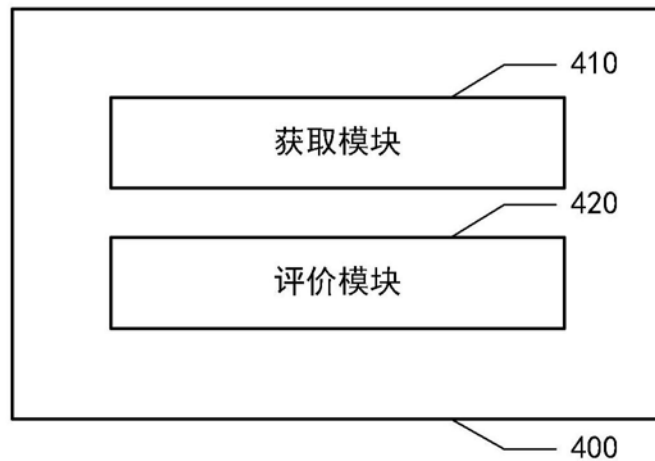


图4

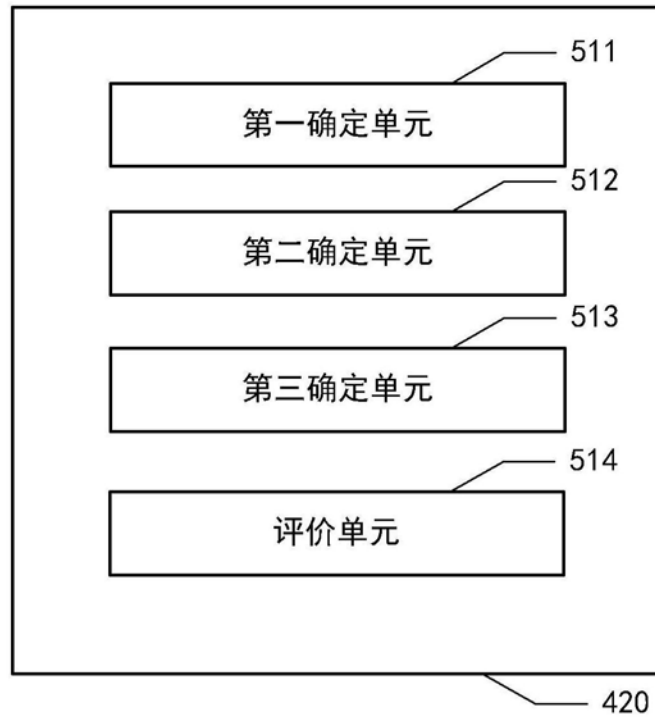


图5A

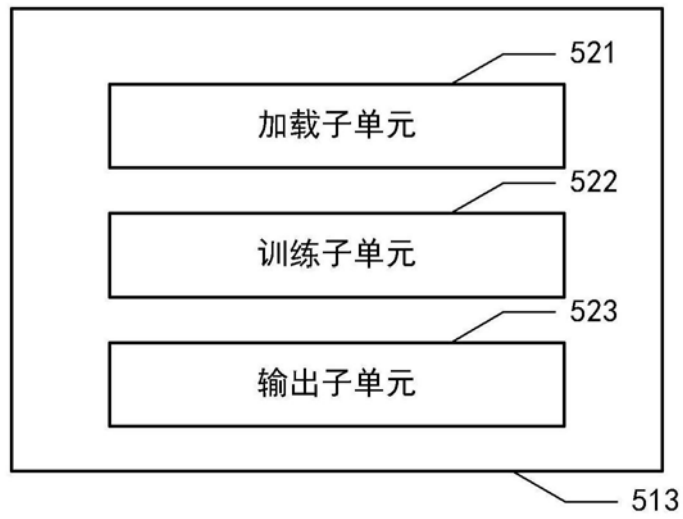


图5B

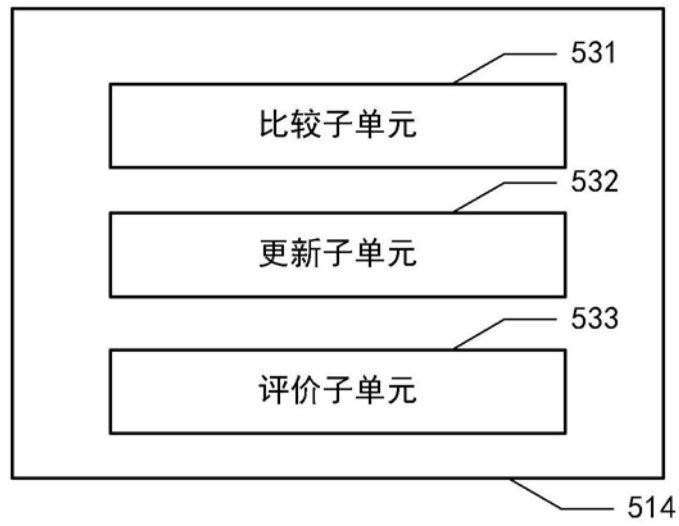


图5C

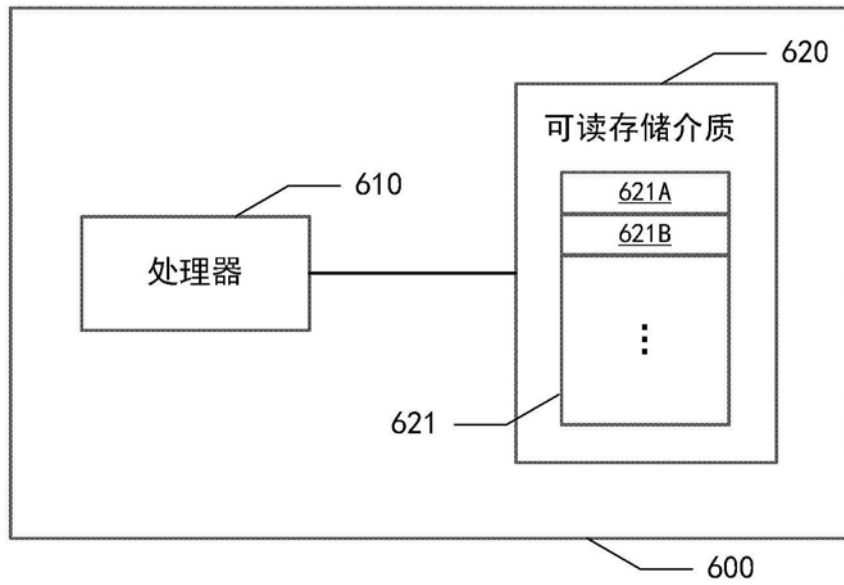


图6