



(12) 发明专利申请

(10) 申请公布号 CN 115578501 A

(43) 申请公布日 2023. 01. 06

(21) 申请号 202211351162.0

(22) 申请日 2022.10.31

(71) 申请人 北京百度网讯科技有限公司

地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 王鹏

(74) 专利代理机构 北京市汉坤律师事务所

11602

专利代理师 姜浩然 吴丽丽

(51) Int. Cl.

G06T 15/04 (2011.01)

G06T 17/00 (2006.01)

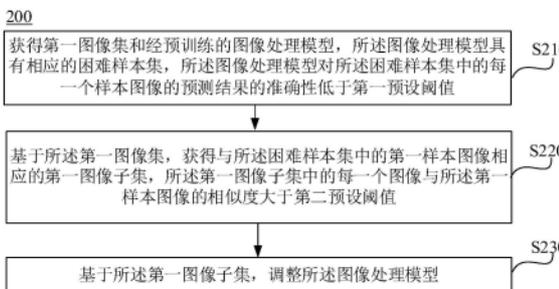
权利要求书4页 说明书12页 附图5页

(54) 发明名称

图像处理方法和装置、电子设备和存储介质

(57) 摘要

本公开提供了图像处理方法、装置、电子设备和存储介质,涉及人工智能技术领域,具体为深度学习、图像处理、计算机视觉技术领域,可应用于OCR等场景。实现方案为:获得第一图像集和经预训练的图像处理模型,图像处理模型具有相应的困难样本集,图像处理模型对困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;基于第一图像集,获得与困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及基于第一图像子集,调整所述图像处理模型。



1. 一种图像处理方法,包括:

获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;

基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及

基于所述第一图像子集,调整所述图像处理模型。

2. 根据权利要求1所述的方法,其中,所述基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集包括:

针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集,所述第二图像集包括与所述第一图像集中的每一图像相应的至少一个增强图像;以及

从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集。

3. 根据权利要求2所述的方法,其中,在所述第二图像集中与所述第一图像集中的每一图像相应的至少一个增强图像中的每一个增强图像与该图像之间的相似度大于第三预设阈值。

4. 根据权利要求2所述的方法,其中,针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集包括:

针对获得所述第一图像集中的第一图像,响应于确定经所述图像处理模型处理后所述第一图像包括至少一个检测框,获得所述第一图像对应的多个增强图像;以及

针对所述多个增强图像中的第一增强图像,响应于确定经所述图像处理模型处理后所述第一增强图像包括至少一个检测框,并且所述第一增强图像中的至少一个检测框与所述第一图像中的至少一个分别对应,确定将所述第一增强图像添加至所述第二图像集。

5. 根据权利要求2所述的方法,其中,针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集包括:

针对获得所述第一图像集中的第二图像,响应于确定经所述图像处理模型处理后所述第二图像包括至少一个检测框,获得所述第二图像对应的多个增强图像;以及

针对所述第二图像和所述多个增强图像中的每一个图像构成的图像对:

分别对该图像对中的两个图像进行第一扰动,以获得两个扰动图像;以及

响应于确定经所述图像处理模型处理后该两个扰动图像分别包括至少一个检测框,并且该两个扰动图像中的第一扰动图像中的至少一个检测框与第二扰动图像中的至少一个检测框分别对应,确定将该图像对中的增强图像添加至所述第二图像集。

6. 根据权利要求2所述的方法,其中,所述从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集包括:

通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征;

通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征;以及

响应于所述第三图像的图像特征与所述第一样本图像的图像特征之间的相似度大于

第四预设阈值,将所述第三图像添加至所述第一图像子集。

7. 根据权利要求6所述的方法,其中,所述通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征包括:

获得所述图像处理模型中的特征提取网络针对所述第一样本图像所提取的第一特征;

响应于确定经所述图像处理模型处理后所述第一样本图像包括至少一个检测框,对所述第一样本图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

基于所述第一特征和所述第一样本图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第一样本图像的图像特征。

8. 根据权利要求6所述的方法,其中,所述通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征包括:

获得所述图像处理模型中的特征提取网络针对所述第三图像所提取的第二特征;

响应于确定经所述图像处理模型处理后所述第三图像包括至少一个检测框,对所述第三图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

基于所述第二特征和所述第三图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第三图像的图像特征。

9. 一种图像处理装置,包括:

第一获取单元,被配置用于获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;

第二获取单元,被配置用于基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及

调整单元,被配置用于基于所述第一图像子集,调整所述图像处理模型。

10. 根据权利要求9所述的装置,其中,所述第二获取单元包括:

数据增强处理单元,被配置用于针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集,所述第二图像集包括与所述第一图像集中的每一图像相应的至少一个增强图像;以及

第一获取子单元,被配置用于从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集。

11. 根据权利要求10所述的装置,其中,在所述第二图像集中与所述第一图像集中的每一图像相应的至少一个增强图像中的每一个增强图像与该图像之间的相似度大于第三预设阈值。

12. 根据权利要求10所述的装置,其中,所述数据增强处理单元包括:

第二获取子单元,被配置用于针对获得所述第一图像集中的第一图像,响应于确定经所述图像处理模型处理后所述第一图像包括至少一个检测框,获得所述第一图像对应的多个增强图像;以及

第一确定单元,被配置用于针对所述多个增强图像中的第一增强图像,响应于确定经

所述图像处理模型处理后所述第一增强图像包括至少一个检测框,并且所述第一增强图像中的至少一个检测框与所述第一图像中的至少一个分别对应,确定将所述第一增强图像添加至所述第二图像集。

13. 根据权利要求10所述的装置,其中,所述数据增强处理单元包括:

第三获取子单元,被配置用于针对获得所述第一图像集中的第二图像,响应于确定经所述图像处理模型处理后所述第二图像包括至少一个检测框,获得所述第二图像对应的多个增强图像;以及

图像对处理单元,被配置用于针对所述第二图像和所述多个增强图像中的每一个图像构成的图像对:

分别对该图像对中的两个图像进行第一扰动,以获得两个扰动图像;以及

响应于确定经所述图像处理模型处理后该两个扰动图像分别包括至少一个检测框,并且该两个扰动图像中的第一扰动图像中的至少一个检测框与第二扰动图像中的至少一个检测框分别对应,确定将该图像对中的增强图像添加至所述第二图像集。

14. 根据权利要求10所述的装置,其中,所述第一获取子单元包括:

第一图像特征获取单元,被配置用于通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征;

第二图像特征获取单元,被配置用于通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征;以及

图像添加单元,被配置用于响应于所述第三图像的图像特征与所述第一样本图像的图像特征之间的相似度大于第四预设阈值,将所述第三图像添加至所述第一图像子集。

15. 根据权利要求14所述的装置,其中,所述第一图像特征获取单元包括:

第一特征获取单元,被配置用于获得所述图像处理模型中的特征提取网络针对所述第一样本图像所提取的第一特征;

第一位置特征获取单元,被配置用于响应于确定经所述图像处理模型处理后所述第一样本图像包括至少一个检测框,对所述第一样本图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

第一图像特征获取子单元,被配置用于基于所述第一特征和所述第一样本图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第一样本图像的图像特征。

16. 根据权利要求14所述的装置,其中,所述第二图像特征获取单元包括:

第二特征获取单元,被配置用于获得所述图像处理模型中的特征提取网络针对所述第三图像所提取的第二特征;

第二位置特征获取单元,被配置用于响应于确定经所述图像处理模型处理后所述第三图像包括至少一个检测框,对所述第三图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

第二图像特征获取子单元,被配置用于基于所述第二特征和所述第三图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第三图像的图像特征。

17. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-8中任一项所述的方法。

18.一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据权利要求1-8中任一项所述的方法。

19.一种计算机程序产品,包括计算机程序,其中,所述计算机程序在被处理器执行时实现权利要求1-8中任一项所述的方法。

图像处理方法、装置、电子设备和存储介质

技术领域

[0001] 本公开涉及人工智能技术领域,具体为深度学习、图像处理、计算机视觉技术领域,可应用于OCR等场景,具体涉及一种图像处理方法、装置、电子设备、计算机可读存储介质和计算机程序产品。

背景技术

[0002] 人工智能是研究使计算机来模拟人的某些思维过程和智能行为(如学习、推理、思考、规划等)的学科,既有硬件层面的技术也有软件层面的技术。人工智能硬件技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理等技术;人工智能软件技术主要包括计算机视觉技术、语音识别技术、自然语言处理技术以及机器学习/深度学习、大数据处理技术、知识图谱技术等几大方向。

[0003] 基于人工智能的图像处理技术中,采用经训练的图像处理模型处理图像,以预测该图像类别或者该图像中所包括的对象的检测框。其中,如何提升图像处理模型的预测能力,以获得准确度较高的预测结果,是人们永远关注的问题。

[0004] 在此部分中描述的方法不一定是之前已经设想到或采用的方法。除非另有指明,否则不应假定此部分中描述的任何方法仅因其包括在此部分中就被认为是现有技术。类似地,除非另有指明,否则此部分中提及的问题不应认为在任何现有技术中已被公认。

发明内容

[0005] 本公开提供了一种图像处理方法、装置、电子设备、计算机可读存储介质和计算机程序产品。

[0006] 根据本公开的一方面,提供了一种图像处理方法,包括:获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及基于所述第一图像子集,调整所述图像处理模型。

[0007] 根据本公开的另一方面,提供了一种图像处理装置,包括:第一获取单元,被配置用于获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;第二获取单元,被配置用于基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及调整单元,被配置用于基于所述第一图像子集,调整所述图像处理模型。

[0008] 根据本公开的另一方面,提供了一种电子设备,包括:至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中所述存储器存储有可被所述至少一个处理器执

行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行根据本公开的实施例所述的方法。

[0009] 根据本公开的另一方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据本公开的实施例所述的方法。

[0010] 根据本公开的另一方面,提供了一种计算机程序产品,包括计算机程序,其中,所述计算机程序在被处理器执行时实现根据本公开的实施例所述的方法。

[0011] 根据本公开的一个或多个实施例,可以针对图像处理模型对应的困难样本集进行精准挖掘,基于第一图像集挖掘出与困难样本集中的样本图像相应的图像,使得通过该图像训练后的图像处理模型,能够提高针对难样本进行预测后所获得的预测结果的准确性。

[0012] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其他特征将通过以下的说明书而变得容易理解。

附图说明

[0013] 附图示例性地示出了实施例并且构成说明书的一部分,与说明书的文字描述一起用于讲解实施例的示例性实施方式。所示出的实施例仅出于例示的目的,并不限制权利要求的范围。在所有附图中,相同的附图标记指代类似但不一定相同的要素。

[0014] 图1示出了根据本公开的实施例的可以在其中实施本文描述的各种方法的示例性系统的示意图;

[0015] 图2示出了根据本公开的实施例的图像处理方法的流程图;

[0016] 图3示出了根据本公开的实施例的图像处理方法中基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集的过程的流程图;

[0017] 图4示出了根据本公开的实施例的图像处理方法中针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集的过程的流程图;

[0018] 图5示出了根据本公开的实施例的图像处理方法中针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集的过程的流程图;

[0019] 图6示出了根据本公开的实施例的图像处理方法中从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集的过程的流程图;

[0020] 图7示出了根据本公开的实施例的图像处理方法中通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征的过程的流程图;

[0021] 图8示出了根据本公开的实施例的图像处理方法中通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征的过程的流程图;

[0022] 图9示出了根据本公开的实施例的图像处理装置的结构框图;以及

[0023] 图10示出了能够用于实现本公开的实施例的示例性电子设备的结构框图。

具体实施方式

[0024] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0025] 在本公开中,除非另有说明,否则使用术语“第一”、“第二”等来描述各种要素不意图限定这些要素的位置关系、时序关系或重要性关系,这种术语只是用于将一个要素与另一要素区分开。在一些示例中,第一要素和第二要素可以指向该要素的同一实例,而在某些情况下,基于上下文的描述,它们也可以指代不同实例。

[0026] 在本公开中对各种所述示例的描述中所使用的术语只是为了描述特定示例的目的,而并非旨在进行限制。除非上下文另外明确地表明,如果不特意限定要素的数量,则该要素可以是一个也可以是多个。此外,本公开中所使用的术语“和/或”涵盖所列出的项目中的任何一个以及全部可能的组合方式。

[0027] 下面将结合附图详细描述本公开的实施例。

[0028] 图1示出了根据本公开的实施例可以将本文描述的各种方法和装置在其中实施的示例性系统100的示意图。参考图1,该系统100包括一个或多个客户端设备101、102、103、104、105和106、服务器120以及将一个或多个客户端设备耦接到服务器120的一个或多个通信网络110。客户端设备101、102、103、104、105和106可以被配置为执行一个或多个应用程序。

[0029] 在本公开的实施例中,服务器120可以运行使得能够执行根据本公开的图像处理方法的一个或多个服务或软件应用。

[0030] 在某些实施例中,服务器120还可以提供其他服务或软件应用,这些服务或软件应用可以包括非虚拟环境和虚拟环境。在某些实施例中,这些服务可以作为基于web的服务或云服务提供,例如在软件即服务(SaaS)模型下提供给客户端设备101、102、103、104、105和/或106的用户。

[0031] 在图1所示的配置中,服务器120可以包括实现由服务器120执行的功能的一个或多个组件。这些组件可以包括可由一个或多个处理器执行的软件组件、硬件组件或其组合。操作客户端设备101、102、103、104、105和/或106的用户可以依次利用一个或多个客户端应用程序来与服务器120进行交互以利用这些组件提供的服务。应当理解,各种不同的系统配置是可能的,其可以与系统100不同。因此,图1是用于实施本文所描述的各种方法的系统的一个示例,并且不旨在进行限制。

[0032] 用户可以使用客户端设备101、102、103、104、105和/或106来接收根据本公开的图像处理方法获得的经调整的图像处理模型预测的预测结果。客户端设备可以提供使客户端设备的用户能够与客户端设备进行交互的接口。客户端设备还可以经由该接口向用户输出信息。尽管图1仅描绘了六种客户端设备,但是本领域技术人员将能够理解,本公开可以支持任何数量的客户端设备。

[0033] 客户端设备101、102、103、104、105和/或106可以包括各种类型的计算机设备,例如便携式手持设备、通用计算机(诸如个人计算机和膝上型计算机)、工作站计算机、可穿戴设备、智能屏设备、自助服务终端设备、服务机器人、游戏系统、瘦客户端、各种消息收发设备、传感器或其他感测设备等。这些计算机设备可以运行各种类型和版本的软件应用程序和操作系统,例如MICROSOFT Windows、APPLE iOS、类UNIX操作系统、Linux或类Linux操作系统(例如GOOGLE Chrome OS);或包括各种移动操作系统,例如MICROSOFT Windows Mobile OS、iOS、Windows Phone、Android。便携式手持设备可以包括蜂窝电话、智能电话、平板电脑、个人数字助理(PDA)等。可穿戴设备可以包括头戴式显示器(诸如智能眼镜)和其

他设备。游戏系统可以包括各种手持式游戏设备、支持互联网的游戏设备等。客户端设备能够执行各种不同的应用程序,例如各种与Internet相关的应用程序、通信应用程序(例如电子邮件应用程序)、短消息服务(SMS)应用程序,并且可以使用各种通信协议。

[0034] 网络110可以是本领域技术人员熟知的任何类型的网络,其可以使用多种可用协议中的任何一种(包括但不限于TCP/IP、SNA、IPX等)来支持数据通信。仅作为示例,一个或多个网络110可以是局域网(LAN)、基于以太网的网络、令牌环、广域网(WAN)、因特网、虚拟网络、虚拟专用网络(VPN)、内部网、外部网、区块链网络、公共交换电话网(PSTN)、红外网络、无线网络(例如蓝牙、WIFI)和/或这些和/或其他网络的任意组合。

[0035] 服务器120可以包括一个或多个通用计算机、专用服务器计算机(例如PC(个人计算机)服务器、UNIX服务器、中端服务器)、刀片式服务器、大型计算机、服务器群集或任何其他适当的布置和/或组合。服务器120可以包括运行虚拟操作系统的一个或多个虚拟机,或者涉及虚拟化的其他计算架构(例如可以被虚拟化以维护服务器的虚拟存储设备的逻辑存储设备的一个或多个灵活池)。在各种实施例中,服务器120可以运行提供下文所描述的功能的一个或多个服务或软件应用。

[0036] 服务器120中的计算单元可以运行包括上述任何操作系统以及任何商业上可用的服务器操作系统的一个或多个操作系统。服务器120还可以运行各种附加服务器应用程序和/或中间层应用程序中的任何一个,包括HTTP服务器、FTP服务器、CGI服务器、JAVA服务器、数据库服务器等。

[0037] 在一些实施方式中,服务器120可以包括一个或多个应用程序,以分析和合并从客户端设备101、102、103、104、105和/或106的用户接收的数据馈送和/或事件更新。服务器120还可以包括一个或多个应用程序,以经由客户端设备101、102、103、104、105和/或106的一个或多个显示设备来显示数据馈送和/或实时事件。

[0038] 在一些实施方式中,服务器120可以为分布式系统的服务器,或者是结合了区块链的服务器。服务器120也可以是云服务器,或者是带人工智能技术的智能云计算服务器或智能云主机。云服务器是云计算服务体系中的一项主机产品,以解决传统物理主机与虚拟专用服务器(VPS,Virtual Private Server)服务中存在的管理难度大、业务扩展性弱的缺陷。

[0039] 系统100还可以包括一个或多个数据库130。在某些实施例中,这些数据库可以用于存储数据和其他信息。例如,数据库130中的一个或多个可用于存储诸如音频文件和视频文件的信息。数据库130可以驻留在各种位置。例如,由服务器120使用的数据库可以在服务器120本地,或者可以远离服务器120且可以经由基于网络或专用的连接与服务器120通信。数据库130可以是不同的类型。在某些实施例中,由服务器120使用的数据库例如可以是关系数据库。这些数据库中的一个或多个可以响应于命令而存储、更新和检索到数据库以及来自数据库的数据。

[0040] 在某些实施例中,数据库130中的一个或多个还可以由应用程序使用来存储应用程序数据。由应用程序使用的数据库可以是不同类型的数据库,例如键值存储库,对象存储库或由文件系统支持的常规存储库。

[0041] 图1的系统100可以以各种方式配置和操作,以使得能够应用根据本公开所描述的各种方法和装置。

[0042] 在相关技术中,通过针对图像处理模型训练过程中的样本分布进行调整,以提升模型的训练效果,例如,通过数据挖掘以调整训练图像集中的正负样本的比例,从而调整样本分布。然而,该方法无法针对模型的困难样本进行精准挖掘,使得经训练的模型针对其进行预测后所获得的预测结果准确性不高的困难样本的预测能力的提升。

[0043] 根据本公开的一方面,提供了一种图像处理方法。参看图2,根据本公开的图像处理方法200包括:

[0044] 步骤S210:获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;

[0045] 步骤S220:基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似性大于第二预设阈值;以及

[0046] 步骤S230:基于所述第一图像子集,调整所述图像处理模型。

[0047] 通过针对图像处理模型的困难样本集,基于第一图像集,挖掘出与之相应的第一图像子集,该第一图像子集中的每一个图像与困难样本集中的样本图像相似,实现针对图像处理模型对应的困难样本集进行精准挖掘,使得通过基于第一图像子集调整后的图像处理模型,能够提高针对困难样本进行预测后所获得的预测结果的准确性。

[0048] 在一些实施例中,图像处理模型可以是任意用于处理图像的模型。例如,人脸识别模型、文字识别模型等等,在此并不限定。

[0049] 在一些实施例中,采用预训练数据集对图像处理模型进行了预训练。例如,文字识别模型,采用预设的字符图像集进行了预训练。

[0050] 预训练数据集往往是从开源数据库中获得的数据集,其中包括各种图像质量的图像,例如模糊的、被涂改的图像等等。即使采用这些图像对图像处理模型进行训练后,经训练的图像处理模型对其预测后所获得的预测结果的置信度仍然不高,即使再采用这些图像对图像处理模型进行训练,也无法提升图像处理模型的预测能力。当图像处理模型在应用过程中遇到同类型的图像时,仍然无法获得准确的预测结果。

[0051] 为此,根据本公开的实施例中,将图像处理模型在训练或者预测过程中针对其进行预测,所获得的预测结果的准确性(或者置信度)低于一定阈值(第一预设阈值)的图像,添加到困难样本集,以针对这些图像区别于该困难样本集的第一图像集中挖掘出与之相似的图像,对模型进行调整(优化训练),使得经调整后的图像处理模型的预测能力得到提升,当再遇到与困难样本集中的图像相应的图像时,能够获得准确的预测结果。

[0052] 在一些实施例中,第一图像集可以包括从网络上爬取的图像,或者采用其他方式收集的图像。例如,通过利用图像处理模型处理图像后,收集的图像处理模型无法获得其准确预测结果的图像。

[0053] 在一些实施例中,在步骤S220中,通过从第一图像集中,获得与困难样本集中的第一样本图像相应的第一图像子集。其中,第一样本图像可以是困难样本集中的任意一个样本图像。

[0054] 在一些实施例中,如图3所示,步骤220、基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集包括:

[0055] 步骤S310:针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集,所述第二图像集包括与所述第一图像集中的每一图像相应的至少一个增强图像;以及

[0056] 步骤S320:从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集。

[0057] 在基于第一图像集进行数据挖掘获得与困难那样本集中的第一样本图像相应的第一图像子集的过程中,通过对第一图像集中的数据进行数据增强后,获得的增强数据进行挖掘,获得第一图像子集,扩大数据挖掘过程中的数据量,提升所获得的难样本的相应数据的数据量。

[0058] 在一些实施例中,在所述第二图像集中与所述第一图像集中的每一图像相应的至少一个增强图像中的每一个增强图像与该图像之间的相似度大于第二预设阈值。

[0059] 在获得第二图像集的过程中,针对第一图像集中的每一个图像,使第二图像集中与该图像相应的每一个增强图像与该图像之间的相似度大于第二预设阈值,避免获得第二图像集的过程中,对第一图像集中的图像所做的增强处理的强度过大,改变该图像的特征,从而使图像标签改变(例如,从对应于正样本变为对应于负样本),进而影响后续采用第一图像集调整图像处理模型的过程。

[0060] 在一些实施例中,通过编码器对第一图像集的图像和对该图像进行增强处理后获得的多个图像分别进行编码,基于各个图像对应的编码向量,计算图像之间的相似度。

[0061] 在一些实施例中,如图4所示,步骤S310、针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集包括:

[0062] 步骤S410:针对获得所述第一图像集中的第一图像,响应于确定经所述图像处理模型处理后所述第一图像包括至少一个检测框,获得所述第一图像对应的多个增强图像;以及

[0063] 步骤S420:针对所述多个增强图像中的第一增强图像,响应于确定经所述图像处理模型处理后所述第一增强图像包括至少一个检测框,并且所述第一增强图像中的至少一个检测框与所述第一图像中的至少一个分别对应,确定将所述第一增强图像添加至所述第二图像集。

[0064] 在获得第二图像集的过程中,针对第一图像集中的第一图像,通过获得图像处理模型对该第一图像和与该第一图像对应的增强图像进行处理后获得各自包括的至少一个检测框,响应于基于该第一图像和与该第一图像对应的增强图像各自包括的至少一个检测框一一对应,将该与第一图像对应的增强图像确定为第二图像集中与该第一图像对应的增强图像,使得在获得第二图像集中与第一图像对应的至少一个增强图像的过程中,是基于图像处理模型处理该第一图像获得的至少一个检测框进行的,即图像增强过程中未对检测框内所图像特征造成影响,进一步避免数据增强过程中改变第一图像中检测框内的图像特征,从而使图像标签改变(例如,从对应于正样本变为对应于负样本),进而影响后续采用第一图像集调整图像处理模型的过程。

[0065] 在一些实施例中,通过对第一图像集中的图像进行旋转、变形、添加噪声、颜色扰动等,获得图像对应的多个增强图像。

[0066] 在一些实施例中,针对第一图像中的至少一个检测框中的每一个检测框,响应于

确定第一增强图像中的至少一个检测框中具有与该检测框的类别、位置分别相同的检测框,确定第一图像中的至少一个检测框与第一增强图像中的至少一个检测框分别对应。

[0067] 在一些实施例中,如图5所示,步骤S310、针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集包括:

[0068] 步骤S510:针对获得所述第一图像集中的第二图像,响应于确定经所述图像处理模型处理后所述第二图像包括至少一个检测框,获得所述第二图像对应的多个增强图像;以及

[0069] 步骤S520:针对所述第二图像和所述多个增强图像中的每一个图像构成的图像对,分别对该图像对中的两个图像进行第一扰动,以获得两个扰动图像,并且响应于确定经所述图像处理模型处理后该两个扰动图像分别包括至少一个检测框,并且该两个扰动图像中的第一扰动图像中的至少一个检测框与第二扰动图像中的至少一个检测框分别对应,确定将该图像对中的增强图像添加至所述第二图像集。

[0070] 在获得第二图像集的过程中,针对第一图像集中的第二图像和与第二图像相应的增强图像分别进行扰动,并利用图像处理模型对经扰动获得的扰动图像进行处理,当两个扰动图分别包括至少一个检测框,并且该至少一个检测框一一对应,则将该第二图像对应的增强图像确定为第二图像集中与该第二图像对应的增强图像。使得所获得的第二图像集中的与第二图像对应的至少一个增强图像,是即使在经过扰动后,仍然未对检测框内所图像特征造成影响的增强图像,进一步避免增强过程中的噪音过高,使使图像标签改变(例如,从对应于正样本变为对应于负样本),进而影响后续采用第一图像集调整图像处理模型的过程。

[0071] 在一些实施例中,如图6所示,步骤S330、从所述第一图像集和所述第二图像集组成的第三图像集中,获得所述第一图像子集包括:

[0072] 步骤S610:通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征;

[0073] 步骤S620:通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征;以及

[0074] 步骤S630:响应于所述第三图像的图像特征与所述第一样本图像的图像特征之间的相似度大于第一相似度阈值,将所述第三图像添加至所述第一图像子集。

[0075] 在从第三图像集中获得与困难样本集中的第一样本图像对应的第一图像子集的过程中,通过利用图像处理模型处理第一样本图像和第三图像集中的图像后获得的图像特征,获得图像和样本之间的相似度,实现第一图像子集的获得。同时,该图像特征是利用图像处理模型获得的,其与图像处理模型的任务相关性高,提升所获得的第一图像子集的准确性。

[0076] 在一些实施例中,在步骤S610和步骤S620中,直接将通过图像处理模型的特征提取网络对第一样本图像和第三图像进行特征提取,所获得的特征作为第一样本图像的图像特征和第三图像的图像特征。

[0077] 在一些实施例中,如图7所示,步骤S610、通过利用所述图像处理模型处理所述第一样本图像,获得所述第一样本图像的图像特征包括:

[0078] 步骤S710:获得所述图像处理模型中的特征提取网络针对所述第一样本图像所提

取的第一特征；

[0079] 步骤S720:响应于确定经所述图像处理模型处理后所述第一样本图像包括至少一个检测框,对所述第一样本图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

[0080] 步骤S730:基于所述第一特征和所述第一样本图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第一样本图像的图像特征。

[0081] 在获得第一样本图像的图像特征的过程中,通过其中所包括的至少一个检测框的位置特征,获得图像特征,使得所获得的图像特征融合了检测框的位置信息,该位置信息对困难样本图像的特征进行了增强,使得基于第一特征和位置特征获得的图像特征具有对第一特征进行了校准的效果,提升所获得的第一样本图像的图像特征的准确度,从而提升所获得的第一图像子集的准确性。

[0082] 在一些实施例中,在步骤S730中,通过将第一特征与至少一个检测框中的每一个检测框的位置特征进行融合,以获得第一样本图像的图像特征。

[0083] 在一些实施例中,如图8所示,步骤S620、通过利用所述图像处理模型处理所述第三图像集中的第三图像,获得所述第三图像的图像特征包括:

[0084] 步骤S810:获得所述图像处理模型中的特征提取网络针对所述第三图像所提取的第二特征;

[0085] 步骤S820:响应于确定经所述图像处理模型处理后所述第三图像包括至少一个检测框,对所述第三图像中的至少一个检测框中的每一个检测框的位置信息进行编码,以获得该检测框的位置特征;以及

[0086] 步骤S830:基于所述第二特征和所述第三图像中的至少一个检测框中的每一个检测框的位置特征,获得所述第三图像的图像特征。

[0087] 在获得第三图像的图像特征的过程中,通过其中所包括的至少一个检测框的位置特征,获得图像特征,使得所获得的图像特征融合了检测框的位置信息,该位置信息对第三图像的特征进行了增强,使得基于第二特征和位置特征获得的图像特征具有对第二特征进行了校准的效果,提升所获得的第三图像的图像特征的准确度,从而提升所获得的第一图像子集的准确性。

[0088] 在一些实施例中,在步骤S230中、在获得第一图像子集之后,对第一图像子集中的每一个个图像进行标注,以获得该图像的标注标签,基于该图像及其标注标签,训练图像处理模型,以调整图像处理模型。

[0089] 根据本公开的另一方面,还提供了一种图像处理装置,如图9所示,装置900包括:第一获取单元910,被配置用于获得第一图像集和经预训练的图像处理模型,所述图像处理模型具有相应的困难样本集,所述图像处理模型对所述困难样本集中的每一个样本图像的预测结果的准确性低于第一预设阈值;第二获取单元920,被配置用于基于所述第一图像集,获得与所述困难样本集中的第一样本图像相应的第一图像子集,所述第一图像子集中的每一个图像与所述第一样本图像的相似度大于第二预设阈值;以及调整单元930,被配置用于基于所述第一图像子集,调整所述图像处理模型。

[0090] 在一些实施例中,所述第二获取单元包括:数据增强处理单元,被配置用于针对所述第一图像集中的每一图像进行数据增强处理,以获得第二图像集,所述第二图像集包括

与所述第一图像集中的每一图像相应的至少一个增强图像；以及第一获取子单元，被配置用于从所述第一图像集和所述第二图像集组成的第三图像集中，获得所述第一图像子集。

[0091] 在一些实施例中，在所述第二图像集中与所述第一图像集中的每一图像相应的至少一个增强图像中的每一个增强图像与该图像之间的相似度大于第三预设阈值。

[0092] 在一些实施例中，所述数据增强处理单元包括：第二获取子单元，被配置用于针对获得所述第一图像集中的第一图像，响应于确定经所述图像处理模型处理后所述第一图像包括至少一个检测框，获得所述第一图像对应的多个增强图像；以及第一确定单元，被配置用于针对所述多个增强图像中的第一增强图像，响应于确定经所述图像处理模型处理后所述第一增强图像包括至少一个检测框，并且所述第一增强图像中的至少一个检测框与所述第一图像中的至少一个分别对应，确定将所述第一增强图像添加至所述第二图像集。

[0093] 在一些实施例中，所述数据增强处理单元包括：第三获取子单元，被配置用于针对获得所述第一图像集中的第二图像，响应于确定经所述图像处理模型处理后所述第二图像包括至少一个检测框，获得所述第二图像对应的多个增强图像；以及图像对处理单元，被配置用于针对所述第二图像和所述多个增强图像中的每一个图像构成的图像对：分别对该图像对中的两个图像进行第一扰动，以获得两个扰动图像；以及响应于确定经所述图像处理模型处理后该两个扰动图像分别包括至少一个检测框，并且该两个扰动图像中的第一扰动图像中的至少一个检测框与第二扰动图像中的至少一个检测框分别对应，确定将该图像对中的增强图像添加至所述第二图像集。

[0094] 在一些实施例中，所述第一获取子单元包括：第一图像特征获取单元，被配置用于通过利用所述图像处理模型处理所述第一样本图像，获得所述第一样本图像的图像特征；第二图像特征获取单元，被配置用于通过利用所述图像处理模型处理所述第三图像集中的第三图像，获得所述第三图像的图像特征；以及图像添加单元，被配置用于响应于所述第三图像的图像特征与所述第一样本图像的图像特征之间的相似度大于第四预设阈值，将所述第三图像添加至所述第一图像子集。

[0095] 在一些实施例中，所述第一图像特征获取单元包括：第一特征获取单元，被配置用于获得所述图像处理模型中的特征提取网络针对所述第一样本图像所提取的第一特征；第一位置特征获取单元，被配置用于响应于确定经所述图像处理模型处理后所述第一样本图像包括至少一个检测框，对所述第一样本图像中的至少一个检测框中的每一个检测框的位置信息进行编码，以获得该检测框的位置特征；以及第一图像特征获取子单元，被配置用于基于所述第一特征和所述第一样本图像中的至少一个检测框中的每一个检测框的位置特征，获得所述第一样本图像的图像特征。

[0096] 在一些实施例中，所述第二图像特征获取单元包括：第二特征获取单元，被配置用于获得所述图像处理模型中的特征提取网络针对所述第三图像所提取的第二特征；第二位置特征获取单元，被配置用于响应于确定经所述图像处理模型处理后所述第三图像包括至少一个检测框，对所述第三图像中的至少一个检测框中的每一个检测框的位置信息进行编码，以获得该检测框的位置特征；以及第二图像特征获取子单元，被配置用于基于所述第二特征和所述第三图像中的至少一个检测框中的每一个检测框的位置特征，获得所述第三图像的图像特征。

[0097] 根据本公开的实施例，还提供了一种电子设备、一种可读存储介质和一种计算机

程序产品。

[0098] 参考图10,现将描述可以作为本公开的服务器或客户端的电子设备1000的结构框图,其是可以应用于本公开的各方面的硬件设备的示例。电子设备旨在表示各种形式的数字电子的计算机设备,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0099] 如图10所示,电子设备1000包括计算单元1001,其可以根据存储在只读存储器(ROM) 1002中的计算机程序或者从存储单元1008加载到随机访问存储器(RAM) 1003中的计算机程序,来执行各种适当的动作和处理。在RAM 1003中,还可存储电子设备1000操作所需的各种程序和数据。计算单元1001、ROM 1002以及RAM 1003通过总线1004彼此相连。输入/输出(I/O)接口1005也连接至总线1004。

[0100] 电子设备1000中的多个部件连接至I/O接口1005,包括:输入单元1006、输出单元1007、存储单元1008以及通信单元1009。输入单元1006可以是能向电子设备1000输入信息的任何类型的设备,输入单元1006可以接收输入的数字或字符信息,以及产生与电子设备的用户设置和/或功能控制有关的键信号输入,并且可以包括但不限于鼠标、键盘、触摸屏、轨迹板、轨迹球、操作杆、麦克风和/或遥控器。输出单元1007可以是能呈现信息的任何类型的设备,并且可以包括但不限于显示器、扬声器、视频/音频输出终端、振动器和/或打印机。存储单元1008可以包括但不限于磁盘、光盘。通信单元1009允许电子设备1000通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据,并且可以包括但不限于调制解调器、网卡、红外通信设备、无线通信收发机和/或芯片组,例如蓝牙TM设备、802.11设备、WiFi设备、WiMax设备、蜂窝通信设备和/或类似物。

[0101] 计算单元1001可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元1001的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元1001执行上文所描述的各个方法和处理,例如方法200。例如,在一些实施例中,方法200可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元1008。在一些实施例中,计算机程序的部分或者全部可以经由ROM 1002和/或通信单元1009而被载入和/或安装到电子设备1000上。当计算机程序加载到RAM 1003并由计算单元1001执行时,可以执行上文描述的方法200的一个或多个步骤。备选地,在其他实施例中,计算单元1001可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行方法200。

[0102] 本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、复杂可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出

装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0103] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码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或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0104] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0105] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0106] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0107] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0108] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本公开中记载的各步骤可以并行地执行、也可以顺序地或以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0109] 虽然已经参照附图描述了本公开的实施例或示例,但应理解,上述的方法、系统和设备仅仅是示例性的实施例或示例,本发明的范围并不由这些实施例或示例限制,而是仅由授权后的权利要求书及其等同范围来限定。实施例或示例中的各种要素可以被省略或者可由其等同要素替代。此外,可以通过不同于本公开中描述的次序来执行各步骤。进一步

地,可以以各种方式组合实施例或示例中的各种要素。重要的是随着技术的演进,在此描述的很多要素可以由本公开之后出现的等同要素进行替换。

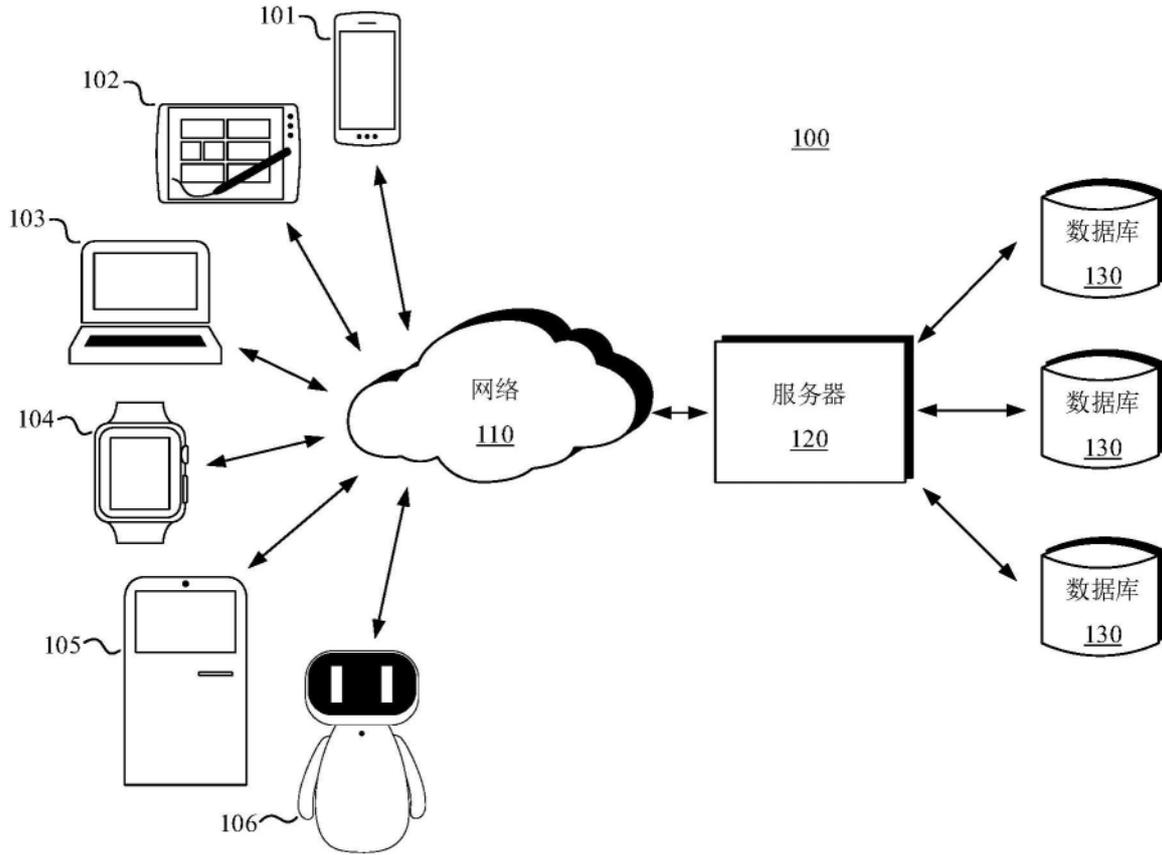


图1

200

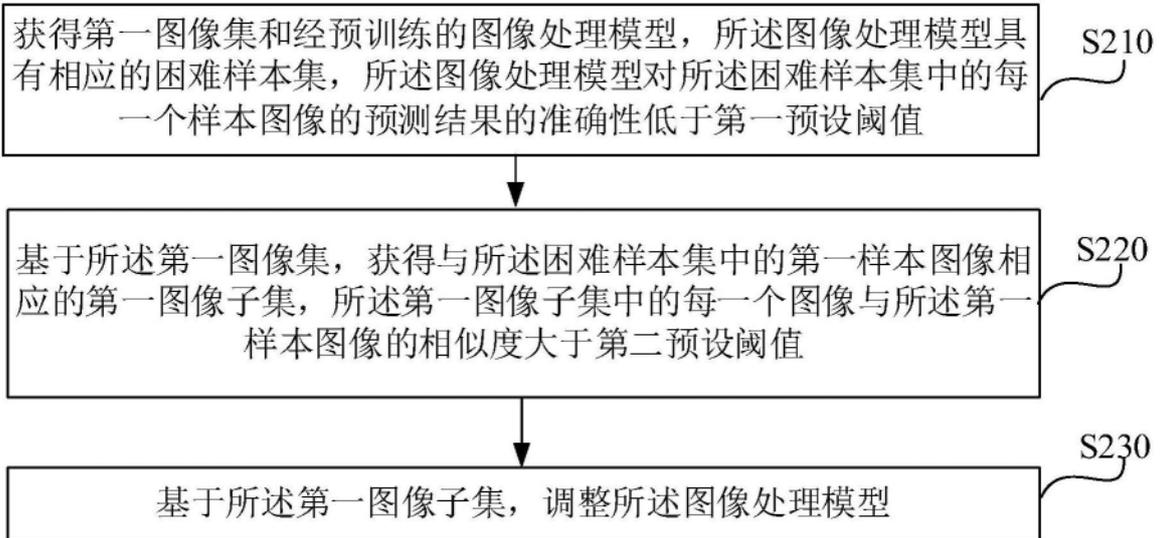


图2

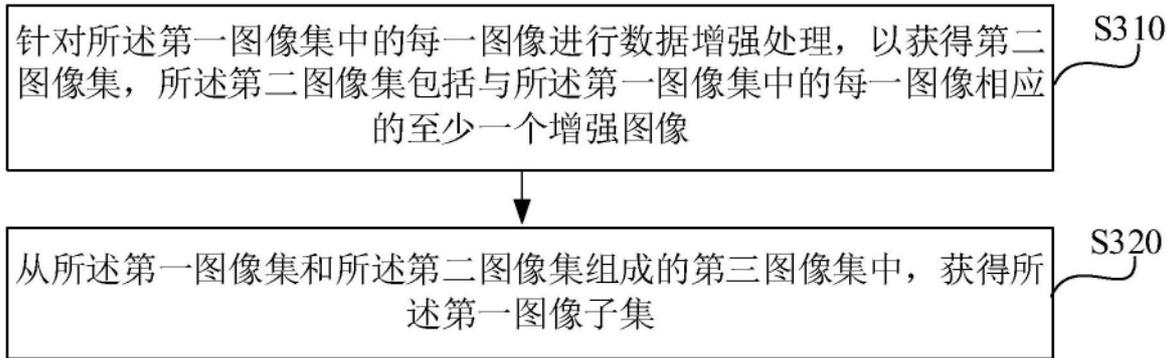


图3

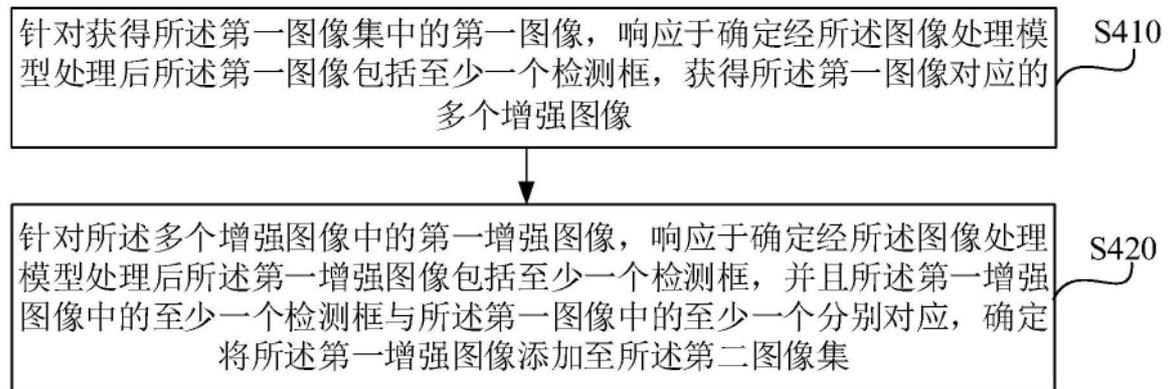


图4

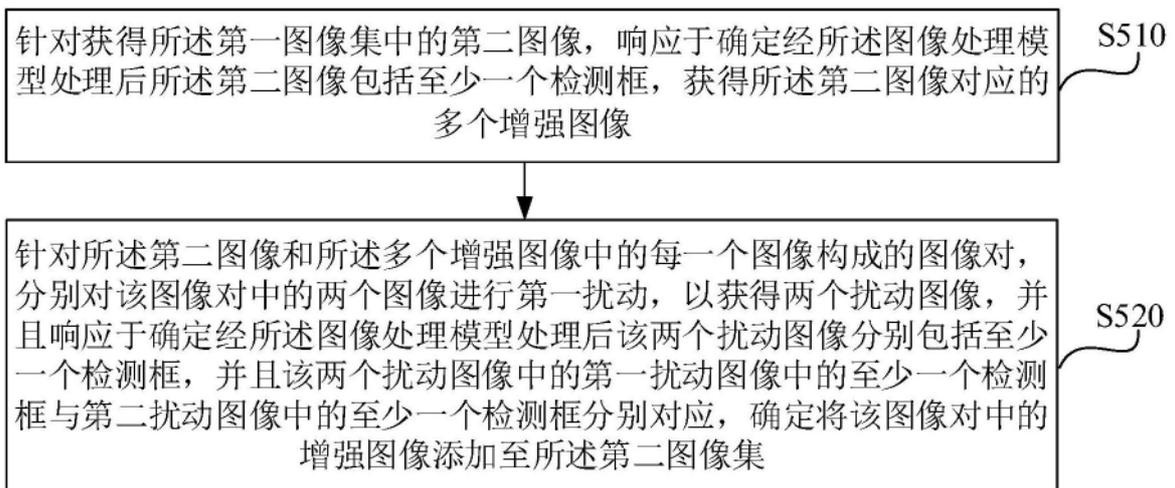


图5

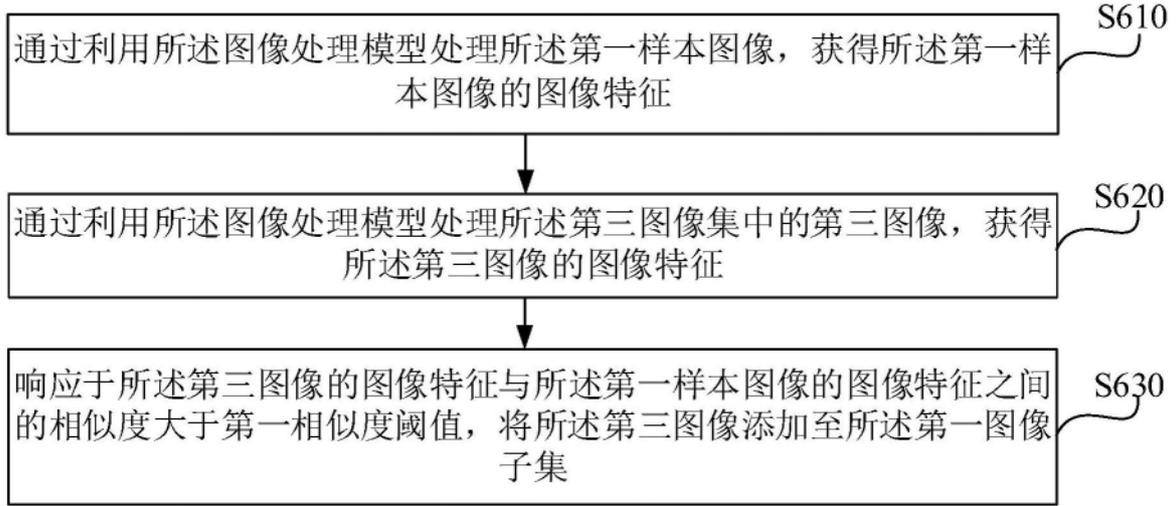


图6

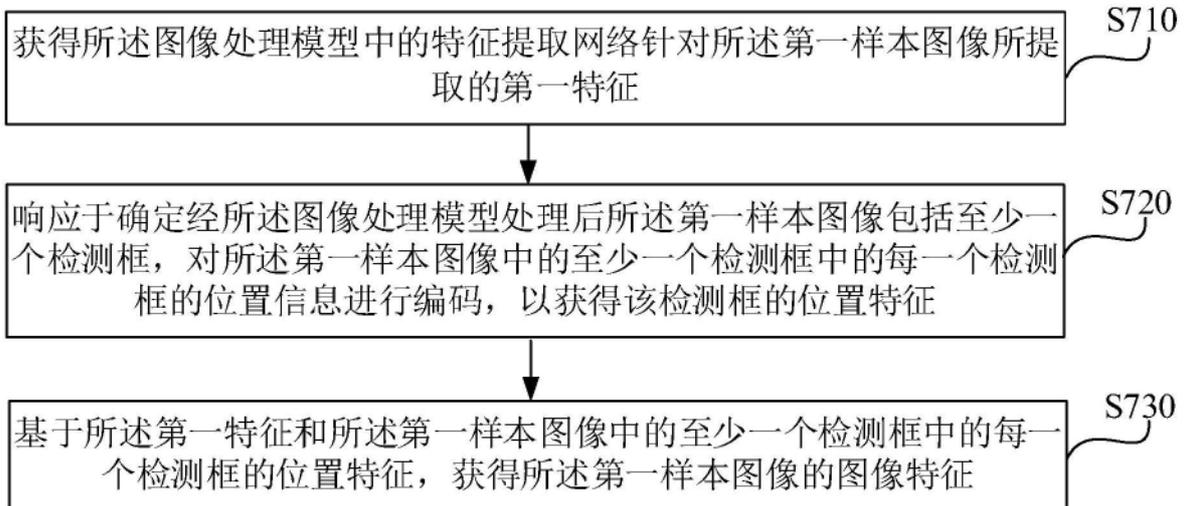


图7

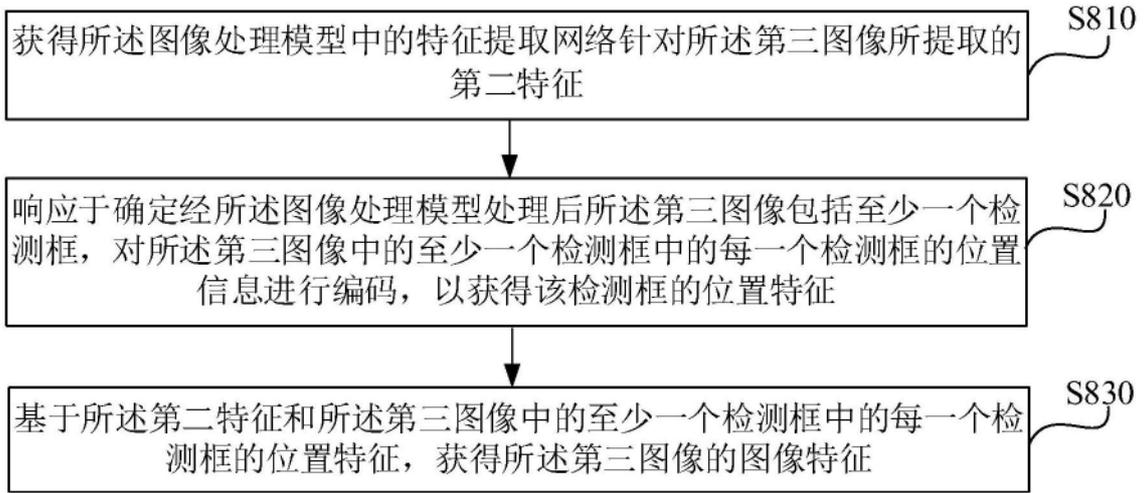


图8

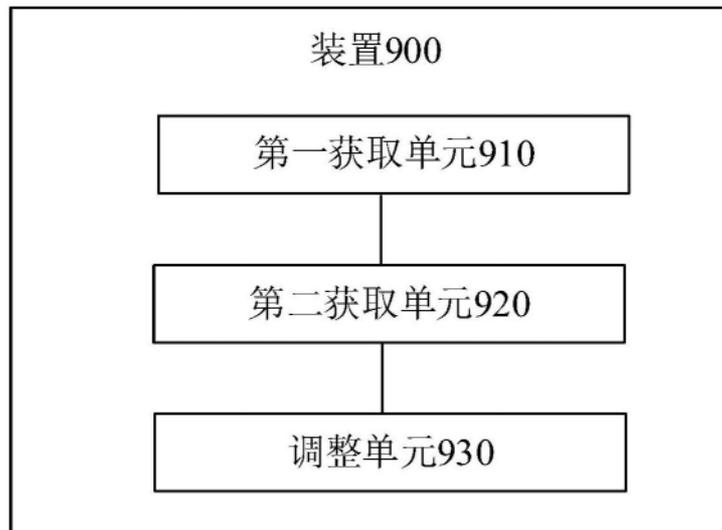


图9

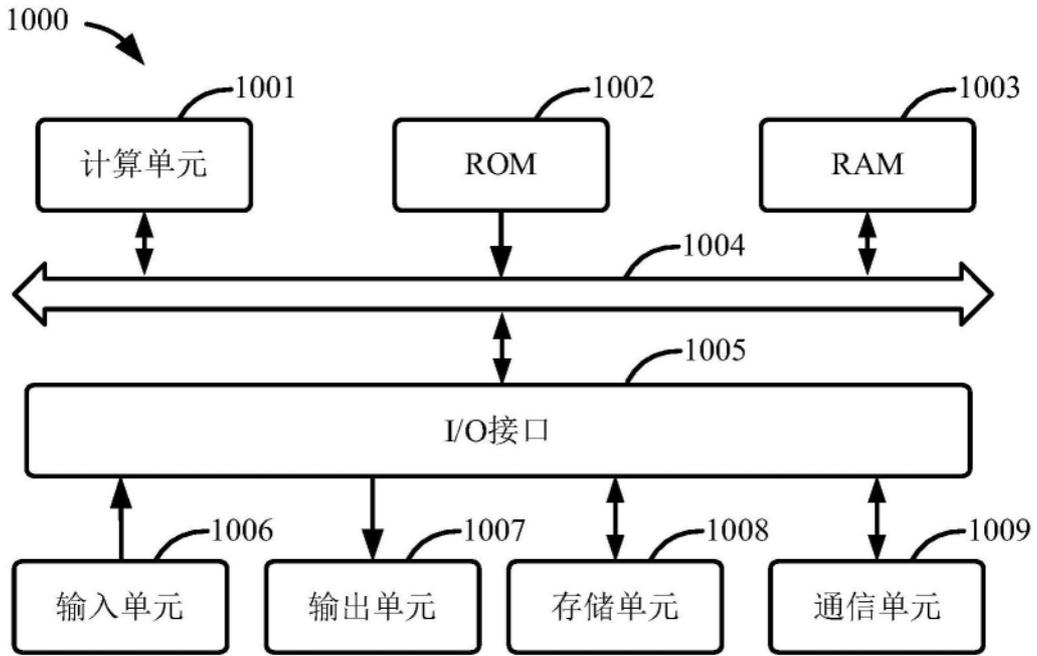


图10