



(12) 发明专利申请

(10) 申请公布号 CN 112883902 A

(43) 申请公布日 2021.06.01

(21) 申请号 202110272132.X

(22) 申请日 2021.03.12

(71) 申请人 百度在线网络技术(北京)有限公司
地址 100085 北京市海淀区上地十街10号
百度大厦三层

(72) 发明人 熊俊峰 王洋 周越 张欢
仲震宇

(74) 专利代理机构 北京市铸成律师事务所
11313
代理人 郭丽祥 杨瑾瑾

(51) Int. Cl.
G06K 9/00 (2006.01)
G06N 3/04 (2006.01)
G06N 3/08 (2006.01)

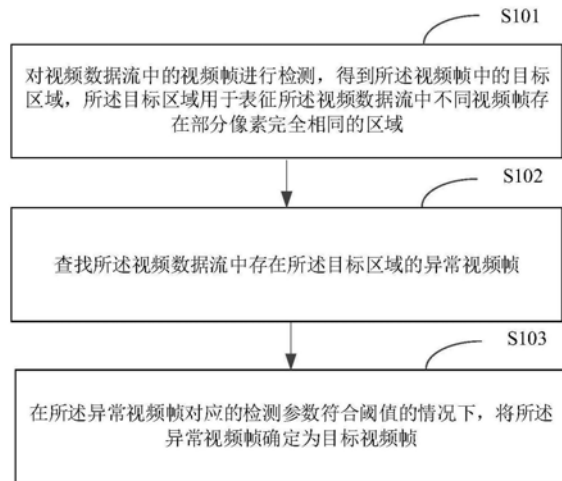
权利要求书2页 说明书9页 附图3页

(54) 发明名称

视频检测方法、装置、电子设备及存储介质

(57) 摘要

本公开提供了一种视频检测方法、装置、电子设备及存储介质,涉及人工智能、深度学习、计算机视觉、图像处理、人脸识别、肢体识别、伪造检测等领域。具体实现方案为:对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;查找所述视频数据流中存在所述目标区域的异常视频帧;在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。采用本公开,可以检测出视频数据流中被人编辑过(或称被伪造过)的异常视频帧。



1. 一种视频检测方法,包括:

对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;

查找所述视频数据流中存在所述目标区域的异常视频帧;

在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

2. 根据权利要求1所述的方法,其中,所述对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,包括:

对所述视频数据流中相邻的至少两个视频帧,分别提取关键区域;

对所述至少两个视频帧分别对应的关键区域进行像素点的特征比对,将比对得到的所述关键区域中存在所述部分像素完全相同的区域,作为所述目标区域。

3. 根据权利要求1或2所述的方法,其中,所述在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧,包括:

对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;

根据所述雷同分值,得到针对所述异常视频帧的像素异常率;

将所述像素异常率作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第一视频序列中存在的异常视频帧确定为目标视频帧。

4. 根据权利要求1或2所述的方法,其中,所述在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧,包括:

对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;

根据所述雷同分值,得到针对所述异常视频帧的像素异常率;

从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值;

将所述视频检测分值作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第二视频序列中存在的异常视频帧确定为目标视频帧。

5. 根据权利要求1或2所述的方法,还包括:

根据视频检测分值配置所述阈值;

其中,所述视频检测分值包括:不同像素异常率构成的第二视频序列在第一视频序列中的分值占比;所述第一视频序列为所述视频数据流中由所述异常视频帧构成的第一视频序列。

6. 根据权利要求5所述的方法,还包括:

根据所述阈值,定位出所述目标视频帧在所述视频数据流中的位置。

7. 一种视频检测装置,所述装置包括:

目标区域检测模块,用于对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;

异常视频查找模块,用于查找所述视频数据流中存在所述目标区域的异常视频帧;

目标视频确定模块,用于在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

8. 根据权利要求7所述的装置,其中,所述目标区域检测模块,用于:

对所述视频数据流中相邻的至少两个视频帧,分别提取关键区域;

对所述至少两个视频帧分别对应的关键区域进行像素点的特征比对,将比对得到的所述关键区域中存在所述部分像素完全相同的区域,作为所述目标区域。

9. 根据权利要求7或8所述的装置,其中,所述目标视频确定模块,用于:

对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;

根据所述雷同分值,得到针对所述异常视频帧的像素异常率;

将所述像素异常率作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第一视频序列中存在的异常视频帧确定为目标视频帧。

10. 根据权利要求7或8所述的装置,其中,所述目标视频确定模块,用于:

对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;

根据所述雷同分值,得到针对所述异常视频帧的像素异常率;

从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值;

将所述视频检测分值作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第二视频序列中存在的异常视频帧确定为目标视频帧。

11. 根据权利要求7或8所述的装置,还包括阈值配置模块,用于:

根据视频检测分值配置所述阈值;

其中,所述视频检测分值包括:不同像素异常率构成的第二视频序列在第一视频序列中的分值占比;所述第一视频序列为所述视频数据流中由所述异常视频帧构成的第一视频序列。

12. 根据权利要求11所述的装置,还包括定位模块,用于:

根据所述阈值,定位出所述目标视频帧在所述视频数据流中的位置。

13. 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1-6中任一项所述的方法。

14. 一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使计算机执行权利要求1-6中任一项所述的方法。

15. 一种计算机程序产品,包括计算机指令,该计算机指令被处理器执行时实现权利要求1-6中任一项所述的方法。

视频检测方法、装置、电子设备及存储介质

技术领域

[0001] 本公开涉及计算机处理领域,本公开尤其涉及人工智能、深度学习、计算机视觉、图像处理、人脸识别、肢体识别、伪造检测等领域。

背景技术

[0002] 随着计算机技术的发展,不仅图片及音频可以被伪造,甚至视频也可以被伪造。在视频的深度伪造处理中,可以用伪造过的图片来生成视频,比如,对一段视频内容通过更新元素(比如更换别人的脸)来达到伪造的目的,以利用各类应用(如银行客户端、门禁系统等)的安全漏洞,来实现攻击,从而给用户带来巨大的损失。

发明内容

[0003] 本公开提供了一种视频检测方法、装置、电子设备及存储介质。

[0004] 根据本公开的一方面,提供了一种视频检测方法,包括:

[0005] 对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;

[0006] 查找所述视频数据流中存在所述目标区域的异常视频帧;

[0007] 在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0008] 根据本公开的另一方面,提供了一种视频检测装置,包括:

[0009] 目标区域检测模块,用于对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;

[0010] 异常视频查找模块,用于查找所述视频数据流中存在所述目标区域的异常视频帧;

[0011] 目标视频确定模块,用于在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0012] 根据本公开的另一方面,提供了一种电子设备,包括:

[0013] 至少一个处理器;以及

[0014] 与该至少一个处理器通信连接的存储器;其中,

[0015] 该存储器存储有可被该至少一个处理器执行的指令,该指令被该至少一个处理器执行,以使该至少一个处理器能够执行本公开任意一实施例所提供的方法。

[0016] 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,该计算机指令用于使该计算机执行本公开任意一实施例所提供的方法。

[0017] 根据本公开的另一方面,提供了一种计算机程序产品,包括计算机指令,该计算机指令被处理器执行时实现本公开任意一实施例所提供的方法。

[0018] 采用本公开,可以对视频数据流中的视频帧进行检测,得到所述视频帧中的目标

区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;查找所述视频数据流中存在所述目标区域的异常视频帧;在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧,从而,可以检测出视频数据流中被人编辑过(或称被伪造过)的异常视频帧。

[0019] 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。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0020] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0021] 图1是根据本公开实施例的视频检测方法的流程示意图;

[0022] 图2是根据本公开实施例的视频检测方法的流程示意图;

[0023] 图3是根据本公开实施例的视频检测装置的组成结构示意图;

[0024] 图4是根据本公开实施例的视频检测装置的组成结构示意图;

[0025] 图5是用来实现本公开实施例的视频检测方法的电子设备的框图。

具体实施方式

[0026] 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和精神。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。

[0027] 本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合,例如,包括A、B、C中的至少一种,可以表示包括从A、B和C构成的集合中选择的任意一个或多个元素。本文中术语“第一”、“第二”表示指代多个类似的技术用语并对其进行区分,并不是限定顺序的意思,或者限定只有两个的意思,例如,第一特征和第二特征,是指代有两类/两个特征,第一特征可以为一个或多个,第二特征也可以为一个或多个。

[0028] 另外,为了更好的说明本公开,在下文的具体实施方式中给出了众多的具体细节。本领域技术人员应当理解,没有某些具体细节,本公开同样可以实施。在一些实例中,对于本领域技术人员熟知的方法、手段、元件和电路未作详细描述,以便于凸显本公开的主旨。

[0029] 针对视频的深度伪造而言,以门禁系统的应用场景为例,可以基于手机应用、门禁,闭路电视等摄像头场景下采集视频,得到视频数据流,该视频数据流中的视频帧为基于三原色(RGB)的图像。即便该图像可以结合包括结构光或者双目摄像头所采集的深度信息(Depth),以及结合红外和颜色照射等多模态的手段,在执行诸如人脸、肢体、动作的安全验证等功能,在RGB上的弱点也是无法规避的,方便了通过图像来伪造视频。尤其是在系统入侵时同时植入所伪造的视频,对用户的攻击威胁更为明显。这种基于图像来驱动视频,从而达到通过伪造的人脸、伪造的人物肢体、伪造的动作等实现欺骗性的攻击,由于攻击效果好,攻击成本低,对各类应用涉及相关安全验证功能来说,危害是巨大的。

[0030] 有鉴于此,基于图像来驱动视频的方案中,可以通过视频理解、视频图像卷积网络

训练的分类器,对视频数据流中伪造的人脸、伪造的人物肢体、伪造的动作等执行相关的检测,但是,该视频图像卷积网络有大量参数,在训练之前需要标注大量的数据,实现的成本高;该视频图像卷积网络通常容易过拟合(即不具备通用性),需要大量的调参操作,以改善其网络性能,效率低且适用场景相对少;运行该视频图像卷积网络的机器比较精密,造价比较昂贵。

[0031] 本公开中,可以通过合理的选择目标区域,基于该目标区域对整个视频数据流进行梳理,从而可以找出该视频数据流中哪些视频帧存在被编辑的痕迹,从而对这种被编辑的异常视频做出正确的判断。考虑到基于图片驱动的视频,是根据关键点驱动部分位置,不同视频帧画面会存在部分像素完全相同的情况,即使视频数据流存在一定随机噪声,但是正常视频(非编辑处理过的伪造视频)下像素同一的可能几乎等于零,则可以认为这种现象属于视频帧存在被该编辑的痕迹。

[0032] 根据本公开的实施例,提供了一种视频检测方法,图1是根据本公开实施例的视频检测方法的流程示意图,该方法可以应用于视频检测装置,例如,该装置可以部署于终端或服务器或其它处理设备执行的情况下,可以实现目标区域检测、异常视频帧查找、以及对异常视频帧评估后确定出目标视频帧等等处理。其中,终端可以为用户设备(UE, User Equipment)、移动设备、个人数字处理(PDA, Personal Digital Assistant)、手持设备、计算设备、车载设备、可穿戴设备等。在一些可能的实现方式中,该方法还可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。如图1所示,包括:

[0033] S101、对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域。

[0034] S102、查找所述视频数据流中存在所述目标区域的异常视频帧。

[0035] S103、在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0036] S101-S103的一示例中,对视频数据流中的视频帧进行检测,以查找该目标区域,比如,该目标区域可以为被伪造者人为编辑的区域,该被伪造者人为编辑的区域为:所述视频数据流中不同视频帧存在部分像素完全相同的区域。基于该目标区域对整个视频数据流进行梳理的过程中,可以查找到该视频数据流中存在所述目标区域的异常视频帧。还可以对该视频数据流中针对多个异常视频帧进行运算,以得到异常视频帧对应的检测参数,比如,视频帧序列像素异常的像素异常率,或进一步对不同像素异常率构成的异常视频序列打分所得到的视频检测分值,从而在检测参数符合阈值的情况下将所述异常视频帧确定为目标视频帧。针对该阈值而言,还可以根据该视频检测分值配置该阈值,以根据该阈值更精确的定位出该视频数据流中该目标视频帧的具体位置。

[0037] 采用本公开,可以对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;查找所述视频数据流中存在所述目标区域的异常视频帧;在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。由于“所述视频数据流中不同视频帧存在部分像素完全相同的区域”为人为编辑区域(或称伪造区域),则这个视频帧是异常的,需要对整个视频数据流进行检测,先找到这类目标区域(即该人为编辑区域)。之后,通过该目标区域梳理整个视频梳理流,将存在该目标区域的视频帧作为异常视频帧。

考虑异常视频帧的检测准确性,还需要在判断出该异常视频帧对应的检测参数符合阈值的情况下,将该异常视频帧最终确定为:通过检测所筛选得出的该目标视频帧,从而,可以准确的检测出视频数据流中被人编辑过(或称被伪造过)的异常视频帧。

[0038] 根据本公开的实施例,提供了一种视频检测方法,图2是根据本公开实施例的视频检测方法的流程示意图,如图2所示,包括:

[0039] S201、对所述视频数据流中相邻的至少两个视频帧,分别提取关键区域。

[0040] S202、对所述至少两个视频帧分别对应的关键区域进行像素点的特征比对,将比对得到的所述关键区域中存在所述部分像素完全相同的区域,作为所述目标区域。

[0041] S203、查找所述视频数据流中存在所述目标区域的异常视频帧。

[0042] S204、在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0043] S201-S204的一示例中,对视频数据流中的视频帧进行检测,以查找该目标区域。考虑到基于图像来驱动视频,是通过伪造的人脸、伪造的人物肢体、伪造的动作等实现欺骗性的攻击,因此,根据应用场景的需求,可以将人脸、人物肢体、伪造的动作(如指定的手势等)等作为该关键区域,而不是对所有的人体及所执行的动作进行检测。其中,该目标区域可以为被伪造者人为编辑的区域,该被伪造者人为编辑的区域为:所述视频数据流中不同视频帧存在部分像素完全相同的区域。查找到该视频数据流中存在所述目标区域的异常视频帧后,还可以对该视频数据流中针对多个异常视频帧进行运算,以得到异常视频帧对应的检测参数,比如,视频帧序列像素异常的像素异常率,或进一步对不同像素异常率构成的异常视频序列打分所得到的视频检测分值,从而在检测参数符合阈值的情况下将所述异常视频帧确定为目标视频帧。针对该阈值而言,还可以根据该视频检测分值配置该阈值,以根据该阈值更精确的定位出该视频数据流中该目标视频帧的具体位置。采用本实施方式,由于针对该关键区域进行检测,因此,不仅提高了处理速度,且检测准确性更高。

[0044] 一实施方式中,所述在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧,包括:对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;根据所述雷同分值,得到针对所述异常视频帧的像素异常率;将所述像素异常率作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第一视频序列中存在的异常视频帧确定为目标视频帧。采用本实施方式,可以通过针对第一视频序列进行像素异常的运算,得到雷同分值,雷同分值作为用于评估像素异常率的一个指标,可以根据该雷同分值,得到针对所述异常视频帧的像素异常率,从而,将像素异常率作为检测参数,在所述检测参数符合阈值的情况下,将所述第一视频序列中存在的异常视频帧确定为目标视频帧。换言之,可以根据像素异常率,从整个视频数据流中筛选出位于第一视频序列中的异常视频帧,并最终将该位于第一视频序列中的异常视频帧确定为目标视频帧,提高了检测准确性。

[0045] 一实施方式中,所述在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧,包括:对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;根据所述雷同分值,得到针对所述异常视频帧的像素异常率;从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值;将所述视频检测分值作为所述检测参数,在所

述检测参数符合阈值的情况下,将所述第二视频序列中存在的异常视频帧确定为目标视频帧。采用本实施方式,可以通过针对第一视频序列进行像素异常的运算,得到雷同分值,雷同分值作为用于评估像素异常率的一个指标,可以根据该雷同分值,得到针对所述异常视频帧的像素异常率。进一步的,还可以从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值,以将所述视频检测分值作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第二视频序列中存在的异常视频帧确定为目标视频帧。换言之,可以根据像素异常率,从整个视频数据流中筛选出位于第一视频序列中的异常视频帧,从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值后,进一步根据该视频检测分值,从第一视频序列中筛选出位于第二视频序列中视频检测分值符合预期的异常视频帧,并最终将该位于第二视频序列中视频检测分值符合预期的异常视频帧确定为目标视频帧,提高了检测准确性。

[0046] 一实施方式中,还包括:根据视频检测分值配置所述阈值。其中,所述视频检测分值包括:不同像素异常率构成的第二视频序列在第一视频序列中的分值占比;所述第一视频序列为所述视频数据流中由所述异常视频帧构成的第一视频序列。配置阈值的一示例中,对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值,根据该雷同分值,得到针对所述异常视频帧的像素异常率,从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值,从而,可以根据该针对所述异常视频帧运算得到的该视频检测分值配置该阈值。采用本实施方式,根据该视频检测分值进行评估的过程中,异常视频的检测准确性主要取决于经打分得到不同像素异常率的异常视频帧在整个视频数据流中的占比,而与异常视频帧的位置无关,从而提高了检测准确性。

[0047] 一实施方式中,还包括:根据所述阈值,定位出所述目标视频帧在所述视频数据流中的位置。针对该阈值而言,还可以根据上述视频检测分值配置阈值,以根据该阈值更精确的定位出该视频数据流中该目标视频帧的具体位置。采用本实施方式,由于该阈值是通过上述视频检测分值配置得到的,因此,该阈值本身就可以用于评估存在异常的目标视频帧,从而,可以根据阈值直接推理出该目标视频帧的位置,不需要涉及神经网络的训练。除此之外,为了更加精确及提高定位效率,还可以根据阈值简单训练得到用于定位的神经网络,根据该用于定位的神经网络该目标视频帧的位置从该视频数据流中定位出来,由于训练所需要的数据少,从而降低了神经网络的复杂度。

[0048] 应用示例:

[0049] 应用本公开实施例一处理流程包括如下内容:

[0050] 1、读取视频数据流,得到多个视频帧,每个视频帧为一帧图片,多个视频帧构成了图片流 $X(i=1,2,\dots,n)$, n 为大于1的正整数,表示图片的个数。

[0051] 2、检测关键区域,可以使用图片关键区域检测器等传感器检测每一帧图片中的感兴趣区域(如人脸区域、包括人物肢体的人体区域、包括用于识别伪造动作的动作区域等),将这些感兴趣的区域作为该关键区域,并在图片流 $X(i=1,2,\dots,n)$ 中予以标识。

[0052] 3、设定超参数采样间隔 d ,得到图片序列 $S(j=1,2,\dots,n/d)$, n 为大于1的正整数,表示图片的个数,图片序列 S 为多个图片序列的统称。

[0053] 4、以每个图片序列为例,将每个图片序列记为图片序列 S_j ,且至少包含前后两张图片,即 $S_1 = (X_1, X_3)$ 、 $S_2 = (X_3, X_5)$ 、 \dots 、 $S(n/d) = (X_{n-2}, X_n)$;对于采样间隔 d 的选取,以前后两张图片为例,可以选取 $d=2$ 。

[0054] 5、根据公式(1)计算针对图片序列的像素异常率:

$$[0055] \quad \text{Ratio} = \frac{\text{Compare}(X_i, X_{i+d})}{\text{单个图片像素数量}} \quad (1)$$

[0056] 其中,Ratio为像素异常率;Compare(X_i, X_{i+d})为雷同分值,可以通过该雷同分值对图片序列中的视频帧进行比较采样,以发现雷同帧数。

[0057] 需要指出的是,雷同分值这一指标,是用于评估像素异常率的其中一个指标,本公开不限于该指标,可以用于评估像素异常率的指标都在本申请的保护范围之内。

[0058] 6、得到以 $R(k=1, 2, \dots, n/d)$ 表示的多个像素异常率序列。其中, n 为大于1的正整数,表示图片的个数, d 为采样间隔。

[0059] 7、针对多个像素异常率序列的评估,可以采用多种方式得出视频检测分值,本应用示例根据如下根据公式(2)或公式(3)来计算视频检测分值,视频检测分值 $\text{score} \in (0, 1)$;其中, score 为0,代表正常视频的标签,在后续基于视频检测分值配置阈值的训练过程中,若某个图片加上该类标签,则说明该视频帧未被编辑过(或称未被伪造过),是“真”的原始视频帧; score 为1,代表异常视频的标签,在后续基于视频检测分值配置阈值的训练过程中,若某个图片加上该标签,则说明该视频帧被编辑过(或称被伪造过),是“假”的伪造视频,即基于图片驱动视频所得到的标签,这类标签对用户具备攻击的可能性。

$$[0060] \quad \text{score} = \text{策略}(R) \quad (2)$$

$$[0061] \quad \text{score} = \frac{1}{1 + e^{-(w^T R + b)}} \quad (3)$$

[0062] 其中,公式(2)中, R 为多个像素异常率序列;策略可以根据不同应用场景的检测需求针对性的配置; score 为视频检测分值。还可以根据 score 配置阈值,比如,将该 score 作为阈值,从而,可以基于策略直接定位出异常视频的所在位置。

[0063] 其中,公式(3)中, R 为多个像素异常率序列; b 为预设的参数;该参数可以根据不同应用场景的检测需求针对性的配置; w 为权重; score 为视频检测分值,可以根据多个像素异常率序列的分布,通过公式(3)的Logistic回归原理(在Logistic回归过程中,可以采用制作训练集和验证集的方法去训练,以得到用于实际应用场景推理的 w 参数)设计一个特征提取器,以便对视频更为精确的打分,而不是如公式(2)只通过预设的策略去打分。需要指出的是,采用Logistic回归可以等价于对视频数据流执行手工设计的卷积核及再做卷积的复杂操作。还可以根据 score 配置阈值,比如,将该 score 作为阈值,从而,基于通过训练得到的神经网络的方式去定位出异常视频的所在位置。

[0064] 采用本应用示例,考虑到视频帧若被伪造过、被编辑过,则存在“不同帧画面存在部分像素完全相同的情况”,因此,可以通过合理的选择待检测区域和对整个视频进行梳理,可以找出这种编辑痕迹,并对视频帧的“真”或“假”做判断,从而快速定位出整个视频数据流中的异常视频位置。计算方便,不需要视觉处理器(GPU, Graphics Processing Unit);计算复杂度低且快速;准确度高,可解释性强,可以直接定位异常视频位置;通过设定阈值可以直接推理,不需要训练,或者训练需要的数据少。而且,通过上述公式计算视频任意序

列中被编辑区域大小,并作为视频级别的判断依据,也不需要设计复杂的卷积神经网络,通过基于策略或简单训练得到的score去配置阈值,从而,用最少的成本,最快的运算即可达到很好的检测效果。

[0065] 根据本公开的实施例,提供了一种视频检测装置,图3是根据本公开实施例的视频检测装置的组成结构示意图,如图3所示,视频检测装置300包括:目标区域检测模块301,用于对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;异常视频查找模块302,用于查找所述视频数据流中存在所述目标区域的异常视频帧;目标视频确定模块303,用于在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0066] 根据本公开的实施例,提供了一种视频检测装置,图4是根据本公开实施例的视频检测装置的组成结构示意图,如图4所示,视频检测装置400包括:目标区域检测模块401,用于对视频数据流中的视频帧进行检测,得到所述视频帧中的目标区域,所述目标区域用于表征所述视频数据流中不同视频帧存在部分像素完全相同的区域;其中,目标区域检测模块401还包括:关键区域提取子模块4011,用于对所述视频数据流中相邻的至少两个视频帧,分别提取关键区域;比对子模块4012,用于对所述至少两个视频帧分别对应的关键区域进行像素点的特征比对,将比对得到的所述关键区域中存在所述部分像素完全相同的区域,作为所述目标区域。以及异常视频查找模块402,用于查找所述视频数据流中存在所述目标区域的异常视频帧;目标视频确定模块403,用于在所述异常视频帧对应的检测参数符合阈值的情况下,将所述异常视频帧确定为目标视频帧。

[0067] 一实施方式中,所述目标视频确定模块,用于对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;根据所述雷同分值,得到针对所述异常视频帧的像素异常率;将所述像素异常率作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第一视频序列中存在的异常视频帧确定为目标视频帧。

[0068] 一实施方式中,所述目标视频确定模块,用于对所述视频数据流中由所述异常视频帧构成的第一视频序列进行像素异常的运算,得到雷同分值;根据所述雷同分值,得到针对所述异常视频帧的像素异常率;从所述第一视频序列中选取由不同像素异常率构成的第二视频序列,并分别进行打分,得到对应的视频检测分值;将所述视频检测分值作为所述检测参数,在所述检测参数符合阈值的情况下,将所述第二视频序列中存在的异常视频帧确定为目标视频帧。

[0069] 一实施方式中,还包括阈值配置模块,用于根据视频检测分值配置所述阈值;其中,所述视频检测分值包括:不同像素异常率构成的第二视频序列在第一视频序列中的分值占比;所述第一视频序列为所述视频数据流中由所述异常视频帧构成的第一视频序列。

[0070] 一实施方式中,还包括定位模块,用于根据所述阈值,定位出所述目标视频帧在所述视频数据流中的位置。

[0071] 本公开实施例各装置中的各模块的功能可以参见上述方法中的对应描述,在此不再赘述。

[0072] 根据本公开的实施例,本公开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。

[0073] 图5是用来实现本公开实施例的视频检测方法的电子设备的框图。该电子设备可以为前述部署设备或代理设备。电子设备旨在表示各种形式的数字计算机,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/或要求的本公开的实现。

[0074] 如图5所示,电子设备500包括计算单元501,其可以根据存储在只读存储器(ROM) 502中的计算机程序或者从存储单元508加载到随机访问存储器(RAM) 503中的计算机程序来执行各种适当的动作和处理。在RAM 503中,还可存储电子设备500操作所需的各种程序和数据。计算单元501、ROM502以及RAM 503通过总线504彼此相连。输入输出(I/O)接口505也连接至总线504。

[0075] 电子设备500中的多个部件连接至I/O接口505,包括:输入单元506,例如键盘、鼠标等;输出单元507,例如各种类型的显示器、扬声器等;存储单元508,例如磁盘、光盘等;以及通信单元509,例如网卡、调制解调器、无线通信收发机等。通信单元509允许电子设备500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。

[0076] 计算单元501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501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501执行上文所描述的各个方法和处理,例如视频检测方法。例如,在一些实施例中,视频检测方法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508。在一些实施例中,计算机程序的部分或者全部可以经由ROM 502和/或通信单元509而被载入和/或安装到电子设备500上。当计算机程序加载到RAM 503并由计算单元501执行时,可以执行上文描述的视频检测方法的一个或多个步骤。备选地,在其他实施例中,计算单元501可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行视频检测方法。

[0077] 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)、负载可编程逻辑设备(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。

[0078] 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单元或控制器,使得程序代码当由处理单元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。

[0079] 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。

[0080] 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。

[0081] 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。

[0082] 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。

[0083] 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执行也可以不同的次序执行,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。

[0084] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

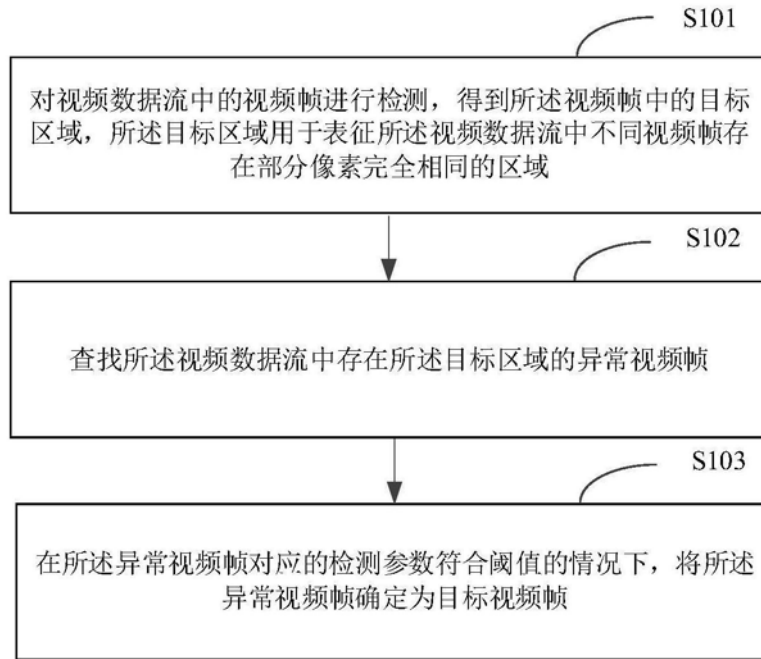


图1

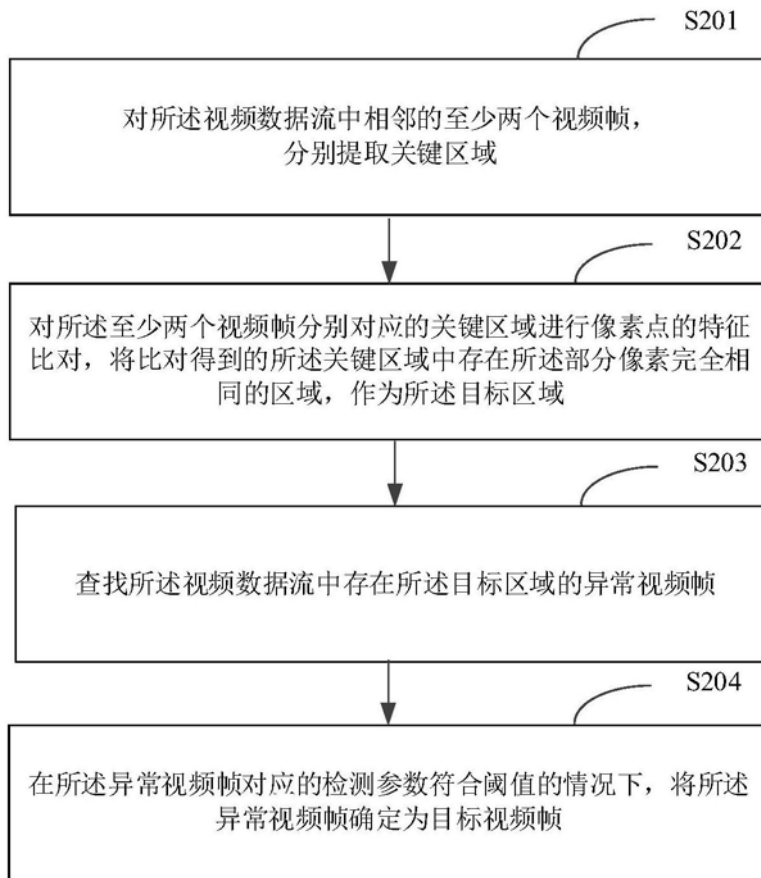


图2

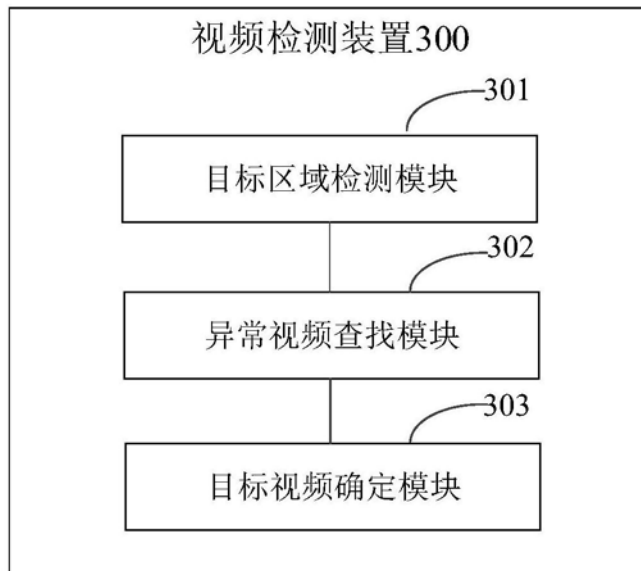


图3

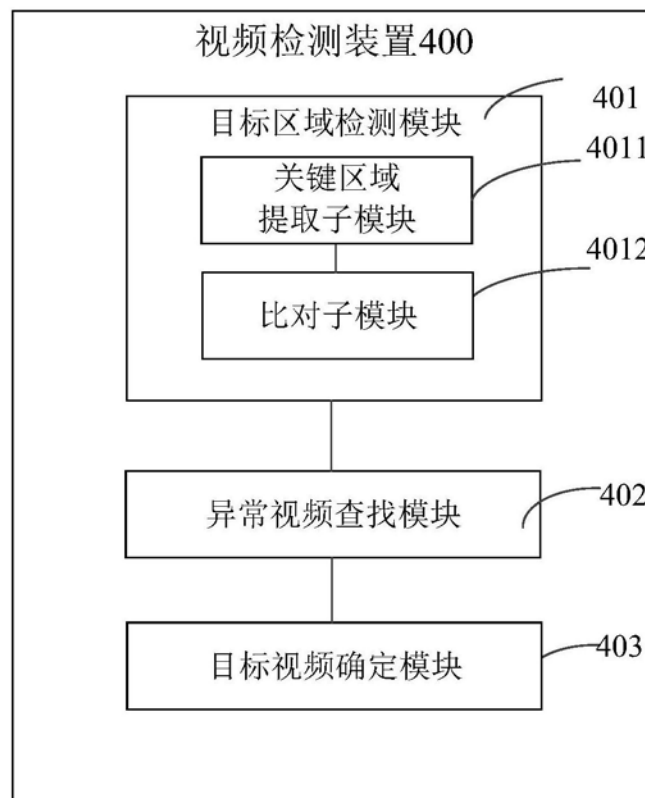


图4

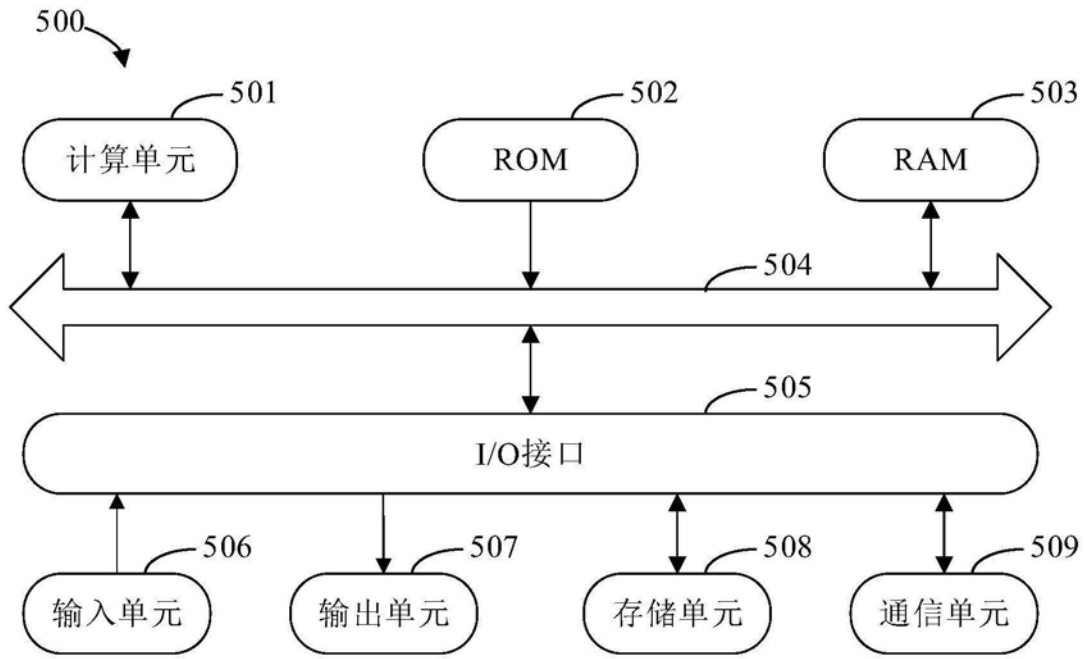


图5