

视频匹配方法、基于区块链的侵权存证方法和装置

技术领域

[01] 本文件涉及计算机技术领域，尤其涉及一种视频匹配方法、基于区块链的侵权存证方法和装置。

背景技术

[02] 目前，在对某个可疑的侵权视频进行侵权定位时，通常是首先从该视频中提取多种特征，再经过搜索引擎检索后，得到与该视频相匹配的多个候选视频的匹配结果。而对该可疑的侵权视频进行最终的侵权定位，还需要分别计算多个候选视频与该可疑的侵权视频的相似度。这需要一种高鲁棒性的算法来应对视频与视频之间的多种特征的误匹配和漏匹配，以及支持多个视频片段的侵权定位。

[03] 然而，现有的视频匹配方法容易受到特征检索结果的噪声影响，且随着视频时长的增加，视频与视频之间匹配的效率也急剧降低。此外，针对侵权视频以及其侵权证据的存证，对于业界来说，也是亟需解决的问题。

发明内容

[04] 本说明书实施例提供了一种视频匹配方法、基于区块链的侵权存证方法和装置，以应对视频与视频之间的多种特征的误匹配和漏匹配，以及支持多个视频片段的侵权定位，提高视频匹配的效率，从而降低人工审核的成本。

[05] 为解决上述技术问题，本说明书实施例是这样实现的：第一方面，提出了一种视频匹配方法，包括：基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[06] 其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[07] 第二方面，一种基于区块链的侵权视频存证方法，包括：获取目标视频的多个特征向量；基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述

深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[08]第三方面，提出了一种基于区块链的侵权存证装置，包括：候选视频检索模块，基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；特征图构建模块，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；模型输出模块，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；证据上传模块，当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[09]其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[10]第四方面，提出了一种视频匹配装置，包括：特征向量获取模块，获取目标视频的多个特征向量；候选视频检索模块，基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；特征图构建模块，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；模型输出模块，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[11]第五方面，提出了一种电子设备，包括：处理器；以及被安排成存储计算机可执行指令的存储器，所述可执行指令在被执行时使所述处理器执行以下操作：基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[12]其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[13]第六方面，提出了一种计算机可读存储介质，所述计算机可读存储介质存储一个或

多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行以下操作：基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[14]其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[15]第七方面，提出了一种电子设备，包括：处理器；以及被安排成存储计算机可执行指令的存储器，所述可执行指令在被执行时使所述处理器执行以下操作：获取目标视频的多个特征向量；基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[16]第八方面，提出了一种计算机可读存储介质，所述计算机可读存储介质存储一个或多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行以下操作：获取目标视频的多个特征向量；基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[17]本说明书实施例采用上述技术方案至少可以达到下述技术效果：在对目标视频进行侵权定位时，能够基于该目标视频的多个特征向量，从视频数据库中检索与目标视频相似的候选视频，再基于目标视频和候选视频，构建目标视频和候选视频之间的时域相似度矩阵特征图，最后，将时域相似度矩阵特征图作为深度学习检测模型的输入，输出得到候选视频中与目标视频相匹配的视频片段和对应的相似度；并在候选视频中与目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有目标视频的摘要、候选视频中与目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块

链中。本说明书实施例提供的方法利用深度学习检测模型，一方面在侵权定位的效率上，能够检测可能的侵权视频的任意多个侵权片段，同时利用向量检索结合检测模型能够极大地提高侵权视频的检测效率；另一方面，也降低了人工审核的成本。此外，还利用区块链不可篡改的特性，将存在侵权的目标视频的摘要、候选视频中与目标视频相匹配的视频片段和对应的相似度上传至区块链中，以备侵权指证时从区块链中获取目标视频侵权的证据。

附图说明

[18] 此处所说明的附图用来提供对本说明书的进一步理解，构成本说明书的一部分，本说明书的示意性实施例及其说明用于解释本说明书，并不构成对本说明书的不当限定。在附图中：

[19] 图1为本说明书一个实施例提供的基于区块链的侵权存证方法的实现流程示意图。

[20] 图2为本说明书一个实施例提供的一种视频匹配方法的实现流程示意图。

[21] 图3为本说明书一个实施例提供的视频匹配方法的应用于一种场景中的流程示意图。

[22] 图4为本说明书一个实施例提供的视频匹配方法中绘制的时域相似度矩阵特征图的示意图。

[23] 图5为本说明书一个实施例提供的基于区块链的侵权存证装置的结构示意图。

[24] 图6为本说明书一个实施例提供的视频匹配装置的结构示意图。

[25] 图7为本说明书一个实施例提供的一种电子设备的结构示意图。

[26] 图8为本说明书一个实施例提供的另一种电子设备的结构示意图。

具体实施方式

[27] 为使本文件的目的、技术方案和优点更加清楚，下面将结合本说明书具体实施例及相应的附图对本说明书技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本文件一部分实施例，而不是全部的实施例。基于本说明书中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本文件保护的范围。

[28] 以下结合附图，详细说明本说明书各实施例提供的技术方案。

[29] 针对侵权视频以及其侵权证据的存证，对于业界来说，也是亟需解决的问题。本说明书实施例通过引入区块链，利用区块链的不可篡改的特性，将目标视频的相关信息、候选视频相关信息及目标视频是否侵权的信息写入到区块链中，从而保证区块链中的侵权信息的可信性，以用于对目标视频是否侵权进行快速取证。具体地，图1是本说明书一个实施例提供的一种基于区块链的侵权视频存证方法的实现流程示意图，包括：

[30] S110，基于目标视频的多个特征向量，从视频数据库中检索与目标视频相似的候选视频。

[31] S120，基于目标视频和所述候选视频，构建目标视频和候选视频之间的时域相似度

矩阵特征图。

[32]S130, 将时域相似度矩阵特征图作为深度学习检测模型的输入, 以输出候选视频中与目标视频相匹配的视频片段和对应的相似度。

[33]其中, 深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的, 其中样本视频包含查询视频和所述查询视频对应的候选视频, 样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[34]S140, 当候选视频中与目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时, 将包含有目标视频的摘要、候选视频中与目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[35]其中, 预设相似度阈值可根据经验值得到, 用于界定目标视频是否存在侵权, 比如可以设置为 60%。应理解, 由于区块链中的存储空间有限, 本说明书实施例对目标视频的侵权证据进行存证时, 可将目标视频通过哈希加密算法转化为一串哈希值, 将该目标视频的哈希值、以及候选视频中与目标视频相匹配的视频片段和对应的相似度上传至区块链中, 由区块链中具备存证权限的节点对该侵权证据进行共识操作, 并在共识之后记录到新生成的区块中。当需要获取该侵权证据时, 可基于目标视频的哈希值从区块链中下载包含有目标视频的哈希值的侵权证据。

[36]图 1 所示实施例相关步骤的具体实现可参考下文所述的图 2 所示实施例中对应的步骤的具体实现, 本说明书一个或多个实施例在此不再赘述。

[37]在对目标视频进行侵权定位时, 能够基于该目标视频的多个特征向量, 从视频数据库中检索与目标视频相似的候选视频, 再基于目标视频和候选视频, 构建目标视频和候选视频之间的时域相似度矩阵特征图, 最后, 将时域相似度矩阵特征图作为深度学习检测模型的输入, 输出得到候选视频中与目标视频相匹配的视频片段和对应的相似度; 并在候选视频中与目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时, 将包含有目标视频的摘要、候选视频中与目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中, 其中, 深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的, 其中样本视频包含查询视频和查询视频对应的候选视频, 样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。本说明书实施例提供的方法利用区块链不可篡改的特性, 将存在侵权的目标视频的摘要、候选视频中与目标视频相匹配的视频片段和对应的相似度上传至区块链中, 以备侵权指证时从区块链中获取目标视频侵权的证据。

[38]如背景技术中所述, 在对侵权视频进行侵权定位时, 需要从侵权视频中提取的多种特征向量, 经过指定搜索引擎检索后, 得到的向量检索结果中会包含 N 个候选视频的匹配结果。这些结果, 需要分别与侵权视频计算相似度并进行侵权定位。这里就需要高鲁棒性的算法来应对特征向量的误匹配和漏匹配, 同时检索结果中如果包含的搜索引擎排序的视频集合大, 则要求高效率。

[39]此外, 该视频匹配算法还要支持多个视频片段的侵权定位, 来降低人工审核的成本。

然而业界常用的动态规划算法,CCF 竞赛等方案都容易受到特征向量检索结果的噪声影响,不够鲁棒,而且随着侵权视频时长的增加,视频匹配的效率也急剧降低。

[40]本说明书实施例为了解决现有的侵权视频的检测效率和准确率较低的问题,还提出一种视频匹配方法,能够获取目标视频的多个特征向量,并基于该目标视频的多个特征向量,从视频数据库中检索与目标视频相似的候选视频,再基于目标视频和候选视频,构建目标视频和候选视频之间的时域相似度矩阵特征图,最后,将时域相似度矩阵特征图作为深度学习检测模型的输入,输出得到候选视频中与目标视频相匹配的视频片段和对应的相似度;其中,深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的,样本视频包含查询视频和查询视频对应的候选视频,样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[41]本说明书实施例提供的方法利用深度学习检测模型,一方面在侵权定位的效率上,能够检测可能的侵权视频的任意多个侵权片段,同时利用向量检索结合检测模型能够极大地提高侵权视频的检测效率;另一方面,也降低了人工审核的成本。

[42]本说明书实施例提供的视频方法,该方法的执行主体,可以但不限于个人电脑、服务器等能够被配置为执行本发明实施例提供的该方法的装置中的至少一种。

[43]为便于描述,下文以该方法的执行主体为能够执行该方法的服务器为例,对该方法的实施方式进行介绍。可以理解,该方法的执行主体为服务器只是一种示例性的说明,并不应理解为对该方法的限定。

[44]具体地,本说明书一个或多个实施例提供的一种视频匹配方法的实现流程示意图如图2所示,包括:S210,获取目标视频的多个特征向量。

[45]其中,目标视频具体可以是可疑的侵权视频,后续所述的候选视频即可以作为该可疑的侵权视频侵权的证据。

[46]可选地,获取目标视频的多个特征向量具体可以将目标视频拆分为多个视频片段,再针对每个视频片段抽取一种或多种特征向量。或者,还可以对目标视频进行抽帧得到多个视频帧,可抽取目标视频中的关键帧也可随机抽取目标视频中的多个视频帧,还可以每隔预设时间段抽取目标视频中的一个视频帧得到多个视频帧,再对抽取的视频帧抽取一种或多种特征向量。其中一种特征向量对应于一种特征提取算法。

[47]因此,目标视频的多个特征向量具体可以包括目标视频的多个视频片段或视频帧对应的多个特征向量,一个视频片段或视频帧对应一个特征向量;或者,目标视频的多个特征向量还可以包括:通过多种特征提取算法提取得到的目标视频的多种特征向量;或者,目标视频的多个特征向量还可以包括:通过多种特征提取算法分别对目标视频的多个视频片段或视频帧提取得到的多个特征向量,一个视频片段或视频帧对应多种特征向量。

[48]S220,基于目标视频的多个特征向量,从视频数据库中检索与目标视频相似的候选视频。

[49]其中,视频数据库中包含了海量的视频,每个视频对应一种或多种特征向量,一种

特征向量对应于一种特征提取算法。

[50]可选地，可分别基于目标视频的多个特征向量中的各特征向量，从视频数据库中检索与目标视频的各特征向量相匹配的特征向量，再确定这些相匹配的特征向量对应的视频即为候选视频。具体地，基于目标视频的多个特征向量，从视频数据库中检索与目标视频相似的候选视频，包括：从视频数据库中，获取与目标视频的多个特征向量相似的特征向量检索结果；基于与目标视频的多个特征向量相似的特征向量检索结果，从视频数据库中获取与目标视频相似的候选视频。

[51]其中，与目标视频的多个特征向量相似的特征向量检索结果具体可包括：与各特征向量相匹配的前几个特征向量，或者与各特征向量最匹配的一个特征向量。比如，可以从视频数据库中，分别获取与目标视频的多个特征向量中的各特征向量相匹配的前 k 个特征向量，再确定这前 k 个特征向量对应的 m 个候选视频，其中 m 小于或等于 k ，且 m 大于或等于 1，当 $m=k$ 时，则表明这 k 个特征向量来自 k 个不同的候选视频，当 $m=1$ 时，则表明这 k 个特征向量来自同一个候选视频，或者也可以从视频数据库中，分别获取与目标视频的多个特征向量中的各特征向量最匹配的一个特征向量，再确定该最匹配的一个特征向量对应的候选视频。也就是说，一个特征向量的检索结果，可能对应于一个候选视频的多个匹配特征向量，也可能是多个候选视频的不同匹配特征向量。

[52]如图 3 所示，为本说明书实施例提供的视频匹配方法应用于一种实际场景中的示意图。在图 3 中， $q_1 \sim q_n$ 为目标视频的多个特征向量， V_3 和 V_1 为从视频数据库中检索到的与目标视频相似的两个候选视频的向量检索结果。图左侧中， V_{3, q_1} 为候选视频 V_3 中与目标视频的特征向量 q_1 匹配位置的相似度值， V_{3, q_2} 为候选视频 V_3 中与目标视频的特征向量 q_2 匹配位置的相似度值， V_{3, q_n} 为候选视频 V_3 中与目标视频的特征向量 q_n 匹配位置的相似度值；图右侧中， V_{1, q_1} 为候选视频 V_1 中与目标视频的特征向量 q_1 匹配位置的相似度值， V_{1, q_n} 为候选视频 V_1 中与目标视频的特征向量 q_n 匹配位置的相似度值。

[53]S230，基于目标视频和候选视频，构建目标视频和所述候选视频之间的时域相似度矩阵特征图。

[54]应理解，由于上述目标视频的多个特征向量、与候选视频中的多个特征向量之间的向量检索结果中会包含有图 3 所示的目标视频的特征向量与候选视频的特征向量之间相匹配的位置（也就是相似的位置）和对应位置的相似度，为了便于深度学习检测模型能够准确地学习到目标视频与候选视频之间相匹配的视频片段和对应的相似度，本说明书实施例可基于目标视频和候选视频之间的向量检索结果构建时域相似度矩阵特征图。具体地，基于目标视频和候选视频，构建目标视频和候选视频之间的时域相似度矩阵特征图，包括：基于目标视频的多个特征向量、与候选视频中的多个特征向量之间的向量检索结果，构建目标视频的多个特征向量与候选视频的多个特征向量之间的相似度矩阵；基于目标视频的多个特征向量与候选视频中的多个特征向量之间的相似度矩阵，在时域维度上，构建目标视频和候选视频之间的时域相似度矩阵特征图。

[55]当候选视频的数量为一个时，可基于目标视频的与候选视频中的多个特征向量之间的相似度矩阵，在同一时域维度上，将目标视频的多个特征向量与候选视频中的多个特

征向量之间的相似度矩阵的分布绘制在二维特征图中。如图4所示,为本说明书实施例提供的视频匹配方法中绘制得到的目标视频和候选视频之间的时域相似度矩阵特征图。在图4中,横坐标为目标视频的时域轴,纵坐标为候选视频的时域值,三角形形状的图样对应于目标视频和候选视频的一种特征向量,方块形状的对应用于目标视频和候选视频的另一种特征向量,各图样的取值为向量检索结果中的相似度分数。在实际应用中,为提高视频匹配的效率,可将这不同的特征向量绘制在同一时域相似度矩阵特征图中。

[56]或者,还可将不同的特征向量绘制在不同的时域相似度矩阵特征图中,即如图3所示的下半部分的左侧,可将每一种特征向量绘制得到的时域相似度矩阵特征图作为深度学习检测模型的一个通道输入,那么目标向量的多个特征向量存在多种特征向量时,则会存在多个时域相似度矩阵特征图作为深度学习检测模型的多个通道的输入。

[57]可选地,为便于深度学习检测模型准确高效地确定目标视频与候选视频之间的相似视频片段和对应的相似度,本说明书实施例可根据目标视频和候选视频之间的时域对应关系,构建目标视频和候选视频之间的时域相似度矩阵特征图。具体地,基于目标视频的多个特征向量与候选视频中的多个特征向量之间的相似度矩阵,在时域维度上,构建目标视频和候选视频之间的时域相似度矩阵特征图,包括:根据目标视频和候选视频之间的时域对应关系,将目标视频的多个特征向量与候选视频中的多个特征向量之间的相似度矩阵,绘制在二维特征图上,得到目标视频和候选视频之间的时域相似度矩阵特征图。

[58]可选地,当候选视频有多个时,根据目标视频和候选视频之间的时域对应关系,将目标视频的多个特征向量与候选视频中的各候选视频的多个特征向量之间的相似度矩阵,绘制在二维特征图上,得到目标视频和候选视频之间的时域相似度矩阵特征图,包括:根据目标视频和多个候选视频之间的时域对应关系,分别将目标视频的多个特征向量与多个候选视频中的各候选视频的多个特征向量之间的相似度矩阵,绘制在多个二维特征图上,得到目标视频和多个候选视频之间的多个时域相似度矩阵特征图;将目标视频和多个候选视频之间的多个时域相似度矩阵特征图进行拼接,得到目标视频和多个候选视频之间的时域相似度矩阵特征图。

[59]当候选视频有多个时,则可分别将目标视频的多个特征向量与多个候选视频中的各候选视频的多个特征向量之间的相似度矩阵,绘制在多个二维特征图上,得到目标视频和多个候选视频之间的多个时域相似度矩阵特征图。为了提高深度学习检测模型的学习效率,可将这多个时域相似度矩阵特征图进行拼接,得到一个时域相似度矩阵特征图。比如,当候选视频有四个时,则会得到目标视频和四个候选视频之间的四个时域相似度矩阵特征图,再将这四个时域相似度矩阵特征图拼接为 2×2 的时域相似度矩阵特征图作为深度学习检测模型的输入。

[60]S240,将时域相似度矩阵特征图作为深度学习检测模型的输入,以输出候选视频中与目标视频相匹配的视频片段和对应的相似度。

[61]其中,深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的,其中样本视频包含查询视频和查询视频对应的候选视频,样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志,

侵权标志用于标识该相匹配的视频片段是否存在侵权。应理解，查询视频对应的候选视频可以是一个也可以是多个。当查询视频对应的候选视频为多个时，样本视频包含查询视频和查询视频对应的多个候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的各个候选视频中所匹配的视频片段及侵权标志。具体地，样本视频对应的标签包括视频片段和对应的 label（标签），label 为侵权或非侵权。

[62] 应理解，在对深度学习检测模型进行训练时，用于训练深度学习检测模型的样本视频对应的标签通常为离散化的标签，即“是”或“否”，对应于本说明书实施例中的标签即为“侵权”或“非侵权”，而在深度学习检测模型进行预测时，输出的则是检测框位置 $[x1, y1, x2, y2]$ ，其中 $[x1, x2]$ 对应于目标视频中的匹配时间片段， $[y1, y2]$ 对应于候选视频中的匹配时间片段，以及 $[x1, y1, x2, y2]$ 对应的置信度，用于表征匹配时间片段的相似度。

[63] 其中，每组样本视频的时域相似度矩阵特征图为该组查询视频和其对应的候选视频之间的时域相似度矩阵特征图，该时域相似度矩阵特征图的获取方式与上文所述的时域相似度矩阵特征图的获取方式一致，本文不再赘述。

[64] 可选地，本说明书实施例中的深度学习检测模型包括但不限于下述模型：更快速的基于卷积神经网络的候选区域检测模型 Faster-Rcnn；带掩码的基于卷积神经网络的候选区域检测模型 Mask-Rcnn；实时物体检测模型 YOLO；单次多框检测模型 SSD。

[65] 其中，深度学习检测模型 Faster-Rcnn 的训练过程为：输入测试图像；将整张图片输入卷积神经网络，进行特征提取；用 RPN 生成一堆锚框，对其进行裁剪过滤后通过 softmax 判断 anchors 属于前景 (foreground) 或者后景 (background)，即是物体或者不是物体，即是一个二分类过程；同时，另一分支边框回归修正锚框，形成较精确的 proposal（注：这里的较精确是相对于后面全连接层的再一次边框回归而言）；把建议窗口映射卷积神经网络的最后一层卷积特征图上；通过 RoI pooling 层使每个 RoI 生成固定尺寸的特征图；利用 Softmax Loss (探测分类概率) 和 Smooth L1 Loss (探测边框回归) 对分类概率和边框回归联合训练。

[66] 深度学习检测模型 Mask-Rcnn 以 Faster RCNN 原型，增加了一个分支用于分割任务，即对于 Faster RCNN 的每个 Proposal Box 都要使用 FCN（全卷积神经网络，将传统卷积神经网络中的全连接层转化成一个个的卷积层）进行语义分割，分割任务与定位、分类任务是同时进行的。

[67] 深度学习检测模型 YOLO (英文全称: You Only Look Once)，是一种对象检测模型，YOLO 有基于 CNN 和锚框的简洁架构，并且是一种针对普遍使用问题的实时对象检测技术。YOLO 将图像分成 13×13 个单元格：每个单元格负责预测 5 个边界框。边界框描述包围对象的矩形。YOLO 还输出一个置信度（即本说明书实施例中的相似度），用于指示预测的边界框实际包含某个对象的程度。先前的检测系统使用分类器或定位器进行检测，将模型应用于图像的多个位置和尺度，图像的高得分区域作为检测目标。YOLO 则采取了完全不同的方法。它将单个神经网络应用于整个图像，该网络将图像分成区域，预测每个区域的边界框和概率，并依据概率大小对所有边框分配权重。

[68] 与 Faster-Rcnn 相比，SSD 没有生成 proposal 的过程，这就极大提高了检测速度。针对不同大小的目标检测，传统的做法是先将图像转换成不同大小（图像金字塔），然后

分别检测，最后将结果综合起来（NMS）。而 SSD 算法则利用不同卷积层的特征图进行综合也能达到同样的效果。算法的主网络结构是 VGG16，将最后两个全连接层改成卷积层，并随后增加了 4 个卷积层来构造网络结构。对其中 5 种不同的卷积层的输出（特征图）分别用两个不同的 3×3 的卷积核进行卷积，一个输出分类用的 confidence，每个 default box 生成 21 个类别 confidence；一个输出回归用的 localization，每个 default box 生成 4 个坐标值 (x, y, w, h) 。

[69] 可选地，将时域相似度矩阵特征图作为深度学习检测模型的输入，以输出候选视频中与目标视频相匹配的视频片段和对应的相似度，包括：将时域相似度矩阵特征图作为深度学习检测模型的输入，以输出候选视频中与目标视频相匹配的视频片段在时域维度上的区间范围、以及相匹配的视频片段之间的相似度。

[70] 具体地，深度学习检测模型输出各个时域相似度矩阵特征图上的检测框位置以及置信度，达到对目标视频的侵权定位的目的。其中，候选视频中与目标视频相匹配的视频片段在时域维度上的区间范围，具体可以是检测框位置： $[x1, y1, x2, y2]$ ，其中 $[x1, x2]$ 为目标视频中的时间片段， $[y1, y2]$ 为候选视频中的时间片段，相匹配的视频片段之间的相似度具体可以用置信度来表征。如图 3 下半部分的右侧所示，将时域相似度矩阵特征图作为深度学习检测模型的输入后，便可输出候选视频中与目标视频相匹配的检测框、以及相匹配的检测框之间的相似度。

[71] 在对目标视频进行侵权定位时，能够获取目标视频的多个特征向量，并基于该目标视频的多个特征向量，从视频数据库中检索与目标视频相似的候选视频，再基于目标视频和候选视频，构建目标视频和候选视频之间的时域相似度矩阵特征图，最后，将时域相似度矩阵特征图作为深度学习检测模型的输入，输出得到候选视频中与目标视频相匹配的视频片段和对应的相似度；其中，深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，样本视频包含查询视频和查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。本说明书实施例提供的方法利用深度学习检测模型，一方面在侵权定位的效率上，能够检测可能的侵权视频的任意多个侵权片段，同时利用向量检索结合检测模型能够极大地提高侵权视频的检测效率；另一方面，也降低了人工审核的成本。

[72] 图 5 是本说明书一个或多个实施例提供的一种基于区块链的侵权存证装置 500 的结构示意图，包括：候选视频检索模块 510，基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；特征图构建模块 520，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；模型输出模块 530，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；证据上传模块 540，当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[73] 其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本

视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[74]视频匹配装置 500 能够实现图 1 和图 2 的方法实施例的方法，具体可参考图 1 和图 2 所示实施例的基于区块链的侵权存证方法和视频匹配方法，不再赘述。

[75]图 6 是本说明书一个或多个实施例提供的一种视频匹配装置 600 的结构示意图，包括：特征向量获取模块 610，获取目标视频的多个特征向量；候选视频检索模块 620，基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；特征图构建模块 630，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；模型输出模块 640，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[76]可选地，在一种实施方式，所述候选视频检索模块 620，用于：从所述视频数据库中，获取与所述目标视频的多个特征向量相似的特征向量检索结果；基于与所述目标视频的多个特征向量相似的特征向量检索结果，从视频数据库中获取与所述目标视频相似的候选视频。

[77]可选地，在一种实施方式，所述特征图构建模块 630，用于：基于所述目标视频的多个特征向量、与所述候选视频中的各候选视频的多个特征向量之间的向量检索结果，构建所述目标视频的多个特征向量与所述候选视频的多个特征向量之间的相似度矩阵；基于所述目标视频的多个特征向量与所述候选视频中的多个特征向量之间的相似度矩阵，在时域维度上，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图。

[78]可选地，在一种实施方式，所述特征图构建模块 630，用于：根据所述目标视频和所述候选视频之间的时域对应关系，将所述目标视频的多个特征向量与所述候选视频的多个特征向量之间的相似度矩阵，绘制在二维特征图上，得到所述目标视频和所述候选视频之间的时域相似度矩阵特征图。

[79]可选地，在一种实施方式，当所述候选视频有多个时，所述特征图构建模块 630，用于：根据所述目标视频和所述多个候选视频之间的时域对应关系，分别将所述目标视频的多个特征向量与所述多个候选视频中的各候选视频的多个特征向量之间的相似度矩阵，绘制在多个二维特征图上，得到所述目标视频和所述多个候选视频之间的多个时域相似度矩阵特征图；将所述目标视频和所述多个候选视频之间的多个时域相似度矩阵特征图进行拼接，得到所述目标视频和所述多个候选视频之间的时域相似度矩阵特征图。

[80]可选地，在一种实施方式，所述模型输出模块 640，用于：将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段在时域维度上的区间范围、以及所述相匹配的视频片段之间的相似度。

[81]可选地，在一种实施方式，所述深度学习检测模型包括下述至少一种：更快速的基

于卷积神经网络的候选区域检测模型 Faster-Rcnn；带掩码的基于卷积神经网络的候选区域检测模型 Mask-Rcnn；实时物体检测模型 YOLO；单次多框检测模型 SSD。

[82] 视频匹配装置 600 能够实现图 2~图 4 的方法实施例的方法，具体可参考图 2~图 4 所示实施例的视频匹配方法，不再赘述。

[83] 图 7 是本说明书的一个实施例提供的电子设备的结构示意图。请参考图 7，在硬件层面，该电子设备包括处理器，可选地还包括内部总线、网络接口、存储器。其中，存储器可能包含内存，例如高速随机存取存储器(Random-Access Memory, RAM)，也可能还包括非易失性存储器(non-volatile memory)，例如至少 1 个磁盘存储器等。当然，该电子设备还可能包括其他业务所需要的硬件。

[84] 处理器、网络接口和存储器可以通过内部总线相互连接，该内部总线可以是 ISA(Industry Standard Architecture, 工业标准体系结构)总线、PCI(Peripheral Component Interconnect, 外设部件互连标准)总线或 EISA(Extended Industry Standard Architecture, 扩展工业标准结构)总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 7 中仅用一个双向箭头表示，但并不表示仅有一根总线或一种类型的总线。

[85] 存储器，用于存放程序。具体地，程序可以包括程序代码，所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器，并向处理器提供指令和数据。

[86] 处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成基于区块链的侵权存证装置。处理器，执行存储器所存放的程序，并具体用于执行以下操作：基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[87] 其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[88] 上述如本说明书图 1 所示实施例揭示的基于区块链的侵权存证方法可以应用于处理器中，或者由处理器实现。处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器(Central Processing Unit, CPU)、网络处理器(Network Processor, NP)等；还可以是数字信号处理器(Digital Signal Processor, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本说明书一个或多个实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可

以是任何常规的处理器等。结合本说明书一个或多个实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

[89]该电子设备还可执行图 1 的基于区块链的侵权存证方法，本说明书在此不再赘述。

[90]当然，除了软件实现方式之外，本说明书的电子设备并不排除其他实现方式，比如逻辑器件抑或软硬件结合的方式等等，也就是说以下处理流程的执行主体并不限于各个逻辑单元，也可以是硬件或逻辑器件。

[91]本说明书实施例还提出了一种计算机可读存储介质，该计算机可读存储介质存储一个或多个程序，该一个或多个程序包括指令，该指令当被包括多个应用程序的便携式电子设备执行时，能够使该便携式电子设备执行图 4 所示实施例的方法，并具体用于执行以下操作：基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中。

[92]其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[93]图 8 是本说明书的一个实施例提供的电子设备的结构示意图。请参考图 8，在硬件层面，该电子设备包括处理器，可选地还包括内部总线、网络接口、存储器。其中，存储器可能包含内存，例如高速随机存取存储器(Random-Access Memory, RAM)，也可能还包括非易失性存储器(non-volatile memory)，例如至少 1 个磁盘存储器等。当然，该电子设备还可能包括其他业务所需要的硬件。

[94]处理器、网络接口和存储器可以通过内部总线相互连接，该内部总线可以是 ISA(Industry Standard Architecture, 工业标准体系结构)总线、PCI(Peripheral Component Interconnect, 外设部件互连标准)总线或 EISA(Extended Industry Standard Architecture, 扩展工业标准结构)总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 8 中仅用一个双向箭头表示，但并不表示仅有一根总线或一种类型的总线。

[95]存储器，用于存放程序。具体地，程序可以包括程序代码，所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器，并向处理器提供指令和数据。

[96]处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行，在逻辑层面上形成视频匹配装置。处理器，执行存储器所存放的程序，并具体用于执行以下操作：

获取目标视频的多个特征向量；基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[97]采用本说明书实施例提供的电子设备可知道：在对目标视频进行侵权定位时，能够获取目标视频的多个特征向量，并基于该目标视频的多个特征向量，从视频数据库中检索与目标视频相似的候选视频，再基于目标视频和候选视频，构建目标视频和候选视频之间的时域相似度矩阵特征图，最后，将时域相似度矩阵特征图作为深度学习检测模型的输入，输出得到候选视频中与目标视频相匹配的视频片段和对应的相似度，一方面在侵权定位的效率上，能够检测可能的侵权视频的任意多个侵权片段的多个特征，同时利用向量检索结合检测模型能够极大地提高侵权视频的检测效率；另一方面，也降低了人工审核的成本。

[98]上述如本说明书图 2~图 4 所示实施例揭示的视频匹配方法可以应用于处理器中，或者由处理器实现。处理器可能是一种集成电路芯片，具有信号的处理能力。在实现过程中，上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器，包括中央处理器（Central Processing Unit, CPU）、网络处理器（Network Processor, NP）等；还可以是数字信号处理器（Digital Signal Processor, DSP）、专用集成电路（Application Specific Integrated Circuit, ASIC）、现场可编程门阵列（Field-Programmable Gate Array, FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本说明书一个或多个实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本说明书一个或多个实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成，或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器，闪存、只读存储器，可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器，处理器读取存储器中的信息，结合其硬件完成上述方法的步骤。

[99]该电子设备还可执行图 2~图 4 的视频匹配方法，本说明书在此不再赘述。

[100] 当然，除了软件实现方式之外，本说明书的电子设备并不排除其他实现方式，比如逻辑器件抑或软硬件结合的方式等等，也就是说以下处理流程的执行主体并不限定于各个逻辑单元，也可以是硬件或逻辑器件。

[101] 本说明书实施例还提出了一种计算机可读存储介质，该计算机可读存储介质存储一个或多个程序，该一个或多个程序包括指令，该指令当被包括多个应用程序的便携式电子设备执行时，能够使该便携式电子设备执行图 2~图 4 所示实施例的方法，并具体用于执行以下操作：获取目标视频的多个特征向量；基于所述目标视频的多个特征向量，

从视频数据库中检索与所述目标视频相似的候选视频；基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

[102] 总之，以上所述仅为本说明书的较佳实施例而已，并非用于限定本说明书的保护范围。凡在本说明书一个或多个实施例的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本说明书一个或多个实施例的保护范围之内。

[103] 上述实施例阐明的系统、装置、模块或单元，具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的，计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

[104] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体(transitory media)，如调制的数据信号和载波。

[105] 还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[106] 本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于系统实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

权利要求书

1、一种基于区块链的侵权存证方法，包括：

基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

2、一种视频匹配方法，包括：

获取目标视频的多个特征向量；

基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

3、如权利要求 2 所述的方法，基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频，包括：

从所述视频数据库中，获取与所述目标视频的多个特征向量相似的特征向量检索结果；

基于与所述目标视频的多个特征向量相似的特征向量检索结果，从视频数据库中获取与所述目标视频相似的候选视频。

4、如权利要求 3 所述的方法，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图，包括：

基于所述目标视频的多个特征向量、与所述候选视频中的各候选视频的多个特征向量之间的向量检索结果，构建所述目标视频的多个特征向量与所述候选视频的多个特征向量之间的相似度矩阵；

基于所述目标视频的多个特征向量与所述候选视频中的多个特征向量之间的相似

度矩阵，在时域维度上，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图。

5、如权利要求4所述的方法，基于所述目标视频的多个特征向量与所述候选视频中的各候选视频的多个特征向量之间的相似度矩阵，在时域维度上，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图，包括：

根据所述目标视频和所述候选视频之间的时域对应关系，将所述目标视频的多个特征向量与所述候选视频的多个特征向量之间的相似度矩阵，绘制在二维特征图上，得到所述目标视频和所述候选视频之间的时域相似度矩阵特征图。

6、如权利要求5所述的方法，当所述候选视频有多个时，根据所述目标视频和所述候选视频之间的时域对应关系，将所述目标视频的多个特征向量与所述候选视频中的各候选视频的多个特征向量之间的相似度矩阵，绘制在二维特征图上，得到所述目标视频和所述候选视频之间的时域相似度矩阵特征图，包括：

根据所述目标视频和所述多个候选视频之间的时域对应关系，分别将所述目标视频的多个特征向量与所述多个候选视频中的各候选视频的多个特征向量之间的相似度矩阵，绘制在多个二维特征图上，得到所述目标视频和所述多个候选视频之间的多个时域相似度矩阵特征图；

将所述目标视频和所述多个候选视频之间的多个时域相似度矩阵特征图进行拼接，得到所述目标视频和所述多个候选视频之间的时域相似度矩阵特征图。

7、如权利要求2或6所述的方法，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度，包括：

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段在时域维度上的区间范围、以及所述相匹配的视频片段之间的相似度。

8、如权利要求2所述的方法，所述深度学习检测模型包括下述至少一种：

更快速的基于卷积神经网络的候选区域检测模型 Faster-Rcnn；

带掩码的基于卷积神经网络的候选区域检测模型 Mask-Rcnn；

实时物体检测模型 YOLO；

单次多框检测模型 SSD。

9、一种基于区块链的侵权存证装置，包括：

候选视频检索模块，基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

特征图构建模块，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

模型输出模块，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

证据上传模块，当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对

的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

10、一种视频匹配装置，包括：

特征向量获取模块，获取目标视频的多个特征向量；

候选视频检索模块，基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

特征图构建模块，基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

模型输出模块，将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

11、一种电子设备，包括：

处理器；以及

被安排成存储计算机可执行指令的存储器，所述可执行指令在被执行时使所述处理器执行以下操作：

基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

12、一种计算机可读存储介质，所述计算机可读存储介质存储一个或多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行以下操作：

基于目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频

中与所述目标视频相匹配的视频片段和对应的相似度；

当所述候选视频中与所述目标视频相匹配的视频片段对应的相似度大于或等于预设相似度阈值时，将包含有所述目标视频的摘要、所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度的侵权证据上传至区块链中；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

13、一种电子设备，包括：

处理器；以及

被安排成存储计算机可执行指令的存储器，所述可执行指令在被执行时使所述处理器执行以下操作：

获取目标视频的多个特征向量；

基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。

14、一种计算机可读存储介质，所述计算机可读存储介质存储一个或多个程序，所述一个或多个程序当被包括多个应用程序的电子设备执行时，使得所述电子设备执行以下操作：

获取目标视频的多个特征向量；

基于所述目标视频的多个特征向量，从视频数据库中检索与所述目标视频相似的候选视频；

基于所述目标视频和所述候选视频，构建所述目标视频和所述候选视频之间的时域相似度矩阵特征图；

将所述时域相似度矩阵特征图作为深度学习检测模型的输入，以输出所述候选视频中与所述目标视频相匹配的视频片段和对应的相似度；

其中，所述深度学习检测模型为基于多组样本视频的时域相似度矩阵特征图和对应的标签训练得到的，其中样本视频包含查询视频和所述查询视频对应的候选视频，样本视频对应的标签包括样本视频中的查询视频在对应的候选视频中所匹配的视频片段及侵权标志。