

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021 年 12 月 16 日 (16.12.2021)



(10) 国际公布号
WO 2021/248733 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2020/116644
- (22) 国际申请日: 2020 年 9 月 22 日 (22.09.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010534822.3 2020年6月12日 (12.06.2020) CN
- (71) 申请人: 浙江大学(ZHEJIANG UNIVERSITY) [CN/CN]; 中国浙江省杭州市西湖区余杭塘路 866号, Zhejiang 310058 (CN)。
- (72) 发明人: 沈海斌(SHEN, Haibin); 中国浙江省杭州市西湖区余杭塘路866号, Zhejiang 310058 (CN)。
欧阳文汉(OUYANG, Wenhan); 中国浙江省杭州市西湖区余杭塘路866号, Zhejiang 310058 (CN)。
- (74) 代理人: 杭州求是专利事务所有限公司(HANGZHOU QIUSHI PATENT OFFICE CO.,

LTD.); 中国浙江省杭州市西湖区玉古路 147 号 黄鸿年科技综合大楼 215 室, Zhejiang 310013 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,

(54) **Title:** LIVE FACE DETECTION SYSTEM APPLYING TWO-BRANCH THREE-DIMENSIONAL CONVOLUTIONAL MODEL, TERMINAL AND STORAGE MEDIUM

(54) 发明名称: 应用双分支三维卷积模型的人脸活体检测系统、终端及存储介质

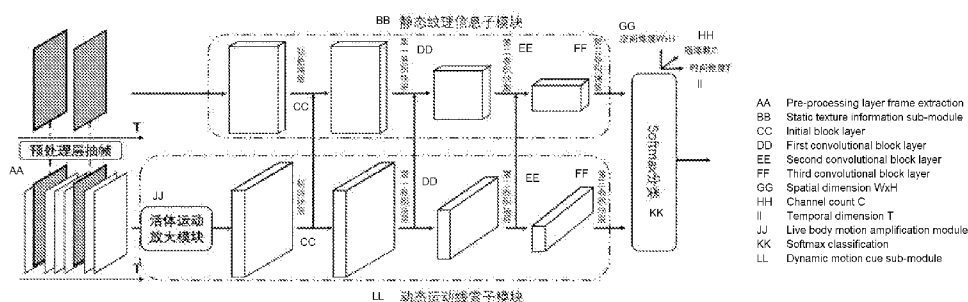


图 2

(57) **Abstract:** A live face detection system applying a two-branch three-dimensional convolutional model, a terminal and a storage medium relate to the technical field of live face detection. The live face detection system comprises a face video capturing module, a face video pre-processing module, a live body annotation module, a live body motion amplification module, a two-branch three-dimensional convolutional model training module and a live body judgement module. The two-branch three-dimensional convolutional model training module is configured with a two-branch three-dimensional convolutional model and comprises a static texture information sub-module, a dynamic motion cue sub-module, a fusion sub-module and a classification sub-module. Outputs of the static texture information sub-module and the dynamic motion cue sub-module are pooled, aggregated and fused through the fusion sub-module, and a detection result is produced through the classification sub-module. The two-branch three-dimensional convolutional model is mimic and biologically significant and has strong robustness and generalization. Live body detection of a human face recognition system is well guaranteed, system security is improved, and information and properties are protected from infringement.

IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：一种应用双分支三维卷积模型的人脸活体检测系统、终端及存储介质，属于人脸活体检测领域。包括人脸视频采集模块、人脸视频预处理模块、活体标注模块、活体运动放大模块、双分支三维卷积模型训练模块、活体判断模块；所述的双分支三维卷积模型训练模块配置有双分支三维卷积模型，包括静态纹理信息子模块、动态运动线索子模块、融合子模块和分类子模块；静态纹理信息子模块和动态运动线索子模块的输出经融合子模块进行池化、汇总、融合，再经分类子模块输出检测结果；所述双分支三维卷积模型是拟态的具有生物学意义的，并且具有很强的鲁棒性和泛化性；可以为人脸识别系统提供极大的活体检测保障，提升系统的安全性，保护信息和财产不受侵害。

应用双分支三维卷积模型的人脸活体检测系统、终端及存储介质

技术领域

本发明涉及一种人脸活体检测领域，具体涉及一种应用双分支三维卷积模型的人脸活体检测系统、终端及存储介质。

背景技术

随着人们越来越多得使用着像笔记本电脑和智能手机之类的电子设备，用他们付款，购物，支付账单以及社交互动，而与之而生的，对于电子身份认证的需求也越来越多。人脸识别验证在众多系统中脱颖而出，并在本发明生活中大规模的展开部署。为了保障安全，预防各种潜在的黑客攻击，人脸活体检测对于人脸验证系统是至关重要的一环。

目前，人脸活体检测算法面临的最大问题就是泛化性的不足，很多训练的模型在训练和对应的测试集上表现很出色，但是在全新的未知数据集中的表现却不尽如人意，这让人脸活体检测算法的实际部署价值大打折扣，所以针对这一现象，发明主要着眼于人脸活体检测模型泛化性的提高。

传统方法的模式众多，不尽相同。局部二值模式方法具有灰度不变性和旋转不变性等显著优点，有简单易算性但是较为简单。加速稳健特征方法使用海森矩阵的行列式值作特征点响应侦测并用积分图加速运算进行检测。但是无论是哪种传统方法，大部分的传统特征方法大多通过人工提取的特征，结合例如 SVM 以及 LDA 等传统的浅层分类器进行活体检测。基于传统人工特征提取方法又往往由于自身方法和训练样本的限制，只能针对特定的攻击方式、或者适用于特定的环境或者光照条件下。哪怕是综合性的多重传统特征提取的方法亦是如此，因为其阈值和参数往往通过人工设定，所以无法做到非常强的适应性和泛化性，不能够适用未知的场景和攻击方法，在实际场景中大多较为脆弱且不稳定。

交互式的方法虽然较为简单有效，不过整个验证时间过程会较长，且在便利性和用户体验上会给用户带来诸多负面感受，而且如果采用视频攻击的方式，往

往可以使基于眨眼检测、以及嘴唇运动等交互式方法失去效用，所以交互式人脸活体检测算法的局限性也较为明显。

目前更多的使用深度学习的方法来解决人脸活体检测问题。二维卷积神经网络是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于图像处理有出色表现。相较于局部二值模式方法等，其能够更好的提取出具有一定泛化性的二维图像特征，从而增加模型准确性。但是深度学习方法同样存在着一定的瓶颈，其模型固然在很多数据集上表现都非常出色，但是在跨数据集测试中仍然表现会不佳。这是由于大部分二维 CNN 模型只重点学习了训练样本的纹理特征，但是样本中的纹理特征因为不同的环境、不同的光线、不同的攻击方式、不同的显示设备材质，也会具有很强的差异性和随机多样性，所以对于为训练集合之外的全新样本的位置纹理特征就不能做到很好的拟合。

另外，会有一些方法试图通过提取人脸深度图，或者其他的辅助监督手段来额外的引入约束，来增强模型的泛化能力，但是这种辅助监督首先只是一种间接的监督手段，和人脸的活体检测的相关性目前尚无定论。而且其提取不仅需要耗费大量的计算量，而且会占用大量的硬盘空间，无论对于训练还是后续测试的都会带来诸多不便。

因此，模型的泛化问题一直是深度学习在活体检测领域应用中亟待解决的问题。

发明内容

为了解决现有的人脸活体检测领域算法泛化性能较差，不能够适用未知的场景和攻击方法，在实际场景中表现较为脆弱且不稳定的问题。本发明提供了一种应用双分支三维卷积模型的人脸活体检测系统、终端及存储介质，本发明采用三维卷积神经网络作为模型骨架，不仅能够抽取出高维的抽象化特征，而且更能从较浅层的网络中把一些具象化的浅层特征同样进行概括，从而获得更全面的时序运动特征。通过同时兼顾高维度、低维度的特征，让模型发挥更好的效果。同时，三维卷积神经网络在时域上的信息抽取能力更强，更适合作为适合处理人脸活体检测技术的技术框架。相较于普通二维卷积网络，能够更好地提取时域信息；而相较于循环神经网络，又能更多的均衡关注低阶和高阶特征信息，从总体上提升整个系统的泛化能力。

本发明的目的在于提供一种应用双分支三维卷积模型的人脸活体检测系统，包括：

人脸视频采集模块：用于采集用户的人脸视频；

人脸视频预处理模块：读取采集到的人脸视频，以 n 帧为一个单位进行分割处理，获得活体识别样本；

活体标注模块：用于对已知活体或非活体的训练样本进行标注，所述的活体标注模块在检测系统处于训练模式时开启，在检测系统处于识别模式时关闭；

活体运动放大模块：根据检测系统的运行模式，对带标签的训练样本或者不带标签的待检测样本进行活体运动信息放大处理，获得运动放大的活体识别样本；

双分支三维卷积模型训练模块：配置有双分支三维卷积模型，包括静态纹理信息子模块、动态运动线索子模块、融合子模块和分类子模块；所述的静态纹理信息子模块和动态运动线索子模块为三维卷积模型的两个分支，当检测系统处于训练模式时，所述活体运动放大模块的输出作为动态运动线索子模块的输入，人脸视频预处理模块的输出作为静态纹理信息子模块的输入，静态纹理信息子模块和动态运动线索子模块的输出经融合子模块进行池化、汇总、融合，再经分类子模块输出检测结果；

活体判断模块：当检测系统处于识别模式时，用于加载双分支三维卷积模型训练模块输出的模型文件，得到训练好的双分支三维卷积模型，并将人脸视频预处理模块输出的不带标签的待检测活体识别样本作为静态纹理信息子模块的输入，将活体运动放大模块输出的经运动放大处理后的不带标签的待检测活体识别样本作为动态运动线索子模块的输入，输出识别结果。

本发明的另一目的在于公开了一种终端，包括存储器和处理器；

所述存储器，用于存储计算机程序；

所述处理器，用于当执行所述计算机程序时，实现上述应用双分支三维卷积模型的人脸活体检测方法和系统的功能。

本发明的另一目的在于公开了一种计算机可读存储介质，其特征在于，所述存储介质上存储有计算机程序，当所述计算机程序被处理器执行时，实现上述应用双分支三维卷积模型的人脸活体检测方法和系统的功能。

本发明具备的有益效果是：

1) 本发明的三维卷积模型采用双分支结构，其中动态运动线索分支运行在一个高帧率的水平（25 帧每秒），更关注于对人脸或者伪造攻击的动态线索收集；静态纹理信息分支运行在低帧率水平（6.25 帧每秒），应用更精细的多尺度卷积方式，专注于区分真实人脸与伪造攻击的静态纹理信息特征提取，为整个系统的高效工作奠定基础，能同时提取静态空间纹理特征和时域运动特征，增强了系统泛化性。

在双分支模型中，对于高帧率的动态运动线索子模块，将通道数设置较小（初始为 8 通道，最终为 128 通道），一方面的目的是节约模型开销；另一方面，模型通道数目越多，则模型的对于静态特征的辨析和抽取能力也就越强，能捕捉更多的纹理和图案细节，而由于动态运动线索分支的通道数目较少，其对于这些静态空间域中的纹理特征的抽取能力也会降低，再考虑到模型高帧率的时序输入，所以该分支能够在降低对于空间域特征信息提取同时，让其更专精于时域信息的提取，获得更为纯粹的时域运动特征信息。而对于低帧率的静态纹理信息子模块，其对于时域变化不太敏感，将通道数设置较大（初始为 64 通道，最终为 1024 通道），由于其模型的输入和计算量本身较小，较高的通道数能够有效提升其对于空间纹理细节信息的提取能力。而对于普通的三维卷积神经网络，由于其模型庞大，且三维卷积核计算量开销昂贵，而实际情况下内存开销又十分有限，所以只能应用较为简单的结构，无法设计复杂的网络结构和训练技巧来对特征提取进行优化，无论在网络深度及特征通道数上都无法设置很大，也很难使用较为复杂的卷积核去提取特征，故模型的效果会受到一定的限制。

2) 对于静态纹理信息分支以及动态运动线索分支，为了引导其分别提取静态空间特征，以及动态时域特征，本发明设置了不同的卷积层结构。

对于动态运动线索分支，本发明将 $3 \times 3 \times 3$ 的时域卷积近似拆分成 $1 \times 1 \times 1$ ， $3 \times 1 \times 1$ ， $1 \times 3 \times 3$ ， $1 \times 1 \times 1$ 的四个卷积，这能够有效的节约模型冗余计算开支，同时保持模块对于时域信息的专注力，因为其相较于原有的 $3 \times 3 \times 3$ 的卷积核，能够有效的减少三维神经网络模型的计算量的冗余开销，为模型中其余计算留有更大的余地。同时，作为一种近似的三维卷积核方法，在获取时域和空间域信息的能力上，并没有比原三维卷积模式差，依旧能保证在人脸活体检测任务中的精度不下降，但是却可以节省 60% 以上的内存计算量，展示了很强的优势。

对于静态纹理信息子模块，本发明对于每一层采用多尺度的卷积核进行特征提取，首先输入到 $1 \times 1 \times 1$ 的卷积中，然后分别输入到 $1 \times 1 \times 1$ ， $1 \times 3 \times 3$ ， $1 \times 5 \times 5$ 的卷积中，最后将输出结果联结并输入到又一个 $1 \times 1 \times 1$ 的卷积中。这种多尺度卷积核的应用使得静态纹理信息子模块能够对于不同尺度的纹理和静态特征信息都能有比较好的提取能力，极大的增强了其对于静态平面空间信息的捕捉能力。这是由于纹理和图案特征的尺寸是不固定的，而使用固定尺寸的卷积核（例如 $1 \times 3 \times 3$ ）则会导致网络只对特定大小尺寸的特征更为敏感，从而更容易忽略一些其他尺寸的特征。相反，如果应用不同大小的卷积核，那么无论对于尺寸较大的全局特征（例如非活体样本表面全局性的莫尔条纹、或者水波纹特征），还是尺寸较小的细微局部特征（例如非活体表面局部的镜面反射纹理以及光斑），都能有合适的卷积核去具有针对性的提取特征，较大的卷积核可以用来提取较为粗糙的整体结构轮廓，而较小的卷积核可以用来提取较为微小的细节特征。

3）为了提升模型在时域维度上的辨别能力，更好的提取时域信息，本方法在最后的全局池化层之前，未在时间维度上进行任何形式的下采样，最大限度的保留和提取时域中有效的信息，这能够最大限度的帮助保留和提取时域中有效的信息，做到时域和空间域特征更好的平衡。而对于开销较为昂贵的普通三维卷积网络来讲，由于其复杂的结构，必须进行下采样来避免计算量溢出问题。

附图说明

图 1 为本发明实施例的简要流程示意图；

图 2 为本发明的双分支三维卷积模型的网络结构示意图；

图 3 为本发明实施例采用的双分支三维卷积模型的网络参数；

图 4 为本发明中动态运动线索子模块的时序卷积块示意图；

图 5 为本发明中静态纹理信息子模块的纹理卷积块示意图。

具体实施方式

下面结合附图和实施例对本发明做进一步的说明。

本发明的一个具体实施案展示了一种应用双分支三维卷积模型的人脸活体检测系统，包括：

人脸视频采集模块：用于采集用户的人脸视频；

人脸视频预处理模块：读取采集到的人脸视频，以 n 帧为一个单位进行分割处理，获得活体识别样本；

活体标注模块：用于对已知活体或非活体的训练样本进行标注，所述的活体标注模块在检测系统处于训练模式时开启，在检测系统处于识别模式时关闭；

活体运动放大模块：根据检测系统的运行模式，对带标签的训练样本或者不带标签的待检测样本进行运动放大处理，获得运动放大的活体识别样本；

双分支三维卷积模型训练模块：配置有双分支三维卷积模型，包括静态纹理信息子模块、动态运动线索子模块、融合子模块和分类子模块；所述的静态纹理信息子模块和动态运动线索子模块为三维卷积模型的两个分支，静态纹理信息子模块和动态运动线索子模块的输出经融合子模块进行池化、汇总、融合，再经分类子模块输出检测结果；

活体判断模块：当检测系统处于识别模式时，用于加载双分支三维卷积模型训练模块输出的模型文件，得到训练好的双分支三维卷积模型，并将人脸视频预处理模块输出的不带标签的待检测活体识别样本作为静态纹理信息子模块的输入，将活体运动放大模块输出的经运动放大处理后的不带标签的待检测活体识别样本作为动态运动线索子模块的输入，输出识别结果。

本发明的三维卷积模型采用了双分支结构，静态纹理信息子模块作为第一分支，活体运动放大模块和动态运动线索子模块作为第二分支。其中，所述的静态纹理信息子模块，包括输入层、时域步长为 k 且空间域步长为 $1*1$ 的预处理层、初始块层、以及 p 个卷积块层，优选的，所述的静态纹理信息子模块中的初始块层由通道数为 $8m$ 的初始化卷积层和初始化池化层构成。每一个卷积块层包括相同数量或不同数量的纹理卷积块，每一个纹理卷积块由卷积核为 $1*1*1$ 的第一卷积层、卷积核为 $1*1*1$ 、 $1*3*3$ 、 $1*5*5$ 的第二卷积层、卷积核为 $1*1*1$ 的第三卷积层构成；其中第一卷积块层的纹理卷积块中三层卷积层的通道数分别为 $8m$ 、 $8m$ 、 $32m$ ，且后一个卷积块层的纹理卷积块中三层卷积层的通道数是前一个卷积块层的 2 倍。初始化卷积层的卷积核为 $1*5*5$ ，初始化池化层的卷积核为 $1*3*3$ ， $2 \leq k \leq 5$ ， k 优选为 4， m 优选为 8。在本发明的一个具体实施中，纹理卷积块中的第二卷积层包含一个 $1*5*5$ 的卷积核，可以拆分为两个串联的 $1*3*3$ 卷积核。

其中，所述的动态运动线索子模块，包括输入层、初始块层、以及 p 个卷积块层，优选的，所述的动态运动线索子模块中的初始块层由通道数为 m 的始化

卷积层和初始化池化层构成。每一个卷积块层包括相同数量或不同数量的时序卷积块，每一个时序卷积块由卷积核为 $1*1*1$ 的第一卷积层、卷积核为 $3*1*1$ 的第二卷积层、卷积核为 $1*3*3$ 的第三卷积层、卷积核为 $1*1*1$ 的第四卷积层构成；其中第一卷积块层的时序卷积块中每一层卷积层的通道数分别为 m 、 m 、 m 、 $4m$ ；且后一个卷积块层的时序卷积块中四层卷积层的通道数是前一个卷积块层的 2 倍。初始化卷积层的卷积核为 $3*5*5$ ，初始化池化层的卷积核为 $1*3*3$ ， m 优选为 8。

所述静态纹理信息子模块中第 i 个卷积块层的输出与动态运动线索子模块中第 i 个卷积块层的输出进行合并后，作为静态纹理信息子模块中第 $i+1$ 个卷积块层的输入；其中， p 为大于 0 的整数， $1 \leq i \leq p-1$ ， p 优选为 3。

其中，所述的活体运动放大模块，在执行样本的运动放大处理过程时，具体为：

1) 通过傅里叶级数的分解，将每一帧中的人脸图像 $f(x + \delta(t))$ 分解为一列正弦函数之和：

$$f(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} A_{\omega} e^{i\omega(x+\delta(t))}$$

其中， $f(x + \delta(t))$ 表示时域中的人脸活体样本图像，即初始时的图像为 $I(x, 0) = f(x)$ ， $\delta(t)$ 为人脸的运动信息函数， A_{ω} 为转化到频域空间后信号的振幅；每个单独的频率 ω 对应一种带宽，而对于特定频率 ω 的带宽是复数的正弦信号：

$$S_{\omega}(x, t) = A_{\omega} e^{i\omega(x+\delta(t))}$$

其中，人脸微小运动的频率 ω 的范围设定为 0.3-3Hz，以提取人脸微小运动。 S_{ω} 为一条正弦曲线，其相位为 $\omega(x + \delta(t))$ 包含有原图像的运动信息；

2) 为了分离出特定的时域对应频带内的微小动作，对相位 $\omega(x + \delta(t))$ 进行过滤，得到滤波后的带通相位，表示如下：

$$B_{\omega}(x, t) = \omega \delta(t)$$

将带通相位 $B_{\omega}(x, t)$ 与 α 相乘， α 为运动信息放大系数，取值为 30，在实际应用中，可以根据需要在 10-50 之间变化，并加上子带 $S_{\omega}(x, t)$ 的相位，从而获取运动放大后的子带 $\hat{s}_{\omega}(x, t)$ ，表示为：

$$\hat{s}_{\omega}(x, t) = S_{\omega}(x, t) e^{i\alpha B_{\omega}} = A_{\omega} e^{i\omega(x+(1+\alpha)\delta(t))}$$

最终获得的结果 $\hat{s}_\omega(x, t)$ 是一条复数的正弦曲线，表示频域空间内经过运动放大后的图像。

3) 根据步骤 2.2) 中的运动放大后的子带 $\hat{s}_\omega(x, t)$ ，获得运动放大后的视频序列 $f(x + (1 + \alpha)\delta(t))$ ，最后将其转换回时域即为放大后的结果。

本发明的一个具体实施案展示了人脸活体检测系统的具体工作流程。

通过人脸视频采集模块获取用户的人脸视频，通过人脸视频预处理模块，以 8 帧为一个单位进行分割，获得活体识别样本。假设原始图像流尺寸为 224x224x8，在由活体运动放大模块进行人脸微小动作放大之后，输入到动态运动线索子模块。所述的动态运动线索子模块，运行在一个高帧率的水平（25 帧每秒），使用三维卷积模块，更关注于对人脸或者伪造攻击的动态线索收集。其通道数设置较小（初始为 8 通道，最终为 128 通道），一方面能够节约模型开销，另一方面也能够让其更专精于时域信息的提取，值得注意的是，整个过程中没有时域上的降采样处理，这在最大限度的保留了时域中的运动信息。

在动态运动线索子模块中的运算流程具体为：首先通过一个尺寸为 3x5x5，步长为 1,2,2 的卷积，获得通道数为 8 的特征；接着通过一个尺寸为 1x3x3，步长为 1,2,2 的初始化池化层，通道数仍为 8；然后通过第二分支中的三个卷积块层，所述的三个卷积块层分别包括 2、3、2 个时序卷积块，每个时序卷积块的结构如图 4 所示，为了节约计算开销和内存，本发明将原始的 3x3x3 的三维卷积核进行拆分，依次通过 1x1x1 的卷积核、3x1x1 的卷积核、1x3x3 的卷积核以及 1x1x1 的卷积核，设置 1x1x1 的卷积核的目的是增强模型的拟合能力。

同样假设原始图像尺寸为 224x224 x8，输入到静态纹理信息子模块中。所述的静态纹理信息子模块运行在低帧率水平（6.25 帧每秒），应用更精细的多尺度卷积方式，专注于区分真实人脸与伪造攻击的静态纹理信息特征提取。其对于时域变化不太敏感，通道数设置较大（初始为 64 通道，最终为 1024 通道），由于其模型的输入和计算量本身较小，较高的通道数能够有效提升其对于空间纹理细节信息的提取能力。

在静态纹理信息子模块中的运算流程具体为：通过预处理层进行抽帧之后，以低频率输入，首先通过一个 1x5x5 的卷积核，步长 1,2,2，获得的特征通道数

为 64, 接着通过一个尺寸为 $1 \times 3 \times 3$, 步长为 1,2,2 的初始化池化层, 通道数不变, 仍为 64。

初始块层的输出结果会与动态运动线索子模块对应层的输出结果进行合并拼接, 然后输入到卷积块层中。

然后经过第一分支中的三个卷积块层, 三个卷积块层分别包括 2、3、2 纹理卷积块, 每个纹理卷积块的结构如图 5 所示, 为了进一步节约模型的内存计算空间, 将图 4 中所示的 $1 \times 5 \times 5$ 的卷积拆分称为两个 $1 \times 3 \times 3$ 的卷积相串联。这种多尺度卷积核的方式使得模块对于不同静态空间尺度的特征都能有很强的信息提取能力。

在两个分支的运算过程中, 相应卷积块层的输出会进行联结后作为下一个第一分支卷积块层的输入。最终, 将运动线索子模块的输出结果与静态纹理信息子模块的输出结果合并后输入到全局池化层, 以及一个 1024 的全连接层, 最终通过 softmax 函数完成分类。

在构建完毕训练时需要的双分支三维卷积神经网络模型后, 针对于训练的样本和对应的标签, 本实施例通过以下方法对该模型进行训练, 并保存下模型文件到存储介质中。对于所有用于训练的样本, 进行批梯度下降训练, 即每次只送入网络模型一个批次共 10 个样本进行训练, 将一个批次内用于训练的样本记作 x , 其对应的标签记作 $\bar{y}$, 训练样本 x 在通过双分支的三维卷积神经网络模型的识别后, 得到模型的识别结果 y 。在本实施例中, 训练的目的就是缩小标签 $\bar{y}$ 和模型的识别结果 y 之间的差异, 因此选择交叉熵损失函数用于描述 $\bar{y}$ 和 y 之间的差异, 其交叉熵损失函数如下:

$$\mathbb{L} = -\frac{1}{10} \sum_{i=1}^{10} \sum_{j=1}^N \bar{y}_{ij} * \log(y_{ij})$$

其中 $\mathbb{L}$ 代表交叉熵损失函数, N 代表训练中识别任务的分类数, 此处为 2。 $\bar{y}_{ij}$ 代表一个批次中的第 i 个样本属于第 j 个类别的概率, y_{ij} 代表一个批次中的第 i 个样本在经过双分支的三维卷积神经网络模型网络后的识别结果属于第 j 个类别的概率。本实施例在 Pytorch toolkit 平台上通过批梯度下降方法, 先单独将第一分支与第二份独立训练两个周期, 然后将双分支模型合并后, 训练网络模型 50 个周期后, 将模型文件保存至存储介质中去, 以供活体判断模块进行人脸活体检测识别任务。所述的一个周期指将所有训练数据都通过批梯度下降方法训练过一次。

在本申请的一个实施案中，提供了一种终端和存储介质。

终端，它包括存储器和处理器；

其中存储器，用于存储计算机程序；

处理器，用于当执行所述计算机程序时，实现前述双分支的三维卷积神经网络模型方法与系统的功能。

需要注意的是，存储器可以包括随机存取存储器(Random Access Memory, RAM)，也可以包括非易失性存储器(Non-Volatile Memory, NVM)，例如至少一个磁盘存储器。上述的处理器为终端的控制中心，利用各种接口和线路连接终端的各个部分，通过执行存储器中的计算机程序来调用存储器中的数据，以执行终端的功能。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 Processing, 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。当然，该终端中还应当具有实现程序运行的必要组件，例如电源、通信总线等等。

示例性的，所述的计算机程序可以被分割为多个模块，每一个模块均被存储在存储器中，分割开来的每一个模块都可以完成具备特定功能的计算机程序指令段，该指令段用于描述计算机程序的执行过程。例如，可以将计算机程序分割成以下模块：

人脸视频采集模块：用于采集用户的人脸视频；

人脸视频预处理模块：读取采集到的人脸视频，以 n 帧为一个单位进行分割处理，获得活体识别样本；

活体标注模块：用于对已知活体或非活体的训练样本进行标注，所述的活体标注模块在检测系统处于训练模式时开启，在检测系统处于识别模式时关闭；

活体运动放大模块：根据检测系统的运行模式，对带标签的训练样本或者不带标签的待检测样本进行运动放大处理，获得运动放大的活体识别样本；

双分支三维卷积模型训练模块：配置有双分支三维卷积模型，包括静态纹理信息子模块、动态运动线索子模块、融合子模块和分类子模块；所述的静态纹理信息子模块和动态运动线索子模块为三维卷积模型的两个分支，静态纹理信息子

模块和动态运动线索子模块的输出经融合子模块进行池化、汇总、融合，再经分类子模块输出检测结果；

活体判断模块：当检测系统处于识别模式时，用于加载双分支三维卷积模型训练模块输出的模型文件，得到训练好的双分支三维卷积模型，并将人脸视频预处理模块输出的不带标签的待检测活体识别样本作为静态纹理信息子模块的输入，将活体运动放大模块输出的经运动放大处理后的不带标签的待检测样本作为动态运动线索子模块的输入，输出识别结果。

以上模块中的程序在执行时均由处理器进行处理。

此外，上述的存储器中的逻辑指令可以通过软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。存储器作为一种计算机可读存储介质，可设置为存储软件程序、计算机可执行程序，如本公开实施例中的系统对应的程序指令或模块。处理器通过运行存储在存储器中的软件程序、指令或模块，从而执行功能应用以及数据处理，即实现上述实施例中的功能。例如，U 盘、移动硬盘、只读存储器(Read-OnlyMemory, ROM)、随机存取存储器(RandomAccessMemory, RAM)、磁碟或者光盘等多种可以存储程序代码的介质，也可以是暂态存储介质。此外，上述存储介质以及终端中的多条指令由处理器加载并执行的具体过程在上述中已经详细说明。

实施例

本实施例用于展示一个具体的实施效果。本实施例中的人脸视频采集模块、人脸视频预处理模块、活体标注模块、活体运动放大模块、双分支三维卷积模型训练模块、活体判断模块均采用上述描述的结构及功能，此处不再赘述。

实施过程为：

包括配置过程和识别过程。首先设置系统处于配置模式，通过人脸视频采集模块获取人脸视频，然后由人脸视频预处理模块对人脸视频进行分割得到活体识别样本，并由活体标注模块进行标记，再通过活体运动放大模块进行处理，最后由活体检测拟态模训练模块根据训练样本集对三维卷积模型进行训练，保存为模型文件。

配置结束后，将系统设置为识别模式，首先通过人脸视频采集模块获取人脸视频，然后由人脸视频预处理模块对人脸视频进行分割得到待检测活体识别样本，

最后由活体判断模块直接加载训练好的模型文件，将待检测样本和运动放大后的待测样本分别作为模型输入，得到识别结果。

本实施例共进行了 6 种测试任务，包括了 OULU-NPU 数据库中的协议 1 测试，OULU-NPU 数据库中的协议 2 测试，OULU-NPU 数据库中的协议 3 测试，NPU 数据库中的协议 4 测试，以及 CASIA-FASD 数据库和回放攻击数据库之间分别进行的交叉测试。其中最困难的是 CASIA-FASD 数据库和回放攻击数据库之间分别进行的交叉测试，因为其对于模型泛化性能力，以及在未知照明、背景、设备条件下的鲁棒性有着很大的挑战。

对于在 OULU-NPU 数据库的四个协议测试，本发明遵循了原协议中的测试指标规则。其采用伪造攻击分类错误（APCER），其用来评估所有攻击方式中错分类错误率最高一种；以及真实活体分类错误率（BPCER），其用于评价真实活体样本分类错误率；及均值分类错误率（ACER），为伪造攻击分类错误和真实人脸分类错误率的均值：

$$ACER = \frac{APCER + BPCER}{2}$$

而在 CASIA 数据库和 Replay Attack 数据库中的测试，本发明遵循原数据库的测试标准，使用总错误率的一半（HTER）作为指标规则，其值为误拒率（FRR）和误接受率（FAR）之后的一半：

$$HTER = \frac{FAR + FRR}{2}$$

为保证公平，所以训练测试均在 Pytorch 基准测试平台上用 GeForce RTX 2080 Ti-NVIDIA GPU 进行，并且所有训练和测试规则均相同。这里所进行对比的模型包括了传统方法中的局部二值模式，循环神经网络中的长短期记忆神经网络，以及卷积神经网络中的二维卷积神经网络。结果如表 1 和表 2 所示。

表 1 各模型在 OULU-NPU 数据库中不同协议上的表现

协议	模型	APCER (%)	BPCER(%)	ACER(%)
1	二维卷积神经网络	20.0	41.7	30.8
	长短期记忆神经网络	18.2	23.4	20.7
	局部二值模式	5.0	20.8	13.1

	双分支三维卷积模型	1.5	3.0	2.2
2	二维卷积神经网络	25.3	23.9	24.6
	局部二值模式	22.5	6.7	14.6
	长短期记忆神经网络	19.8	10.3	10.1
	双分支三维卷积模型	5.6	3.2	4.4
3	局部二值模式	14.2±9.2	8.6±5.9	11.4±4.6
	二维卷积神经网络	21.4±7.7	25.3±9.6	23.3±2.3
	长短期记忆神经网络	15.5±5.9	23.6±2.8	19.5±4.4
	双分支三维卷积模型	2.0±0.5	3.7±1.3	2.8±0.9
4	局部二值模式	29.2±37.5	23.3±13.3	26.3±16.9
	二维卷积神经网络	21.7±8.2	44.2±11.1	32.9±5.8
	长短期记忆神经网络	15.1±8.9	18.4±3.3	16.7±6.2
	双分支三维卷积模型	12.1±5.3	10.5±7.3	11.6±6.6

表 2 各模型在 CASIA 和回放攻击数据库上的跨数据测试

模型	训练集 CASIA 测试集 Replay Attack HTER(%)	训练集 Replay Attack 测试集 CASIA HTER(%)
局部二值模式	39.79	42.04
二维卷积神经网络	48.6	41.81
长短期记忆神经网络	33.55	27.86
双分支三维卷积模型	18.2	20.4

可以看到本发明的双分支三维卷积模型在所有的数据测试中,在 OULU-NPU 数据库的四种测试协议下,通过和普通的二维卷积神经网络模型,以及传统的纹理特征模型进行比较,发明提出的三维拟态模型都占据了绝对的性能优势,由于 OULU-NPU 是一种接近于目前正在应用的实际手机场景的数据库,所以也可以侧面上证明了三维拟态模型在接近于实际移动支付场景中能够有效的防范各种非活体伪造攻击,具有很强的实用价值。无论相较于传统方法,还是二维卷积神

神经网络以及循环神经网络，都有非常巨大的优势。而在更具挑战的跨数据集测试中，模型性能的鲁棒性和优越性也得到了很好得体现。这说明本发明所述模型的结构是有效的，也是具有先进性的。

由于 OULU-NPU 虽然是一种考虑到跨场景跨设备的泛化性测试数据集，但是考虑到拍摄场景、光线的限制，以及同一批拍摄人员的摄影习惯是固定的、同一批攻击者的攻击方式习惯等细节也是相对单一的等客观条件，使得该数据库在测试过程中仍存在很多相似性的可能，所以并不能完全接近实际的复杂应用场景。所以将模型在 CASIA-FASD 数据库以及 Idiap Replay Attack 数据库中进行跨数据集的交叉测试，从而对模型进行更具挑战性的、完全接近实际场景的泛化性测试。在这些对比模型中，涵盖了很多不同种类的实验模型，包括一些传统的纹理提取算法，以及深度学习中的 CNN 以及 RNN 时序模型。

CASIA 与 Replay Attack 数据集上的交叉测试是一种模型泛化性的最高标准考验，因为两种数据集在采集设备，活体 ID，采集环境还是采集人员的拍摄习惯，都存在着极大的不同，彼此差异很大，所以非常符合现实中的检测场景。从表中可以看出，这里采用半总错误率 HTER 作为性能考察指标，通过和各类模型的综合性比较，包括传统的纹理特征提取模型、深度卷积学习中的 CNN 以及 LSTM 时序模型，发明提出的双分支三维卷积模型都展现出了更为优越的性能。

发明提出的双分支三维卷积模型无论相比于传统手工提取特征的方法，还是最为先进的复杂深度学习网络，都展现出了远超其它模型的泛化性表现，在两种规则的跨数据集互相测试中都是如此，由此可以看出提出的双分支三维卷积模型是更为鲁棒和强健的，所以也印证了其在实际场景测试中的表现也会更加优异。

以上列举的仅是本发明的具体实施例。显然，本发明不限于以上实施例，还可以有许多变形。本领域的普通技术人员能从本发明公开的内容直接导出或联想到的所有变形，均应认为是本发明的保护范围。

权利要求书

1. 一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，包括：

人脸视频采集模块：用于采集用户的人脸视频；

人脸视频预处理模块：读取采集到的人脸视频，以 n 帧为一个单位进行分割处理，获得活体识别样本；

活体标注模块：用于对已知活体或非活体的训练样本进行标注，所述的活体标注模块在检测系统处于训练模式时开启，在检测系统处于识别模式时关闭；

活体运动放大模块：根据检测系统的运行模式，对带标签的训练样本或者不带标签的待检测样本进行活体运动信息放大处理，获得运动放大的活体识别样本；

双分支三维卷积模型训练模块：配置有双分支三维卷积模型，包括静态纹理信息子模块、动态运动线索子模块、融合子模块和分类子模块；

当检测系统处于训练模式时，所述活体运动放大模块的输出作为动态运动线索子模块的输入，人脸视频预处理模块的输出作为静态纹理信息子模块的输入，静态纹理信息子模块和动态运动线索子模块的输出经融合子模块进行池化、汇总、融合，再经分类子模块输出检测结果；

所述静态纹理信息子模块包括输入层、时域步长为 k 且空间域步长为 $1*1$ 的预处理抽帧层、通道数为 $8m$ 的初始块层、以及 p 个卷积块层；所述动态运动线索子模块包括输入层、通道数为 m 的初始块层、以及 p 个卷积块层；

所述静态纹理信息子模块的初始块层的输出与动态运动线索子模块的初始块层的输出进行合并后，作为静态纹理信息子模块第 1 个卷积块层的输入；静态纹理信息子模块第 i 个卷积块层的输出与动态运动线索子模块第 i 个卷积块层的输出进行合并后，作为静态纹理信息子模块第 $i+1$ 个卷积块层的输入；所述静态纹理信息子模块的卷积块层和动态运动线索子模块的卷积块层均包括若干个由多层卷积层构成的卷积子模块，每一个卷积子模块的输出卷积层通道数大于输入卷积层通道数；其中 m 、 p 和 k 为大于 0 的整数， $2 \leq k \leq 5$ ， $1 \leq i \leq p-1$ ；

活体判断模块：当检测系统处于识别模式时，用于加载双分支三维卷积模型训练模块输出的模型文件，得到训练好的双分支三维卷积模型，并将人脸视频预处理模块输出的不带标签的待检测活体识别样本作为静态纹理信息子模块的输

入, 将活体运动放大模块输出的经运动放大处理后的不带标签的待检测活体识别样本作为动态运动线索子模块的输入, 输出识别结果。

2. 如权利要求 1 所述的一种应用双分支三维卷积模型的人脸活体检测系统, 其特征在于, 所述的活体运动放大模块具体为:

2.1) 通过傅里叶级数的分解, 将每一帧中的人脸图像 $f(x + \delta(t))$ 分解为一系列正弦函数之和:

$$f(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} A_{\omega} e^{i\omega(x+\delta(t))}$$

其中, $f(x + \delta(t))$ 表示时域中的人脸活体样本图像, 即初始时的图像为 $I(x, 0) = f(x)$, $\delta(t)$ 为人脸的运动信息函数, A_{ω} 为转化到频域空间后信号的振幅; i 代表复频域内图像所对应的虚部; 每个单独的频率 ω 对应一种带宽, 对于特定频率 ω 的带宽是复数的正弦信号:

$$S_{\omega}(x, t) = A_{\omega} e^{i\omega(x+\delta(t))}$$

其中, 人脸微小运动的频率 ω 的范围设定为 0.3-3Hz, 以提取人脸微小运动; S_{ω} 为一条正弦曲线, 其相位 $\omega(x + \delta(t))$ 包含有原图像的运动信息; 通过调节相位来调整运动的幅度;

2.2) 对上式中的 $\omega(x + \delta(t))$ 通过直流互补滤波器进行过滤, 得到滤波后的带通相位, 表示如下:

$$B_{\omega}(x, t) = \omega\delta(t)$$

将带通相位 $B_{\omega}(x, t)$ 与 α 相乘, 其中 α 为运动放大系数, 并加上子带 $S_{\omega}(x, t)$ 的相位, 从而获取运动放大后的子带 $\hat{s}_{\omega}(x, t)$, 表示为:

$$\hat{s}_{\omega}(x, t) = S_{\omega}(x, t) e^{i\alpha B_{\omega}} = A_{\omega} e^{i\omega(x+(1+\alpha)\delta(t))}$$

其中 $\hat{s}_{\omega}(x, t)$ 是一条复数的正弦曲线, 且正好为输入正弦曲线的 $(1 + \alpha)$ 倍;

2.3) 根据步骤 2.2) 中的运动放大后的子带 $\hat{s}_{\omega}(x, t)$, 获得运动放大后的视频序列 $f(x + (1 + \alpha)\delta(t))$, 最后将其转换回时域即为放大后的结果。

3. 如权利要求 1 所述的一种应用双分支三维卷积模型的人脸活体检测系统, 其特征在于, 所述的初始块层包括初始化卷积层和初始化池化层; 静态纹理信息子模块的初始化卷积层中的卷积核为 $1*5*5$, 初始化池化层的卷积核为 $1*3*3$; 动态运动线索子模块的初始化卷积层中的卷积核为 $3*5*5$, 初始化池化层的卷积核为 $1*3*3$ 。

4. 如权利要求 3 所述的一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，所述的静态纹理信息子模块包括三个卷积块层，三个卷积块层分别包括 2、3、2 个纹理卷积子模块，每一个纹理卷积子模块由卷积核为 $1*1*1$ 的第一卷积层、卷积核为 $1*1*1$ 、 $1*3*3$ 、 $1*5*5$ 的第二卷积层、卷积核为 $1*1*1$ 的第三卷积层构成，其中第一卷积块层的纹理卷积子模块中三层卷积层的通道数分别为 $8m$ 、 $8m$ 、 $32m$ ，且后一个卷积块层的纹理卷积子模块中三层卷积层的通道数是前一个卷积块层的 2 倍。
5. 如权利要求 4 所述的一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，将所述静态纹理信息子模块卷积块层中第二卷积层的 $1*5*5$ 卷积核拆分为两个串联的 $1*3*3$ 卷积核。
6. 如权利要求 1 所述的一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，所述的动态运动线索子模块包括三个卷积块层，三个卷积块层分别包括 2、3、2 个时序卷积子模块，每一个时序卷积子模块由卷积核为 $1*1*1$ 的第一卷积层、卷积核为 $3*1*1$ 的第二卷积层、卷积核为 $1*3*3$ 的第三卷积层、卷积核为 $1*1*1$ 的第四卷积层构成，其中第一卷积块层的时序卷积子模块中每一层卷积层的通道数分别为 m 、 m 、 m 、 $4m$ ；且后一个卷积块层的时序卷积子模块中四层卷积层的通道数是前一个卷积块层的 2 倍。
7. 如权利要求 1 所述的一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，所述的 m 取值为 8。
8. 如权利要求 1 所述的一种应用双分支三维卷积模型的人脸活体检测系统，其特征在于，所述的 p 取值为 3， k 取值为 4。
9. 一种终端，其特征在于，包括存储器和处理器；
所述存储器，用于存储计算机程序；
所述处理器，用于当执行所述计算机程序时，实现如权利要求 1~8 任一项所述应用双分支三维卷积模型的人脸活体检测系统。
10. 一种计算机可读存储介质，其特征在于，所述存储介质上存储有计算机程序，当所述计算机程序被处理器执行时，实现如权利要求 1~8 任一项应用双分支三维卷积模型的人脸活体检测系统。

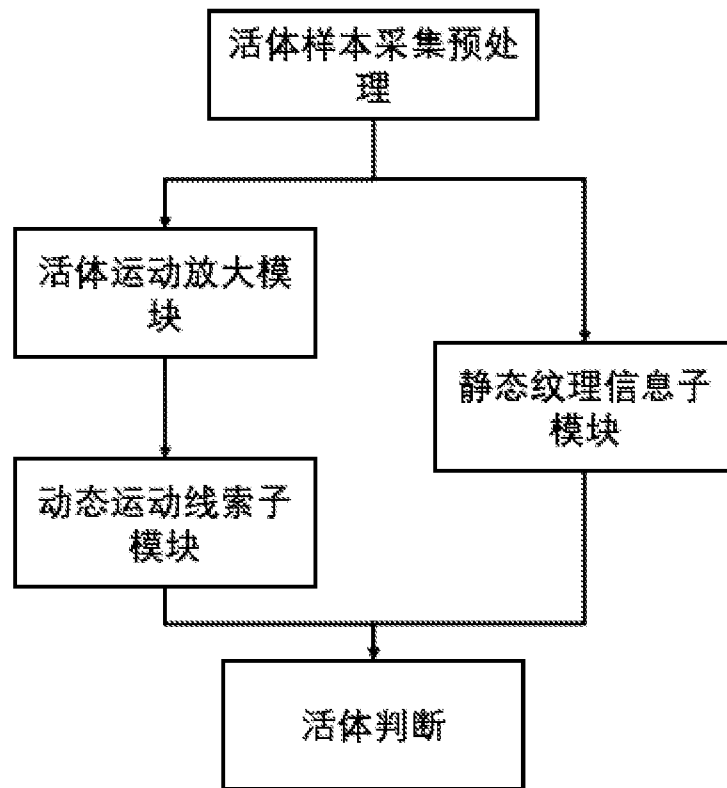


图 1

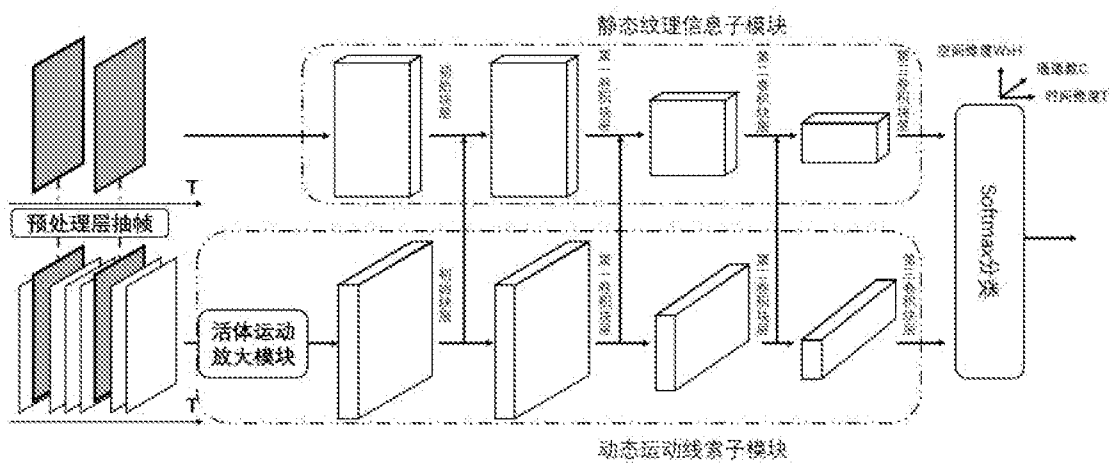


图 2

名称	静态纹理信息分支	动态运动线索分支 (帧率为25fps)
预处理层	步长4,1,1 (抽帧处理, 帧率为6.25fps)	
初始块层	卷积核1x5x5, 通道数64 步长1,2,2	卷积核3x5x5, 通道数8 步长1,2,2
	池化核1x3x3, 通道数64 步长1,2,2	卷积核1x3x3, 通道数8 步长1,2,2
第一卷积块层	$\begin{bmatrix} 1\times1\times1 & \text{通道数64} \\ 1\times1\times1, 1\times3\times3, 1\times5\times5 & \text{通道数64} \\ 1\times1\times1 & \text{通道数256} \end{bmatrix} \times 2$	$\begin{bmatrix} 1\times1\times1 & \text{通道数8} \\ 3\times1\times1 & \text{通道数8} \\ 1\times3\times3 & \text{通道数8} \\ 1\times1\times1 & \text{通道数32} \end{bmatrix} \times 2$
第二卷积块层	$\begin{bmatrix} 1\times1\times1 & \text{通道数128} \\ 1\times1\times1, 1\times3\times3, 1\times5\times5 & \text{通道数128} \\ 1\times1\times1 & \text{通道数512} \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1\times1 & \text{通道数16} \\ 3\times1\times1 & \text{通道数16} \\ 1\times3\times3 & \text{通道数16} \\ 1\times1\times1 & \text{通道数64} \end{bmatrix} \times 3$
第三卷积块层	$\begin{bmatrix} 1\times1\times1 & \text{通道数256} \\ 1\times1\times1, 1\times3\times3, 1\times5\times5 & \text{通道数256} \\ 1\times1\times1 & \text{通道数1024} \end{bmatrix} \times 2$	$\begin{bmatrix} 1\times1\times1 & \text{通道数32} \\ 3\times1\times1 & \text{通道数32} \\ 1\times3\times3 & \text{通道数32} \\ 1\times1\times1 & \text{通道数128} \end{bmatrix} \times 2$
全局均值池化, 特征合并, 全连接层		

图 3

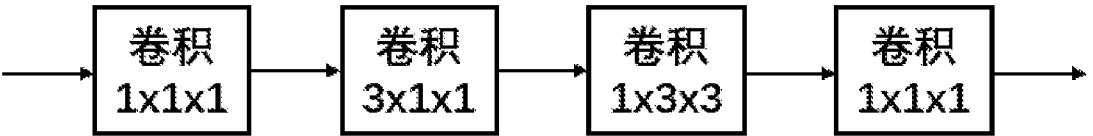


图 4

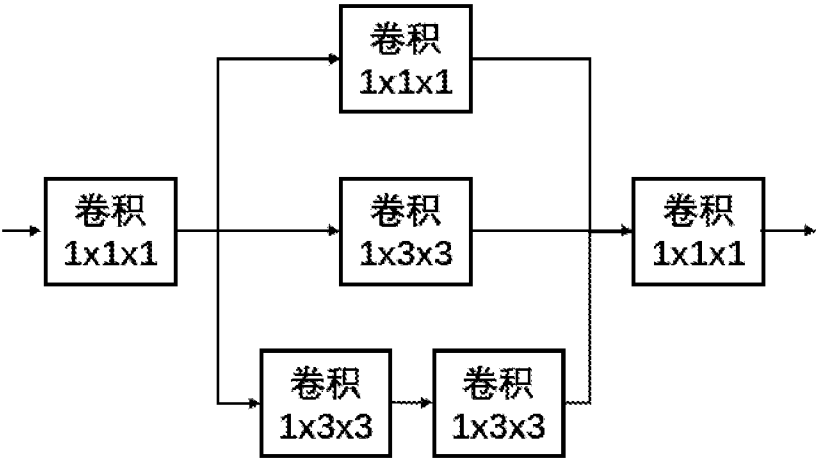


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/116644

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE: 人脸, 欺诈, 活体, 纹理, 运动, 放大, 三维, 卷积神经网络, 双分支, face, spoofing, living body, texture, motion, magnification, 3d, three-dimensional, convolutional neural network, two, double, branch, stream

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 110516576 A (XIDIAN UNIVERSITY) 29 November 2019 (2019-11-29) description paragraphs [0045]-[0111]	1-10
A	CN 109344716 A (CLOUDMINDS (SHENZHEN) ROBOTICS SYSTEMS CO., LTD.) 15 February 2019 (2019-02-15) entire document	1-10
A	CN 109886244 A (BEIJING SHIZHEN INTELLIGENT TECHNOLOGY CO., LTD.) 14 June 2019 (2019-06-14) entire document	1-10
A	US 9202105 B1 (AMAZON TECHNOLOGIES, INC.) 01 December 2015 (2015-12-01) entire document	1-10
A	BHARADWAJ, Samarth et al. "Computationally Efficient Face Spoofing Detection with Motion Magnification" 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 12 September 2013 (2013-09-12), pp. 105-110	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

24 February 2021

Date of mailing of the international search report

10 March 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088

China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/116644**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHEN, Haonan. "Attention-Based Two-Stream Convolutional Networks for Face Spoofing Detection" <i>IEEE Transactions on Information Forensics and Security</i> , Vol. 15, 17 June 2019 (2019-06-17), pp. 578-593	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/116644

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	110516576	A	29 November 2019	None	
CN	109344716	A	15 February 2019	None	
CN	109886244	A	14 June 2019	None	
US	9202105	B1	01 December 2015	None	

国际检索报告

国际申请号

PCT/CN2020/116644

A. 主题的分类

G06K 9/00 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06K

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

WPI, EPDOC, CNPAT, CNKI, IEEE: 人脸, 欺诈, 活体, 纹理, 运动, 放大, 三维, 卷积神经网络, 双分支, face, spoofing, living body, texture, motion, magnification, 3d, three-dimensional, convolutional neural network, two, double, branch, stream

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 110516576 A (西安电子科技大学) 2019年 11月 29日 (2019 - 11 - 29) 说明书第[0045]-[0111]段	1-10
A	CN 109344716 A (深圳前海达闼云端智能科技有限公司) 2019年 2月 15日 (2019 - 02 - 15) 全文	1-10
A	CN 109886244 A (北京视甄智能科技有限公司) 2019年 6月 14日 (2019 - 06 - 14) 全文	1-10
A	US 9202105 B1 (AMAZON TECHNOLOGIES, INC.) 2015年 12月 1日 (2015 - 12 - 01) 全文	1-10
A	BHARADWAJ, Samarth 等. "Computationally Efficient Face Spoofing Detection with Motion Magnification" 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013年 9月 12日 (2013 - 09 - 12), 第105-110页	1-10

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2021年 2月 24日

国际检索报告邮寄日期

2021年 3月 10日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

吴琼

电话号码 86-(10)-53961338

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CHEN, Haonan. "Attention-Based Two-Stream Convolutional Networks for Face Spoofing Detection" IEEE Transactions on Information Forensics and Security, 第15卷, 2019年 6月 17日 (2019 - 06 - 17), 第578-593页	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/116644

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	110516576	A	2019年 11月 29日	无	
CN	109344716	A	2019年 2月 15日	无	
CN	109886244	A	2019年 6月 14日	无	
US	9202105	B1	2015年 12月 1日	无	