

(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
2019 年 7 月 4 日 (04.07.2019)



(10) 国际公布号
WO 2019/128367 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2018/109169

(22) 国际申请日: 2018 年 9 月 30 日 (30.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711436879.4 2017 年 12 月 26 日 (26.12.2017) CN

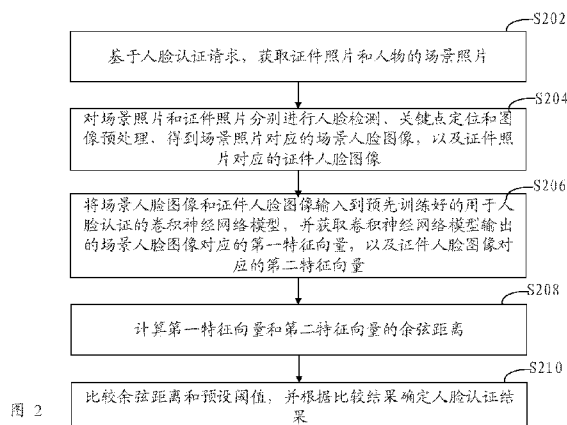
(71) 申请人: 广州广电运通金融电子股份有限公司 (GRG BANKING EQUIPMENT CO., LTD.) [CN/CN]; 中国广东省广州市高新技术产业开发区科学城科林路 9、11 号, Guangdong 510663 (CN)。

(72) 发明人: 许丹丹(XU, Dandan); 中国广东省广州市高新技术产业开发区科学城科林路 9、11 号, Guangdong 510663 (CN)。梁添才(LIANG, Tiancai); 中国广东省广州市高新技术产业开发区科学城科林路 9、11 号, Guangdong 510663 (CN)。章烈剽(ZHANG, Liepiao); 中国广东省广州市高新技术产业开发区科学城科林路 9、11 号, Guangdong 510663 (CN)。龚文川(GONG, Wenchuan); 中国广东省广州市高新技术产业开发区科学城科林路 9、11 号, Guangdong 510663 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房

(54) **Title:** FACE VERIFICATION METHOD AND APPARATUS BASED ON TRIPLET LOSS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 基于 Triplet Loss 的人脸认证方法、装置、计算机设备和存储介质



S202 BASED ON A FACE VERIFICATION REQUEST, ACQUIRE A CERTIFICATE PHOTOGRAPH AND A SCENARIO PHOTOGRAPH OF A PERSON
S204 RESPECTIVELY PERFORM FACE DETECTION, KEY POINT POSITIONING AND IMAGE PRE-PROCESSING ON THE SCENARIO PHOTO AND THE CERTIFICATE PHOTO, SO AS TO OBTAIN A SCENARIO FACE IMAGE CORRESPONDING TO THE SCENARIO PHOTOGRAPH AND A CERTIFICATE FACE IMAGE CORRESPONDING TO THE CERTIFICATE PHOTOGRAPH
S206 INPUT THE SCENARIO FACE IMAGE AND THE CERTIFICATE FACE IMAGE INTO A PRE-TRAINED CONVOLUTIONAL NEURAL NETWORK MODEL USED FOR FACE VERIFICATION, AND ACQUIRE A FIRST FEATURE VECTOR CORRESPONDING TO THE SCENARIO FACE IMAGE AND A SECOND FEATURE VECTOR CORRESPONDING TO THE CERTIFICATE FACE IMAGE, WHICH FEATURE VECTORS ARE OUTPUT BY THE CONVOLUTIONAL NEURAL NETWORK MODEL
S208 CALCULATE A COSINE DISTANCE BETWEEN THE FIRST FEATURE VECTOR AND THE SECOND FEATURE VECTOR
S210 COMPARE THE COSINE DISTANCE WITH A PRE-SET THRESHOLD VALUE, AND DETERMINE A FACE VERIFICATION RESULT ACCORDING TO A COMPARISON RESULT

(57) **Abstract:** The present invention relates to a face verification method and apparatus based on Triplet Loss, and a computer device and a storage medium. The method comprises: based on a face verification request, acquiring a certificate photograph and a scenario photograph of a person; respectively performing face detection, key point positioning and image pre-processing on the scenario photograph and the certificate photograph, so as to obtain a scenario face image corresponding to the scenario photograph and a certificate face image corresponding to the certificate photograph; inputting the scenario face image and the certificate face image into a pre-trained



(部位: 自编01-03和08-12单元) (仅限办公用途), Guangdong 510623 (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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, 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))。

convolutional neural network model used for face verification, and acquiring a first feature vector corresponding to the scenario face image and a second feature vector corresponding to the certificate face image, which feature vectors are output by the convolutional neural network model; calculating a cosine distance between the first feature vector and the second feature vector; and comparing the cosine distance with a pre-set threshold value, and determining a face verification result according to a comparison result. The method improves the reliability of face verification.

(57) 摘要: 本发明涉及一种基于Triplet Loss的人脸认证方法、装置、计算机设备和存储介质, 该方法包括: 基于人脸认证请求, 获取证件照片和人物的场景照片; 对场景照片和证件照片分别进行人脸检测、关键点定位和图像预处理, 得到场景照片对应的场景人脸图像, 以及证件照片对应的证件人脸图像; 将场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型, 并获取卷积神经网络模型输出的场景人脸图像对应的第一特征向量, 以及证件人脸图像对应的第二特征向量; 计算第一特征向量和第二特征向量的余弦距离; 比较余弦距离和预设阈值, 并根据比较结果确定人脸认证结果。该方法提高了人脸认证的可靠性。

基于 Triplet Loss 的人脸认证方法、装置、计算机设备和存储介质

技术领域

本发明涉及图像处理技术领域,特别是涉及一种基于 Triplet Loss 的人脸认证方法、装置、计算机设备和存储介质。

背景技术

人脸认证,是指对比现场采集的人物场景照片以及身份信息中的证件照片,判断是否为同一个人。人脸认证的关键技术为人脸识别。

随着深度学习技术的兴起,人脸识别的相关问题不断突破传统的技术瓶颈,性能水平得到较大的提升。在运用深度学习解决人脸识别问题的研究工作中,主要有两派主流的方法:基于分类学习的方法和基于度量学习的方法。其中,基于分类学习的方法主要是在深度卷积网络提取特征之后计算样本的分类损失函数(比如 softmax loss、center loss 及相关变体)来对网络进行优化,网络最后一层是用于分类的全连接层,其输出节点的数量往往要与训练数据集的总类别数保持一致,该方法适用于训练样本较多,尤其是同一类别的训练样本比较丰富的情况,网络可以得到较好的训练效果和泛化能力。但当类别数达到数十万或更高数量级时,网络最后的分类层(全连接层)参数量会呈线性增长而相当庞大,导致网络难以训练。

另一类方法是基于度量学习的方法,该方法以元组的方式组织训练样本(比如二元组 pair 或者三元组 triplet),在深度卷积网络之后无需通过分类层,而是直接基于卷积特征向量计算样本间的度量损失(比如 contrastive loss、triplet loss 等)来对网络进行优化,该方法不需要训练分类层,因此网络参数量不受类别数增长的影响,对训练数据集的类别数无限定,只需要根据相应策略选取同类或异类样本构造合适的元组即可。相比分类学习方法,度量学习方法更适用于训练数据广度较大但深度不足(样本类别数多,但同类样本少)的情况,通过样本之间的不同组合,可以构造相当丰富的元组数据用于训练,同时度量学习方式更加关注元组内部关系,对于 1:1 人脸验证这类判断是与不是的问题有其先天的优势。

在实际应用中,许多的机构都要求实名制登记,例如,银行开户,手机号码登记、金融账号开户等等。实名制登记要求用户携带身份证到指定的地点,由工作人员验证本人与身份证的照片对应后,方可开户成功。而随着互联网技术的发展,越来越多的机构推出了便民服务,不再强制要求客户到指定网点。用户的地理位置不受限制,上传身份证,并利用移动终端的图像采集装置采集现场的人物场景照片,由系统进行人脸认证,并在人脸认证通过后,

即可开户成功。而传统地基于度量的学习方法，使用欧式距离来度量样本之间的相似度，而欧氏距离衡量的是空间各点的绝对距离，跟各个点所在的位置坐标直接相关，这并不符合人脸特征空间的分布属性，导致人脸识别的可靠性较低。

发明内容

基于此，有必要针对传统的人脸认证方法可靠性低的问题，提供一种基于 Triplet Loss 的人脸认证方法、装置、计算机设备和存储介质。

一种基于 Triplet Loss 的人脸认证方法，包括：

基于人脸认证请求，获取证件照片和人物的场景照片；

对所述场景照片和所述证件照片分别进行人脸检测、关键点定位和图像预处理，得到所述场景照片对应的场景人脸图像，以及所述证件照片对应的证件人脸图像；

将所述场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型，并获取所述卷积神经网络模型输出的所述场景人脸图像对应的第一特征向量，以及所述证件人脸图像对应的第二特征向量；其中，所述卷积神经网络模型基于三元组损失函数的监督训练得到；

计算所述第一特征向量和所述第二特征向量的余弦距离；

比较所述余弦距离和预设阈值，并根据比较结果确定人脸认证结果。

在一个实施例中，所述方法还包括：

获取带标记的训练样本，所述训练样本包括标记了属于每个标记对象的一张证件人脸图像和至少一张场景人脸图像；

根据所述训练样本训练卷积神经网络模块，通过 OHEM 产生各训练样本对应的三元组元素；所述三元组元素包括参考样本、正样本和负样本；

根据各训练样本的三元组元素，基于三元组损失函数的监督，训练所述卷积神经网络模型；该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数；

将验证集数据输入所述卷积神经网络模型，达到训练结束条件时，得到训练好的用于人脸认证的卷积神经网络模型。

在另一个实施例中，根据所述训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素的步骤，包括：

随机选择一个图像作为参考样本，选择属于同一标签对象、与参考样本类别不同的图像作为正样本；

根据 OHEM 策略, 利用当前训练的卷积神经网络模型提取特征之间的余弦距离, 对于每一个参考样本, 从其它不属于所述标签对象的图像中, 选择距离最小、与所述参考样本属于不同类别的图像, 作为该参考样本的负样本。

在另一个实施例中, 所述三元组损失函数包括对同类样本的余弦距离的限定, 以及对异类样本的余弦距离的限定。

在另一个实施例中, 所述三元组损失函数为:

$$\tilde{L} = \sum_i^N [\cos(f(x_i^a), f(x_i^n)) - \cos(f(x_i^a), f(x_i^p)) + \alpha_1]_+ + \sum_i^N [\alpha_2 - \cos(f(x_i^a), f(x_i^p))]_+$$

其中, $\cos(\bullet)$ 表示余弦距离, 其计算方式为 $\cos(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$, N 是三元组数量, $f(x_i^a)$ 表示参考样本的特征向量, $f(x_i^p)$ 表示同类正样本的特征向量, $f(x_i^n)$ 表示异类负样本的特征向量, $[\bullet]_+$ 的含义如下: $[x]_+ = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$, α_1 为类间间隔参数, α_2 为类内间隔参数。

在另一个实施例中, 所述方法还包括: 利用基于海量开源人脸数据训练好的基础模型参数进行初始化, 在特征输出层后添加归一化层及三元组损失函数层, 得到待训练的卷积神经网络模型。

一种基于 Triplet Loss 的人脸认证装置, 包括: 图像获取模块、图像预处理模块、特征获取模块、计算模块和认证模块;

所述图像获取模块, 用于基于人脸认证请求, 获取证件照片和人物的场景照片;

所述图像预处理模块, 用于对所述场景照片和所述证件照片分别进行人脸检测、关键点定位和图像预处理, 得到所述场景照片对应的场景人脸图像, 以及所述证件照片对应的证件人脸图像;

所述特征获取模块, 用于将所述场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型, 并获取所述卷积神经网络模型输出的所述场景人脸图像对应的第一特征向量, 以及所述证件人脸图像对应的第二特征向量; 其中, 所述卷积神经网络模型基于三元组损失函数的监督训练得到;

所述计算模块, 用于计算所述第一特征向量和所述第二特征向量的余弦距离;

所述认证模块, 用于比较所述余弦距离和预设阈值, 并根据比较结果确定人脸认证结果。

在另一个实施例中，所述装置还包括：样本获取模块、三元组获取模块、训练模块和验证模块；

所述样本获取模块，用于获取带标记的训练样本，所述训练样本包括标记了属于每个标记对象的一张证件人脸图像和至少一张场景人脸图像；

所述三元组获取模块，用于根据所述训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素；所述三元组元素包括参考样本、正样本和负样本；

所述训练模块，用于根据各训练样本的三元组元素，基三元组损失函数的监督，训练所述卷积神经网络模型；该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数；

所述验证模块，用于将验证集数据输入所述卷积神经网络模型，达到训练结束条件时，得到训练好的用于人脸认证的卷积神经网络模型。

一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述的基于 Triplet Loss 的人脸认证方法的步骤。

一种存储介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时，实现上述的基于 Triplet Loss 的人脸认证方法的步骤。

本发明所述的基于 Triplet Loss 的人脸认证方法、装置、计算机设备和存储介质，利用预先训练的卷积神经网络进行人脸认证，由于卷积神经网络模型基于三元组损失函数的监督训练得到，而场景人脸图像和证件人脸图像的相似度根据场景人脸图像对应的第一特征向量和证件人脸图像对应的第二特征向量的余弦距离计算得到，余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，从而更符合人脸特征空间的分布属性，提高了人脸认证的可靠性。

附图说明

图 1 为一个实施例的基于 Triplet Loss 的人脸认证系统的结构示意图；

图 2 为一个实施例中基于 Triplet Loss 的人脸认证方法的流程图；

图 3 为一个实施例中训练得到用于人脸认证的卷积神经网络模型的步骤的流程图；

图 4 为在类间间隔一致、类内方差较大情况下，样本错分的概率示意图；

图 5 为在类间间隔一致、类内方差较小情况下，样本错分的概率示意图；

图 6 为一个实施例中基于 Triplet Loss 的人脸认证的迁移学习过程的示意图；

图 7 为一个实施例中用于人脸认证的卷积神经网络模型的结构示意图；

图 8 为一个实施例中基于 Triplet Loss 的人脸认证方法的流程示意图；

图 9 为一个实施例中基于 Triplet Loss 的人脸认证装置的结构框图；

图 10 为另一个实施例中基于 Triplet Loss 的人脸认证装置的结构框图。

具体实施方式

图 1 为一个实施例的基于 Triplet Loss 的人脸认证系统的结构示意图。如图 1 所示，人脸认证系统包括服务器 101 和图像采集装置 102。其中，服务器 101 与图像采集装置 102 网络连接。图像采集装置 102 采集待认证用户的实时场景照片，以及证件照片，并将采集的实时场景照片和证件照片发送至服务器 101。服务器 101 判断场景照片的人物与证件照中的人物是否为同一人，对待认证用户的身份进行认证。基于具体的应用场景，图像采集装置 102 可以为摄像头，或是具有摄像功能的用户终端。以在开户现场为例，图像采集装置 102 可以为摄像头；以通过互联网进行金融账号开户为例，图像采集装置 102 可以为具有摄像功能的移动终端。

在其它的实施例中，人脸认证系统还可以包括读卡器，用于读取证件（如身份证等）芯片内的证件照。

图 2 为一个实施例中基于 Triplet Loss 的人脸认证方法的流程图。如图 2 所示，该方法包括：

S202，基于人脸认证请求，获取证件照片和人物的场景照片。

其中，证件照片是指能够证明人物身份的证件所对应的照片，例如身份证上所印制的证件照或芯片内的证件照。证件照片的获取方式可以采用对证件进行拍照获取，也可以通过读卡器读取证件芯片所存储的证件照片。本实施例中的证件可以为身份证，驾驶证或社会保障卡等。

人物的场景照片是指待认证用户在认证时所采集，该待认证用户在现场环境的照片。现场环境是指用户在拍照时的所处环境，现场环境不受限制。场景照片的获取方式可以为，利用具有摄像功能的移动终端采集场景照片并发送至服务器。

人脸认证，是指对比现场采集的人物场景照片以及身份信息中的证件照片，判断是否为同一个人。人脸认证请求基于实际的应用操作触发，例如，基于用户的开户请求，触发人脸认证请求。应用程序在用户终端的显示界面提示用户进行照片的采集操作，并在照片采集完成后，将采集的照片发送至服务器，进行人脸认证。

S204，对场景照片和证件照片分别进行人脸检测、关键点定位和图像预处理，得到场景照片对应的场景人脸图像，以及证件照片对应的证件人脸图像。

人脸检测是指识别照片并获取照片中的人脸区域。

关键点定位，是指对照片中检测的人脸区域，获取人脸关键点在每幅照片中的位置。人脸关键点包括眼睛，鼻尖、嘴角尖、眉毛以及人脸各部件轮廓点。

本实施例中，可采用基于多任务联合学习的级联卷积神经网络 MTCNN 方法同时完成人脸检测和人脸关键点检测，亦可采用基于 LBP 特征的人脸检测方法和基于形状回归的人脸关键点检测方法。

图像预处理是指将根据检测的人脸关键点在每张图片中的位置，进行人像对齐和剪切处理，从而得到尺寸归一化的场景人脸图像和证件人脸图像。其中，场景人脸图像是指对场景照片进行人脸检测、关键点定位和图像预处理后得到的人脸图像，证件人脸图像是指对证件照片进行人脸检测、关键点定位和图像预处理后得到的人脸图像。

S206，将场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型，并获取卷积神经网络模型输出的场景人脸图像对应的第一特征向量，以及证件人脸图像对应的第二特征向量。

其中，卷积神经网络模型基于三元组损失函数的监督预先根据训练样本提前训练好的。卷积神经网络包括卷积层、池化层、激活函数层和全连接层，每层的各个神经元参数通过训练确定。利用训练好的卷积神经网络，通过网络前向传播，获取卷积神经网络模型的全连接层输出的场景人脸图像的第一特征向量，以及证件人脸图像对应的第二特征向量。

三元组 (triplet) 是指从训练数据集中随机选一个样本，该样本称为参考样本，然后再随机选取一个和参考样本属于同一人的样本作为正样本，选取不属于同一人的样本作为负样本，由此构成一个 (参考样本、正样本、负样本) 三元组。由于人证比对主要是基于证件照与场景照的比对，而不是证件照与证件照、或者场景照与场景照的比对，因此三元组的模式主要有两种组合：以证件照图像为参考样本时，正样本和负样本均为场景照；以场景照图像为参考样本时，正样本和负样本均为证件照。

针对三元组中的每个样本，训练一个参数共享的网络，得到三个元素的特征表达。改进三元组损失 (triplet loss) 的目的就是通过学习，让参考样本和正样本的特征表达之间的距离尽可能小，而参考样本和负样本的特征表达之间的距离尽可能大，并且要让参考样本和正样本的特征表达之间的距离和参考样本和负样本的特征表达之间的距离之间有一个最小的间隔。

S208，计算第一特征向量和第二特征向量的余弦距离。

余弦距离，也称为余弦相似度，是用向量空间中两个向量夹角的余弦值作为衡量两个个体间差异的大小的度量。第一特征向量和第二特征向量的余弦距离越大，表示场景人脸图像

和证件人脸图像的相似度越大，第一特征向量和第二特征向量的余弦距离越小，表示场景人脸图像和证件人脸图像的相似度越小。当场景人脸图像和证件人脸图像的余弦距离越接近于 1 时，两张图像属于同一人的机率越大，当场景人脸图像和证件人脸图像的余弦距离越小，两张图像属于同一人的机率越小。

传统的三元组损失（triplet loss）方法中，使用欧式距离来度量样本之间的相似度。而欧氏距离衡量的是空间各点的绝对距离，跟各个点所在的位置坐标直接相关，这并不符合人脸特征空间的分布属性。本实施例中，考虑人脸特征空间的分布属性和实际应用场景，采用余弦距离来度量样本之间的相似度。余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，而不是位置，从而更符合人脸特征空间的分布属性。

具体地，余弦距离的计算公式为：

$$\cos(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$$

其中，x 表示第一特征向量，y 表示第二特征向量。

S210，比较余弦距离和预设阈值，并根据比较结果确定人脸认证结果。

认证结果包括认证通过，即证件照片和场景照片属于同一人。认证结果还包括认证失败，即证件照片和场景照片不属于同一人。

具体地，将余弦距离与预设阈值进行比较，当余弦距离大于预设阈值时，表示即证件照片与场景照片的相似度大于预设阈值，认证成功，当余弦距离小于预设阈值时，表示即证件照片与场景照片的相似度小于预设阈值，认证失败。

上述的基于 Triplet Loss 的人脸认证方法，利用预先训练的卷积神经网络进行人脸认证，由于卷积神经网络模型基于三元组损失函数的监督训练得到，而场景人脸图像和证件人脸图像的相似度根据场景人脸图像对应的第一特征向量和证件人脸图像对应的第二特征向量的余弦距离计算得到，余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，从而更符合人脸特征空间的分布属性，提高了人脸认证的可靠性。

在另一个实施例中，人脸认证方法还包括训练得到用于人脸认证的卷积神经网络模型的步骤。图 3 为一个实施例中训练得到用于人脸认证的卷积神经网络模型的步骤的流程图。如图 3 所示，该步骤包括：

S302，获取带标记的训练样本，训练样本包括标记了属于每个标记对象的一张证件人脸图像和至少一张场景人脸图像。

本实施例中，标记对象即人，训练样本以人为单位，标记了同属于一个人的场景人脸图

像和证件人脸图像。具体地，场景人脸图像和证件人脸图像可通过对带标记的场景照片和证件照片进行人脸检测、关键点定位和图像预处理得到。

人脸检测是指识别照片并获取照片中的人脸区域。

关键点定位，是指对照片中检测的人脸区域，获取人脸关键点在每幅照片中的位置。人脸关键点包括眼睛，鼻尖、嘴角尖、眉毛以及人脸各部件轮廓点。

本实施例中，可采用基于多任务联合学习的级联卷积神经网络 MTCNN 方法同时完成人脸检测和人脸关键，亦可采用基于 LBP 特征的人脸检测方法和基于形状回归的人脸关键点检测方法。

图像预处理是指将根据检测的人脸关键点在每张图片中的位置，进行人像对齐和剪切处理，从而得到尺寸归一化场景人脸图像和证件人脸图像。其中，场景人脸图像是指对场景照片进行人脸检测、关键点定位和图像预处理后得到的人脸图像，证件人脸图像是指对证件照片进行人脸检测、关键点定位和图像预处理后得到的人脸图像。

S304，根据训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素；三元组元素包括参考样本、正样本和负样本。

三元组有两种组合方式：以证件照图像为参考样本时，正样本和负样本均为场景照图像；以场景照图像为参考样本时，正样本和负样本均为证件照图像。

具体地，以证件照为参考图像为例，从训练数据集中随机选一个人的证件照样本，该样本称为参考样本，然后再随机选取一个和参考样本属于同一人的场景照样本作为正样本，选取不属于同一人的场景照样本作为负样本，由此构成一个（参考样本、正样本、负样本）三元组。

即正样本与参考样本为同类样本，即属于同一人图像。负样本是参考样本的异类样本，即不属于同一人的图像。其中，三元组元素中的参考样本和正样本是训练样本中已标记的，负样本在卷积神经网络的训练过程中，采用 OHEM（Online Hard Example Mining）策略在线构造三元组，即在网络每次迭代优化的过程中，利用当前网络对候选三元组进行前向计算，选择训练样本中与参考样本不属于同一用户，且余弦距离最近的图像作为负样本，从而得到各训练样本对应的三元组元素。

一个实施例中，根据训练样本训练卷积神经网络，并产生各训练样本对应的三元组元素的步骤，包括以下步骤 S1 和 S2：

S1：随机选择一个图像作为参考样本，选择属于同一标签对象、与参考样本类别不同的图像作为正样本。

类别是指所属的图像类型，本实施例中，训练样本的类别包括场景人脸图像和证件人脸图像。因为人脸认证主要是证件照和场景照之间的对比，因此，参考样本和正样本应当属于不同的类别，若参考样本为场景人脸图像，则正样本为证件人脸图像；若参考样本为证件人脸图像，则正样本为场景人脸图像。

S2：根据 OHEM 策略，利用当前训练的卷积神经网络模型提取特征之间的余弦距离，对于每一个参考样本，从其它不属于同一标签对象的图像中，选择距离最小、与参考样本属于不同类别的图像，作为该参考样本的负样本。

负样本从与参考样本不属于同一人的标签的人脸图像中选择，具体地，负样本在卷积神经网络的训练过程中，采用 OHEM 策略在线构造三元组，即在网络每次迭代优化的过程中，利用当前网络对候选三元组进行前向计算，选择训练样本中与参考样本不属于同一用户，且余弦距离最近、与参考样本不属于同一类别的图像作为负样本。即，负样本与参考样本的类别不同。可以认为，三元组中若以证件照为参考样本，则正样本和负样本均是场景照；反之若以场景照为参考样本，则另外正样本和负样本均是证件照。

S306，根据各训练样本的三元组元素，基于三元组损失函数的监督，训练卷积神经网络模型，该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数。

人证核验终端通过比对用户证件芯片照与场景照是否一致来对用户身份进行验证，后台采集到的数据往往是单个人的样本只有两张图，即证件照与比对时刻抓拍到的场景照，而不同个体的数量却可以成千上万。这种类别数量较大而同类样例少的数据如果用基于分类的方法来进行训练，分类层参数会过于庞大而导致网络非常难以学习，因此考虑用度量学习的方法来解决。其中度量学习的典型的一般是用三元组损失（triplet loss）方法，通过构造图像三元组来学习一种有效的特征映射，在该映射下同类样本的特征距离小于异类样本的特征距离，从而达到正确比对的目的。

三元组损失（triplet loss）的目的就是通过学习，让参考样本和正样本的特征表达之间的距离尽可能小，而参考样本和负样本的特征表达之间的距离尽可能大，并且要让参考样本和正样本的特征表达之间的距离和参考样本和负样本的特征表达之间的距离之间有一个最小的间隔。

在另一个实施例中，三元组损失函数包括对同类样本的余弦距离的限定，以及对异类样本的余弦距离的限定。

其中，同类样本是指参考样本和正样本，异类样本是指参考样本和负样本。同类样本的

余弦距离是指参考样本和正样本的余弦距离，异类样本的余弦距离是指参考样本和负样本的余弦距离。

一方面，原始的 triplet loss 方法只是考虑了类间差距而没有考虑类内差距，如果类内分布不够聚敛，网络的泛化能力就会减弱，对场景适应性也会随之降低。另一方面，原始的 triplet loss 方法采用的是欧式距离来度量样本之间的相似度，实际上人脸模型部署后在特征比对环节，更多地会采用余弦距离来进行度量。欧氏距离衡量的是空间各点的绝对距离，跟各个点所在的位置坐标直接相关；而余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，而不是位置，从而更符合人脸特征空间的分布属性。

采用 triplet loss 方法，通过在线构造三元组数据输入网络，然后反向传播三元组的度量损失来进行迭代优化。每一个三元组包含三张图像，分别是一个参考样本，一个与参考样本同类的正样本，以及一个与参考样本异类的负样本，标记为 (anchor, positive, negative)。原始 triplet loss 的基本思想是，通过度量学习使得参考样本与正样本之间的距离小于参考样本与负样本之间的距离，并且距离之差大于一个最小间隔参数 α 。因此原始的 triplet loss 损失函数如下：

$$L = \sum_i^N \left[\left\| f(x_i^a) - f(x_i^p) \right\|_2^2 - \left\| f(x_i^a) - f(x_i^n) \right\|_2^2 + \alpha \right]_+$$

其中，N 是三元组数量， $f(x_i^a)$ 表示参考样本 (anchor) 的特征向量， $f(x_i^p)$ 表示同类正样本 (positive) 的特征向量， $f(x_i^n)$ 表示异类负样本 (negative) 的特征向量。 $\|\cdot\|_2$ 表示 L2 范数，即欧氏距离。 $[\cdot]_+$ 的含义如下：
$$[x]_+ = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

从上式可看出，原始的 triplet loss 函数只限定了同类样本(anchor, positive)与异类样本(anchor, negative)之间的距离，即通过间隔参数 α 尽可能增大类间距离，而对类内距离未作任何限定，即对同类样本之间的距离未作任何约束。如果类内距离比较分散，方差过大，网络的泛化能力就会减弱，样本被错分的概率就会更大。图 4 为在类间间隔一致、类内方差较大情况下，样本错分的概率示意图，图 5 为在类间间隔一致、类内方差较小情况下，样本错分的概率示意图，如图 4 和图 5 所示，阴影部分表示样本错分的概率，在类间间隔一致、类内方差较大情况下，样本错分的概率明显大于类间间隔一致、类内方差较小情况下，样本错分的概率。

针对上述问题, 本发明提出改进的 triplet loss 方法, 一方面保留了原始方法中对类间距离的限定, 同时增加了对类内距离的约束项, 使得类内距离尽可能聚敛。其 loss 函数表达式为:

$$\tilde{L} = \sum_i^N [\cos(f(x_i^a), f(x_i^n)) - \cos(f(x_i^a), f(x_i^p)) + \alpha_1]_+ + \sum_i^N [\alpha_2 - \cos(f(x_i^a), f(x_i^p))]_+$$

其中, $\cos(\cdot)$ 表示余弦距离, 其计算方式为 $\cos(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$, N 是三元组数量, $f(x_i^a)$ 表示参考样本的特征向量, $f(x_i^p)$ 表示同类正样本的特征向量, $f(x_i^n)$ 表示异类负样本的特征向量, $[\cdot]_+$ 的含义如下: $[x]_+ = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$, α_1 为类间间隔参数, α_2 为类内间隔参数。

相比原始的 triplet loss 函数, 改进后的 triplet loss 函数的度量方式由欧氏距离改为余弦距离, 这样可以保持训练阶段与部署阶段度量方式的一致性, 提高特征学习的连续性。同时新的 triplet loss 函数第一项与原始的 triplet loss 作用一致, 用于增大类间差距, 第二项增加了对同类样本对(正元组)的距离约束, 用于缩小类内差距。 α_1 为类间间隔参数, 取值范围为 0~0.2, α_2 为类内间隔参数, 取值范围为 0.8~1.0。值得注意的是, 由于是用余弦方式度量, 得到的度量值对应两个样本之间的相似度, 因此在 $\tilde{L}$ 表达式中, 只有负元组余弦相似度在 α_1 范围内大于正元组余弦相似度的样本, 才会真正参与训练。

基于改进后的三元组损失函数来训练模型, 通过类间损失与类内损失的联合约束来对模型进行反向传播的优化训练, 使得同类样本在特征空间尽可能接近而异类样本在特征空间尽可能远离, 提高模型的辨识力, 从而提高人脸认证的可靠性。

S308, 将验证集数据输入卷积神经网络, 达到训练结束条件时, 得到训练好的用于人脸认证的卷积神经网络。

具体地, 从人证图像数据池中取 90% 数据作为训练集, 剩余 10% 作为验证集。基于上式计算出改进后的 triplet loss 值, 反馈到卷积神经网络中进行迭代优化。同时观测模型在验证集中的性能表现, 当验证性能不再升高时, 模型达到收敛状态, 训练阶段终止。

上述的人脸认证方法, 一方面在原始 triplet loss 的损失函数中增加了对类内样本距离的约束, 从而在增大类间差距的同时减小类内差距, 提升模型的泛化能力; 另一方面, 将原始 triplet loss 的度量方式由欧氏距离改为余弦距离, 保持训练与部署的度量一致性, 提高特征学

习的连续性。

在另一个实施例中，训练卷积神经网络的步骤还包括：利用基于海量开源人脸数据训练好的基础模型参数进行初始化，在特征输出层后添加归一化层及改进后的三元组损失函数层，得到待训练的卷积神经网络。

具体地，在用深度学习解决人证合一问题时，常规的基于互联网海量人脸数据训练得到的深度人脸识别模型在特定场景下的人证比对应用上性能会大幅下降，而特定应用场景下的人证数据来源又比较有限，直接地学习往往由于样本不足导致训练结果不理想，因此极需要研发一种有效地针对小数据集的场景数据进行扩展训练的方法，以提升人脸识别模型在特定应用场景下的准确率，满足市场应用需求。

深度学习算法往往依赖于海量数据的训练，在人证合一应用中，证件照与场景照比对属于异质样本比对问题，常规的基于海量互联网人脸数据训练得到的深度人脸识别模型在人证比对应用上性能会大幅下降。然而人证数据来源有限（需要同时具备同一个人的身份证图像及相应的场景图像），可用于训练的数据量较少，直接训练会由于样本不足导致训练效果不理想，因此在运用深度学习进行人证合一的模型训练时，往往是利用迁移学习的思想，先基于海量的互联网人脸数据训练一个在开源测试集上性能可靠的基础模型，然后再在有限的人证数据上进行二次扩展训练，使模型能自动学习特定模态的特征表示，提升模型性能。此过程如图 6 所示。

在二次训练的过程中，整个网络用预训练好的基础模型参数进行初始化，然后在网络的特征输出层之后添加一个 L2 归一化层以及改进后的 triplet loss 层，待训练的卷积神经网络结构图如图 7 所示。

一个实施例中，一种人脸认证方法的流程示意图如图 8 所示，包括三个阶段，分别为数据采集与预处理阶段、训练阶段和部署阶段。

数据采集与预处理阶段，由人证核验终端设备的读卡器模块读取证件芯片照，以及前置摄像头抓取现场照片，经过人脸检测器、关键点检测器、人脸对齐与剪切模块之后得到尺寸归一化的证件人脸图像和场景人脸图像。

训练阶段，从人证图像数据池中取 90% 数据作为训练集，剩余 10% 作为验证集。由于人证比对主要是证件照与场景照之间的比对，因为三元组中若以证件照为参考图（anchor），则另外两张图均是场景照；反之若以场景照为参考图，则另外两张图均是证件照。采用 OHEM

在线构造三元组的策略，即在网络每次迭代优化的过程中，利用当前网络对候选三元组进行前向计算，筛选满足条件的有效三元组，按照上式计算出改进后的 triplet loss 值，反馈到网络中进行迭代优化。同时观测模型在验证集中的性能表现，当验证性能不再升高时，模型达到收敛状态，训练阶段终止。

部署阶段，将训练好的模型部署到人证核验终端进行使用时，设备采集到的图像经过与训练阶段相同的预处理程序，然后通过网络前向计算得到每张人脸图像的特征向量，通过计算余弦距离得到两张图像的相似度，然后根据预设阈值进行判决，大于预设阈值的为同一人，反之为不同人。

上述的人脸认证方法，原始 triplet loss 函数只限制了类间距离的学习关系，上述的人脸认证方法，通过改进原始 triplet loss 损失函数增加了类内距离的约束项，可以使得网络在训练过程中增大类间差距的同时尽可能减小类内差距，从而提高网络的泛化能力，进而提升模型的场景适应性。另外，用余弦距离替代了原始 triplet loss 中的欧氏距离度量方式，更符合人脸特征空间的分布属性，保持了训练阶段与部署阶段度量方式的一致性，使得比对结果更加可靠。

在一个实施例中，提供一种人脸认证装置，如图 9 所示，包括：图像获取模块 902、图像预处理模块 904、特征获取模块 906、计算模块 908 和认证模块 910。

图像获取模块 902，用于基于人脸认证请求，获取证件照片和人物的场景照片。

图像预处理模块 904，用于对场景照片和证件照片分别进行人脸检测、关键点定位和图像预处理，得到场景照片对应的场景人脸图像，以及证件照片对应的证件人脸图像。

特征获取模块 906，用于将场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型，并获取卷积神经网络模型输出的场景人脸图像对应的第一特征向量，以及证件人脸图像对应的第二特征向量；其中，卷积神经网络模型基于三元组损失函数的监督训练得到。

计算模块 908，用于计算第一特征向量和第二特征向量的余弦距离。

认证模块 910，用于比较余弦距离和预设阈值，并根据比较结果确定人脸认证结果。

上述的人脸认证装置，利用预先训练的卷积神经网络进行人脸认证，由于卷积神经网络模型基于改进后的三元组损失函数的监督训练得到，而场景人脸图像和证件人脸图像的相似度根据场景人脸图像对应的第一特征向量和证件人脸图像对应的第二特征向量的余弦距离计

算得到，余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，而不是位置，从而更符合人脸特征空间的分布属性，提高了人脸认证的可靠性。

如图 9 所示，在另一个实施例中，人脸认证装置还包括：样本获取模块 912、三元组获取模块 914、训练模块 916 和验证模块 918。

样本获取模块 912，用于获取带标记的训练样本，所述训练样本包括标记了属于每个标记对象的一张证件人脸图像和至少一张场景人脸图像。

三元组获取模块 914，用于根据训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素；三元组元素包括参考样本、正样本和负样本。

具体地，三元组获取模块 914，用于随机选择一个图像作为参考样本，选择属于同一标签对象、与参考样本类别不同的图像作为正样本，还用于根据 OHEM 策略，利用当前训练的卷积神经网络模型提取特征之间的余弦距离，对于每一个参考样本，从其它具有不属于同一标签对象的人脸图像中，选择距离最小、与参考样本属于不同类别的图像，作为该参考样本的负样本。

具体地，以证件照作为参考样本时，正样本和负样本均为场景照；以场景照作为参考样本时，正样本和负样本均为证件照。

训练模块 916，用于根据各训练样本的三元组元素，基于三元组损失函数的监督，训练卷积神经网络模型，该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数。

具体地，改进型三元组损失函数包括对同类样本的余弦距离的限定，以及对异类样本的余弦距离的限定。

改进型三元组损失函数为：

$$\tilde{L} = \sum_i^N \left[\cos(f(x_i^a), f(x_i^n)) - \cos(f(x_i^a), f(x_i^p)) + \alpha_1 \right]_+ + \sum_i^N \left[\alpha_2 - \cos(f(x_i^a), f(x_i^p)) \right]_+$$

其中， $\cos(\bullet)$ 表示余弦距离，其计算方式为 $\cos(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$ ，N 是三元组数量， $f(x_i^a)$ 表示参考样本的特征向量， $f(x_i^p)$ 表示同类正样本的特征向量， $f(x_i^n)$ 表示异类负样本的特征向量， $[\cdot]_+$ 的含义如下：

$[x]_+ = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$ ， α_1 为类间间隔参数， α_2 为类内间隔参数。

验证模块 918，用于将验证集数据输入卷积神经网络模型，达到训练结束条件时，得到训练好的用于人脸认证的卷积神经网络模型。

在另一个实施例中，人脸认证装置还包括模型初始化模块 920，用于利用基于海量开源人脸数据训练好的基础模型参数进行初始化，在特征输出层后添加归一化层及三元组损失函数层，得到待训练的卷积神经网络。上述的人脸认证装置，一方面在原始 triplet loss 的损失函数中增加了对类内样本距离的约束，从而在增大类间差距的同时减小类内差距，提升模型的泛化能力；另一方面，将原始 triplet loss 的度量方式由欧氏距离改为余弦距离，保持训练与部署的度量一致性，提高特征学习的连续性。

一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，处理器执行计算机程序时实现上述各实施例的人脸认证方法的步骤。

一种存储介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时，实现上述各实施例的人脸认证方法的步骤。

以上所述实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本发明的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明专利的保护范围应以所附权利要求为准。

权利要求书

1、一种基于 Triplet Loss 的人脸认证方法，包括：

基于人脸认证请求，获取证件照片和人物的场景照片；

对所述场景照片和所述证件照片分别进行人脸检测、关键点定位和图像预处理，得到所述场景照片对应的场景人脸图像，以及所述证件照片对应的证件人脸图像；

将所述场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型，并获取所述卷积神经网络模型输出的所述场景人脸图像对应的第一特征向量，以及所述证件人脸图像对应的第二特征向量；其中，所述卷积神经网络模型基于三元组损失函数的监督训练得到；

计算所述第一特征向量和所述第二特征向量的余弦距离；

比较所述余弦距离和预设阈值，并根据比较结果确定人脸认证结果。

2、根据权利要求 1 所述的方法，其特征在于，所述方法还包括：

获取带标记的训练样本，所述训练样本包括标记了属于每个标记对象的一张证件人脸图像和至少一张场景人脸图像；

根据所述训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素；所述三元组元素包括参考样本、正样本和负样本；

根据各训练样本的三元组元素，基于三元组损失函数的监督，训练所述卷积神经网络模型；该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数；

将验证集数据输入所述卷积神经网络模型，达到训练结束条件时，得到训练好的用于人脸认证的卷积神经网络模型。

3、根据权利要求 2 所述的方法，其特征在于，根据所述训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素的步骤，包括：

随机选择一个图像作为参考样本，选择属于同一标签对象、与参考样本类别不同的图像作为正样本；

根据 OHEM 策略，利用当前训练的卷积神经网络模型提取特征之间的余弦距离，对于每一个参考样本，从其它不属于同一标签对象的图像中，选择距离最小、与所述参考样本属于不同类别的图像，作为该参考样本的负样本。

4、根据权利要求2所述的方法，其特征在于，所述三元组损失函数包括对同类样本的余弦距离的限定，以及对异类样本的余弦距离的限定。

5、根据权利要求4所述的方法，其特征在于，所述三元组损失函数为：

$$\tilde{L} = \sum_i^N [\cos(f(x_i^a), f(x_i^n)) - \cos(f(x_i^a), f(x_i^p)) + \alpha_1]_+ + \sum_i^N [\alpha_2 - \cos(f(x_i^a), f(x_i^p))]_+$$

其中， $\cos(\bullet)$ 表示余弦距离，其计算方式为 $\cos(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$ ， N 是三元组数量， $f(x_i^a)$ 表示参考样本的特征向量， $f(x_i^p)$ 表示同类正样本的特征向量， $f(x_i^n)$ 表示异类负样本的特征向量， $[\bullet]_+$ 的含义如下： $[x]_+ = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$ ， α_1 为类间间隔参数， α_2 为类内间隔参数。

6、根据权利要求2所述的方法，其特征在于，所述方法还包括：利用基于海量开源人脸数据训练好的基础模型参数进行初始化，在特征输出层后添加归一化层及三元组损失函数层，得到待训练的卷积神经网络模型。

7、一种基于 Triplet Loss 的人脸认证装置，包括：图像获取模块、图像预处理模块、特征获取模块、计算模块和认证模块；

所述图像获取模块，用于基于人脸认证请求，获取证件照片和人物的场景照片；

所述图像预处理模块，用于对所述场景照片和所述证件照片分别进行人脸检测、关键点定位和图像预处理，得到所述场景照片对应的场景人脸图像，以及所述证件照片对应的证件人脸图像；

所述特征获取模块，用于将所述场景人脸图像和证件人脸图像输入到预先训练好的用于人脸认证的卷积神经网络模型，并获取所述卷积神经网络模型输出的所述场景人脸图像对应的第一特征向量，以及所述证件人脸图像对应的第二特征向量；其中，所述卷积神经网络模型基于三元组损失函数的监督训练得到；

所述计算模块，用于计算所述第一特征向量和所述第二特征向量的余弦距离；

所述认证模块，用于比较所述余弦距离和预设阈值，并根据比较结果确定人脸认证结果。

8、根据权利要求7所述的装置，其特征在于，所述装置还包括：样本获取模块、三元组获取模块、训练模块和验证模块；

所述样本获取模块，用于获取带标记的训练样本，所述训练样本包括标记了属于每个标

记对象的一张证件人脸图像和至少一张场景人脸图像；

所述三元组获取模块，用于根据所述训练样本训练卷积神经网络模型，通过 OHEM 产生各训练样本对应的三元组元素；所述三元组元素包括参考样本、正样本和负样本；

所述训练模块，用于根据各训练样本的三元组元素，基于三元组损失函数的监督，训练所述卷积神经网络模型；该三元组损失函数，以余弦距离作为度量方式，通过随机梯度下降算法来优化模型参数；

所述验证模块，用于将验证集数据输入所述卷积神经网络模型，达到训练结束条件时，得到训练好的用于人脸认证的卷积神经网络模型。

9、一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现权利要求 1 至 6 任一项所述的基于 Triplet Loss 的人脸认证方法的步骤。

10、一种存储介质，其上存储有计算机程序，其特征在于，该计算机程序被处理器执行时，实现权利要求 1 至 6 任一项所述的基于 Triplet Loss 的人脸认证方法的步骤。

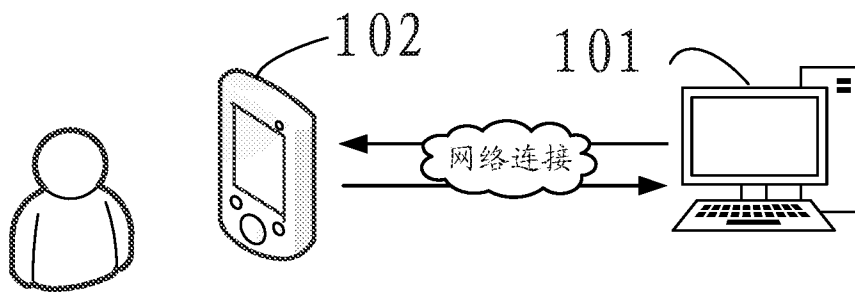


图 1

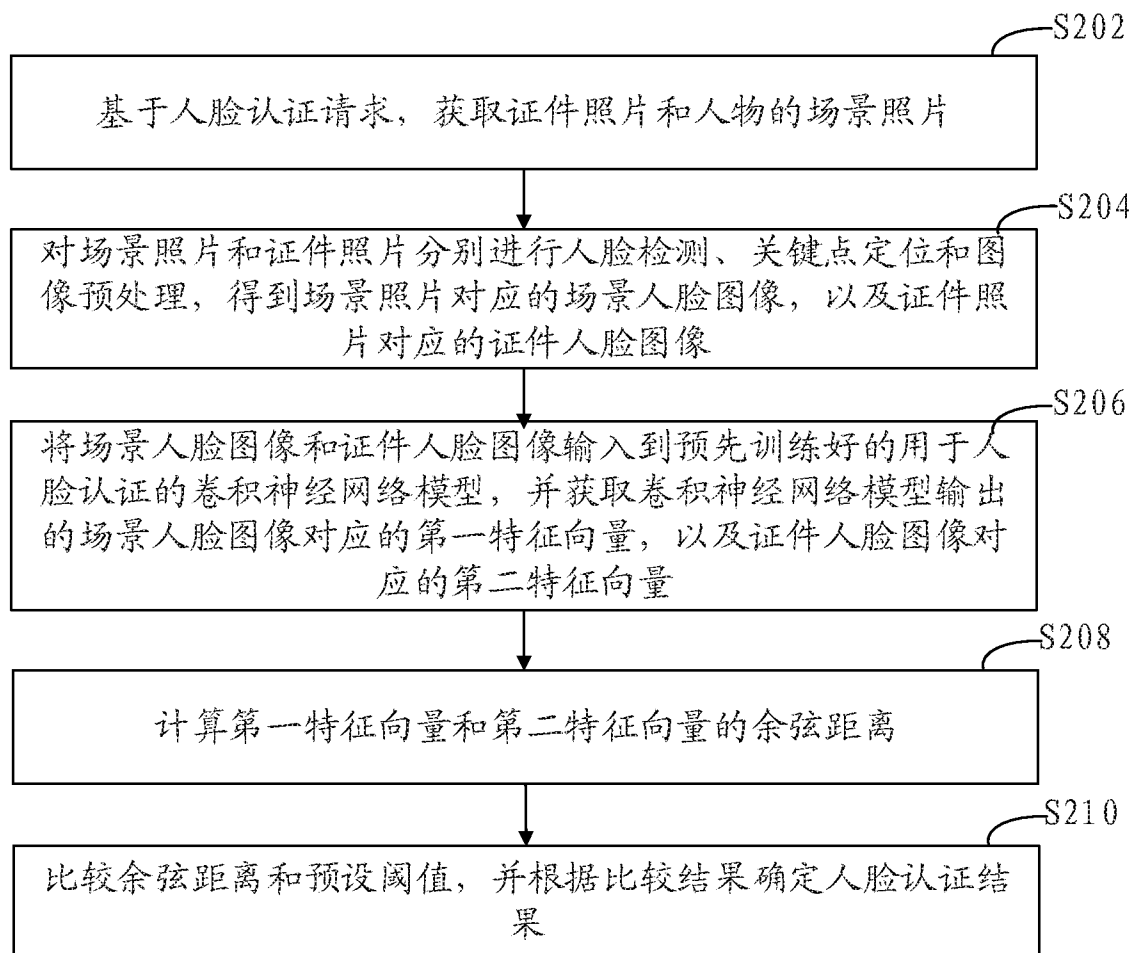


图 2

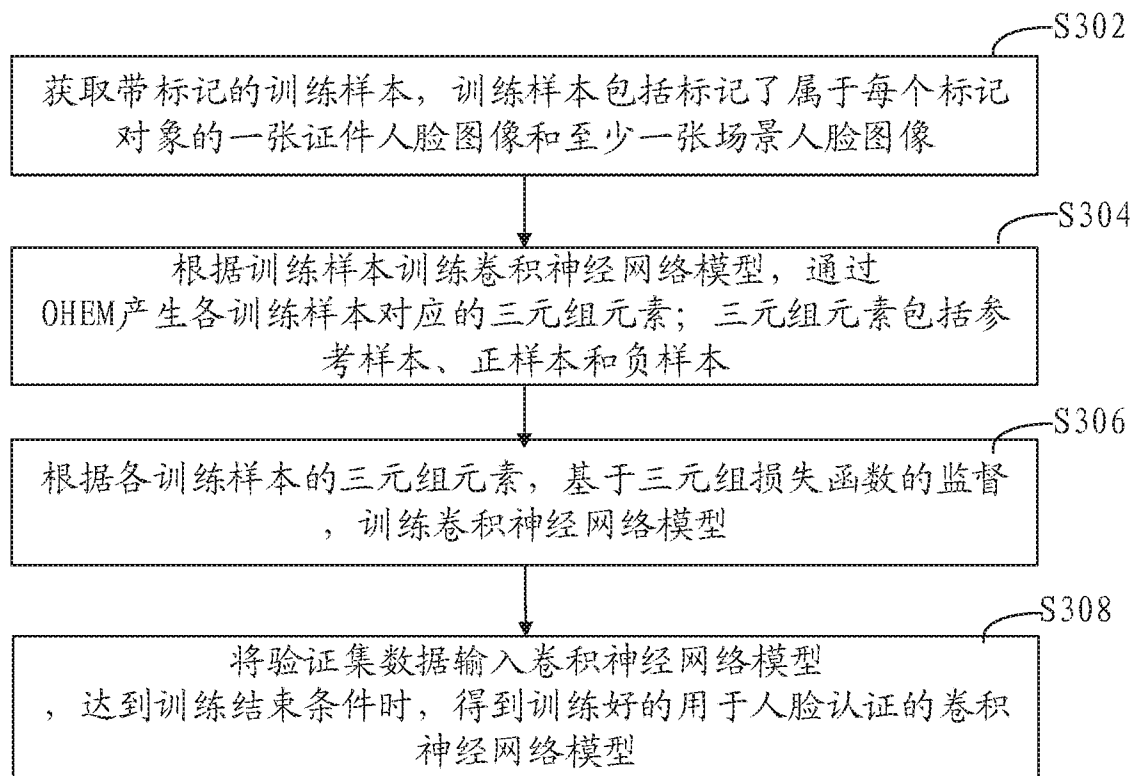


图 3

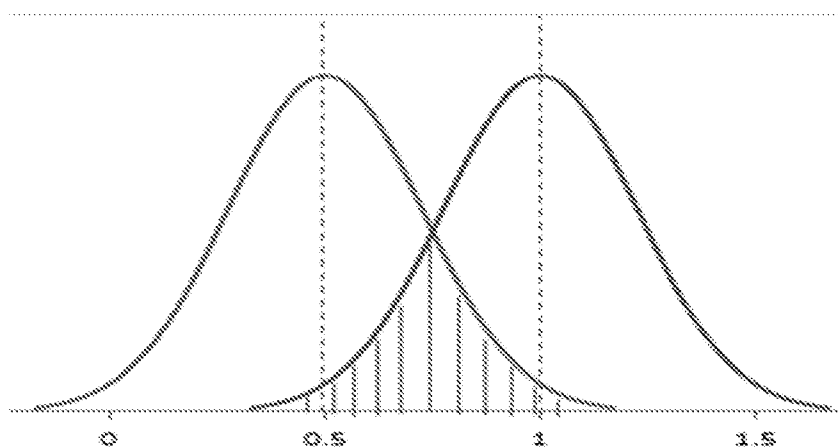


图 4

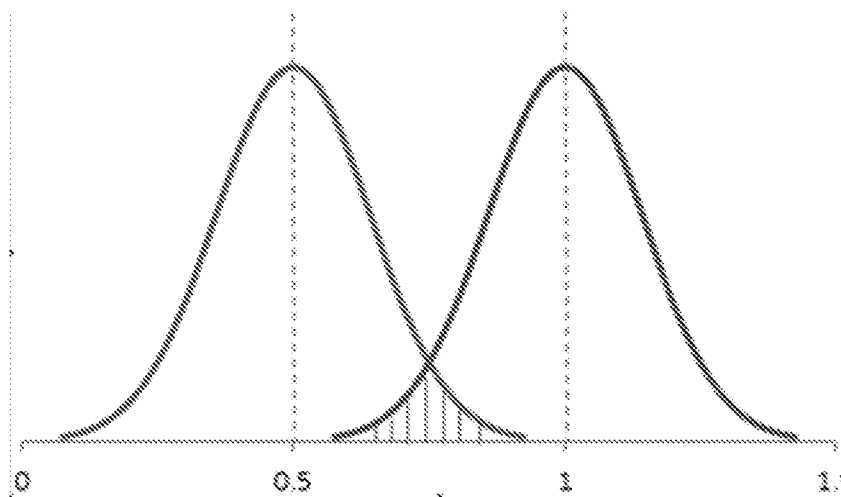


图 5

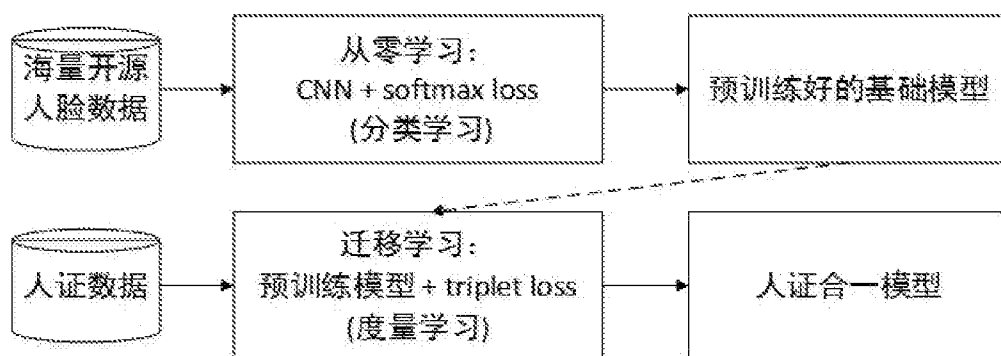


图 6

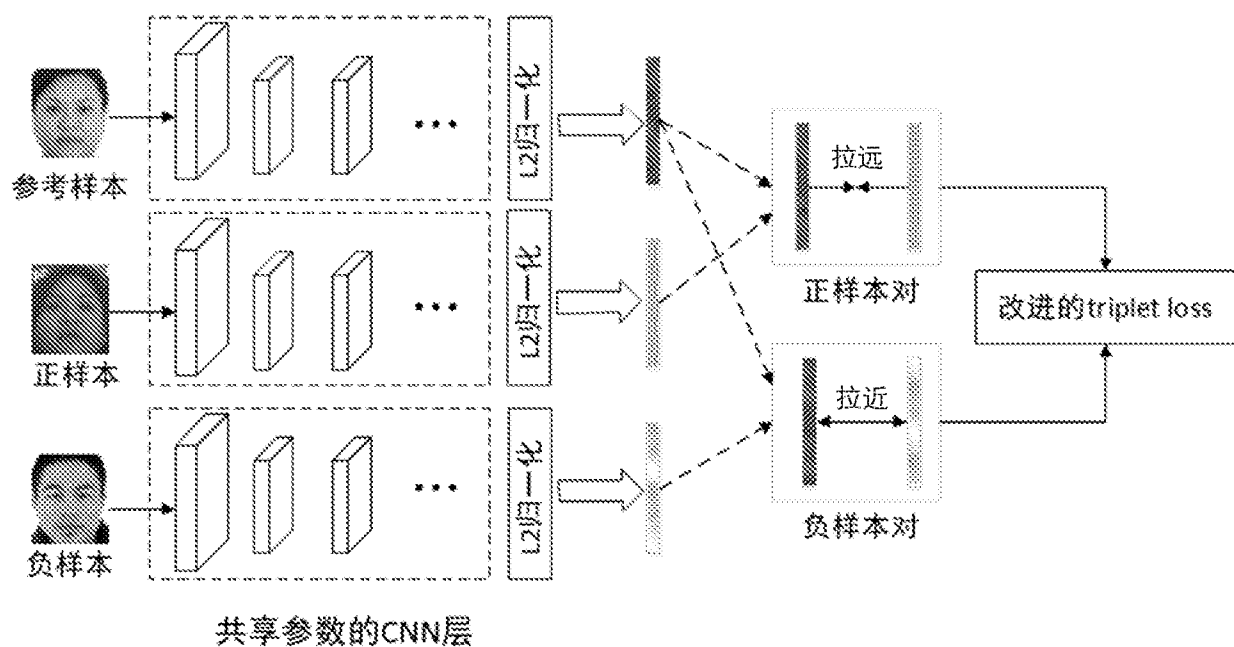


图 7

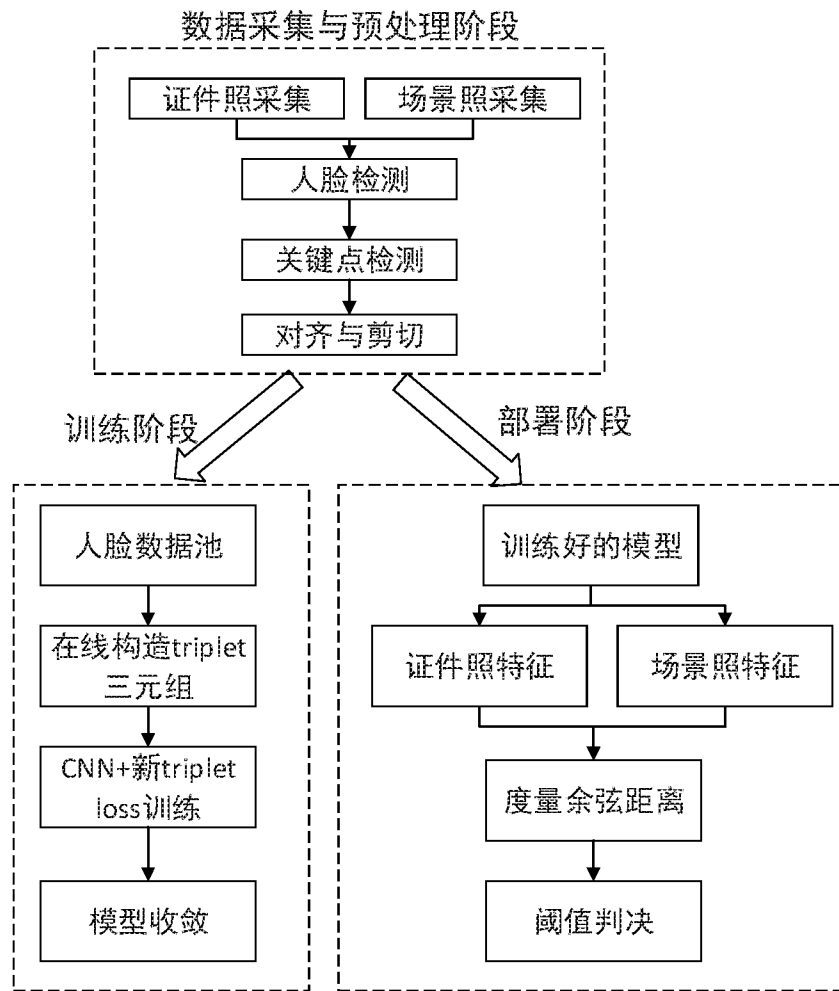


图 8

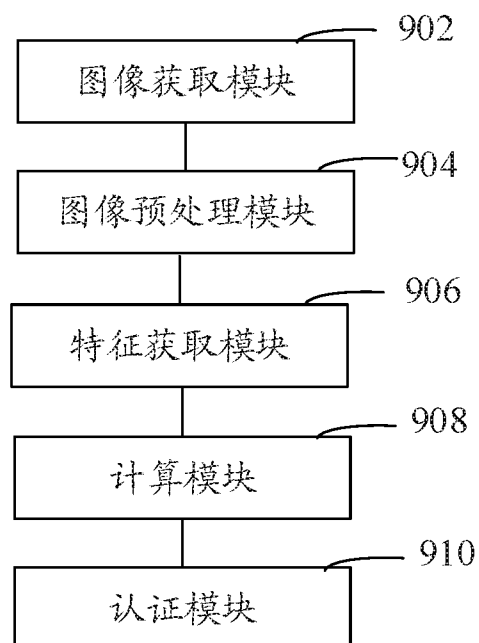


图 9

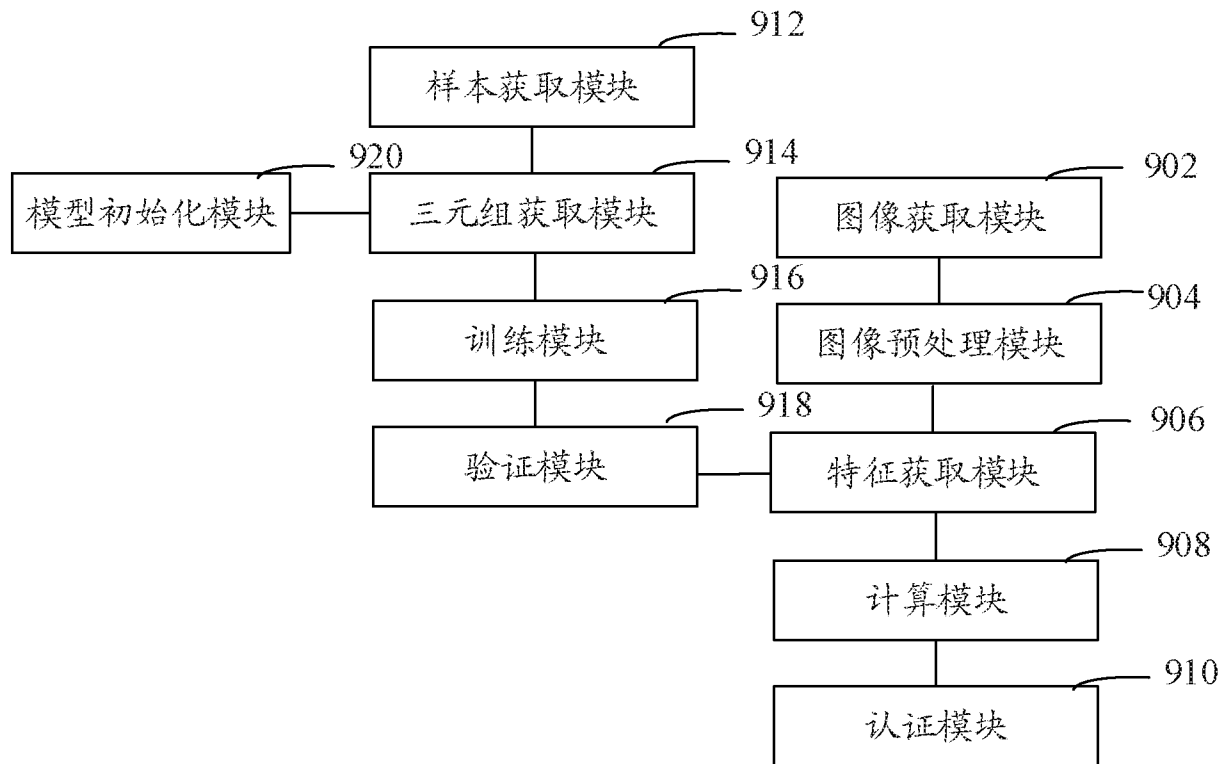


图 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/109169

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 9/-

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)

USTXT; EPTXT; CNTXT; CNABS; WOTXT; VEN; CNKI: 人脸, 认证, Triplet 损失函数, 特征向量, 图像, 余弦距离, 证件, 照片, 训练, 模型, 阈值, 样本, face, recognition, Triplet loss function, feature vector, cosine distance, identification, photo, train, model, threshold, sample

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 108009528 A (GRG BANKING EQUIPMENT CO., LTD.) 08 May 2018 (08.05.2018), entire document	1-10
Y	CN 107194341 A (XIDIAN UNIVERSITY) 22 September 2017 (22.09.2017), description, paragraphs [0033]-[0039] and [0059]-[0072]	1-10
Y	BRANDON, Amos et al. OpenFace: A General-purpose Face Recognition Library with Mobile Applications. CMU School of Computer Science, Tech. Rep., 2016. 30 June 2016 (30.06.2016), pp. 16 and 17	1-10
Y	CN 106599827 A (ZHEJIANG GONGSHANG UNIVERSITY) 26 April 2017 (26.04.2017), description, paragraphs [0054] and [0055]	2-6, 8

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“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	“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
“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)	“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
“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	“&” document member of the same patent family

Date of the actual completion of the international search 12 November 2018	Date of mailing of the international search report 25 December 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer XIA, Xue Telephone No. (86-20) 28950718

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/109169

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107423690 A (GUANGDONG UNIVERSITY OF TECHNOLOGY) 01 December 2017 (01.12.2017), entire document	1-10
A	US 9129216 B1 (XDROID KFT.) 08 September 2015 (08.09.2015), entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/109169

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 108009528 A	08 May 2018	None	
CN 107194341 A	22 September 2017	None	
CN 106599827 A	26 April 2017	None	
CN 107423690 A	01 December 2017	None	
US 9129216 B1	08 September 2015	None	

国际检索报告

国际申请号

PCT/CN2018/109169

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) G06K9/- 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)) USTXT;EPTXT;CNTXT;CNABS;WOTXT;VEN;CNKI:人脸, 认证, Triplet损失函数, 特征向量, 图像, 余弦距离, 证件, 照片, 训练, 模型, 阈值, 样本, face, recognition, Triplet loss function, feature vector, cosine distance, identification, photo, train, model, threshold, sample																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 108009528 A (广州广电运通金融电子股份有限公司) 2018年 5月 8日 (2018 - 05 - 08) 全文</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 107194341 A (西安电子科技大学) 2017年 9月 22日 (2017 - 09 - 22) 说明书第33-39、59-72段</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>BRANDON, Amos等. "OpenFace:A general-purpose face recognition library with mobile applications" CMU School of Computer Science, Tech. Rep., 2016, 2016年 6月 30日 (2016 - 06 - 30), 第16-17页</td> <td>1-10</td> </tr> <tr> <td>Y</td> <td>CN 106599827 A (浙江工商大学) 2017年 4月 26日 (2017 - 04 - 26) 说明书第54-55段</td> <td>2-6、8</td> </tr> <tr> <td>A</td> <td>CN 107423690 A (广东工业大学) 2017年 12月 1日 (2017 - 12 - 01) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 9129216 B1 (XDROID KFT) 2015年 9月 8日 (2015 - 09 - 08) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 108009528 A (广州广电运通金融电子股份有限公司) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-10	Y	CN 107194341 A (西安电子科技大学) 2017年 9月 22日 (2017 - 09 - 22) 说明书第33-39、59-72段	1-10	Y	BRANDON, Amos等. "OpenFace:A general-purpose face recognition library with mobile applications" CMU School of Computer Science, Tech. Rep., 2016, 2016年 6月 30日 (2016 - 06 - 30), 第16-17页	1-10	Y	CN 106599827 A (浙江工商大学) 2017年 4月 26日 (2017 - 04 - 26) 说明书第54-55段	2-6、8	A	CN 107423690 A (广东工业大学) 2017年 12月 1日 (2017 - 12 - 01) 全文	1-10	A	US 9129216 B1 (XDROID KFT) 2015年 9月 8日 (2015 - 09 - 08) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
PX	CN 108009528 A (广州广电运通金融电子股份有限公司) 2018年 5月 8日 (2018 - 05 - 08) 全文	1-10																					
Y	CN 107194341 A (西安电子科技大学) 2017年 9月 22日 (2017 - 09 - 22) 说明书第33-39、59-72段	1-10																					
Y	BRANDON, Amos等. "OpenFace:A general-purpose face recognition library with mobile applications" CMU School of Computer Science, Tech. Rep., 2016, 2016年 6月 30日 (2016 - 06 - 30), 第16-17页	1-10																					
Y	CN 106599827 A (浙江工商大学) 2017年 4月 26日 (2017 - 04 - 26) 说明书第54-55段	2-6、8																					
A	CN 107423690 A (广东工业大学) 2017年 12月 1日 (2017 - 12 - 01) 全文	1-10																					
A	US 9129216 B1 (XDROID KFT) 2015年 9月 8日 (2015 - 09 - 08) 全文	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 2018年 11月 12日		国际检索报告邮寄日期 2018年 12月 25日																					
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (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 夏雪 电话号码 86-(20)-28950718																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/109169

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	108009528	A	2018年 5月 8日	无	
CN	107194341	A	2017年 9月 22日	无	
CN	106599827	A	2017年 4月 26日	无	
CN	107423690	A	2017年 12月 1日	无	
US	9129216	B1	2015年 9月 8日	无	