

手掌图像处理方法、装置、设备、存储介质及程序产品

相关申请

本申请要求 2022 年 11 月 21 日申请的，申请号为 2022114733356，名称为“掌部图像的识别方法、装置、设备、存储介质及程序产品”的中国专利申请的优先权，在此将其全文引入作为参考。

技术领域

本申请实施例涉及计算机技术领域，进一步涉及计算机视觉技术领域，特别涉及一种手掌图像处理方法、装置、设备、存储介质及程序产品。

背景技术

随着计算机技术的发展，手掌识别技术应用越来越广泛，可以应用于多种场景下。例如，支付场景或上班打卡场景等，通过手掌识别，可以对用户身份进行验证。

相关技术中，用户提供实体手掌，手掌识别设备对用户掌部采集手掌图像，从而基于手掌图像识别设备将手掌图像进行加密，加密后的手掌图像用于身份验证。

然而，对手掌图像进行加密、解密需要消耗的时间较多，降低了手掌识别的效率。

发明内容

根据本申请的各种实施例，提供一种手掌图像处理方法、装置、设备、存储介质及程序产品。

一方面，提供了一种手掌图像处理方法，由手掌识别设备执行，包括：

获取手掌图像，对所述手掌图像提取特征，得到多个尺度的图像特征；

将所述多个尺度的图像特征融合，得到图像融合特征；

基于所述图像融合特征，确定标识所述手掌图像中掌部的掌部框；

基于所述掌部框，从所述手掌图像中裁剪出掌部图像；及

将所述掌部图像进行图像加密，得到掌部图像加密数据，所述掌部图像加密数据用于传输至服务器，由所述服务器基于所述掌部图像加密数据进行身份识别。

另一方面，本申请提供一种掌部检测模型的训练方法，由计算机设备执行，包括：

获取样本手掌图像和样本掌部图像；

对所述样本手掌图像提取特征，得到多个尺度的样本图像特征；

将所述多个尺度的样本图像特征融合，得到样本图像融合特征；

基于所述样本图像融合特征，确定所述标识样本手掌图像中掌部的样本预测掌部框；

基于所述样本预测掌部框，从所述样本手掌图像中裁剪出样本预测掌部图像；

根据所述样本掌部图像及所述样本预测掌部图像，计算损失函数值；及

基于所述损失函数值对所述掌部检测模型的模型参数进行更新。

另一方面，本申请提供一种手掌图像处理装置，包括：

特征提取模块，用于获取手掌图像，对所述手掌图像提取特征，得到多个尺度的图像特征；

特征融合模块，用于将所述多个尺度的图像特征融合，得到图像融合特征；

预测模块，用于基于所述图像融合特征，确定标识所述手掌图像中掌部的掌部框；

裁剪模块，用于基于所述掌部框，从所述手掌图像中裁剪出掌部图像；及

加密模块，用于将所述掌部图像进行图像加密，得到掌部图像加密数据，所述掌部图像加密数据用于传输至服务器，由所述服务器基于所述掌部图像加密数据进行身份验证。

另一方面，本申请提供一种掌部检测模型的训练装置，包括掌部检测模型，还包括获取模块、裁剪模块、计算模块和更新模块，所述掌部检测模型包括特征提取模块、特征融合模块和预测模块；

所述获取模块，用于获取样本手掌图像和样本掌部图像；

所述特征提取模块，用于对所述样本手掌图像提取特征，得到多个尺度的样本图像特

征；

所述特征融合模块，用于将所述多个尺度的样本图像特征融合，得到样本图像融合特征；

所述预测模块，用于基于所述样本图像融合特征，确定标识所述样本手掌图像中掌部的样本预测掌部框；

所述裁剪模块，用于基于所述样本预测掌部框，从所述样本手掌图像中裁剪出样本预测掌部图像；

所述计算模块，用于根据所述样本掌部图像及所述样本预测掌部图像，计算损失函数值；

所述更新模块，用于基于所述损失函数值对掌部检测模型的模型参数进行更新。

另一方面，本申请还提供了一种计算机设备，包括存储器和处理器，所述存储器存储有计算机可读指令，所述处理器执行所述计算机可读指令时执行本申请各方法实施例的步骤。

另一方面，本申请还提供了一种计算机可读存储介质，存储有计算机可读指令，所述计算机可读指令被处理器执行时执行本申请各方法实施例的步骤。

另一方面，本申请还提供了一种计算机程序产品，包括计算机可读指令，所述计算机可读指令被处理器执行时执行本申请各方法实施例的步骤。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其它特征、目的和优点将从说明书、附图以及权利要求书变得明显。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1是本申请一个示例性实施例提供的一种手掌图像处理方法的示意图；

图2是本申请一个示例性实施例提供的计算机系统的架构示意图；

图3是本申请一个示例性实施例提供的手掌图像处理方法的流程图；

图4是本申请一个示例性实施例提供的手掌图像处理方法的流程图；

图5是本申请一个示例性实施例提供的图像特征融合的示意图；

图6是本申请一个示例性实施例提供的手掌中手指缝点的示意图；

图7是本申请一个示例性实施例提供的手掌图像处理方法的示意图；

图8是本申请一个示例性实施例提供的基于手掌图像处理方法应用于跨设备支付场景的流程示意图；

图9是本申请一个示例性实施例提供的基于手掌图像处理方法应用于跨设备身份验证场景的流程示意图；

图10是本申请一个示例性实施例提供的手掌图像处理方法的流程示意图；

图11是本申请一个示例性实施例提供的掌部检测模型生成阶段和掌部框的预测阶段的框架图；

图12是本申请一个示例性实施例提供的掌部检测模型的训练方法的流程图；

图13是本申请一个示例性实施例提供的掌部检测模型的训练方法的流程图；

图14是本申请一个示例性实施例提供的手掌图像处理装置的框图；

图15是本申请一个示例性实施例提供的掌部检测模型的训练装置的框图；

图16是本申请一个示例性实施例提供的计算机设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本

申请中的实施例，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，都属于本申请保护的范围。

计算机视觉技术（Computer Vision, CV）是一门研究如何使机器“看”的科学，更进一步的讲，就是指用摄影机和电脑代替人眼对目标进行识别和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科，计算机视觉研究相关的理论和技术，试图建立能够从图像或者多维数据中获取信息的人工手掌图像识别系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、视频处理、视频语义理解、视频内容/行为识别、三维物体重建、3D 技术、虚拟现实、增强现实、同步定位与地图构建等技术，还包括常见的生物特征识别技术。

本申请实施例提供了一种手掌图像处理方法的示意图，如图 1 所示，该方法可以由手掌识别设备 10 执行，该手掌识别设备 10 可以具有图像加密功能。

示例性地，手掌识别设备 10 包括摄像头，手掌识别设备 10 通过摄像头获取手掌图像 11。手掌识别设备 10 也可以连接图像采集器，图像采集器具有摄像头，手掌识别设备 10 可获取由该图像采集器采集的手掌图像 11。手掌识别设备 10 对手掌图像 11 进行掌部预测，得到手掌图像 11 中掌部对应的掌部图像 12。手掌识别设备 10 将掌部图像 12 进行图像加密，得到掌部图像加密数据 13。手掌识别设备 10 将掌部图像加密数据 13 发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像 11 对应的用户标识 14。

示例性地，手掌识别设备 10 获取手掌图像 11，并对手掌图像 11 提取特征，得到多个尺度的图像特征。手掌识别设备 10 将多个尺度的图像特征进行特征融合，得到图像融合特征。手掌识别设备 10 基于图像融合特征，确定手掌图像 11 中的掌部框。手掌识别设备 10 基于掌部框，从手掌图像 11 中裁剪出掌部图像 12。手掌识别设备 10 将掌部图像 12 进行图像加密，得到掌部图像加密数据 13。

掌部图像加密数据 13 用于被发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像 11 对应的用户标识 14。

掌部图像 12 是手掌图像 11 中的有效识别区域，或，掌部图像 12 是指手掌图像 11 中掌部所在的区域，或，掌部图像 12 是指手掌图像 11 中能够用于手掌识别的区域。

可选地，手掌识别设备 10 将掌部图像加密数据 13 发送至手掌识别服务器的方式包括网络传输、数据线传输、蓝牙传输中的至少一种，但不限于此，本申请实施了对此不作具体限定。

示例性地，手掌识别设备 10 对手掌图像 11 提取特征，得到图像特征。手掌识别设备 10 将图像特征进行特征融合，得到图像融合特征。手掌识别设备 10 确定手掌图像中的掌部框，基于掌部框，从手掌图像中裁剪出掌部图像 12。

示例性地，掌部框可以由掌部检测模型预测得到，掌部检测模型包括主干网络、颈部网络和预测网络。

手掌识别设备 10 将手掌图像 11 输入至主干网络；由主干网络对手掌图像进行切片操作，得到多个尺度的切片图像；手掌识别设备 10 对多个尺度的切片图像进行特征提取，得到手掌图像 11 对应的多个尺度的图像特征。尺度表示特征的尺寸大小，多个尺度是指多个尺寸。

例如，输入的手掌图像 11 的尺寸为 640*640，输出的多个尺度的图像特征的尺寸为：80*80、40*40、20*20。图像或特征的尺度、尺寸，可以是表达该图像或特征的矩阵的维度。

示例性地，手掌识别设备 10 将多个尺度的图像特征输入至颈部网络进行特征融合，得到图像融合特征。手掌识别设备 10 将图像融合特征输入至预测网络进行预测，得到手掌图像 11 中的掌部框。手掌识别设备 10 在手掌图像 11 中将掌部框之外的图像区域裁剪掉，将裁剪剩余的图像作为掌部图像 12。

综上，本实施例提供的方法，通过获取手掌图像并进行特征提取，得到手掌图像对应的多个尺度的图像特征；将多个尺度的图像特征进行特征融合，得到图像融合特征；基于

图像融合特征确定手掌图像中掌部对应的掌部框；基于掌部框，在手掌图像中裁剪得到掌部对应的掌部图像；将掌部图像进行图像加密，得到掌部图像加密数据，将掌部图像加密数据发送至手掌识别服务器中进行解密和掌部对比识别，进行身份识别。本申请通过从手掌图像中裁剪得到掌部对应的掌部图像，且仅对掌部图像进行图像加密、传输和解密，节省了手掌图像的识别时间，提高了手掌识别的识别效率。

图2示出了本申请一个实施例提供的计算机系统的架构示意图。该计算机系统可以包括：手掌识别设备100和手掌识别服务器200。

手掌识别设备100可以是诸如手机、平板电脑、车载终端（车机）、可穿戴设备、个人计算机（Personal Computer, PC）、具有手掌图像识别功能的语音交互设备、具有手掌图像识别功能的家电、车载终端、飞行器、无人售货终端等电子设备。手掌识别设备100中可以安装运行应用程序的客户端，该应用程序可以是专门进行手掌图像识别的应用程序，也可以是提供有手掌图像识别功能的其他应用程序，本申请对此不作限定。另外，本申请对该应用程序的形式不作限定，包括但不限于安装在手掌识别设备100中的应用程序（Application, App）、子应用程序、网页程序等。其中，子应用程序是运行在母应用程序提供的运行环境中的程序，母应用程序则是一种独立的原生应用程序，子应用程序依赖于母应用程序运行。

手掌识别服务器200可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云计算服务的云服务器、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络（Content Delivery Network, CDN）、以及大数据和人工手掌图像识别平台等基础云计算服务的云服务器。手掌识别服务器200可以是上述应用程序的后台服务器，用于为应用程序的客户端提供后台服务。

其中，云技术（Cloud technology）是指在广域网或局域网内将硬件、软件、网络等系列资源统一起来，实现数据的计算、储存、处理和共享的一种托管技术。云技术基于云计算商业模式应用的网络技术、信息技术、整合技术、管理平台技术、应用技术等的总称，可以组成资源池，按需所用，灵活便利。云计算技术将变成重要支撑。技术网络系统的后台服务需要大量的计算、存储资源，如视频网站、图片类网站和更多的门户网站。伴随着互联网行业的高度发展和应用，将来每个物品都有可能存在自己的识别标志，都需要传输到后台系统进行逻辑处理，不同程度级别的数据将会分开处理，各类行业数据皆需要强大的系统后盾支撑，只能通过云计算来实现。

在一些实施例中，上述服务器还可以实现为区块链系统中的节点。区块链（Blockchain）是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性（防伪）和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层。

手掌识别设备100和手掌识别服务器200之间可以通过网络进行通信，如有线或无线网络。

图3是本申请一个示例性实施例提供的手掌图像处理方法的流程图。该方法由手掌识别设备执行。可以理解，该方法也可以由服务器执行。该方法包括：

步骤302：获取手掌图像，对手掌图像提取特征，得到多个尺度的图像特征。

其中，手掌图像是对实体手掌进行拍摄得到的图像。实体手掌包括掌部和手指，掌部则是连接实体胳膊和实体手指的这部分有机体。手掌图像包括实体手掌全部或部分的影像，为了方便描述，后续将实体手掌在手掌图像中的影像，称之为手掌。手掌图像包含的手掌，可以是待进行身份识别的用户的实体手掌。手掌图像中的手掌，至少包括掌部，还可以包括手指或部分手臂，还可以包括对实体手掌拍摄时实体手掌所处环境的影像。

尺度，是图像特征的尺寸，可以用矩阵维度数表示，如100*100、80*80、60*60或40*40。

多个尺度是指两个或两个以上的尺度，且每个尺度不同。多个尺度中最大尺度可以与手掌图像的尺度相同，一般比手掌图像的尺度小。

该手掌图像可以是由该手掌识别设备对待验证身份的用户实体手掌进行拍摄得到的，也可以是由其他设备发送的。手掌识别设备可直接采集实体手掌的图像，得到手掌图像。手掌识别设备也可以通过手掌图像连接的图像采集器对实体手掌采集图像，得到手掌图像。手掌识别设备可对获取的手掌图像进行特征提取，得到手掌图像对应的多个尺度的图像特征。

例如，手掌识别设备为商店收银设备，商店收银设备通过摄像头拍摄用户的实体手掌，得到该手掌图像；或者，手掌识别设备为手掌图像识别服务器，商店收银设备拍摄到用户的手掌图像后，将该手掌图像发送至该手掌图像识别服务器。

示例性地，多个尺度的图像特征，可以通过将高分辨率的手掌图像下采样为多个尺度的低分辨率的图像后，从多个尺度的低分辨率的图像分别提取的。其中高分辨率和低分辨率是相对的概念，高分辨率是比低分辨率高的分辨率。

例如，输入的手掌图像的尺寸为 640*640，输出的多个尺度的图像特征的尺寸为：80*80、40*40、20*20。

步骤 304：将多个尺度的图像特征融合，得到图像融合特征。

手掌识别设备可将多个尺度的图像特征进行特征融合，得到图像融合特征。

多个尺度的图像特征是指对多个尺度的图像进行特征提取得到的特征。不同尺度的图像特征可以表达不同的信息。其中，多个尺度的图像特征中，尺度越大的图像特征，表达的位置信息更多，尺度越小的图像特征，表达的语义信息更多。

示例性地，手掌识别设备提取多个尺度、不同感受野下、不同类别的特征，并将多个尺度的图像特征进行融合，得到手掌图像对应的图像融合特征。

步骤 306：基于图像融合特征，确定标识手掌图像中掌部的掌部框。

掌部框是指在手掌图像中用于表示掌部所在位置的标识。掌部框包含手掌图像中的掌部。掌部框的边界将手掌图像中的掌部围合。

可选地，掌部框的形状可以是规则形状，如可以为矩形、正方形、菱形、六边形、圆形或三角形中的至少一种。掌部框的形状也可以是不规则形状。规则形状是能够用固定参数基于固定规则绘制的形状。不规则形状则无法用固定参数并基于固定规则绘制，一般只能用构成不规则形状的点的集合来表示。

手掌识别设备可采用人工智能的方式，基于图像融合特征进行预测，确定手掌图像中掌部对应的掌部框。

在一些实施例中，计算机设备可基于图像融合特征，确定标识手掌图像中掌部的掌部框在手掌图像中的位置。当掌部框是规则形状，该位置可以用绘制该规则形状所需的固定参数的值表示，如可以用掌部框的固定点的位置和掌部框的尺寸参数表示。固定点是相对于掌部框位置固定的点。尺寸参数是表示掌部框尺寸大小的参数。如，当掌部框是矩形，固定点可以是任意的角点或者矩形的中心点，尺寸参数则是掌部框相邻边的边长。

步骤 308：基于掌部框，从手掌图像中裁剪出掌部图像。

掌部图像是手掌图像中的有效识别区域，或，掌部图像是指手掌图像中掌部所在的区域，或，掌部图像是指手掌图像中能够用于手掌识别的区域。

示例性地，手掌识别设备基于掌部框的位置，在手掌图像中将位于掌部框之外的图像区域裁剪掉，将裁剪后剩余的图像作为掌部图像。

步骤 310：将掌部图像进行图像加密，得到掌部图像加密数据，掌部图像加密数据用于传输至服务器，由服务器基于掌部图像加密数据进行身份识别。

掌部图像加密数据，是对掌部图像加密后得到的数据，将掌部图像加密数据解密，可以得到掌部图像。其中，掌部图像加密数据可以是比特流或字串，也可以是肉眼难以辨识内容的图像。掌部图像用于被发送至服务器，由服务器基于掌部图像加密数据进行身份识别。

别。

手掌识别设备可接入网络，通过网络将掌部图像加密数据发送至服务器。手掌识别设备也可以通过与终端之间的点对点连接方式，将掌部图像加密数据发送至终端，再由接入网络的终端将掌部图像加密数据发送至服务器。当手掌识别设备本身是服务器，则可以通过直连线缆或网络将掌部图像加密数据发送至另一服务器。

在一些实施例中，服务器可将掌部图像加密数据解密，得到掌部图像，再基于解密得到的掌部图像进行身份识别。

在一些实施例中，服务器可将掌部图像加密数据解密，得到掌部图像，从数据库中查找与解密得到的掌部图像匹配的掌部图像，当查找到匹配的掌部图像，则确定与匹配的掌部图像关联的用户标识，以进行身份识别。当未查找到匹配的掌部图像，则可确定身份识别失败。

在一些实施例中，服务器可将掌部图像加密数据解密，得到掌部图像，从掌部图像中提取掌部特征，再从数据库中查找与提取的掌部特征匹配的掌部特征，当查找到匹配的掌部特征，则确定与匹配的掌部特征关联的用户标识，以进行身份识别。

在一些实施例中，服务器在基于掌部图像加密数据进行身份识别确定用户标识后，可进一步验证用户标识对应的权限。当经验证用户标识具有预设功能的权限，则服务器可进一步执行或触发该预设功能。该预设功能如门禁控制或其它需要验证用户权限的功能。

可选地，手掌识别设备将掌部图像加密数据发送至手掌识别服务器的方式包括网络传输、数据线传输、蓝牙传输中的至少一种，但不限于此，本申请实施了对此不作具体限定。

手掌识别设备可基于掌部图像加密数据进行解密以及掌部对比识别，以进行身份识别。掌部对比识别，具体可以是将掌部图像中的掌部特征与数据库中的预设掌部特征进行对比识别。

预设掌部特征为存储的用户标识掌部的掌部特征，每个预设掌部特征具有对应的用户标识，表示该预设掌部特征属于该用户标识，是该用户掌部的掌部特征。该用户标识可以为任意的用户标识，如，该用户标识为支付应用中注册的用户标识，或，该用户标识为企业中登记的用户标识。

在本申请实施例中，手掌识别服务器中包括数据库，该数据库中包括多个预设掌部特征，及每个预设掌部特征对应的用户标识。在该数据库中，预设掌部特征与用户标识可以是一一对应，也可以是一个用户标识对应至少两个预设掌部特征。

例如，多个用户在支付应用中进行注册，通过将每个用户的预设掌部特征与对应的用户标识进行绑定，将多个用户的掌部特征与对应的用户标识对应存储于数据库中，后续用户使用支付应用时，通过对红外图像的第一掌部区域、彩色图像的第二掌部区域与数据库中的预设掌部特征进行掌部对比识别，来确定目标用户标识，实现对用户的身份验证。

示例性地，手掌识别设备将掌部图像加密数据发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像对应的用户标识。

综上，本实施例提供的方法，通过获取手掌图像并进行特征提取，得到手掌图像对应的多个尺度的图像特征；将多个尺度的图像特征进行特征融合，得到图像融合特征；基于图像融合特征进行预测，得到手掌图像中掌部对应的掌部框；基于掌部框，在手掌图像中裁剪得到掌部对应的掌部图像；将掌部图像进行图像加密，得到掌部图像加密数据，将掌部图像加密数据发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像对应的用户标识。本申请通过从手掌图像中裁剪得到掌部对应的掌部图像，且仅对掌部图像进行图像加密、传输和解密，节省了手掌图像的识别时间，提高了手掌识别的识别效率。

图4是本申请一个示例性实施例提供的手掌图像处理方法的流程图。该方法由手掌识别设备执行。该方法包括：

步骤402：获取手掌图像，通过掌部检测模型中的主干网络，对手掌图像提取特征，得到多个尺度的图像特征。

手掌图像为待确定用户标识的手掌图像。

示例性地，手掌识别设备对用户的实体手掌进行拍摄，得到手掌图像。其中，该手掌图像中包含该手掌，该手掌可以为用户的左手掌，也可以为用户的右手掌。

例如，该手掌识别设备为物联网设备，该物联网设备通过摄像头拍摄用户的左手掌，得到手掌图像，该物联网设备可以为商店收银设备。再例如，用户在商店购物进行交易时，用户将手掌伸向商店支付终端的摄像头，该商店支付终端通过该摄像头拍摄该用户的实体手掌，得到手掌图像。

在一些实施例中，手掌识别设备与其他设备建立通信连接，通过该通信连接，接收其他设备发送的手掌图像。例如，该手掌识别设备为支付应用服务器，其他设备可以为支付终端，支付终端拍摄用户的实体手掌，得到手掌图像后，通过该支付终端与支付应用服务器之间的通信连接，将该手掌图像发送至支付应用服务器，以使该支付应用服务器能够确定该手掌图像的用户标识。

在一些实施例中，掌部框通过掌部检测模型预测得到，掌部检测模型包括主干网络；对手掌图像提取特征，得到多个尺度的图像特征，包括：通过主干网络对手掌图像进行切片操作，得到多个尺度的切片图像；通过主干网络对多个尺度的切片图像分别提取特征，得到多个尺度的图像特征。

其中，掌部检测模型是人工智能模型，可以是深度学习模型，可以是神经网络模型。掌部检测模型至少具有基于图像融合特征，确定标识手掌图像中掌部的掌部框的功能。主干网络是掌部检测模型中用于从手掌图像提取多个尺度的图像特征的结构。切片操作是对手掌图像进行切分、以生成更小尺度图像的操作。更小尺度是指生成的图像的尺度小于手掌图像的尺度。切片图像的像素点来源于手掌图像。

手掌识别设备可通过主干网络进行切片操作，得到多个尺度的切片图像。手掌识别设备可对多个尺度的切片图像进行特征提取，得到手掌图像对应的多个尺度的图像特征。

在一些实施例中，通过主干网络对手掌图像进行切片操作，得到多个尺度的切片图像，包括：通过主干网络，基于手掌图像，确定多个尺度中最大尺度的切片图像；通过主干网络，将最大尺度的切片图像作为首层，开始逐层下采样，获得包括首层的切片图像在内的多个尺度的切片图像。

计算机设备可先基于手掌图像，确定多个尺度中最大尺度的切片图像。该最大尺度的图像特征的尺度，可以小于或等于手掌图像的尺度。计算机设备可对手掌图像进行一次下采样，从而生成多个尺度中最大尺度的切片图像。

对手掌图像进行下采样，是在手掌图像中每隔至少一个像素点采样的过程，将下采样得到的像素点按照在手掌图像中的顺序拼接，就可以得到尺度变小的切片图像。

若多个尺度中尺度的数量为 N ，即多个尺度为 N 个尺度。 N 个尺度都是不同的尺度，因此 N 个尺度中存在最大尺度。将最大尺度的切片图像作为第 1 层，从第 1 层开始逐层下采样，对第 n 层进行下采样得到第 $n+1$ 层的切片图像，直到对 $N-1$ 层下采样得到第 N 层的切片图像。 n 是从 1 到 $N-1$ 的正整数。最后可以得到从 1 到 N 层共 N 个切片图像。

本实施例中，通过逐层下采样获得多尺度的切片图像，提供了从手掌图像中切分出多尺度的切片图像的新途径。

在一些实施例中，通过主干网络对手掌图像进行切片操作，得到多个尺度的切片图像，包括：通过主干网络，将手掌图像中的像素点，在多个尺度下采样并拼接，得到多个尺度的切片图像。其中，不同尺度的切片图像中相邻的两个像素点，在手掌图像中采样处的两个像素点间隔的像素点数量不同。

本实施例中，手掌识别设备可在每种尺度下，采用该尺度下的采样间隔，从手掌图像中采样像素点，将采样的像素点按照在其在手掌图像中的相对位置关系拼接，从而得到每种尺度下的切片图像。其中，采样间隔是间隔的像素点的数量，包括行列两个方向上像素点的间隔。其中，尺度与采样间隔负相关，尺度越大，采样间隔越小。尺度越小，采样间

隔越大。手掌图像没有采样的像素点，则可以集中到图像特征的通道。

例如，在多个尺度中的最大尺度下，在手掌图像中，每隔一列进行取样，即，在行方向上每隔一个像素选取像素点，将选取的像素点进行拼接，从而得到切片图像。拼接起来的切片图像中的信息并没有丢失，但将手掌图像中的尺寸信息集中于通道，将得到的新图片再经过卷积操作，便可得到没有信息丢失情况下的下采样特征图。

再例如，在多个尺度中的第二大尺度下，在手掌图像中，每隔两列进行取样，即，每隔两个像素选取像素点，将选取的像素点进行拼接，从而得到切片图像。拼接起来的切片图像中的信息并没有丢失，但将手掌图像中的尺寸信息集中于通道，将得到的新图片再经过卷积操作，便可得到没有信息丢失情况下的下采样特征图。

例如，原始图像的尺寸为：640*640*3，通过切片操作，得到 320*320*12 的特征图。其中，原始图像中的 640*640 用以表示宽*高，原始图像中的 3 用以表示每个像素对应的特征向量的长度（也称为通道数）。

步骤 404：通过掌部检测模型的颈部网络，将多个尺度的图像特征进行特征融合，得到图像融合特征。

掌部框通过掌部检测模型预测得到，掌部检测模型包括颈部网络；将多个尺度的图像特征融合，得到图像融合特征，包括：通过颈部网络，将多个尺度的图像特征进行特征融合，得到图像融合特征。

其中，颈部网络属于掌部检测模型，是用于将多个尺度的图像特征进行特征融合以得到图像融合特征的结构。颈部网络连接主干网络，用于接收主干网络输出的多个尺度的图像特征。

多个尺度的图像特征是指对多个尺度的图像进行特征提取得到的特征。不同尺度的图像特征可以表达不同的信息。其中，多个尺度的图像特征中，尺度越大的图像特征，表达的位置信息更多，尺度越小的图像特征，表达的语义信息更多。

示例性地，手掌识别设备提取多个尺度、不同感受野下、不同类别的特征，并将多个尺度的图像特征进行融合，得到手掌图像对应的图像融合特征。

在一些实施例中，通过颈部网络，将多个尺度的图像特征进行特征融合，得到图像融合特征，包括：通过颈部网络将多个尺度的图像特征在相同尺度融合，得到图像融合特征。其中，相同尺度的图像特征的融合，可以是求和或求平均值。

手掌识别设备将多个尺度的图像特征进行融合，增加了小目标的检测精度，从而可以更准确地检测出掌部框，从而能够更准确地裁剪出掌部图像，进一步减少掌部图像加密数据的数据量，提高身份识别效率，并保障身份识别的安全性。

在一些实施例中，通过颈部网络，将多个尺度的图像特征进行特征融合，得到图像融合特征，包括：通过颈部网络，基于多个尺度的图像特征进行特征融合，得到多个尺度的第一中间特征；通过颈部网络，基于多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征；通过颈部网络，将多个尺度的第二中间特征进行特征融合，得到图像融合特征。

其中，多个尺度的特征融合，是指在不同尺度上分别进行特征融合。多个尺度的第一中间特征，可以与多个尺度的图像特征一一对应。多个尺度的第一中间特征，可以与多个尺度的第二中间特征一一对应。图像特征的多个尺度，可以与第一中间特征的多个尺度相一致。第一中间特征的多个尺度，可以与第二中间特征的多个尺度相一致。将多个尺度的第二中间特征进行特征融合，可以是将多个尺度的第二中间特征统一到相同尺度下进行融合，该融合可以是求和或者求平均值。

本实施例中，不同尺度行进行特征融合，最后再融合为图像融合特征，可以充分地表达出手掌图像中的细节特征，从而便于更准确地检测出掌部框，从而能够更准确地裁剪出掌部图像，进一步减少掌部图像加密数据的数据量，提高身份识别效率，并保障身份识别的安全性。

在一些实施例中，多个尺度为 N 个尺度， N 为大于 1 的正整数，基于多个尺度的图像特征，提取多个尺度的第一中间特征，包括：基于按尺度正序第 1 位的图像特征，确定按尺度正序第 1 位的第一中间特征；从按尺度正序第 1 位的第一中间特征开始，逐层上采样，将按尺度正序第 n 位的第一中间特征上采样的结果，与按尺度正序第 $n+1$ 位的图像特征融合，获得按尺度正序第 $n+1$ 位的第一中间特征；当获得第 N 层的第一中间特征，确定从第 1 层到第 N 层的多个尺度的第一中间特征。其中， n 为 1 到 $N-1$ 的正整数；按尺度正序是按尺度从小到大的顺序。

其中，将从小到大的顺序称为正序，按尺度正序，即按照尺度从小到大的顺序。上采样是放大尺度的过程。上采样可以下一层的尺度为目标，对每一层的特征中相邻的特征元素进行插值，得到下一层的特征。

在一些实施例中，基于按尺度正序第 1 位的图像特征，确定按尺度正序第 1 位的第一中间特征包括：对按尺度正序第 1 位的图像特征进行至少一次卷积，得到按尺度正序第 1 位的第一中间特征。除了进行卷积，还可以进行通道调整。通道调整后第一中间特征的通道数可以一致。

在一些实施例中，按尺度正序第 n 位的第一中间特征上采样的结果的尺度，可以与按尺度正序第 $n+1$ 位的图像特征的尺度相同。将按尺度正序第 n 位的第一中间特征上采样的结果，与按尺度正序第 $n+1$ 位的图像特征融合，可以是将两者拼接，或者在相同尺度下加和或求平均值。

示例性地，如图 5 所示出的图像特征融合的示意图。手掌识别设备对手掌图像 501 进行特征提取，得到手掌图像 501 对应的多个尺度的图像特征 502。图像特征 502 包括特征层 T1、特征层 T2、特征层 T3、特征层 T4。在获取图像特征 502 后，将特征层 T4 进行 1 次卷积并调整通道后得到特征层 P4。将特征层 P4 进行上采样并与特征层 T3 结合，再进行卷积并调整通道后得到特征层 P3。将特征层 P3 进行上采样并与特征层 T2 结合，再进行卷积并调整通道后得到特征层 P2。将特征层 P2 进行上采样并与特征层 T1 结合，再进行 1×1 卷积调整通道后得到特征层 P1。

在一些实施例中，多个尺度为 N 个尺度， N 为大于 1 的正整数，基于多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征，包括：基于按尺度倒序第 1 位的第一中间特征，确定按尺度倒序第 1 位的第二中间特征；从按尺度倒序第 1 位的第二中间特征开始，逐层下采样，将按尺度倒序第 m 位的第二中间特征下采样的结果，与按尺度倒序第 $m+1$ 位的第一中间特征融合，获得按尺度倒序第 $m+1$ 位的第二中间特征； m 为 1 到 $N-1$ 的整数；按尺度倒序是按尺度从大到小的顺序；当获得第 N 层的第一中间特征，确定从第 1 层到第 N 层的多个尺度的第一中间特征。

其中，将从大到小的顺序称为倒序，按尺度倒序，即按照尺度从大到小的顺序。下采样是缩小尺度的过程。下采样可以下一层的尺度为目标，对每一层的特征中的特征元素进行间隔采用，得到下一层的特征。其中，第一中间特征可以称为语义图像特征。第二中间特征可以称为位置图像特征。语义图像特征更多体现语义信息，语义图像特征更多体现位置信息。

在一些实施例中，基于按尺度倒序第 1 位的第一中间特征，确定按尺度倒序第 1 位的第二中间特征包括：对按尺度倒序第 1 位的第一中间特征进行至少一次卷积，得到按尺度倒序第 1 位的第二中间特征。除了进行卷积，还可以进行通道调整。通道调整后第一中间特征的通道数可以一致。

按尺度倒序第 m 位的第二中间特征下采样的结果的尺度，可以与按尺度倒序第 $m+1$ 位的第一中间特征的尺度一致。将按尺度倒序第 m 位的第二中间特征下采样的结果，与按尺度倒序第 $m+1$ 位的第一中间特征融合，可以是将两者拼接，或将两者在相同尺度下加和或求平均值。

示例性地，手掌识别设备可将特征层 P1 进行 1 次卷积并调整通道后得到特征层 N1。

将特征层 N1 进行下采样并与特征层 P2 结合，再进行卷积并调整通道后得到特征层 N2；将特征层 N2 进行下采样并与特征层 P3 结合，再进行卷积并调整通道后得到特征层 N3；将特征层 N3 进行下采样并与特征层 P4 结合，再进行卷积并调整通道后得到特征层 N4。手掌识别设备将特征层 N1-4 进行特征融合，得到图像融合特征 503。

可选地，卷积中的激活函数为 logistic 激活函数。

在一些实施例中，手掌识别设备将多个尺度的图像特征由小尺度到大尺度进行上采样，得到第一中间特征；手掌识别设备将多个尺度的图像特征由大尺度到小尺度进行下采样，得到第二中间特征；手掌识别设备将第一中间特征和第二中间特征进行特征融合，得到图像融合特征。

步骤 406：通过掌部检测模型的预测网络，基于图像融合特征，确定标识手掌图像中掌部的掌部框。

掌部框是指在手掌图像中用于表示掌部所在位置的标识。

在一些实施例中，掌部框通过掌部检测模型确定，掌部检测模型包括预测网络；基于图像融合特征，确定标识手掌图像中掌部的掌部框，包括：将图像融合特征进行网格划分，得到图像融合特征的多个网格特征；将多个网格特征输入预测网络，对多个网格特征进行预测，得到掌部框。

其中，网格特征是将图像融合特征进行网格划分后得到的网格化的图像融合特征。示例性地，手掌识别设备将图像融合特征进行网格划分，得到多个网格特征；手掌识别设备通过预测网络，基于每个网格特征分别进行预测，得到每个网格特征对应的预测结果，进一步可以综合多个网格特征各自对应的预测结果，确定掌部框在手掌图像中的位置。

在一些实施例中，手掌识别设备装配有红外摄像头，手掌图像处理方法还包括：获取对与手掌图像相同掌部采集的红外图像，红外图像是指红外摄像头基于红外光对掌部成像所得到的图像；从红外图像识别掌部区域。基于掌部框，从手掌图像中裁剪出掌部图像，包括：确定手掌图像中的掌部框，以及红外图像中掌部区域的交集；基于交集，从手掌图像中裁剪出掌部图像。

本实施例中，手掌识别设备还设置有红外摄像头，手掌识别设备获取同一实体手掌对应的红外图像，手掌识别设备对红外图像进行区域识别，确定红外图像上的掌部区域。红外图像是指红外摄像头基于红外光对手掌成像所得到的图像。

在一些实施例中，从红外图像识别掌部区域，包括：检测红外图像中的手指缝点；基于手指缝点，确定红外图像中的掌部区域。

示例性地，手掌识别设备对红外图像中的手指缝点进行检测，基于手指缝点确定红外图像上的掌部区域。

如图 6 所示出的手掌中手指缝点的示意图，手指缝点为食指与中指之间的第一手指缝点 601，或，手指缝点为中指与无名指之间的第二手指缝点 602，或，手指缝点为无名指与小指之间的第三手指缝点 603。

由于在手掌图像中的掌部区域可能存在于该手掌图像中的任一区域，为了能够确定掌部区域在该手掌图像中的位置，通过对手掌图像进行手指缝点检测，从而得到手掌图像的至少一个手指缝点，以便后续能够根据该至少一个手指缝点，确定掌部区域。

步骤 408：基于掌部框，从手掌图像中裁剪出掌部图像。

掌部图像是手掌图像中的有效识别区域，或，掌部图像是指手掌图像中掌部所在的区域，或，掌部图像是指手掌图像中能够用于手掌识别的区域。

在一些实施例中，手掌识别设备获取红外图像中的至少三个手指缝点；顺序连接手指缝点，得到手指缝点连线；手掌识别设备基于手指缝点连线，在红外图像中裁剪得到掌部区域。

在一些实施例中，确定手掌图像中的掌部框，以及红外图像中掌部区域的交集，包括：获取手掌图像中掌部框的坐标参数，以及获取红外图像中掌部区域的坐标参数；基于掌部

框的坐标参数与掌部区域的坐标参数，确定掌部框和掌部区域的交集。

手掌识别设备获取手掌图像中的掌部框的坐标参数，以及获取红外图像中的掌部区域的坐标参数；手掌识别设备基于掌部框的坐标参数与掌部区域的坐标参数的交集，在手掌图像中裁剪得到掌部对应的掌部图像。

步骤 410：将掌部图像进行图像加密，得到掌部图像加密数据，掌部图像加密数据用于传输至服务器，由服务器基于掌部图像加密数据进行身份识别。

其中，掌部图像加密数据用于被发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像对应的用户标识。掌部对比识别是指将掌部图像中的掌部特征与数据库中的预设掌部特征进行对比识别。

需要说明的是，掌部作为生物特征的一种，具有生物唯一性与区分性。相对于目前被广泛应用于核身、支付、门禁、乘车等领域的面部识别，掌部不会受化妆、口罩、墨镜等影响，可以提高用户身份验证的准确率。在某些场景下，如疫情防控场景下，需要佩戴口罩遮住口鼻，这种情况下使用手掌图像进行身份验证可以作为一种更好的选择。

跨设备注册识别是一种对于用户体验非常重要的能力。对于关联的两种类型的设备，用户可以在一种类型的设备中进行注册，将用户的用户标识与该用户的掌部特征进行绑定，之后该用户可以在另一种类型的设备上身份验证。由于手机和物联网设备在图像风格和图像质量上差别大，通过跨设备注册识别，可以使用户在手机端注册后，直接可以在物联网设备端使用，无需用户在两种类型的设备上注册，例如，用户通过手机端注册后，可以在商店的设备上直接进行身份验证，无需用户在该商店的设备上注册，避免了用户的信息泄露。

综上，本实施例提供的方法，通过获取手掌图像并进行特征提取，得到手掌图像对应的多个尺度的图像特征；将多个尺度的图像特征进行特征融合，得到图像融合特征；基于图像融合特征进行预测，得到手掌图像中掌部对应的掌部框；基于掌部框，在手掌图像中裁剪得到掌部对应的掌部图像；将掌部图像进行图像加密，得到掌部图像加密数据，将掌部图像加密数据发送至手掌识别服务器中进行解密和掌部对比识别，得到手掌图像对应的用户标识。本申请通过从手掌图像中裁剪得到掌部对应的掌部图像，且仅对掌部图像进行图像加密、传输和解密，节省了手掌图像的识别时间，提高了手掌识别的识别效率。

本实施例提供的方法，通过主干网络对手掌图像进行切片操作并进行特征提取，得到多个尺度的图像特征，基于上述方法，得到了多个尺度、不同感受野下、不同类别的特征，提高了掌部框预测的精确度。

本实施例提供的方法，通过颈部网络对多个尺度的图像特征由小尺度到大尺度进行上采样，得到第一中间特征；将多个尺度的图像特征由大尺度到小尺度进行下采样，得到第二中间特征；将第一中间特征和第二中间特征进行特征融合，得到图像融合特征，基于上述方法，能够将多个尺度的图像特征进行融合，增加了手掌图像中的小目标的检测精度。

本实施例提供的方法，通过将图像融合特征进行网格划分，得到网格化的图像融合特征；通过预测网络，基于网格化的图像融合特征预测，得到掌部框，基于上述方法，提高了掌部框预测的精确度，进一步保证了基于手掌进行身份识别的安全性。

本实施例提供的方法，通过获取同一掌部对应的红外图像，对红外图像进行区域识别，确定红外图像上的掌部区域，基于掌部框和红外图像上的掌部区域的交集，在手掌图像中裁剪得到掌部对应的掌部图像，基于上述方法，通过红外图像上的掌部区域和掌部框，共同得到掌部对应的掌部图像，提高了掌部图像的裁剪精度，使得得到的掌部图像更加精确。

图 7 是本申请一个示例性实施例提供的手掌图像处理方法的示意图。该方法由手掌识别设备执行。该方法包括：

手掌识别设备获取掌部对应的手掌图像 701，手掌识别设备将获取的手掌图像 701 输入至主干网络 702 进行特征提取，得到手掌图像 701 对应的多个尺度的图像特征。

手掌识别设备通过颈部网络 703，将多个尺度的图像特征进行特征融合，得到图像融

合特征。

手掌识别设备通过预测网络 704，基于图像融合特征进行预测，得到手掌图像 701 中掌部对应的掌部框。手掌识别设备基于掌部框，在手掌图像 701 中裁剪得到掌部对应的掌部图像 705。

手掌识别设备将掌部图像 705 输入至加密网络 706 进行图像加密，得到掌部图像加密数据 707。

手掌识别设备将掌部图像加密数据 707 通过网络传输 708 至手掌识别服务器中的解密网络 709 进行图像解密并通过验证网络 710 进行掌部对比识别，得到手掌图像 701 对应的用户标识 711。

综上，本实施例提供的方法，通过从手掌图像中裁剪得到掌部对应的掌部图像，且仅对掌部图像进行图像加密、传输和解密，节省了手掌图像的识别时间，提高了手掌识别的识别效率。

图 8 是本申请一个示例性实施例提供的基于手掌图像处理方法的跨设备支付的示意图。该方法涉及用户终端 801、商户终端 803 及支付应用服务器 802。

其中，用户终端 801 安装有支付应用，用户终端 801 基于用户标识登录支付应用，与支付应用服务器 802 建立通信连接，通过该通信连接，用户终端 801 与支付应用服务器 802 可以进行交互；商户终端 803 均安装有支付应用，商户终端 803 基于商户标识登录支付应用，与支付应用服务器 802 建立通信连接，通过该通信连接，商户终端 803 与支付应用服务器 802 可以进行交互。

该跨设备支付流程包括：

1、用户在家中手持用户终端 801，通过该用户终端 801 拍摄用户自己的实体手掌，得到该用户的手掌图像，基于用户标识登录支付应用，向支付应用服务器 802 发送手掌图像注册请求，该手掌图像注册请求携带该用户标识及手掌图像。

2、支付应用服务器 802 接收到用户终端 801 发送的手掌图像注册请求，对手掌图像进行特征提取，得到该手掌图像的掌部特征，将该掌部特征与该用户标识进行对应存储，向用户终端 801 发送手掌图像绑定成功通知。

其中，支付应用服务器 802 将掌部特征与用户标识进行对应存储后，将该掌部特征作为预设掌部特征，后续可以通过存储的预设掌部特征，来确定对应的用户标识。

3、用户终端 801 接收到手掌图像绑定成功通知，显示该手掌图像绑定成功通知，提示用户手掌图像与用户标识绑定。

其中，用户通过自己的用户终端 801 与支付应用服务器 802 之间的交互，完成手掌图像注册，后续可以通过手掌图像来实现自动支付。

4、用户在商店购买商品进行交易时，商户终端 803 拍摄该用户的实体手掌，得到手掌图像，基于商户标识登录的支付应用，向支付应用服务器 802 发送支付请求，该支付请求携带该商户标识、消费金额及手掌图像。

5、支付应用服务器 802 接收到支付请求后，对手掌图像进行掌部对比识别，确定该手掌图像的用户标识，确定该用户标识在支付应用中的账号，通过该账号完成转账，在转账完成后，向商户终端 803 发送支付完成通知。

其中，用户在利用用户终端 801 进行手掌图像注册后，可以直接在商户终端 803 通过掌部进行支付，无需用户在商户终端 803 上进行手掌图像注册，从而实现了跨设备手掌图像识别的效果，提高了便捷性。

6、商户终端 803 接收到支付完成通知，显示该支付完成通知，提示用户支付完成，以使用户与商户完成物品的交易，用户可以将物品带走。

另外，上述实施例以通过用户终端 801 与商户终端 803 实现跨设备支付的过程，还可以将上述商户终端 803 替换为公交车上的支付设备，按照上述步骤，实现跨设备乘车支付的方案。

图9是本申请一个示例性实施例提供的基于手掌图像处理方法的跨设备身份验证的示意图。该方法涉及用户终端901、门禁设备903及门禁服务器902。

其中，用户终端901与门禁服务器902建立通信连接，通过该通信连接，用户终端901与门禁服务器902可以进行交互；门禁设备903与门禁服务器902建立通信连接，通过该通信连接，门禁设备903与门禁服务器902可以进行交互。

该跨设备身份验证流程包括：

1、用户在家中手持用户终端901，通过该用户终端901拍摄用户自己的实体手掌，得到该用户的手掌图像，并向门禁服务器902发送掌部注册请求，该掌部注册请求携带该用户标识及手掌图像。

2、门禁服务器902接收到用户终端901发送的掌部注册请求，对手掌图像进行特征提取，得到该手掌图像的掌部特征，将该掌部特征与该用户标识进行对应存储，向用户终端901发送掌部绑定成功通知。

其中，门禁服务器902将掌部特征与用户标识进行对应存储后，将该掌部特征可以作为预设掌部特征，后续可以通过存储的预设掌部特征，来确定对应的用户标识。

3、用户终端901接收到掌部绑定成功通知，显示该掌部绑定成功通知，提示用户手掌图像与用户标识绑定。

其中，用户通过自己的用户终端901与门禁服务器之间的交互，完成手掌图像注册，后续可以通过手掌图像来实现自动开门。

4、当用户外出回家时，门禁设备903拍摄该用户的实体手掌，得到该用户的手掌图像，向门禁服务器902发送身份验证请求，该身份验证请求携带该验证手掌图像。

5、门禁服务器902接收门禁设备903发送的身份验证请求，对该验证手掌图像进行识别，得到该手掌图像的用户标识，确定该用户为注册用户，向门禁设备903发送验证通过通知。

6、门禁设备903接收门禁服务器902发送的验证通过通知，根据该验证通过通知，控制家门打开，以使用户能够进入到室内。

上述实施例是以通过用户终端901与门禁设备903实现跨设备身份验证的过程。

通过上述跨设备身份验证场景可知，无论是用户终端901与门禁服务器902之间交互的掌部注册阶段，还是在通过其他终端设备与服务器进行交互的手掌图像处理阶段，均是在用户终端901或其他终端设备在获取到手掌图像后，对手掌图像中掌部所在区域进行裁剪、加密，将掌部所在区域的手掌图像发送至服务器，由服务器进行掌部对比识别。且在掌部对比识别阶段，门禁服务器902通过将该掌部特征与预设掌部特征进行比对，得到当前用户的识别结果。

图10是本申请一个示例性实施例提供的手掌图像处理方法的示意图。该方法由手掌识别设备执行。该方法包括：

手掌识别设备获取掌部对应的手掌图像1001，将手掌图像1001输入至主干网络中的调焦层1002，调焦层1002将手掌图像1001进行切片操作，得到多个尺度的图像特征。将图像特征输入至卷积层1003进行卷积调整通道后输入至瓶颈层1004进行特征提取。池化层1005用于对图像特征进行池化。

例如，手掌图像1001输入至调焦层1002，调焦层1002将手掌图像1001进行切片操作，得到多个尺度的图像特征表示为(3, 64, 1, 1)，经过卷积层1003进行卷积并调整通道后的图像特征表示为(64, 128, 3, 2)，经过瓶颈层1004后，图像特征表示为(128, 128)*3，经过卷积层1003进行卷积调整通道后的图像特征表示为(128, 256, 3, 2)，经过瓶颈层1004后的图像特征表示为(256, 256)*3。图像特征(256, 256)*3经过卷积层1003进行卷积调整通道后的图像特征表示为(256, 512, 3, 2)，经过瓶颈层1004后的图像特征表示为(512, 512)*9，再次经过卷积层1003进行卷积调整通道后的图像特征表示为(512, 1024, 3, 2)，经过池化层1005对图像池化并经过瓶颈层1004后的图像特征

表示为 $(1024, 1024) * 3$ 。图像特征 $(1024, 1024) * 3$ 经过上采样 1006 并与图像特征 $(512, 512) * 9$ 进行拼接 1007 后, 经过瓶颈层 1004 后的图像特征表示为 $(1024, 512) * 3$ 。图像特征 $(1024, 512) * 3$ 经过卷积层 1003 进行卷积调整通道后的图像特征表示为 $(512, 256, 1, 1)$ 。图像特征 $(512, 256, 1, 1)$ 经过上采样 1006 并与图像特征 $(256, 512, 3, 2)$ 进行拼接 1007 后, 经过瓶颈层 1004 后的图像特征表示为 $(512, 256) * 3$ 。图像特征 $(512, 256) * 3$ 经过卷积层 1003 进行卷积调整通道后的图像特征表示为 $(256, 256, 3, 2)$ 。图像特征 $(256, 256, 3, 2)$ 与图像特征 $(512, 256, 1, 1)$ 进行拼接 1007 后, 经过瓶颈层 1004 后的图像特征表示为 $(512, 512) * 3$ 。图像特征 $(512, 512) * 3$ 经过卷积层 1003 进行卷积调整通道后的图像特征表示为 $(512, 512, 3, 2)$ 。图像特征表示为 $(512, 512, 3, 2)$ 与图像特征表示为 $(1024, 512, 1, 1)$ 进行拼接 1007 后, 经过瓶颈层 1004 后的图像特征表示为 $(1024, 1024) * 3$ 。最终输出的多个尺度的图像特征 1008 为: $(512, 256)$ 、 $(512, 512)$ 、 $(1024, 1024)$ 。其中, (x, y, a, b) , x, y 可以分别表示特征矩阵的两个维度, a 和 b 可以分别表示通道数以及尺度。 $(x, y) * b$, x, y 可以分别表示特征矩阵的两个维度, b 可以表示通道数。 (x, y) , x, y 可以分别表示特征矩阵的两个维度。

手掌识别设备将多个尺度的图像特征 1008 进行特征融合, 得到图像融合特征。手掌识别设备将图像融合特征进行预测, 得到手掌图像 1001 中掌部对应的掌部框。手掌识别设备基于掌部框, 在手掌图像 1001 中裁剪得到掌部对应的掌部图像。手掌识别设备将掌部图像进行图像加密, 并通过网络传输至手掌识别服务器中进行图像解密及掌部对比识别, 得到手掌图像 1001 对应的用户标识。

综上, 本实施例提供的方法, 通过主干网络对手掌图像进行切片操作并进行特征提取, 得到多个尺度的图像特征, 基于上述方法, 得到了多个尺度、不同感受野下、不同类别的特征, 提高了掌部框预测的精确度。

本申请涉及的掌部框的预测可以基于掌部检测模型实现, 该方案包括掌部检测模型生成阶段和掌部框的预测阶段。图 11 是本申请一示例性实施例示出的一种掌部检测模型生成阶段和掌部框的预测阶段的框架图, 如图 11 所示, 在掌部检测模型生成阶段, 掌部检测模型生成设备 1110 通过预先设置好的训练样本数据集 (包括样本手掌图像和样本掌部图像), 得到掌部检测模型之后, 基于该掌部检测模型生成掌部框。在掌部框的预测阶段, 掌部框的预测处理设备 1120 基于该掌部检测模型, 对输入的手掌图像进行处理, 获得该手掌图像的预测结果, 比如, 确定该手掌图像中掌部的面积和位置等等。

其中, 上述掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 可以是计算机设备, 比如, 该计算机设备可以是个人电脑、服务器等固定式计算机设备, 或者, 该计算机设备也可以是平板电脑、电子书阅读器等移动式计算机设备。

可选的, 上述掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 可以是同一个设备, 或者, 掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 也可以是不同的设备。并且, 当掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 是不同设备时, 掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 可以是同一类型的设备, 比如掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 可以都是服务器; 或者掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 也可以是不同类型的设备, 比如掌部框的预测处理设备 1120 可以是个人电脑或者终端, 而掌部检测模型生成设备 1110 可以是服务器等。本申请实施例对掌部检测模型生成设备 1110 和掌部框的预测处理设备 1120 的具体类型不做限定。

上实施例对手掌图像处理方法进行了描述, 接下来将就掌部检测模型的训练方法进行描述。

图 12 是本申请一个示例性实施例提供的掌部检测模型的训练方法的流程图。该方法由手掌识别设备执行。该方法包括:

步骤 1201: 获取样本手掌图像和样本掌部图像。

样本手掌图像为待确定用户标识的图像，该样本手掌图像中包含手掌，该手掌为待验证身份的用户的实体手掌的影像，该样本手掌图像还可以包含其他的信息，如用户的手指、拍摄用户手掌时所处的场景等。该样本手掌图像可以是由该手掌识别设备对待验证身份的用户的实体手掌进行拍摄得到的，也可以是由其他设备发送的。

例如，手掌识别设备为商店收银设备，商店收银设备通过摄像头拍摄用户的实体手掌，得到该样本手掌图像；或者，手掌识别设备为手掌图像识别服务器，商店收银设备拍摄到用户的样本手掌图像后，将该样本手掌图像发送至该手掌图像识别服务器。

样本掌部图像是指样本手掌图像中标注的有效识别区域，或，样本掌部图像是指样本手掌图像中标注的掌部所在的区域，或，样本掌部图像是指样本手掌图像中标注的能够用于识别的区域。

步骤 1202：对样本手掌图像提取特征，得到多个尺度的样本图像特征。

其中，样本手掌图像是对实体手掌进行拍摄得到的图像。样本手掌图像包括实体手掌全部或部分的影像，为了方便描述，后续将实体手掌在样本手掌图像中的影像，称之为手掌。样本手掌图像包含的手掌，可以是待进行身份识别的用户的实体手掌。样本手掌图像中的手掌，至少包括掌部，还可以包括手指或部分手臂，还可以包括对实体手掌拍摄时实体手掌所处环境的影像。

尺度，是样本图像特征的尺寸，可以用矩阵维度数表示，如 100*100、80*80、60*60 或 40*40。多个尺度是指两个或两个以上的尺度，且每个尺度不同。多个尺度中最大尺度可以与样本手掌图像的尺度相同，一般比样本手掌图像的尺度小。

该样本手掌图像可以是由该手掌识别设备对待验证身份的用户的实体手掌进行拍摄得到的，也可以是由其他设备发送的。手掌识别设备可直接采集实体手掌的图像，得到样本手掌图像。手掌识别设备也可以通过样本手掌图像连接的图像采集器对实体手掌采集图像，得到样本手掌图像。手掌识别设备可对获取的样本手掌图像进行特征提取，得到样本手掌图像对应的多个尺度的样本图像特征。

例如，手掌识别设备为商店收银设备，商店收银设备通过摄像头拍摄用户的实体手掌，得到该样本手掌图像；或者，手掌识别设备为样本手掌图像识别服务器，商店收银设备拍摄到用户的样本手掌图像后，将该样本手掌图像发送至该样本手掌图像识别服务器。

示例性地，多个尺度的样本图像特征，可以通过将高分辨率的样本手掌图像下采样为多个尺度的低分辨率的图像后，从多个尺度的低分辨率的图像分别提取的。其中高分辨率和低分辨率是相对的概念，高分辨率是比低分辨率高的分辨率。

例如，输入的样本手掌图像的尺寸为 640*640，输出的多个尺度的样本图像特征的尺寸为：80*80、40*40、20*20。

步骤 1203：将多个尺度的样本图像特征融合，得到样本图像融合特征。

手掌识别设备可将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征。

多个尺度的样本图像特征是指对多个尺度的图像进行特征提取得到的特征。不同尺度的样本图像特征可以表达不同的信息。其中，多个尺度的样本图像特征中，尺度越大的样本图像特征，表达的位置信息更多，尺度越小的样本图像特征，表达的语义信息更多。

示例性地，手掌识别设备提取多个尺度、不同感受野下、不同类别的特征，并将多个尺度的样本图像特征进行融合，得到样本手掌图像对应的样本图像融合特征。

步骤 1204：基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框。

样本预测掌部框是指在样本手掌图像中用于表示掌部所在位置的标识。样本预测掌部框包含样本手掌图像中的掌部。样本预测掌部框的边界将样本手掌图像中的掌部围合。

可选地，样本预测掌部框的形状可以是规则形状，如可以为矩形、正方形、菱形、六边形、圆形或三角形中的至少一种。样本预测掌部框的形状也可以是不规则形状。规则形状是能够用固定参数基于固定规则绘制的形状。不规则形状则无法用固定参数并基于固定规则绘制，一般只能用构成不规则形状的点的集合来表示。

手掌识别设备可采用人工智能的方式，基于样本图像融合特征进行预测，确定标识样本手掌图像中掌部的样本预测掌部框。

在一些实施例中，计算机设备可基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框在样本手掌图像中的位置。当样本预测掌部框是规则形状，该位置可以用绘制该规则形状所需的固定参数的值表示，如可以用样本预测掌部框的固定点的位置和样本预测掌部框的尺寸参数表示。固定点是相对于样本预测掌部框位置固定的点。尺寸参数是表示样本预测掌部框尺寸大小的参数。如，当样本预测掌部框是矩形，固定点可以是任意的角点或者矩形的中心点，尺寸参数则是样本预测掌部框相邻边的边长。

步骤 1205：基于样本预测掌部框，从样本手掌图像中裁剪出样本预测掌部图像。

样本预测掌部图像是样本手掌图像中的有效识别区域，或，样本预测掌部图像是指样本手掌图像中掌部所在的区域，或，样本预测掌部图像是指样本手掌图像中能够用于手掌识别的区域。

示例性地，手掌识别设备基于样本预测掌部框的位置，在样本手掌图像中将位于样本预测掌部框之外的图像区域裁剪掉，将裁剪后剩余的图像作为样本预测掌部图像。

步骤 1206：基于样本掌部图像及样本预测掌部图像，计算损失函数值。

示例性地，手掌识别设备基于样本掌部图像及样本预测掌部图像，计算损失函数值。

步骤 1207：基于损失函数值对掌部检测模型的模型参数进行更新。

示例性地，手掌识别设备基于损失函数值对掌部检测模型的模型参数进行更新。

模型参数更新是指对掌部检测模型里面的网络参数进行更新，或对模型里面的各个网络模块的网络参数进行更新，或对模型里面的各个网络层的网络参数进行更新，但不限于此，本申请实施例对此不作限定。

掌部检测模型的模型参数包括掌部检测模型中主干网络的网络参数、颈部网络的网络参数及预测网络的网络参数中的至少一种。

综上，本实施例提供的方法，通过对获取的样本手掌图像进行特征提取，得到样本手掌图像对应的多个尺度的样本图像特征；将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征；基于样本图像融合特征进行预测，得到标识样本手掌图像中掌部的样本预测掌部框；基于样本预测掌部框，在样本手掌图像中裁剪得到掌部对应的样本预测掌部图像；基于样本预测掌部图像及样本预测掌部图像，计算损失函数值；基于损失函数值对掌部检测模型的模型参数进行更新，使得训练好的掌部检测模型能够具备更高的样本预测掌部框预测精度，从而得到更加准确的样本预测掌部框。

图 13 是本申请一个示例性实施例提供的掌部检测模型的训练方法的流程图。该方法由手掌识别设备执行。该方法包括：

步骤 1301：获取样本手掌图像和样本掌部图像。

样本手掌图像为待确定用户标识的图像，该样本手掌图像中包含手掌，该手掌为待验证身份的用户的实体手掌的影像，该样本手掌图像还可以包含其他的信息，如用户的手指、拍摄用户手掌时所处的场景等。该样本手掌图像可以由该手掌识别设备对待验证身份的用户的实体手掌进行拍摄得到的，也可以是由其他设备发送的。

样本掌部图像是指样本手掌图像中标注的有效识别区域，或，样本掌部图像是指样本手掌图像中标注的掌部所在的区域，或，样本掌部图像是指样本手掌图像中标注的能够用于识别的区域。

示例性地，样本手掌图像的采集方式为：获取 50-100 个样本手掌图像。通过查找重复值、缺失值、异常值的方式对样本手掌图像进行清理，从而得到有效的样本手掌图像。对样本手掌图像进行清理的代码可表示为：

```
allocate output Pixel Value [image width][image height]
allocate window [window width × window height]
edge x= (window width / 2) rounded down
```

```

edge y= (window height / 2) rounded down
  for x from edge x to image width – edge x do
    for y from edge y to image height – edge y do
      i = 0
      for fx from 0 to window width do
        for fy from 0 to window height do
          window[i] = input Pixel Value[x + fx – edge x][y + fy – edge y]
          i= i + 1
        sort entries in window []
        output Pixel Value[x][y] = window [window width * window height / 2]

```

样本手掌图像的采集距离要求：将掌部放置于距手掌识别设备 5cm、8cm、10cm、12cm、15cm 的位置进行采集，在每个距离下采集 10-15 秒。

样本掌部图像的获取方式为：通过标注工具 Labeling 对样本手掌图像中的掌部区域进行标注。掌部区域进行标注的代码可表示为：

```

<object>
  <name>plam</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
  <bndbox>
    <xmin>45</xmin>
    <ymin>34</ymin>
    <xmax>104</xmax>
    <ymax>84</ymax>
  </bndbox>
</object>

```

上述代码中，45 表示掌部区域在 x 轴上的最小值，104 表示掌部区域在 x 轴上的最大值，34 表示掌部区域在 y 轴上的最小值，84 表示掌部区域在 y 轴上的最大值。

标注工具 Labeling 对样本手掌图像中的掌部区域通过结合均匀分布定理，用更新后的标签向量 $\hat{y}_i$ 代替传统的独热 one-hot 编码的标签向量 y_{hot} 。

$$\hat{y}_i = y_{hot}(1 - \alpha) + \alpha/k$$

式中，k 为多分类的类别总个数， α 是一个较小的超参数（一般取 0.1）。

即，标签向量 $\hat{y}_i$ 可表示为：

$$\hat{y}_i = \begin{cases} 1 - \alpha, & i = \text{类别个数} \\ \frac{\alpha}{k}, & i \neq \text{类别个数} \end{cases}$$

标签向量 $\hat{y}_i$ 相当于往真实分布中加入了噪声，避免模型对正确标签过于自信，使得预测正负样本的输出值差别不那么大，从而避免过拟合，提高模型的泛化能力。

步骤 1302：通过掌部检测模型中的主干网络，对样本手掌图像提取特征，得到多个尺度的样本图像特征。

示例性地，手掌识别设备对用户的实体手掌进行拍摄，得到样本手掌图像。其中，该样本手掌图像中包含该手掌，该手掌可以为用户的左手掌，也可以为用户的右手掌。

例如，该手掌识别设备为物联网设备，该物联网设备通过摄像头拍摄用户的左手掌，得到样本手掌图像，该物联网设备可以为商店收银设备。再例如，用户在商店购物进行交易时，用户将手掌伸向商店支付终端的摄像头，该商店支付终端通过该摄像头拍摄该用户的实体手掌，得到样本手掌图像。

在一些实施例中，手掌识别设备与其他设备建立通信连接，通过该通信连接，接收其他设备发送的样本手掌图像。例如，该手掌识别设备为支付应用服务器，其他设备可以为支付终端，支付终端拍摄用户的实体手掌，得到样本手掌图像后，通过该支付终端与支付应用服务器之间的通信连接，将该样本手掌图像发送至支付应用服务器，以使该支付应用服务器能够确定该样本手掌图像的用户标识。

在一些实施例中，样本预测掌部框通过掌部检测模型预测得到，掌部检测模型包括主干网络；对样本手掌图像提取特征，得到多个尺度的样本图像特征，包括：将样本手掌图像输入至主干网络；通过主干网络对样本手掌图像进行切片操作，得到多个尺度的切片图像；通过主干网络对多个尺度的切片图像分别提取特征，得到多个尺度的样本图像特征。

其中，掌部检测模型是人工智能模型，可以是深度学习模型，可以是神经网络模型。掌部检测模型至少具有基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框的功能。主干网络是掌部检测模型中用于从样本手掌图像提取多个尺度的样本图像特征的结构。切片操作是对样本手掌图像进行切分，以生成更小尺度图像的操作。更小尺度是指生成的图像的尺度小于样本手掌图像的尺度。切片图像的像素点来源于样本手掌图像。

手掌识别设备可将样本手掌图像输入至主干网络，由主干网络进行切片操作，得到多个尺度的切片图像。手掌识别设备可对多个尺度的切片图像进行特征提取，得到样本手掌图像对应的多个尺度的样本图像特征。

在一些实施例中，通过主干网络对样本手掌图像进行切片操作，得到多个尺度的切片图像，包括：通过主干网络，基于样本手掌图像，确定多个尺度中最大尺度的切片图像；通过主干网络，将最大尺度的切片图像作为首层，开始逐层下采样，获得包括首层的切片图像在内的多个尺度的切片图像。

计算机设备可先基于样本手掌图像，确定多个尺度中最大尺度的切片图像。该最大尺度的样本图像特征的尺度，可以小于或等于样本手掌图像的尺度。计算机设备可对样本手掌图像进行一次下采样，从而生成多个尺度中最大尺度的切片图像。

对样本手掌图像进行下采样，是在样本手掌图像中每隔至少一个像素点采样的过程，将下采样得到的像素点按照在样本手掌图像中的顺序拼接，就可以得到尺度变小的切片图像。

若多个尺度中尺度的数量为 N ，即多个尺度为 N 个尺度。 N 个尺度都是不同的尺度，因此 N 个尺度中存在最大尺度。将最大尺度的切片图像作为第 1 层，从第 1 层开始逐层下采样，对第 n 层进行下采样得到第 $n+1$ 层的切片图像，直到对 $N-1$ 层下采样得到第 N 层的切片图像。 n 是从 1 到 $N-1$ 的正整数。最后可以得到从 1 到 N 层共 N 个切片图像。

本实施例中，通过逐层下采样获得多尺度的切片图像，提供了从样本手掌图像中切分出多尺度的切片图像的新途径。

在一些实施例中，通过主干网络对样本手掌图像进行切片操作，得到多个尺度的切片图像，包括：通过主干网络，将样本手掌图像中的像素点，在多个尺度下采样并拼接，得到多个尺度的切片图像。其中，不同尺度的切片图像中相邻的两个像素点，在样本手掌图像中采样处的两个像素点间隔的像素点数量不同。

本实施例中，手掌识别设备可在每种尺度下，采用该尺度下的采样间隔，从样本手掌图像中采样像素点，将采样的像素点按照其在样本手掌图像中的相对位置关系拼接，从而得到每种尺度下的切片图像。其中，采样间隔是间隔的像素点的数量，包括行列两个方向上像素点的间隔。其中，尺度与采样间隔负相关，尺度越大，采样间隔越小。尺度越小，采样间隔越大。样本手掌图像没有采样的像素点，则可以集中到样本图像特征的通道。

例如，在多个尺度中的最大尺度下，在样本手掌图像中，每隔一行进行取样，即，在行方向上每隔一个像素选取像素点，将选取的像素点进行拼接，从而得到切片图像。拼接起来的切片图像中的信息并没有丢失，但将样本手掌图像中的尺寸信息集中于通道，将得

到的新图片再经过卷积操作，便可得到没有信息丢失情况下的下采样特征图。

再例如，在多个尺度中的第二大尺度下，在样本手掌图像中，每隔两列进行取样，即，每隔两个像素选取像素点，将选取的像素点进行拼接，从而得到切片图像。拼接起来的切片图像中的信息并没有丢失，但将样本手掌图像中的尺寸信息集中于通道，将得到的新图片再经过卷积操作，便可得到没有信息丢失情况下的下采样特征图。

例如，原始图像的尺寸为：640*640*3，通过切片操作，得到 320*320*12 的特征图。其中，原始图像中的 640*640 用以表示宽*高，原始图像中的 3 用以表示每个像素对应的特征向量的长度（也称为通道数）。

步骤 1303：通过掌部检测模型的颈部网络，将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征。

样本预测掌部框通过掌部检测模型预测得到，掌部检测模型包括颈部网络；将多个尺度的样本图像特征融合，得到样本图像融合特征，包括：通过颈部网络，将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征。

其中，颈部网络属于掌部检测模型，是用于将多个尺度的样本图像特征进行特征融合以得到样本图像融合特征的结构。颈部网络连接主干网络，用于接收主干网络输出的多个尺度的样本图像特征。

多个尺度的样本图像特征是指对多个尺度的图像进行特征提取得到的特征。不同尺度的样本图像特征可以表达不同的信息。其中，多个尺度的样本图像特征中，尺度越大的样本图像特征，表达的位置信息更多，尺度越小的样本图像特征，表达的语义信息更多。

示例性地，手掌识别设备提取多个尺度、不同感受野下、不同类别的特征，并将多个尺度的样本图像特征进行融合，得到样本手掌图像对应的样本图像融合特征。

在一些实施例中，通过颈部网络，将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征，包括：通过颈部网络将多个尺度的样本图像特征在相同尺度融合，得到样本图像融合特征。其中，相同尺度的样本图像特征的融合，可以是求和或求平均值。

手掌识别设备将多个尺度的样本图像特征进行融合，增加了小目标的检测精度，从而可以更准确地检测出样本预测掌部框，从而能够更准确地裁剪出样本预测掌部图像，进一步减少样本预测掌部图像加密数据的数据量，提高身份识别效率，并保障身份识别的安全性。

在一些实施例中，通过颈部网络，将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征，包括：通过颈部网络，基于多个尺度的样本图像特征进行特征融合，得到多个尺度的第一中间特征；通过颈部网络，基于多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征；通过颈部网络，将多个尺度的第二中间特征进行特征融合，得到样本图像融合特征。

其中，多个尺度的特征融合，是指在不同尺度上分别进行特征融合。多个尺度的第一中间特征，可以与多个尺度的样本图像特征一一对应。多个尺度的第一中间特征，可以与多个尺度的第二中间特征一一对应。样本图像特征的多个尺度，可以与第一中间特征的多个尺度相一致。第一中间特征的多个尺度，可以与第二中间特征的多个尺度相一致。将多个尺度的第二中间特征进行特征融合，可以是将多个尺度的第二中间特征统一到相同尺度下进行融合，该融合可以是求和或者求平均值。

本实施例中，不同尺度行进行特征融合，最后再融合为样本图像融合特征，可以充分地表达出样本手掌图像中的细节特征，从而便于更准确地检测出样本预测掌部框，从而能够更准确地裁剪出样本预测掌部图像，进一步减少样本预测掌部图像加密数据的数据量，提高身份识别效率，并保障身份识别的安全性。

在一些实施例中，多个尺度为 N 个尺度，N 为大于 1 的正整数，基于多个尺度的样本图像特征，提取多个尺度的第一中间特征，包括：基于按尺度正序第 1 位的样本图像特征，确定按尺度正序第 1 位的第一中间特征；从按尺度正序第 1 位的第一中间特征开始，逐层

上采样, 将按尺度正序第 n 位的第一中间特征上采样的结果, 与按尺度正序第 $n+1$ 位的样本图像特征融合, 获得按尺度正序第 $n+1$ 位的第一中间特征; 当获得第 N 层的第一中间特征, 确定从第 1 层到第 N 层的多个尺度的第一中间特征。其中, n 为 1 到 $N-1$ 的正整数; 按尺度正序是按尺度从小到大的顺序。

其中, 将从小到大的顺序称为正序, 按尺度正序, 即按照尺度从小到大的顺序。上采样是放大尺度的过程。上采样可以下一层的尺度为目标, 对每一层的特征中相邻的特征元素进行插值, 得到下一层的特征。

在一些实施例中, 基于按尺度正序第 1 位的样本图像特征, 确定按尺度正序第 1 位的第一中间特征包括: 对按尺度正序第 1 位的样本图像特征进行至少一次卷积, 得到按尺度正序第 1 位的第一中间特征。除了进行卷积, 还可以进行通道调整。通道调整后第一中间特征的通道数可以一致。

在一些实施例中, 按尺度正序第 n 位的第一中间特征上采样的结果的尺度, 可以与按尺度正序第 $n+1$ 位的样本图像特征的尺度相同。将按尺度正序第 n 位的第一中间特征上采样的结果, 与按尺度正序第 $n+1$ 位的样本图像特征融合, 可以是将两者拼接, 或将两者在相同尺度下加和或求平均。

示例性地, 如图 5 所示出的样本图像特征融合的示意图。手掌识别设备对样本手掌图像 501 进行特征提取, 得到样本手掌图像 501 对应的多个尺度的样本图像特征 502。样本图像特征 502 包括特征层 T1、特征层 T2、特征层 T3、特征层 T4。在获取样本图像特征 502 后, 将特征层 T4 进行 1 次卷积并调整通道后得到特征层 P4。将特征层 P4 进行上采样并与特征层 T3 结合, 再进行卷积并调整通道后得到特征层 P3。将特征层 P3 进行上采样并与特征层 T2 结合, 再进行卷积并调整通道后得到特征层 P2。将特征层 P2 进行上采样并与特征层 T1 结合, 再进行 $1*1$ 卷积调整通道后得到特征层 P1。

在一些实施例中, 多个尺度为 N 个尺度, N 为大于 1 的正整数, 基于多个尺度的第一中间特征进行多个尺度的特征融合, 得到多个尺度的第二中间特征, 包括: 基于按尺度倒序第 1 位的第一中间特征, 确定按尺度倒序第 1 位的第二中间特征; 从按尺度倒序第 1 位的第二中间特征开始, 逐层下采样, 将按尺度倒序第 m 位的第二中间特征下采样的结果, 与按尺度倒序第 $m+1$ 位的第一中间特征融合, 获得按尺度倒序第 $n+1$ 位的第二中间特征; m 为 1 到 $N-1$ 的整数; 按尺度倒序是按尺度从大到小的顺序; 当获得第 N 层的第一中间特征, 确定从第 1 层到第 N 层的多个尺度的第一中间特征。

其中, 将从大到小的顺序称为倒序, 按尺度倒序, 即按照尺度从大到小的顺序。下采样是缩小尺度的过程。下采样可以下一层的尺度为目标, 对每一层的特征中的特征元素进行间隔采用, 得到下一层的特征。其中, 第一中间特征可以称为语义样本图像特征。第二中间特征可以称为位置样本图像特征。语义样本图像特征更多体现语义信息, 语义样本图像特征更多体现位置信息。

在一些实施例中, 基于按尺度倒序第 1 位的第一中间特征, 确定按尺度倒序第 1 位的第二中间特征包括: 对按尺度倒序第 1 位的第一中间特征进行至少一次卷积, 得到按尺度倒序第 1 位的第二中间特征。除了进行卷积, 还可以进行通道调整。通道调整后第一中间特征的通道数可以一致。

按尺度倒序第 m 位的第二中间特征下采样的结果的尺度, 可以与按尺度倒序第 $m+1$ 位的第一中间特征的尺度一致。将按尺度倒序第 m 位的第二中间特征下采样的结果, 与按尺度倒序第 $m+1$ 位的第一中间特征融合, 可以是将两者拼接, 或者将两者在相同尺度下加和或求平均。

示例性地, 手掌识别设备可将特征层 P1 进行 1 次卷积并调整通道后得到特征层 N1。将特征层 N1 进行下采样并与特征层 P2 结合, 再进行卷积并调整通道后得到特征层 N2; 将特征层 N2 进行下采样并与特征层 P3 结合, 再进行卷积并调整通道后得到特征层 N3; 将特征层 N3 进行下采样并与特征层 P4 结合, 再进行卷积并调整通道后得到特征层 N4。

手掌识别设备将特征层 N1-4 进行特征融合，得到样本图像融合特征 503。

可选地，卷积中的激活函数为 logistic 激活函数。

在一些实施例中，手掌识别设备将多个尺度的样本图像特征由小尺度到大尺度进行上采样，得到第一中间特征；手掌识别设备将多个尺度的样本图像特征由大尺度到小尺度进行下采样，得到第二中间特征；手掌识别设备将第一中间特征和第二中间特征进行特征融合，得到样本图像融合特征。

步骤 1304：通过掌部检测模型的预测网络，基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框。

样本预测掌部框是指在样本手掌图像中用于表示掌部所在位置的标识。

在一些实施例中，样本预测掌部框通过掌部检测模型确定，掌部检测模型包括预测网络；基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框，包括：将样本图像融合特征进行网格划分，得到样本图像融合特征的多个网格特征；将多个网格特征输入预测网络，对多个网格特征进行预测，得到样本预测掌部框。

示例性地，手掌识别设备将样本图像融合特征进行网格划分，得到网格化的样本图像融合特征；手掌识别设备将网格化的融合特征输入至预测网络中，由预测网络对每个网格的进行预测，得到样本预测掌部框。对多个网格特征进行预测，得到样本预测掌部框，具体可以综合对多个网格特征的预测结果，确定样本预测掌部框的位置。

在一些实施例中，手掌识别设备装配有红外摄像头，样本手掌图像处理方法还包括：获取对与样本手掌图像相同掌部采集的红外图像，红外图像是指红外摄像头基于红外光对掌部成像所得到的图像；从红外图像识别掌部区域。基于样本预测掌部框，从样本手掌图像中裁剪出样本预测掌部图像，包括：确定样本手掌图像中的样本预测掌部框，以及红外图像中掌部区域的交集；基于交集，从样本手掌图像中裁剪出样本预测掌部图像。

本实施例中，手掌识别设备还设置有红外摄像头，手掌识别设备获取同一实体手掌对应的红外图像，手掌识别设备对红外图像进行区域识别，确定红外图像上的掌部区域。红外图像是指红外摄像头基于红外光对手掌成像所得到的图像。

在一些实施例中，从红外图像识别掌部区域，包括：检测红外图像中的手指缝点；基于手指缝点，确定红外图像中的掌部区域。

示例性地，手掌识别设备对红外图像中的手指缝点进行检测，基于手指缝点确定红外图像上的掌部区域。

如图 6 所示出的手掌中手指缝点的示意图，手指缝点为食指与中指之间的第一手指缝点 601，或，手指缝点为中指与无名指之间的第二手指缝点 602，或，手指缝点为无名指与小指之间的第三手指缝点 603。

由于在样本手掌图像中的掌部区域可能存在于该样本手掌图像中的任一区域，为了能够确定掌部区域在该样本手掌图像中的位置，通过该样本手掌图像进行手指缝点检测，从而得到样本手掌图像的至少一个手指缝点，以便后续能够根据该至少一个手指缝点，确定掌部区域。

步骤 1305：基于样本预测掌部框，从样本手掌图像中裁剪出样本预测掌部图像。

样本预测掌部图像是指样本手掌图像中预测的有效识别区域，或，样本预测掌部图像是指样本手掌图像中预测的掌部所在的区域，或，样本预测掌部图像是指样本手掌图像中预测的能够用于识别的区域。

在一些实施例中，手掌识别设备获取红外图像中的至少三个手指缝点；顺序连接手指缝点，得到手指缝点连线；手掌识别设备基于手指缝点连线，在红外图像中裁剪得到掌部区域。

在一些实施例中，确定样本手掌图像中的样本预测掌部框，以及红外图像中掌部区域的交集，包括：获取样本手掌图像中样本预测掌部框的坐标参数，以及获取红外图像中掌部区域的坐标参数；基于样本预测掌部框的坐标参数与掌部区域的坐标参数，确定样本预

测掌部框和掌部区域的交集。

手掌识别设备获取样本手掌图像中的样本预测掌部框的坐标参数，以及获取红外图像的中的掌部区域的坐标参数；手掌识别设备基于样本预测掌部框的坐标参数与掌部区域的坐标参数的交集，在样本手掌图像中裁剪得到掌部对应的样本预测掌部图像。

步骤 1306：根据样本掌部图像及样本预测掌部图像，计算损失函数值。

示例性地，手掌识别设备基于样本掌部图像及样本预测掌部图像，计算损失函数值。

步骤 1307：基于损失函数值对掌部检测模型的模型参数进行更新。

模型参数更新是指对掌部检测模型里面的网络参数进行更新，或对模型里面的各个网络模块的网络参数进行更新，或对模型里面的各个网络层的网络参数进行更新，但不限于此，本申请实施例对此不作限定。

掌部检测模型的模型参数包括主干网络的网络参数、颈部网络的网络参数及预测网络的网络参数中的至少一种。

在一些实施例中，掌部检测模型的模型参数更新包括更新掌部检测模型中的所有网络模块的网络参数，或，固定掌部检测模型中的部分网络模块的网络参数，仅更新剩余部分的网络模块的网络参数。比如，对掌部检测模型的模型参数进行更新时，固定掌部检测模型中的主干网络的网络参数，仅对颈部网络的网络参数及预测网络的网络参数进行更新。

基于损失函数值，将损失函数值作为训练指标对掌部检测模型中的主干网络的网络参数、颈部网络的网络参数及预测网络的网络参数进行更新，直至损失函数值发生收敛，从而得到训练完成的掌部检测模型。

损失函数值发生收敛是指损失函数值不再发生改变，或，掌部检测模型的训练时相邻两次迭代之间的误差差别小于预设值，或，掌部检测模型的训练次数达到预设次数中的至少一种，但不限于此，本申请实施例对此不作限定。

可选的，训练满足的目标条件可以为初始模型的训练迭代次数达到目标次数，技术人员可以预先设定训练迭代次数。或者，训练满足的目标条件可以为损失值满足目标阈值条件，但不限于此，本申请实施例对此不作限定。

综上，本实施例提供的方法，通过对获取的样本手掌图像进行特征提取，得到样本手掌图像对应的多个尺度的样本图像特征；将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征；基于样本图像融合特征进行预测，得到标识样本手掌图像中掌部的样本预测掌部框；基于样本预测掌部框，在样本手掌图像中裁剪得到掌部对应的样本预测掌部图像；基于样本掌部图像及样本预测掌部图像，计算损失函数值；基于损失函数值对掌部检测模型的模型参数进行更新，使得训练好的掌部检测模型能够具备更高的掌部框预测精度，从而得到更加准确的掌部框。

示意性的，本申请实施例提供的手掌图像处理方法的应用场景包括但不限于以下场景：例如，手掌图像识别支付场景下：

商户的手掌识别设备通过拍摄用户的实体手掌，获取到该用户的手掌图像，采用本申请实施例提供的手掌图像处理方法，确定该手掌图像的目标用户标识，将该目标用户标识对应的资源账户中的部分资源，转入到商户资源账户中，实现通过掌部自动支付。

又如，跨设备支付场景下：

用户可以在家或其他私密空间使用个人手机完成身份注册，将该用户的账号与该用户的手掌图像进行绑定，之后可以到店内设备上对该用户的手掌图像进行识别，确定该用户的账号，通过该账号直接支付。

再例如，上班打卡场景下：

手掌识别设备通过拍摄用户的实体手掌，获取到该用户的手掌图像，采用本申请实施例提供的手掌图像处理方法，确定该手掌图像的目标用户标识，为该目标用户标识建立打卡标记，确定该目标用户标识在当前时间已完成上班打卡。

当然，除了应用于上述场景外，本申请实施例提供方法还可以应用于其他需要手掌图

像处理的场景，本申请实施例并不对具体的应用场景进行限定。

图 14 示出了本申请一个示例性实施例提供的手掌图像处理装置的结构示意图。该装置可以通过软件、硬件或者两者的结合实现成为手掌识别设备的全部或一部分，该装置包括：

特征提取模块 1401，用于获取手掌图像，对手掌图像提取特征，得到多个尺度的图像特征。

特征融合模块 1402，用于将多个尺度的图像特征融合，得到图像融合特征。

预测模块 1403，用于基于图像融合特征，确定标识手掌图像中掌部的掌部框。

裁剪模块 1404，用于基于掌部框，从手掌图像中裁剪出掌部图像。

加密模块 1405，用于将掌部图像进行图像加密，得到掌部图像加密数据，掌部图像加密数据用于传输至服务器，由服务器基于掌部图像加密数据进行身份验证。

其中，特征提取模块 1401 可以是掌部检测模型的主干网络，特征融合模块 1402 可以是掌部检测模型的颈部网络，预测模块 1403 可以是掌部检测模型的预测网络的至少一部分，裁剪模块 1404 可以是预测网络的一部分，也可以是掌部检测模型中独立的模块，也可以是掌部检测模型之外独立的模块。

在一些实施例中，特征提取模块 1401 是掌部检测模型的主干网络，用于对手掌图像进行切片操作，得到多个尺度的切片图像；对多个尺度的切片图像分别提取特征，得到多个尺度的图像特征。

在一些实施例中，特征提取模块 1401 是掌部检测模型的主干网络，用于基于手掌图像，确定多个尺度中最大尺度的切片图像；将最大尺度的切片图像作为首层，开始逐层下采样，获得包括首层的切片图像在内的多个尺度的切片图像。

在一些实施例中，特征提取模块 1401 是掌部检测模型的主干网络，用于将手掌图像中的像素点，在多个尺度下采样并拼接，得到多个尺度的切片图像；其中，不同尺度的切片图像中相邻的两个像素点，在手掌图像中采样处的两个像素点间隔的像素点数量不同。

在一些实施例中，特征融合模块 1402 是掌部检测模型的颈部网络，用于将多个尺度的图像特征进行特征融合，得到图像融合特征。

在一些实施例中，特征融合模块 1402 是掌部检测模型的颈部网络，用于基于多个尺度的图像特征进行特征融合，得到多个尺度的第一中间特征；基于多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征；将多个尺度的第二中间特征进行特征融合，得到图像融合特征。

在一些实施例中，多个尺度为 N 个尺度， N 为大于 1 的正整数，特征融合模块 1402 是掌部检测模型的颈部网络，用于基于按尺度正序第 1 位的图像特征，确定按尺度正序第 1 位的第一中间特征；从按尺度正序第 1 位的第一中间特征开始，逐层上采样，将按尺度正序第 n 位的第一中间特征上采样的结果，与按尺度正序第 $n+1$ 位的图像特征融合，获得按尺度正序第 $n+1$ 位的第一中间特征； n 为 1 到 $N-1$ 的正整数；按尺度正序是按尺度从小到大的顺序；当获得第 N 层的第一中间特征，确定从第 1 层到第 N 层的多个尺度的第一中间特征。

在一些实施例中，多个尺度为 N 个尺度， N 为大于 1 的正整数，特征融合模块 1402 是掌部检测模型的颈部网络，用于基于按尺度倒序第 1 位的第一中间特征，确定按尺度倒序第 1 位的第二中间特征；从按尺度倒序第 1 位的第二中间特征开始，逐层下采样，将按尺度倒序第 m 位的第二中间特征下采样的结果，与按尺度倒序第 $m+1$ 位的第一中间特征融合，获得按尺度倒序第 $n+1$ 位的第二中间特征； m 为 1 到 $N-1$ 的整数；按尺度倒序是按尺度从大到小的顺序；当获得第 N 层的第一中间特征，确定从第 1 层到第 N 层的多个尺度的第一中间特征。

在一些实施例中，预测模块 1403 是掌部检测模型的预测网络的至少一部分，用于将图像融合特征进行网格划分，得到图像融合特征的多个网格特征；基于多个网格特征进行预

测，得到掌部框。

在一些实施例中，手掌识别设备装配有红外摄像头，获取模块 1406，用于获取对与手掌图像相同掌部采集的红外图像，红外图像是指红外摄像头基于红外光对掌部成像所得到的图像。

区域识别模块 1407，用于从红外图像识别掌部区域。

在一些实施例中，裁剪模块 1404，用于确定手掌图像中的掌部框，以及红外图像中掌部区域的交集；基于交集，从手掌图像中裁剪出掌部图像。

在一些实施例中，区域识别模块 1407，用于检测红外图像中的手指缝点；基于手指缝点，确定红外图像中的掌部区域。

在一些实施例中，区域识别模块 1407，用于获取红外图像中的至少三个手指缝点；顺序连接手指缝点，得到手指缝点连线；基于手指缝点连线，在红外图像中裁剪得到掌部区域。

在一些实施例中，裁剪模块 1404，用于获取手掌图像中掌部框的坐标参数，以及获取红外图像中掌部区域的坐标参数；基于掌部框的坐标参数与掌部区域的坐标参数，确定掌部框和掌部区域的交集。

图 15 示出了本申请一个示例性实施例提供的掌部检测模型的训练装置的结构示意图。该装置可以通过软件、硬件或者两者的结合实现成为手掌识别设备的全部或一部分，该装置包括掌部检测模型、获取模块 1501、计算模块 1506 和更新模块 1507，掌部检测模型包括特征提取模块 1502、特征融合模块 1503 和预测模块 1504，还可以包括裁剪模块 1505。其中，特征提取模块 1502 可以是掌部检测模型的主干网络，特征融合模块 1503 可以是掌部检测模型的颈部网络，预测模块 1504 可以是预测网络的至少一部分，裁剪模块 1505 可以是预测网络的一部分，也可以是掌部检测模型中独立的模块，也可以是掌部检测模型之外独立的模块。

获取模块 1501，用于获取样本手掌图像和样本掌部图像。

特征提取模块 1502，用于对样本手掌图像提取特征，得到多个尺度的样本图像特征。

特征融合模块 1503，用于将多个尺度的样本图像特征融合，得到样本图像融合特征。

预测模块 1504，用于基于样本图像融合特征，确定标识样本手掌图像中掌部的样本预测掌部框。

裁剪模块 1505，用于基于样本预测掌部框，从样本手掌图像中裁剪出样本预测掌部图像。

计算模块 1506，用于根据样本掌部图像及样本预测掌部图像，计算损失函数值。

更新模块 1507，用于基于损失函数值对掌部检测模型的模型参数进行更新。

在一些实施例中，特征提取模块 1502 是掌部检测模型的主干网络，用于对样本手掌图像进行切片操作，得到多个尺度的切片图像；对多个尺度的切片图像分别提取特征，得到多个尺度的样本图像特征。

在一些实施例中，特征提取模块 1502 是掌部检测模型的主干网络，用于基于样本手掌图像，确定多个尺度中最大尺度的切片图像；将最大尺度的切片图像作为首层，开始逐层下采样，获得包括首层的切片图像在内的多个尺度的切片图像。

在一些实施例中，特征提取模块 1502 是掌部检测模型的主干网络，用于将样本手掌图像中的像素点，在多个尺度下采样并拼接，得到多个尺度的切片图像；其中，不同尺度的切片图像中相邻的两个像素点，在样本手掌图像中采样处的两个像素点间隔的像素点数量不同。

在一些实施例中，特征融合模块 1503 是掌部检测模型的颈部网络，用于将多个尺度的样本图像特征进行特征融合，得到样本图像融合特征。

在一些实施例中，特征融合模块 1503 是掌部检测模型的颈部网络，用于将多个尺度的样本图像特征输入至颈部网络；基于多个尺度的样本图像特征进行特征融合，得到多个尺

度的第一中间特征；基于多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征；将多个尺度的第二中间特征进行特征融合，得到样本图像融合特征。

在一些实施例中，多个尺度为 N 个尺度， N 为大于1的正整数，特征融合模块1503是掌部检测模型的颈部网络，用于基于按尺度正序第1位的样本图像特征，确定按尺度正序第1位的第一中间特征；从按尺度正序第1位的第一中间特征开始，逐层上采样，将按尺度正序第 n 位的第一中间特征上采样的结果，与按尺度正序第 $n+1$ 位的样本图像特征融合，获得按尺度正序第 $n+1$ 位的第一中间特征； n 为1到 $N-1$ 的正整数；按尺度正序是按尺度从小到大的顺序；当获得第 N 层的第一中间特征，确定从第1层到第 N 层的多个尺度的第一中间特征。

在一些实施例中，多个尺度为 N 个尺度， N 为大于1的正整数，特征融合模块1503，用于基于按尺度倒序第1位的第一中间特征，确定按尺度倒序第1位的第二中间特征；从按尺度倒序第1位的第二中间特征开始，逐层下采样，将按尺度倒序第 m 位的第二中间特征下采样的结果，与按尺度倒序第 $m+1$ 位的第一中间特征融合，获得按尺度倒序第 $n+1$ 位的第二中间特征； m 为1到 $N-1$ 的整数；按尺度倒序是按尺度从大到小的顺序；当获得第 N 层的第一中间特征，确定从第1层到第 N 层的多个尺度的第一中间特征。

在一些实施例中，预测模块1504是掌部检测模型的预测网络的至少一部分，用于将样本图像融合特征进行网格划分，得到样本图像融合特征的多个网格特征；对多个网格特征进行预测，得到样本预测掌部框。

在一些实施例中，手掌识别设备装配有红外摄像头，获取模块1501用于获取对与样本手掌图像相同掌部采集的红外图像，红外图像是指红外摄像头基于红外光对掌部成像所得到的图像。

区域识别模块1508，用于从红外图像识别掌部区域。

裁剪模块1505，用于确定样本手掌图像中的样本预测掌部框，以及红外图像中掌部区域的交集；基于交集，从样本手掌图像中裁剪出掌部图像。

在一些实施例中，区域识别模块1508用于检测红外图像中的手指缝点；基于手指缝点，确定红外图像中的掌部区域。

在一些实施例中，裁剪模块1505用于获取样本手掌图像中样本预测掌部框的坐标参数，以及获取红外图像中掌部区域的坐标参数；基于样本预测掌部框的坐标参数与掌部区域的坐标参数，确定样本预测掌部框和掌部区域的交集。

在一些实施例中，区域识别模块1508，用于对红外图像中的手指缝点进行检测，基于手指缝点确定红外图像上的掌部区域。

在一些实施例中，区域识别模块1508，用于获取红外图像中的至少三个手指缝点；顺序连接手指缝点，得到手指缝点连线；基于手指缝点连线，在红外图像中裁剪得到掌部区域。

图16示出了本申请一示例性实施例示出的计算机设备1600的结构框图。该计算机设备可以实现为本申请上述方案中的服务器。图像计算机设备1600包括中央处理单元（Central Processing Unit，CPU）1601、包括随机存取存储器（Random Access Memory，RAM）1602和只读存储器（Read-Only Memory，ROM）1603的系统存储器1604，以及连接系统存储器1604和中央处理单元1601的系统总线1605。图像计算机设备1600还包括用于存储操作系统1609、应用程序1610和其他程序模块1611的大容量存储设备1606。

大容量存储设备1606通过连接到系统总线1605的大容量存储控制器（未示出）连接到中央处理单元1601。大容量存储设备1606及其相关联的计算机可读介质为图像计算机设备1600提供非易失性存储。也就是说，大容量存储设备1606可以包括诸如硬盘或者只读光盘（Compact Disc Read-Only Memory，CD-ROM）驱动器之类的计算机可读介质（未示出）。

不失一般性，计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质

包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括 RAM、可擦除可编程只读存储器（Erasable Programmable Read Only Memory, EPROM）、电子抹除式可复写只读存储器（Electrically-Erasable Programmable Read-Only Memory, EEPROM）闪存或其他固态存储其技术，CD-ROM、数字多功能光盘（Digital Versatile Disc, DVD）或其他光学存储、磁带盒、磁带、磁盘存储或其他磁性存储设备。当然，本领域技术人员可知计算机存储介质不局限于上述几种。上述的系统存储器 1604 和大容量存储设备 1606 可以统称为存储器。

根据本公开的各种实施例，图像计算机设备 1600 还可以通过诸如因特网等网络连接到网络上的远程计算机运行。也即图像计算机设备 1600 可以通过连接在系统总线 1605 上的网络接口单元 1607 连接到网络 1608，或者说，也可以使用网络接口单元 1607 来连接到其他类型的网络或远程计算机系统（未示出）。

存储器还包括至少一段计算机可读指令，至少一段计算机可读指令存储于存储器中，中央处理器 1601 通过执行该至少一段程序来实现上述各个实施例所示的手掌图像处理方法，或，掌部检测模型的训练方法中的全部或部分步骤。

本申请实施例还提供一种计算机设备，该计算机设备包括处理器和存储器，该存储器中存储有至少一条程序，该至少一条程序由处理器加载并执行以实现上述各方法实施例提供的手掌图像处理方法，或，掌部检测模型的训练方法。

本申请实施例还提供一种计算机可读存储介质，该存储介质中存储有至少一条计算机可读指令，该至少一条计算机可读指令由处理器加载并执行以实现上述各方法实施例提供的手掌图像处理方法，或，掌部检测模型的训练方法。

本申请实施例还提供一种计算机程序产品，计算机程序产品包括计算机可读指令，计算机可读指令存储在计算机可读存储介质中；计算机可读指令由计算机设备的处理器从计算机可读存储介质读取并执行，使得计算机设备执行以实现上述各方法实施例提供的手掌图像处理方法，或，掌部检测模型的训练方法。

可以理解的是，在本申请的具体实施方式中，涉及到的数据，手掌图像，历史数据，以及画像等与用户身份或特性相关的用户数据处理等相关的数据，当本申请以上实施例运用到具体产品或技术中时，需要获得用户许可或者同意，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

应当理解的是，在本文中提及的“多个”是指两个或两个以上。“和/或”，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。字符“/”一般表示前后关联对象是一种“或”的关系。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，上述提到的存储介质可以是只读存储器，磁盘或光盘等。

以上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申请的几种实施方式，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，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求

- 1.一种手掌图像处理方法，由手掌识别设备执行，包括：
获取手掌图像，对所述手掌图像提取特征，得到多个尺度的图像特征；
将所述多个尺度的图像特征融合，得到图像融合特征；
基于所述图像融合特征，确定标识所述手掌图像中掌部的掌部框；
基于所述掌部框，从所述手掌图像中裁剪出掌部图像；及
将所述掌部图像进行图像加密，得到掌部图像加密数据，所述掌部图像加密数据用于传输至服务器，由所述服务器基于所述掌部图像加密数据进行身份识别。
- 2.根据权利要求1所述的方法，所述掌部框通过掌部检测模型预测得到，所述掌部检测模型包括主干网络；
所述对所述手掌图像提取特征，得到多个尺度的图像特征，包括：
通过所述主干网络对所述手掌图像进行切片操作，得到多个尺度的切片图像；
通过所述主干网络对所述多个尺度的切片图像分别提取特征，得到多个尺度的图像特征。
- 3.根据权利要求2所述的方法，所述通过所述主干网络对所述手掌图像进行切片操作，得到多个尺度的切片图像，包括：
通过所述主干网络，基于所述手掌图像，确定所述多个尺度中最大尺度的切片图像；
通过所述主干网络，将最大尺度的切片图像作为首层，开始逐层下采样，获得包括首层的切片图像在内的多个尺度的切片图像。
- 4.根据权利要求2所述的方法，所述通过所述主干网络对所述手掌图像进行切片操作，得到多个尺度的切片图像，包括：
通过所述主干网络，将所述手掌图像中的像素点，在多个尺度下采样并拼接，得到多个尺度的切片图像；
其中，不同尺度的切片图像中相邻的两个像素点，在所述手掌图像中采样处的两个像素点间隔的像素点数量不同。
- 5.根据权利要求1至3任一项所述的方法，所述掌部框通过掌部检测模型预测得到，所述掌部检测模型包括颈部网络；
所述将所述多个尺度的图像特征融合，得到图像融合特征，包括：
将所述多个尺度的图像特征输入至所述颈部网络进行特征融合，得到图像融合特征。
- 6.根据权利要求5所述的方法，所述将所述多个尺度的图像特征输入至所述颈部网络进行特征融合，得到图像融合特征，包括：
将所述多个尺度的图像特征输入至所述颈部网络；
通过所述颈部网络，基于所述多个尺度的图像特征进行特征融合，得到多个尺度的第一中间特征；
通过所述颈部网络，基于所述多个尺度的第一中间特征进行多个尺度的特征融合，得到多个尺度的第二中间特征；
通过所述颈部网络，将所述多个尺度的第二中间特征进行特征融合，得到图像融合特征。
- 7.根据权利要求6所述的方法，所述多个尺度为N个尺度，N为大于1的正整数，所述基于所述多个尺度的图像特征，提取多个尺度的第一中间特征，包括：
基于按尺度正序第1位的图像特征，确定按尺度正序第1位的第一中间特征；
从按尺度正序第1位的第一中间特征开始，逐层上采样，将按尺度正序第n位的第一中间特征上采样的结果，与按尺度正序第n+1位的图像特征融合，获得按尺度正序第n+1位的第一中间特征；n为1到N-1的正整数；按尺度正序是按尺度从小到大的顺序；
当获得第N层的第一中间特征，确定从第1层到第N层的多个尺度的第一中间特征。

8.根据权利要求6或7所述的方法,所述多个尺度为 N 个尺度, N 为大于1的正整数,所述基于所述多个尺度的第一中间特征进行多个尺度的特征融合,得到多个尺度的第二中间特征,包括:

基于按尺度倒序第1位的第一中间特征,确定按尺度倒序第1位的第二中间特征;

从按尺度倒序第1位的第二中间特征开始,逐层下采样,将按尺度倒序第 m 位的第二中间特征下采样的结果,与按尺度倒序第 $m+1$ 位的第一中间特征融合,获得按尺度倒序第 $n+1$ 位的第二中间特征; m 为1到 $N-1$ 的整数;按尺度倒序是按尺度从大到小的顺序;

当获得第 N 层的第一中间特征,确定从第1层到第 N 层的多个尺度的第一中间特征。

9.根据权利要求1至8任一项所述的方法,所述掌部框通过掌部检测模型确定,所述掌部检测模型包括预测网络;

所述基于所述图像融合特征,确定标识所述手掌图像中掌部的掌部框,包括:

将所述图像融合特征进行网格划分,得到图像融合特征的多个网格特征;

将所述多个网格特征输入所述预测网络,对所述多个网格特征进行预测,得到所述掌部框。

10.根据权利要求1至9任一所述的方法,所述手掌识别设备装配有红外摄像头,所述方法还包括:

获取对与所述手掌图像相同掌部采集的红外图像,所述红外图像是指所述红外摄像头基于红外光对所述掌部成像所得到的图像;

从所述红外图像识别掌部区域;

所述基于所述掌部框,从所述手掌图像中裁剪出掌部图像,包括:

确定所述手掌图像中的所述掌部框,以及所述红外图像中所述掌部区域的交集;

基于所述交集,从所述手掌图像中裁剪出掌部图像。

11.根据权利要求10所述的方法,所述从所述红外图像识别掌部区域,包括:

检测所述红外图像中的手指缝点;

基于所述手指缝点,确定所述红外图像中的所述掌部区域。

12.根据权利要求10或11所述的方法,所述确定所述手掌图像中的所述掌部框,以及所述红外图像中所述掌部区域的交集,包括:

获取所述手掌图像中所述掌部框的坐标参数,以及获取所述红外图像中所述掌部区域的坐标参数;

基于所述掌部框的坐标参数与所述掌部区域的坐标参数,确定所述掌部框和所述掌部区域的交集。

13.一种掌部检测模型的训练方法,由计算机设备执行,包括:

获取样本手掌图像和样本掌部图像;

对所述样本手掌图像提取特征,得到多个尺度的样本图像特征;

将所述多个尺度的样本图像特征融合,得到样本图像融合特征;

基于所述样本图像融合特征,确定所述标识样本手掌图像中掌部的样本预测掌部框;

基于所述样本预测掌部框,从所述样本手掌图像中裁剪出样本预测掌部图像;

根据所述样本掌部图像及所述样本预测掌部图像,计算损失函数值;及

基于所述损失函数值对所述掌部检测模型的模型参数进行更新。

14.一种手掌图像处理装置,包括:

特征提取模块,用于获取手掌图像,对所述手掌图像提取特征,得到多个尺度的图像特征;

特征融合模块,用于将所述多个尺度的图像特征融合,得到图像融合特征;

预测模块,用于基于所述图像融合特征,确定标识所述手掌图像中掌部的掌部框;

裁剪模块,用于基于所述掌部框,从所述手掌图像中裁剪出掌部图像;及

加密模块,用于将所述掌部图像进行图像加密,得到掌部图像加密数据,所述掌部图

像加密数据用于传输至服务器，由所述服务器基于所述掌部图像加密数据进行身份验证。

15.一种掌部检测模型的训练装置，包括掌部检测模型，还包括获取模块、裁剪模块、计算模块和更新模块，所述掌部检测模型包括特征提取模块、特征融合模块和预测模块；

所述获取模块，用于获取样本手掌图像和样本掌部图像；

所述特征提取模块，用于对所述样本手掌图像提取特征，得到多个尺度的样本图像特征；

所述特征融合模块，用于将所述多个尺度的样本图像特征融合，得到样本图像融合特征；

所述预测模块，用于基于所述样本图像融合特征，确定所述标识样本手掌图像中掌部的样本预测掌部框；

所述裁剪模块，用于基于所述样本预测掌部框，从所述样本手掌图像中裁剪出样本预测掌部图像；

所述计算模块，用于根据所述样本掌部图像及所述样本预测掌部图像，计算损失函数值；

所述更新模块，用于基于所述损失函数值对掌部检测模型的模型参数进行更新。

16.一种计算机设备，包括处理器和存储器，所述存储器中存储有至少一条计算机可读指令，至少一条所述计算机可读指令由所述处理器加载并执行以实现如权利要求 1 至 13 中任一项所述的方法。

17.一种计算机存储介质，所述计算机可读存储介质中存储有至少一条计算机可读指令，至少一条计算机可读指令由处理器加载并执行以实现如权利要求 1 至 13 中任一项所述的方法。

18.一种计算机程序产品，所述计算机程序产品包括计算机可读指令，所述计算机可读指令由处理器加载并执行以实现如权利要求 1 至 13 中任一项所述的方法。