

(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
2020 年 1 月 9 日 (09.01.2020)



(10) 国际公布号
WO 2020/006961 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01)

中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。

(21) 国际申请号: PCT/CN2018/116334

(22) 国际申请日: 2018 年 11 月 20 日 (20.11.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201810715195.6 2018年7月3日 (03.07.2018) CN

(71) 申请人: 北京字节跳动网络技术有限公司 (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市石景山区实兴大街30号院3号楼2层B-0035房间, Beijing 100041 (CN)。

(72) 发明人: 周恺卉(ZHOU, Kaihui); 中国北京市海淀区知春路63号中国卫星通信大厦今日头条小邮局, Beijing 100190 (CN)。 王长虎(WANG, Changhu);

(74) 代理人: 北京英赛嘉华知识产权代理有限公司 (INSIGHT INTELLECTUAL PROPERTY LIMITED); 中国北京市海淀区知春路甲48号盈都大厦A座19A, Beijing 100098 (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,

(54) Title: IMAGE EXTRACTION METHOD AND DEVICE

(54) 发明名称: 用于提取图像的方法和装置

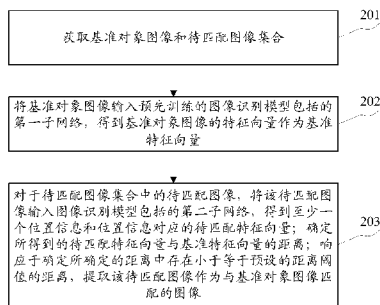


图 2

- 201 ACQUIRE A REFERENCE OBJECT IMAGE AND A SET OF IMAGES TO BE MATCHED
- 202 INPUT THE REFERENCE OBJECT IMAGE INTO A FIRST SUB-NETWORK COMPRISED IN A PRE-TRAINED IMAGE RECOGNITION MODEL, AND ACQUIRE A FEATURE VECTOR OF THE REFERENCE OBJECT IMAGE AS A REFERENCE FEATURE VECTOR
- 203 FOR IMAGES TO BE MATCHED IN THE SET OF IMAGES TO BE MATCHED: INPUT THE IMAGE TO BE MATCHED INTO A SECOND SUB-NETWORK COMPRISED IN THE IMAGE RECOGNITION MODEL, SO AS TO ACQUIRE AT LEAST ONE PIECE OF POSITION INFORMATION AND A FEATURE VECTOR TO BE MATCHED CORRESPONDING TO THE POSITION INFORMATION; DETERMINE THE DISTANCE BETWEEN THE ACQUIRED FEATURE VECTOR TO BE MATCHED AND THE REFERENCE FEATURE VECTOR; AND IN RESPONSE TO DETERMINING THAT THE DETERMINED DISTANCES COMPRISE A DISTANCE LESS THAN OR EQUAL TO A PRESET DISTANCE THRESHOLD, EXTRACT THE IMAGE TO BE MATCHED AS AN IMAGE MATCHING THE REFERENCE OBJECT IMAGE

(57) Abstract: Embodiments of the present application disclose an image extraction method and device. A specific embodiment of the method comprises: acquiring a reference object image and a set of images to be matched; and inputting the reference object image into a first sub-network comprised in a pre-trained image recognition model, and acquiring a feature vector of the reference object image as a reference feature vector, wherein the following extraction steps are executed on images to be matched in the set of images to be matched: inputting the image to be matched into a second sub-network comprised in the image recognition model, so as to acquire at least one piece of position information and a feature vector to be matched corresponding to the position information; determining the distance between the acquired feature vector to be matched and the reference feature vector; and in response to determining that the determined distances comprise a distance less than or equal to a preset distance threshold, extracting the image to be matched as an image matching the reference object image. The embodiments improve the flexibility of image recognition, and enrich the means of image recognition.

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))。

(57) 摘要：本申请实施例公开了用于提取图像的方法和装置。该方法的一具体实施方式包括：获取基准对象图像和待匹配图像集合；将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量；对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。该实施方式提高了图像识别的灵活性，丰富了图像识别的手段。

用于提取图像的方法和装置

本专利申请要求于 2018 年 7 月 3 日提交的、申请号为 201810715195.6、申请人为北京字节跳动网络技术有限公司、发明名称为“用于提取图像的方法和装置”的中国专利申请的优先权，该申请的全文以引用的方式并入本申请中。

技术领域

本申请实施例涉及计算机技术领域，具体涉及用于提取图像的方法和装置。

背景技术

目前，图像识别技术的应用领域越来越广泛，利用图像识别模型对图像进行识别，是图像识别技术的一种常用手段。图像识别模型通常是利用大量的训练样本进行训练得到的模型，为了使图像识别模型能够识别出某图像中的目标图像（例如水印图像、人物图像、物体图像等），通常需要利用包含目标图像的样本图像进行训练以得到图像识别模型。

发明内容

本申请实施例提出了用于提取图像的方法和装置。

第一方面，本申请实施例提供了一种用于提取图像的方法，该方法包括：获取基准对象图像和待匹配图像集合；将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量；对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信

息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。

在一些实施例中，提取步骤还包括：确定小于等于距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

在一些实施例中，提取步骤还包括：基于输出的位置信息和该待匹配图像，生成包括位置标记的匹配后图像，其中，位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。

在一些实施例中，第二子网络包括用于将特征向量变换到目标维度的维度变换层；以及将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个待匹配特征向量，包括：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个与基准特征向量的维度相同的待匹配特征向量。

在一些实施例中，图像识别模型通过以下步骤训练得到：获取训练样本集，其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像的标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置；从训练样本集中选取训练样本，执行如下训练步骤：将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量；从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量；基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成；响应于确定训练完成，将初始模型确定为图像识别模型。

在一些实施例中，基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成，包括：根据预设的权重值，将第一损失

值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。

在一些实施例中，训练得到图像识别模型的步骤还包括：响应于确定初始模型未训练完成，调整初始模型的参数，以及从训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行训练步骤。

第二方面，本申请实施例提供了一种用于提取图像的装置，该装置包括：获取单元，被配置成获取基准对象图像和待匹配图像集合；生成单元，被配置成将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量；提取单元，被配置成对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。

在一些实施例中，提取单元包括：输出模块，被配置成确定小于等于距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

在一些实施例中，提取单元还包括：生成模块，被配置成基于输出的位置信息和该待匹配图像，生成包括位置标记的匹配后图像，其中，位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。

在一些实施例中，第二子网络包括用于将特征向量变换到目标维度的维度变换层；以及提取单元进一步被配置成：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个与基准特征向量的维度相同的待匹配特征向量。

在一些实施例中，图像识别模型通过以下步骤训练得到：获取训

训练样本集，其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像的标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置；从训练样本集中选取训练样本，执行如下训练步骤：将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量；从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量；基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成；响应于确定训练完成，将初始模型确定为图像识别模型。

在一些实施例中，基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成，包括：根据预设的权重值，将第一损失值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。

在一些实施例中，训练得到图像识别模型的步骤还包括：响应于确定初始模型未训练完成，调整初始模型的参数，以及从训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行训练步骤。

第三方面，本申请实施例提供了一种电子设备，该电子设备包括：一个或多个处理器；存储装置，其上存储有一个或多个程序；当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如第一方面中任一实现方式描述的方法。

第四方面，本申请实施例提供了一种计算机可读介质，其上存储有计算机程序，该计算机程序被处理器执行时实现如第一方面中任一实现方式描述的方法。

本申请实施例提供的用于提取图像的方法和装置，通过利用预先训练的图像识别模型，得到基准图像的基准特征向量和待匹配图像的

至少一个待匹配特征向量，再通过比较基准特征向量和待匹配特征向量的距离，得到与基准图像匹配的图像，从而实现了在训练图像识别模型所需的训练样本不包括基准图像的情况下，利用图像识别模型提取出与基准图像匹配的图像，提高了图像识别的灵活性，丰富了图像识别的手段。

附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本申请的其它特征、目的和优点将会变得更明显：

图 1 是本申请的一个实施例可以应用于其中的示例性系统架构图；

图 2 是根据本申请的用于提取图像的方法的一个实施例的流程图；

图 3 是根据本申请的用于提取图像的方法的训练得到图像识别模型的流程图；

图 4 是根据本申请的用于提取图像的方法的一个应用场景的示意图；

图 5 是根据本申请的用于提取图像的方法的又一个实施例的流程图；

图 6 是根据本申请的用于提取图像的装置的一个实施例的结构示意图；

图 7 是适于用来实现本申请实施例的电子设备的计算机系统的结构示意图。

具体实施方式

下面结合附图和实施例对本申请作进一步的详细说明。可以理解的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

需要说明的是，在不冲突的情况下，本申请中的实施例及实施例

中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

图 1 示出了可以应用本申请实施例的用于提取图像的方法或用于提取图像的装置的示例性系统架构 100。

如图 1 所示，系统架构 100 可以包括终端设备 101、102、103，网络 104 和服务器 105。网络 104 用以在终端设备 101、102、103 和服务器 105 之间提供通信链路的介质。网络 104 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备 101、102、103 通过网络 104 与服务器 105 交互，以接收或发送消息等。终端设备 101、102、103 上可以安装有各种通讯客户端应用，例如图像处理类应用、拍摄类应用、社交平台软件等。

终端设备 101、102、103 可以是硬件，也可以是软件。当终端设备 101、102、103 为硬件时，可以是具有显示屏的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。当终端设备 101、102、103 为软件时，可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块（例如用来提供分布式服务的软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

服务器 105 可以是提供各种服务的服务器，例如对终端设备 101、102、103 上的各种应用提供支持的后台服务器。后台服务器可以对获取的图像进行分析等处理，并将处理结果（例如提取的与基准图像匹配的图像）输出。

需要说明的是，本申请实施例所提供的用于提取图像的方法可以由服务器 105 执行，也可以由终端设备 101、102、103 执行。相应地，用于提取图像的装置可以设置于服务器 105 中，也可以设置于终端设备 101、102、103 中。

需要说明的是，服务器可以是硬件，也可以是软件。当服务器为硬件时，可以实现成多个服务器组成的分布式服务器集群，也可以实现成单个服务器。当服务器为软件时，可以实现成多个软件或软件模

块（例如用来提供分布式服务的软件或软件模块），也可以实现成单个软件或软件模块。在此不做具体限定。

应该理解，图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

继续参考图 2，示出了根据本申请的用于提取图像的方法的一个实施例的流程 200。该用于提取图像的方法，包括以下步骤：

步骤 201，获取基准对象图像和待匹配图像集合。

在本实施例中，用于提取图像的方法的执行主体（例如如图 1 所示的服务器或终端设备）可以通过有线连接方式或者无线连接方式从远程或者从本地获取基准对象图像和待匹配图像集合。其中，基准对象图像可以是待将其与其他图像进行对比的图像，基准对象图像可以是表征某对象的图像。对象可以是各种事物，例如水印、标志、人脸、物体等。待匹配图像集合可以是预先存储的某类图像（例如包含商标的图像）的集合。

步骤 202，将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量。

在本实施例中，上述执行主体可以将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量。其中，第一子网络用于表征图像和图像的特征向量的对应关系。在本实施例中，图像识别模型可以是基于机器学习技术而创建的各种神经网络模型。该神经网络模型可以具有各种神经网络（例如 DenseBox、VGGNet、ResNet、SegNet 等）的结构。上述基准特征向量可以由神经网络模型包括的第一子网络（例如由神经网络模型包括的某个或某些卷积层组成的网络）提取的、表征图像的特征（例如形状、颜色、纹理等特征）的数据组成的向量。

步骤 203，对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存

在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图片。

在本实施例中，对于待匹配图像集合中的每个待匹配图像，上述执行主体可以对该待匹配图像执行如下提取步骤：

步骤 2031，将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量。其中，第二子网络用于表征图像与图像的位置信息、图像的待匹配特征向量的对应关系。位置信息用于表征待匹配特征向量对应的区域图像在该待匹配图像中的位置。待匹配特征向量是该待匹配图像包括的区域图像的特征向量。在本实施例中，第二子网络（例如由神经网络模型包括的某个或某些卷积层组成的网络）可以根据所确定的至少一个位置信息，从待匹配图像中确定每个位置信息表征的区域图像，以及确定每个区域图像的特征向量。区域图像可以是表征某个对象（例如水印、标志等）的图像。可选地，位置信息可以包括坐标信息和标识信息。其中，坐标信息（例如区域图像角点坐标、区域图像的大小等）用于指示区域图像在待匹配图像中的位置，标识信息（例如区域图像的序号、区域图像的类别等）用于标识区域图像。作为示例，假设某待匹配图像中包括两个水印图像，第二子网络确定的位置信息分别为“(1, x1, y1, w1, h1)”和“(2, x2, y2, w2, h2)”，其中，1、2分别为两个水印图像的序号，(x1, y1)、(x2, y2)分别为两个水印图像的左上角点的坐标，w1、w2分别为两个水印图像的宽度，h1、h2分别为两个水印图像的高度。通过使用第二子网络，上述执行主体可以提取该待匹配图像的特征向量，以及从该待匹配图像的特征向量中，分别提取与上述两个位置信息对应的特征向量作为待匹配特征向量。

实践中，第二子网络可以是基于现有的目标检测网络（例如 SSD（Single Shot MultiBox Detector）、R-CNN（Region-based Convolutional Neural Networks）、Faster R-CNN 等）建立的神经网络。通过使用第二子网络，可以从待匹配图像中提取出待匹配区域图像的特征向量，从而提高图像之间进行匹配的针对性，有利于提高图像识别的效率和准确性。

在本实施例的一些可选的实现方式中，上述第二子网络包括用于将特征向量变换到目标维度的维度变换层。其中，维度变换层可以对特征向量进行处理（例如对特征向量包括的某些维度的数值以取平均值的方式合并）的公式；也可以是第二子网络包括的池化（pooling）层。池化层可以用于对输入的数据进行下采样（Down Sample）或上采样（Up Sample），以压缩或增加数据量。上述目标维度可以是技术人员设置的各种维度，例如与基准特征向量的维度相同的维度。上述执行主体可以将该待匹配图像输入图像识别模型包括的第二子网络，由第二子网络提取出该待匹配图像的至少一个特征向量，再由第二子网络包括的维度变换层对提取的各个特征向量进行维度变换，得到至少一个与基准特征向量的维度相同的待匹配特征向量。实践中，可以采用 ROI Pooling（Region Of Interest Pooling，感兴趣区域池化）层，使得每个待匹配特征向量与基准特征向量的维度相同。其中，ROI Pooling 层是目前广泛研究和应用的公知技术，在此不再赘述。

步骤 2032，确定所得到的待匹配特征向量与基准特征向量的距离。具体地，上述执行主体可以确定所得到的至少一个待匹配特征向量中的每个待匹配特征向量与基准特征向量的距离。其中，距离可以是以下任意一种：欧式距离、马氏距离(Mahalanobis Distance)等。上述预设距离可以是大于等于 0 的任意数值。其中，距离可以表征两个特征向量之间的相似程度，也即可以表征两个图像之间的相似程度。作为示例，如果两个特征向量的距离越大，则这两个特征向量各自对应的图像之间越不相似。

步骤 2033，响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。其中，距离阈值可以是技术人员根据经验设置的数值，也可以是上述执行主体根据历史数据（例如记录的历史距离阈值）计算（例如计算平均值）出的数值。具体地，如果所确定的各个距离中存在小于距离阈值的距离，则表示该待匹配图像中存在与基准对象图像相似的区域图像，也即表示该待匹配图像与基准对象图像匹配。

通过执行本步骤，可以在训练图像识别模型所使用的训练样本中

不包括基准对象图像的情况下，从待匹配图像集合中提取出与基准对象图像匹配的图像，并且，将待匹配图像中包括的区域图像与基准对象进行对比，可以提高对图像进行匹配的针对性，进而提高图像识别的准确性。

在本实施例的一些可选的实现方式中，如图 3 所示，图像识别模型可以预先通过以下步骤训练得到：

步骤 301，获取训练样本集。其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置。样本对象图像可以是表征某对象（例如水印、标志、人脸、物体等）的图像。标注位置信息的数量可以是至少一个，每个标注信息可以对应一个区域图像，各个区域图像中，包括表征的对象与样本对象图像表征的对象相同的区域图像。

步骤 302，从训练样本集中选取训练样本。其中，训练样本的选取方式和选取数量在本申请中并不限制。例如，可以采用随机选取或按训练样本的编号顺序选取的方式，从训练样本集中选取训练样本。

步骤 303，将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量。

其中，初始模型可以是基于机器学习技术而创建的现有的各种神经网络模型。该神经网络模型可以具有现有的各种神经网络的结构（例如 DenseBox、VGGNet、ResNet、SegNet 等）。上述各个特征向量可以是神经网络模型包括的某些层（例如卷积层）中提取的数据组成的向量。上述第一子网络、第二子网络分别与步骤 202 和步骤 203 中描述的第一子网络、第二子网络相同，在此不再赘述。

步骤 304，从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量。

具体地，上述目标区域图像可以是表征的对象与样本对象图像表征的对象相同的区域图像。执行本步骤的执行主体可以按照技术人员

指定的位置信息作为目标位置信息，以及将目标位置信息表征的区域图像作为目标区域图像，将目标区域图像的第二特征向量作为目标第二特征向量；或者，执行本步骤的执行主体可以根据所得到的位置信息，确定每个位置信息对应的区域图像与样本对象图像的相似度，以及将与样本对象图像的相似度最大的区域图像确定为目标区域图像，将目标区域图像的位置信息确定为目标位置信息，将目标区域图像的第二特征向量确定为目标第二特征向量。

步骤 305，基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成。

其中，第一损失值可以表征目标位置信息与目标区域图像对应的标注位置信息的差距。通常，第一损失值越小，目标位置信息与目标区域图像对应的标注位置信息的差距越小，也即目标位置信息越接近标注位置信息。实践中，第一损失值可以根据以下任意一种损失函数得到：Softmax 损失函数、Smooth L1（平滑 L1 范数）损失函数等。

第二损失值可以表征目标第二特征向量与第一特征向量的距离。通常第二损失值越大，目标第二特征向量与第一特征向量的距离越大，即目标区域图像与样本对象图像越不相似。作为示例，第二损失值可以是目标第二特征向量与第一特征向量的距离（例如欧式距离、马氏距离等）。

作为另一示例，第二损失值可以由 Triplet loss 损失函数得到，其中，Triplet loss 误差函数如下式所示：

$$L = \sum_i \left[\left\| x_i^a - x_i^p \right\|^2 - \left\| x_i^a - x_i^n \right\|^2 + threshold \right]_+,$$

其中， L 为第二损失值， Σ 为求和符号， i 为本次选取的各个训练样本的序号， a 表征样本对象图像， p 表征正样本图像（即目标区域图像）。 n 表征负样本图像（即样本匹配图像中的、除目标区域图像以外的其他区域图像；或者，预设的、表征的对象与样本对象图像表征的对象不同的图像）。 x_i^a 表征序号为 i 的训练样本包括的样本对象图像的特征向量， x_i^p 表征序号为 i 的训练样本对应的正样本图像（例如目标区域图像）的特征向量， x_i^n 表征序号为 i 的训练样本对应的负样本图

像（例如样本匹配图像中的、除目标区域图像以外的其他区域图像）的特征向量。 $threshold$ 表征预设距离， $\|x_i^a - x_i^p\|^2$ 表征第一距离（即第一特征向量与正样本图像的特征向量的距离）， $\|x_i^a - x_i^n\|^2$ 表征第二距离（即第一特征向量与负样本图像的特征向量的距离）。上式中方括号右下侧的“+”表示取正值，即当方括号中的表达式的计算结果为正时，取该正值，当为负时，取 0。实践中，在训练过程中，可以根据反向传播算法，调整初始模型的参数，使得 L 值最小或 L 值收敛，则表示训练完成。

在本实施例，执行本步骤的执行主体可以基于第一损失值和第二损失值，得到总损失值，将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。其中，目标值可以是预设的损失值阈值，当总损失值与目标值之差小于等于损失值阈值时，确定训练完成。

在本实施例的一些可选的实现方式中，执行本步骤的执行主体可以根据预设的权重值，将第一损失值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。上述权重值可以调整第一损失值和第二损失值占总损失值的比例，实现不同的应用场景下，使图像识别模型实现不同的功能（例如某些场景侧重提取位置信息，某些场景侧重比较特征向量的距离）。

步骤 306，响应于确定训练完成，将初始模型确定为图像识别模型。

在本实施例的一些可选的实现方式中，训练得到图像识别模型的执行主体可以响应于确定初始模型未训练完成，调整初始模型的参数，以及从训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行训练步骤。例如，假设初始模型是卷积神经网络，可以采用反向传播算法调整初始模型中各卷积层中的权重。然后，可以从训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行步骤 303-步骤 306。

需要说明的是，上述训练得到图像识别模型的执行主体可以与用于提取图像的方法的执行主体相同或者不同。如果相同，则训练得到图像识别模型的执行主体可以在训练得到图像识别模型后将训练好的图像识别模型的结构信息和参数的参数值存储在本地。如果不同，则训练得到图像识别模型的执行主体可以在训练得到图像识别模型后将训练好的图像识别模型的结构信息和参数的参数值发送给用于提取图像的方法的执行主体。

继续参见图 4，图 4 是根据本实施例的用于提取图像的方法的应用场景的一个示意图。在图 4 的应用场景中，服务器 401 首先获取终端设备 408 上传的水印图像 402（即基准对象图像），以及从本地获取待匹配图像集合 403。服务器 401 将水印图像 402 输入预先训练的图像识别模型 404 包括的第一子网络 4041，得到水印图像 402 的特征向量作为基准特征向量 405。

然后，服务器 401 从待匹配图像集合 403 中选择一个待匹配图像 4031，将待匹配图像 4031 输入图像识别模型 404 包括的第二子网络 4041，得到位置信息 4061、4062、4063，以及位置信息对应的待匹配特征向量 4071、4072、4073。其中，待匹配特征向量 4071、4072、4073 分别是待匹配图像 4031 包括的水印图像 40311、40312、40313 的特征向量。

最后，服务器 401 确定待匹配特征向量 4071 与基准特征向量 405 的距离小于等于预设的距离阈值，提取待匹配图像 4031 作为与基准对象图像匹配的图像并将匹配的图像发送至终端设备 408。服务器 401 通过反复从待匹配图像集合 403 中选择待匹配图像与水印图像 402 进行匹配，从而从待匹配图像集合 403 中提取出多个与水印图像 402 匹配的图像。

本申请的上述实施例提供的方法，通过利用预先训练的图像识别模型，得到基准图像的基准特征向量和待匹配图像的至少一个待匹配特征向量，再通过比较基准特征向量和待匹配特征向量的距离，得到与基准图像匹配的图像，从而提高了与基准图像进行匹配的针对性，以及实现了在训练图像识别模型所需的训练样本不包括基准图像的情

况下，利用图像识别模型提取出与基准图像匹配的图像，提高了图像识别的灵活性，丰富了图像识别的手段。

进一步参考图 5，其示出了用于提取图像的方法的又一个实施例的流程 500。该用于提取图像的方法的流程 500，包括以下步骤：

步骤 501，获取基准对象图像和待匹配图像集合。

在本实施例中，步骤 501 与图 2 对应实施例中的步骤 501 基本一致，这里不再赘述。

步骤 502，将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量。

在本实施例中，步骤 502 与图 2 对应实施例中的步骤 502 基本一致，这里不再赘述。

步骤 503，对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像；确定小于等于距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

在本实施例中，对于待匹配图像集合中的每个待匹配图像，上述执行主体可以对该待匹配图像执行如下提取步骤：

步骤 5031，将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量。步骤 5031 与图 2 对应实施例中的步骤 2031 基本一致，这里不再赘述。

步骤 5032，确定所得到的待匹配特征向量与基准特征向量的距离。步骤 5032 与图 2 对应实施例中的步骤 2032 基本一致，这里不再赘述。

步骤 5033，响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。步骤 5033 与图 2 对应实施例中的步骤 2033 基本一致，这里不再赘述。

步骤 5034, 确定小于等于距离阈值的距离对应的区域图像的位置信息, 以及将所确定的位置信息输出。

在本实施例中, 用于提取图像的方法的执行主体 (例如图 1 所示的服务器或终端设备) 可以基于步骤 5032 中确定的小于等于预设的距离阈值的距离, 从步骤 5031 中得到的至少一个位置信息中, 确定小于等于距离阈值的距离对应的位置信息, 以及将小于等于距离阈值的距离对应的位置信息输出。上述执行主体可以以各种方式将位置信息输出, 例如, 可以在与上述执行主体连接的显示器上显示位置信息包括的区域图像的标识信息、坐标信息等信息。

在本实施例的一些可选的实现方式中, 上述执行主体可以在输出位置信息之后, 基于输出的位置信息和该待匹配图像, 生成包括位置标记的匹配后图像。其中, 位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。具体地, 上述执行主体可以在待匹配图像中, 根据输出的位置信息绘制出预设形状的框, 将所绘制的框作为位置标记, 将包括位置标记的该待匹配图像作为匹配后图像。

从图 5 中可以看出, 与图 2 对应的实施例相比, 本实施例中的用于提取图像的方法的流程 500 突出了输出位置信息的步骤。由此, 本实施例描述的方案可以进一步确定出待匹配图像包括的目标区域图像的位置, 提高图像识别的针对性。

进一步参考图 6, 作为对上述各图所示方法的实现, 本申请提供了一种用于提取图像的装置的一个实施例, 该装置实施例与图 2 所示的方法实施例相对应, 该装置具体可以应用于各种电子设备中。

如图 6 所示, 本实施例的用于提取图像的装置 600 包括: 获取单元 601, 被配置成获取基准对象图像和待匹配图像集合; 生成单元 602, 被配置成将基准对象图像输入预先训练的图像识别模型包括的第一子网络, 得到基准对象图像的特征向量作为基准特征向量; 提取单元 603, 被配置成对于待匹配图像集合中的待匹配图像, 执行如下提取步骤: 将该待匹配图像输入图像识别模型包括的第二子网络, 得到至少

一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。

在本实施例中，获取单元 601 可以通过有线连接方式或者无线连接方式从远程或者从本地获取基准对象图像和待匹配图像集合。其中，基准对象图像可以是待将其与其他图像进行对比的图像，基准对象图像是表征某对象的图像。对象可以是各种事物，例如水印、标志、人脸、物体等。待匹配图像集合可以是预先存储的某类图像（例如包含商标的图像）的集合。

在本实施例中，生成单元 602 可以将所述基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到所述基准对象图像的特征向量作为基准特征向量。其中，第一子网络用于表征图像和图像的特征向量的对应关系。在本实施例中，图像识别模型可以是基于机器学习技术而创建的各种神经网络模型。该神经网络模型可以具有各种神经网络（例如 DenseBox、VGGNet、ResNet、SegNet 等）的结构。上述基准特征向量可以由神经网络模型包括的第一子网络（例如由神经网络模型包括的某个或某些卷积层组成的网络）提取的、表征图像的特征（例如形状、颜色、纹理等特征）的数据组成的向量。

在本实施例中，提取单元 603 可以对该待匹配图像执行如下步骤：

首先，将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量。其中，第二子网络用于表征图像与图像的位置信息、图像的待匹配特征向量的对应关系。位置信息用于表征待匹配特征向量对应的区域图像在该待匹配图像中的位置。待匹配特征向量是该待匹配图像包括的区域图像的特征向量。

然后，确定所得到的待匹配特征向量与所述基准特征向量的距离。具体地，上述提取单元 603 可以确定所得到的至少一个待匹配特征向

量中的每个待匹配特征向量与基准特征向量的距离。其中，距离可以是以下任意一种：欧式距离、马氏距离(Mahalanobis Distance)等。

最后，响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与所述基准对象图像匹配的图像。其中，距离阈值可以是技术人员根据经验设置的数值，也可以是上述提取单元 603 根据历史数据（例如记录的历史距离阈值）计算（例如计算平均值）出的数值。

在本实施例的一些可选的实现方式中，提取单元 603 可以包括：输出模块，被配置成确定小于等于距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

在本实施例的一些可选的实现方式中，提取单元 603 还可以包括：生成模块，被配置成基于输出的位置信息和该待匹配图像，生成包括位置标记的匹配后图像，其中，位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。

在本实施例的一些可选的实现方式中，第二子网络可以包括用于将特征向量变换到目标维度的维度变换层；以及提取单元 603 可以进一步被配置成：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个与基准特征向量的维度相同的待匹配特征向量。

在本实施例的一些可选的实现方式中，图像识别模型通过以下步骤训练得到：获取训练样本集，其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像的标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置；从训练样本集中选取训练样本，执行如下训练步骤：将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量；从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量；基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的第二损失值，确定初始模型是否训练完成；响应于确

定训练完成，将初始模型确定为图像识别模型。

在本实施例的一些可选的实现方式中，训练得到图像识别模型的执行主体可以根据预设的权重值，将第一损失值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。

在本实施例的一些可选的实现方式中，训练得到图像识别模型的步骤还可以包括：响应于确定初始模型未训练完成，调整初始模型的参数，以及从训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行训练步骤。

本申请的上述实施例提供的装置，通过利用预先训练的图像识别模型，得到基准图像的基准特征向量和待匹配图像的至少一个待匹配特征向量，再通过比较基准特征向量和待匹配特征向量的距离，得到与基准图像匹配的图像，从而提高了与基准图像进行匹配的针对性，以及实现了在训练图像识别模型所需的训练样本不包括基准图像的情况下，利用图像识别模型提取出与基准图像匹配的图像，提高了图像识别的灵活性，丰富了图像识别的手段。

下面参考图 7，其示出了适于用来实现本申请实施例的电子设备（例如图 1 所示的服务器或终端设备）的计算机系统 700 的结构示意图。图 7 示出的电子设备仅仅是一个示例，不应对本申请实施例的功能和使用范围带来任何限制。

如图 7 所示，计算机系统 700 包括中央处理单元（CPU）701，其可以根据存储在只读存储器（ROM）702 中的程序或者从存储部分 708 加载到随机访问存储器（RAM）703 中的程序而执行各种适当的动作和处理。在 RAM 703 中，还存储有系统 700 操作所需的各种程序和数据。CPU 701、ROM 702 以及 RAM 703 通过总线 704 彼此相连。输入/输出（I/O）接口 705 也连接至总线 704。

以下部件连接至 I/O 接口 705：包括键盘、鼠标等的输入部分 706；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 707；包括硬盘等的存储部分 708；以及包括诸如 LAN 卡、

调制解调器等网络接口卡的通信部分 709。通信部分 709 经由诸如因特网的网络执行通信处理。驱动器 710 也根据需要连接至 I/O 接口 705。可拆卸介质 711，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 710 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 708。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分 709 从网络上被下载和安装，和/或从可拆卸介质 711 被安装。在该计算机程序被中央处理单元（CPU）701 执行时，执行本申请的方法中限定的上述功能。需要说明的是，本申请所述的计算机可读介质可以是计算机可读信号介质或者计算机可读介质或者是上述两者的任意组合。计算机可读介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中，计算机可读介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

可以以一种或多种程序设计语言或其组合来编写用于执行本申请的操作的计算机程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本申请各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本申请实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中，例如，可以描述为：一种处理器包括获取单元、生成单元、提取单元。其中，这些单元的名称在某种情况下并不构成对该单元本身的限定，例如，获取单元还可以被描述为“获取基准对象图像和待匹配图像集合的单元”。

作为另一方面，本申请还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的电子设备中所包含的；也可以是单独存在，而未装配入该电子设备中。上述计算机可读介质承载有一

个或者多个程序，当上述一个或者多个程序被该电子设备执行时，使得该电子设备：获取基准对象图像和待匹配图像集合；将基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到基准对象图像的特征向量作为基准特征向量；对于待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与基准对象图像匹配的图像。

以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本申请中所涉及的发明范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述发明构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的（但不限于）具有类似功能的技术特征进行互相替换而形成的技术方案。

权 利 要 求 书

1、一种用于提取图像的方法，包括：

获取基准对象图像和待匹配图像集合；

将所述基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到所述基准对象图像的特征向量作为基准特征向量；

对于所述待匹配图像集合中的待匹配图像，执行如下提取步骤：
将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与所述基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与所述基准对象图像匹配的图像。

2、根据权利要求 1 所述的方法，其中，所述提取步骤还包括：

确定小于等于所述距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

3、根据权利要求 2 所述的方法，其中，所述提取步骤还包括：

基于输出的位置信息和该待匹配图像，生成包括位置标记的匹配后图像，其中，位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。

4、根据权利要求 1 所述的方法，其中，所述第二子网络包括用于将特征向量变换到目标维度的维度变换层；以及

所述将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个待匹配特征向量，包括：

将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个与基准特征向量的维度相同的待匹配特征向量。

5、根据权利要求 1-4 之一所述的方法，其中，所述图像识别模型通过以下步骤训练得到：

获取训练样本集，其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像的标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置；

从所述训练样本集中选取训练样本，执行如下训练步骤：将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量；从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量；基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成；响应于确定训练完成，将初始模型确定为图像识别模型。

6、根据权利要求 5 所述的方法，其中，所述基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成，包括：

根据预设的权重值，将第一损失值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。

7、根据权利要求 5 所述的方法，其中，训练得到所述图像识别模型的步骤还包括：

响应于确定初始模型未训练完成，调整初始模型的参数，以及从所述训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行所述训练步骤。

8、一种用于提取图像的装置，包括：

获取单元，被配置成获取基准对象图像和待匹配图像集合；

生成单元，被配置成将所述基准对象图像输入预先训练的图像识别模型包括的第一子网络，得到所述基准对象图像的特征向量作为基准特征向量；

提取单元，被配置成对于所述待匹配图像集合中的待匹配图像，执行如下提取步骤：将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个位置信息和位置信息对应的待匹配特征向量，其中，待匹配特征向量是该待匹配图像包括的区域图像的特征向量，位置信息用于表征区域图像在该待匹配图像中的位置；确定所得到的待匹配特征向量与所述基准特征向量的距离；响应于确定所确定的距离中存在小于等于预设的距离阈值的距离，提取该待匹配图像作为与所述基准对象图像匹配的图像。

9、根据权利要求 8 所述的装置，其中，所述提取单元包括：

输出模块，被配置成确定小于等于所述距离阈值的距离对应的区域图像的位置信息，以及将所确定的位置信息输出。

10、根据权利要求 9 所述的装置，其中，所述提取单元还包括：

生成模块，被配置成基于输出的位置信息和该待匹配图像，生成包括位置标记的匹配后图像，其中，位置标记用于标记输出的位置信息对应的待匹配区域图像在匹配后图像中的位置。

11、根据权利要求 8 所述的装置，其中，所述第二子网络包括用于将特征向量变换到目标维度的维度变换层；以及

所述提取单元进一步被配置成：

将该待匹配图像输入所述图像识别模型包括的第二子网络，得到至少一个与基准特征向量的维度相同的待匹配特征向量。

12、根据权利要求 8-11 之一所述的装置，其中，所述图像识别模

型通过以下步骤训练得到：

获取训练样本集，其中，训练样本包括样本对象图像、样本匹配图像、样本匹配图像的标注位置信息，标注位置信息表征样本匹配图像中包括的区域图像的位置；

从所述训练样本集中选取训练样本，执行如下训练步骤：将选取的训练样本包括的样本对象图像输入初始模型包括的第一子网络，得到第一特征向量，将样本匹配图像输入初始模型包括的第二子网络，得到至少一个位置信息和与位置信息对应的第二特征向量；从所得到的至少一个位置信息中，确定表征样本匹配图像中的目标区域图像的位置信息作为目标位置信息，确定目标位置信息对应的第二特征向量作为目标第二特征向量；基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成；响应于确定训练完成，将初始模型确定为图像识别模型。

13、根据权利要求 12 所述的装置，其中，所述基于表征目标位置信息的误差的第一损失值和表征目标第二特征向量与第一特征向量的距离的差距的第二损失值，确定初始模型是否训练完成，包括：

根据预设的权重值，将第一损失值与第二损失值的加权求和结果作为总损失值，以及将总损失值与目标值进行比较，根据比较结果确定初始模型是否训练完成。

14、根据权利要求 12 所述的装置，其中，训练得到所述图像识别模型的步骤还包括：

响应于确定初始模型未训练完成，调整初始模型的参数，以及从所述训练样本集中的、未被选取的训练样本中，选取训练样本，使用参数调整后的初始模型作为初始模型，继续执行所述训练步骤。

15、一种电子设备，包括：

一个或多个处理器；

存储装置，其上存储有一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-7 中任一所述的方法。

16、一种计算机可读介质，其上存储有计算机程序，其中，该程序被处理器执行时实现如权利要求 1-7 中任一所述的方法。

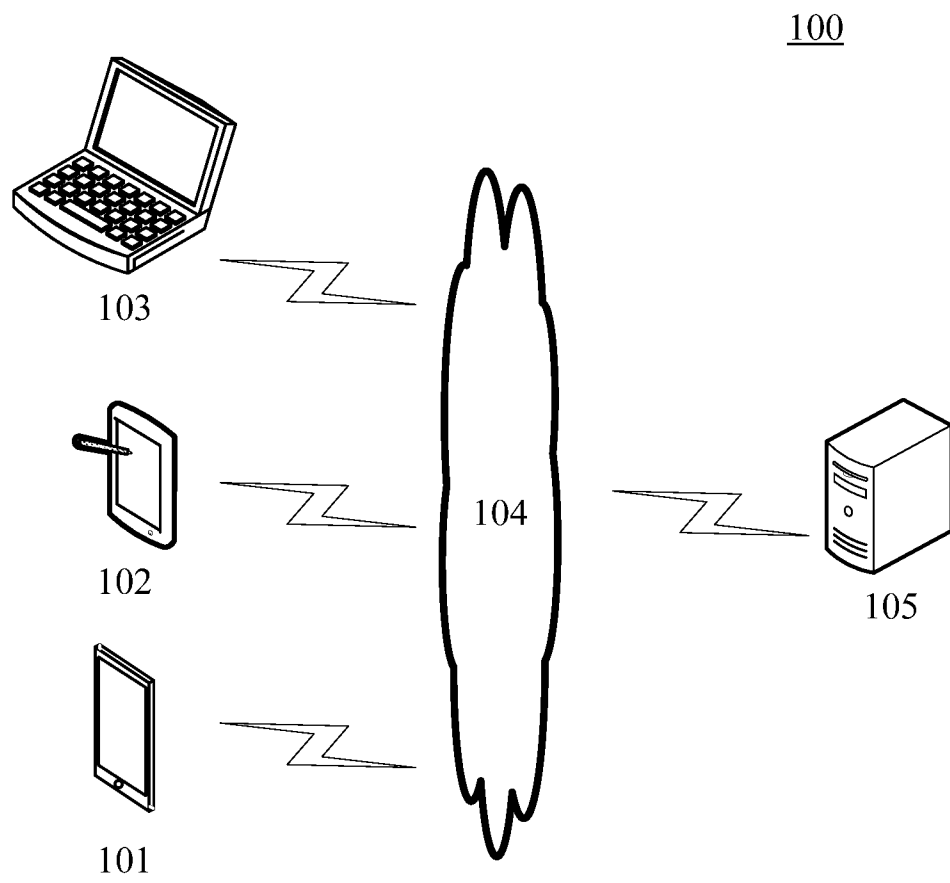


图 1

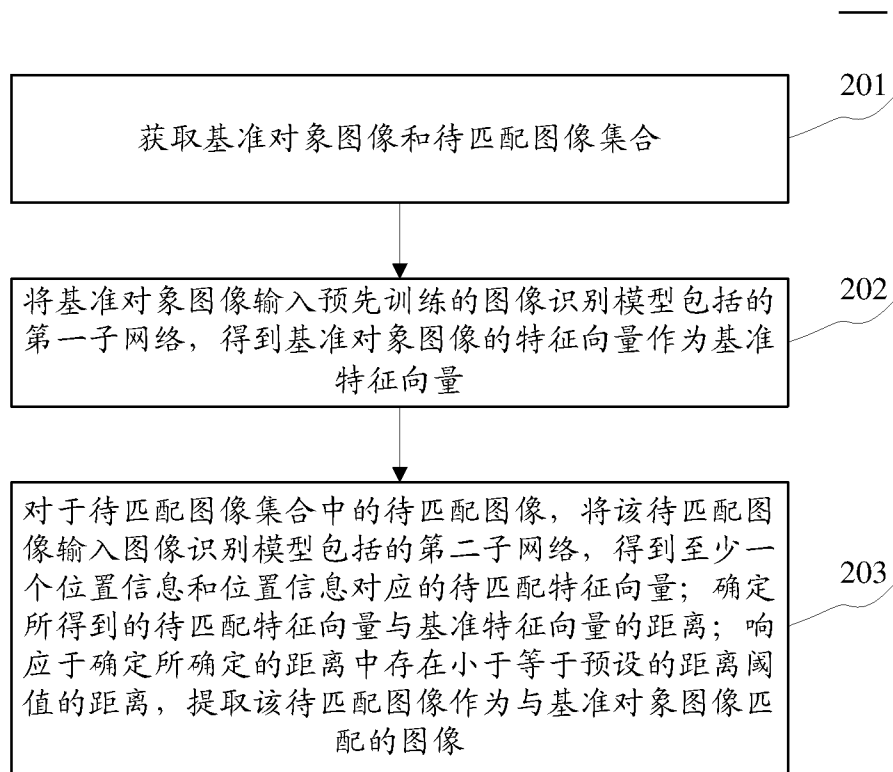


图 2

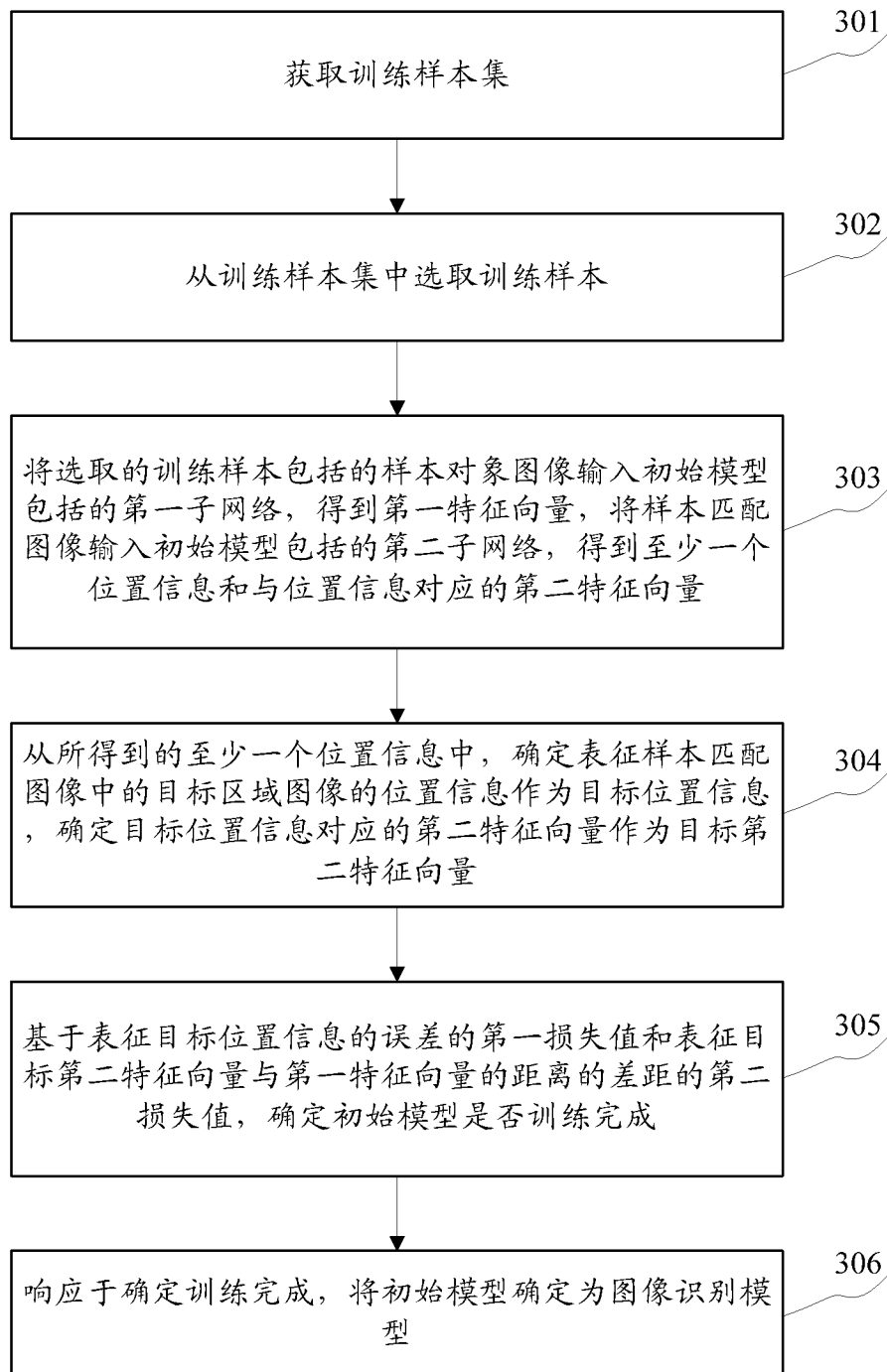


图 3

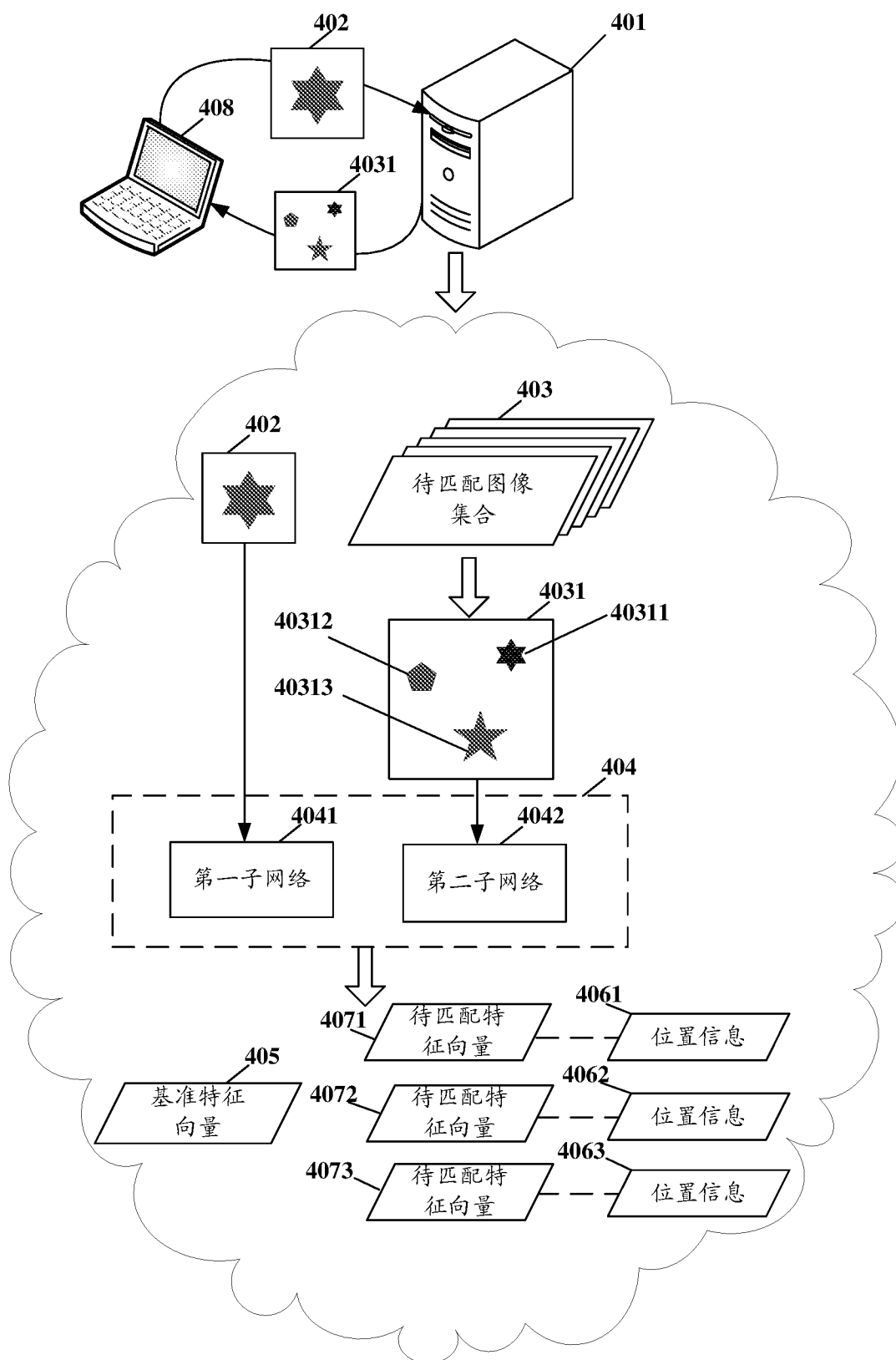


图 4

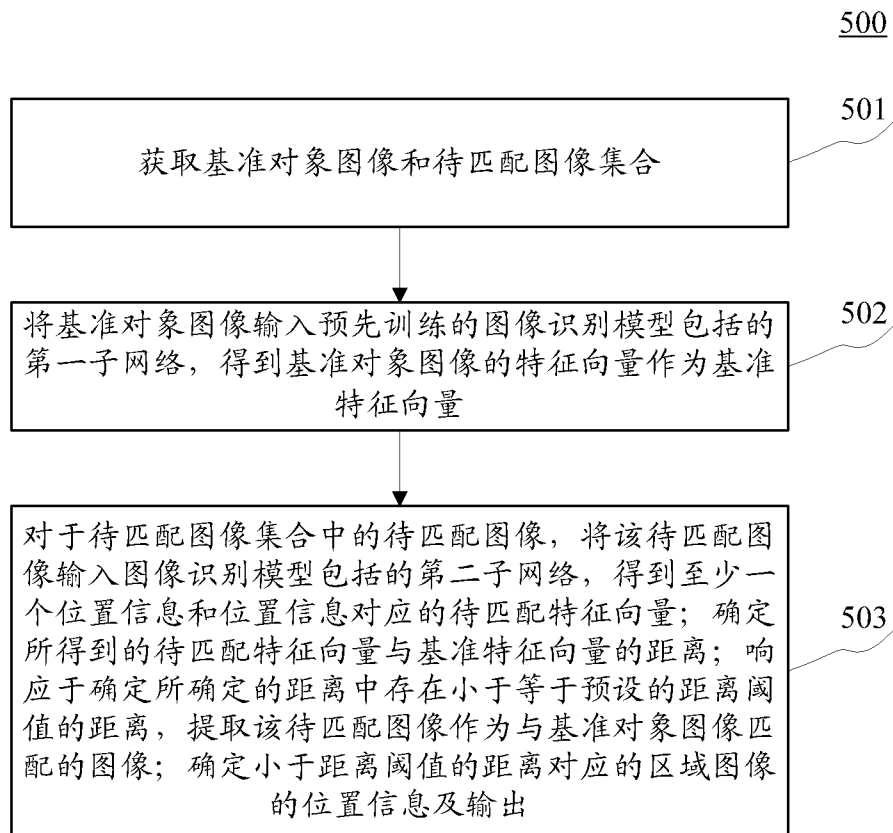


图 5

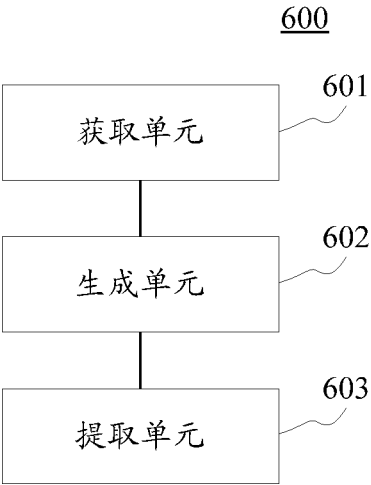


图 6

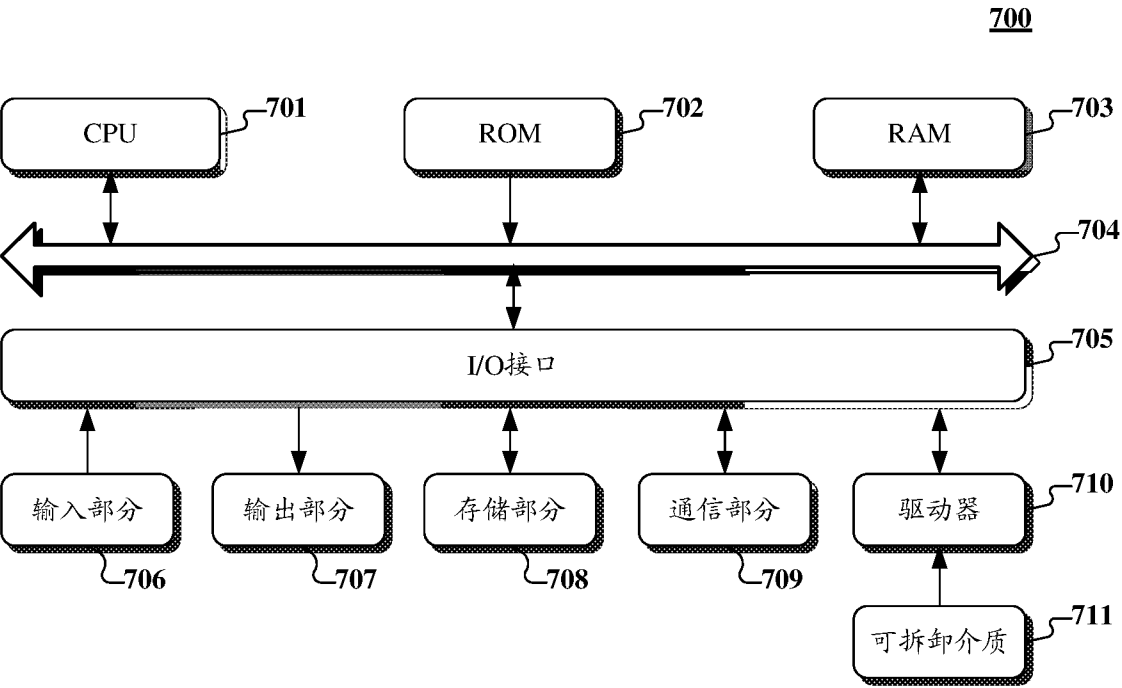


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/116334

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(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)

CNPAT, CNKI, WPI, EPODOC, IEEE, GOOGLE: 字节跳动, 周恺卉, 王长虎, 图像, 图片, 提取, 目标, 基准, 训练, 样本, 匹配, 特征向量, 位置, 区域, 距离, 神经网络, 识别, picture, image, target, extract, reference, match, recognize, feature, vector, model, position, location, area, distance, neural net

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 107679466 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 09 February 2018 (2018-02-09) abstract, description, paragraphs [0049]-[0083], and figures 2-4	1-16
X	CN 107239535 A (BEIJING XIAOMI MOBILE SOFTWARE CO., LTD.) 10 October 2017 (2017-10-10) description, paragraphs [0070]-[0086], and figure 1	1-16
A	CN 104376326 A (JILIN UNIVERSITY) 25 February 2015 (2015-02-25) entire document	1-16
A	CN 108038880 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 15 May 2018 (2018-05-15) entire document	1-16
A	CN 108154196 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 12 June 2018 (2018-06-12) entire document	1-16
A	US 2014105509 A1 (CANON KABUSHIKI KAISHA) 17 April 2014 (2014-04-17) entire document	1-16

☐ 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

20 March 2019

Date of mailing of the international search report

10 April 2019

Name and mailing address of the ISA/CN

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

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/116334

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	107679466	A	09 February 2018	None	
CN	107239535	A	10 October 2017	None	
CN	104376326	A	25 February 2015	None	
CN	108038880	A	15 May 2018	None	
CN	108154196	A	12 June 2018	None	
US	2014105509	A1	17 April 2014	None	

国际检索报告

国际申请号

PCT/CN2018/116334

A. 主题的分类

G06K 9/62 (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))

CNPAT, CNKI, WPI, EPODOC, IEEE, GOOGLE: 字节跳动, 周恺卉, 王长虎, 图像, 图片, 提取, 目标, 基准, 训练, 样本, 匹配, 特征向量, 位置, 区域, 距离, 神经网络, 识别, picture, image, target, extract, reference, match, recognize, feature, vector, model, position, location, area, distance, neural net

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
X	CN 107679466 A (百度在线网络技术北京有限公司) 2018年 2月 9日 (2018 - 02 - 09) 说明书摘要、说明书第[0049]-[0083]段和附图2-4	1-16
X	CN 107239535 A (北京小米移动软件有限公司) 2017年 10月 10日 (2017 - 10 - 10) 说明书第[0070]-[0086]段和附图1	1-16
A	CN 104376326 A (吉林大学) 2015年 2月 25日 (2015 - 02 - 25) 全文	1-16
A	CN 108038880 A (百度在线网络技术北京有限公司) 2018年 5月 15日 (2018 - 05 - 15) 全文	1-16
A	CN 108154196 A (百度在线网络技术北京有限公司) 2018年 6月 12日 (2018 - 06 - 12) 全文	1-16
A	US 2014105509 A1 (CANON KABUSHIKI KAISHA) 2014年 4月 17日 (2014 - 04 - 17) 全文	1-16

☐ 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

2019年 3月 20日

国际检索报告邮寄日期

2019年 4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

陈茜茜

电话号码 86-(10)-53961695

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/116334

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	107679466	A	2018年 2月 9日	无	
CN	107239535	A	2017年 10月 10日	无	
CN	104376326	A	2015年 2月 25日	无	
CN	108038880	A	2018年 5月 15日	无	
CN	108154196	A	2018年 6月 12日	无	
US	2014105509	A1	2014年 4月 17日	无	