

(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 1 月 10 日 (10.01.2019)



(10) 国际公布号
WO 2019/007253 A1

- (51) 国际专利分类号:
G06K 9/62 (2006.01)
- (21) 国际申请号: PCT/CN2018/093350
- (22) 国际申请日: 2018 年 6 月 28 日 (28.06.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201710546203.4 2017 年 7 月 6 日 (06.07.2017) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 李博 (LI, Bo); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 张伦 (ZHANG, Lun);

中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 楚汝峰 (CHU, Rufeng); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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,

(54) Title: IMAGE RECOGNITION METHOD, APPARATUS AND DEVICE, AND READABLE MEDIUM

(54) 发明名称: 图像识别方法、装置及设备、可读介质

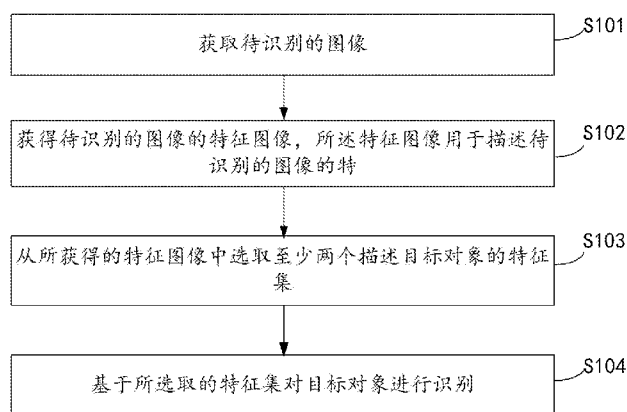


图 1

- S101 Obtain an image to be recognized
S102 Obtain a feature image of the image to be recognized, the feature image being used for describing features of the image to be recognized
S103 Select at least two feature sets describing target objects from the obtained feature image
S104 Recognize the target objects on the basis of the selected feature sets

(57) Abstract: The present application provides an image recognition method, apparatus and device, and a readable medium. The method comprises: obtaining an image to be recognized; obtaining a feature image of the image to be recognized, the feature image being used for describing features of the image to be recognized; selecting at least two feature sets describing target objects from the obtained feature image; and recognizing the target objects on the basis of the selected feature sets. By implementing the present application, when features capable of describing a target object are selected from a feature image, multiple feature sets are selected from different regions in the feature image and can effectively represent the same type of target objects at different positions in the image, and therefore, the target objects can be more accurately recognized.

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))。

(57) 摘要: 本申请提供一种图像识别方法、装置及设备、可读介质, 所述方法包括: 获取待识别的图像; 获得待识别的图像的特征图像, 所述特征图像用于描述待识别的图像的特征; 从所获得的特征图像中选取至少两个描述目标对象的特征集; 基于所选取的特征集对目标对象进行识别。实施本申请, 从特征图像中选取能够描述目标对象的特征时, 从所述特征图像内不同区域选取多个特征集, 能有效代表图像中不同位置的同类目标对象, 因此能更准确地识别出目标对象。

图像识别方法、装置及设备、可读介质

本申请要求 2017 年 07 月 06 日递交的申请号为 201710546203.4、发明名称为“图像识别方法、装置及设备、可读介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及图像处理技术领域，尤其涉及图像识别方法、装置及设备、可读介质。

背景技术

10 随着计算机技术的发展和计算机视觉原理的广泛应用，利用图像识别技术对目标对象进行分类，在无人驾驶汽车、智能冰箱等产品上具有广泛的应用价值。目前的图像识别技术对目标对象进行识别时，一般先采用特征提取模型对含有目标对象的整张图像进行特征提取，生成整张图像的特征图像，该特征图像由所提取的特征组成，所提取的特征包括颜色特征、纹理特征、形状特征、空间关系特征等图像特征中的至少一项，再采用
15 单个固定大小的矩形框将特征图像中描述目标对象（如车，食材等）的特征框出来，进而选取被框出的特征为目标特征，然后将目标特征输入分类模型进行分类。

但是，在某些产品上应用图像识别技术时，拍摄所得的含有目标对象的图像的不同区域可能出现同一目标对象，例如，智能冰箱内的食材大多会被用户随机摆放在冰箱内，拍摄智能冰箱内的食材所得的图像内。目前的图像识别技术在识别这种图像时，易导致
20 错误的识别结果。

发明内容

有鉴于此，本申请提供一种图像识别方法、装置及设备、可读介质。

25 根据本申请实施例的第一方面，提供一种图像识别方法，包括步骤：
获取待识别的图像；
获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征；
从所获得的特征图像中选取至少两个描述目标对象的特征集；
基于所选取的特征集对目标对象进行识别。

30

根据本申请实施例的第二方面，提供一种电子设备，包括：

处理器；

存储处理器可执行指令的存储器；

其中，所述处理器耦合于所述存储器，用于读取所述存储器存储的程序指令，并作

5 为响应，执行如下操作：

获取待识别的图像；

获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征；

从所获得的特征图像中选取至少两个描述目标对象的特征集；

基于所选取的特征集对目标对象进行识别。

10

根据本申请实施例的第三方面，提供一种图像识别装置，包括：

图像获取模块，用于获取待识别的图像；

特征提取模块，用于获得待识别的图像的特征图像，所述特征图像用于描述待识别
的图像的特征；

15 特征选取模块，用于从所获得的特征图像中选取至少两个描述目标对象的特征集；

目标识别模块，用于基于所选取的特征集对目标对象进行识别。

根据本申请实施例的第四方面，提供一个或多个机器可读介质，其上存储有指令，
当由一个或多个处理器执行时，使得终端设备执行以上所述的方法。

20

实施本申请提供的实施例，从特征图像中选取能够描述目标对象的特征时，从所述
特征图像内不同区域选取多个特征集，能有效代表图像中不同位置的同类目标对象，因
此能更准确地识别出目标对象。

25 附图说明

图 1 是本申请一示例性实施例示出的图像识别方法的流程图；

图 2a 是本申请一示例性实施例示出的用于图像识别的系统的框图；

图 2b 是本申请另一示例性实施例示出的图像识别方法的交互示意图；

图 2c 是本申请示例性实施例示出的图像识别方法中的池化操作和调整像素的实现

30 过程的示意图；

图 2d 是从本申请示例性实施例示出的图像识别方法中的目标识别过程的示意图；

图 3 是本申请一示例性实施例示出的图像识别装置的逻辑框图；

图 4 是本申请一示例性实施例示出的图像识别装置的硬件结构图。

5 具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

10 在本申请使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

应当理解，尽管在本申请可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

请参阅图 1，图 1 是本申请一示例性实施例示出的图像识别方法的流程图，该实施
20 例能应用于具有图像处理功能的各种电子设备，可以包括以下步骤 S101-S104：

步骤 S101、获取待识别的图像。

步骤 S102、获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征。

步骤 S103、从所获得的特征图像中选取至少两个描述目标对象的特征集。

25 步骤 S104、基于所选取的特征集对目标对象进行识别。

本申请实施例，所获取的图像可以是图像采集模块（如摄像头）直接采集的图像，也可以是经过图像预处理后的图像数据，这里提到的图像预处理可以包括对提高图像识别准确度有益的图像处理，例如：对场景文本图像进行的颜色空间变换，对文本词图像中词图像的位置校正，对字符图像进行的去噪处理等。

30 对于获得的图像，可以采用卷积神经网络模型、分类器、多级网络结构等特征提取

算法, 对其进行特征提取, 生成特征图像, 该特征图像的各区域含有所提取的各种特征。

在某些例子中, 对目标对象进行识别时需要定位目标对象的具体位置, 为了准确定位到目标对象, 可以采用能够有效保留目标对象的位置信息的全卷积神经网络对所获取的图像进行特征提取, 该全卷积神经网络可以包括 AlexNet、GoogleNet、VGGNet、ResNet
5 或其他卷积神经网络模型的全卷积层。

获得特征图像后, 考虑到图像中可能存在大小、形状各异的目标对象, 在提取描述目标对象的特征集时可以基于区域, 从所获得的特征图像中选取描述目标对象的特征, 从特征图像的同一区域内选取的特征构成描述目标对象的一个特征集, 区域尺寸不同时, 特征集所含的特征量不同, 选取的特征集在所述特征图像中所属的区域的尺寸不同。

10 一般情况下含有目标对象的区域的卷积运算结果大于预定阈值, 该预定阈值可以由设计人员在训练分类器和特征提取模型时, 通过验证集上正负样本的卷积运算结果的分布来设定, 其数值一般大于或等于 0 且小于 1, 例如 0.3, 0.5 等, 在某些例子中, 可以通过以下滑动窗口技术将卷积运算结果大于预定阈值的特征集作为描述目标对象的特征集:

15 采用多种尺寸的滑动窗口从所获得的特征图像中选出多个候选特征集, 滑动窗口的尺寸可以包括 8*16, 8*8, 16*16, 16*32, 16*8, 32*64, 32*32, 32*16、以及 64*32。

对选出的候选特征集进行卷积运算。

将卷积运算结果大于预定阈值的候选特征集选取为描述目标对象的特征集。

实际应用中, 卷特征图像可以是全卷积神经网络对所获取的图像进行特征提取后输出
20 出的卷积特征块, 采用滑动窗口选取描述目标对象的特征集时, 可以在卷积特征块上寻找可能存在目标对象的区域。对于卷积特征块上的每一个位置, 分别采用尺寸为 8*16、8*8、16*16、16*32、16*8、32*64、32*32、32*16 和 64*32 的矩形框 (滑动窗口) 标记出各矩形区域, 其中, 矩形框可以用 4 维向量 $[c_x, c_y, w_0, h_0]$ 等区域标识标记各矩形区域, (c_x, c_y) 表示矩形框的中心点坐标, w_0, h_0 分别表示矩形框的宽度和高度, 与矩形框的尺寸
25 对应。

然后提取标记出的每个矩形区域内的特征, 将一个矩形区域内的特征组成一个候选特征集, 该矩形区域即为该特征集在特征图像 (卷积特征块) 内所属的区域, 然后对每个候选特征集进行卷积运算, 确定卷积运算结果大于预定阈值的候选特征集所属的矩形区域为可能存在目标对象的区域, 并将卷积运算结果大于预定阈值的候选特征集选取为

描述目标对象的特征集，这些最终被选取的特征集还可以基于其所属的矩形区域的尺寸的不同，被划分为不同种类的特征集。在其他例子中，本申请的设计人员还可以根据图像识别的具体应用场景，将滑动窗口的边长、以及不同边长间的比值设置为其他数值，本申请对此不做限制。

- 5 在某些例子中，卷积特征块的高度和宽度均较大，如果对于卷积特征块上的每一个位置，分别选用尺寸为 8*16、8*8、16*16、16*32、16*8、32*64、32*32、32*16 和 64*32 的矩形框（滑动窗口）标记出各矩形区域，然后提取每个矩形区域内的特征，将构成大量的特征集，进行卷积运算的计算量较大，选取到的描述目标对象的特征集的数目可能也较多，进而会增大图像识别过程的计算量，降低图像识别效率，为了解决这些问题，
- 10 考虑到描述目标对象的特征集的卷积运算结果大于其他非目标对象的特征集的卷积运算结果，可以将卷积运算结果超过预定阈值的特征集确定为候选特征集，将所确定的候选特征集中，卷积运算结果为前 N 大的特征集选取为描述目标对象的特征集，N 大于 1 且小于所确定的候选特征集的总数，候选特征集的数目较大时，N 可以为 300。

- 实际应用中，可以将从各矩形区域内的特征组成的各特征集输入预定的特征集筛选
- 15 模型，由该特征集筛选模型对每个特征集进行卷积运算，将卷积运算结果超过预定阈值的特征集确定为候选特征集，将所确定的候选特征集中，卷积运算结果为前 N 大的特征集选取为描述目标对象的特征集，N 大于 1 且小于所确定的候选特征集的总数，本申请设计人员可以根据应用场景、以及运行本申请的图像识别方法的电子设备的计算能力确定 N 的具体数值，如 300。这里提到的特征集筛选模型可以是深度神经网络模型、多级
- 20 结构的网络模型、或者基于图像颜色、边缘、超像素特征的概率模型。

在某些场景，所获得的图像内除包括目标对象外，可能还含多种多样的背景信息，这些背景信息会对目标识别造成一定程度的困扰，为了降低背景信息对识别过程的负面影响，可以在目标识别过程中增加上下文特征，所增加的上下文特征包括局部上下文特征和/或全局上下文特征。

- 25 实际应用中，可以在从所获得的特征图像中选取描述目标对象的特征集后，从所述特征图像中选取特征集的上下文特征，然后根据所选取的特征集和上下文特征对目标对象进行识别。添加上下文特征后，目标识别过程可以处理与目标对象相关的更多特征，一方面便于识别体积比较小的目标对象，另一方面可以排除不可能的目标对象，比如：船和大海总是在一起，如果检测到船和树木在一起，就说明目标对象识别错误。

- 30 在某些例子中，从所述特征图像中选取特征集的上下文特征（特征集的上下文特征

可以指特征集中不同描述目标对象的特征分别对应的上下文特征)时,可以在选取好描述目标对象的特征集后,针对选取的每个特征集,可以该特征集所属的区域的中心点为参照点,将该特征集所属的区域的边长增大 0.5 倍,构成局部上下文特征所属的区域,然后提取该区域的特征为该特征集的局部上下文特征。这样提取局部上下文特征后,特征集的局部上下文特征所属的区域的边长,为所述特征集所属的区域的边长的 1.5 倍,可以包含与目标对象相关的更多特征,便于识别体积比较小的目标对象。

在提取到描述目标对象的特征集和上下文特征后,可以基于这些特征对目标对象进行识别,如将所提取的描述目标对象的特征集和上下文特征输入训练后的分类器进行目标分类,但是这样操作会面临巨大的计算量的挑战,而且,分类器对超过较大数量的特征进行运算时还容易出现过拟合。为了解决这个问题,本申请可以对特征集、以及所述特征集的上下文特征分别进行池化操作,然后根据所述池化操作所得的特征集和上下文特征,对目标对象进行识别。这里提到的池化,用于降低特征集的维度和出现过拟合的概率,一般是对不同位置的特征进行聚合统计,例如,进行池化操作时,可以用特征集的某个区域的各特征的平均值 (或最大值)代替各特征。

在其他例子中,为了进一步提高目标识别效率,可以在对特征集、以及所述特征集的上下文特征分别进行池化操作时,对各特征提取通道所提取的指定特征分别进行池化操作;不同特征提取通道所提取的指定特征在所述特征图像中的坐标不同。如:特征图像为全卷积网络提取的卷积特征块,可以将每个特征集在卷积特征块内所属的区域分割为三部分,在进行池化操作时,单独对全卷积网络的第一特征提取通道提取的第一部分区域内的特征(指定特征)进行池化操作,单独对全卷积网络的第二特征提取通道提取的第二部分区域内的特征(指定特征)进行池化操作,单独对全卷积网络的第三特征提取通道提取的第三部分区域内的特征(指定特征)进行池化操作。这样操作后,一方面对目标进行识别所需的深度神经网络的层数较少,另一方面可以识别到目标对象的相对位置关系,便于准确定位目标对象。

此外为了提高目标识别效率,还可以将池化操作所得的特征集和上下文特征调整到匹配的像素;然后根据调整后的特征集和上下文特征,对目标对象进行识别。该匹配的像素一般小于每个特征集的像素,本申请设计人员可以根据应用场景、以及运行本申请的图像识别方法的电子设备的计算能力确定匹配的像素,在某些场景,考虑到图像中可能存在大小、形状各异的目标对象,匹配的像素可以包括 3*12、12*3、5*10、10*5 和 7*7 中的至少两种。

此外，本申请实施例也可以在基于所选取的特征集对目标对象进行识别时，将所选取的特征集调整到匹配的像素，然后根据调整后的特征集对目标对象进行识别。

如果匹配的像素有多种数值，可以将一种匹配的像素的特征集合上下文特征作为描述目标对象的一个分支的特征，该分支的特征数量为 $H_0 \times W_0 \times ((3 \times h_i \times w_i) \times (C+1))$ ，其中，
 5 H_0 、 W_0 、 C_0 分别代表特征图像（如卷积特征块）的高、宽和通道数，
 $h_i \times w_i \in \{3 \times 12, 12 \times 3, 5 \times 10, 10 \times 5, 7 \times 7\}$ ， C 表示目标对象的类别数，+1 是将背景也算作一个目标类别， $h_i \times w_i$ 上每一个位置点都是一个 $3 \times (C+1)$ 维的向量，该向量包括 3 个 $(C+1)$ 维的向量。

在对提取的特征集和上下文特征顺次进行池化操作和像素调整后，可以将像素调整
 10 后的特征输入目标识别模型，目标识别模型在识别目标的过程中对应每一个分支的每个特征集、以及该特征集的上下文特征，都会产生一个类别向量和该特征集所属区域的位置偏移向量。这里提到的目标识别模型可以是分类器之类的分类模型。

其中，类别向量的长度可以为 $(C+1)$ ，每一维向量元素可以表示目标对象属于某一类别的概率 P_j ， $j \in \{0, \dots, C\}$ ，其中 0 表示背景类。然后目标识别模型根据预定的向量筛选
 15 准则来确定一个最终的目标类别向量和目标位置偏移向量。

位置偏移向量可以是 4 维向量，这个向量的每一维元素表示特征集所属区域的位置偏移向量 $[\Delta_x, \Delta_y, \Delta_w, \Delta_h]$ ，该位置偏移向量对应于 4 维向量 $[c_x, c_y, w_0, h_0]$ ，这里 Δ_x 、 Δ_y 、 Δ_w 和 Δ_h 就是 c_x 、 c_y 、 w_0 和 h_0 分别需要调整的偏移量。目标对象所在位置被调整后对应的向量即 $[c_x + w_0 \cdot \Delta_x, c_y + h_0 \cdot \Delta_y, w_0 \cdot \Delta_w, h_0 \cdot \Delta_h]$ 。

20 如果预定的向量筛选准则为选取最大的向量元素，可以从每个分支的各特征集对应的类别向量中选取向量元素最大的类别向量为最终识别的目标类别向量，然后将该目标类别向量对应的特征集的位置偏移向量为最终识别的位置偏移向量，在某些例子中，选取向量元素最大的类别向量为最终识别的目标类别向量时，可以根据如下公式选取：

$$S = \max_{a \in \{1, \dots, A\}} (\max_{c \in \{1, \dots, C\}} score_c^a) ;$$

25 其中，score 表示类别向量，向量的每一维上的元素表示目标对象属于相应类别的概率， C 表示类别数目， A 表示分指数（预定像素的种类数）。在某个例子中，有 2 类目标对象，一类为狗，另一类为猫，那么 $C=2$ ，类别向量的第一维表示目标对象属于猫这

一类别的可能性，第二维表示目标对象属于狗这一类别的可能性，类别向量可以用 $\text{score} = [0.3, 0.9]$ 来表示，通过上述公式括号里的第一次求 $\max$ ，可以取 0.3 和 0.9 中的最大值。经过第一次求 $\max$ 后，每个分支都有一个最大值，假设为 $\text{score}_2, \text{score}_3, \dots$ ，通过上述公式中括号外的第二次求 $\max$ ，可以从不同的分支的 $\text{score}_2, \text{score}_3, \dots$ 中选取最大值为终

5 极最大大值，将选取的终极最大值所属的类别向量确定为能确定目标对象所属类别的目标类别向量。在其他例子中，还可以根据类别向量的所有维度的平均值（mean）、最小值、中位数(median)来确定目标类别向量。

在确定好目标类别向量后，目标识别模型可以将目标类别向量和目标位置偏移向量作为识别结果输出，所述目标类别向量为各特征集对应的类别向量中最大的向量元素所

10 属的类别向量，所述目标位置偏移向量为所述目标类别向量对应的特征集的位置偏移向量。

在某些例子中，目标类别向量为 $c_i, i \in \{1, \dots, C+1\}$ ，其向量元素并非目标对象属于对应类别的概率值，在输出目标类别向量前，可以通过求 softmax 将目标类别向量转成概率形式目标类别向量，Softmax 的公式如下：

$$p_i = \frac{\exp(c_i)}{\sum_{j=1}^K \exp(c_j)}$$

p_i 为概率形式目标类别向量。

在目标识别模型输出识别结果后，可以得到目标对象所属的类别，进而结合目标类别向量对应的特征集所属区域的初始位置 $[c_x, c_y, w_0, h_0]$ 和识别结果中的目标位置偏移向量，可以得到目标对象在图像中所处的位置： $[c_x + w_0 \cdot \Delta_x, c_y + h_0 \cdot \Delta_y, w_0 \cdot \Delta_w, h_0 \cdot \Delta_h]$ 。

20 综上所述，通过本申请实施例的图像识别，可以检测出一幅图像中所有目标对象的类别和位置，如果待识别的图像为拍摄智能冰箱的储物间所得的图像，目标对象为智能冰箱的储物间储存的食材，那么基于图像识别结果，在智能冰箱领域可以进一步进行相关信息的统计，如：统计同一类别的食材的数目、所有类别的食材的数目等，然后基于统计结果可以对食品进行准确有效地智能化管理，如：可以调换冰箱运行模式，始终让

25 食品保持最佳存储状态，可以让用户通过手机或计算机等随时随地了解冰箱里食物的数量、保鲜保质信息，可以提醒用户定时补充食品等。

此外，本申请实施例的图像识别应用于无人驾驶汽车中时，可以准确识别汽车前面

的路况，基于路况进行相应的驾驶操作，如：无人驾驶时绕开障碍物等。

请参阅图 2a,图 2a 是本申请一示例性实施例示出的用于实现图像识别的系统 200 的框图，该系统 200 适用于具有图像处理功能的各种电子设备，可以包括顺次连接的摄像头 210、全卷积神经网络 220、特征集生成模块 230、特征集筛选模型 240、池化操作模
5 块 260、像素调整模块 270 和目标识别模型 280，还包括分别与全卷积神经网络 220、特征集生成模块 230、特征集筛选模型 240 和池化操作模块 260 连接的上下文获取模块 250。

其中，摄像头 210，直接拍摄对应场景的图像，在其他例子中，还可以用图像搜集设备代替摄像头 210，从对应区域搜集对应场景的图像。

全卷积神经网络 220，对图像采集模块 210 获取的图像进行特征提取，生成卷积特
10 征块（特征图像）。

特征集生成模块 230，用于从卷积特征块中可能存在目标对象的区域提取特征，构成特征集。

特征集筛选模型 240，用于从提取的各特征集中筛选出能够较好的描述目标对象的特征集。

上下文获取模块 250，用于基于筛选出的每个特征集的所属区域，从卷积特征块中
15 提取筛选出的每个特征集的上下文特征。

池化操作模块 260，用于分别对描述目标对象的特征集和进行池化操作，以减少特征量，提高目标识别过程的计算量，进而提高图像识别的准确度。

像素调整模块 270，用于分别将池化操作后的特征集和上下文特征调整到匹配的像
20 素。

目标识别模型 280，用于基于像素调整后的特征识别出目标对象的类别，在某些例子中，还可以进一步用于定位目标对象在图像内的位置。

以下结合图 2a 至图 2d 列举一个应用实例。

在该实例中，本申请的设计人员预先将图像识别应用于智能冰箱，分别选用尺寸为
25 8*16、8*8、16*16、16*32、16*8、32*64、32*32、32*16 和 64*32 的矩形框（滑动窗口），将匹配的像素设置为 3*12、12*3、5*10、10*5 和 7*7 这五种，将预定的向量筛选准则定为选取最大的向量元素。

安装在智能冰箱内的摄像头 210 拍摄智能冰箱内的画面生成待识别的图像（步骤 S201），并将图像发送到全卷积神经网络 220（步骤 S202），全卷积神经网络 220 对图
30 像进行特征提取生成卷积特征块（S203），并发送卷积特征块到特征集生成模块 230 和

上下文获取模块 250 (S204、S205)，特征集生成模块 230 采用各种尺寸的矩形框从卷积特征块中提取特征，获得各特征集，存储各特征集在卷积特征块中所属的区域的区域标识 (S206)，并将获得的特征集发送到特征集筛选模型 240 (S207)，特征筛选模型 240 通过计算各特征集的卷积选取出能描述目标对象的特征集 (S208)，并将选取出的特征集发送到上下文获取模块 250 和池化操作模块 260 (S209、S210)，上下文获取模块 250 向特征集生成模块 230 请求能描述目标对象的特征集所属的区域的区域标识 (S211)，特征集生成模块 230 响应该请求将对应的区域标识发送给上下文获取模块 250 (S212)，然后上下文获取模块 250 基于所接收的各区域标识确定能描述目标对象的特征集的局部上下文特征在卷积特征块中所属的区域的区域标识 (S213)，确定局部上下文特征所属的区域的区域标识时，可以描述目标对象的特征集所属的区域的中心点为中心，将其边长扩大 0.5 倍。

在步骤 S214 中，上下文获取模块 250 基于所确定的区域标识从卷积特征块的对应区域提取局部上下文特征。在其他例子中，还可以将卷积特征块确定为全局上下特征。

在步骤 S215 中，上下文获取模块 250 将所提取的上下文特征发送到池化操作模块 260。

在步骤 S216 中，池化操作模块 260 分别对所接收的特征集和上下文特征进行池化操作。

在步骤 S217 中，池化操作模块 260 将池化操作后的特征集和上下文特征输送到像素调整模块 270。

在步骤 S218 中，像素调整模块 270 分别对应各种匹配的像素，分别将接收的特征集合上下文特征调整到匹配的像素。

在某些例子中，池化操作和调整像素的过程可以参照图 2c，图 2c 中的 w 和 h 的乘积代表匹配的像素的具体数值，另外图 2c 仅示出了一个描述目标对象的特征集的池化操作和调整像素过程，该特征集由第一组特征 510、第二组特征 520 和第三组特征 530 构成，第一组特征 510 为全卷积神经网络 220 的第一特征提取通道提取并输出的特征，第二组特征 520 为全卷积神经网络 220 的第二特征提取通道提取并输出的特征，第三组特征 530 为全卷积神经网络 220 的第三特征提取通道提取并输出的特征，在对该特征集进行池化操作前，分别按区域将三组特征分别等分为三部分，如图中两条虚线所划分出的三个区域内的特征，位于顶层的为第一部分，两条虚线中间的为第二部分，位于底层的为第三部分。

在对该特征集进行池化操作时，单独对第一组特征 510 中的第一部分特征进行池化操作，单独对第二组特征 520 中的第二部分特征进行池化操作，单独对第三组特征 530 中的第三部分特征进行池化操作。然后分别将单独池化操作生成的特征调整到匹配的像素，生成像素调整后的特征集，由图 2c 所示的第四组特征 540、第五组特征 550 和第六组特征 560 组成，第一组特征 510 被池化操作和调整像素后变成第四组特征 540，第二组特征 520 被池化操作和调整像素后变成第五组特征 550，第三组特征 530 被池化操作和调整像素后变成第六组特征 540，其他特征集和上下文特征的池化操作和调整像素过程与图 2c 所示类似，在此不再赘述。

在步骤 S219 中，像素调整模块 270 将调整像素后的特征集和上下文特征输送到目标识别模型 280。

在步骤 S220 中，目标识别模型 280 基于输入的特征集和上下文特征对目标对象进行识别，并输出目标对象的目标类别向量和目标位置偏移向量。

具体的目标识别过程可以参阅图 2d,图 2d 中的类别向量 611、612、613、以及位置偏移向量 614、615、616，分别为目标识别模型 280 对应三个分支（本例子仅示出三个分支）的每个特征集、以及该特征集的上下文特征，所产生的类别向量和位置偏移向量。其中，类别向量 611、612、613 的长度可以为 $(C+1)$ ，每一维向量元素可以表示目标对象属于某一类别的概率 P_j ， $j \in \{0, \dots, C\}$ ，0 表示背景类；位置偏移向量 614、615、616 可以是 4 维向量，这个向量的每一维元素表示特征集所属区域的位置偏移向量 $[\Delta_x, \Delta_y, \Delta_w, \Delta_h]$ 。

目标识别模型 280 基于预定的向量筛选准则分别对类别向量 611、612、613 以及未示出的其他类别向量进行筛选，确定一个最终的目标类别向量 621 和目标位置偏移向量 622。

由上述实施例可知，本申请的图像识别方法，从特征图像中选取能够描述目标对象的特征时，从所述特征图像内不同区域选取多个特征集，能有效代表图像中不同位置的同类目标对象，因此能更准确地识别出目标对象。

再者，本申请的图像识别方法，可以针对各目标对象的形状、大小差异，在提取描述目标对象的特征时，从特征图像中大小不同的多个区域选取多个特征集，用像素不同的特征集来分别描述大小和形状各异的目标对象，还可以结合描述目标对象的特征集的局部上下文特征和全局上下文特征对目标对象进行识别，因此，能更准确地获得目标对

象的类别、和/或定位到目标对象。而且，通过对描述目标对象的特征集、特征集的局部上下文特征和全局上下文特征进行池化操作和像素调整，可以进一步降低图像识别过程的计算量，提高识别效率。

此外，将本申请实施例的图像识别方法应用到各场景中时，很可能需要面对互联网数据类似的大规模数据，且应用的实时性要求较高，为了满足这些，可以采用 C/C++ 或汇编语言实现本申请的图像识别方法对应的程序指令。

与前述图像识别方法的实施例相对应，本申请还提供了图像识别装置的实施例。

参见图 3，图 3 是本申请一示例性实施例示出的图像识别装置的逻辑框图，该装置可以包括：图像获取模块 310、特征提取模块 320、特征选取模块 330 和目标识别模块 340。

其中，图像获取模块 310，用于获取待识别的图像。

特征提取模块 320，用于获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征。

特征选取模块 330，用于从所获得的特征图像中选取至少两个描述目标对象的特征集。

目标识别模块 340，用于基于所选取的特征集对目标对象进行识别。

一些例子中，选取的特征集在所述特征图像中所属的区域的尺寸不同。

作为例子，各特征集在所述特征图像中所属的区域的尺寸可以包括：

8*16，8*8，16*16，16*32，16*8，32*64，32*32，32*16，64*32。

一些例子中，本申请的图像识别装置还可以包括：

上下文选取模块，用于从所述特征图像中选取特征集的上下文特征。

目标识别模块 340 还可以用于根据所选取的特征集和上下文特征对目标对象进行识别。

作为例子，所述上下文特征包括局部上下文特征和/或全局上下文特征。

作为例子，特征集的局部上下文特征所属的区域的边长，为该特征集所属的区域的边长的 1.5 倍。

一些例子中，本申请的图像识别装置还可以包括：

池化操作模块，用于对选取的特征集、以及所述特征集的上下文特征分别进行池化操作。

目标识别模块 340 还可以用于根据所述池化操作所得的特征集和上下文特征对目标

对象进行识别。

一些例子中，本申请的图像识别装置还可以包括：

像素调整模块，用于将所述池化操作所得的特征集和上下文特征调整到匹配的像素。

5 目标识别模块 340 还可以用于根据调整后的特征集和上下文特征对目标对象进行识别。

作为例子，所述池化操作模块在对选取的特征集、以及所述特征集的上下文特征分别进行池化操作时，还用于对各特征提取通道所提取的指定特征分别进行池化操作；不同特征提取通道所提取的指定特征在所述特征图像中的坐标不同。

一些例子中，所述目标识别模块 340 还可以用于：

10 将所选取的特征集调整到匹配的像素；

根据调整后的特征集对目标对象进行识别。

作为例子，匹配的像素包括以下至少两种：

3*12，12*3，5*10，10*5，7*7。

15 上述装置中各个单元（或模块）的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程，在此不再赘述。

对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元或模块可以是或者也可以不是物理上分开的，作为单元或模块显示的部件可以是或者也可以不是物理单元或模块，即可以位于一个地方，或者也可以分布到多个网络单元或模块上。可以根据实际的需要选择其中的部分或者全部模块来实现本申请方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

20 本申请图像识别装置的实施例可以应用在电子设备上。具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。一种典型的实现中，电子设备为计算机，计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备、互联网电视、智能机车、无人驾驶汽车、智能冰箱、其他智能家居设备或者这些设备中的任意几种设备的组合。

30 装置实施例可以通过软件实现，也可以通过硬件或者软硬件结合的方式实现。以软件实现为例，作为一个逻辑意义上的装置，是通过其所在电子设备的处理器将非易失性存储器等可读介质中对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言，

如图 4 所示，为本申请图像识别装置所在电子设备的一种硬件结构图，除了图 4 所示的处理器、内存、网络接口、以及非易失性存储器之外，实施例中装置所在的电子设备通常根据该电子设备的实际功能，还可以包括其他硬件，对此不再赘述。电子设备的存储处理器可以是可执行指令的存储器；处理器可以耦合存储器，用于读取所述存储器存储的程序指令，并作为响应，执行如下操作：获取待识别的图像；获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征；从所获得的特征图像中选取至少两个描述目标对象的特征集；基于所选取的特征集对目标对象进行识别。

此外，本申请实施例还提供一种计算机存储介质，所述存储介质中存储有程序指令，所述程序指令包括：

- 10 获取待识别的图像；
- 获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征；
- 从所获得的特征图像中选取至少两个描述目标对象的特征集；
- 基于所选取的特征集对目标对象进行识别。

- 本申请实施例可采用在一个或多个其中包含有程序代码的存储介质（包括但不限于
- 15 磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。计算机可用存储介质包括永久性和非永久性、可移动和非可移动媒体，可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括但不限于：相变内存（PRAM）、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、
 - 20 电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。

在其他实施例中，处理器所执行的操作可以参考上文方法实施例中相关的描述，在此不予赘述。

- 25 以上所述仅为本申请的较佳实施例而已，并不用以限制本申请，凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请保护的范围之内。

权 利 要 求 书

1.一种图像识别方法，其特征在于，包括步骤：

获取待识别的图像；

获得待识别的图像的特征图像，所述特征图像用于描述待识别的图像的特征；

5 从所获得的特征图像中选取至少两个描述目标对象的特征集；

基于所选取的特征集对目标对象进行识别。

2.根据权利要求1所述的方法，其特征在于，在从所获得的特征图像中选取描述目标对象的特征集后，所述方法还包括：

从所述特征图像中选取特征集的上下文特征；

10 所述基于所选取的特征集对目标对象进行识别，包括：

根据所选取的特征集和上下文特征对目标对象进行识别。

3.根据权利要求2所述的方法，其特征在于，所述上下文特征包括局部上下文特征和/或全局上下文特征。

4.根据权利要求3所述的方法，其特征在于，特征集的局部上下文特征所属的区域的边长，为所述特征集所属的区域的边长的1.5倍。

5.根据权利要求2所述的方法，其特征在于，所述根据所选取的特征集和上下文特征对目标对象进行识别，包括：

对选取的特征集、以及所述特征集的上下文特征分别进行池化操作；

根据所述池化操作所得的特征集和上下文特征对目标对象进行识别。

20 6.根据权利要求5所述的方法，其特征在于，在对选取的特征集、以及所述特征集的上下文特征分别进行池化操作时，对各特征提取通道所提取的指定特征分别进行池化操作；不同特征提取通道所提取的指定特征在所述特征图像中的坐标不同。

7.根据权利要求5所述的方法，其特征在于，在对选取特征集、以及特征集的上下文特征分别进行池化操作后，所述方法还包括：

25 将所述池化操作所得的特征集和上下文特征调整到匹配的像素；

所述根据所述池化操作所得的特征集和上下文特征对目标对象进行识别，包括：

根据调整后的特征集和上下文特征对目标对象进行识别。

8.根据权利要求1所述的方法，其特征在于，所述基于所选取的特征集对目标对象进行识别，包括：

30 将所选取的特征集调整到匹配的像素；

根据调整后的特征集对目标对象进行识别。

9.根据权利要求 7 或 8 所述的方法，其特征在于，匹配的像素包括以下至少两种：

3*12，12*3，5*10，10*5，7*7。

10.根据权利要求 1 所述的方法，其特征在于，选取的特征集在所述特征图像中所属
5 的区域尺寸不同。

11.根据权利要求 10 所述的方法，其特征在于，特征集在所述特征图像中所属的区域尺寸包括：

8*16，8*8，16*16，16*32，16*8，32*64，32*32，32*16，64*32。

12.根据权利要求 1 所述的方法，其特征在于，所述特征图像通过以下任一算法获得：
10 卷积神经网络模型，分类器，多级网络结构。

13.根据权利要求 1 所述的方法，其特征在于，描述目标对象的特征集的卷积运算结果大于预定阈值。

14.根据权利要求 1 所述的方法，其特征在于，描述目标对象的特征集的卷积运算结果大于其他非目标对象的特征集的卷积运算结果。

15.根据权利要求 1 所述的方法，其特征在于，识别结果包括目标对象的目标类别向量和目标位置偏移向量，所述目标类别向量为各特征集对应的类别向量中最大的向量元素所属的类别向量，所述目标位置偏移向量为所述目标类别向量对应的特征集的位置偏移向量。

16.根据权利要求 1 所述的方法，其特征在于，待识别的图像为拍摄智能冰箱的储物
20 间所得的图像，所述目标对象为所述智能冰箱的储物间储存的食材。

17.一种电子设备，其特征在于，包括：

处理器；

存储处理器可执行指令的存储器；

其中，所述处理器耦合于所述存储器，用于读取所述存储器存储的程序指令，并作为响应，执行如下操作：

获取待识别的图像；

获得待识别的图像的特征图像；

从所获得的特征图像中选取至少两个描述目标对象的特征集，选取的特征集在所述特征图像中所属的区域尺寸不同；

30 基于所选取的特征集对目标对象进行识别。

18.根据权利要求 17 所述的电子设备, 其特征在于, 所述处理器还被配置为执行以下操作:

从所述特征图像中选取特征集的上下文特征;

根据所选取的特征集和上下文特征对目标对象进行识别。

5 19.根据权利要求 18 所述的电子设备, 其特征在于, 所述上下文特征包括局部上下文特征和/或全局上下文特征。

20.根据权利要求 19 所述的电子设备, 其特征在于, 特征集的局部上下文特征所属的区域的边长, 为该特征集所属的区域的边长的 1.5 倍。

10 21.根据权利要求 18 所述的电子设备, 其特征在于, 所述处理器还被配置为执行以下操作:

对选取的特征集、以及所述特征集的上下文特征分别进行池化操作;

根据所述池化操作所得的特征集和上下文特征对目标对象进行识别。

22.根据权利要求 21 所述的电子设备, 其特征在于, 所述处理器还被配置为执行以下操作:

15 在对选取的特征集、以及所述特征集的上下文特征分别进行池化操作时, 对各特征提取通道所提取的指定特征分别进行池化操作; 不同特征提取通道所提取的指定特征在所述特征图像中的坐标不同。

23.根据权利要求 21 所述的电子设备, 其特征在于, 所述处理器还被配置为执行以下操作:

20 将所述池化操作所得的特征集和上下文特征调整到匹配的像素;

根据调整后的特征集和上下文特征对目标对象进行识别。

24.根据权利要求 17 所述的电子设备, 其特征在于, 所述处理器还被配置为执行以下操作:

将所选取的特征集调整到匹配的像素;

25 根据调整后的特征集对目标对象进行识别。

25.根据权利要求 23 或 24 所述的电子设备, 其特征在于, 匹配的像素包括以下至少两种:

3*12, 12*3, 5*10, 10*5, 7*7。

30 26.根据权利要求 17 所述的电子设备, 其特征在于, 选取的特征集在所述特征图像中所属的区域的尺寸不同。

27.根据权利要求 26 所述的电子设备, 其特征在于, 特征集在所述特征图像中所属的区域的尺寸包括:

8*16, 8*8, 16*16, 16*32, 16*8, 32*64, 32*32, 32*16, 64*32。

28.根据权利要求 17 所述的电子设备, 其特征在于, 所述特征图像通过以下任一算法获得:

卷积神经网络模型, 分类器, 多级网络结构。

29.根据权利要求 17 所述的电子设备, 其特征在于, 描述目标对象的特征集的卷积运算结果大于预定阈值。

30.根据权利要求 17 所述的电子设备, 其特征在于, 描述目标对象的特征集的卷积运算结果大于其他非目标对象的特征集的卷积运算结果。

31.根据权利要求 17 所述的电子设备, 其特征在于, 识别结果包括目标对象的目标类别向量和目标位置偏移向量, 所述目标类别向量为各特征集对应的类别向量中最大的向量元素所属的类别向量, 所述目标位置偏移向量为所述目标类别向量对应的特征集的位置偏移向量。

32.根据权利要求 17 所述的电子设备, 其特征在于, 待识别的图像为拍摄智能冰箱的储物间所得的图像, 所述目标对象为所述智能冰箱的储物间储存的食材。

33.一种图像识别装置, 其特征在于, 包括:

图像获取模块, 用于获取待识别的图像;

特征提取模块, 用于获得待识别的图像的特征图像, 所述特征图像用于描述待识别的图像的特征;

特征选取模块, 用于从所获得的特征图像中选取至少两个描述目标对象的特征集;

目标识别模块, 用于基于所选取的特征集对目标对象进行识别。

34.根据权利要求 33 所述的装置, 其特征在于, 选取的特征集在所述特征图像中所属的区域的尺寸不同。

35.根据权利要求 34 所述的装置, 其特征在于, 各特征集在所述特征图像中所属的区域的尺寸包括:

8*16, 8*8, 16*16, 16*32, 16*8, 32*64, 32*32, 32*16, 64*32。

36.根据权利要求 33 所述的装置, 其特征在于, 所述装置还包括:

上下文选取模块, 用于从所述特征图像中选取特征集的上下文特征;

所述目标识别模块还用于根据所选取的特征集和上下文特征对目标对象进行识别。

37.根据权利要求 36 所述的装置，其特征在于，所述上下文特征包括局部上下文特征和/或全局上下文特征。

38.根据权利要求 36 所述的装置，其特征在于，所述装置还包括：

池化操作模块，用于对选取的特征集、以及所述特征集的上下文特征分别进行池化操作；

所述目标识别模块还用于根据所述池化操作所得的特征集和上下文特征对目标对象进行识别。

39.根据权利要求 38 所述的装置，其特征在于，所述池化操作模块在对选取的特征集、以及所述特征集的上下文特征分别进行池化操作时，还用于对各特征提取通道所提取的指定特征分别进行池化操作；不同特征提取通道所提取的指定特征在所述特征图像中的坐标不同。

40.根据权利要求 38 所述的装置，其特征在于，所述目标识别模块还用于：

将所述池化操作所得的特征集和上下文特征调整到匹配的像素；

根据调整后的特征集和上下文特征对目标对象进行识别。

41.根据权利要求 33 所述的装置，其特征在于，所述目标识别模块还用于：

将所选取的特征集调整到匹配的像素；

根据调整后的特征集对目标对象进行识别。

42.根据权利要求 40 或 41 所述的装置，其特征在于，匹配的像素包括以下至少两种：

3*12，12*3，5*10，10*5，7*7。

43.一个或多个机器可读介质，其特征在于，其上存储有指令，当由一个或多个处理器执行时，使得终端设备执行如权利要求 1-16 中任一项所述的方法。

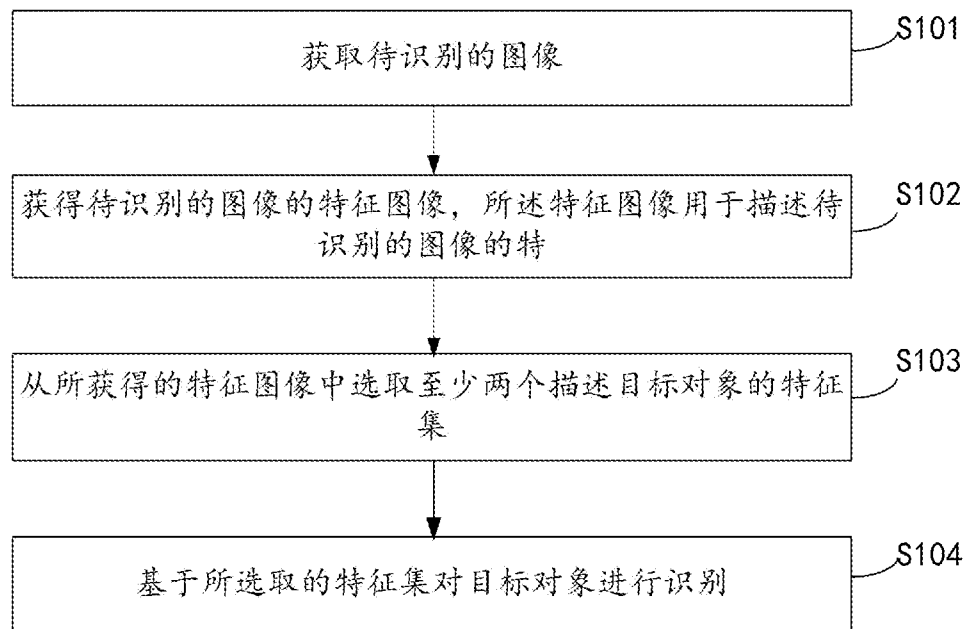


图 1

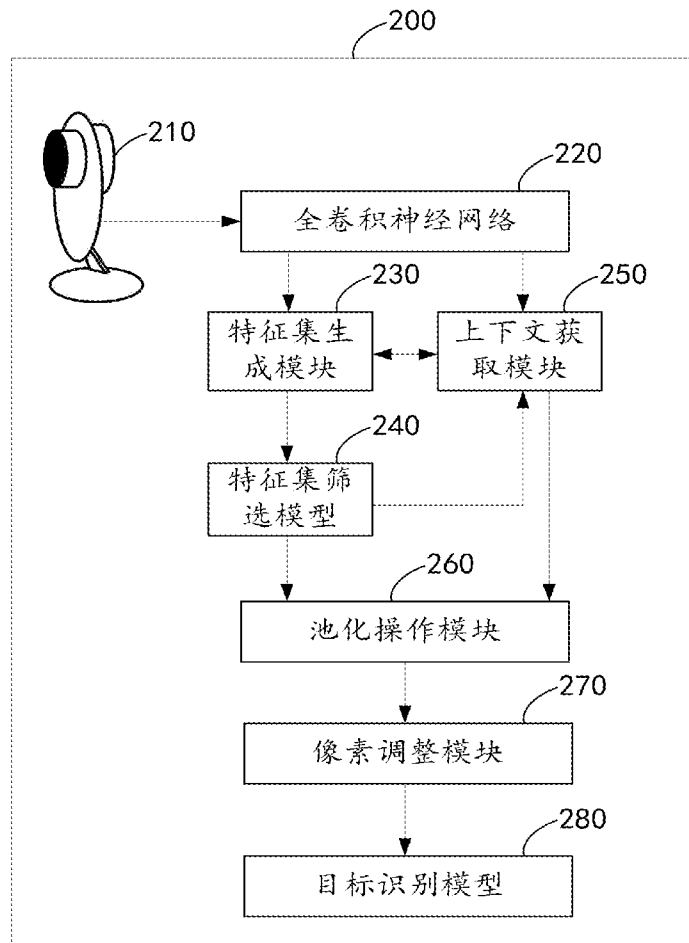


图 2a

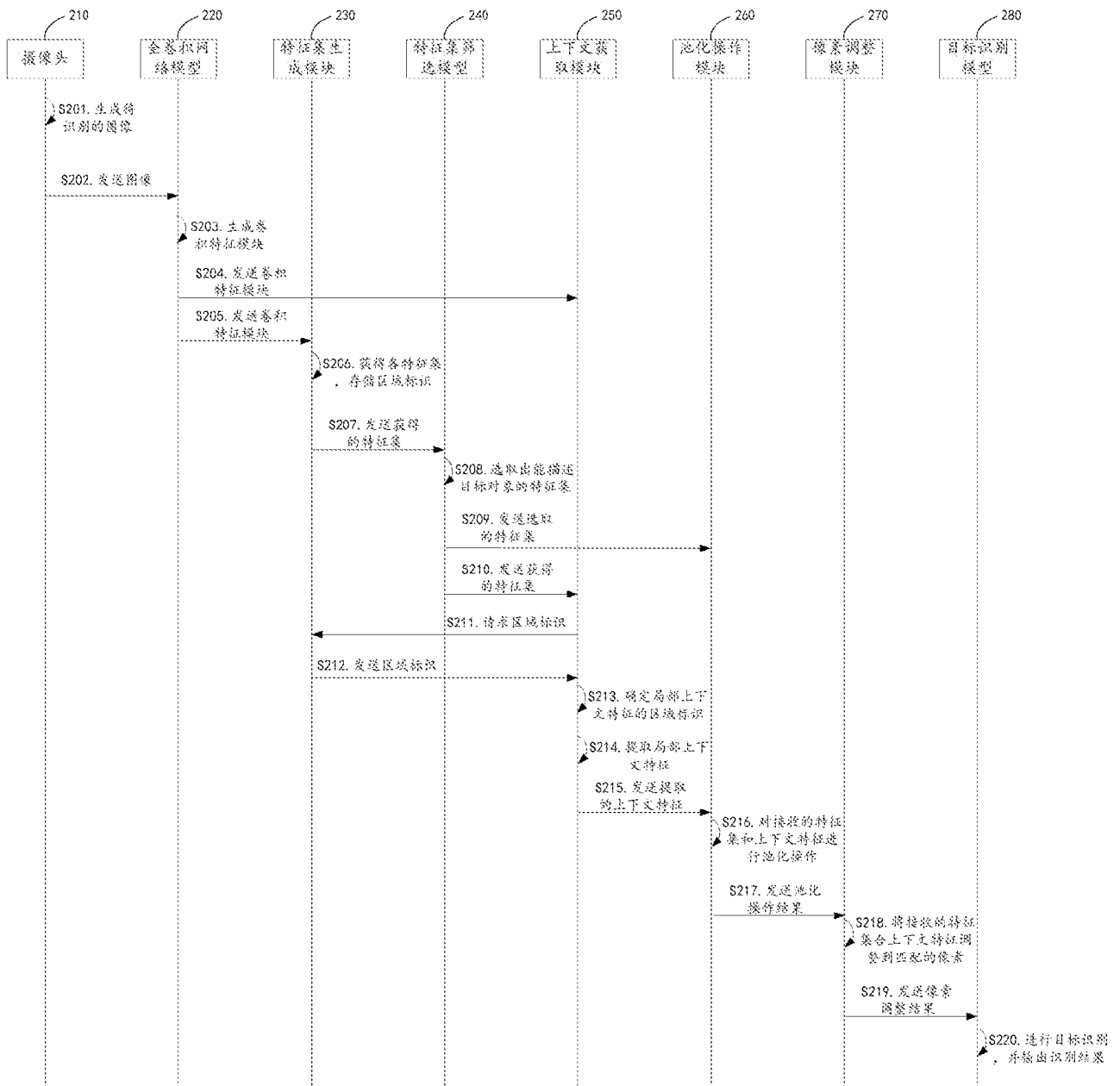


图 2b

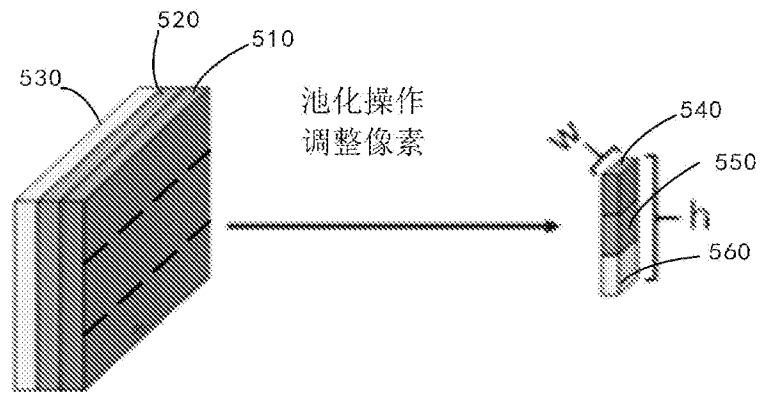


图 2c

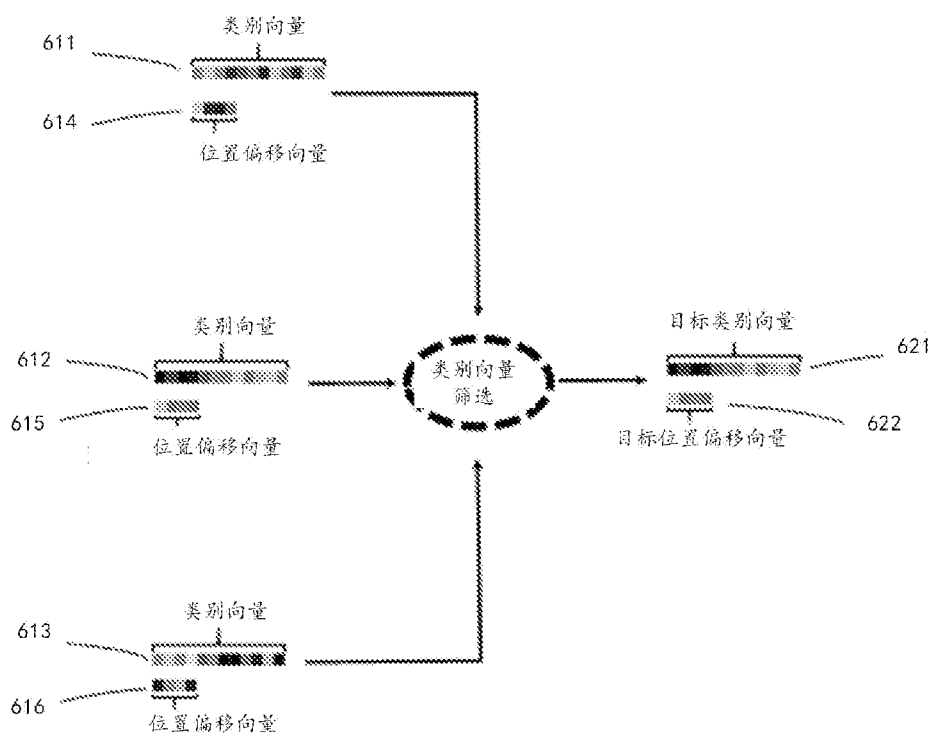


图 2d

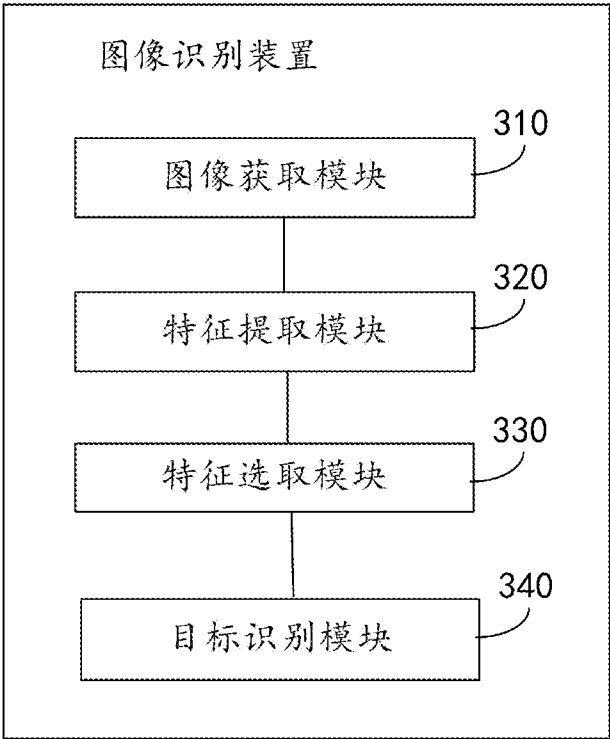


图 3

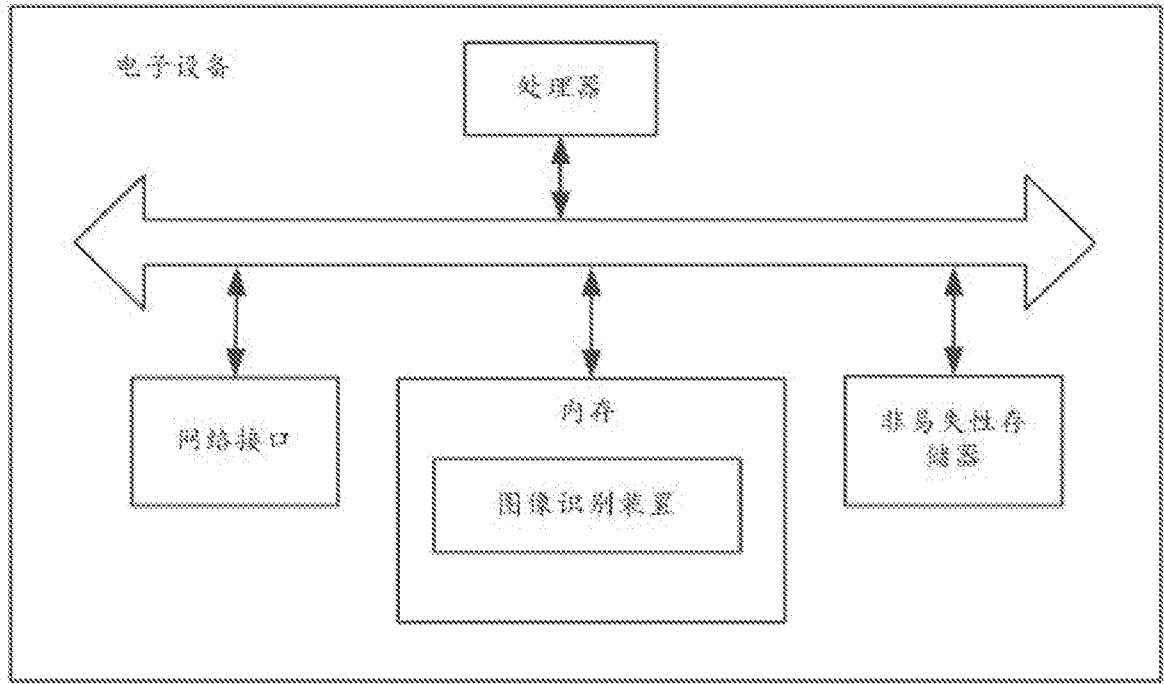


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/093350

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)

G06K9/-

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)

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT: 图像, 识别, 特征, 卷积神经网络, image, picture, recognize, recognise, feature, characteristic, CNN, convolution neural network

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106529467 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 22 March 2017 (2017-03-22) description, paragraphs 4-13 and 30-35	1-43
X	CN 106803090 A (CHINA UNIONPAY CO., LTD.) 06 June 2017 (2017-06-06) description, paragraphs 66-88	1-43
X	US 2003179931 A1 (SUN, HUNGMIN) 25 September 2003 (2003-09-25) description, paragraphs 22-43	1-43



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

12 September 2018

Date of mailing of the international search report

29 September 2018

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/CN2018/093350

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106529467	A	22 March 2017	None			
CN	106803090	A	06 June 2017	None			
US	2003179931	A1	25 September 2003	TW	I281126	B	11 May 2007

国际检索报告

国际申请号

PCT/CN2018/093350

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)

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))

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT: 图像, 识别, 特征, 卷积神经网络, image, picture, recognize, recognise, feature, characteristic, CNN, convolution neural network

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 106529467 A (南京邮电大学) 2017年 3月 22日 (2017 - 03 - 22) 说明书第4-13、30-35段	1-43
X	CN 106803090 A (中国银联股份有限公司) 2017年 6月 6日 (2017 - 06 - 06) 说明书第66-88段	1-43
X	US 2003179931 A1 (SUN, HUNG MING) 2003年 9月 25日 (2003 - 09 - 25) 说明书第22-43段	1-43

☐ 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 9月 12日

国际检索报告邮寄日期

2018年 9月 29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

受权官员

王国海

传真号 (86-10) 62019451

电话号码 86-(20)-28958137

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/093350

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	106529467	A	2017年 3月 22日	无			
CN	106803090	A	2017年 6月 6日	无			
US	2003179931	A1	2003年 9月 25日	TW	I281126	B	2007年 5月 11日