

(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 09.2018
2018 年 9 月 20 日 20.

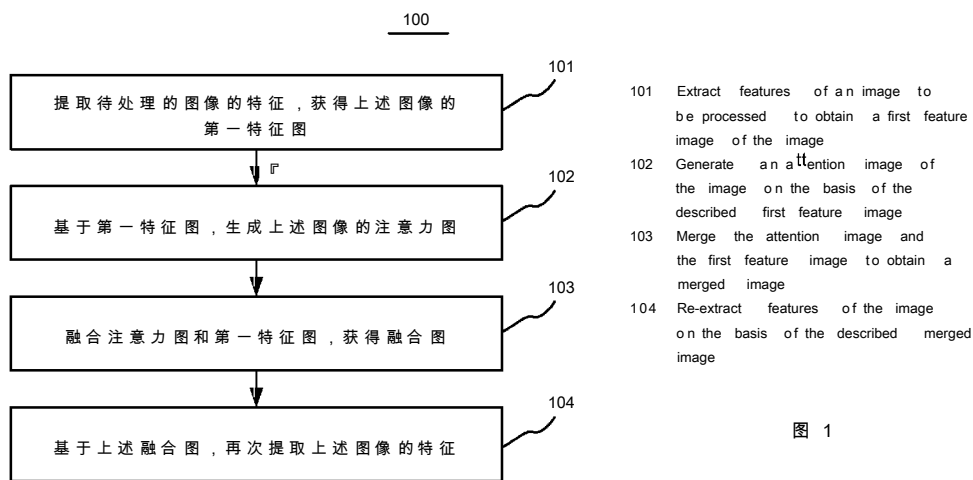
(10) 国际公布号

WO 2018/166438 A1

- (51) 国际专利分类号 :
G06K 9/32 (2006.01) G06T 7/00 (2017.01)
G06K 9/62 (2006 .01)
- (21) 国际申请号 : PCT/CN201 8/0788 10
- (22) 国际申请日 : 2018 年 3 月 13 日 (13.03.2018)
- (25) 申请语言 : 中文
- (26) 公布语言 : 中文
- (30) 优先权 :
201710145253.1 2017 年 3 月 13 日 (13.03.2017) CN
- (71) 申请人 北京市商汤科技开发有限公司 (BEIJING SENSETIME TECHNOLOGY DEVELOPMENT CO., LTD.) [CN/CN] ; 中国北京市海淀区中关村东路 1 号院 3 号楼 7 层 710-712 房间 , Beijing 100084 (CN) 。
- (72) 发明人 :王飞 (WANG, Fei) ; 中国北京市海淀区中关村东路 1 号院 3 号楼 7 层 710-712 房间 , Beijing 100084 (CN), 钱晨 (QIAN, Chen) ; 中国北京市海淀区中关村东路 1 号院 3 号楼 7 层 710-712 房间 ,Beijing 100084 (CN) 。
- (74) 代理人 :北京思源智汇知识产权代理有限公司 (BEIJING SIYUANZHUI INTELLECTUAL PROPERTY AGENCY) ; 中国北京市海淀区中关村东路 18 号财智国际大厦 B 座 502 室 ,Beijing 100083 (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,

(54) Title: IMAGE PROCESSING METHOD AND DEVICE AND ELECTRONIC DEVICE

(54) 发明名称 : 图像处理方法、装置及电子设备



(57) Abstract: Disclosed in the present application are an image processing method and device, and an electronic device. One embodiment of the image processing method comprises: extracting features of an image to be processed to obtain a first feature image of the image; generating an attention image of the image on the basis of the described first feature image; merging the attention image and the first feature image to obtain a merged image; and re-extracting features of the image on the basis of the described merged image. The embodiment introduces an attention mechanism into image processing, thus effectively improving the efficiency of acquiring information from the image.

(57) 摘要 : 本申请公开了图像处理方法、装置及电子设备。所述图像处理方法的一实施方式包括 : 提取待处理的图像的特征, 获得上述图像的第一特征图 ; 基于上述第一特征图, 生成上述图像的注意力图 ; 融合上述注意力图和上述第一特征图, 获得融合图 ; 基于上述融合图, 再次提取上述图像的特征。该实施方式实现了将注意力机制引入到图像处理中, 有效提高了从图像中获取信息的效率。



WO 2018/166438 A1

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, , MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

图像处理方法、装置及电子设备

本申请要求在 2017 年 03 月 13 日提交中国专利局、申请号为 CN201710145253.1 、发明名称为“图像处理方法、装置及电子设备”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。

技术领域

本申请涉及计算机视觉技术，具体涉及一种图像处理方法、装置及电子设备。

背景技术

计算机视觉是使用计算机及相关设备对生物视觉的一种模拟。在计算机视觉领域，视觉注意力机制一直为人们所关注。人类可以通过视觉注意力机制快速扫描视野范围内的整体区域，并将与目标物体无关的区域滤除，而仅仅关注目标物体所在的区域。因此，注意力机制极大的提高了人类对于目标物体的信息获取效率。

发明内容

本申请实施例提出了一种图像处理的技术方案。

根据本申请实施例的一个方面，提供了一种图像处理方法，包括：提取待处理的图像的特征，获得上述图像的第一特征图；基于上述第一特征图，生成上述图像的注意力图；融合上述注意力图和上述第一特征图，获得融合图；基于上述融合图，再次提取上述图像的特征。

在一些实施例中，上述基于上述第一特征图，生成上述图像的注意力图，包括：对上述第一特征图依次进行 N 次下采样处理，其中， N 为大于或等于 1 的整数；对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得上述图像的注意力图，其中，上述注意力图的分辨率与上述第一特征图的分辨率相同。

在一些实施例中，上述对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，包括：对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，其中， n 为大于 1 且小于 N 的整数；对卷积操作后的特征图进行第 $n+1$ 次上采样处理。

在一些实施例中，上述对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，包括：对上述第 $N-n$ 次下采样处理后的特征图进行卷积处理，得到卷积图；将上述卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；对相加后的特征图进行卷积操作。

在一些实施例中，上述对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，还包括：对第 N 次下采样处理后的特征图进行至少一次卷积操作；对最后一次卷积操作后的特征图进行第 1 次上采样处理。

在一些实施例中，上述融合上述注意力图和上述第一特征图，获得融合图，包括：对上述第一特征图进行至少一次卷积操作；融合上述注意力图和最后一次卷积操作后的第一特征图，获得上述融合图。

在一些实施例中，上述融合上述注意力图和上述第一特征图，获得融合图，包括：至

少对上述注意力图进行归一化处理；融合归一化处理后的注意力图和上述第一特征图，获得上述融合图。

在一些实施例中，上述至少对上述注意力图进行归一化处理，包括：对上述注意力图依次进行至少一次卷积处理；对最后一次卷积处理后的注意力图进行归一化处理。

在一些实施例中，上述融合上述注意力图和上述第一特征图，获得融合图，包括：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得上述融合图。

在一些实施例中，上述融合上述注意力图和上述第一特征图，获得融合图，包括：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得相乘图；将上述相乘图中至少一个像素的特征值与上述第一特征图中相应像素的特征值相加，获得上述融合图。

在一些实施例中，在上述基于上述融合图，再次提取上述图像的特征之后，上述方法还包括以下至少之一：根据再次提取的上述图像的特征，检测或识别上述图像中包括的物体；根据再次提取的上述图像的特征，确定上述图像中包括的物体的类别；根据再次提取的上述图像的特征，对上述图像进行分割。

根据本申请实施例的另一个方面，提供了一种图像处理装置，上述装置包括：第一特征提取单元，用于提取待处理的图像的特征，获得上述图像的第一特征图；注意力提取单元，用于基于上述第一特征图，生成上述图像的注意力图；融合单元，用于融合上述注意力图和上述第一特征图，获得融合图；第二特征提取单元，用于基于上述融合图，再次提取上述图像的特征。

在一些实施例中，上述注意力提取单元包括：下采样模块，用于对上述第一特征图依次进行 N 次下采样处理，其中， N 为大于或等于1的整数；上采样模块，用于对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得上述图像的注意力图，其中，上述注意力图的分辨率与上述第一特征图相同。

在一些实施例中，上述上采样模块用于：对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，其中， n 为大于1且小于 N 的整数；对卷积操作后的特征图进行第 $n+1$ 次上采样处理，获得所述图像的注意力图。

在一些实施例中，上述上采样模块对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作时，用于：对上述第 $N-n$ 次下采样处理后的特征图进行卷积处理，得到卷积图；将上述卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；对相加后的特征图进行卷积操作，获得所述图像的注意力图。

在一些实施例中，上述上采样模块用于：对第 N 次下采样处理后的特征图进行至少一次卷积操作；对最后一次卷积操作后的特征图进行第1次上采样处理。

在一些实施例中，上述装置还包括：第二卷积单元，用于对上述第一特征图进行至少一次卷积操作；上述融合单元，用于融合上述注意力图和最后一次卷积操作后的第一特征图，获得上述融合图。

在一些实施例中，上述装置还包括：归一化单元，用于至少对上述注意力图进行归一化处理；上述融合单元，用于融合归一化处理后的注意力图和上述第一特征图，获得上述融合图。

在一些实施例中，上述装置还包括：第二卷积单元，用于对上述注意力图依次进行至

少一次卷积处理；上述归一化单元，用于对最后一次卷积处理后的注意力图进行归一化处理。

在一些实施例中，上述融合单元用于：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得上述融合图。

在一些实施例中，上述融合单元用于：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得相乘图；将上述相乘图中至少一个像素的特征值与上述第一特征图中相应像素的特征值相加，获得上述融合图。

在一些实施例中，上述装置还包括以下至少之一：检测单元，用于根据再次提取的上述图像的特征，检测或识别上述图像中包括的物体；分类单元，用于根据再次提取的上述图像的特征，确定上述图像中包括的物体的类别；分割单元，用于根据再次提取的上述图像的特征，对上述图像进行分割。

根据本申请实施例的又一个方面，提供了一种计算机可读存储介质，其上存储有计算机指令，该指令被执行时实现本申请任一实施方式所述图像处理方法中各步骤的操作。

根据本申请实施例的再一个方面，提供了一种电子设备，包括：处理器和存储器；所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行本申请任一实施方式所述图像处理方法对应的。

根据本申请实施例的再一个方面，提供了一种计算机程序，包括计算机可读代码，当所述计算机可读代码在设备上运行时，所述设备中的处理器执行用于实现本申请任一实施方式所述图像处理方法中各步骤的指令。

本申请实施例提供的图像处理方法、装置及电子设备、程序和介质，首先提取待处理的图像的特征，得到上述图像的第一特征图，基于上述第一特征图，生成上述图像的注意力图，再将注意力图和第一特征图融合，最后再基于得到的融合图再次提取上述图像的特征，从而实现了将注意力机制引入到图像处理中，有效地提高了从图像中获取信息的效率。

下面通过附图和实施例，对本申请实施例的技术方案做进一步的详细描述。

附图说明

构成说明书的一部分的附图描述了本申请的实施例，并且连同描述一起用于解释本申请的原理。

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 是根据本申请的用于检测目标物体的方法的一个实施例的流程图；

图 2 是根据本申请的用于检测目标物体的方法的生成注意力图的流程示意图；

图 3a 是图 2 所示的流程对应的一种网络结构的示意图；

图 3b 是图 2 所示的流程对应的另一种网络结构的示意图；

图 4 是根据本申请的用于检测目标物体的方法的融合注意力图和第一特征图的流程示意图；

图 5a 是图 4 所示的流程对应的神经网络的结构示意图；

图 5b 是图 5a 所示的神经网络的处理过程示意图；

图 6 是图 5a 所示的神经网络构成的深度卷积神经网络的结构示意图；

图 7 是根据本申请的图像处理装置的一个实施例的结构示意图；

图 8 是适于用来实现本申请实施例的终端设备或服务器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实施例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申请，而非对该申请实施例的限定。同时，应当明白，为了便于描述，附图中所示出的至少一个部分的尺寸并不是按照实际的比例关系绘制的。

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申请实施例相关的部分。

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。对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论，但在适当情况下，所述技术、方法和设备应当被视为说明书的一部分。

应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步讨论。

本申请实施例可以应用于终端设备、计算机系统、服务器等电子设备，其可与众多其它通用或专用计算系统环境或配置一起操作。适于与终端设备、计算机系统、服务器等电子设备一起使用的众所周知的终端设备、计算系统、环境和/或配置的例子包括但不限于：个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境，等等。

终端设备、计算机系统、服务器等电子设备可以在由计算机系统执行的计算机系统可执行指令（诸如程序模块）的一般语境下描述。通常，程序模块可以包括例程、程序、目标程序、组件、逻辑、数据结构等等，它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施，分布式云计算环境中，任务是由通过通信网络链接的远程处理设备执行的。在分布式云计算环境中，程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

下面将参考附图并结合实施例来详细说明本申请。

参考图 1，示出了根据本申请的图像处理方法的一个实施例的流程 100。本实施例的图像处理方法，包括以下步骤：

步骤 101，提取待处理的图像的特征，获得上述图像的第一特征图。

在本申请实施例中，待处理的图像可以是包含各种物体、建筑、人物、景色的图像，其可以是静态图像，也可以是视频中的一帧图像。在提取上述待处理的图像的特征时，可以利用神经网络中的一个或多个卷积层来实现。提取上述图像的特征，得到上述图像的第一特征图。

在一个可选示例中，该步骤 101 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的第一特征提取单元 701 执行。

步骤 102，基于第一特征图，生成上述图像的注意力图。

在获得上述待处理的图像的第一特征图后，可以对上述特征图进行一系列的处理，得到上述图像的注意力图。上述一系列的处理例如可以是：对第一特征图进行多次下采样处理、对第一特征图交替进行下采样和上采样处理、对第一特征图进行多次下采样处理后再进行多次上采样处理、对第一特征图进行卷积或平均池化等等。基于特征图生成注意力图的方式可采用本申请实施例下文提供的任一种方法，也可采用基于注意力机制生成注意力

图的其他现有方法，本申请实施例并不限制。基于计算机视觉技术的注意力机制生成的注意力图可以包含上述待处理的图像的全局信息，并且对上述全局信息中还包括注意力关注的特征的权重信息，可以模拟人的视觉系统，重点关注图像当中权重大的特征信息，而不丧失图像的全局信息。

在一个可选示例中，该步骤 102 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的注意力提取单元 702 执行。

步骤 103，融合注意力图和第一特征图，获得融合图。

在得到注意力图和第一特征图后，可以将二者进行融合，以获得上述待处理的图像中包含的物体、人物、景色的有效信息，也就是说，利用上述融合图，能够更有效地表达待处理的图像中的物体、人物、景色等信息。

在一个可选示例中，该步骤 103 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的融合单元 703 执行。

步骤 104，基于上述融合图，再次提取上述图像的特征。

本申请各实施例中，在得到上述融合图后，可以再次提取上述图像的特征，得到的特征可以进行进一步的应用。在再次提取上述图像的特征时，可以利用多个级联的卷积层或残差单元来实现。

在一个可选示例中，该步骤 104 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的第二特征提取单元 704 执行。

在一种可选示例中，本申请各实施例的图像处理方法，可以由神经网络来实现。可以理解的是，为了实现更好的提取上述待处理的图像的特征，可以多次重复上述神经网络，以形成更深的神经网络。这样，能够得到待处理的图像的更全面的全局信息，从而能够提高对待处理的图像的特征表达能力。

可以理解的是，上述神经网络在使用前可以利用带有标注信息的图片来训练神经网络，并根据训练结果反向传播以修改神经网络的参数，完成神经网络的训练，从而得到上述神经网络。

本申请的上述实施例提供的图像处理方法，首先提取待处理的图像的特征，得到上述图像的第一特征图，基于上述第一特征图，生成上述图像的注意力图，再将注意力图和第一特征图融合，最后再基于得到的融合图再次提取上述图像的特征，从而实现了将注意力机制引入到图像处理中，有效地提高了从图像中获取信息的效率。

参考图 2，其示出了根据本申请的图像处理方法的生成注意力图的流程 200。如图 2 所示，本实施例中通过以下步骤生成待处理的图像的注意力图。

步骤 201，对第一特征图依次进行 N 次下采样处理。

本申请各实施例中，N 为整数，且 $N \geq 1$ 。在对步骤 101 中得到的第一特征图进行下采样处理，可以获得第一特征图的全局信息。但下采样的次数越多，得到的全局信息图的维度与第一特征图的维度相差就越大。本实施例中，上述下采样操作可以通过但不限于以下方式来实现：利用不同步长的池化层、不同步长的卷积层、平均池化层，来进行下采样操作。例如，利用步长为 2 的池化层对第一特征图进行下采样时，假设第一特征图的分辨率为 224×224 ，经过 3 次下采样后，得到的特征图的分辨率为 28×28 。由于经过 N 次下采样后得到的特征图与第一特征图的分辨率不同，虽然经过 N 次下采样后得到的特征图中包含了第一特征图的全局信息，但其无法指导分辨率为 224×224 的特征进行学习。

在一个可选示例中，该步骤 201 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的下采样模块执行。

步骤 202，对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得图像的注意力图。

本实施例中，在得到 N 次下采样处理后的特征图后，可以对上述特征图进行 N 次上采样处理。这样经 N 次上采样处理后的特征图的分辨率与第一特征图的分辨率相同。本实施例中，上述上采样操作可以通过但不限于以下方式来实现：利用解卷积层、最邻近插值层、线性插值层，进行上采样操作。例如，经过 3 次下采样后，得到的特征图的分辨率为 28×28 ，此特征图再经过 3 次上采样处理后，得到的注意力图的分辨率与第一特征图的分辨率相同。

在本实施例的一些可选的实现方式中，在每次下采样及每次上采样处理后，还可以对经下采样处理得到的特征图以及对经上采样处理得到的特征图进行卷积操作。在对经上采样处理得到的特征图进行卷积操作后，再进行下一次的上采样处理。即对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作；对卷积操作后的特征图进行第 n+1 次上采样处理。其中，n 为正整数，且 $1 < n < N$ 。

在一个可选示例中，该步骤 202 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的上采样模块执行。

可以理解的是，本实现方式中，在对经第 N-1 次上采样处理得到的特征图进行卷积操作后，进行第 N 次上采样处理，并且，不需要对经第 N 次上采样处理得到的注意力图进行卷积操作。本实现方式中的卷积操作，可以利用卷积层来实现，也可以利用残差单元来实现，本实现方式对此不做限定。上述残差单元可以是包括两个以上卷积层的网络结构。

本实现方式的图像处理方法，不仅注意力图的分辨率与第一特征图的分辨率相同，得到的注意力图可以用于指导后续对第一特征图中的特征进行学习；同时，通过在每次下采样处理和上采样处理后，对得到的特征图进行卷积操作，能够更好的学习得到的不同维度下特征图中的特征。

在本实施例的一些可选的实现方式中，在对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作时，还可以通过以下步骤来实现：

对第 N-n 次下采样处理后的特征图进行卷积处理，得到卷积图；将卷积图中至少一个像素（例如，各像素）的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；对相加后的（即：相加得到的）特征图进行卷积操作。

本实现方式中，n 为正整数，且 $1 < n < N$ 。N 可以为一个预设值，也可以是根据第一特征图的分辨率计算得到的值，并且可通过如下计算方法来确定 N 的值：设置下采样处理后得到的特征图的最小分辨率，根据第一特征图的分辨率以及上述最小分辨率来确定能够进行的下采样的次数，即确定 N 的值。例如第一特征图的分辨率为 56×56 ，设置的最小分辨率为 7×7 ，每次下采样操作后，得到的特征图的分辨率降为下采样之前的特征图的四分之一，则确定 N 的值最大为 3。

由于本实现方式可以对第一特征图进行 N 次下采样处理，然后再进行 N 次上采样处理，则在下采样处理的过程中以及上采样处理的过程中会得到分辨率相同的特征图。为了获得待处理的图像的更深层次的信息，可以对第 N-n 次下采样处理后的特征图进行卷积处理，得到卷积图。然后将卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相

应像素的特征值相加，然后对相加后的特征图进行卷积操作。

本实现方式的图像处理方法，将在下采样处理的过程中以及上采样处理的过程中得到的分辨率相同的特征图相加，能够获得待处理的图像的更深层次的信息。

在本实施例的一些可选的实现方式中，在对第 N 次下采样处理后的特征图依次进行 N 次上采样处理时，还可以包括以下步骤：

对第 N 次下采样处理后的特征图进行至少一次卷积操作，并对最后一次卷积操作后的特征图进行第 1 次上采样处理。

本实现方式中，在最后一次下采样处理后，对得到的特征图进行卷积操作，得到全局信息图，然后对上述全局信息图再次进行卷积操作，然后对再次进行卷积操作后的特征图进行第一次上采样处理。这样，下采样处理过程与上采样处理过程为对称的两个处理过程，最后得到的注意力图更能够反映待处理的图像中包含的特征信息。

在一种可选示例中，本实现方式的图像处理方法可以采用图 3a 所示的网络结构来实现。如图 3a 所示，本实现方式的网络结构包括输入层 301、级联的多个卷积层 302、多个下采样单元 303、多个上采样单元 304、多个残差单元 305 以及输出层 306。可以理解的是，本实现方式中的卷积操作由残差单元来实现。

其中，输入层 301 用于输入待处理的图像。级联的多个卷积层 302 用于提取待处理的图像的特征，得到第一特征图。可以理解的是，上述级联的多个卷积层 302 也可以由残差单元来实现。

每个下采样单元 303 包括一个下采样层 3031 和一个残差单元 3032，下采样层 3031 用于对级联的多个卷积层 302 得到的第一特征图进行下采样处理，每经过一次下采样处理得到的特征图的分辨率就相当于第一特征图分辨率的四分之一。每个残差单元 3032 用于在每次下采样处理后对下采样处理后的特征图进行卷积操作，以提取下采样处理后的特征图的特征。举例来说，第一特征图的分辨率为 56×56 ，经过上述下采样层 3031 的一次下采样处理后，得到的图像的分辨率为 28×28 ，残差单元 3032 提取上述 28×28 的图像的特征。假如上述网络结构中包括 3 个下采样单元 303，则经过第 3 个下采样单元处理后的得到的特征图的分辨率为 7×7 ，且第 3 个下采样单元的残差单元提取了上述 7×7 的图像的特征，得到了第一特征图的全局信息图。可以理解的是，上述网络结构中的下采样单元 303 的数目可以是任意的，本实现方式并不对此进行限定。同时，可以理解的是，且每个下采样单元中的残差单元的结构可以相同，即包括相同数量的卷积层，但各个卷积层的参数不同。

在得到第一特征图的全局信息图后，利用上采样单元 304 对上述全局信息图继续处理。上述上采样单元 304 可以包括一个残差单元 3041 和上采样层 3042，残差单元 3041 与残差单元 3032 的结构可以相同，但参数不同。残差单元 3041 用于提取上述残差单元 3032 得到的全局信息图的特征，在提取了上述特征后，经上采样层 3042 的上采样处理，得到的特征图的分辨率为全局信息图分辨率的四倍。经过与下采样次数相同的上采样处理后，最后得到的注意力图的分辨率与第一特征图的分辨率相同。

可以理解的是，图 3a 中，下采样单元 303 中的下采样层可以通过最大池化层来实现，上采样单元 304 中的上采样层可以通过双线性插值层来实现。另外，每个下采样单元 303 和上采样单元 304 中还可以包括多个残差单元。可参考图 3b，如图 3b 所示，下采样单元 303' 包括一个最大池化层和 r 个级联的残差单元，上采样单元 304' 包括 r 个级联的残差单元和一个内插层，且最后一个最大池化层和第一个双线性插值层之间包括 $2r$ 个级联的残差

单元。并且上采样过程中与下采样过程中得到的相同分辨率的特征图之间通过一个残差单元卷积后可以相加，因此，在最后一个最大池化层之前和第一个双线性插值层之后连接有一个残差单元 305'。其中， r 为大于或等于 1 的整数。

由于在下采样和上采样的过程中，会得到分辨率相同的特征图，例如下采样过程中会得到分辨率为 28×28 、 14×14 和 7×7 的特征图，同样的，在上采样的过程中会得到分辨率为 14×14 、 28×28 和 56×56 的特征图。在下采样过程中得到的分辨率为 14×14 特征图，可以经一个残差单元 305 处理后与上采样过程中得到的分辨率为 14×14 特征图中对应像素的特征值相加，用于后续的上采样；在下采样过程得到的分辨率为 28×28 特征图，可以经一个残差单元 304 处理后与上采样过程中得到的分辨率为 28×28 特征图中对应像素的特征值相加，用于后续的上采样。这样的处理，可以捕捉待处理的图像中包含的物体的多尺度特征，同时可以增强注意力所关注的至少一个物体的特征的强度，抑制注意力不关注的其它物体的特征的强度。

参考图 4，其示出了根据本申请的图像处理方法中融合注意力和第一特征图所采用的流程 400。如图 4 所示，本实施例的融合操作可以通过以下步骤来实现：

步骤 401，至少对注意力图进行归一化处理。

本实施例中，对注意力图进行归一化处理，可以将注意力图中至少一个像素（例如，各像素）的权重值限定在 $[0,1]$ 之间。上述归一化操作，例如可以通过 sigmoid 函数来实现，sigmoid 函数是神经网络的阈值函数，其可以将变量映射到 $[0,1]$ 之间。

在一个可选示例中，该步骤 401 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的归一化单元执行。

步骤 402，融合归一化处理后的注意力图和第一特征图，获得融合图。

在对注意力图进行归一化处理，然后将归一化处理后的注意力图和第一特征图融合，得到融合图。

本实施例中，对注意力进行归一化操作一方面可以方便后续的数据处理，另一方面可以促进后续的数据处理得到更准确的结果。

在一个可选示例中，该步骤 104 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的融合单元 703 执行。

在本实施例的一些可选的实现方式中，在步骤 401 中对注意力图进行归一化处理前，可以先对注意力图进行至少一次卷积处理，然后对最后一次卷积处理后的注意力图再进行归一化处理。在一个可选示例中，上述操作可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的第二卷积单元和归一化单元执行。

在一种可选示例中，上述卷积操作可以通过卷积层来实现，可选的，可以设置该卷积层的卷积核为 1×1 ，这样可以增强对注意力图中所包含的特征的表达能力。

在本实施例的一些可选的实现方式中，在步骤 402 中融合注意力图和第一特征图时，可以将归一化处理后的注意力图中至少一个像素（例如，各像素）的权重值与第一特征图中相应像素的特征值相乘，得到融合图。在一个可选示例中，上述操作可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的融合单元 703 执行。

本实现方式中，由于注意力图与第一特征图的分辨率相同，因此注意力图中的至少一个像素可以与第一特征图中的至少一个像素一一对应。并且，步骤 401 中已对注意力图中至少一个像素的权重值进行归一化处理，可以将归一化处理后的权重值与第一特征图中对

应像素的特征值利用相乘的方法进行融合，将得到的相乘图作为融合图。

在本实施例的一些可选的实现方式中，在步骤 402 中得到上述相乘图后，还可以将得到的相乘图中至少一个像素的特征值与第一特征图中相应像素的特征值相加，将相加后得到的特征图作为融合图。在一个可选示例中，上述操作可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的融合单元 703 执行。

由于得到的相乘图中包含了待处理的图像中的特征信息，这些特征信息可称为有用信息。并且，不论是对注意力图的处理，还是对特征图的处理，都会降低待处理的图像中的特征信息的信号强度，即降低了第一特征图中至少一个像素的特征值。信号强度的衰减不利于神经网络对特征的学习，而对于上述有用信息的衰减则会直接影响神经网络的特征学习能力。

本实现方式中，将相乘图中至少一个像素的特征值与第一特征图中相应像素的特征值相加，一方面可以增加上述有用信息在整个融合图至少一个像素的特征值的比例，相当于抑制了有用信息之外的其他信息，另一方面还可以防止信号强度的衰减。

在图 3 所示的网络结构的基础上，结合图 1 所示实施例所描述的方案，可以得到本实施例的神经网络结构如图 5a 所示。图 5a 中，神经网络 500 包括第一特征提取单元 501、第一卷积单元 502、注意力提取单元 503、第二卷积单元 504、归一化单元 505、融合单元 506 以及第二特征提取单元 507。其中，第一特征提取单元 501、第一卷积单元 502 以及第二特征提取单元 507 都由多个残差单元形成，第一特征提取单元 501 包括 p 个级联的残差单元，第一卷积单元 502 包括 t 个级联的残差单元，第二特征提取单元 507 包括 p 个级联的残差单元。其中， p 、 t 为大于 1 的整数。

第一特征提取单元 501 的作用相当于图 3 中级联的多个卷积层 302，用于提取待处理的图像的特征，得到第一特征图。第一卷积单元 502 可以进一步提取第一特征图的特征。注意力提取单元 503 的作用相当于图 3 中多个下采样单元 303、多个上采样单元 304 以及多个残差单元 305，获取注意力图。第二卷积单元 504 用于在对注意力图进行归一化处理前对注意力图进行至少一次卷积操作。归一化单元 505 用于对注意力图进行归一化处理。融合单元 506 用于融合归一化处理后的注意力图和第一特征图，得到融合图。第二特征提取单元 507 用于再次提取融合图的特征。

图 5a 所示的神经网络的处理过程可参考图 5b，如图 5b 所示，以 x 表示输入的特征，即第一特征图。利用图 5a 所示的注意力提取单元 503 的感受野以及第一卷积单元 502 的感受野分别模拟人类视觉的注意力。图 5b 中的左侧分支相当于注意力提取单元 503，右侧分支相当于第一卷积单元 502。

图 5b 中的左侧分支包括两次下采样与两次上采样，经第一次下采样后，得到的特征图的分辨率为第一特征图 x 分辨率的四分之一；经第二次下采样后，得到的特征图的分辨率为第一特征图 x 分辨率的十六分之一；然后经第一次上采样处理，得到的特征图与第一次下采样处理后得到的特征图的分辨率相同；经第二次上采样处理，得到的特征图与第一特征图的分辨率相同。同时，经过上述两次下采样处理和两次上采样处理，即在遍历整个特征图后，确定了图像中注意力关注的特征的权重 $M(x)$ 。

图 5b 中的右侧分支包括对第一特征图 x 的卷积操作，得到特征 $T(x)$ 。

最后，将得到的权重 $M(x)$ 与特征 $T(x)$ 进行融合，得到融合图，上述融合图中包括融合后的特征 $(1 + M(x)) \cdot T(x)$ 。

可以理解的是，为了构造更深层次的神经网络，在本申请的可选实施例中，还可以将上述神经网络 500 作为子神经网络，多次重复执行，并可以将不同参数的子神经网络堆叠，得到如图 6 所示的深度卷积神经网络 600。深度卷积神经网络 600 可以包括多个子神经网络，图 6 中示意性的示出了三个子神经网络，分别为子神经网络 601、子神经网络 602 以及子神经网络 603。每个子神经网络的参数可以相同，也可以不同。此处所指的子神经网络的参数可以包括：注意力提取单元中的下采样和上采样的次数、第一卷积单元中残差单元的个数等等。另外，每个子神经网络可以重复多次，例如当子神经网络 601、子神经网络 602 以及子神经网络 603 至少一个不相同，深度卷积神经网络 600 可以包括 m 个子神经网络 601、 k 个子神经网络 602 以及 j 个子神经网络 603，其中， m 、 k 、 j 均为正整数。

基于本实施例提出的图像处理方法构建的神经网络，在神经网络训练过程可有效减小神经网络训练过所需调整的参数量，提高了对特征的学习效率；同时基于训练完成后的神经网络进行图像处理的过程中，无需进行调整参数的操作；通过相同次数的下采样和上采样操作，实现了全局信息的反向传递，从而促进了注意力关注的有用信息的传递。

在本实施例的一些可选的实现方式中，待处理的图像中可以包含多个物体，且上述多个物体可以为同一种类的物体，也可以为不同种类的物体。上述物体可以是至少一个类别的物体，例如可以包括飞机、自行车、汽车等各种交通工具，还可以包括鸟类、狗、狮子等各种动物。

在基于融合图，再次提取了上述待处理的图像的特征后，可以利用再次提取的特征，检测或识别上述图像中包括的物体。

进一步的，还可以利用再次提取的特征，确定上述图像中包括的物体的类别。

进一步的，还可以利用再次提取的特征，对上述图像进行分割，将包含物体的部分分割出来。

本实现方式的图像处理方法，在再次提取了待处理的图像的特征之后，可以利用再次提取的特征实现不同的应用，可以满足不同任务的图像处理需求。例如，可以对图像中包含的物体进行检测或识别，可应用到无人驾驶或导盲装置中；可以对图像中包含的物体进行分类，可以应用到军事领域的侦测装置中；可以对图像进行分割，可以应用到进一步的对物体的分析中。

本申请实施例提供的任一种图像处理方法可以由任意适当的具有数据处理能力的设备执行，包括但不限于：终端设备和服务器等。或者，本申请实施例提供的任一种图像处理方法可以由处理器执行，如处理器通过调用存储器存储的相应指令来执行本申请实施例提及的任一种图像处理方法。下文不再赘述。

本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等至少一个种可以存储程序代码的介质。

作为对上述至少一个图所示方法的实现，本申请提供了一种图像处理装置的一个实施例，该装置实施例与图 1 所示的方法实施例相对应，该装置可选可以应用于至少一个种电子设备中。

如图 7 所示，本实施例的图像处理装置 700 包括：第一特征提取单元 701、注意力提取单元 702、融合单元 703 以及第二特征提取单元 704。

其中，第一特征提取单元 701，用于提取待处理的图像的特征，获得上述图像的第一特征图。

注意力提取单元 702，用于基于上述第一特征图，生成上述图像的注意力图。

融合单元 703，用于融合上述注意力图和上述第一特征图，获得融合图。

第二特征提取单元 704，用于基于上述融合图，再次提取上述图像的特征。

在本实施例的一些可选的实现方式中，上述注意力提取单元 702 可以进一步包括图 7 中未示出的下采样模块和上采样模块。

下采样模块，用于对上述第一特征图依次进行 N 次下采样处理，其中， N 为大于或等于 1 的整数；

上采样模块，用于对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得上述图像的注意力图，其中，上述注意力图的分辨率与上述第一特征图相同。

在本实施例的一些可选的实现方式中，上述上采样模块可以用于：对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作；对卷积操作后的特征图进行第 $n+1$ 次上采样处理，获得所述图像的注意力图。其中， n 为大于 1 且小于 N 的整数。

在本实施例的一些可选的实现方式中，上述上采样模块对第 $N-n$ 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作时，用于：对上述第 $N-n$ 次下采样处理后的特征图进行卷积处理，得到卷积图；将上述卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；对相加后的特征图进行卷积操作。

在本实施例的一些可选的实现方式中，上述上采样模块可以用于：对第 N 次下采样处理后的特征图进行至少一次卷积操作；对最后一次卷积操作后的特征图进行第 1 次上采样处理，获得图像的注意力图。在本实施例的一些可选的实现方式中，上述图像处理装置 700 还可以包括图 7 中未示出的第二卷积单元，用于对上述第一特征图进行至少一次卷积操作。相应的，融合单元 703，用于融合上述注意力图和最后一次卷积操作后的第一特征图，获得上述融合图。

在本实施例的一些可选的实现方式中，上述图像处理装置 700 还可以包括图 7 中未示出的归一化单元，用于至少对上述注意力图进行归一化处理。相应的，融合单元 703，用于融合归一化处理后的注意力图和上述第一特征图，获得上述融合图。

在本实施例的一些可选的实现方式中，上述图像处理装置 700 还可以包括图 7 中未示出的第二卷积单元，用于对上述注意力图依次进行至少一次卷积处理。相应的，上述归一化单元，用于对最后一次卷积处理后的注意力图进行归一化处理。

在本实施例的一些可选的实现方式中，上述融合单元 703 还可以进一步用于：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得上述融合图。

在本实施例的一些可选的实现方式中，上述融合单元 703 还可以进一步用于：将归一化处理后的注意力图中至少一个像素的权重值与上述第一特征图中相应像素的特征值相乘，获得相乘图；将上述相乘图中至少一个像素的特征值与上述第一特征图中相应像素的特征值相加，获得上述融合图。

在本实施例的一些可选的实现方式中，上述图像处理装置 700 还可以包括图 7 中未示出的以下至少一项：检测单元、分类单元、分割单元。

其中，检测单元，用于根据再次提取的上述图像的特征，检测或识别上述图像中包括

的物体。

分类单元，用于根据再次提取的上述图像的特征，确定上述图像中包括的物体的类别。

分割单元，用于根据再次提取的上述图像的特征，对上述图像进行分割。

本申请的上述实施例提供的图像处理装置，首先提取待处理的图像的特征，得到上述图像的第一特征图，基于上述第一特征图，生成上述图像的注意力图，再将注意力图和第一特征图融合，最后再基于得到的融合图再次提取上述图像的特征，从而实现了将注意力机制引入到图像处理中，有效地提高了从图像中获取信息的效率。

附图中的流程图和框图，图示了按照本申请至少一个种实施例的系统、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，上述模块、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：一种处理器包括第一特征提取单元、注意力提取单元、融合单元及第二特征提取单元。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，第一特征提取单元还可以被描述为“提取待处理的图像的特征，获得上述图像的第一特征图的单元”。

另外，本申请实施例还提供了另一种电子设备，包括：处理器和存储器，该存储器用于存放至少一可执行指令，可执行指令使处理器执行本申请前述任一实施例的图像处理方法对应的操作。

本申请实施例还提供了一种电子设备，例如可以是移动终端、个人计算机（PC）、平板电脑、服务器等。下面参考图8，其示出了适于用来实现本申请实施例的终端设备或服务器的电子设备800的结构示意图：如图8所示，计算机系统800包括一个或多个处理器、通信部等，上述一个或多个处理器例如：一个或多个中央处理单元（CPU）801，和/或一个或多个图像处理器（GPU）813等，处理器可以根据存储在只读存储器（ROM）802中的可执行指令或者从存储部分808加载到随机访问存储器（RAM）803中的可执行指令而执行至少一个种适当的动作和处理。通信部812可包括但不限于网卡，上述网卡可包括但不限于IB（Infiniband）网卡。

处理器可与ROM 802和/或RAM 803通信以执行可执行指令，通过总线804与通信部812相连、并经通信部812与其他目标设备通信，从而完成本申请实施例提供的任一项方法对应的操作，例如，提取待处理的图像的特征，获得上述图像的第一特征图；基于上述第一特征图，生成上述图像的注意力图；融合上述注意力图和上述第一特征图，获得融合图；基于上述融合图，再次提取上述图像的特征。

此外，在RAM 803中，还可存储有装置操作所需的至少一个种程序和数据。CPU 801、ROM 802以及RAM 803通过总线804彼此相连。在有RAM 803的情况下，ROM 802为可选模块。RAM 803存储可执行指令，或在运行时向ROM 802中写入可执行指令，可执

行指令使 CPU 801 执行上述通信方法对应的操作。输入/输出 (I/O) 接口 805 也连接至总线 804。通信部 812 可以集成设置,也可以设置为具有多个子模块(例如多个 IB 网卡),并在总线 804 链接上。

以下部件连接至 I/O 接口 805: 包括键盘、鼠标等的输入部分 806; 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 807; 包括硬盘等的存储部分 808; 以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 809。通信部分 809 经由诸如因特网的网络执行通信处理。驱动器 810 也根据需要连接至 I/O 接口 805。可拆卸介质 811, 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 810 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 808。

需要说明的, 如图 8 所示的架构仅为一种可选实现方式, 在可选实践过程中, 可根据实际需要对上述图 8 的部件数量和类型进行选择、删减、增加或替换; 在不同功能部件设置上, 也可采用分离设置或集成设置等实现方式, 例如 GPU 813 和 CPU 801 可分离设置或者可将 GPU 813 集成在 CPU 801 上, 通信部可分离设置, 也可集成设置在 CPU 801 或 GPU 813 上, 等等。这些可替换的实施方式均落入本申请公开的保护范围。

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, 其包括有形地包含在机器可读介质上的计算机程序, 计算机程序包含用于执行流程图所示的方法的程序代码, 程序代码可包括对应执行本申请实施例提供的方法步骤对应的指令, 例如, 提取待处理的图像的特征, 获得上述图像的第一特征图; 基于上述第一特征图, 生成上述图像的注意力图; 融合上述注意力图和上述第一特征图, 获得融合图; 基于上述融合图, 再次提取上述图像的特征。在这样的实施例中, 该计算机程序可以通过通信部分 809 从网络上被下载和安装, 和/或从可拆卸介质 811 被安装。在该计算机程序被 CPU 801 执行时, 执行本申请的方法中限定的上述功能。

另外, 本申请实施例还提供了一种计算机程序, 包括计算机可读代码, 当计算机可读代码在设备上运行时, 设该设备中的处理器执行用于实现本申请前述任一实施例的图像处理方法中各步骤的指令。

另外, 本申请实施例还提供了一种计算机可读存储介质, 用于存储计算机可读的指令, 该指令被执行时实现本申请前述任一实施例的图像处理方法中各步骤的操作。

可能以许多方式来实现本申请实施例的方法和装置、设备。例如, 可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本申请实施例的方法和装置、设备。用于方法的步骤的上述顺序仅是为了进行说明, 本申请实施例的方法的步骤不限于以上可选描述的顺序, 除非以其它方式特别说明。此外, 在一些实施例中, 还可将本申请实施为记录在记录介质中的程序, 这些程序包括用于实现根据本申请实施例的方法的机器可读指令。因而, 本申请实施例还覆盖存储用于执行根据本申请实施例的方法的程序的记录介质。

本申请实施例的描述是为了示例和描述起见而给出的, 而并不是无遗漏的或者将本申请实施例限于所公开的形式。很多修改和变化对于本领域的普通技术人员而言是显然的。选择和描述实施例是为了更好说明本申请实施例的原理和实际应用, 并且使本领域的普通技术人员能够理解本申请实施例从而设计适于特定用途的带有至少一个种修改的至少一个种实施例。

权利要求

1、一种图像处理方法，其特征在于，包括：

提取待处理的图像的特征，获得所述图像的第一特征图；

基于所述第一特征图，生成所述图像的注意力图；

融合所述注意力图和所述第一特征图，获得融合图；

基于所述融合图，再次提取所述图像的特征。

2、根据权利要求 1 所述的方法，其特征在于，所述基于所述第一特征图，生成所述图像的注意力图，包括：

对所述第一特征图依次进行 N 次下采样处理，其中，N 为大于或等于 1 的整数；

对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得所述图像的注意力图，其中，所述注意力图的分辨率与所述第一特征图的分辨率相同。

3、根据权利要求 2 所述的方法，其特征在于，所述对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，包括：

对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，其中，n 为大于 1 且小于 N 的整数；

对卷积操作后的特征图进行第 n+1 次上采样处理。

4、根据权利要求 3 所述的方法，其特征在于，所述对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，包括：

对所述第 N-n 次下采样处理后的特征图进行卷积处理，得到卷积图；

将所述卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；

对相加后的特征图进行卷积操作。

5、根据权利要求 3 或 4 所述的方法，其特征在于，所述对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，还包括：

对第 N 次下采样处理后的特征图进行至少一次卷积操作；

对最后一次卷积操作后的特征图进行第 1 次上采样处理。

6、根据权利要求 1-5 任一项所述的方法，其特征在于，所述融合所述注意力图和所述第一特征图，获得融合图，包括：

对所述第一特征图进行至少一次卷积操作；

融合所述注意力图和最后一次卷积操作后的第一特征图，获得所述融合图。

7、根据权利要求 1-6 任一项所述的方法，其特征在于，所述融合所述注意力图和所述第一特征图，获得融合图，包括：

至少对所述注意力图进行归一化处理；

融合归一化处理后的注意力图和所述第一特征图，获得所述融合图。

8、根据权利要求 7 所述的方法，其特征在于，所述至少对所述注意力图进行归一化处理，包括：

对所述注意力图依次进行至少一次卷积处理；

对最后一次卷积处理后的注意力图进行归一化处理。

9、根据权利要求 7 或 8 所述的方法，其特征在于，所述融合所述注意力图和所述第

一特征图，获得融合图，包括：

将归一化处理后的注意力图中至少一个像素的权重值与所述第一特征图中相应像素的特征值相乘，获得所述融合图。

10、根据权利要求 7 或 8 所述的方法，其特征在于，所述融合所述注意力图和所述第一特征图，获得融合图，包括：

将归一化处理后的注意力图中至少一个像素的权重值与所述第一特征图中相应像素的特征值相乘，获得相乘图；

将所述相乘图中至少一个像素的特征值与所述第一特征图中相应像素的特征值相加，获得所述融合图。

11、根据权利要求 1-10 任一项所述的方法，其特征在于，在所述基于所述融合图，再次提取所述图像的特征之后，所述方法还包括以下至少之一：

根据再次提取的所述图像的特征，检测或识别所述图像中包括的物体；

根据再次提取的所述图像的特征，确定所述图像中包括的物体的类别；

根据再次提取的所述图像的特征，对所述图像进行分割。

12、一种图像处理装置，其特征在于，包括：

第一特征提取单元，用于提取待处理的图像的特征，获得所述图像的第一特征图；

注意力提取单元，用于基于所述第一特征图，生成所述图像的注意力图；

融合单元，用于融合所述注意力图和所述第一特征图，获得融合图；

第二特征提取单元，用于基于所述融合图，再次提取所述图像的特征。

13、根据权利要求 11 所述的装置，其特征在于，所述注意力提取单元包括：

下采样模块，用于对所述第一特征图依次进行 N 次下采样处理，其中，N 为大于或等于 1 的整数；

上采样模块，用于对第 N 次下采样处理后的特征图依次进行 N 次上采样处理，获得所述图像的注意力图，其中，所述注意力图的分辨率与所述第一特征图的分辨率相同。

14、根据权利要求 13 所述的装置，其特征在于，所述上采样模块用于：

对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作，其中，n 为大于 1 且小于 N 的整数；

对卷积操作后的特征图进行第 n+1 次上采样处理，获得所述图像的注意力图。

15、根据权利要求 14 所述的装置，其特征在于，所述上采样模块对第 N-n 次下采样处理后的特征图和第 n 次上采样处理后的特征图进行卷积操作时，用于：

对所述第 N-n 次下采样处理后的特征图进行卷积处理，得到卷积图；

将所述卷积图中至少一个像素的特征值与第 n 次上采样处理后的特征图中相应像素的特征值相加；

对相加后的特征图进行卷积操作。

16、根据权利要求 14 或 15 所述的装置，其特征在于，所述上采样模块用于：

对第 N 次下采样处理后的特征图进行至少一次卷积操作；

对最后一次卷积操作后的特征图进行第 1 次上采样处理，获得所述图像的注意力图。

17、根据权利要求 12-16 任一项所述的装置，其特征在于，还包括：

第二卷积单元，用于对所述第一特征图进行至少一次卷积操作；

所述融合单元，用于融合所述注意力图和最后一次卷积操作后的第一特征图，获得所

述融合图。

18、根据权利要求 12-17 任一项所述的装置，其特征在于，还包括：

归一化单元，用于至少对所述注意力图进行归一化处理；

所述融合单元，用于融合归一化处理后的注意力图和所述第一特征图，获得所述融合图。

19、根据权利要求 18 所述的装置，其特征在于，还包括：

第二卷积单元，用于对所述注意力图依次进行至少一次卷积处理；以及

所述归一化单元，用于对最后一次卷积处理后的注意力图进行归一化处理。

20、根据权利要求 18 或 19 所述的装置，其特征在于，所述融合单元用于：

将归一化处理后的注意力图中至少一个像素的权重值与所述第一特征图中相应像素的特征值相乘，获得所述融合图。

21、根据权利要求 18 或 19 所述的装置，其特征在于，所述融合单元用于：

将归一化处理后的注意力图中至少一个像素的权重值与所述第一特征图中相应像素的特征值相乘，获得相乘图；

将所述相乘图中至少一个像素的特征值与所述第一特征图中相应像素的特征值相加，获得所述融合图。

22、根据权利要求 14-21 任一项所述的装置，其特征在于，还包括以下至少之一：

检测单元，用于根据再次提取的所述图像的特征，检测或识别所述图像中包括的物体；

分类单元，用于根据再次提取的所述图像的特征，确定所述图像中包括的物体的类别；

分割单元，用于根据再次提取的所述图像的特征，对所述图像进行分割。

23、一种计算机可读存储介质，其上存储有计算机指令，其特征在于，所述指令被执行时实现权利要求 1-11 中任一项所述图像处理方法中各步骤的操作。

24、一种电子设备，其特征在于，包括：处理器和存储器；

所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行权利要求 1-11 中任一项所述图像处理方法对应的操作。

25、一种计算机程序，包括计算机可读代码，其特征在于，当所述计算机可读代码在设备上运行时，所述设备中的处理器执行用于实现权利要求 1-11 中任一项所述的图像处理方法中各步骤的指令。

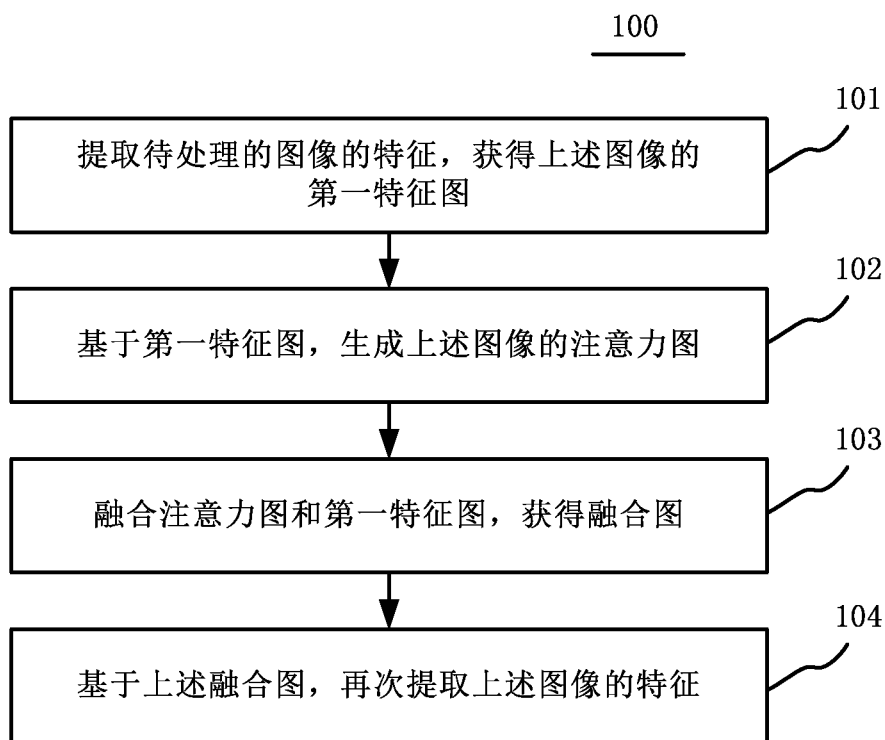


图 1

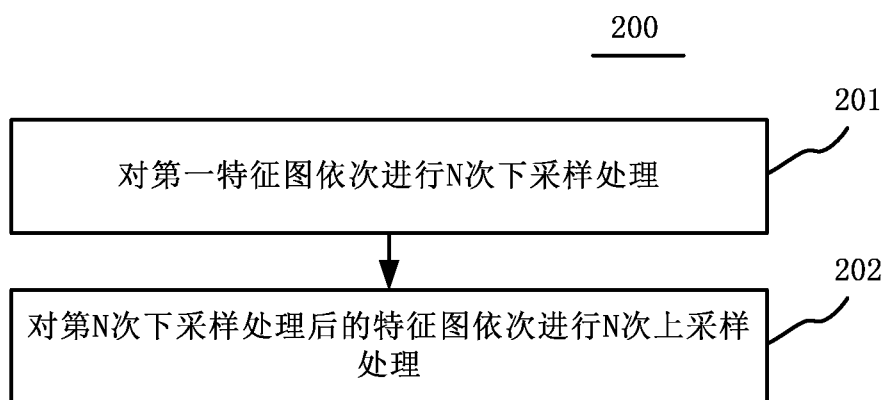


图 2

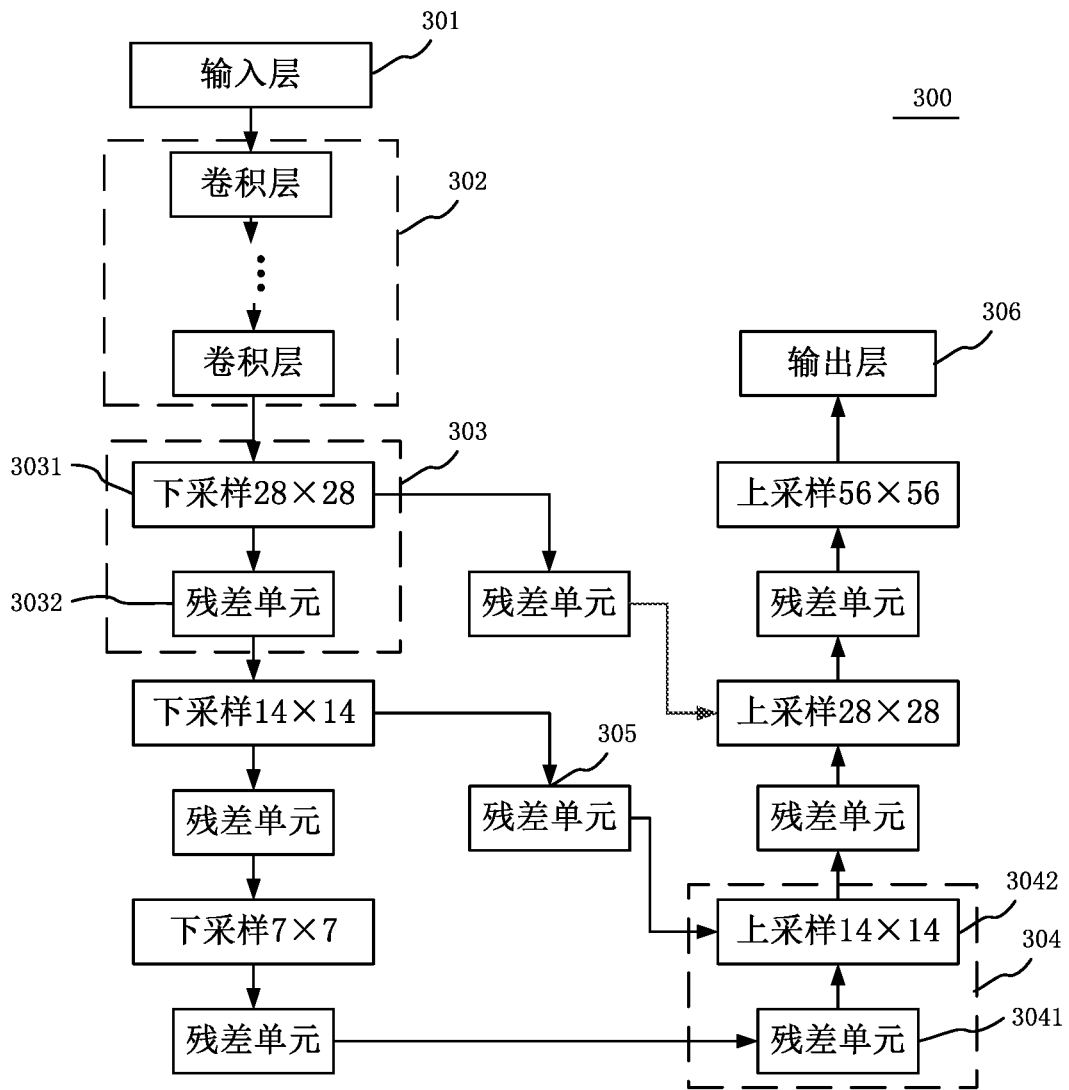


图 3a

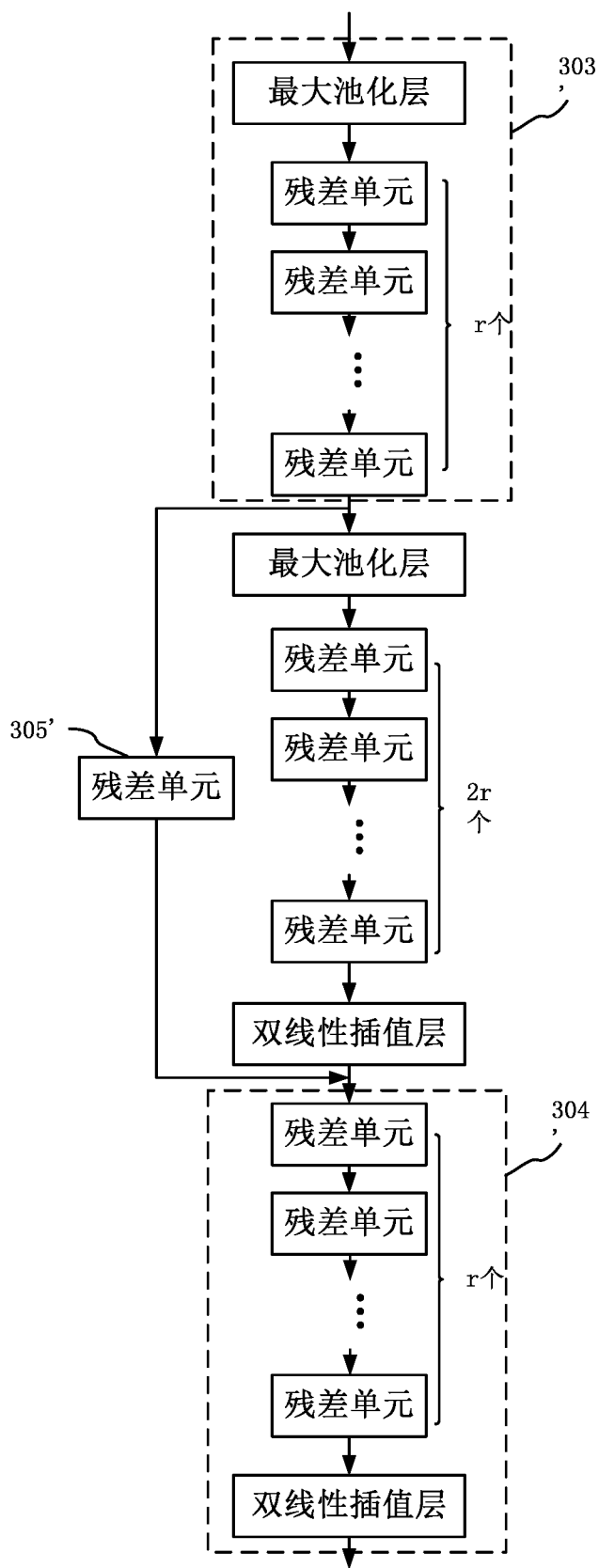


图 3b

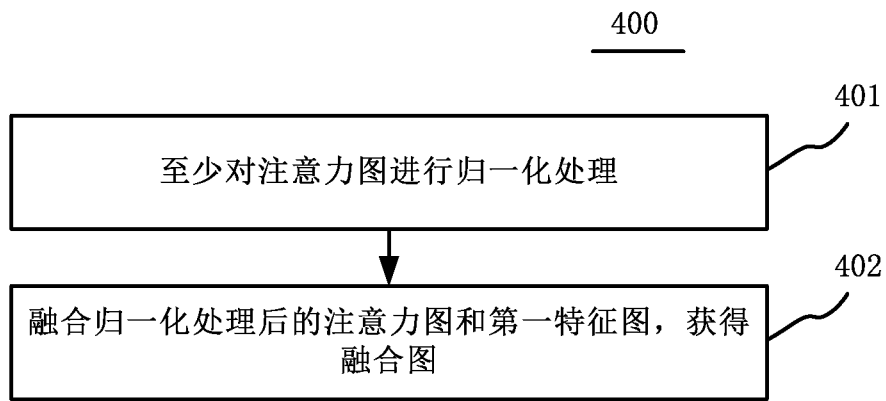


图 4

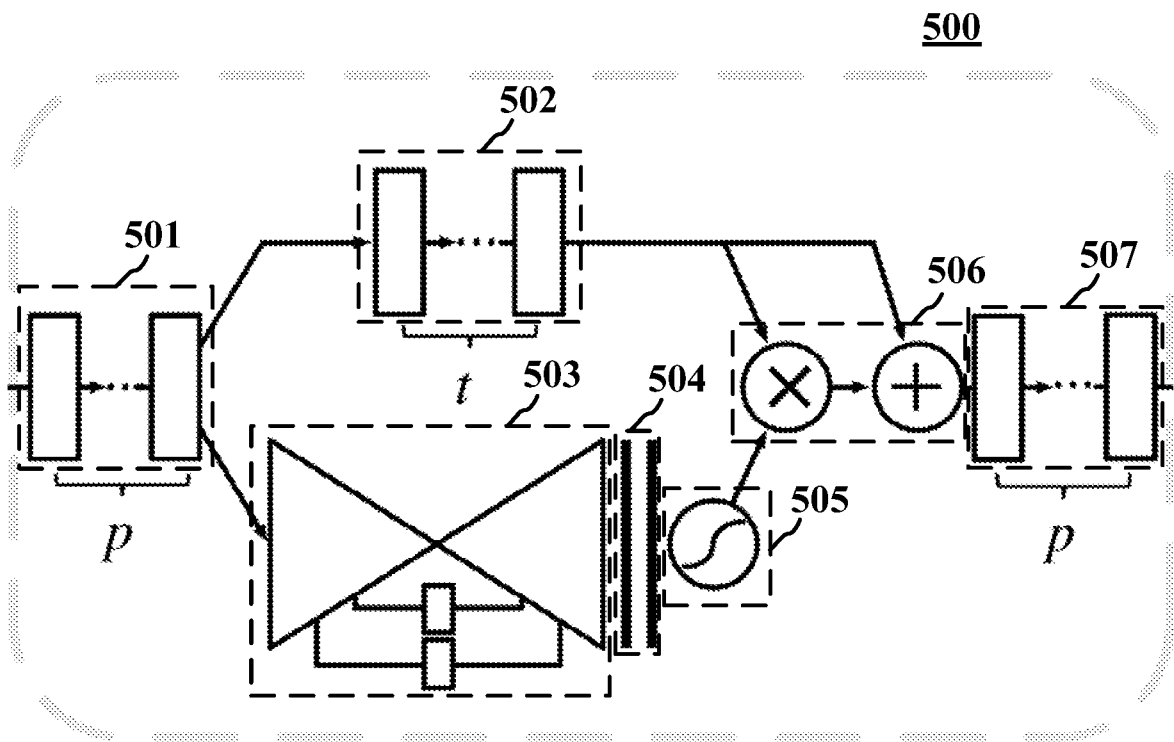


图 5a

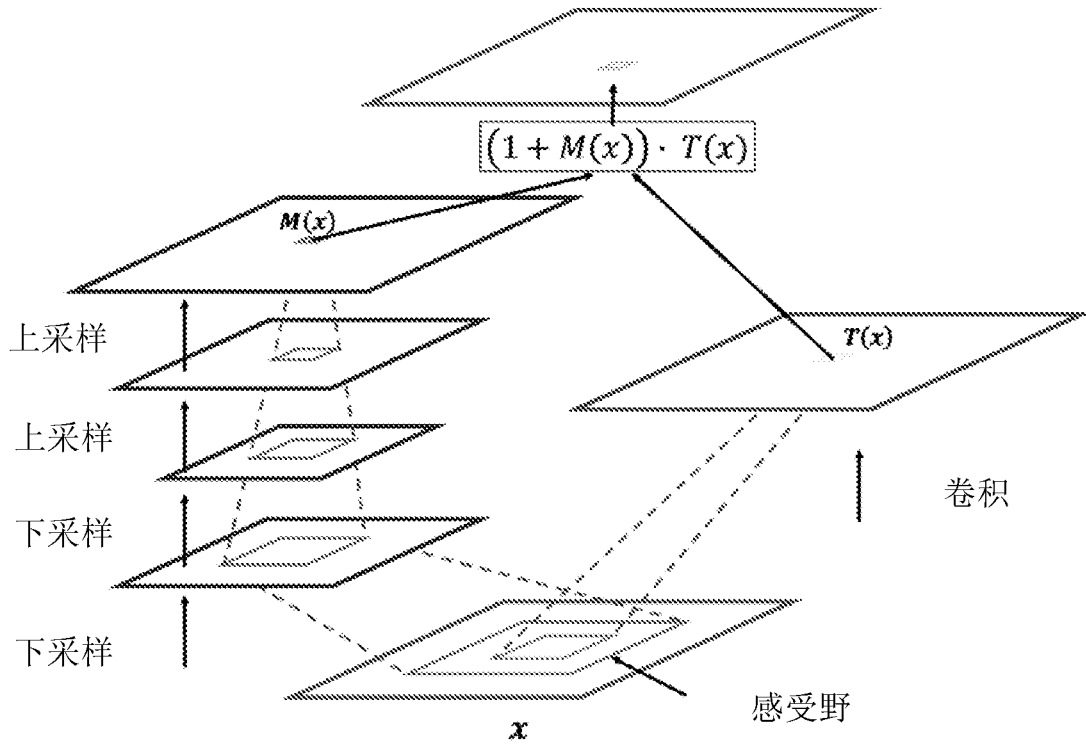


图 5b

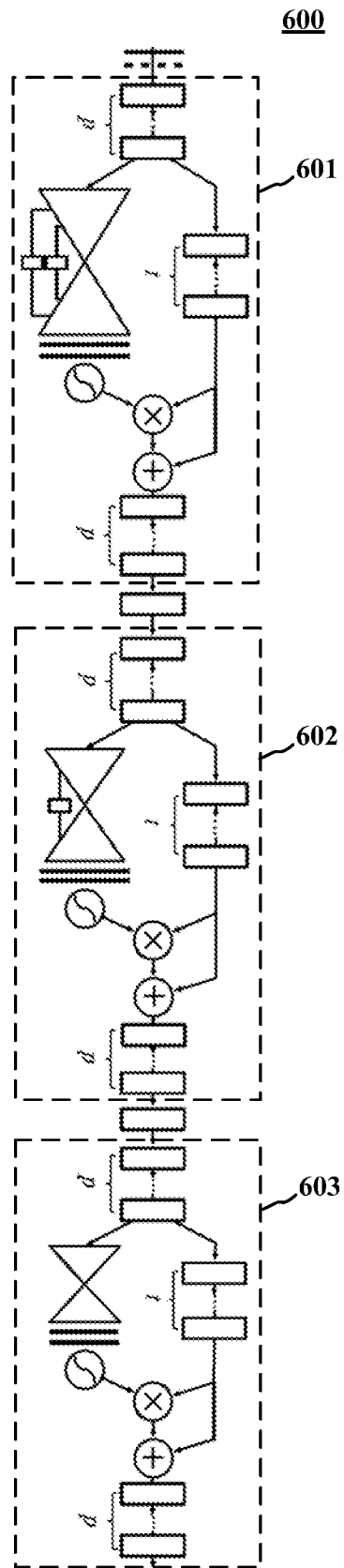


图 6

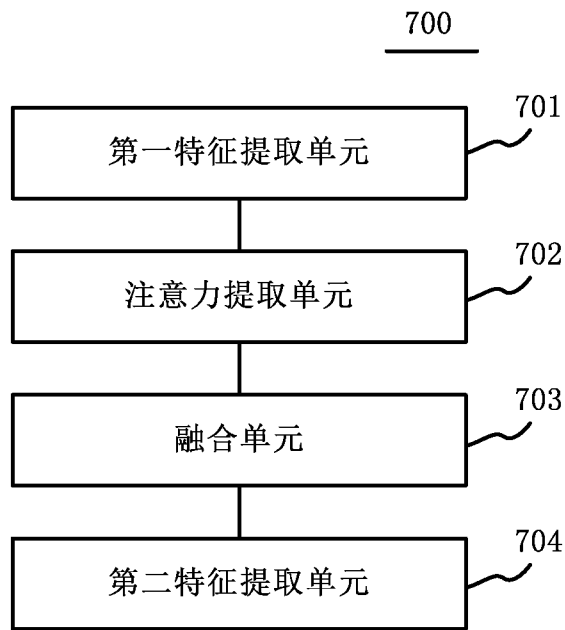


图 7

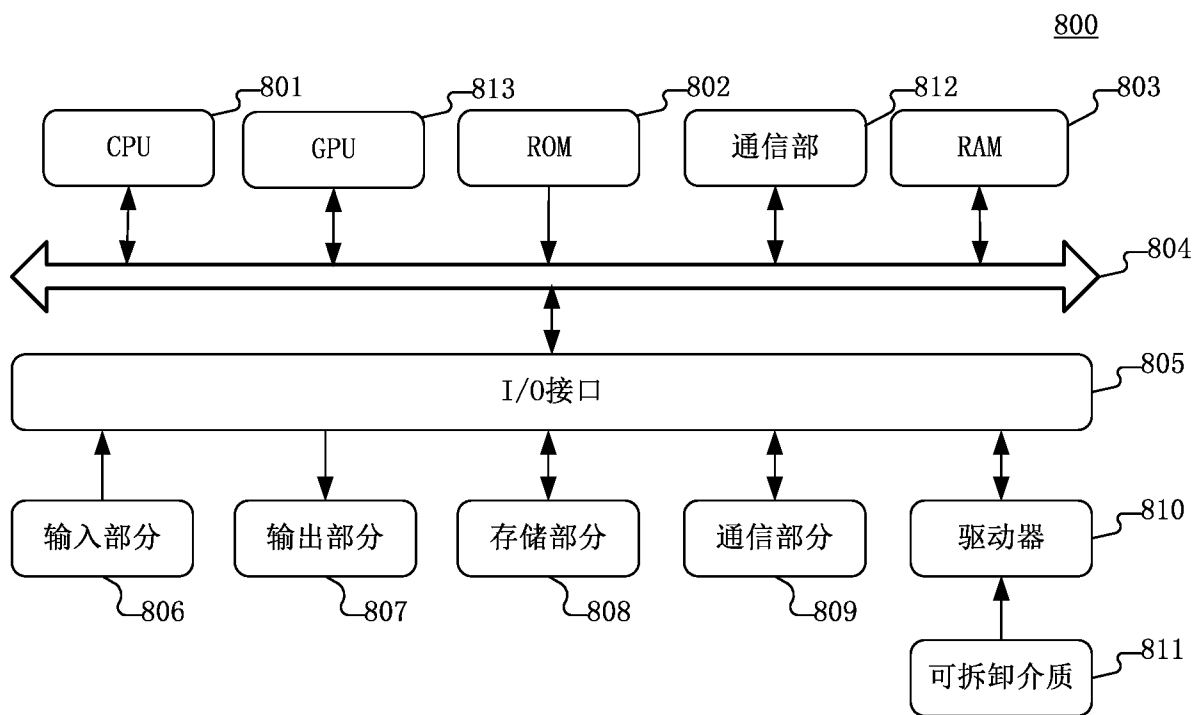


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/078810

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/32 (2006.01) i; G06K 9/62 (2006.01) i; G06T 7/00 (2017.01) 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; G06T

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: 图像, 特征, 提取, 注意力, 融合, 卷积, 相乘, 采样, image, feature, characteristic, obtain., attention, fusion, convolution, multiplication, sample

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106934397 A (BEIJING SENSETIME SCIENCE AND TECHNOLOGY DEVELOPMENT CO., LTD.) 07 July 2017 (07.07.2017), claims 1-10, and description, paragraphs [0126]-[0134]	1-25
PX	CN 107729901 A (ALIBABA GROUP HOLDING LTD.) 23 February 2018 (23.02.2018), the abstract, and claims 1, 2 and 7-10	1, 12, 23-25
X	CN 103996185 A (CHONGQING UNIVERSITY) 20 August 2014 (20.08.2014), the abstract, claims 1 and 4, description, paragraphs [0035]-[0074], and figure 1	1, 7-12, 17-25
A	CN 103679718 A (HOHAI UNIVERSITY) 26 March 2014 (26.03.2014), entire document	1-25
A	CN 101866422 A (SHANGHAI UNIVERSITY) 20 October 2010 (20.10.2010), entire document	1-25
A	CN 101980248 A (XIDIAN UNIVERSITY) 23 February 2011 (23.02.2011), entire document	1-25

 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

11 April 2018

Date of mailing of the international search report

28 April 2018

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer

YU, Lina

Telephone No. (86-10) 53961756

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN201 8/0788 10

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 105228033 A (LENOVO (BEIJING) CO., LTD.) 06 January 2016 (06.01 .2016), entire document	1-25
A	CN 101697593 A (WUHAN UNIVERSITY) 21 April 2010 (21.04.2010), entire document	1-25
A	CN 106157319 A (HARBIN INSTITUTE OF TECHNOLOGY) 23 November 2016 (23.11 .2016), entire document	1-25
A	US 2014153651 A I (THOMSON LICENSING) 05 June 2014 (05.06.2014), entire document	1-25

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/078810

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 106934397 A	07 July 2017	None	
CN 107729901 A	23 February 2018	None	
CN 103996185 A	20 August 2014	None	
CN 103679718 A	26 March 2014	None	
CN 101866422 A	20 October 2010	None	
CN 101980248 A	23 February 2011	CN 101980248 B	05 December 2012
CN 105228033 A	06 January 2016	None	
CN 101697593 A	21 April 2010	CN 101697593 B	10 October 2012
CN 106157319 A	23 November 2016	None	
US 2014153651 A I	05 June 2014	W O 2013010751 A I	24 January 2013
		US 9641795 B2	02 May 2017
		JP 5974089 B2	23 August 2016
		EP 2735146 A I	28 May 2014
		EP 2549754 A I	23 January 2013
		JP 2014521272 A	25 August 2014
		W O 2013010751 A 8	15 August 2013
		CN 103688538 A	26 March 2014
		CN 103688538 B	31 May 2017
		KR 20140042860 A	07 April 2014

国际检索报告

国际申请号

PCT/CN2018/078810

<p>A. 主题的分类</p> <p>G06K 9/32 (2006. 01) i; G06K 9/62 (2006. 01) i; G06T 7/00 (2017. 01) i</p> <p>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</p>																													
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>G06K; G06T</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNPAT, CNKI, WPI, EPODOC, IEEE : 图像, 特征, 提取, 注意力, 融合, 卷积, 相乘, 采样, image, feature, characteristic, obtain, attention, fusion, convolution, multiplication, sample</p>																													
<p>C 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 106934397 A (北京市商汤科技开发有限公司) 2017 年 7 月 7 日 (2017 - 07 - 07) 权利要求 1-10, 说明书第 126-134 段</td> <td>1-25</td> </tr> <tr> <td>PX</td> <td>CN 107729901 A (阿里巴巴集团控股有限公司) 2018 年 2 月 23 日 (2018 - 02 - 23) 说明书摘要, 权利要求 1、2、7-10</td> <td>1、12、23-25</td> </tr> <tr> <td>X</td> <td>CN 103996185 A (重庆大学) 2014 年 8 月 20 日 (2014 - 08 - 20) 说明书摘要, 权利要求 1、4, 说明书第 35-74 段, 附图 1</td> <td>1、7-12、17-25</td> </tr> <tr> <td>A</td> <td>CN 103679718 A (河海大学) 2014 年 3 月 26 日 (2014 - 03 - 26) 全文</td> <td>1-25</td> </tr> <tr> <td>A</td> <td>CN 101866422 A (上海大学) 2010 年 10 月 20 日 (2010 - 10 - 20) 全文</td> <td>1-25</td> </tr> <tr> <td>A</td> <td>CN 101980248 A (西安电子科技大学) 2011 年 2 月 23 日 (2011 - 02 - 23) 全文</td> <td>1-25</td> </tr> <tr> <td>A</td> <td>CN 105228033 A (联想北京有限公司) 2016 年 1 月 6 日 (2016 - 01 - 06) 全文</td> <td>1-25</td> </tr> <tr> <td>A</td> <td>CN 101697593 A (武汉大学) 2010 年 4 月 21 日 (2010 - 04 - 21) 全文</td> <td>1-25</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 106934397 A (北京市商汤科技开发有限公司) 2017 年 7 月 7 日 (2017 - 07 - 07) 权利要求 1-10, 说明书第 126-134 段	1-25	PX	CN 107729901 A (阿里巴巴集团控股有限公司) 2018 年 2 月 23 日 (2018 - 02 - 23) 说明书摘要, 权利要求 1、2、7-10	1、12、23-25	X	CN 103996185 A (重庆大学) 2014 年 8 月 20 日 (2014 - 08 - 20) 说明书摘要, 权利要求 1、4, 说明书第 35-74 段, 附图 1	1、7-12、17-25	A	CN 103679718 A (河海大学) 2014 年 3 月 26 日 (2014 - 03 - 26) 全文	1-25	A	CN 101866422 A (上海大学) 2010 年 10 月 20 日 (2010 - 10 - 20) 全文	1-25	A	CN 101980248 A (西安电子科技大学) 2011 年 2 月 23 日 (2011 - 02 - 23) 全文	1-25	A	CN 105228033 A (联想北京有限公司) 2016 年 1 月 6 日 (2016 - 01 - 06) 全文	1-25	A	CN 101697593 A (武汉大学) 2010 年 4 月 21 日 (2010 - 04 - 21) 全文	1-25
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																											
PX	CN 106934397 A (北京市商汤科技开发有限公司) 2017 年 7 月 7 日 (2017 - 07 - 07) 权利要求 1-10, 说明书第 126-134 段	1-25																											
PX	CN 107729901 A (阿里巴巴集团控股有限公司) 2018 年 2 月 23 日 (2018 - 02 - 23) 说明书摘要, 权利要求 1、2、7-10	1、12、23-25																											
X	CN 103996185 A (重庆大学) 2014 年 8 月 20 日 (2014 - 08 - 20) 说明书摘要, 权利要求 1、4, 说明书第 35-74 段, 附图 1	1、7-12、17-25																											
A	CN 103679718 A (河海大学) 2014 年 3 月 26 日 (2014 - 03 - 26) 全文	1-25																											
A	CN 101866422 A (上海大学) 2010 年 10 月 20 日 (2010 - 10 - 20) 全文	1-25																											
A	CN 101980248 A (西安电子科技大学) 2011 年 2 月 23 日 (2011 - 02 - 23) 全文	1-25																											
A	CN 105228033 A (联想北京有限公司) 2016 年 1 月 6 日 (2016 - 01 - 06) 全文	1-25																											
A	CN 101697593 A (武汉大学) 2010 年 4 月 21 日 (2010 - 04 - 21) 全文	1-25																											
<p><input checked="" type="checkbox"/> 其余文件在 c 栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																													
<p>* 引用文件的具体类型:</p> <p>"A" 认为不特别相关的表示了现有技术一般状态的文件</p> <p>"E" 在国际申请日的当天或之后公布的在先申请或专利</p> <p>"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)</p> <p>"O" 涉及口头公开、使用、展览或其他方式公开的文件</p> <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</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&" 同族专利的文件</p>																													
<p>国际检索实际完成的日期</p> <p>2018 年 4 月 11 日</p>		<p>国际检索报告邮寄日期</p> <p>2018 年 4 月 28 日</p>																											
<p>ISA/CN 的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局 (ISA/CN)</p> <p>中国北京市海淀区蓟门桥西土城路 6 号 100088</p> <p>传真号 (86-10) 62019451</p>		<p>授权官员</p> <p>于利娜</p> <p>电话号码 (86-10) 53961756</p>																											

C 相关文件		
类型 ^k	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 106157319 A (哈尔滨工业大学) 2016 年 11 月 23 日 (2016 - 11 - 23) 全文	1—25
A	US 2014153651 A1 (THOMSON LICENSING) 2014 年 6 月 5 日 (2014 - 06 - 05) 全文	1—25

国际检索报告

关于同族专利的信息

国际申请号

PCT/CN2018/0788 10

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 106934397 A	2017 年 7 月 7 日	无	
CN 107729901 A	2018 年 2 月 23 日	无	
CN 103996185 A	2014 年 8 月 20 日	无	
CN 103679718 A	2014 年 3 月 26 日	无	
CN 101866422 A	2010 年 10 月 20 日	无	
CN 101980248 A	2011 年 2 月 23 日	CN 101980248 B	2012 年 12 月 5 日
CN 105228033 A	2016 年 1 月 6 日	无	
CN 101697593 A	2010 年 4 月 21 日	CN 101697593 B	2012 年 10 月 10 日
CN 106157319 A	2016 年 11 月 23 日	无	
US 2014153651 A1	2014 年 6 月 5 日	wo 2013010751 A1	2013 年 1 月 24 日
		us 9641795 B2	2017 年 5 月 2 日
		JP 5974089 B2	2016 年 8 月 23 日
		EP 2735146 A1	2014 年 5 月 28 日
		EP 2549754 A1	2013 年 1 月 23 日
		JP 2014521272 A	2014 年 8 月 25 日
		Wo 2013010751 A8	2013 年 8 月 15 日
		CN 103688538 A	2014 年 3 月 26 日
		CN 103688538 B	2017 年 5 月 31 日
		KR 20140042860 A	2014 年 4 月 7 日