

分割模型的训练方法及装置、图像识别方法及装置

技术领域

[01]本说明书一个或多个实施例涉及人工智能技术，尤其涉及分割模型的训练方法及装置、图像识别方法和装置。

背景技术

[02]图像识别，是指利用计算机对图像进行处理、分析和理解，以识别各种不同模式的目标对象的技术。图像识别技术一般分为人脸识别与商品识别等，人脸识别主要运用在安全检查、身份核验与移动支付中；商品识别主要运用在商品流通过程中，特别是无人货架、智能零售柜等无人零售领域。

[03]在图像识别技术中，需要从图像包括的各种对象中识别出目标对象。比如，在人脸识别方案中，交互的屏幕采用的是实时展示相机原始采集的数据，但是这个过程在排队场景中，会导致不想刷脸的人员的脸也出现在屏幕上，这无形中会对排队用户产生一个隐私不友好的感受，部分用户甚至会感觉隐私被侵犯，因此需要通过图像识别分割出目标人脸。再如，在商品识别中，拍摄的图像中可能存在用户手握的之前已付款商品及当前待付款商品，因此，需要通过图像识别分割出当前待付款的目标商品。

[04]然而相关技术的图像识别方法不能准确地分割出目标对象。

发明内容

[05]本说明书一个或多个实施例描述了分割模型的训练方法及装置、图像识别方法和装置，能够更加准确地得到图像中的目标对象的分割信息。

[06]根据第一方面，提供了一种分割模型的训练方法，该分割模型包括：第一网络模型、第二网络模型以及第三网络模型，其中包括：获取样本图像对；其中，所述样本图像对包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像；将所述深度图像输入第一网络模型，得到该第一网络模型输出的第一深度特征提取结果；将所述深度图像与所述 RGB 图像的组图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；将所述目标对象的边缘特征以及所述第一深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果；根据所述样本图像对的标签以及该目标对象的分割结果，对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整。

[07]其中，所述样本图像对的标签包括：第一标签以及第二标签；其中，第一标签为预

先由人工形成的对所述 RGB 图像或所述深度图像的分割结果；第二标签为对所述第一标签进行高斯模糊处理后得到的；所述对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整，包括：根据第一标签与目标对象的分割结果之间的差异，对第二网络模型以及第三网络模型进行参数调整；根据第二标签与第一深度特征提取结果之间的差异，对第一网络模型进行参数调整。

[08]其中，在所述将所述深度图像输入第一网络模型中之后，进一步包括：得到第一网络模型包括的中间层神经网络所提取出的目标对象的轮廓信息，将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给所述第二网络模型；在所述将所述深度图像与所述 RGB 图像的组合图像输入第二网络模型之后，并在得到该第二网络模型输出的目标对象的边缘特征之前，进一步包括：由所述第二网络模型包括的前端各层神经网络对所述组合图像进行特征提取，得到初级边缘特征；由第二网络模型包括的后端各层神经网络对该初级边缘特征以及第二深度特征提取结果进行处理，以便得到并输出所述目标对象的边缘特征。

[09]其中，调整所述第一网络模型及所述第二网络模型的卷积核以及卷积步长，使得所述初级边缘特征与所述第二深度特征提取结果对应的图像尺寸相同。

[10]其中，在将所述深度图像与所述 RGB 图像进行组合之前，进一步包括：对所述 RGB 图像的像素值和所述深度图像的像素值进行归一化，并且将深度图像中值为空的像素的像素值归一化为 0。

[11]根据第二方面，提供了图像识别方法，其中包括：获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像；将待处理的深度图像输入第一网络模型，得到该第一网络模型输出的深度特征提取结果；将待处理的深度图像与待处理的 RGB 图像的组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；将该目标对象的边缘特征以及该深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果。

[12]根据第三方面，提供了分割模型的训练装置，其中包括：样本图像获取模块，配置为获取样本图像对；其中，样本图像对中包括对同一视觉范围拍摄后得到的所述 RGB 图像以及所述深度图像；第一网络模型训练模块，配置为将所述深度图像输入第一网络模型，得到该第一网络模型输出的第一深度特征提取结果；第二网络模型训练模块，配置为将所述深度图像与所述 RGB 图像的组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；第三网络模型训练模块，配置为将所述目标对象的边缘特征以及所述第一深度特征提取结果输入第三网络模型，得到该第三网络模型输出的

目标对象的分割结果；调整模块，配置为根据所述样本图像对的标签以及该目标对象的分割结果，对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整。

[13]其中，第一网络模型训练模块，进一步被配置为得到第一网络模型包括的中间层神经网络所提取出的目标对象的轮廓信息，将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给所述第二网络模型；第二网络模型训练模块，进一步被配置为控制所述第二网络模型包括的前端各层神经网络对所述组合图像进行特征提取以得到初级边缘特征，以及控制第二网络模型包括的后端各层神经网络对该初级边缘特征以及第二深度特征提取结果进行处理，以由第二网络模型输出所述目标对象的边缘特征。

[14]根据第四方面，提供了图像识别装置，其中，包括：图像输入模块，配置为获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像；第一网络模型，配置为接收所述待处理的深度图像，得到深度特征提取结果；第二网络模型，配置为接收所述待处理的深度图像与所述待处理的 RGB 图像的组合图像，得到目标对象的边缘特征；第三网络模型，配置为接收该目标对象的边缘特征以及该深度特征提取结果，得到目标对象的分割结果。

[15]根据第五方面，提供了一种计算设备，包括存储器和处理器，所述存储器中存储有可执行代码，所述处理器执行所述可执行代码时，实现本说明书任一实施例所述的方法。

[16]本说明书实施例提供的分割模型的训练方法及装置、图像识别方法及装置，不仅是在训练过程的初始阶段利用了深度图像（即，将深度图像与 RGB 图像进行组合，利用组合图像得到目标对象的边缘特征），而且还在训练过程的后续阶段利用了对深度图像的深度特征提取结果，也就是说，是同时利用组合图像以及深度特征提取结果来训练分割模型，可见，在训练过程中的不同阶段分别利用了深度图像提供的深度信息，从而使训练出的分割模型能够更加准确地得到图像中的目标对象的分割信息。

附图说明

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

[18]图 1 是本说明书一个实施例所应用的系统架构的示意图。

[19]图 2 是本说明书一个实施例中分割模型的训练方法的流程图。

[20]图 3A 是本说明书一个实施例中对第一种分割模型进行训练的训练方法的示意图。

[21]图 3B 是本说明书一个实施例中对第一种分割模型的应用结构示意图。

[22]图 4A 本说明书一个实施例中对第二种分割模型进行训练的训练方法的示意图。

[23]图 4B 是本说明书一个实施例中对第二种分割模型的应用结构的示意图。

[24]图 5 是本说明书一个实施例中采用方式 2 进行分割模型训练的方法的流程图。

[25]图 6 是本说明书一个实施例中的图像识别方法的流程图。

[26]图 7 是本说明书一个实施例中分割模型的训练装置的结构示意图。

[27]图 8 是本说明书一个实施例中图像识别装置的结构示意图。

具体实施方式

[28]如前所述，需要从图像中准确地分割出目标对象。比如在人脸识别过程中，采集的图像中包括 2 个人像的信息，需要分割出其中最靠前且居中的目标人像的信息，从而进行人脸支付等业务流程。再如，在商品识别过程中，采集的图像中包括 3 个商品的信息，需要分割出其中最靠前且居中的目标商品的信息，从而进行目标商品的付款处理等业务流程。

[29]为了方便对本说明书提供的方法进行理解，首先对本说明书所涉及和适用的系统架构进行描述。如图 1 中所示，该系统架构中主要包括：RGB 图像拍摄装置、深度图像拍摄装置、图像识别装置。

[30]其中，RGB 图像拍摄装置可以拍摄出 RGB 图像，深度图像拍摄装置可以拍摄出深度图像，图像识别装置可以基于 RGB 图像以及深度图像，进行前景分割，分割出目标对象的信息。在实际的应用中，RGB 图像拍摄装置、深度图像拍摄装置通常设置在同一地点，以便针对同一视觉范围进行拍摄。RGB 图像拍摄装置、深度图像拍摄装置以及图像识别装置中的任意一个或多个可以被设置于一个独立的设备中，也可以集成在位于业务场景中的 POS（销售终端）机内部。

[31]应该理解，图 1 中的 RGB 图像拍摄装置、深度图像拍摄装置、图像识别装置的数目仅仅是示意性的。根据实现需要，可以选择和布设任意数目。

[32]图 1 中各装置通过网络交互。其中，网络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。

[33]本说明书实施例的方法包括：首先基于 RGB 图像以及深度图像训练出一个分割模型，然后将分割模型设置在图像识别装置中，这样在实际应用中，就可以将待分割的 RGB

图像以及深度图像输入图像识别装置中的分割模型，从而得到目标对象的分割信息。

[34]下面说明在本说明书实施例中分割模型的训练方法。

[35]图 2 是本说明书一个实施例中分割模型的训练方法的流程图。可以理解，该方法可以通过任何具有计算、处理能力的装置、设备、平台、设备集群来执行。参见图 2，在本说明书实施例中，该分割模型可以是一种由多个网络模型组成的联合模型，具体包括：第一网络模型、第二网络模型以及第三网络模型，该训练方法包括：步骤 201：获取样本图像对；其中，样本图像对中包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像。

[36]步骤 203：将深度图像输入第一网络模型，得到该第一网络模型输出的第一深度特征提取结果。

[37]步骤 205：将深度图像与 RGB 图像的组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征。

[38]步骤 207：将目标对象的边缘特征以及第一深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果。

[39]步骤 209：根据样本图像对的标签以及该目标对象的分割结果，对第一网络模型、第二网络模型以及第三网络模型进行参数调整。

[40]从上述图 2 所示的流程中可以看出，为了能够更加准确地得到图像中的目标对象的分割信息，需要训练出分割模型。在对分割模型进行训练时，不仅利用了 RGB 图像而且还利用了深度图像，通过深度图像可以得到深度特征提取结果即目标对象的大致轮廓信息，通过 RGB 图像可以得到边缘细节信息，这样，综合深度图像及 RGB 图像，就可以在大致轮廓的基础上补充边缘的细节信息，从而得到更为准确的目标对象的分割信息。

[41]并且，在上述图 2 所示过程中，不仅是在训练过程的初始阶段利用了深度图像（即，将深度图像与 RGB 图像进行组合，利用得到目标对象的边缘特征），而且还在训练过程的后续阶段利用了对深度图像的深度特征提取结果，通过使用第一网络模型、第二网络模型以及第三网络模型，实现了同时利用组合图像以及深度特征提取结果来训练分割模型，可见，在训练过程中的不同阶段分别利用了深度图像提供的深度信息，从而使得训练出的分割模型能够更加准确地得到图像中的目标对象的分割信息。

[42]在本说明书的实施例中，如前所述，会在训练过程中的不同阶段分别利用深度图像提供的深度信息，该不同阶段至少可以包括如下两种：方式 1：初始阶段+最后阶段。

[43]在该方式 1 中，在一次训练过程中的初始阶段以及最后阶段中两次利用深度图像提供的深度信息。具体地，参见图 3A，首先，在初始阶段，将拍摄得到的深度图像与 RGB

图像进行组合后，输入网络模型 2，通过该网络模型 2 利用了一次深度图像提供的深度信息；其次，因为拍摄得到的深度图像还会同时被输入网络模型 1，由该网络模型 1 输出第一深度特征提取结果，该第一深度特征提取结果会与网络模型 2 最终输出的图像中目标对象的边缘特征一起被输入网络模型 3，通过该网络模型 3 再次利用深度图像提供的深度信息。

[44]方式 2：初始阶段+中间阶段+最后阶段。

[45]在该方式 2 中，在一次训练过程中的初始阶段、中间阶段以及最后阶段共三次利用深度图像提供的深度信息。具体地，参见图 4A，首先，在初始阶段，将拍摄得到的深度图像与 RGB 图像进行组合后，输入网络模型 2，通过该网络模型 2 利用了一次深度图像提供的深度信息；其次，拍摄得到的深度图像还会同时被输入网络模型 1，该网络模型 1 的中间层神经网络也会得到一个初步的目标对象的轮廓信息，该初步的目标对象的轮廓信息虽然不是网络模型 1 的最终输出，但是也可以从一种分割精度上体现目标对象的轮廓信息，因此，该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果（即深度特征提取的中间结果）提供给网络模型 2，由网络模型 2 利用根据组合图像得到的边缘特征以及该第二深度特征提取结果进行处理，从而得到网络模型 2 的最终输出，可见，在网络模型 2 的中间处理过程中通过利用第二深度特征提取结果第二次利用了深度图像提供的深度信息；最后，因为拍摄得到的深度图像还会同时被输入网络模型 1，由该网络模型 1 输出第一深度特征提取结果，该第一深度特征提取结果会与网络模型 2 最终输出的图像中目标对象的边缘特征一起被输入网络模型 3，通过该网络模型 3 第三次利用了深度图像提供的深度信息。

[46]无论采用上述方式 1 还是方式 2，在本说明书实施例中，网络模型 3 会最终输出目标对象的分割信息，也即得到了分割模型的最终输出结果，根据该分割信息以及样本图像对的标签就可以调整分割模型的参数，以便实现对分割模型的训练。

[47]样本图像对的标签可以包括：第一标签以及第二标签；其中，第一标签为预先由人工形成的对所述 RGB 图像或所述深度图像的分割结果，即分割结果的真实值；第二标签为对所述第一标签进行高斯模糊处理后得到的。

[48]在对分割模型进行参数调整时，参见图 3A 和图 4A，根据第一标签与目标对象的分割结果之间的差异，对网络模型 2 以及网络模型 3 进行参数调整；根据第二标签与第一深度特征提取结果之间的差异，对网络模型 1 进行参数调整。

[49]需要说明的是，在上述中，是根据第一标签与目标对象的分割结果之间的差异，对网络模型 2 进行参数调整。在本说明书的其他实施例中，也可以将第一标签即分割结果

的真实值的边缘部分进行腐蚀，得到第三标签，然后利用该第三标签与网络模型 2 输出的目标对象的边缘特征之间的差异，对网络模型 2 进行参数调整。

[50]训练过程通常是通过多轮训练完成的，直至分割模型收敛。基于上述方式 1，训练出的分割模型也即后续在图像识别业务过程中应用的分割模型的应用结构如图 3B 所示。基于上述方式 2，训练出的分割模型也即后续在图像识别业务过程中应用的分割模型的应用结构如图 4B 所示。

[51]下面结合具体的实施例并以采用上述方式 2 为例来对图 2 所示的过程进行详细说明。参见图 4A、图 4B 以及图 5，具体包括：步骤 501：获取样本图像对；其中，样本图像对中包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像。

[52]通常，RGB 拍摄装置与深度图像拍摄装置安装在同一地点，以便能拍摄到相同视觉范围内的图像。比如，均安装在收银处的 POS 机上。以人脸支付场景为例，从大致相同的位置利用 RGB 图像拍摄装置及深度图像拍摄装置对当前待付款的人进行拍摄，得到 RGB 图像及深度图像，RGB 图像及深度图像中均包括人像的信息，并且很可能会包括多个人像的信息。

[53]步骤 503：将深度图像输入网络模型 1。

[54]网络模型 1 的作用是：从深度图像中提取深度结构信息，也即目标对象的大致轮廓信息。网络模型 1 的结构可以是多层的卷积神经网络。

[55]网络模型 1 可以包括 MobileNetV2。在网络模型 1 中，MobileNetV2 提取出图像中各个对象的深度数据特征（比如深度人脸数据特征），之后再使用反卷积操作，将卷积结果上采样到输入的深度图像的 1/4 大小。

[56]步骤 505：将深度图像与 RGB 图像进行组合，将得到的组合图像输入网络模型 2。

[57]网络模型 2 的作用是：利用 RGB 图像的信息对通过深度图像得到的目标对象轮廓信息中的边缘细节信息进行补全，从而使得分割出的目标对象的轮廓更加清晰准确。网络模型 2 的结构可以是多层的卷积神经网络。

[58]在本步骤 505 中，组合图像实际上是将深度图像与 RGB 图像拼接在一起后生成的，比如原 RGB 图像为 3 通道，将深度图像拼接在第 4 个通道上，则得到了组合图像。

[59]并且，在本步骤 505 中，为了便于处理，可以首先对 RGB 图像的像素值和深度图像的像素值进行归一化，比如归一化为 0-1 中的一个值，其中，对于深度图像，将深度图像中值为空的像素的像素值归一化为 0。

[60]步骤 507：在网络模型 1 中，中间层神经网络提取出目标对象的轮廓信息；将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给网络

模型 2。

[61]步骤 509: 网络模型 1 最终输出图像中目标对象的轮廓信息, 将该目标对象的轮廓信息作为第一深度特征提取结果输入到网络模型 3。

[62]步骤 511: 在网络模型 2 中, 前端各层神经网络对组合图像进行特征提取, 得到初级边缘特征。

[63]步骤 513: 在网络模型 2 中, 后端各层神经网络对该初级边缘特征以及接收到的第二深度特征提取结果进行处理, 得到并向网络模型 3 输出图像中目标对象的边缘特征。

[64]在本步骤 513 中, 因为网络模型 1 的中间处理结果 (即第二深度特征提取结果) 与网络模型 2 的中间处理结果 (即初级边缘特征) 需要一起作为网络模型 2 中后端各层神经网络的输入, 因此, 需要使得该两个中间结果的大小相同, 即图像的大小相同。在本说明书实施例中, 可以调整网络模型 1 及网络模型 2 的卷积核以及卷积步长, 使得该两个中间结果 (即, 初级边缘特征及第二深度特征提取结果) 对应的图像尺寸相同。

[65]需要说明的是, 如果是利用上述方式 1 来实现分割模型的训练, 那么, 本步骤 513 中, 在网络模型 2 中, 后端各层神经网络对该初级边缘特征进行处理 (不再利用第二深度特征提取结果), 得到并向网络模型 3 输出图像中目标对象的边缘特征。

[66]步骤 515: 在网络模型 3 中, 对输入的图像中目标对象的边缘特征以及第一深度特征提取结果进行处理, 得到并输出目标对象的分割信息。

[67]步骤 517: 根据第一标签与目标对象的分割结果之间的差异, 对网络模型 2 以及网络模型 3 进行参数调整。

[68]在本步骤 517 中, 第一标签是预先由人工形成的对上述 RGB 图像或者深度图像中目标对象的分割信息, 因此, 根据该第一标签与网络模型 3 输出的目标对象的分割信息之间的差异, 就可以同时调整网络模型 2、网络模型 3 的参数, 以便优化分割模型。

[69]步骤 519: 对第一标签进行高斯模糊处理, 得到第二标签; 根据第二标签与第一深度特征提取结果之间的差异, 对网络模型 1 进行参数调整。

[70]参见图 3A 和图 4A, 可以首先对网络模型 1 输出的第一深度特征提取结果进行反卷积操作, 然后再利用反卷积操作的结果与第二标签之间的差异, 对网络模型 1 进行参数调整。

[71]参见上述步骤 503, 因为网络模型 1 的最终输出通常为输入图像的 1/4 大小 (步骤 503 中将卷积结果上采样到输入的深度图像的 1/4 大小), 因此, 本步骤 519 中, 可以将第一标签缩小到原始大小的 1/4 之后, 再进行高斯模糊处理, 得到第二标签。

[72]在本步骤 519 中, 第二标签是根据人工标签即第一标签生成的, 因此, 根据该第二

标签与网络模型 1 输出的第一深度特征提取结果之间的差异,就可以调整网络模型 1 的参数,以便优化网络模型 1,从而能够在后续训练过程中,利用优化后的网络模型 1 来继续训练分割模型。

[73]在本说明书的实施例中,可以利用多组样本图像执行多次步骤 501 至步骤 519 的训练过程,直至分割模型收敛。

[74]根据上述图 5 所示过程可以看出,本说明书实施例不仅是在训练过程的初始阶段利用了原始的深度图像(即,将深度图像与 RGB 图像进行组合,利用组合图像训练分割模型),而且还在训练过程的后续阶段两次利用了对深度图像的处理结果(即,利用深度特征提取结果训练分割模型),也就是说,是同时利用组合图像以及深度特征提取结果来训练分割模型,可见,在训练过程中的不同阶段分别利用了深度图像提供的深度信息,从而使得训练出的分割模型能够更加准确地得到图像中的目标对象的分割信息。

[75]可以利用本说明书任一实施例方法所训练出的分割模型进行图像识别。参见图 6,在本说明书一个实施例中,图像识别方法包括:步骤 601:获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像。

[76]步骤 603:将待处理的深度图像输入分割模型中的网络模型 1,得到该网络模型 1 输出的深度特征提取结果。

[77]步骤 605:将待处理的深度图像与待处理的 RGB 图像的组合图像输入分割模型中的网络模型 2,得到该网络模型 2 输出的目标对象的边缘特征。

[78]步骤 607:将该目标对象的边缘特征以及该深度特征提取结果输入分割模型中的网络模型 3,得到该网络模型 3 输出的目标对象的分割结果。

[79]在利用上述图 6 所示方法从图像中得到目标对象的分割信息(比如当前待收款的人像信息,或者当前待收款的商品信息)之后,则可以利用该目标对象的分割信息进行后续的业务处理,比如,因为已经从多个人像中分割出目标人像的信息,因此,就可以有针对性的对该目标人像的人脸进行活体检测及人脸支付,通过这种方式,可以避免人脸识别过成中不想刷脸的人的脸出现在屏幕上的问题,同时解决排队人员出现在屏幕上导致的隐私不安全感的问题,整体流程对用户隐私更友好。

[80]在本说明书的一个实施例中,提供了一种分割模型的训练装置,参见图 7,该装置包括:样本图像获取模块 701,配置为获取样本图像对;其中,样本图像对中包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像;第一网络模型训练模块 702,配置为将所述深度图像输入第一网络模型,得到该第一网络模型输出的第一深度特征提取结果;第二网络模型训练模块 703,配置为将所述深度图像与所述 RGB 图像组合,将该

组合图像输入第二网络模型,得到该第二网络模型输出的目标对象的边缘特征;第三网络模型训练模块 704,配置为将所述目标对象的边缘特征以及所述第一深度特征提取结果输入第三网络模型,得到该第三网络模型输出的目标对象的分割结果;调整模块 705,配置为根据所述样本图像对的标签以及该目标对象的分割结果,对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整。

[81]在上述图 7 所示的本说明书的训练装置的一个实施例中,样本图像对的标签包括:第一标签以及第二标签;其中,第一标签为预先由人工形成的对所述 RGB 图像或所述深度图像的分割结果;第二标签为对所述第一标签进行高斯模糊处理后得到的;调整模块 705 被配置为执行:根据第一标签与目标对象的分割结果之间的差异,对第二网络模型以及第三网络模型进行参数调整;根据第二标签与第一深度特征提取结果之间的差异,对第一网络模型进行参数调整。

[82]在上述图 7 所示的本说明书的训练装置的一个实施例中,第一网络模型训练模块 702 进一步被配置为执行:得到第一网络模型包括的中间层神经网络所提取出的目标对象的轮廓信息,将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给所述第二网络模型;则第二网络模型训练模块 703 进一步被配置为执行:控制第二网络模型包括的前端各层神经网络对组合图像进行特征提取以得到初级边缘特征,以及控制第二网络模型包括的后端各层神经网络对该初级边缘特征以及第二深度特征提取结果进行处理,以由第二网络模型输出目标对象的边缘特征。

[83]在上述图 7 所示的本说明书的训练装置的一个实施例中,调整模块 705 被配置为执行:调整第一网络模型及第二网络模型的卷积核以及卷积步长,使得初级边缘特征与第二深度特征提取结果对应的图像尺寸相同。

[84]在上述图 7 所示的本说明书的训练装置的一个实施例中,第二网络模型训练模块 703 进一步被配置为执行:在将深度图像与 RGB 图像进行组合之前,对 RGB 图像的像素值和深度图像的像素值进行归一化,并且将深度图像中值为空的像素的像素值归一化为 0。

[85]在本说明书一个实施例中,提出了一种图像识别装置。参见图 8,该装置包括:图像输入模块 801,配置为获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像;第一网络模型 802,配置为接收所述待处理的深度图像,得到深度特征提取结果;第二网络模型 803,配置为接收所述待处理的深度图像与所述待处理的 RGB 图像的组合图像,得到目标对象的边缘特征;第三网络模型 804,配置为接收该目标对象的边缘特征以及该深度特征提取结果,得到目标对象的分割结果。

[86]本说明书一个实施例提供了一种计算机可读存储介质,其上存储有计算机程序,当

计算机程序在计算机中执行时，令计算机执行说明书中任一个实施例中的方法。

[87]本说明书一个实施例提供了一种计算设备，包括存储器和处理器，存储器中存储有可执行代码，处理器执行可执行代码时，实现执行说明书中任一个实施例中的方法。

[88]可以理解的是，本说明书实施例示意的结构并不构成对本说明书实施例的装置的具体限定。在说明书的另一些实施例中，上述装置可以包括比图示更多或者更少的部件，或者组合某些部件，或者拆分某些部件，或者不同的部件布置。图示的部件可以以硬件、软件或者软件和硬件的组合来实现。

[89]上述装置、系统内的各模块之间的信息交互、执行过程等内容，由于与本说明书方法实施例基于同一构思，具体内容可参见本说明书方法实施例中的叙述，此处不再赘述。

[90]本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[91]本领域技术人员应该可以意识到，在上述一个或多个示例中，本公开所描述的功能可以用硬件、软件、挂件或它们的任意组合来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。

[92]以上的具体实施方式，对本公开的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上仅为本公开的具体实施方式而已，并不用于限定本公开的保护范围，凡在本公开的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本公开的保护范围之内。

权利要求书

1、一种分割模型的训练方法，该分割模型包括：第一网络模型、第二网络模型以及第三网络模型，其中，所述方法包括：

获取样本图像对；其中，所述样本图像对中包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像；

将所述深度图像输入第一网络模型，得到该第一网络模型输出的第一深度特征提取结果；

将所述深度图像与所述 RGB 图像的组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；

将所述目标对象的边缘特征以及所述第一深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果；

根据所述样本图像对的标签以及该目标对象的分割结果，对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整。

2、根据权利要求 1 所述的方法，其中，

所述样本图像对的标签包括：第一标签以及第二标签；其中，第一标签为预先由人工形成的对所述 RGB 图像或所述深度图像的分割结果；第二标签为对所述第一标签进行高斯模糊处理后得到的；

所述对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整，包括：

根据第一标签与目标对象的分割结果之间的差异，对第二网络模型以及第三网络模型进行参数调整；

根据第二标签与第一深度特征提取结果之间的差异，对第一网络模型进行参数调整。

3、根据权利要求 1 所述的方法，其中，

在所述将所述深度图像输入第一网络模型中之后，进一步包括：得到第一网络模型包括的中间层神经网络所提取出的目标对象的轮廓信息，将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给所述第二网络模型；

在所述将所述深度图像与所述 RGB 图像的组合图像输入第二网络模型之后，并在得到该第二网络模型输出的目标对象的边缘特征之前，进一步包括：

由所述第二网络模型包括的前端各层神经网络对所述组合图像进行特征提取，得到初级边缘特征；

由第二网络模型包括的后端各层神经网络对该初级边缘特征以及第二深度特征提取结果进行处理，以便得到并输出所述目标对象的边缘特征。

4、根据权利要求 3 所述的方法，其中，调整所述第一网络模型及所述第二网络模型的卷积核以及卷积步长，使得所述初级边缘特征与所述第二深度特征提取结果对应的图像尺寸相同。

5、根据权利要求 1 所述的方法，其中，在将所述深度图像与所述 RGB 图像进行组合之前，进一步包括：

对所述 RGB 图像的像素值和所述深度图像的像素值进行归一化，并且将深度图像中值为空的像素的像素值归一化为 0。

6、一种图像识别方法，包括：

获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像；

将待处理的深度图像输入第一网络模型，得到该第一网络模型输出的深度特征提取结果；

将待处理的深度图像与待处理的 RGB 图像的组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；

将该目标对象的边缘特征以及该深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果。

7、一种分割模型的训练装置，包括：

样本图像获取模块，配置为获取样本图像对；其中，样本图像对中包括对同一视觉范围拍摄后得到的 RGB 图像以及深度图像；

第一网络模型训练模块，配置为将所述深度图像输入第一网络模型，得到该第一网络模型输出的第一深度特征提取结果；

第二网络模型训练模块，配置为将所述深度图像与所述 RGB 图像组合，将该组合图像输入第二网络模型，得到该第二网络模型输出的目标对象的边缘特征；

第三网络模型训练模块，配置为将所述目标对象的边缘特征以及所述第一深度特征提取结果输入第三网络模型，得到该第三网络模型输出的目标对象的分割结果；

调整模块，配置为根据所述样本图像对的标签以及该目标对象的分割结果，对所述第一网络模型、第二网络模型以及第三网络模型进行参数调整。

8、根据权利要求 7 所述的装置，其中，

第一网络模型训练模块进一步被配置为执行：得到第一网络模型包括的中间层神经网络所提取出的目标对象的轮廓信息，将该中间层神经网络所提取出的目标对象的轮廓信息作为第二深度特征提取结果输出给所述第二网络模型；

第二网络模型训练模块进一步被配置为执行：控制所述第二网络模型包括的前端各

层神经网络对所述组合图像进行特征提取以得到初级边缘特征,以及控制第二网络模型包括的后端各层神经网络对该初级边缘特征以及第二深度特征提取结果进行处理,以由第二网络模型输出所述目标对象的边缘特征。

9、一种图像识别装置,包括:

图像输入模块,配置为获取对同一视觉范围拍摄后得到的待处理的 RGB 图像以及待处理的深度图像;

第一网络模型,配置为接收所述待处理的深度图像,得到深度特征提取结果;

第二网络模型,配置为接收所述待处理的深度图像与所述待处理的 RGB 图像的组合图像,得到目标对象的边缘特征;

第三网络模型,配置为接收该目标对象的边缘特征以及该深度特征提取结果,得到目标对象的分割结果。

10、一种计算设备,包括存储器和处理器,所述存储器中存储有可执行代码,所述处理器执行所述可执行代码时,实现权利要求 1-6 中任一项所述的方法。