



(12) 发明专利申请

(10) 申请公布号 CN 112967251 A

(43) 申请公布日 2021.06.15

(21) 申请号 202110239211.0

(22) 申请日 2021.03.03

(71) 申请人 网易(杭州)网络有限公司

地址 310052 浙江省杭州市滨江区长河街
道网商路599号4幢7层

(72) 发明人 唐吉霖 袁焱 范长杰 胡志鹏

(74) 专利代理机构 北京超凡宏宇专利代理事务
所(特殊普通合伙) 11463

代理人 钟扬飞

(51) Int. Cl.

G06T 7/00 (2017.01)

G06K 9/46 (2006.01)

G06K 9/62 (2006.01)

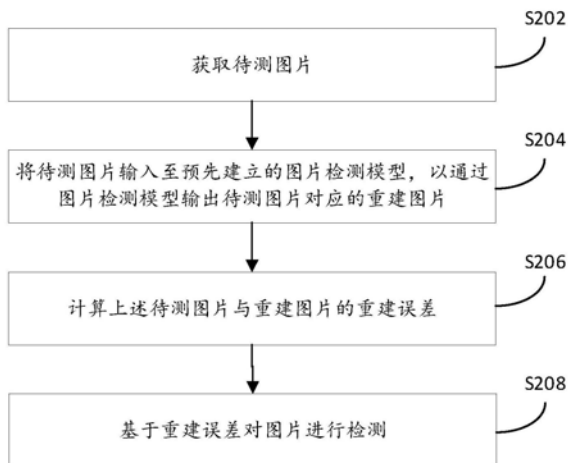
权利要求书3页 说明书11页 附图4页

(54) 发明名称

图片检测方法、图片检测模型的训练方法及装置

(57) 摘要

本发明提供了一种图片检测方法、图片检测模型的训练方法及装置,涉及图片检测的技术领域,该图片检测方法包括:获取待测图片;将待测图片输入至预先建立的图片检测模型,以通过图片检测模型输出待测图片对应的重建图片;计算待测图片与重建图片的重建误差;基于重建误差对图片进行检测。本发明提供的图片检测方法、图片检测模型的训练方法及装置,由于图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构,从而提高了从输入图片到特征空间的映射复杂性,使得异常样本更易落在正常样本的特征空间之外,减少正常样本跟异常样本之间的混淆,提高了异常检测的准确率。



1. 一种图片检测方法,其特征在于,包括:

获取待测图片;

将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片;其中,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

计算所述待测图片与所述重建图片的重建误差;

基于所述重建误差对所述图片进行检测。

2. 根据权利要求1所述的方法,其特征在于,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,其中,所述多层次自编码器结构包括多层次特征空间;

将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片的步骤包括:

将所述待测图片输入至所述多层次自编码器结构,通过所述多层次自编码器结构将所述待测图片嵌入到所述多层次特征空间,得到所述待测图片的特征分布;

将所述待测图片的特征分布与所述多层次特征空间中预先配置的先验分布进行特征拟合;

将特征拟合的结果输入至所述多层次解码器结构进行图片重组,以生成所述待测图片对应的重建图片。

3. 根据权利要求2所述的方法,其特征在于,所述多层次自编码器结构为基于卷积神经网络的编码器结构,且,所述多层次自编码器结构包括至少一个卷积结构,所述卷积结构包括预设卷积层和激活层,且,所述卷积层设置有预设的卷积核和步长;

将所述待测图片输入至所述多层次自编码器结构,通过所述多层次自编码器结构将所述待测图片嵌入到所述多层次特征空间,得到所述待测图片的特征分布的步骤包括:

将所述待测图片输入至所述多层次自编码器结构,通过至少一个所述卷积结构得到所述待测图片的层次特征信息;

将所述层次特征信息嵌入至所述多层次特征空间,以得到所述待测图片的特征分布,其中,所述特征分布包括所述待测图片的低阶表示和所述待测图片的高阶表示。

4. 根据权利要求3所述的方法,其特征在于,所述多层次特征空间还包括高阶特征判别器和低阶特征判别器;

将所述待测图片的特征分布与所述多层次特征空间中预先配置的先验分布进行特征拟合的步骤包括:

将所述待测图片的特征分布中所包括的低阶表示输入至所述低阶特征判别器,以及,将所述待测图片的特征分布中所包括的高阶表示输入至所述高阶特征判别器;

通过所述高阶特征判别器、所述低阶特征判别器和所述多层次特征空间中预先配置的先验分布对所述待测图片的特征分布进行拟合。

5. 根据权利要求1所述的方法,其特征在于,计算所述待测图片与所述重建图片的重建误差的步骤包括:

分别计算所述待测图片和所述重建图片的第一范数距离,以及,所述待测图片和所述重建图片在所述多层次特征空间的第二范数距离;

基于所述第一范数距离和所述第二范数距离计算所述待测图片与所述重建图片的重建误差。

6. 根据权利要求5所述的方法,其特征在于,所述第一范数距离和所述第二范数距离均为L2范数举例;

所述重建误差表示为:

$$A(\hat{X}) = \alpha \|X - \hat{X}\|_2 + \beta \sum_i \|E_i(X) - E_i(\hat{X})\|_2$$

其中, $E_i(\cdot)$ 表示由所述自编码器网络结构中的多层次自编码器结构的第i层的特征图所求得特征表示, α 和 β 表示权重超参数。

7. 根据权利要求1所述的方法,其特征在于,基于所述重建误差对所述图片进行检测的步骤包括:

判断所述重建误差是否小于预先设置的误差阈值;

如果是,确定所述待测图片为正常图片;

如果不是,确定所述待测图片为异常图片。

8. 一种图片检测模型的训练方法,其特征在于,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

其中,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,且,所述多层次自编码器结构包括多层次特征空间;

所述方法包括:

获取预先建立的图片集,其中,所述图片集包括预设数量的正常图像样本;

将所述正常图像样本输入至待训练的图片检测模型,对所述正常图像样本进行图片重建,得到所述正常图像样本对应的重建图像样本;

计算所述正常图像样本和所述重建图像样本的损失函数;

基于所述损失函数对所述待训练的图片检测模型的模型参数进行迭代优化,以对所述待训练的图片检测模型进行训练。

9. 根据权利要求8所述的方法,其特征在于,所述方法还包括:

记录所述迭代优化的迭代次数;

如果所述迭代次数达到预设的迭代阈值,则结束所述迭代优化;

保存当前待训练的图片检测模型的模型参数。

10. 一种图片检测装置,其特征在于,包括:

第一获取模块,用于获取待测图片;

重建模块,用于将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片;其中,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

第一计算模块,用于计算所述待测图片与所述重建图片的重建误差;

检测模块,用于基于所述重建误差对所述图片进行检测。

11. 一种图片检测模型的训练装置,其特征在于,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

其中,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构

以及多层次解码器结构,且,所述多层次自编码器结构包括多层次特征空间;

所述装置包括:

第二获取模块,用于获取预先建立的图片集,其中,所述图片集包括预设数量的正常图像样本;

输入模块,用于将所述正常图像样本输入至待训练的图片检测模型,对所述正常图像样本进行图片重建,得到所述正常图像样本对应的重建图像样本;

第二计算模块,用于计算所述正常图像样本和所述重建图像样本的损失函数;

优化模块,用于基于所述损失函数对所述待训练的图片检测模型的模型参数进行迭代优化,以对所述待训练的图片检测模型进行训练。

12. 一种电子设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现上述权利要求1-9任一项所述方法的步骤。

13. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行上述权利要求1-9任一项所述方法的步骤。

图片检测方法、图片检测模型的训练方法及装置

技术领域

[0001] 本发明涉及图片检测的技术领域,尤其是涉及一种图片检测方法、图片检测模型的训练方法及装置。

背景技术

[0002] 异常图片检测旨在识别、检测出给定图片集合中不符合预期模式或不同于数据集中其他正常图片的异常图片样本。在实际应用中,图片异常存在着无数种不同的情况,无法提前准确、全面地预测每一种可能出现的情况。

[0003] 此外,数据集合中绝大多数都是正常样本,异常样本的数量相比之下十分稀少,正负样本的比例严重失衡且获得异常样本的成本非常高昂。因此,不同于一般的可以依靠大量标记训练数据以监督方式解决的计算机视觉任务,异常图片检测任务通常采用无监督的方式解决。

[0004] 近年来,在无监督异常图片检测领域,主要使用的是基于图片重建的方法,但是这种方式通常仅考虑给定的图片样本在单一层次下的特征表示,难以对同一图片所具有的不同层次的信息进行充分考虑,因而也难以在单一层次特征空间中有效地区分正常样本和异常样本,易造成混淆,引发误判,降低异常检测的准确率。

发明内容

[0005] 有鉴于此,本发明的目的在于提供一种图片检测方法、图片检测模型的训练方法及装置,以缓解上述技术问题。

[0006] 第一方面,本发明实施例提供了一种图片检测方法,包括:获取待测图片;将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片;其中,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;计算所述待测图片与所述重建图片的重建误差;基于所述重建误差对所述图片进行检测。

[0007] 优选地,在一种可能的实施方式中,上述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,其中,所述多层次自编码器结构包括多层次特征空间;将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片的步骤包括:将所述待测图片输入至所述多层次自编码器结构,通过所述多层次自编码器结构将所述待测图片嵌入到所述多层次特征空间,得到所述待测图片的特征分布;将所述待测图片的特征分布与所述多层次特征空间中预先配置的先验分布进行特征拟合;将特征拟合的结果输入至所述多层次解码器结构进行图片重组,以生成所述待测图片对应的重建图片。

[0008] 优选地,在一种可能的实施方式中,上述多层次自编码器结构为基于卷积神经网络的编码器结构,且,所述多层次自编码器结构包括至少一个卷积结构,所述卷积结构包括预设卷积层和激活层,且,所述卷积层设置有预设的卷积核和步长;将所述待测图片输入至

所述多层次自编码器结构,通过所述多层次自编码器结构将所述待测图片嵌入到所述多层次特征空间,得到所述待测图片的特征分布的步骤包括:将所述待测图片输入至所述多层次自编码器结构,通过至少一个所述卷积结构得到所述待测图片的层次特征信息;将所述层次特征信息嵌入至所述多层次特征空间,以得到所述待测图片的特征分布,其中,所述特征分布包括所述待测图片的低阶表示和所述待测图片的高阶表示。

[0009] 优选地,在一种可能的实施方式中,上述多层次特征空间还包括高阶特征判别器和低阶特征判别器;将所述待测图片的特征分布与所述多层次特征空间中预先配置的先验分布进行特征拟合的步骤包括:将所述待测图片的特征分布中所包括的低阶表示输入至所述低阶特征判别器,以及,将所述待测图片的特征分布中所包括的高阶表示输入至所述高阶特征判别器;通过所述高阶特征判别器、所述低阶特征判别器和所述多层次特征空间中预先配置的先验分布对所述待测图片的特征分布进行拟合。

[0010] 优选地,在一种可能的实施方式中,上述计算所述待测图片与所述重建图片的重建误差的步骤包括:分别计算所述待测图片和所述重建图片的第一范数距离,以及,所述待测图片和所述重建图片在所述多层次特征空间的第二范数距离;基于所述第一范数距离和所述第二范数距离计算所述待测图片与所述重建图片的重建误差。

[0011] 优选地,在一种可能的实施方式中,上述第一范数距离和所述第二范数距离均为L2范数举例;所述重建误差表示为:

$$[0012] \quad A(\hat{X}) = \alpha \|X - \hat{X}\|_2 + \beta \sum_i \|E_i(X) - E_i(\hat{X})\|_2$$

[0013] 其中, $E_i(\cdot)$ 表示由所述自编码器网络结构中的多层次自编码器结构的第i层的特征图所求得特征表示, α 和 β 表示权重超参数。

[0014] 优选地,在一种可能的实施方式中,上述基于所述重建误差对所述图片进行检测的步骤包括:判断所述重建误差是否小于预先设置的误差阈值;如果是,确定所述待测图片为正常图片;如果不是,确定所述待测图片为异常图片。

[0015] 第二方面,本发明实施例还提供一种图片检测模型的训练方法,该图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;其中,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,且,所述多层次自编码器结构包括多层次特征空间;该方法包括:获取预先建立的图片集,其中,所述图片集包括预设数量的正常图像样本;将所述正常图像样本输入至待训练的图片检测模型,对所述正常图像样本进行图片重建,得到所述正常图像样本对应的重建图像样本;计算所述正常图像样本和所述重建图像样本的损失函数;基于所述损失函数对所述待训练的图片检测模型的模型参数进行迭代优化,以对所述待训练的图片检测模型进行训练。

[0016] 优选地,在一种可能的实施方式中,上述方法还包括:记录所述迭代优化的迭代次数;如果所述迭代次数达到预设的迭代阈值,则结束所述迭代优化;保存当前待训练的图片检测模型的模型参数。

[0017] 第三方面,本发明实施例还提供了一种图片检测装置,包括:第一获取模块,用于获取待测图片;重建模块,用于将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片;其中,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;第一计算模块,用于计算所述待测图片与所

述重建图片的重建误差;检测模块,用于基于所述重建误差对所述图片进行检测。

[0018] 第四方面,本发明实施例还提供了一种图片检测模型的训练装置,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;其中,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,且,所述多层次自编码器结构包括多层次特征空间;该装置包括:第二获取模块,用于获取预先建立的图片集,其中,所述图片集包括预设数量的正常图像样本;输入模块,用于将所述正常图像样本输入至待训练的图片检测模型,对所述正常图像样本进行图片重建,得到所述正常图像样本对应的重建图像样本;第二计算模块,用于计算所述正常图像样本和所述重建图像样本的损失函数;优化模块,用于基于所述损失函数对所述待训练的图片检测模型的模型参数进行迭代优化,以对所述待训练的图片检测模型进行训练。

[0019] 第五方面,本发明实施例还提供了一种电子设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述第一方面和第二方面所述方法的步骤。

[0020] 第六方面,本发明实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行上述第一方面和第二方面所述方法的步骤。

[0021] 本发明实施例带来了以下有益效果:

[0022] 本发明实施例提供的图片检测方法、图片检测模型的训练方法及装置,在进行图片检测时可以获取待测图片,将待测图片输入至预先建立的图片检测模型,以通过图片检测模型输出待测图片对应的重建图片,并计算待测图片与重建图片的重建误差;以便于基于重建误差对图片进行检测,而由于图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构,使得在图片检测过程中可以利用多层次特征空间及分布约束的自编码器网络结构将输入的图片样本嵌入到多层次特征空间中以更准确地编码、区分不同的图片样本,从而提高了从输入图片到特征空间的映射复杂性,使得异常样本更易落在正常样本的特征空间之外,减少正常样本跟异常样本之间的混淆,提高了异常检测的准确率。

[0023] 本发明的其他特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本发明而了解。本发明的目的和其他优点在说明书、权利要求书以及附图中所特别指出的结构来实现和获得。

[0024] 为使本发明的上述目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附附图,作详细说明如下。

附图说明

[0025] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术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发明的一些实施方式,对于本领域技术人员来讲,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。

[0026] 图1为一种所示无监督异常图片检测的示意图;

[0027] 图2为本发明实施例提供的一种图片检测方法的流程图;

[0028] 图3为本发明实施例提供的一种图片检测的过程示意图;

- [0029] 图4为本发明实施例提供的一种自编码器网络结构的示意图；
- [0030] 图5为本发明实施例提供的一种判别器结构示意图；
- [0031] 图6为本发明实施例提供的一种图片检测模型的训练方法的流程图；
- [0032] 图7为本发明实施例提供的一种图片检测装置的结构示意图；
- [0033] 图8为本发明实施例提供的一种图片检测模型的训练装置的结构示意图；
- [0034] 图9为本发明实施例提供的一种电子设备的结构示意图。

具体实施方式

[0035] 为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合附图对本发明的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0036] 目前，近年来，在无监督异常图片检测领域，主要使用基于图片重建的方法。该类方法多利用编码器-解码器的整体架构(如：自编码器，GAN(Generative Adversarial Networks,生成式对抗网络))来学习正常图片样本的参数化地特征嵌入和图片重建，并通过计算和比较输入图片跟重建图片之间的重建误差来对异常样本进行检测。

[0037] 为便于理解，图1示出了一种所示无监督异常图片检测的示意图，该类方法主要包含编码器网络和解码器网络两个部分，即图1中简单示出的编码器和解码器。

[0038] 具体地，编码器网络主要接受给定的待测图片样本(即图1中的输入图片)作为输入，通过一个卷积神经网络计算得到该样本的特征表示；解码器网络则以编码器输出的图片样本的特征表示作为输入，利用一个卷积神经网络以重建、生成输入的待测样本。在训练过程中，图1所示的模型仅能观察到正常样本，通过梯度下降算法来迭代地更新、优化模型的权重参数以最小化输入的正常样本跟重建的正常样本之间的损失函数，不断提高正常样本的重建质量，降低正常样本的重建误差，从而最终得到期望的重建模型。

[0039] 具体而言，模型的损失函数定义了重建的正常图片跟输入的正常图片之间的相似程度，损失函数越小则表示两张图像越相似。给定训练好的模型，在测试过程中，由于该模型仅以正常样本作为输入进行训练，其参数无法准确地重建异常样本，使得异常样本将产生较高的重建误差，而正常样本将具有较低的重建误差。这样，利用上述差别，便可将计算得到的重建误差跟预设的阈值相比较以检测、区分出图片集中的异常样本。

[0040] 但是，上述方法通常仅考虑了给定的图片样本在单一层次下的特征表示，没有充分考虑同一图片所具有的不同层次的信息，如，底层的结构信息、高层的类别语义信息等等，因此，无法在单一层次特征空间中有效地区分正常样本和异常样本，易造成混淆，引发误判，降低异常检测的准确率。同时，上述方法对正常样本的特征表示的分布没有任何显式地先验约束，使得其自由度过大，难以编码得到紧实的特征空间，也难以有效地限制解码器的泛化能力。

[0041] 基于此，本发明实施例提供一种图片检测方法、图片检测模型的训练方法及装置，可以有效缓解上述问题。

[0042] 为便于对本实施例进行理解，首先对本发明实施例所公开的一种图片检测方法进行详细介绍。

[0043] 在一种可能的实施方式中,本发明实施例提供了一种图片检测方法,图2示出了一种图片检测方法的流程图,具体地,如图2所示,该方法包括以下步骤:

[0044] 步骤S202,获取待测图片;

[0045] 步骤S204,将待测图片输入至预先建立的图片检测模型,以通过图片检测模型输出待测图片对应的重建图片;

[0046] 其中,上述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

[0047] 具体地,上述图片检测模型通常是基于神经网络模型实现的检测模型,且由于本发明实施例中使用的图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构,使得本发明实施例在构建上述图片检测模型时,需同时构建自编码器网络结构所需的编码器网络和解码器网络,以及,自编码器网络结构对应的多层次特征空间,以便于在图片检测时,可以将输入的待测图片嵌入到一个具有互补的多层次结构的特征空间中,进而得到上述待测图片对应的重建图片。

[0048] 步骤S206,计算上述待测图片与重建图片的重建误差;

[0049] 步骤S208,基于重建误差对图片进行检测。

[0050] 在实际使用时,上述重建误差也可以称为待测图片与重建图片的异常分数,通常,计算得到的重建误差主要表征待测图片与重建图片在原始像素空间和多层次特征空间下的相似性,有利于更好地区分正常样本和异常样本,进而提高异常检测性能,因此,利用上述重建误差可以实现对图片进行检测。具体地,上述步骤S208中,对图片进行检测时,可以判断重建误差是否小于预先设置的误差阈值;如果是,确定待测图片为正常图片;如果否,则确定待测图片为异常图片。

[0051] 因此,本发明实施例提供的图片检测方法,在进行图片检测时可以获取待测图片,将待测图片输入至预先建立的图片检测模型,以通过图片检测模型输出待测图片对应的重建图片,并计算待测图片与重建图片的重建误差;以便于基于重建误差对图片进行检测,而由于图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构,使得在图片检测过程中可以利用多层次特征空间及分布约束的自编码器网络结构将输入的图片样本嵌入到多层次特征空间中以更准确地编码、区分不同的图片样本,从而提高了从输入图片到特征空间的映射复杂性,使得异常样本更易落在正常样本的特征空间之外,减少正常样本跟异常样本之间的混淆,提高了异常检测的准确率。

[0052] 在实际使用时,本发明实施例中的多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,其中,多层次自编码器结构还包括多层次特征空间;

[0053] 具体地,考虑到卷积神经网络(CNN,Convolutional Neural Networks,卷积神经网络)所具有的金字塔形的特征层次结构,使得网络中不同层的特征图(feature map)可以包含输入图片的不同层次的信息。而通常,较浅层的特征图通常包含更多的低级结构信息,例如边缘,直线和拐角。同时,更深层的特征图通常捕获了更多的有关类别标签的高级语义信息。因此,网络不同层的特征实际上编码了输入图片不同层次的信息表示,并且低层和高层特征的互补性能够有效提高识别、区分不同图片样本的性能。因此,本发明实施例中自编码器网络结构所包括的多层次自编码器结构,可以将输入的图片样本嵌入到一个具有互补

的层次结构的特征空间中以更准确地编码、区分不同的图片样本。换言之，多层次的特征空间提高了从输入图片到特征空间的映射复杂性，可以使得异常样本更易落在正常样本的特征空间之外。

[0054] 基于上述理论，上述步骤S204中，通过图片检测模型输出重建图片的步骤可以包括以下过程：将待测图片输入至多层次自编码器结构，通过该多层次自编码器结构将待测图片嵌入到多层次特征空间，得到待测图片的特征分布；将待测图片的特征分布与多层次特征空间中预先配置的先验分布进行特征拟合；以将特征拟合的结果输入至多层次解码器结构进行图片重组，进而生成待测图片对应的重建图片。

[0055] 为了便于理解，图3示出了本发明实施例提供的一种图片检测的过程示意图，具体地，图3中，自编码器网络结构为多层次自编码器结构，且，该多层次自编码器结构为基于卷积神经网络的编码器结构，且，多层次自编码器结构包括至少一个卷积结构，卷积结构包括预设卷积层和激活层，且，卷积层设置有预设的卷积核和步长；

[0056] 具体地，由于自编码器网络结构包括多层次自编码器结构以及多层次解码器结构，因此，图3中示出了多层次自编码器结构和多层次解码器结构，其中，为了便于说明，图3中多层次自编码器结构简单表示为编码器，且，图3中仅简单示出了有限个卷积结构，包括预设的卷积层和激活层。

[0057] 进一步，图3中的多层次解码器结构通常也简单表示为解码器，且，为了便于说明，图3中也简单示出了有限个解码器所包括的卷积结构，即有限数量个卷积层和激活层。图3中的中间部分则是多层次特征空间部分。

[0058] 具体地，图4示出了一种可能的自编码器网络结构的示意图，包括多层次自编码器结构和多层次解码器结构，其中，图4中，示出了包含四个卷积层的编码器，且，图4中，以ConvDown模块表示其中一个卷积层，也即，图4中的编码器包括四个ConvDown模块，分别是ConvDown64、ConvDown128、ConvDown256和ConvDown512，其中，上述各个ConvDown模块后的数字则表示该ConvDown模块的卷积核的个数，通常，上述ConvDown模块可以对特征图进行空间下采样，同时还可以逐渐增加特征图的个数，减小空间分辨率，增加深度。并且，在实际使用时，每个ConvDown模块可以包含一个3x3（卷积核大小）且步长为2的卷积层，其后跟一个ReLU激活层，具体的ConvDown模块的设置方式和结构可以根据实际使用情况进行设置，本发明实施例对此不进行限制。

[0059] 进一步，图4中，也示出了包括多个卷积结构的解码器，其中，图4中的解码器的卷积结构则包括多个ConvUp模块，多个ConvUp模块可以在空间上采样，同时逐渐减少特征图的个数，最终输出重建图片。其中，图4中解码器所包括的每个ConvUp模块也可以包含一个3x3（卷积核大小）且步长为2的反卷积层，其后跟一个ReLU激活层，具体的ConvUp模块的设置方式和结构也可以根据实际使用情况进行设置，本发明实施例对此不进行限制。

[0060] 图4中中间部分则为全连接层通常用FC表示，即图4中的FC128、FC1024、FC256和FC2048，这部分的全连接层也称为特征嵌入部分，该特征嵌入部分通常以编码器网络较浅层和较深层的特征图分别作为输入，利用全连接层（FC）来分别地生成输入图片样本的低阶（low-level）和高阶（high-level）特征表示，其中，FC128、FC1024对应低阶的特征表示，FC256和FC2048则对应高阶的特征表示，后面数字表示全连接层的参数，具体可以根据实际使用情况进行设置，本发明实施例对此不进行限制。

[0061] 因此,基于图4所示的网络结构,利用该自编码器网络结构得到待测图片的特征分布的步骤则包括一下过程:

[0062] 将待测图片输入至多层次自编码器结构,即输入至图3和图4中左侧的多层次自编码器结构(编码器),通过至少一个卷积结构得到待测图片的层次特征信息;将层次特征信息嵌入至多层次特征空间,以得到待测图片的特征分布,其中,特征分布包括待测图片的低阶表示和待测图片的高阶表示。

[0063] 通常,低阶表示包括待测图片的低级结构信息,例如边缘,直线和拐角等等,高阶表示则包括待测图片的高级结构信息,如更多有关类别标签的高级语义信息等等。当编码器将上述待测图片嵌入到多层次特征空间,得到待测图片的特征分布后,则可以将待测图片的特征分布与多层次特征空间中预先配置的先验分布进行特征拟合。

[0064] 具体地,为了便于对待测图片的特征分布与多层次特征空间中预先配置的先验分布进行特征拟合,本发明实施例中的多层次特征空间还包括高阶特征判别器和低阶特征判别器;即,图3中所示的高阶特征判别器和低阶特征判别器,在进行特征拟合时,需要将待测图片的特征分布中所包括的低阶表示输入至低阶特征判别器,以及,将待测图片的特征分布中所包括的高阶表示输入至高阶特征判别器;然后通过高阶特征判别器、低阶特征判别器和多层次特征空间中预先配置的先验分布对待测图片的特征分布进行拟合。

[0065] 在实际使用时,上述预先配置的先验分布可以是正态分布,例如,可以是标准正态分布 $N(0, 1)$,以显式地约束、控制正常样本的特征分布。在其他实施方式中,上述预先配置的先验分布还可以是其他的形式,具体以实际使用为准,本发明实施例对此不进行限制。

[0066] 上述通过在嵌入的多层次特征空间中引入一系列特征判别器,如上述高阶特征判别器和低阶特征判别器,来实现上述特征拟合的过程,旨在正确区分从输入的正常样本由编码器网络编码得到的特征表示以及从先验的正态分布中随机采样得到的特征表示;同时,编码器网络则要尽可能地使其编码结果模仿、接近从先验的正态分布中采样得到的结果,以欺骗特征判别器。基于上述过程,两个网络相互对抗、不断迭代优化、调整各自参数,最终使得特征判别器网络无法准确判断编码器网络的输出结果的真实来源。

[0067] 综上,基于上述图3所示自编码器网络结构的示意图,在进行图片检测时,可以使用自编码器网络结构来分别地学习正常图片样本的参数化地特征嵌入和特征重建,并通过计算、比较输入图片跟重建图片之间的重建误差来对异常样本进行检测、区分。具体而言,自编码器网络结构包含多层次自编码器结构和多层次解码器结构两个部分。其中,多层次自编码器结构负责将输入的待测图片嵌入到一个具有互补的层次结构的特征空间中,得到待测图片的特征分布,同时将待测图片的特征分布与多层次特征空间中预先配置的先验分布进行特征拟合;多层次解码器结构,即上述解码器,则旨在相应地学习从该先验特征空间到正常图片空间的特定映射,生成重建图片,进而计算得到重建误差以实现基于阈值比较的异常检测。

[0068] 并且,本发明实施例中,考虑到自编码器网络结构的卷积神经网络不同层的特征实际上编码了输入图片(如上述待测图片)的不同层次的身份信息,并且低层和高层特征的互补能够有效地提高识别、区分不同图片样本的性能。因此,上述图3中所示的自编码器网络结构,可以将输入的图片样本嵌入到一个具有互补的多层次结构的特征空间中,该多层次的特征空间提高了从输入图片到特征空间的映射复杂性,使得异常样本更易落在正常图

片样本的特征空间之外。

[0069] 此外,上述多层次特征空间中预先配置的先验分布,一方面可以有效提高编码得到的正常样本特征空间的紧凑度,降低将异常样本映射到正常样本特征空间而引发混淆的可能性和偶然性;另一方面,一旦定义了这样一个参数化的先验特征空间,上述多层次解码器结构(即,解码器网络)将相应地随之学习从该参数化的先验特征空间到正常图片空间的特定映射,而不是从其他所有可能的未指定的混合特征空间学习,从而将解码器的泛化能力限制在有界的流形范围内,有效抑制网络对异常样本的泛化性能,使异常样本更易产生较高的重建误差。

[0070] 为了便于理解,图5示出了一种判别器结构示意图,如图5所示,分别包括高阶特征判别器和低阶特征判别器,两者的结构类似,仅在全连接层FC的参数有所不同,在实际使用时,上述高阶特征判别器和低阶特征判别器的作用主要是判断给定的特征表示是否是“真实的”,即区分由编码器网络编码得到的特征表示(fake)和从先验的正态分布中随机采样得到的特征表示(real)。通常编码器由多个全连接层,即,FC层构成,其输入是特定的特征表示,其输出为一个标量,代表了该特征表示是“真实”(即,不是由编码器编码得到)的概率。将上述判别器嵌入到图3所示的多层次特征空间中,以正确区分从输入的正常样本由编码器网络编码得到的特征表示以及从先验的正态分布中随机采样得到的特征表示,进而输出对应的重建图片,进行后续的重建误差的计算。

[0071] 进一步,本发明实施例中,在进行重建误差的计算时,通常是分别计算待测图片和重建图片的第一范数距离,以及,待测图片和重建图片在多层次特征空间的第二范数距离;然后基于该第一范数距离和第二范数距离计算待测图片与重建图片的重建误差。

[0072] 具体地,第一范数距离和第二范数距离均为L2范数举例;上述重建误差表示为:

$$[0073] \quad A(\hat{X}) = \alpha \|X - \hat{X}\|_2 + \beta \sum_i \|E_i(X) - E_i(\hat{X})\|_2$$

[0074] 其中, $E_i(\cdot)$ 表示由自编码器网络结构中的多层次自编码器结构的第i层的特征图所求得特征表示, α 和 β 表示权重超参数。 X 表示待测图片, \hat{X} 则表示重建图片, $A(\cdot)$ 表示的则是上述重建误差,也称为异常分数。

[0075] 基于上述自编码器网络结构,本发明实施例提供的图片检测方法,可以利用多层次的自编码器结构将输入的图片样本嵌入到一个具有互补的层次结构的特征空间中以更准确地编码、区分不同的图片样本,从而提高了从输入图片到特征空间的映射复杂性,使得异常样本更易落在正常样本的特征空间之外,减少正常样本跟异常样本之间的混淆,提高了异常检测的准确率。

[0076] 并且,在图片检测过程中,可以显式地约束正常样本的特征分布,有效提高了编码得到的正常样本特征空间的紧凑度,减少了将异常样本映射到正常样本特征空间的可能性同时限制了解码器的泛化能力,进一步提升了异常检测的性能。

[0077] 在实际使用时,上述图片检测模型可以通过对抗训练的方法得到,因此,本发明实施例中还提供了一种图片检测模型的训练方法,且该图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构。

[0078] 其中,上述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,且,多层次自编码器结构包括多层次特征空间。

[0079] 具体地,如图6所示的一种图片检测模型的训练方法的流程图,该方法包括以下步骤:

[0080] 步骤S602,获取预先建立的图片集,其中,图片集包括预设数量的正常图像样本;

[0081] 步骤S604,将正常图像样本输入至待训练的图片检测模型,对正常图像样本进行图片重建,得到正常图像样本对应的重建图像样本;

[0082] 步骤S606,计算正常图像样本和重建图像样本的损失函数;

[0083] 步骤S608,基于损失函数对待训练的图片检测模型的模型参数进行迭代优化,以对待训练的图片检测模型进行训练。

[0084] 进一步,在迭代优化过程中还需记录迭代优化的迭代次数;如果迭代次数达到预设的迭代阈值,则结束迭代优化;保存当前待训练的图片检测模型的模型参数。

[0085] 在实际使用时,通过上述方法训练得到的图片检测模型可以在图片检测中进行调用,具体地,可以记载上述图片检测模型的模型参数进行图片检测,然后输入待测的正常或异常图片,此时,图片检测模型可以根据输入的待测图片生成相应的重建图片,进而计算重建误差,跟预先设置的误差阈值相比较以检测出异常图片。

[0086] 在上述实施方式的基础上,本发明实施例还提供了一种图片检测装置,如图7所示的一种图片检测装置的结构示意图,该装置包括:

[0087] 第一获取模块70,用于获取待测图片;

[0088] 重建模块72,用于将所述待测图片输入至预先建立的图片检测模型,以通过所述图片检测模型输出所述待测图片对应的重建图片;其中,所述图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;

[0089] 第一计算模块74,用于计算所述待测图片与所述重建图片的重建误差;

[0090] 检测模块76,用于基于所述重建误差对所述图片进行检测。

[0091] 进一步,本发明实施例还提供了一种图片检测模型的训练装置,该图片检测模型配置有多层次特征空间及分布约束的自编码器网络结构;且,所述多层次特征空间及分布约束的自编码器网络结构包括多层次自编码器结构以及多层次解码器结构,且,所述多层次自编码器结构包括多层次特征空间;具体地,如图8所示的一种图片检测模型的训练装置的结构示意图,该装置包括:

[0092] 第二获取模块80,用于获取预先建立的图片集,其中,所述图片集包括预设数量的正常图像样本;

[0093] 输入模块82,用于将所述正常图像样本输入至待训练的图片检测模型,对所述正常图像样本进行图片重建,得到所述正常图像样本对应的重建图像样本;

[0094] 第二计算模块84,用于计算所述正常图像样本和所述重建图像样本的损失函数;

[0095] 优化模块86,用于基于所述损失函数对所述待训练的图片检测模型的模型参数进行迭代优化,以对所述待训练的图片检测模型进行训练。

[0096] 本发明实施例提供的种图片检测装置和图片检测模型的训练装置,与上述实施例提供的图片检测方法和图片检测模型的训练方法具有相同的技术特征,所以也能解决相同的技术问题,达到相同的技术效果。

[0097] 本发明实施例还提供一种电子设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述方法

的步骤。

[0098] 本发明实施例还提供一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行上述方法的步骤。

[0099] 进一步,本发明实施例还提供了一种电子设备的结构示意图,如图9所示,为该电子设备的结构示意图,其中,该电子设备包括处理器91和存储器90,该存储器90存储有能够被该处理器91执行的计算机可执行指令,该处理器91执行该计算机可执行指令以实现上述图片检测方法或者图片检测模型的训练方法。

[0100] 在图9示出的实施方式中,该电子设备还包括总线92和通信接口93,其中,处理器91、通信接口93和存储器90通过总线92连接。

[0101] 其中,存储器90可能包含高速随机存取存储器(RAM,Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过至少一个通信接口93(可以是有线或者无线)实现该系统网元与至少一个其他网元之间的通信连接,可以使用互联网,广域网,本地网,城域网等。总线92可以是ISA(Industry Standard Architecture,工业标准体系结构)总线、PCI(Peripheral Component Interconnect,外设部件互连标准)总线或EISA(Extended Industry Standard Architecture,扩展工业标准结构)总线等。所述总线92可以分为地址总线、数据总线、控制总线等。为便于表示,图9中仅用一个双向箭头表示,但并不表示仅有一根总线或一种类型的总线。

[0102] 处理器91可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器91中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器91可以是通用处理器,包括中央处理器(Central Processing Unit,简称CPU)、网络处理器(Network Processor,简称NP)等;还可以是数字信号处理器(Digital Signal Processor,简称DSP)、专用集成电路(Application Specific Integrated Circuit,简称ASIC)、现场可编程门阵列(Field-Programmable Gate Array,简称FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本发明实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器,处理器91读取存储器中的信息,结合其硬件完成前述实施例的图片检测方法或者图片检测模型的训练方法。

[0103] 本发明实施例所提供的图片检测方法、图片检测模型的训练方法及装置的计算机程序产品,包括存储了程序代码的计算机可读存储介质,所述程序代码包括的指令可用于执行前面方法实施例中所述的方法,具体实现可参见方法实施例,在此不再赘述。

[0104] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的装置的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0105] 另外,在本发明实施例的描述中,除非另有明确的规定和限定,术语“安装”、“相连”、“连接”应做广义理解,例如,可以是固定连接,也可以是可拆卸连接,或一体地连接;可以是机械连接,也可以是电连接;可以是直接相连,也可以通过中间媒介间接相连,可以是两个元件内部的连通。对于本领域技术人员而言,可以具体情况理解上述术语在本发明中

的具体含义。

[0106] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备）执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器（ROM, Read-Only Memory）、随机存取存储器（RAM, Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

[0107] 在本发明的描述中，需要说明的是，术语“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。此外，术语“第一”、“第二”、“第三”仅用于描述目的，而不能理解为指示或暗示相对重要性。

[0108] 最后应说明的是：以上实施例，仅为本发明的具体实施方式，用以说明本发明的技术方案，而非对其限制，本发明的保护范围并不局限于此，尽管参照前述实施例对本发明进行了详细的说明，本领域技术人员应当理解：任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化，或者对其中部分技术特征进行等同替换；而这些修改、变化或者替换，并不使相应技术方案的本质脱离本发明实施例技术方案的精神和范围，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以权利要求的保护范围为准。

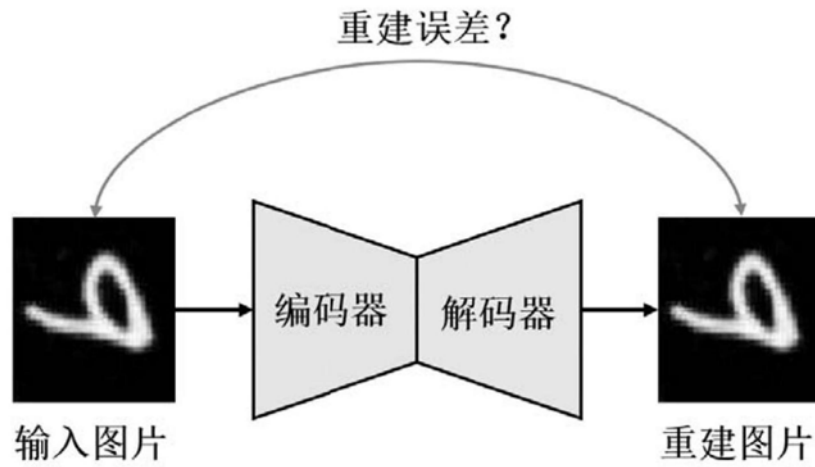


图1

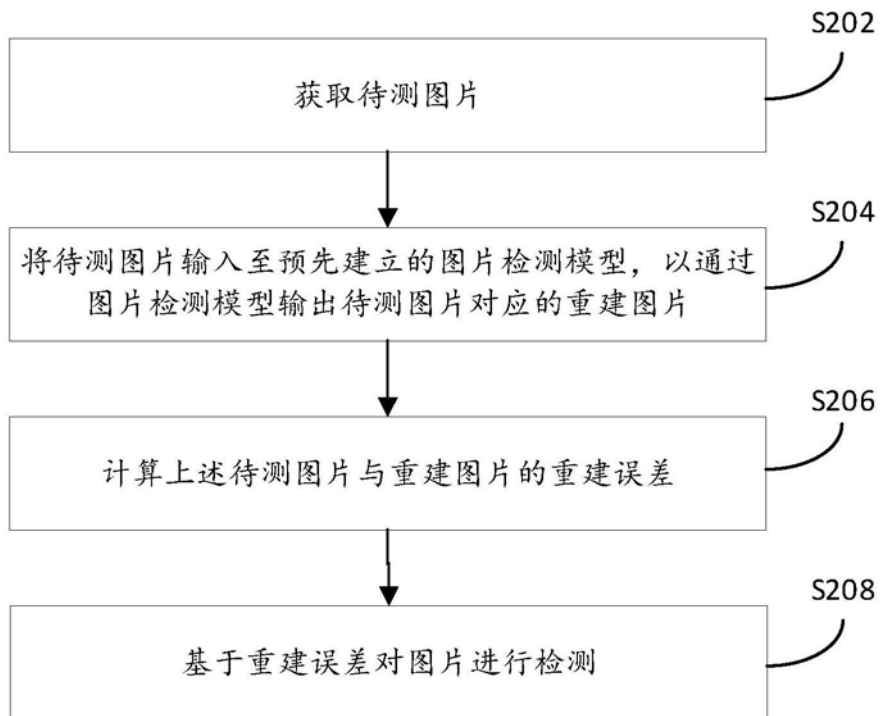


图2

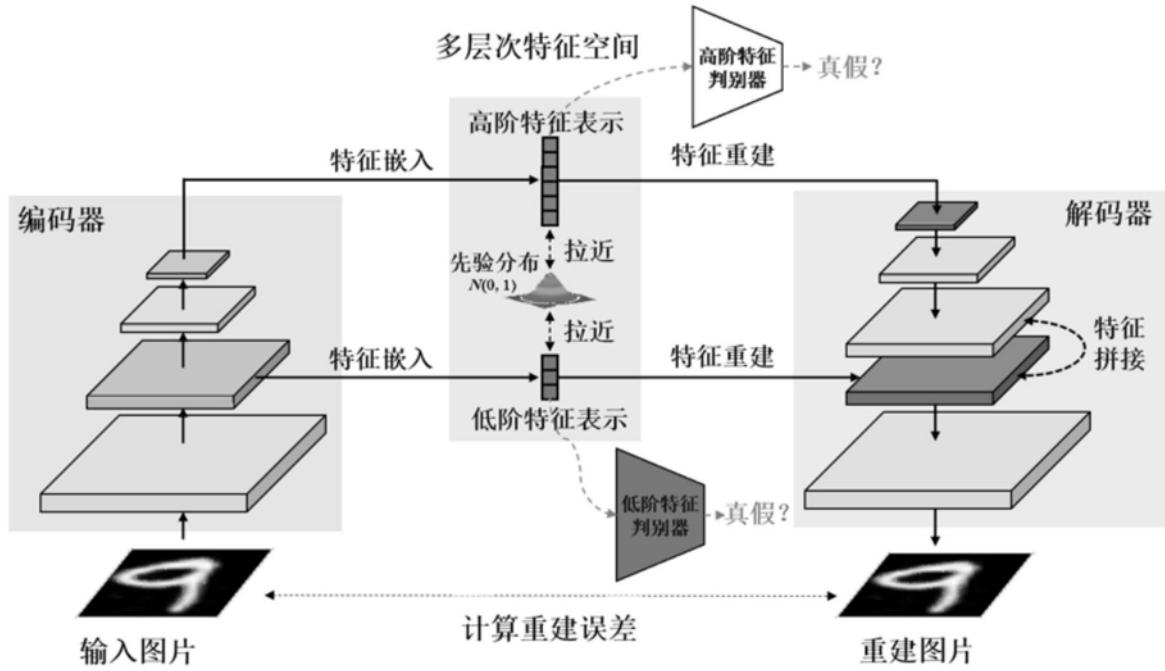


图3

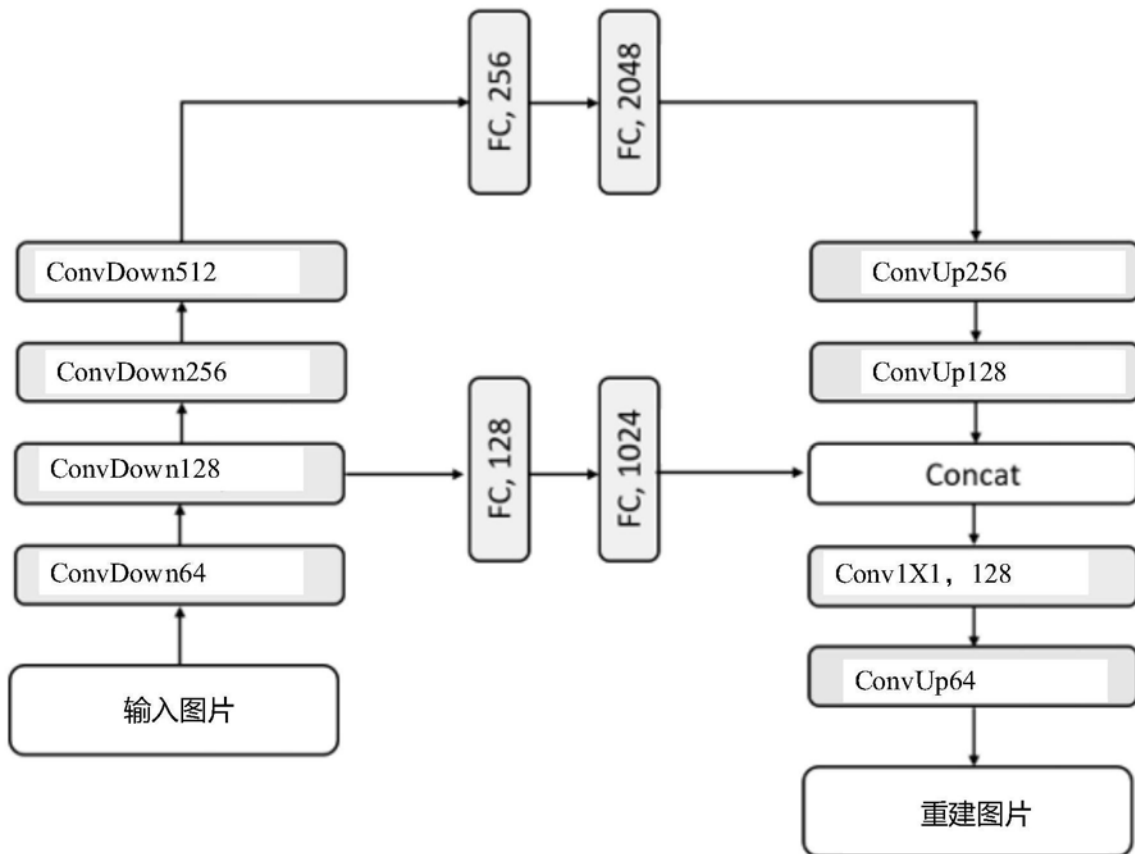


图4

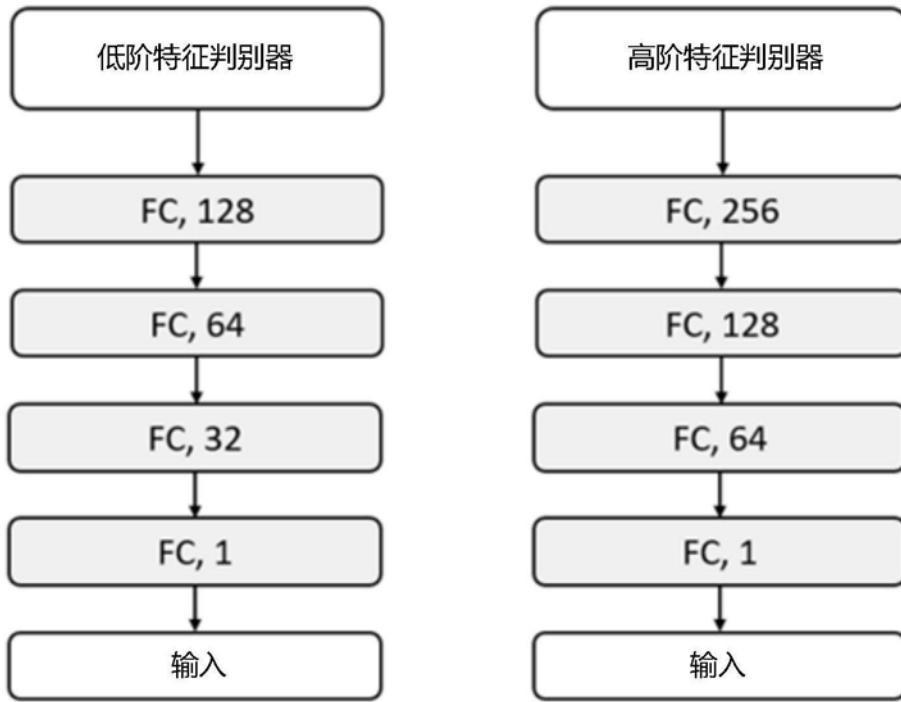


图5

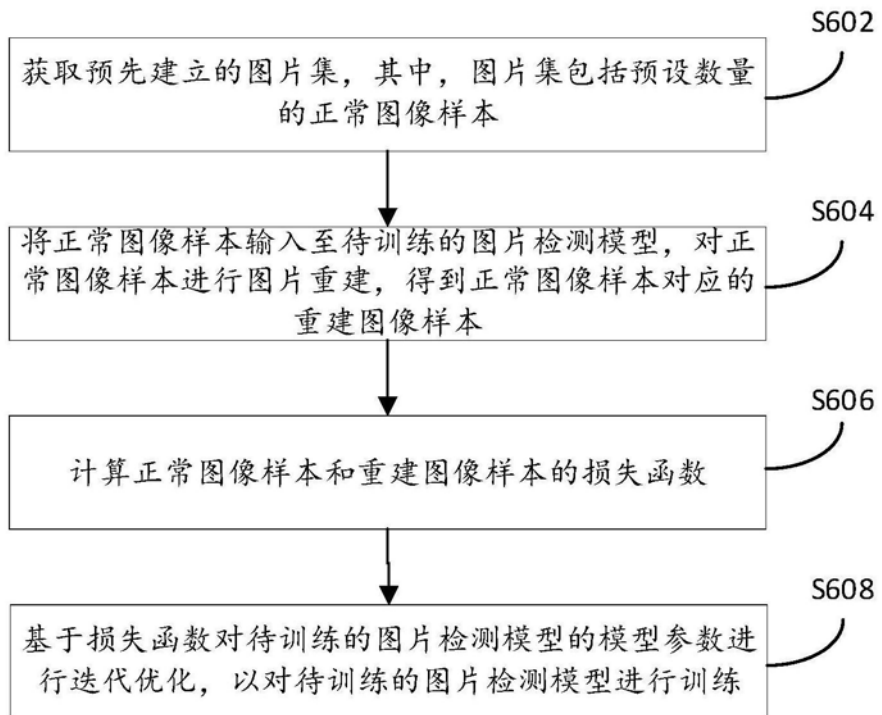


图6

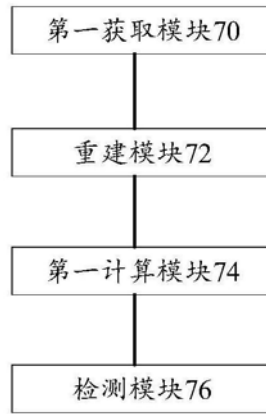


图7



图8

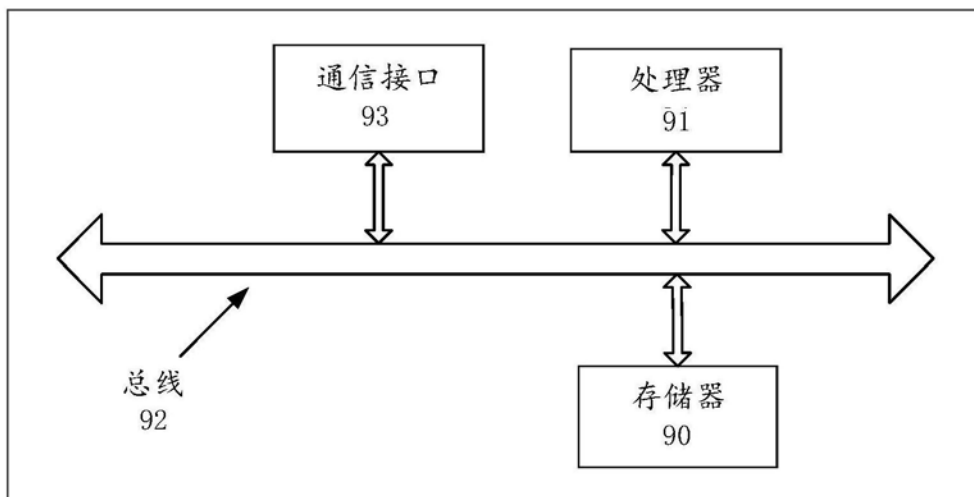


图9