



(12) 发明专利申请

(10) 申请公布号 CN 114067018 A

(43) 申请公布日 2022. 02. 18

(21) 申请号 202111376380.5

(22) 申请日 2021.11.19

(71) 申请人 长春理工大学

地址 130022 吉林省长春市朝阳区卫星路
7186号

(72) 发明人 詹伟达 桂婷婷 朱德鹏 陈宇
李锐

(74) 专利代理机构 北京中理通专利代理事务所
(普通合伙) 11633

代理人 刘慧宇

(51) Int. Cl.

G06T 11/00 (2006.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

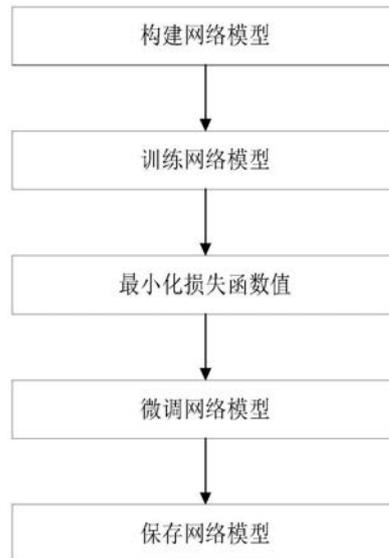
权利要求书1页 说明书6页 附图3页

(54) 发明名称

一种基于扩张残差生成对抗网络的红外图像彩色化方法

(57) 摘要

一种轻量化的单张图像超分辨率重建方法,属于图像超分辨率重建领域,为了解决现有的超分辨率方法空间复杂度和时间复杂度高的问题,该方法包括如下步骤:构建网络模型:整个网络包括四个主要模块:浅层特征提取模块、深层特征提取模块、信息融合模块和上采样模块;准备数据集:对所用数据集进行模拟降质,所得高低分辨率图像对用以训练整个卷积神经网络;训练网络模型;最小化损失值;微调模型;保存模型:将最后得到的模型参数进行固化,之后需要进行超分辨率重建操作时,直接将图像和网络参数载入网络中即可得到最终的超分辨率图像。在保持较高重建质量的前提下,大大减少了网络的参数量和计算量,更适于在嵌入式设备上实现。



1. 一种基于扩张残差生成对抗网络的红外图像彩色化方法,其特征是,该方法包括如下步骤:

步骤1,构建网络模型:整个生成对抗网络包括生成器和鉴别器;

所述的生成器由三个模块组成,分别是浅层特征提取模块、下采样模块和上采样模块,将红外图像输入生成器中,红外图像先经过浅层特征提取模块提取图像浅层信息,再经过下采样模块提取图像深层次的语义信息,最后经过上采样模块重建红外图像色度图像;

鉴别器采用小型VGG网络结构,由六个去掉批归一化的卷积块和两个全连接层组成;生成器输出的色度图像与红外图像融合即是生成的彩色红外图像,将生成的彩色红外图像与数据集中的可见光彩色图像输入鉴别器中,鉴别器输出真假概率信息判断输入图像是否真实;

步骤2,训练网络模型:训练红外图像着色模型,对整个生成对抗网络用可见光数据集ImageNet采用生成器和鉴别器进行对抗训练;调整数据集中每个图像的尺寸,固定输入图像的大小,分离出彩色图像的亮度通道L,将其作为整个网络的输入,将彩色图像作为标签,进行有监督的训练;

步骤3,最小化损失函数值:生成器采用色度损失和对抗损失,鉴别器采用对抗损失,通过最小化生成器和鉴别器的损失函数交替优化生成器和鉴别器,直到生成器和鉴别器的对抗达到平衡状态即可认为模型参数已训练完成,保存模型参数;

步骤4,微调网络模型:用红外图像数据集对模型进行训练和微调,得到稳定可靠的模型参数;最终使得模型着色的效果更好;

步骤5,保存网络模型:将最终确定的模型参数进行固化。

2. 根据权利要求1所述的一种基于扩张残差生成对抗网络的红外图像彩色化方法,其特征在于,所述浅层特征提取模块由卷积块一和卷积块二组成,提取图像初级结构信息;下采样模块由残差卷积块一至残差卷积块三、扩张残差卷积块一至二和卷积块三至六组成,残差卷积块一至三完成下采样操作,扩张残差块一至二在保持图像分辨率不变的前提下维持感受野,进一步提取图像语义高级特征,卷积块三至六消除扩张残差卷积导致的网格效应。

3. 根据权利要求2所述的一种基于扩张残差生成对抗网络的红外图像彩色化方法,其特征在于,所述残差卷积块和扩张残差卷积块网络结构是相同的,都是将上一层网络中输出的特征图与其经过两个残差结构输出的特征图进行拼接操作,增加对于不同网络层特征图信息的有效利用;二者区别在于残差卷积块用于实现下采样操作,所以跳跃连接中存在卷积操作进而实现拼接操作,而扩张残差卷积块是恒等残差块,可直接进行拼接操作。

4. 根据权利要求1所述的一种基于扩张残差生成对抗网络的红外图像彩色化方法,其特征在于,所述上采样模块由卷积块七至九和两次像素重组组成,将下采样模块输出的特征图先经过卷积块七扩张特征图通道,再经过第一次像素重组进行特征图二倍重建,卷积块八继续扩大特征图通道,第二次像素重组进行特征图四倍重建,最后卷积块九输出重建的色度图像。

一种基于扩张残差生成对抗网络的红外图像彩色化方法

技术领域

[0001] 本发明涉及一种基于扩张残差生成对抗网络的红外图像彩色化方法,属于图像彩色化领域。

背景技术

[0002] 图像彩色化又名图像着色,是将灰度图转化成彩色图的过程,在遥感图像、医学图像、夜视图像、电影和动漫等众多领域具有较大的应用空间。红外图像由于其能在极度恶劣的条件下成像,在人们生活中的运用越来越广,但是红外图像与可见光灰度图像一样都是单色图像,相较于彩色图像,缺乏更多的色彩特征,不利于人眼的直接分辨、识别,制约了场景中有效信息的获取。因此,对灰度图像进行着色有利于对图像的特征提取和人眼辨识其中的信息,方便人们工作。然而,现有的图像着色方法大多基于无监督学习的神经网络,会产生三个关键问题,即着色图像的颜色溢出、细节损失、缺乏对比度。

[0003] 中国专利公开号为“CN109712203B”,名称为“一种基于自注意力生成对抗网络的图像着色方法”,该方法首先利用Perceptual Loss损失函数和Hinge Loss损失函数训练灰度图像着色模型;其次,输入待着色的灰度图像;接着,通过卷积运算、谱归一化、批归一化和激活函数运算得到特征提取层输出结果;然后,利用特征融合模块将特征提取卷积模块的输出与反卷积计算阶段联系起来;最后通过自注意力网络输出着色图像。该方法得到的着色图像对比度较低,颜色生硬不自然,不符合人眼视觉效果,同时实现过程复杂和效率低下。

发明内容

[0004] 本发明为了解决现有的彩色化方法得到的着色图像颜色溢出、细节损失、对比度较低的问题,提供了一种基于扩张残差生成对抗网络的红外图像彩色化方法。该方法使得到的着色图像较好的抑制了颜色溢出问题,图像的颜色生动自然,更加符合人眼视觉观察,同时本发明提出的方法实现过程简单,图像着色效率更高。

[0005] 本发明采用的技术方案是:

[0006] 一种基于扩张残差生成对抗网络的红外图像彩色化方法,包含如下步骤:

[0007] 步骤1,构建网络模型:整个生成对抗网络包括生成器和鉴别器;

[0008] 所述的生成器由三个模块组成,分别是浅层特征提取模块、下采样模块和上采样模块,将红外图像输入生成器中,红外图像先经过浅层特征提取模块提取图像浅层信息,再经过下采样模块提取图像深层次的语义信息,最后经过上采样模块重建红外图像色度图像;

[0009] 鉴别器采用小型VGG网络结构,由六个去掉批归一化的卷积块和两个全连接层组成;生成器输出的色度图像与红外图像融合即是生成的彩色红外图像,将生成的彩色红外图像与数据集中的可见光彩色图像输入鉴别器中,鉴别器输出真假概率信息判断输入图像是否真实;

[0010] 步骤2,训练网络模型:训练红外图像着色模型,对整个生成对抗网络用可见光数据集ImageNet采用生成器和鉴别器进行对抗训练。调整数据集中每个图像的尺寸,固定输入图像的大小,分离出彩色图像的亮度通道L,将其作为整个网络的输入,将彩色图像作为标签,进行有监督的训练。

[0011] 步骤3,最小化损失函数值:生成器采用色度损失和对抗损失,鉴别器采用对抗损失,通过最小化生成器和鉴别器的损失函数交替优化生成器和鉴别器,直到生成器和鉴别器的对抗达到平衡状态即可认为模型参数已训练完成,保存模型参数;

[0012] 步骤4,微调网络模型:用红外图像数据集对模型进行训练和微调,得到稳定可靠的模型参数。最终使得模型着色的效果更好;

[0013] 步骤5,保存网络模型:将最终确定的模型参数进行固化。

[0014] 如果需要对红外图像进行着色时,直接将图像输入到以上方法得到的网络模型中即可得到最终的彩色红外图像。

[0015] 浅层特征提取模块由卷积块一和卷积块二组成,提取图像初级结构信息。下采样模块由残差卷积块一至残差卷积块三、扩张残差卷积块一至二和卷积块三至六组成,残差卷积块一至三完成下采样操作,扩张残差块一至二在保持图像分辨率不变的前提下维持感受野,进一步提取图像语义高级特征,卷积块三至六消除扩张残差卷积导致的网格效应。

[0016] 残差卷积块和扩张残差卷积块网络结构是相同的,都是将上一层网络中输出的特征图与其经过两个残差结构输出的特征图进行拼接操作,增加对于不同网络层特征图信息的有效利用。二者区别在于残差卷积块用于实现下采样操作,所以跳跃连接中存在卷积操作进而实现拼接操作,而扩张残差卷积块是恒等残差块,可直接进行拼接操作。

[0017] 上采样模块由卷积块七至九和两次像素重组组成,将下采样模块输出的特征图先经过卷积块七扩张特征图通道,再经过第一次像素重组进行特征图二倍重建,卷积块八继续扩大特征图通道,第二次像素重组进行特征图四倍重建,最后卷积块九输出重建的色度图像。

[0018] 本发明的有益效果如下:

[0019] 1、所申请的方法是基于深度学习的方法,采用的网络结构是生成对抗网络结构,利用生成器和鉴别器之间的博弈特性,加强对图像深层信息的提取,增强着色图像的自然度和真实度,动态提高彩色化图像质量。

[0020] 2、本发明改进了下采样模块的结构,下采样模块使用的是扩张残差卷积网络结构,该结构可以有效减少网络的参数,并且使网络深度变浅,提高网络着色效率。在特征提取过程中,扩张卷积的使用,可以在保持图像分辨率不变的前提下维持感受野,提高网络的特征提取和表达能力;残差结构的使用增加对不同特征图信息的有效利用,加强网络层之间的联系使得网络更具有鲁棒性。

[0021] 3、本发明改进了上采样模块的结构,用卷积块加像素重组替换传统反卷积进行图像重建上色,改善图像重建过程中的细节、颜色信息损失问题。

[0022] 4、本发明去除了鉴别器中的批归一化层,批归一化层会破坏了图像原本的对比度信息,导致不可预测的伪影以及限制模型的泛化能力,去掉批归一化层可以提高网络的表现性能和减少计算量,增加训练的稳定性。同时本发明改进了损失函数,损失函数的设计有助于学习图像更清晰的边缘和更详细的纹理,使得输出图像具有更高的真实性。

附图说明

[0023] 图1为本发明一种基于扩张残差生成对抗网络的红外图像彩色化方法流程图。

[0024] 图2为本发明一种基于扩张残差生成对抗网络的红外图像彩色化方法的生成器网络结构图。

[0025] 图3为本发明生成器网络中残差卷积块和扩张残差卷积块的网络结构图。

[0026] 图4为本发明一种基于扩张残差生成对抗网络的红外图像彩色化方法的鉴别器网络结构图。

具体实施方式

[0027] 下面结合附图对本发明做进一步详细说明。

[0028] 如图1所示,一种基于扩张残差生成对抗网络的红外图像彩色化方法,该方法具体包括如下步骤:

[0029] 步骤1,构建网络模型。

[0030] 整个生成对抗网络包括生成器和鉴别器,生成器根据输入的目标红外图像生成彩色红外图像,鉴别器判断生成的彩色红外图像是否足够真实。

[0031] 如图2所示,所述的生成器由三个模块组成,分别是浅层特征提取模块、下采样模块和上采样模块。将红外图像输入生成器中,红外图像先经过浅层特征提取模块提取图像浅层信息,再经过下采样模块提取图像深层次的语义信息,最后经过上采样模块重建红外图像色度图像。

[0032] 浅层特征提取模块由卷积块一和卷积块二组成,提取图像初级结构信息。下采样模块由残差卷积块一至残差卷积块三、扩张残差卷积块一至二和卷积块三至六组成,残差卷积块一至三完成下采样操作,扩张残差块一至二在保持图像分辨率不变的前提下维持感受野,进一步提取图像语义高级特征,卷积块三至六消除扩张残差卷积导致的网格效应。

[0033] 如图3所示,差卷积块和扩张残差卷积块网络结构是相同的,都是将上一层网络中输出的特征图与其经过两个残差结构输出的特征图进行拼接操作,增加对于不同网络层特征图信息的有效利用。二者区别在于残差卷积块用于实现下采样操作,所以跳跃连接中存在卷积操作进而实现拼接操作,而扩张残差卷积块是恒等残差块,可直接进行拼接操作。

[0034] 采样模块由卷积块七至九和两次像素重组组成,将下采样模块输出的特征图先经过卷积块七扩张特征图通道,再经过第一次像素重组进行特征图二倍重建,卷积块八继续扩大特征图通道,第二次像素重组进行特征图四倍重建,最后卷积块九输出重建的色度图像。

[0035] 如图4所示,鉴别器采用小型VGG网络结构,由6个去掉批归一化的卷积块和两个全连接层组成。生成器输出的色度图像与红外图像融合即是生成的彩色红外图像,将生成的彩色红外图像与数据集中的可见光彩色图像输入鉴别器中,鉴别器输出真假概率信息判断输入图像是否真实。

[0036] 步骤2,训练网络模型。

[0037] 训练红外图像着色模型,对整个生成对抗网络用可见光数据集ImageNet采用生成器和鉴别器进行对抗训练。调整数据集中每个图像的尺寸,固定输入图像的大小,分离出彩色图像的亮度通道L,将其作为整个网络的输入,将彩色图像作为标签,进行有监督的训练。

[0038] 步骤3,最小化损失函数值。

[0039] 生成器采用色度损失和相对平均鉴别器的对抗损失,鉴别器采用相对平均鉴别器的对抗损失,通过最小化生成器和鉴别器的损失函数交替优化生成器和鉴别器,直到生成器和鉴别器的对抗达到平衡状态即可认为模型参数已训练完成,保存模型参数。生成器使用L1范数计算色度损失,通过最小化重建的着色图像与数据集中原始彩色图像之间的色度损失避免着色结果过于跳跃;使用相对平均鉴别器的对抗损失,通过最小化对抗损失提升着色效果的自然度与真实性。鉴别器使用相对平均鉴别器的对抗损失,最小化对抗损失有利于加快着色进程,提升着色效果。

[0040] 步骤4,微调网络模型。用红外图像数据集FLIR对模型进行训练和微调,得到稳定可靠的模型参数。最终使得模型着色的效果更好。

[0041] 步骤5,保存网络模型。将最终确定的模型参数进行固化。

[0042] 如果需要对红外图像进行着色时,直接将图像输入到以上方法得到的网络模型中即可得到最终的彩色红外图像。

[0043] 实施例:

[0044] 一种基于扩张残差生成对抗网络的红外图像彩色化方法,该方法具体包括如下步骤:

[0045] 步骤1,构建网络模型。

[0046] 如图2所示,生成器中,卷积块由卷积层、批归一化层和激活函数组成,卷积核大小为 3×3 ,步长为1,填充为1;如图3所示,残差卷积块一至三中,第一个卷积层步长为2用于实现下采样操作,第二个卷积层步长为1用于提取特征,卷积核大小均为 3×3 ;传统下采样操作以改变图像分辨率来提取更深层次的信息,不利于图像恢复,使用扩张残差卷积块一至二替代传统下采样操作,其扩张率分别为2和4,目的是为了在保持图像分辨率为 28×28 的前提下维持感受野不变,进一步提取图像细节、语义等深层信息,但是扩张残差卷积会导致网格化效应,因此在网络后面添加扩张率为2的卷积块三至四和扩张率为1的卷积块五至六,通过网络学习,消除网格现象。卷积块七至八用于扩张特征图通道,两次像素重组实现上采样,其放大尺度分别为2和4,用于恢复红外图像原始尺寸、细节、色彩,卷积块九用于输出色度图像。

[0047] 总的来说,着色过程是输入红外图像,经过3次步幅为2的下采样操作提取特征,经过2次像素重组重建色度图像,最后将输出的色度图像与红外图像融合成红外彩色图像输入鉴别器。

[0048] 如图4所示,鉴别器中挖掘信息部分是由重复的卷积层、激活函数组成,最后接全连接层和激活函数输出图像真假概率信息。本发明去除了鉴别器中的批归一化层,批归一化层会破坏图像原本的对比度信息,导致不可预测的伪影以及限制模型的泛化能力,去掉批归一化层可以提高网络的表现性能和减少计算量,增加训练的稳定性。鉴别器中卷积层的卷积核大小均为 3×3 。

[0049] 为了保证网络的鲁棒性,保留更多的结构信息,充分提取图像特征,本发明使用两种激活函数,分别为Sigmoid函数和LeakyReLU函数,卷积块九中的激活函数和鉴别器最后一个激活函数使用Sigmoid函数,其余激活函数均为LeakyReLU函数。Sigmoid函数和LeakyReLU函数定义如下所示:

$$[0050] \quad f(x)_{sigmoid} = \frac{1}{1 + e^{-x}}$$

$$[0051] \quad f(x)_{LeakyReLU} = \begin{cases} x, & \text{if } x \geq 0 \\ \frac{x}{a}, & \text{if } x < 0 \end{cases}$$

[0052] 步骤2,训练网络模型。

[0053] 训练红外图像着色模型,对整个生成对抗网络用可见光数据集ImageNet采用生成器和鉴别器进行对抗训练。ImageNet包含1000个场景分类,120万张图片,是一个庞大的数据集,我们通过随机选取其中每种类别若干张图片组合成24000张ImageNet数据集,将这24000张图像尺寸调整为 224×224 作为整个网络的输入。可见光数据集的对抗训练可以确定一组初始化参数,加快后续网络训练进程。

[0054] 步骤3,最小化损失函数值。

[0055] 通过最小化生成器和鉴别器的损失函数交替优化生成器和鉴别器。通过固定生成器参数使鉴别器损失函数最小化更新鉴别器参数,通过固定鉴别器参数使生成器损失函数最小化更新生成器参数,获得最优网络模型。生成器使用L1范数计算色度损失,使用相对平均鉴别器的对抗损失约束色彩真实性。L1色度损失学习每个像素的两个色度值(a,b),定义L1计算公式为:

$$[0056] \quad L_1 = E_{x_i} \|G(x_i) - y_i\|_1$$

[0057] 它是评价生成图像 $G(x_i)$ 和实际图像 y_i 之间的L1距离的内容损失,避免着色结果过于平淡和跳跃。

[0058] 定义生成器的对抗损失 L_G^{Ra} 计算公式为:

$$[0059] \quad L_G^{Ra} = -E_{x_r} [\log(1 - D_{Ra}(x_r, x_f))] - E_{x_f} [\log(D_{Ra}(x_f, x_r))]$$

[0060] 其中, D_{Ra} 表示相对平均鉴别器, $D_{Ra}(x_r, x_f) = \sigma(C(x_r) - E_{x_f}[C(x_f)])$, E_{x_f} 表示小批量中所有假数据取平均值的操作, σ 是Sigmoid函数, $C(x)$ 是非变换鉴别器输出, $x_f = G(x_i)$, x_i 为输入的灰度图像, x_r 为实际的彩色图像, x_f 是生成图像即色度图像与红外图像融合图像。

[0061] 因此生成器的总损失 L_G 为:

$$[0062] \quad L_G = L_G^{Ra} + \gamma L_1$$

[0063] 其中, γ 是平衡不同损失项的系数,控制各个损失项的权重。

[0064] 鉴别器损失函数定义为:

$$[0065] \quad L_D^{Ra} = -E_{x_r} [\log(D_{Ra}(x_r, x_f))] - E_{x_f} [\log(1 - D_{Ra}(x_f, x_r))]$$

[0066] 通过优化生成器和鉴别器损失函数,有助于网络学习更清晰的边缘和更详细的纹理,使得着色图像的颜色自然,真实度更高,视觉效果更好。

[0067] 设定训练次数为100,并设置批处理大小为10,损失函数权重 γ 设置为100,训练过程的初始学习速率为0.0003。网络参数优化器选择自适应矩估计算法,优点主要在于经过偏置校正后,每一次迭代学习率都有个确定范围,使得参数比较平稳。当鉴别器判断假图的

能力与生成器生成图像欺骗过鉴别器的能力平衡时,认为网络已基本训练完成。

[0068] 步骤4,微调网络模型。用红外图像数据集FLIR对模型进行训练和微调,得到稳定可靠的模型参数。最终使得模型着色的效果更好。

[0069] 步骤5,保存网络模型。网络训练完成后,将最终确定的模型参数进行固化。

[0070] 如果需要对红外图像进行着色时,直接将图像输入到以上方法得到的网络模型中即可得到最终的彩色红外图像。该网络模型固定了输入图像的大小为 224×224 。

[0071] 其中,卷积,跳跃连接,拼接操作,残差结构,扩张卷积和像素重组的实现是本领域技术人员公知的算法,具体流程和方法可在相应的教科书或者技术文献中查阅到。

[0072] 本发明通过构建一种基于扩张残差生成对抗网络的红外图像彩色化网络模型,不仅提高了特征利用率,而且可以自动给红外图像着色,避免需要人工辅助着色的繁琐。主观上,本发明的着色图像色彩鲜明,颜色自然生动,很好的改善了颜色溢出、细节损失和缺乏对比度的问题;客观上,通过计算与现有方法得到图像的相关指标,进一步验证了该方法的可行性和优越性。现有技术和本发明提出方法的相关指标对比如表1所示:

[0073] 表1现有技术和本发明提出方法的相关指标对比

	方法	峰值信噪比	结构相似性
[0074]	现有方法	27.36	0.897
	本发明	28.48	0.912

[0075] 从表中可知,本发明提出的方法有更高的峰值信噪比和结构相似性两个指标,进一步说明本发明具有更好的着色图像质量。

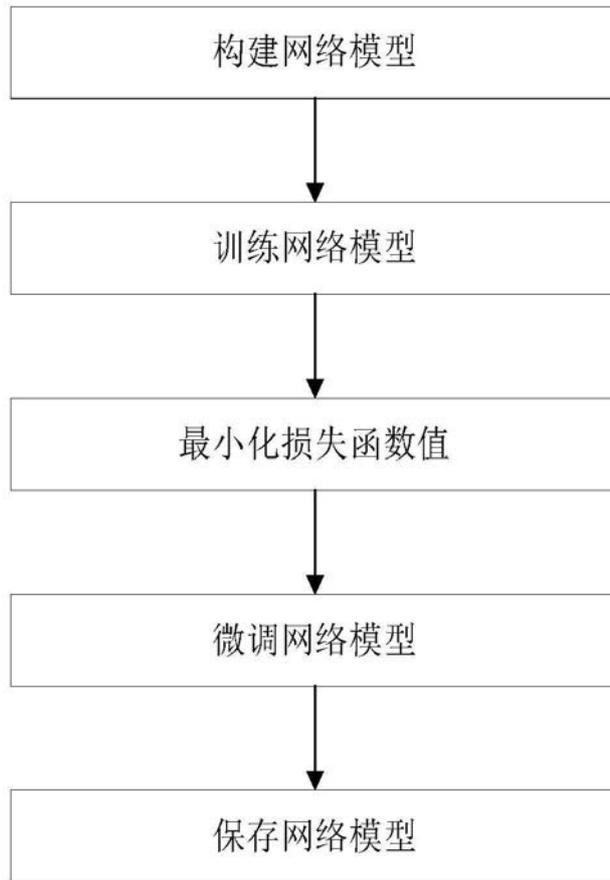


图1

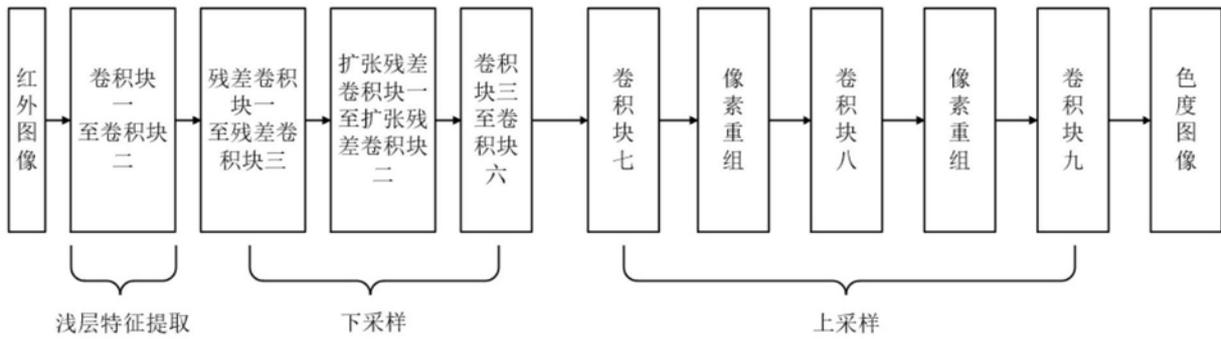


图2

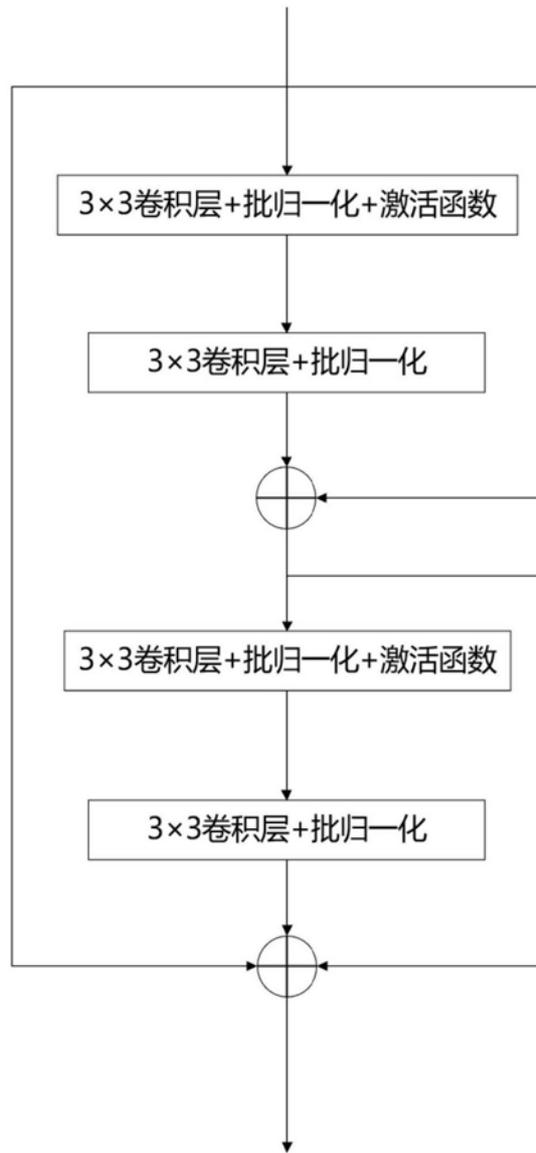


图3

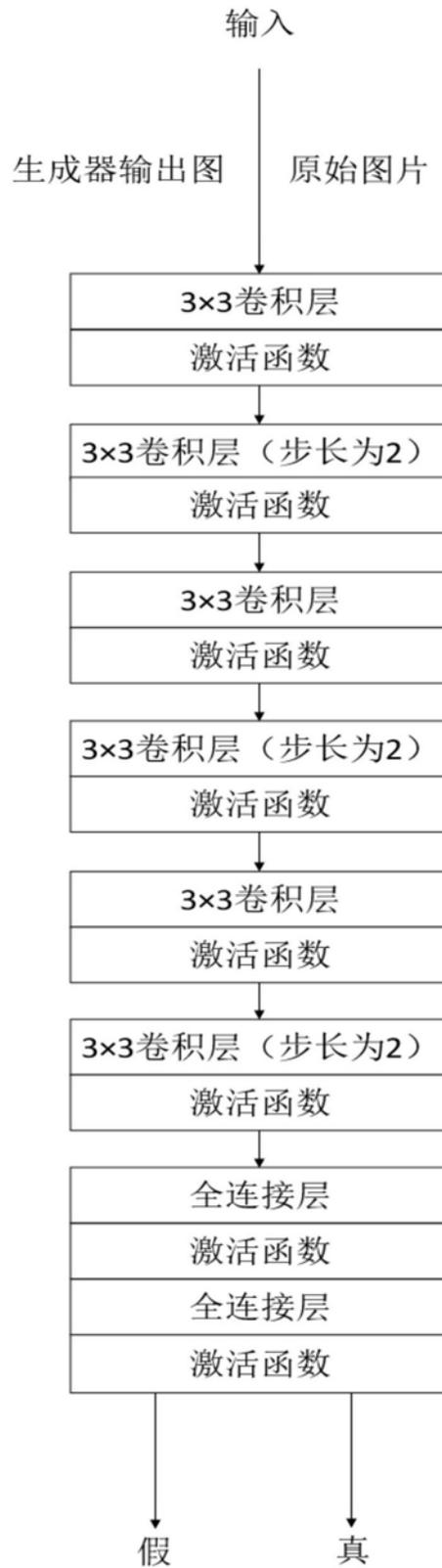


图4