



(12) 发明专利

(10) 授权公告号 CN 110070483 B

(45) 授权公告日 2023. 10. 20

(21) 申请号 201910235651.1

CN 109376582 A, 2019.02.22

(22) 申请日 2019.03.26

CN 108596839 A, 2018.09.28

(65) 同一申请的已公布的文献号

CN 1458791 A, 2003.11.26

申请公布号 CN 110070483 A

WO 2015014131 A1, 2015.02.05

(43) 申请公布日 2019.07.30

JP H08251404 A, 1996.09.27

(73) 专利权人 中山大学

Yifan Liu 等. Auto-painter: Cartoon image generation from sketch by using conditional Wasserstein generative adversarial networks. 《Neurocomputing》. 2018, 第311卷78-87.

地址 510006 广东省广州市海珠区新港西路135号

(72) 发明人 曾坤 区炳坚 周凡

审查员 李想

(51) Int. Cl.

G06T 3/00 (2006.01)

G06V 40/16 (2022.01)

(56) 对比文件

CN 109377448 A, 2019.02.22

CN 107577985 A, 2018.01.12

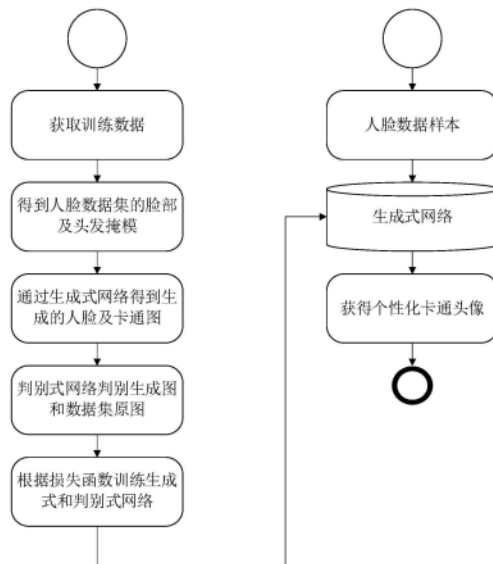
权利要求书3页 说明书6页 附图2页

(54) 发明名称

一种基于生成式对抗网络的人像卡通化方法

(57) 摘要

本发明公开了一种基于生成式对抗网络的人像卡通化方法。本发明通过生成式网络将人脸与背景分割、将人脸转化为卡通的人脸、将背景转化为卡通背景,之后合成得到卡通图,并用判别式网络进行判别;然后通过损失函数对生成式网络与判别式网络进行训练;最后把待处理人脸图像输入训练好的生成式网络即可生成对应的卡通图像。本发明有助于根据输入的人脸图片全自动的生成人像卡通图片,或者根据用户输入人脸图像,给出推荐的卡通化方案,供用户选择或修改,节省用户选取素材拼接的时间。



1. 一种基于生成式对抗网络的人像卡通化方法,其特征在于,所述方法包括:

步骤一,获取人脸数据训练集和卡通数据训练集;

步骤二,对人脸数据训练集进行预处理,得到头发掩模、脸部掩模以及五官掩模;

步骤三,构建生成式网络,将人脸数据训练集中的人脸图像转化为卡通图像,以及将卡通数据训练集中的卡通图像转化为人脸图像;

步骤四,构建判别式网络,对转化而成的卡通图像以及转化而成的人脸图像分别进行判别;

步骤五,根据步骤二生成的掩模、步骤三生成式网络生成的人脸和卡通图像、步骤四得到的判别结果,计算损失函数值并优化生成式网络和判别式网络;

步骤六,重复步骤三到步骤五,循环迭代多轮,得到训练好的卡通图像生成式网络;

步骤七,把待处理的人脸图像输入到最终得到的卡通生成式网络,则可得到对应的具有个人特色的卡通图像;

其中,所述构建生成式网络,将人脸数据训练集中的人脸图像转化为卡通图像,具体为:

输入的人脸图记作 x ,在训练GAN的时候,使用到的卡通图记作 y ; x 作为一个分割网络 G_{lattn} 的输入,该网络用于分割前景,得到前景掩模,其输出 $G_{lattn}(x)$ 是一个掩模,其大小与输入图片大小一致,通道数为1,每个像素取值在0-1之间,越接近0,表示该像素为背景的可能性越高,越接近1,表示该像素是脸或头发的概率越高;从而背景掩模表示为 $1-G_{lattn}(x)$;

同时,人脸图 x 也输入到人脸特征编码网络 e_1 中,得到高层语义信息编码向量 $e_1(x)$;

利用所述高层语义信息编码向量 $e_1(x)$ 作为卡通背景解码网络 d_{lbg} 的输入, d_{lbg} 专注于生成与卡通数据集中的背景相近的背景,而不关注卡通人脸的生成;

利用所述高层语义信息编码向量 $e_1(x)$ 作为卡通人脸解码网络 d_{lcnt} 的输入, d_{lcnt} 专注于生成卡通数据集中的人脸,而不关注卡通数据集的背景生成;

利用从所述 $G_{lattn}(x)$ 得到的人脸及头发掩模以及背景掩模,以及从所述 d_{lbg} 和 d_{lcnt} 得到的卡通背景以及卡通脸的生成结果,得到最终生成的卡通图:

$$G_1(x) = G_{lattn}(x) \odot d_{lcnt}(e_1(x)) + (1 - G_{lattn}(x)) \odot d_{lbg}(e_1(x));$$

\odot 表示逐像素相乘;

其中,所述构建判别式网络,对转化而成的卡通图像以及转化而成的人脸图像分别进行判别,具体为:

将所述生成的卡通图以及从卡通数据训练集里取样的卡通图送入卡通图原图判别网络 D_1 中,该网络用于判断输入的图像是否为非生成的卡通图样本;

训练时,通过减少卡通图判别模型的影响权重,同时加入一个新的卡通图判别网络 D_{1Edg} ,该判别网络 D_{1Edg} 以经过边缘提取网络 $EdgeExtraNet$ 后的卡通边缘图作为输入; $EdgeExtraNet$ 网络由两个卷积层组成,第一个卷积层用于得到灰度图,另一个卷积层用于边缘提取,卷积核为图像处理领域中的边缘算子,卷积核参数如下所示:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix};$$

人脸图像判别器网络由原图判别网络构成,与卡通图判别器不同的是,不需要使用边缘判别网络;

其中,所述计算损失函数值并优化生成式网络和判别式网络,具体为:

计算颜色损失:在人变卡通的过程中,为了保持低层语义信息,将输入图像x与生成的卡通图输入到一个平滑网络SmoothNet中,使用20个参数一样的卷积层组成平滑网络,卷积核参数如下所示:

$$\frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix};$$

根据所述预处理得到的粗略的头发掩模A_{hair},以及除去鼻子,眼睛和嘴巴区域的脸部部分A_{face2},两个区域是结构和颜色不需要发生太大改变的区域,对这两个区域计算损失:

$$L_{\text{color}} = ||\text{SmoothNet}(x) \odot (\text{A}_{\text{hair}} + \text{A}_{\text{face2}}) - \text{SmoothNet}(G_1(x)) \odot (\text{A}_{\text{hair}} + \text{A}_{\text{face2}})||_1;$$

计算循环一致性损失:在人变卡通再还原回人的这个循环中,循环一致性损失表达式如下:

$$L_{\text{cyc1}} = ||x \odot (\text{A}_{\text{face1}} + \text{A}_{\text{face2}}) - G_2(G_1(x)) \odot (\text{A}_{\text{face1}} + \text{A}_{\text{face2}})||_1 + ||\text{SmoothNet}(x) \odot \text{A}_{\text{hair}} - \text{SmoothNet}(G_2(G_1(x))) \odot \text{A}_{\text{hair}}||_1$$

对于卡通输入图y,在卡通转化为人再还原回卡通这一个循环中,沿用CycleGAN的循环一致性损失:

$$L_{\text{cyc2}} = ||y - G_1(G_2(y))||_1;$$

计算掩模循环损失:为使人脸照片的前景区域与生成卡通的前景区域保持一致,掩模循环损失定义如下:

$$L_{\text{cycattn1}} = ||G_1\text{attn}(x) - G_2\text{attn}(G_1(x))||_1;$$

在卡通变人过程中,掩模循环一致性 L_{cycattn2} 与 L_{cycattn1} 定义类似;

计算掩模监督损失:

$$L_{\text{msk_sup1}} = ||G_1\text{attn}(x) - (\text{A}_{\text{hair}} + \text{A}_{\text{face2}} + \text{A}_{\text{face1}})||_1;$$

计算GAN对抗损失:希望生成模型能够生成可以欺骗判别模型的图片,而判别模型能够正确区分哪些图像是生成模型生成的,设X为人脸图像分布,Y为卡通图像分布,优化目标为:

$$\min_{\theta_{G1}} \max_{\theta_{D1}} E_{x \sim X}(\log(1 - D_1(G_1(x)))) + E_{y \sim Y}(\log(D_1(y)))$$

其中 θ_{G1} 表示G1模型中的参数, θ_{D1} 表示模型D1中的参数,对应的由卡通到人脸的GAN对抗损失类似;

计算GAN梯度图对抗损失:在人变卡通的过程中,加入了卡通梯度图判别器,使得生成器模型更加关注卡通的结构而非颜色,其优化目标为:

$$\min_{\theta_{G1}} \max_{\theta_{D1\text{Edg}}} E_{x \sim X}(\log(1 - D_1\text{Edg}(\text{EdgeExtraNet}(G_1(x)))))) \\ + E_{y \sim Y}(\log(D_1\text{Edg}(\text{EdgeExtraNet}(y))))$$

其中 θ_{D1Edg} 为判别模型D1Edg的参数；

最终损失函数值为上述所有损失结果的线性组合,首先固定生成式网络,通过反向传播优化判别器网络,然后固定判别器网络,优化生成式网络。

一种基于生成式对抗网络的人像卡通化方法

技术领域

[0001] 本发明涉及计算机视觉领域,具体涉及一种基于生成式对抗网络的人像卡通化方法。

背景技术

[0002] 人脸卡通化,即将人脸图片作为输入,得到对应的有个人特色的卡通人脸。该技术具有较广的应用场景和意义。近年来,也涌现了一批如“脸萌”、“魔漫相机”等受到广泛欢迎的应用。然而,人脸卡通化是一个涉及到多个图像领域的复杂问题,要快速自动地生成一张具有人肖像特点的卡通画具有一定的挑战性。

[0003] 人脸卡通化难点在于:许多卡通风格要保持发型,脸型等结构,而改变如眼睛、鼻子等结构,同时要除掉像头发、脸部皮肤等细节,并保持总体的色彩。也就是人脸卡通化涉及到了部分结构的改变、细节纹理的去除以及颜色的保留,与传统的风格转化有很大的不同,一些经典的风格转化算法在此问题上无法使用。

[0004] 在人像卡通化领域中,商业的产品主要以素材拼接的方法为主。以“脸萌”为例子,首先需要有画师预先绘制大量的卡通画素材,且素材的摆放位置也是固定好的,需要用户手动选取素材进行拼接,以形成一张卡通画像。因此类似的商业软件并不能做到自动生成人像卡通,且素材固定也限制了其效果,不利于生成与人脸相似的卡通画。

[0005] 在科研领域中,全自动的人像卡通化生成有很大一部分可归类为基于组件的卡通生成方法,其主要是将人脸五官和头发等结构切割,通过匹配素材库中最接近的素材,并进行形变等处理融合成卡通头像。这类方法首先需要准备大量的头发、脸、五官等素材,这个过程需要耗费大量的时间,且需要专业的知识。最终生成的卡通头像较为生硬,素材中的卡通人脸的各部分结构必须与人脸结构相似。

[0006] 近年来生成对抗网络(GAN)在跨域的图像转化中取得了很大的关注,为人像卡通化提供了新的思路。现阶段使用GAN进行人脸卡通化的研究还比较少,并存在不少问题:(1)使用GAN进行卡通化的研究中,多使用标准的CycleGAN或其变种,CycleGAN在将马转化为斑马图片、白天转化为黑夜图片等场景中有不错的效果,但在两个域的结构差异较大的情况下,如人脸变卡通,表现效果不佳。(2)很多使用GAN进行人脸卡通化的研究效果,只能生成一张“像卡通头像”的图片,丢失了个性特征,在发型,脸型以及颜色的保持上效果都很差。

[0007] 现有的一个方法是域转化网络(DTN)。对于两个域的图像(人脸图片及卡通人脸图片),生成模型G包含两个网络:预训练好的编码器f以及解码器g,将人脸图片x送入f,提取得到一个具有输入图像高层语义信息的编码 $f(x)$,这个编码可能包含了如发型、颜色等脸部信息。解码器g负责将前面得到的编码解码为一张卡通图像。判别模型D需要判断输入的图片是否为样本中的卡通图,还是由生成模型生成的卡通图。与传统的GAN不一样的地方,该网络也将来自卡通头像的图片y输入到生成模型中,得到生成的卡通图 $G(y)$,通过损失函数限制 $G(y)$,与输入卡通图y要一致。此外,通过损失函数限制 $f(G(x))$ 要与 $f(x)$ 一致,从而保证根据人脸图片x生成的卡通图 $G(x)$ 保留了x的高层语义特征。

[0008] 这种方法的缺点在于：

[0009] 从DTN的技术方案可以看到，对于两个域的图像，使用同一个编码器 f 去提取图像的高层语义信息，这就限制了两个域的图像的差别不能太大。经实验，当两个域的图像区别较大，也就是卡通的风格更加偏离人的风格的情况下，这种方法便不能取得满意的效果。

[0010] 此外，通过限制输入人脸图和生成的卡通图在高层语义信息上保持一致，是一个很弱的监督，导致生成的图片在颜色、脸型、发型等跟输入的人脸图经常不能保持一致，也就是说个性化效果并不强。

发明内容

[0011] 本发明的目的是克服现有方法的不足，提出了一种基于生成式对抗网络的人像卡通化方法。本发明解决的问题主要有两个：一是仅保持高层语义的一致性，导致生成的卡通头像丢失了个性特征，发型、脸型、发色及肤色等偏低层的语义信息无法很好的保持；二是现有的工作将人脸分割、人脸转化为卡通的人脸、背景转化为卡通图中的背景三个子任务用同一个网络处理，导致生成的卡通图质量不佳。

[0012] 为了解决上述问题，本发明提出了一种基于生成式对抗网络的人像卡通化方法，所述方法包括：

[0013] 步骤一，获取人脸数据训练集和卡通数据训练集；

[0014] 步骤二，对人脸数据训练集进行预处理，得到头发掩模、脸部掩模以及五官掩模；

[0015] 步骤三，构建生成式网络，将人脸数据训练集中的人脸图像转化为卡通图像，以及将卡通数据训练集中的卡通图像转化为人脸图像；

[0016] 步骤四，构建判别式网络，对转化而成的卡通图像以及转化而成的人脸图像分别进行判别；

[0017] 步骤五，根据步骤二生成的掩模、步骤三生成式网络生成的人脸和卡通图像、步骤四得到的判别结果，计算损失函数值并优化生成式网络和判别式网络；

[0018] 步骤六，重复步骤三到步骤五，循环迭代多轮，得到训练好的卡通图像生成式网络；

[0019] 步骤七，把待处理的人脸图像输入到最终得到的卡通生成式网络，则可得到对应的具有个人特色的卡通图像。

[0020] 优选地，所述将人脸数据训练集中的人脸图像转化为卡通图像的步骤，具体包括：

[0021] 使用一个分割网络获取前景掩模；

[0022] 使用一个编码网络对人脸图像进行编码；

[0023] 使用背景解码网络对人脸图像编码进行解码，得到人脸图像背景；

[0024] 使用前景解码网络对人脸图像编码进行解码，得到人脸图像前景；

[0025] 使用前景掩模、人脸图像背景、人脸图像前景，得到生成的卡通图。

[0026] 优选地，所述对转化而成的卡通图像进行判别的步骤，具体包括：

[0027] 将生成的卡通图以及从卡通数据训练集里取样的卡通图送入卡通图原图判别网络，该网络用于判断输入的图像是否为非生成的卡通图样本；

[0028] 获取生成卡通图的边缘图以及获取卡通数据训练集里取样的卡通图的边缘图，使用边缘判别式网络对边缘图进行判别。

[0029] 本发明提出的一种基于生成式对抗网络的人像卡通化方法,解决现有的使用GAN进行人脸卡通化工作中普遍存在的问题:底层语义信息无法保持以及生成卡通头像效果欠佳。有助于根据输入的人脸图片全自动的生成人像卡通图片,或者根据用户输入人脸图像,给出推荐的卡通化方案,供用户选择或修改,节省用户选取素材拼接的时间。

附图说明

[0030] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例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,下面描述中的附图仅仅是本发明的一些实施例,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它的附图。

[0031] 图1是本发明实施例的人像卡通化方法总体流程图;

[0032] 图2是本发明实施例的生成式网络的结构图;

[0033] 图3是本发明实施例的判别式网络的结构图。

具体实施方式

[0034] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅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,本领域普通技术人员在没有作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。

[0035] 图1是本发明实施例的人像卡通化方法总体流程图,如图1所示,该方法包括:

[0036] S1,获取人脸数据训练集和卡通数据训练集;

[0037] S2,对人脸数据训练集进行预处理,得到头发掩模、脸部掩模以及五官掩模;

[0038] S3,构建生成式网络,将人脸数据训练集中的人脸图像转化为卡通图像,以及将卡通数据训练集中的卡通图像转化为人脸图像;

[0039] S4,构建判别式网络,对转化而成的卡通图像以及转化而成的人脸图像分别进行判别;

[0040] S5,根据S2生成的掩模、S3生成的人脸和卡通图像、S4得到的判别结果,计算损失函数值并优化生成式网络和判别式网络;

[0041] S6,重复S3到S5,循环迭代多轮,得到训练好的卡通图像生成式网络;

[0042] S7,把待处理的人脸图像输入到最终得到的卡通生成式网络,则可得到对应的具有个人特色的卡通图像。

[0043] 步骤S1,具体如下:

[0044] 从公开的对齐人脸数据集CelebA中获取对齐的人脸数据构建人脸训练集,通过网络爬虫或者使用公开数据集的方式获取卡通头像图片作为卡通训练集,本实施例中使用的是谷歌公开的卡通数据集。

[0045] 步骤S2,具体如下:

[0046] S2-1:将输入的人脸图片输入预先训练好的语义分割网络,得到粗糙的头发和脸的掩模。本实施例中的预训练网络结构使用的是谷歌提出的语义分割模型,训练数据来自一个人部件剖析数据集;

[0047] S2-2:使用主动形状模型(Active Shape Model,ASM),对输入的人脸进行人脸特征点提取,并分别计算眉毛,眼睛,鼻子,嘴巴特征点的凸包,作为它们的掩模;

[0048] S2-3:由于许多卡通风格,相比于真实人脸,在眼睛、鼻子、嘴巴等位置会出现较大的形变,因此计算眼睛、鼻子以及嘴巴这些特征点的一个凸包,得到一个脸部结构变化区域,记作Aface1,根据S2-1得到的粗糙的脸部掩模,将该区域中除掉Aface1的区域记作Aface2,Aface2通常是结构不需要发生改变及颜色需要保持的部分。最后记S21中得到的头发区域为Ahair。

[0049] 步骤S3,具体如下:

[0050] 将人脸转化为个性化卡通图像的生成式网络,以及将卡通图像转化为人脸的生成式网络,两个网络结构是对称的。下面主要描述将人脸图像转化为卡通图像的生成流程,网络结构如图2所示,从输入的人脸图像x得到输出的个性化卡通图像,需要经过以下步骤:

[0051] S3-1:输入的人脸图,记作x,在训练GAN的时候,会使用到卡通图,记作y。x作为一个分割网络G1attn的输入,该网络用于分割前景,得到前景掩模,其输出G1attn(x)是一个掩模,其大小与输入图片大小一致,通道数为1,每个像素取值在0-1之间,越接近0,表示该像素为背景的可能性越高,越接近1,表示该像素是脸或头发的概率越高。从而背景掩模则可表示为 $1-G1attn(x)$;

[0052] S3-2:同时,人脸输入图x也会输入到一个人脸特征编码网络中(记作e1),得到高层语义信息编码向量 $e1(x)$;

[0053] S3-3:利用S3-2中得到的高层语义编码向量作为卡通背景解码网络(记作d1bg)的输入,我们希望d1bg专注于生成与卡通数据集中的背景相近的背景,而不用关注卡通人脸的生成;

[0054] S3-4:利用S3-2中得到的高层语义编码向量作为卡通人脸解码网络(记作d1cnt)的输入,我们希望d1cnt专注于生成与卡通数据集中的人脸,而不用关注卡通数据集的背景生成;

[0055] S3-5:利用S3-1中得到的人脸及头发掩模以及背景掩模,以及S3-3和S3-4得到的具有较好的卡通背景以及卡通脸的生成结果,得到最终生成的卡通图:

[0056] $G1(x) = G1attn(x) \odot d1cnt(e1(x)) + (1-G1attn(x)) \odot d1bg(e1(x))$

[0057] 其中, \odot 表示逐像素相乘;

[0058] S3-6:从卡通到人脸的判别模型结构类似,生成模型G2包含分割网络G2attn,编码网络e2,背景解码网络d2bg,人脸解码网络d2cnt,这里不再赘述。

[0059] 步骤S4,如图3所示,具体如下:

[0060] S4-1:将S3-5中得到的生成卡通图以及从卡通数据训练集里取样的卡通图送入卡通图原图判别网络D1中,该网络用于判断输入的图像是否为非生成的卡通图样本。

[0061] S4-2:此外,生成的卡通的颜色也受到卡通数据集颜色分布的影响,生成式网络很难生成卡通数据集中没有的颜色,这是因为判别网络会根据卡通数据集的颜色分布对图片进行分类,导致生成网络生成的颜色跟卡通数据集的颜色接近。因此,在训练的时候,通过减少卡通图判别模型的影响权重,同时加入一个新的卡通图判别网络D1Edg,这个判别网络以经过边缘提取网络(EdgeExtraNet)后的卡通边缘图作为输入。本实施例中,EdgeExtraNet网络由两个卷积层组成,第一个卷积层用于得到灰度图,另一个卷积层用于

边缘提取,卷积核为图像处理领域中的边缘算子,卷积核参数如下所示:

$$[0062] \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

[0063] S4-3:人脸图像判别器网络由原图判别网络构成,与卡通图判别器不同的是,这里不需要使用边缘判别网络。

[0064] 步骤S5,具体如下:

[0065] S5-1:计算颜色损失。在人边卡通的过程中,为了保持发色,肤色等低层语义信息,将输入图像x与S3-5中得到的生成卡通图输入到一个平滑网络(SmoothNet)中,在本发明中简单的使用了20个参数一样的卷积层组成平滑网络,卷积核参数如下所示:

$$[0066] \quad \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

[0067] 根据S-2中的数据预处理模块,我们已经得到一个粗略的头发掩模A_{hair},以及除去鼻子,眼睛和嘴巴区域的脸部部分A_{face2},两个区域是结构和颜色不需要发生太大改变的区域,对这两个区域计算损失:

$$[0068] \quad L_{color} = ||\text{SmoothNet}(x) \odot (A_{hair}+A_{face2}) - \text{SmoothNet}(G1(x)) \odot (A_{hair}+A_{face2}) ||_1$$

[0069] S5-2:计算循环一致性损失,参考CycleGAN,希望经过G1得到的生成卡通图G1(x),经过卡通到人的转化G2后,能还原回x本身。在人脸变卡通这种应用中,在人变卡通,也就是x转化为G1(x)的时候,通常来说会丢失不少细节信息,特别是头发比较多以及背景部分比较大的人脸照片,在转化为卡通图的时候头发的纹理常常会丢失,人脸照片背景转化后也常常会丢失很多信息。因此再用G2(x)还原回人脸照片的时候,这些信息几乎是不可能也是没必要完全恢复的,为此,在人变卡通再还原回人的这个循环中,本发明的循环一致性损失表达式如下:

$$[0070] \quad L_{cyc1} = ||x \odot (A_{face1}+A_{face2}) - G2(G1(x)) \odot (A_{face1}+A_{face2}) ||_1 + ||\text{SmoothNet}(x) \odot A_{hair} - \text{SmoothNet}(G2(G1(x))) \odot A_{hair} ||_1$$

[0071] 对于卡通输入图y,在人变卡通再还原回人的这个循环中,我们更加关注人脸部的恢复以及头发颜色的恢复。而在卡通转化为人再还原回卡通这一个循环中,由于卡通转化为人通常是添加一些背景信息和头发纹理等信息,人脸再还原会卡通则只需要再把这些添加的信息删掉,因此可以沿用CycleGAN的循环一致性损失:

$$[0072] \quad L_{cyc2} = ||y - G1(G2(y)) ||_1$$

[0073] S5-3:计算掩模循环损失,我们希望在人变卡通过程中,人脸照片的前景区域与生成卡通的前景区域保持一致,因此掩模循环损失定义如下:

$$[0074] \quad L_{cycattn1} = ||G1attn(x) - G2attn(G1(x)) ||_1$$

[0075] 在卡通变人过程中,掩模循环一致性 $L_{cycattn2}$ 与 $L_{cycattn1}$ 定义类似。

[0076] S5-4:计算掩模监督损失,由于我们有预处理模块S2得到的粗糙掩模,其结果可以为掩模提供一定程度的监督:

$$[0077] \quad L_{\text{masksup1}} = ||G1attn(x) - (A\text{hair}+A\text{face2}+A\text{face1})||_1$$

[0078] 同样的,若有方法可以对卡通数据集的前后景进行语义分割,也可以对其加入掩模监督损失,帮助分割网络更好地生成正确的掩模。

[0079] S5-5:计算GAN对抗损失,我们希望生成模型能够生成可以欺骗判别模型的图片,而判别模型能够正确区分哪些图像是生成模型生成的,设X为人脸图像分布,Y为卡通图像分布,优化目标为:

$$[0080] \quad \min_{\theta_{G1}} \max_{\theta_{D1}} E_{x \sim X}(\log(1 - D1(G1(x)))) + E_{y \sim Y}(\log(D1(y)))$$

[0081] 其中 θ_{G1} 表示G1模型中的参数, θ_{D1} 表示模型D1中的参数,对应的由卡通到人脸的GAN对抗损失类似。

[0082] S5-6:计算GAN梯度图对抗损失,在人变卡通的过程中,本发明加入了卡通梯度图判别器,使得生成器模型更加关注卡通的结构而非颜色,其优化目标为:

$$[0083] \quad \min_{\theta_{G1}} \max_{\theta_{D1Edg}} E_{x \sim X}(\log(1 - D1Edg(EdgeExtraNet(G1(x)))))) \\ + E_{y \sim Y}(\log(D1Edg(EdgeExtraNet(y))))$$

[0084] 其中 θ_{D1Edg} 为判别模型D1Edg的参数。

[0085] S5-7:最终损失函数值为S5-1到S5-7的结果的线性组合,首先固定生成式网络,通过反向传播优化判别器网络,然后固定判别器网络,优化生成式网络。

[0086] 步骤S6,具体如下:

[0087] 重复S3-S5,在本实施例中,循环迭代200轮,得到训练好的卡通图像生成式网络。

[0088] 本发明实施例提出的一种基于生成式对抗网络的人像卡通化方法,解决现有的使用GAN进行人脸卡通化工作中普遍存在的问题:底层语义信息无法保持以及生成卡通头像效果欠佳。有助于根据输入的人脸图片全自动的生成人像卡通图片,或者根据用户输入人脸图像,给出推荐的卡通化方案,供用户选择或修改,节省用户选取素材拼接的时间。

[0089] 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成,该程序可以存储于一计算机可读存储介质中,存储介质可以包括:只读存储器(ROM,Read Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁盘或光盘等。

[0090] 另外,以上对本发明实施例所提供的一种基于生成式对抗网络的人像卡通化方法进行了详细介绍,本文中应用了具体个例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

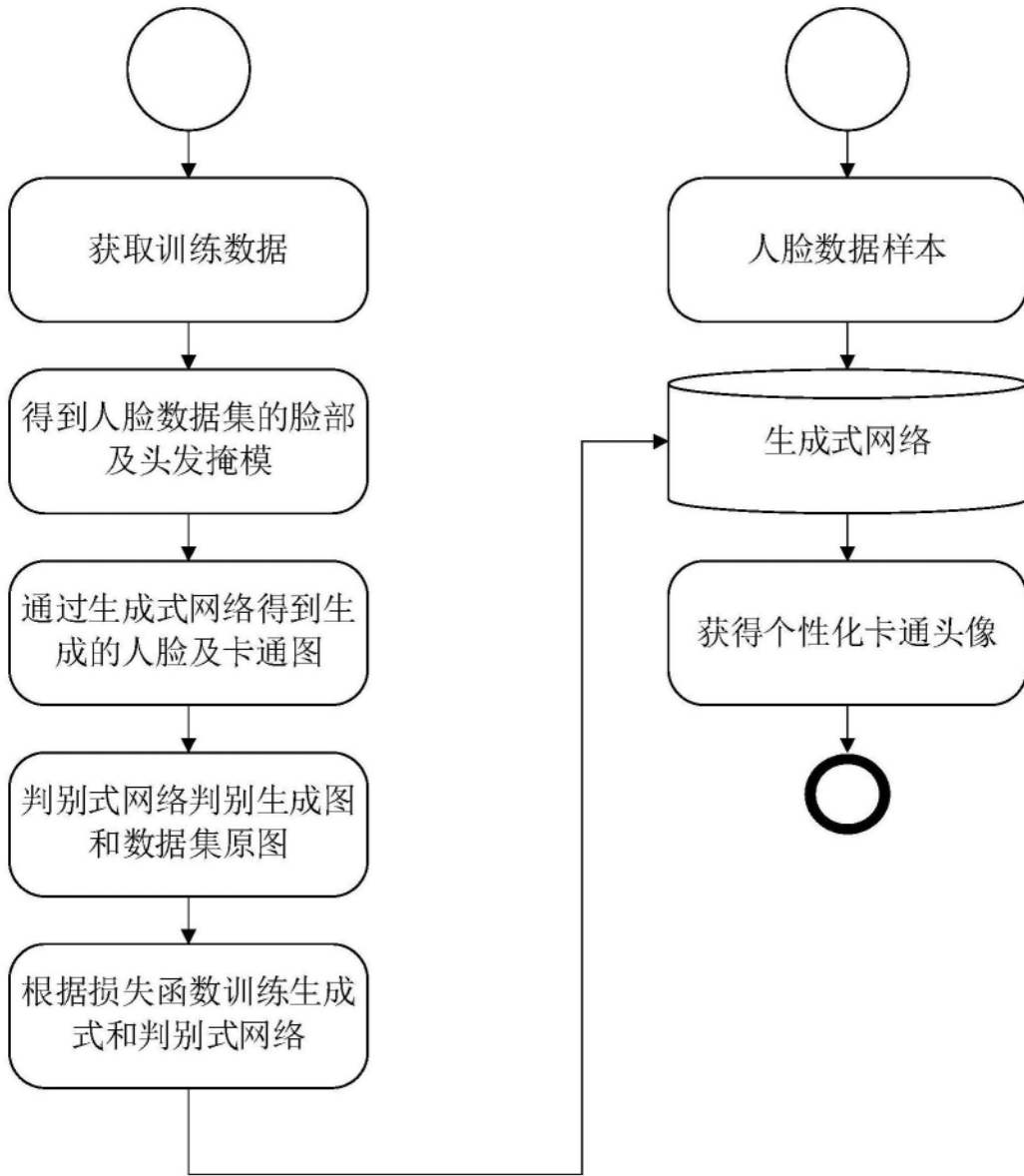


图1

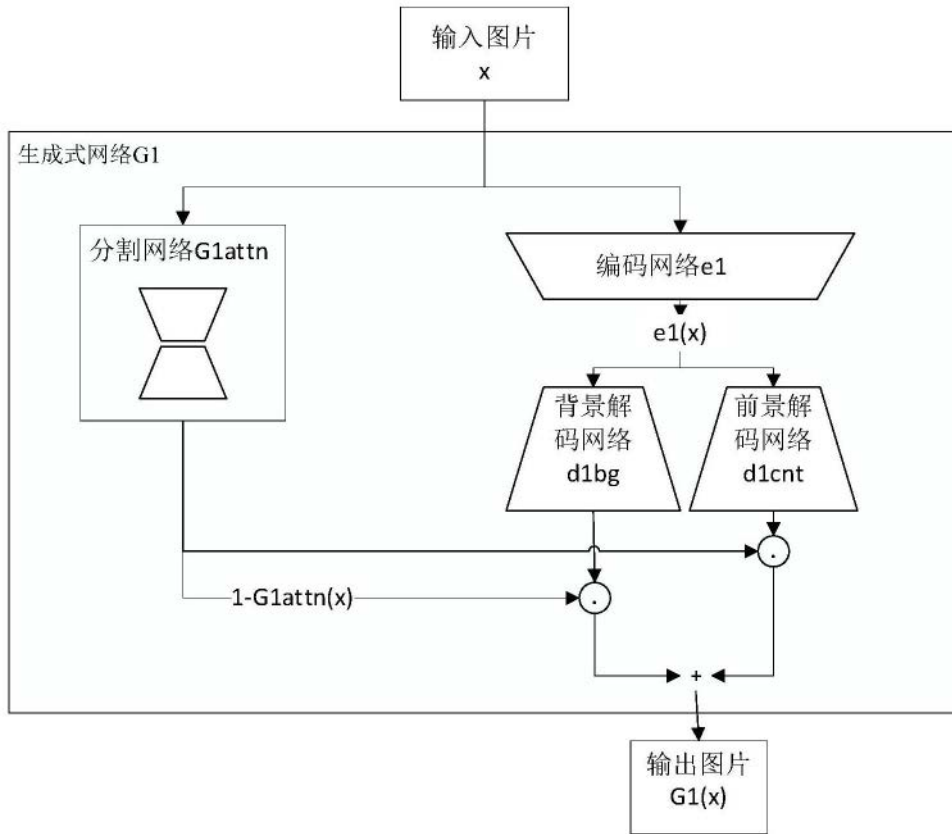


图2

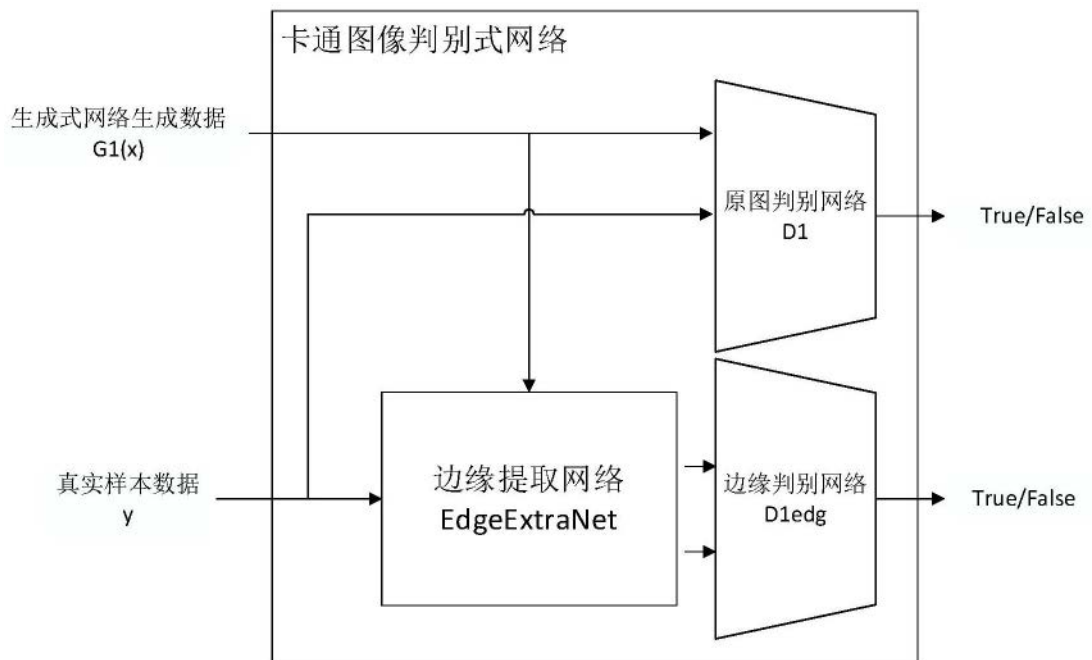


图3