



(12) 发明专利

(10) 授权公告号 CN 113129409 B

(45) 授权公告日 2023. 07. 25

(21) 申请号 202110512222.1

(22) 申请日 2021.05.11

(65) 同一申请的已公布的文献号
申请公布号 CN 113129409 A

(43) 申请公布日 2021.07.16

(66) 本国优先权数据
202110485802.6 2021.04.30 CN

(73) 专利权人 华南农业大学
地址 510642 广东省广州市天河区五山路
483号

(72) 发明人 崔金荣 练俊健 刘海龙 黄诚
钟浩伟

(74) 专利代理机构 广州粤高专利商标代理有限公司 44102

专利代理师 刘俊

(51) Int.Cl.

G06T 11/40 (2006.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 111862270 A, 2020.10.30

CN 109544662 A, 2019.03.29

CN 110223359 A, 2019.09.10

US 2010085372 A1, 2010.04.08

李媛等. 注意力引导的交互式工笔花卉上色算法. 模式识别与人工智能. 2020, 第33卷 (第07期), 第575-587段.

审查员 李健壮

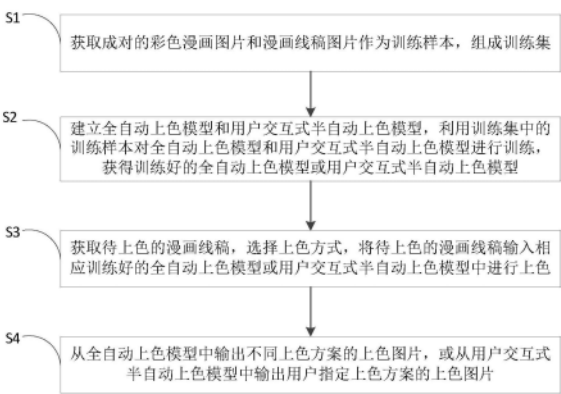
权利要求书3页 说明书8页 附图4页

(54) 发明名称

一种基于深度学习的漫画线稿上色方法

(57) 摘要

本发明提供一种基于深度学习的漫画线稿上色方法, 包括: 获取成对的彩色漫画图片和漫画线稿图片作为训练样本; 建立全自动上色模型和用户交互式半自动上色模型, 利用训练集中的训练样本对全自动上色模型和用户交互式半自动上色模型进行训练, 获得训练好的全自动上色模型或用户交互式半自动上色模型; 选择上色方式, 将待上色的漫画线稿输入训练好的全自动上色模型或用户交互式半自动上色模型中, 从全自动上色模型中输出不同上色方案的上色图片, 或从用户交互式半自动上色模型中输出用户指定上色方案的上色图片。本发明能够实现输入一张待上色的漫画线稿, 输出多张不同上色方案的上色图片, 上色效率高, 上色效果好, 并且还能输出用户指定上色方案的上色图片。



1.一种基于深度学习的漫画线稿上色方法,其特征在于,包括以下步骤:

S1:获取成对的彩色漫画图片和漫画线稿图片作为训练样本,组成训练集;

S2:建立全自动上色模型和用户交互式半自动上色模型,利用训练集中的训练样本对全自动上色模型和用户交互式半自动上色模型进行训练,获得训练好的全自动上色模型或用户交互式半自动上色模型;具体的:

获得训练好的全自动上色模型的方法为:

建立全自动上色模型,包括第一特征编码器、第一空间自适应归一化模块、第一生成器、第一判别器和第二判别器;

将漫画线稿图片输入至第一特征编码器中进行特征提取,获得漫画线稿图片的特征图;

将特征图输入第一空间自适应归一化模块进行归一化,获得归一化的特征图;

将归一化的特征图输入第一生成器中,并向第一生成器输入随机隐变量,获得上色图片;

将上色图片和漫画线稿图片对应的彩色漫画图片输入第一判别器和第二判别器中,设置全自动上色模型损失函数,改变随机隐变量,对全自动上色模型进行训练,直到全自动上色模型损失函数收敛,获得训练好的全自动上色模型;

获得训练好的用户交互式半自动上色模型的方法为:

建立用户交互式半自动上色模型,包括第二特征编码器、第二空间自适应归一化模块、第二生成器、第三判别器、第四判别器和预训练VGG16网络模块;

将漫画线稿图片输入至第二特征编码器中进行特征提取,获得漫画线稿图片的特征图;

将特征图输入第二空间自适应归一化模块进行归一化,获得归一化的特征图;

将归一化的特征图输入第二生成器中,并向第二生成器输入模拟颜色提示,获得上色图片;

将上色图片和漫画线稿图片对应的彩色漫画图片输入第三判别器、第四判别器和预训练VGG16网络模块中,设置用户交互式半自动上色模型损失函数,输入不同的模拟颜色提示,对用户交互式半自动上色模型进行训练,直到用户交互式半自动上色模型收敛,获得训练好的用户交互式半自动上色模型;

获得归一化的特征图的方法为:

$$\gamma_{n,c,i,j}^u \frac{h_{n,c,i,j}^u - \mu_{n,c}^u}{\sigma_{n,c}^u} + \beta_{n,c,i,j}^u$$

其中,c表示输入特征图的通道维度,n表示batchsize的维度,u表示不同的空间自适应归一化模块,i和j分别表示特征图中元素的横纵坐标; $h_{n,c,i,j}^u$ 表示归一化处理前的输入, $\mu_{n,c}^u$ 表示特征图的均值, $\sigma_{n,c}^u$ 表示特征图的标准差; $\gamma_{n,c,i,j}^u$ 和 $\beta_{n,c,i,j}^u$ 表示空间自适应归一化模块的学习参数;

S3:获取待上色的漫画线稿,选择上色方式,将待上色的漫画线稿输入相应训练好的全自动上色模型或用户交互式半自动上色模型中进行上色;

S4:从全自动上色模型中输出不同上色方案的上色图片,或从用户交互式半自动上色

模型中输出用户指定上色方案的上色图片。

2. 根据权利要求1所述的基于深度学习的漫画线稿上色方法,其特征在于,所述步骤S1中,漫画线稿图片为合成模拟漫画线稿图片,获得合成模拟漫画线稿图片的方法:

使用边界检测器XDoG对彩色漫画图片进行线稿提取,通过改变边界检测器XDoG的参数,获得不同级别线宽的漫画线稿图片;对漫画线稿图片进行位图-矢量-位图处理,统一线条后得到合成模拟漫画线稿图片。

3. 根据权利要求1所述的基于深度学习的漫画线稿上色方法,其特征在于,所述全自动上色模型损失函数为:

$$\min_{G_1} \max_{D_1, D_2} \sum L_{GAN}(G_1, D_k)$$

式中:

$$L_{GAN}(G_1, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_1}(G_1)$$

$$L_{adv, D_k}(D_k) = E_{x \sim p_A} [(D_k(x) + 1)^2] + E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) \right)^2 \right]$$

$$L_{adv, G_1}(G_1) = E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) - 1 \right)^2 \right]$$

其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_1 表示随机隐变量, $z_1 \sim p_z$ 表示随机隐变量 z_1 采样自标准高斯分布; G_1 表示第一生成器, D_k 表示第 k 判别器, $k=1, 2$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G_1}(G_1)$ 表示第一生成器对抗损失函数, E (*) 表示分布函数的期望值。

4. 根据权利要求1所述的基于深度学习的漫画线稿上色方法,其特征在于,获得所述模拟颜色提示的方法为:

将彩色漫画图片转化为灰度图;使用二值化方法提取灰度图中图像的区域二值图 mask;

生成随机矩阵 S 和随机值 α , 将随机矩阵 S 中小于 α 的元素设置为 1, 大于 α 的元素设置为 0;

将随机矩阵 S 与区域二值图 mask 相乘, 获得采样矩阵 K , 并将采样矩阵 K 扩展到 3 通道与彩色漫画图片相乘, 获得模拟颜色提示。

5. 根据权利要求1所述的基于深度学习的漫画线稿上色方法,其特征在于,所述用户交互式半自动上色模型损失函数为:

$$\min_{G_2} \max_{D_3, D_4} \sum L_{GAN}(G_2, D_k)$$

式中:

$$L_{GAN}(G_2, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_2}(G_2) + \lambda_1 L_{cont}(G_2)$$

$$L_{adv, D_k}(D_k) = E_{x \sim p_A} [(D_k(x) + 1)^2] + E_{z_2 \sim p_z, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) \right)^2 \right]$$

$$L_{adv, G_2}(G_2) = E_{z_2 \sim p_z, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) - 1 \right)^2 \right]$$

$$L_{cont} = \frac{1}{hwc} \|\phi(x) - \phi(G_2(y, z_2))\|_1$$

其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_2 表示模拟颜色提示, $z_2 \sim p_z$ 表示模拟颜色提示 z_2 采样自标准高斯分布; G_2 表示第二生成器, D_k 表示第 k 判别器, $k=3, 4$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G_2}(G_2)$ 表示第二生成器对抗损失函数, L_{cont} 表示感知损失函数, $E(*)$ 表示分布函数的期望值; h, w, c 分别代表预训练 VGG16 网络模块的高、宽、通道数。

6. 根据权利要求1所述的基于深度学习的漫画线稿上色方法, 其特征在于, 所述将上色图片和漫画线稿图片对应的彩色漫画图片输入第二判别器或第四判别器之前, 还需要对上色图片和彩色漫画图片进行下采样。

7. 根据权利要求1所述的基于深度学习的漫画线稿上色方法, 其特征在于, 对全自动上色模型和用户交互式半自动上色模型进行训练时, 采用Adam梯度下降算法进行训练。

一种基于深度学习的漫画线稿上色方法

技术领域

[0001] 本发明涉及图像处理的技术领域,更具体地,涉及一种基于深度学习的漫画线稿上色方法。

背景技术

[0002] 在人工智能的许多领域,深度网络已经取得了远远超过传统方法的表现,包括语音、自然语言、视觉、游戏等领域。在许多任务中,经典的机器学习方法无法与深度学习比较。过去对于漫画自动上色的方法大多基于Pix2PixGAN模型,该模型架构一个重要的缺点在于:在不加入颜色提示的情况下,对于一个线稿输入,只能对应唯一的输出。对于漫画线稿上色任务来说,漫画的上色策略应该是多样化的,对于同一个漫画线稿,我们可以选择不同的配色策略(例如不同的发色、服装颜色、肤色等)来对线稿进行着色。Paintschainer是现有基于深度学习的漫画线稿上色应用,先后发表了3个模型来实现漫画线稿上色,分别是Tanpopo、Satsuki和Canna,分别具有不同的上色风格。这三种模型都具有不同的缺点: Tanpopo模型存在明显的噪声, Satsuki模型在颜色交汇的位置会产生“光晕”现象, Canna提供了一个粉色的上色风格,上色位置、颜色不精准,在某些颜色上存在严重的色差,存在颜色溢出、分界模糊、不自然的问题。

[0003] 2019年3月29日公开的中国专利CN109544662A中提供了一种基于SRUnet的动漫风格线稿上色方法及系统。首先建立由激励层和残差U网组成的动漫风格线稿上色网络模型SRUnet;然后采用训练集中的图片对所述SRUnet模型进行训练,获得SRUnet生成的动漫风格彩色图像;根据实际彩色图像和所述动漫风格彩色图像确定SRUnet的总损失值;根据所述总损失值优化SRUnet的网络参数,得到训练后的SRUnet,采用所述训练后的SRUnet即可对待上色的黑白线稿图进行上色。该方法输入一张待上色的黑白线稿图,只能输出一张上色图片,上色效率低,上色效果不好。

发明内容

[0004] 本发明为克服上述现有上色方法输出唯一上色图片上色效率低并且上色效果差的缺陷,提供一种基于深度学习的漫画线稿上色方法,实现了全自动上色时输出不同上色方案的上色图片,用户交互式半自动上色时输出用户指定上色方案的上色图片,上色效率高并且上色图片具有优秀的上色效果。

[0005] 为解决上述技术问题,本发明的技术方案如下:

[0006] 本发明提供一种基于深度学习的漫画线稿上色方法,所述方法包括以下步骤:

[0007] S1:获取成对的彩色漫画图片和漫画线稿图片作为训练样本,组成训练集;

[0008] S2:建立全自动上色模型和用户交互式半自动上色模型,利用训练集中的训练样本对全自动上色模型和用户交互式半自动上色模型进行训练,获得训练好的全自动上色模型或用户交互式半自动上色模型;

[0009] S3:获取待上色的漫画线稿,选择上色方式,将待上色的漫画线稿输入相应训练好

的全自动上色模型或用户交互式半自动上色模型中进行上色；

[0010] S4:从全自动上色模型中输出不同上色方案的上色图片,或从用户交互式半自动上色模型中输出用户指定上色方案的上色图片。

[0011] 优选地,所述步骤S1中,漫画线稿图片为合成模拟漫画线稿图片,获得合成模拟漫画线稿图片的方法:

[0012] 使用边界检测器XDoG对彩色漫画图片进行线稿提取,通过改变边界检测器XDoG的参数,获得不同级别线宽的漫画线稿图片;对漫画线稿图片进行位图-矢量-位图处理,统一线条后得到合成模拟漫画线稿图片。手工漫画线稿通常不具有固定的风格与线条样式,使用其来训练模型,很容易会导致模型过拟合,所以采用模拟漫画线稿图片作为训练集。

[0013] 优选地,所述获得训练好的全自动上色模型的方法为:

[0014] 建立全自动上色模型,包括第一特征编码器、第一空间自适应归一化模块、第一生成器、第一判别器和第二判别器;

[0015] 将漫画线稿图片输入至第一特征编码器中进行特征提取,获得漫画线稿图片的特征图;

[0016] 将特征图输入第一空间自适应归一化模块进行归一化,获得归一化的特征图;

[0017] 将归一化的特征图输入第一生成器中,并向第一生成器输入随机隐变量,获得上色图片;

[0018] 将上色图片和漫画线稿图片对应的彩色漫画图片输入第一判别器和第二判别器中,设置全自动上色模型损失函数,改变随机隐变量,对全自动上色模型进行训练,直到全自动上色模型损失函数收敛,获得训练好的全自动上色模型。

[0019] 优选地,所述全自动上色模型损失函数为:

$$[0020] \min_{G_1} \max_{D_1, D_2} \sum L_{GAN}(G_1, D_k)$$

[0021] 式中:

$$[0022] L_{GAN}(G_1, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_1}(G_1)$$

$$[0023] L_{adv, D_k}(D_k) = E_{x \sim p_A} [(D_k(x) + 1)^2] + E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) \right)^2 \right]$$

$$[0024] L_{adv, G_1}(G_1) = E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) - 1 \right)^2 \right]$$

[0025] 其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_1 表示随机隐变量, $z_1 \sim p_z$ 表示随机隐变量 z_1 采样自标准高斯分布; G_1 表示第一生成器, D_k 表示第 k 判别器, $k=1, 2$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G}(G)$ 表示第一生成器对抗损失函数, $E(*)$ 表示分布函数的期望值。

[0026] 优选地,所述获得训练好的用户交互式半自动上色模型的方法为:

[0027] 建立用户交互式半自动上色模型,包括第二特征编码器、第二空间自适应归一化模块、第二生成器、第三判别器、第四判别器和预训练VGG16网络模块;

[0028] 将漫画线稿图片输入至第二特征编码器中进行特征提取,获得漫画线稿图片的特征图;

[0029] 将特征图输入第二空间自适应归一化模块进行归一化,获得归一化的特征图;

[0030] 将归一化的特征图输入第二生成器中,并向第二生成器输入模拟颜色提示,获得上色图片;

[0031] 将上色图片和漫画线稿图片对应的彩色漫画图片输入第三判别器、第四判别器和预训练VGG16网络模块中,设置用户交互式半自动上色模型损失函数,输入不同的模拟颜色提示,对用户交互式半自动上色模型进行训练,直到用户交互式半自动上色模型收敛,获得训练好的用户交互式半自动上色模型。

[0032] 优选地,获得所述模拟颜色提示的方法为:

[0033] 将彩色漫画图片转化为灰度图;使用二值化方法提取灰度图中图像的区域二值图mask;

[0034] 生成随机矩阵S和随机值 α ,将随机矩阵S中小于 α 的元素设置为1,大于 α 的元素设置为0;

[0035] 将随机矩阵S与区域二值图mask相乘,获得采样矩阵K,并将采样矩阵K扩展到3通道与彩色漫画图片相乘,获得模拟颜色提示。

[0036] 优选地,用户交互式半自动上色模型损失函数为:

$$[0037] \min_{G_2} \max_{D_3, D_4} \sum L_{GAN}(G_2, D_k)$$

[0038] 式中:

$$[0039] L_{GAN}(G_2, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_2}(G_2) + \lambda_1 L_{cont}(G_2)$$

$$[0040] L_{adv, D_k}(D_k) = E_{x \sim p_A} [(D_k(x) + 1)^2] + E_{z_2 \sim p_{z_2}, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) \right)^2 \right]$$

$$[0041] L_{adv, G_2}(G_2) = E_{z_2 \sim p_{z_2}, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) - 1 \right)^2 \right]$$

$$[0042] L_{cont} = \frac{1}{hwc} \|\phi(x) - \phi(G_2(y, z_2))\|_1$$

[0043] 其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_2 表示模拟颜色提示, $z_2 \sim p_{z_2}$ 表示模拟颜色提示 z_2 采样自标准高斯分布; G_2 表示第二生成器, D_k 表示第 k 判别器, $k=3,4$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G_2}(G_2)$ 表示第二生成器对抗损失函数, L_{cont} 表示感知损失函数, $E(*)$ 表示分布函数的期望值; h, w, c 分别代表预训练VGG16网络模块的高、宽、通道数。

[0044] 优选地,获得归一化的特征图的方法为:

$$[0045] \gamma_{n,c,i,j}^u \frac{h_{n,c,i,j}^u - \mu_{n,c}^u}{\sigma_{n,c}^u} + \beta_{n,c,i,j}^u$$

[0046] 其中, c 表示输入特征图的通道维度, n 表示batchsize的维度, u 表示不同的空间自适应归一化模块, i 和 j 分别表示特征图中元素的横纵坐标; $h_{n,c,i,j}^u$ 表示归一化处理前的输入, $\mu_{n,c}^u$ 表示特征图的均值, $\sigma_{n,c}^u$ 表示特征图的标准差; $\gamma_{n,c,i,j}^u$ 和 $\beta_{n,c,i,j}^u$ 表示空间自适应归一化模块的学习参数。

[0047] 优选地,所述将上色图片和漫画线稿图片对应的彩色漫画图片输入第二判别器或第四判别器之前,还需要对上色图片和彩色漫画图片进行下采样。

[0048] 优选地,对全自动上色模型和用户交互式半自动上色模型进行训练时,采用Adam梯度下降算法进行训练。

[0049] 与现有技术相比,本发明技术方案的有益效果是:

[0050] 本发明通过获取成对的彩色漫画图片和漫画线稿图片作为训练样本,对建立的全自动上色模型和用户交互式半自动上色模型进行训练,获得训练好的全自动上色模型或用户交互式半自动上色模型;选择上色方式,将待上色的漫画线稿输入训练好的全自动上色模型中输出不同上色方案的上色图片,或输入训练好的用户交互式半自动上色模型中输出用户指定上色方案的上色图片。本发明能够实现输入一张待上色的漫画线稿,输出多张不同上色方案的上色图片,上色效率高,上色效果好,并且还能输出用户指定上色方案的上色图片。

附图说明

[0051] 图1为实施例所述的基于深度学习的漫画线稿上色方法的流程图;

[0052] 图2为实施例所述的合成模拟漫画线稿图片的示意图;

[0053] 图3为实施例所述的全自动上色模型的数据流向图;

[0054] 图4为实施例所述的用户交互式半自动上色模型的数据流向图;

[0055] 图5为实施例所述的获得模拟颜色提示的数据流向图

[0056] 图6为实施例所述的全自动上色模型的上色效果图;

[0057] 图7为实施例所述的用户交互式半自动上色模型的上色效果图。

具体实施方式

[0058] 附图仅用于示例性说明,不能理解为对本专利的限制;

[0059] 为了更好说明本实施例,附图某些部件会有省略、放大或缩小,并不代表实际产品的尺寸;

[0060] 对于本领域技术人员来说,附图中某些公知结构及其说明可能省略是可以理解的。

[0061] 下面结合附图和实施例对本发明的技术方案做进一步的说明。

[0062] 实施例

[0063] 本实施例提供一种基于深度学习的漫画线稿上色方法,如图1所示所述方法包括以下步骤:

[0064] S1:获取成对的彩色漫画图片和漫画线稿图片作为训练样本,组成训练集;

[0065] 所述漫画线稿图片为合成模拟漫画线稿图片,获得合成模拟漫画线稿图片的方法:

[0066] 使用边界检测器XDoG对彩色漫画图片进行线稿提取,通过改变边界检测器XDoG的参数,获得不同级别线宽的漫画线稿图片;对不同级别线宽的漫画线稿图片进行位图-矢量-位图处理,统一线条后得到合成模拟漫画线稿图片;

[0067] 手工线稿通常不具有固定的风格与线条样式,使用手工线稿来训练模型,很容易

会导致模型对于手工线稿形式的过拟合。为了缓解这种现象,我们使用多种不同的方法合成多种不同样式的线稿图片作为训练集。具体地,应用边界检测器XDoG以生成更多具有不同样式的漫画线稿图片,为了模拟艺术家绘制的线稿图,我们将XDoG算法的参数设置为 $\varphi = 1 \times 10^9$,以使草图线条的边界处保持逐步过渡;将其他默认值设置为 $\tau = 0.95, \kappa = 4.5$,参数 σ 设置为0.3/0.4/0.5,以获得不同级别的线宽,对不同级别线宽的漫画线稿图片进行位图-矢量-位图处理,统一线条后得到合成模拟漫画线稿图片。如图2所示,分别是使用XDoG算法,参数 σ 设定为0.3/0.4/0.5提取的漫画线稿图片。

[0068] S2:建立全自动上色模型和用户交互式半自动上色模型,利用训练集中的训练样本对全自动上色模型和用户交互式半自动上色模型进行训练,获得训练好的全自动上色模型或用户交互式半自动上色模型;训练时,采用Adam梯度下降算法进行训练;

[0069] 建立全自动上色模型并训练的具体方法为:建立全自动上色模型,包括第一特征编码器、第一空间自适应归一化模块、第一生成器、第一判别器和第二判别器;

[0070] 如图3所示,将漫画线稿图片输入至第一特征编码器中进行特征提取,获得漫画线稿图片的特征图;第一特征编码器由Conv2d-InstanceNormal-LeakyRelu($\alpha = 0.2$)模块组成;

[0071] 将特征图输入第一空间自适应归一化模块进行归一化,获得归一化的特征图;获得归一化的特征图的方法为:

$$[0072] \quad \gamma_{n,c,i,j}^u \frac{h_{n,c,i,j}^u - \mu_{n,c}^u}{\sigma_{n,c}^u} + \beta_{n,c,i,j}^u$$

[0073] 其中,c表示输入特征图的通道维度,n表示入batchsize的维度,u表示不同的空间自适应归一化模块,i和j分别表示特征图中元素的横纵坐标; $h_{n,c,i,j}^u$ 表示归一化处理前的输入, $\mu_{n,c}^u$ 表示特征图的均值, $\sigma_{n,c}^u$ 表示特征图的标准差; $\gamma_{n,c,i,j}^u$ 和 $\beta_{n,c,i,j}^u$ 表示空间自适应归一化模块的学习参数;

[0074] 将归一化的特征图输入第一生成器中,并向第一生成器输入随机隐变量,获得上色图片;

[0075] 将上色图片和漫画线稿图片对应的彩色漫画图片输入第一判别器和第二判别器中,设置全自动上色模型损失函数,改变随机隐变量,对全自动上色模型进行训练,直到全自动上色模型损失函数收敛,获得训练好的全自动上色模型;

[0076] 随机隐变量是一个256维服从标准高斯分布的随机变量,对生成器输入随机隐变量的目的是控制上色图片的生成;

[0077] 使用两个不同尺度的PatchGAN判别器,在不同的尺度下对生成图片进行评价,获得更加准确的评价结果。第一判别器和第二判别器均由Conv2d-InstanceNomalization-Relu Block组成,卷积层步长均为2;第一判别器中输入的上色图片和彩色漫画图片为256*256*3,第二判别器中输入的上色图片和彩色漫画图片需要下采样为128*128*3。

[0078] 所述全自动上色模型损失函数为:

$$[0079] \quad \min_{G_1} \max_{D_1, D_2} \sum L_{GAN}(G_1, D_k)$$

[0080] 式中:

$$[0081] \quad L_{GAN}(G_1, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_1}(G_1)$$

$$[0082] \quad L_{adv, D_k}(D_k) = E_{x \sim p_A}[(D_k(x) + 1)^2] + E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) \right)^2 \right]$$

$$[0083] \quad L_{adv, G_1}(G_1) = E_{z_1 \sim p_z, y \sim p_B} \left[\left(D_k(G_1(y, z_1)) - 1 \right)^2 \right]$$

[0084] 其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_1 表示随机隐变量, $z_1 \sim p_z$ 表示随机隐变量 z_1 采样自标准高斯分布; G_1 表示第一生成器, D_k 表示第 k 判别器, $k=1, 2$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G_1}(G)$ 表示第一生成器对抗损失函数, $E(*)$ 表示分布函数的期望值。

[0085] 建立用户交互式半自动上色模型并训练的具体方法为:

[0086] 建立用户交互式半自动上色模型, 包括第二特征编码器、第二空间自适应归一化模块、第二生成器、第三判别器、第四判别器和预训练VGG16网络模块;

[0087] 如图4所示, 将漫画线稿图片输入至第二特征编码器中进行特征提取, 获得漫画线稿图片的特征图; 第二特征编码器由Conv2d-InstanceNormal-LeakyRelu ($\alpha=0.2$) 模块组成;

[0088] 将特征图输入第二空间自适应归一化模块进行归一化, 获得归一化的特征图; 获得归一化的特征图的方法与全自动上色模型中的方法相同;

[0089] 将归一化的特征图输入第二生成器中, 并向第二生成器输入模拟颜色提示, 模拟颜色提示与归一化的特征图在通道C维度上连接, 获得上色图片; 如图5所示, 获得所述模拟颜色提示的方法为:

[0090] 将彩色漫画图片转化为灰度图; 使用二值化方法提取灰度图中图像的区域二值图mask, 彩色漫画图片均为纯白背景, 所以二值化阈值取250, 即图片 >255 的值设置为1, ≤ 255 的值设置为0;

[0091] 生成服从均匀分布值域为0~1的256*256维随机矩阵S和取值范围为0~0.002随机值 α , 将随机矩阵S中小于 α 的元素设置为1, 大于 α 的元素设置为0;

[0092] 将随机矩阵S与区域二值图mask相乘, 获得采样矩阵K, 并将采样矩阵K扩展到3通道与彩色漫画图片相乘, 获得256*256*3的模拟用户颜色提示矩阵, 作为模拟颜色提示。

[0093] 将上色图片和漫画线稿图片对应的彩色漫画图片输入第三判别器、第四判别器和预训练VGG16网络模块中, 设置用户交互式半自动上色模型损失函数, 输入不同的模拟颜色提示, 对用户交互式半自动上色模型进行训练, 直到用户交互式半自动上色模型收敛, 获得训练好的用户交互式半自动上色模型。

[0094] 使用两个不同尺度的PatchGAN判别器, 在不同的尺度下对生成图片进行评价, 获得更加准确的评价结果。第三判别器和第四判别器均由Conv2d-InstanceNormalization-Relu Block组成, 卷积层步长均为2; 第三判别器中输入的上色图片和彩色漫画图片为256*256*3, 第四判别器中输入的上色图片和彩色漫画图片需要下采样为128*128*3。

[0095] 所述用户交互式半自动上色模型损失函数为:

$$[0096] \quad \min_{G_2} \max_{D_3, D_4} \sum L_{GAN}(G_2, D_k)$$

[0097] 式中:

$$[0098] \quad L_{GAN}(G_2, D_k) = L_{adv, D_k}(D_k) + L_{adv, G_2}(G_2) + \lambda_1 L_{cont}(G_2)$$

$$[0099] \quad L_{adv, D_k}(D_k) = E_{x \sim p_A} [(D_k(x) + 1)^2] + E_{z_2 \sim p_z, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) \right)^2 \right]$$

$$[0100] \quad L_{adv, G_2}(G_2) = E_{z_2 \sim p_z, y \sim p_B} \left[\left(D_k(G_2(y, z_2)) - 1 \right)^2 \right]$$

$$[0101] \quad L_{cont} = \frac{1}{hwc} \|\phi(x) - \phi(G_2(y, z_2))\|_1$$

[0102] 其中, p_A 表示彩色漫画图片数据分布, p_B 表示漫画线稿图片数据分布, x, y 表示采样的图片, $x \sim p_A$ 表示 x 采样自彩色漫画图片, $y \sim p_B$ 表示 y 采样自漫画线稿图片, z_2 表示模拟颜色提示, $z_2 \sim p_z$ 表示模拟颜色提示 z_2 采样自标准高斯分布; G_2 表示第二生成器, D_k 表示第 k 判别器, $k=3, 4$; $L_{adv, D_k}(D_k)$ 表示第 k 判别器对抗损失函数, $L_{adv, G_2}(G_2)$ 表示第二生成器对抗损失函数, L_{cont} 表示感知损失函数, $E(*)$ 表示分布函数的期望值; h, w, c 分别代表预训练 VGG16 网络模块的高、宽、通道数。

[0103] S3: 获取待上色的漫画线稿, 选择上色方式, 将待上色的漫画线稿输入相应训练好的全自动上色模型或用户交互式半自动上色模型中进行上色;

[0104] S4: 从全自动上色模型中输出不同上色方案的上色图片, 或从用户交互式半自动上色模型中输出用户指定上色方案的上色图片。

[0105] 在实际应用过程中, 使用 Anime sketch colorization Pair 数据集中的彩色漫画图片和漫画线稿图片作为训练集, 该数据集的图片全部都是漫画人物图片, 均不存在背景。选用 tensorflow 2.3.0 框架建立全自动上色模型和用户交互式半自动上色模型, 其中全自动上色模型和用户交互式半自动上色模型可以共用一套特征编码器、空间自适应归一化模块、生成器和两个不同尺度的判别器; 训练过程中, 采用 Adam 优化器, 利用梯度下降算法对全自动上色模型和用户交互式半自动上色模型进行训练, 生成器和判别器的学习率分别设置为 0.0001 和 0.0004, 每批大小设置为 16。

[0106] 对于全自动上色, 将彩色漫画图片和漫画线稿图片输入全自动上色模型中, 并输入随机隐变量, 在训练集上训练 130 轮后, 全自动上色模型损失函数收敛, 获得训练好的全自动上色模型, 模型会表现出多模态合成的性质; 对于用户交互式半自动上色, 将彩色漫画图片和漫画线稿图片输入用户交互式半自动上色模型中, 并输入模拟颜色提示, 在训练集上训练 130 轮后, 用户交互式半自动上色模型损失函数收敛, 获得训练好的用户交互式半自动上色模型;

[0107] 训练好后, 获取待上色的漫画线稿, 选择上色方式, 将待上色的漫画线稿输入相应的训练好的上色模型中。选择全自动上色时, 将待上色的漫画线稿输入训练好的全自动上色模型, 改变输入的随机隐变量的数值, 如图 6 所示, 获得不同的上色图片, 还可以获得从一种颜色到另一种颜色过渡的一组上色图片; 选择用户交互式上色时, 将待上色的漫画线稿输入训练好的用户交互式半自动上色模型中, 并且模拟颜色提示通过交互软件以用户手动打点的形式生成后输入模型中, 如图 7 所示, 生成与用户选择的模拟颜色提示相同颜色的上

色图片。

[0108] FID用于评估真实图像和生成图像的特征向量之间的距离。较小的FID分数表示两组图像的分布更紧密。我们使用FID将我们的全自动上色模型和用户交互式半自动上色模型与相同数据集下传统的上色的模型AutoPainter和AlacGAN进行比较,评估上色效果。FID分数如下表所示:

[0109]

model	val set without hint	val set with hints	Real Line Art without hint
AutoPainter	47.70	47.55	125.12
AlacGAN	31.67	31.33	68.36
Ours(auto-color)	28.80	\	62.88
Ours(user-guide)	34.14	17.85	35.81

[0110] 从表中可以看出,第二列是基于没有模拟颜色提示的合成模拟漫画线稿图片,第三列是基于有模拟颜色提示的合成模拟漫画线稿图片,第四列是基于没有模拟颜色提示的手工线稿图片;全自动上色模型(Ours(auto-color))在没有模拟颜色提示的情况下,FID分数小于传统的上色的模型AutoPainter和AlacGAN的FID分数;用户交互式半自动上色模型(Ours(user-guide))在有模拟颜色提示的情况下,FID分数小于传统的上色的模型AutoPainter和AlacGAN的FID分数。综上所述,本实施例提供的基于深度学习的漫画线稿上色方法具有优秀的上色效果。

[0111] 显然,本发明的上述实施例仅仅是为清楚地说明本发明所作的举例,而并非是对本发明的实施方式的限定。对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其它不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本发明权利要求的保护范围之内。

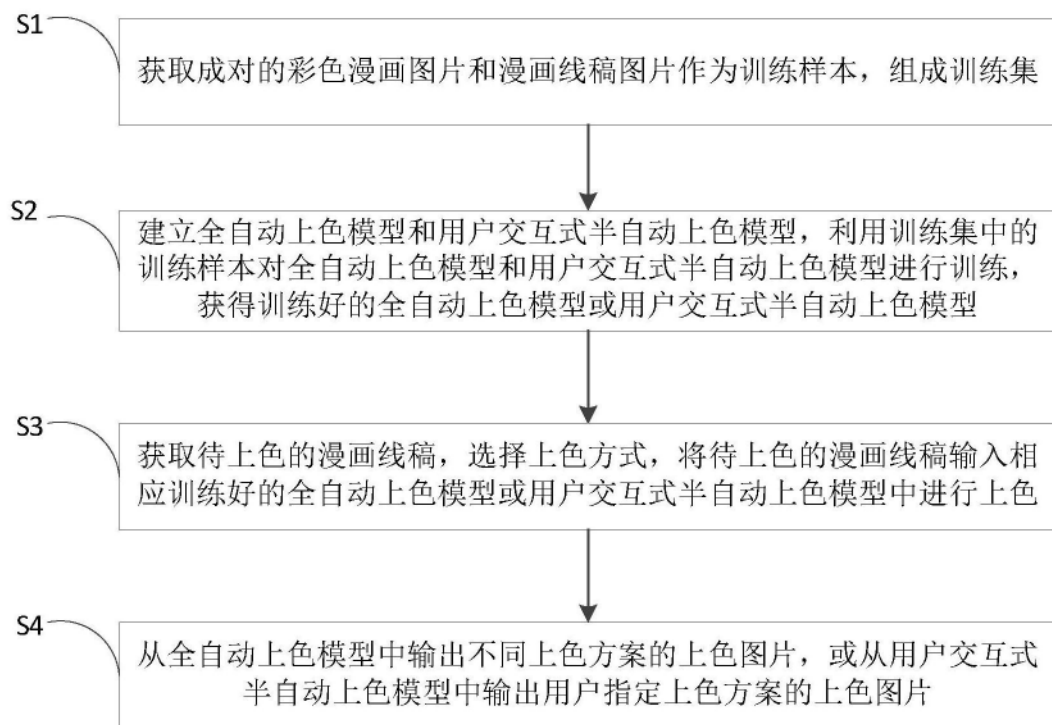


图1



图2

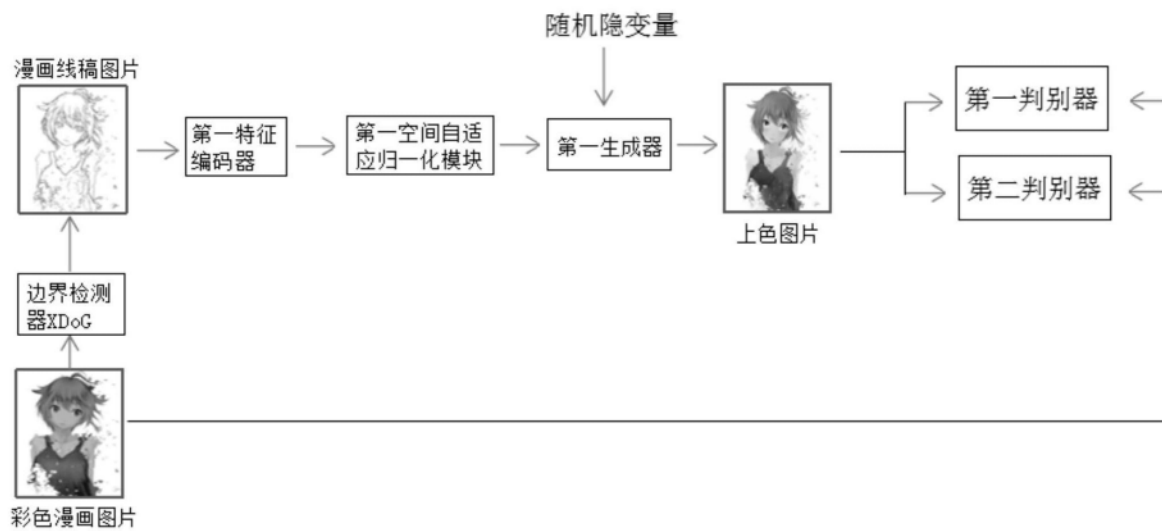


图3

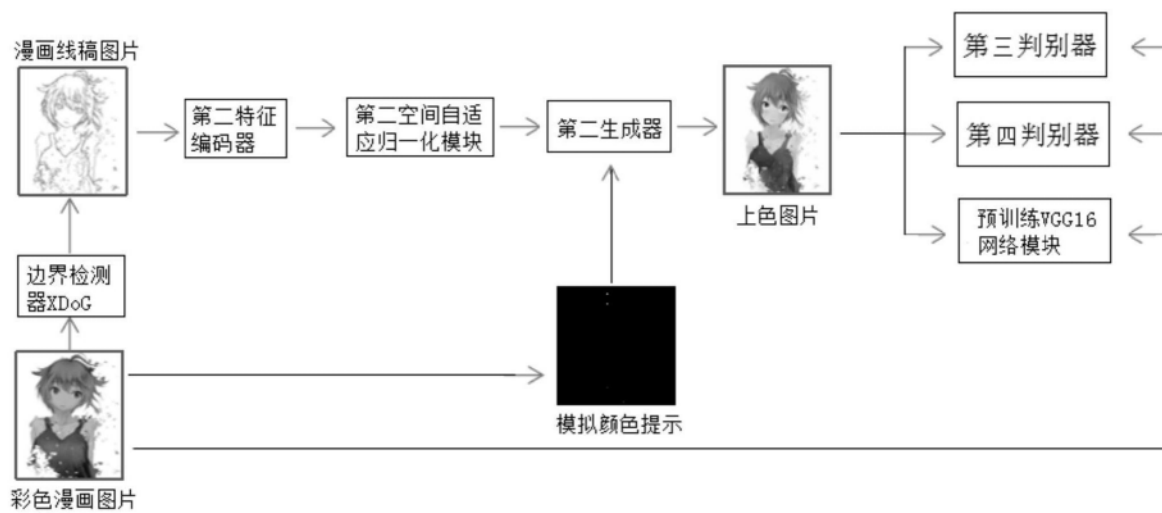


图4

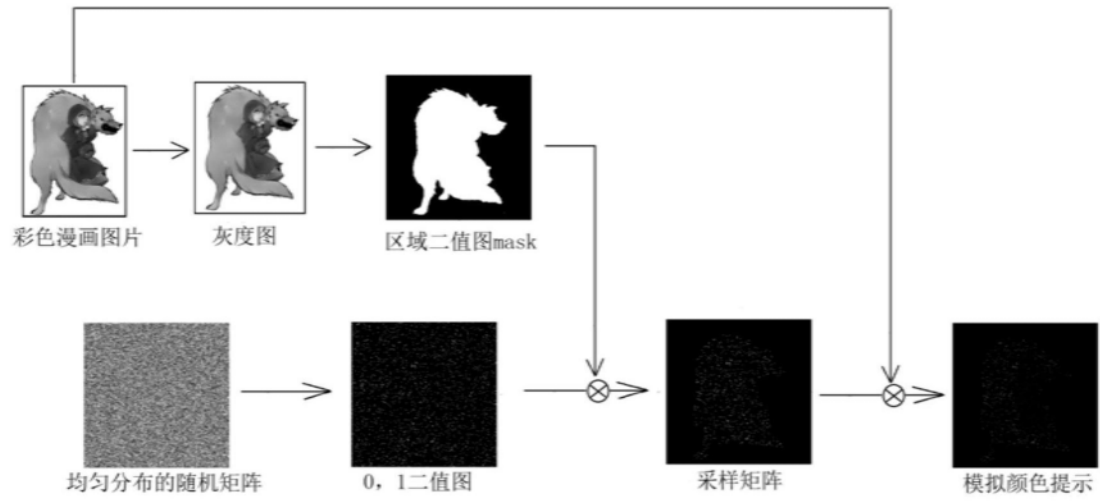


图5



图6

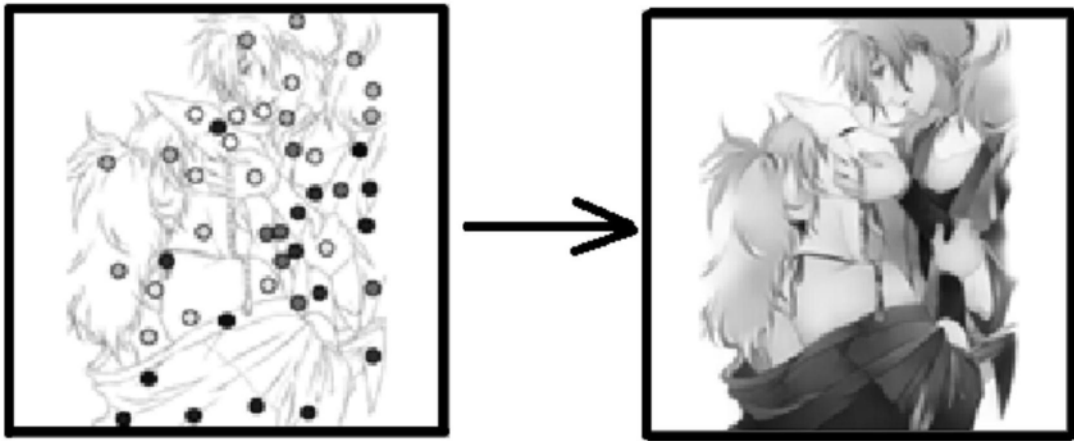


图7