

图像生成模型的训练方法、生成方法、装置及设备

本申请要求于 2021 年 02 月 02 日提交的、申请号为 202110143927.0、发明名称为“图像生成模型的训练方法、生成方法、装置及设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请实施例涉及机器学习和图像处理技术领域，特别涉及一种图像生成模型的训练方法、生成方法、装置及设备。

背景技术

目前，机器学习技术越来越广泛地应用于图像处理技术领域，例如，基于机器学习技术变换图像的风格。

在相关技术中，先采集大量的训练样本对，每一训练样本对包括真实人脸图像与对应的人脸漫画图像；再采用大量的训练样本对，直接对学习模型进行大量的迭代训练，训练完成的学习模型即为图像生成模型。

发明内容

本申请实施例提供了一种图像生成模型的训练方法、生成方法、装置及设备。所述技术方案如下：

根据本申请实施例的一个方面，提供了一种图像生成模型的训练方法，所述方法包括：

训练得到第一变换模型，所述第一变换模型用于基于第一噪声样本生成第一训练图像，所述第一训练图像为第一类风格的图像；

基于所述第一变换模型，训练得到重构模型，所述重构模型用于将原始图像样本映射为所述原始图像样本对应的隐变量；

训练得到第二变换模型，所述第二变换模型用于基于第二噪声样本生成第二训练图像，所述第二训练图像为第二类风格的图像；

将所述第一变换模型与所述第二变换模型进行嫁接，生成嫁接后的变换模型；

基于所述重构模型和所述嫁接后的变换模型，生成图像生成模型，所述图像生成模型用于将所述第一类风格的待变换图像，变换为所述第二类风格的目标图像。

根据本申请实施例的一个方面，提供了一种图像生成方法，所述方法包括：

将第一类风格的待变换图像输入重构模型，生成所述待变换图像对应的隐变量；

基于所述待变换图像对应的隐变量，采用嫁接后的变换模型生成所述待变换图像对应的目标图像，所述目标图像为第二类风格的图像；

其中，所述嫁接后的变换模型是通过将第一变换模型与第二变换模型进行嫁接生成的模型；所述第一变换模型用于基于第一噪声样本生成所述第一类风格的图像；所述第二变换模型用于基于第二噪声样本生成所述第二类风格的图像。

根据本申请实施例的一个方面，提供了一种图像生成模型的训练装置，所述装置包括：

模型训练模块，用于训练得到第一变换模型，所述第一变换模型用于基于第一噪声样本生成第一训练图像，所述第一训练图像为第一类风格的图像；

所述模型训练模块，还用于基于所述第一变换模型，训练得到重构模型，所述重构模型用于将原始图像样本映射为所述原始图像样本对应的隐变量；

所述模型训练模块，还用于训练得到第二变换模型，所述第二变换模型用于基于第二噪

声样本生成第二训练图像，所述第二训练图像为第二类风格的图像；

模型生成模块，用于将所述第一变换模型与所述第二变换模型进行嫁接，生成嫁接后的变换模型；

所述模型生成模块，还用于基于所述重构模型和所述嫁接后的变换模型，生成图像生成模型，所述图像生成模型用于将所述第一类风格的待变换图像，变换为所述第二类风格的目标图像。

根据本申请实施例的一个方面，提供了一种图像生成装置，所述装置包括：

变量生成模块，用于将第一类风格的待变换图像输入重构模型，生成所述待变换图像对应的隐变量；

图像生成模块，用于基于所述待变换图像对应的隐变量，采用嫁接后的变换模型生成所述待变换图像对应的目标图像，所述目标图像为第二类风格的图像；

其中，所述嫁接后的变换模型是通过将第一变换模型与第二变换模型进行嫁接生成的模型；所述第一变换模型用于基于第一噪声样本生成所述第一类风格的图像；所述第二变换模型用于基于第二噪声样本生成所述第二类风格的图像。

根据本申请实施例的一个方面，提供了一种计算机设备，所述计算机设备包括处理器和存储器，所述存储器中存储有计算机程序，所述计算机程序由所述处理器加载并执行以实现上述图像生成模型的训练方法，或实现上述图像生成方法。

根据本申请实施例的一个方面，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中存储有计算机程序，所述计算机程序由处理器加载并执行以实现上述图像生成模型的训练方法，或实现上述图像生成方法。

根据本申请实施例的一个方面，提供了一种计算机程序产品，所述计算机程序产品由处理器加载并执行以实现上述图像生成模型的训练方法，或实现上述图像生成方法。

附图说明

图1是本申请一个实施例提供的图像生成模型的训练方法的流程图；

图2是本申请一个实施例提供的模型嫁接的示意图；

图3是本申请另一个实施例提供的图像生成模型的训练方法的流程图；

图4是本申请一个实施例提供的第一变换模型的结构示意图；

图5是本申请一个实施例提供的重构模型的结构示意图；

图6是本申请一个实施例提供的图像生成方法的流程图；

图7是本申请一个实施例提供的图像生成方法的示意图；

图8是本申请一个实施例提供的图像生成模型的训练装置的框图；

图9是本申请另一个实施例提供的图像生成模型的训练装置的框图；

图10是本申请一个实施例提供的图像生成装置的框图。

具体实施方式

人脸变换是计算机视觉中常见的图像处理方式。采用本申请实施例提供的方法训练得到的图像生成模型，能够应用于目标应用程序中（如视频类应用程序、社交类应用程序、直播类应用程序、购物类应用程序、图像处理类应用程序等）中，通过将待变换图像（图片或视频）输入目标应用程序中，生成变换后的目标图像。其中，图像生成模型写入了目标应用程序，或写入了为目标应用程序提供后台服务的服务器中。可选地，人脸变换还可以应用于人脸漫画化、人脸风格化、人脸属性编辑（如从视觉上改变年龄、改变性别等）等场景中，本申请实施例对此不作具体限定。

本申请实施例提供的方法，各步骤的执行主体可以是计算机设备，该计算机设备是指具备数据计算、处理和存储能力的电子设备。该计算机设备可以是诸如PC（Personal Computer，个人计算机）、平板电脑、智能手机、可穿戴设备、智能机器人等终端；也可以是服务器。其

中，服务器可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云计算服务的云服务器。

下面，通过几个实施例对本申请技术方案进行介绍说明。

请参考图 1，其示出了本申请一个实施例提供的图像生成模型的训练方法的流程图。在本实施例中，主要以该方法应用于上文介绍的计算机设备中来举例说明。该方法可以包括如下几个步骤（101~105）：

步骤 101，训练得到第一变换模型，第一变换模型用于基于第一噪声样本生成第一训练图像，第一训练图像为第一类风格的图像。

在一些实施例中，将第一噪声样本输入第一变换模型，得到对应的第一训练图像，再基于该对应的第一训练图像调整第一变换模型的参数；之后，继续将剩余的第一噪声样本输入第一变换模型，对第一变换模型进行下一次的迭代训练，直到第一变换模型训练完成。其中，第一噪声样本是经过向量化处理、并以向量形式或矩阵形式输入第一变换模型的。

在本实施例中，第一变换模型用于生成符合第一特征的第一训练图像。第一特征为客观存在的人或物体或场景的特征。在一个示例中，第一特征为真实人脸特征，也即，第一训练图像为模拟真实人脸的图像。在训练过程中，通过判别第一训练图像中的人脸符合真实人脸特征可能性，调整第一变换模型的权重参数，使得第一变换模型能够生成符合真实人脸特征的图像。

可选地，第一噪声样本为随机生成的噪声。在一些实施例中，第一噪声样本的概率密度函数服从高斯分布。

步骤 102，基于第一变换模型，训练得到重构模型，重构模型用于将原始图像样本映射为原始图像样本对应的隐变量。

可选地，隐变量是指未能观测到的变量。在一些实施例中，重构模型能够将原始图像样本通过编码映射为对应的隐变量，从而将高维的原始图像样本转换为较低维度的隐变量，减少了模型的计算量和计算时间，进一步提升了模型训练的时间成本。

步骤 103，训练得到第二变换模型，第二变换模型用于基于第二噪声样本生成第二训练图像，第二训练图像为第二类风格的图像。

在一些实施例中，将第二噪声样本输入第二变换模型，得到对应的第二训练图像，再基于该对应的第二训练图像调整第二变换模型的参数；之后，继续将剩余的第二噪声样本输入第二变换模型，对第二变换模型进行下一次的迭代训练，直到第二变换模型训练完成。

在本实施例中，第二类风格为漫画风格（如波普漫画风格、简笔漫画风格等）、水墨画风格、工笔画风格、油画风格、水彩画风格、素描风格、抽象画风格、映像画风格等等，本申请实施例对此不作具体限定。

可选地，在训练过程中，第二变换模型的初始权重参数为第一变换模型的权重参数。也即，第二变换模型基于第一变换模型进一步训练得到的。

步骤 104，将第一变换模型与第二变换模型进行嫁接，生成嫁接后的变换模型。

可选地，分别从第一变换模型和第二变换模型中选取不同部分的权重参数进行嫁接，生成嫁接后的变换模型。嫁接后的变换模型同时具备第一变换模型生成或保持图像的第一特征的特性，以及第二变换模型生成第二类风格的图像的特征。

在一些实施例中，将第一变换模型的多层权重网络中的 n 层权重网络，与第二变换模型的多层权重网络中的 m 层权重网络进行嫁接，生成嫁接后的变换模型。其中， n 层权重网络与 m 层权重网络的层次不相同， n 为正整数， m 为正整数。可选地，第一变换模型和第二变换模型为结构相同或相近的模型，其权重网络的层数相同，将第一变换模型的后 n 层权重网络拼接在第二变换模型的前 m 层权重网络之后，生成嫁接后的变换模型。可选地，嫁接后的变换模型与第一变换模型或第二变换模型的权重网络的层数相同。嫁接后的变换模型与第一

变换模型或第二变换模型的结构相同或相近。

需要说明的是， n 为 2、3、4、5、6、7、8、9 等等。可选地， n 的具体取值由相关技术人员根据实际情况进行设定，本申请实施例对此不作具体限定。 m 为 2、3、4、5、6、7、8、9 等等。可选地， m 的具体取值由相关技术人员根据实际情况进行设定，本申请实施例对此不作具体限定。

可选地， n 与 m 之和为第一变换模型或第二变换模型的权重网络的层数。

在一个示例中，如图 2 所示，第一变换模型 21 和第二变换模型 22 为结构相同的模型，其权重网络的层数均为 14 层。将第一变换模型 21 的后 6 层权重网络 23 拼接在第二变换模型 22 的前 8 层权重网络 24 之后，得到嫁接后的变换模型 25。

在另一些实施例中，将第一变换模型的多层权重网络中的权重参数，与第二变换模型的多层权重网络中对应的权重参数进行求和或平均或做差运算，生成嫁接后的变换模型。可选地，第一变换模型、第二变换模型和嫁接后的变换模型为结构相同或相近的模型，其权重网络的结构和层数完全相同。在一个示例中，将第一变换模型中权重网络的第 i 个权重参数，与第二变换模型中权重网络的第 i 个权重参数进行求和或平均或做差运算，得到嫁接后的变换模型中权重网络的第 i 个权重参数， i 为正整数；以此类推，得到嫁接后的变换模型中权重网络的所有权重参数。可选地，求和运算为加权求和运算，计算嫁接后的变换模型中权重网络的各个权重参数所采用的权重系数均相同。

步骤 105，基于重构模型和嫁接后的变换模型，生成图像生成模型。

可选地，图像生成模型用于将第一类风格的待变换图像，变换为第二类风格的目标图像。在一些实施例中，将重构模型与嫁接后的变换模型组合，即得到图像生成模型。待变换图像为第一类风格的图像，目标图像为第二类风格的图像，将待变换图像输入图像生成模型，能够生成目标图像，从而将待变换图像由第一类风格变换为第二类风格，使待变换图像实现风格变换。

在一些可能的实现方式中，该步骤 105 包括如下子步骤：

- 1、将重构模型和嫁接后的变换模型进行组合，得到组合后的变换模型；
- 2、获取第四训练样本集，第四训练样本集包括至少一个原始图像样本，以及至少一个原始图像样本分别对应的第二类风格的图像；
- 3、采用第四训练样本集，对组合后的变换模型进行微调(fine-tune)，生成图像生成模型。

在一些实施例中，将嫁接后的变换模型拼接在重构模型之后，得到组合后的变换模型；之后，通过模型训练对组合后的变换模型的权重参数进行微调，得到图像生成模型。

综上所述，本申请实施例提供的技术方案中，通过分别训练第一变换模型和第二变换模型，由于第一变换模型具备良好地保持图像的目标特征的优势，第二变换模型具备生成风格较为强烈的第二类风格的图像的优势，训练过程中，分别针对第一变换模型和第二变换模型各自的目标优势进行训练，再将训练得到的第一变换模型和第二变换模型进行嫁接，得到嫁接后的变换模型，基于重构模型和嫁接后的变换模型生成图像生成模型，较大程度地降低了模型训练的复杂程度，因而所需训练样本和训练迭代次数较少，节省了模型训练的时间成本。

请参考 3 图，其示出了本申请另一个实施例提供的图像生成模型的训练方法的流程图。在本实施例中，主要以该方法应用于上文介绍的计算机设备中来举例说明。该方法可以包括如下几个步骤（步骤 301~312）：

步骤 301，获取第一训练样本集，第一训练样本集包括多个第一噪声样本。

可选地，第一噪声样本是在训练第一变换模型时使用的噪声样本。在一些实施例中，通过随机噪声生成函数，生成多个随机的第一噪声样本，得到第一训练样本集。

可选地，第一变换模型为 StyleGAN (StyleGenerative Adversarial Networks，基于样式的生成对抗网络)-V2 模型。如图 4 所示，第一变换模型包括第一映射网络 (Mapping Network)

41 和第一合成网络 42。映射网络能够对输入进行解缠。第一映射网络 41 包括 8 个全连接层 (FC, Fully Connected Layer)。第一合成网络 42 包括至少一个可学习的仿射变换层 (A, Affine Transformation)、至少一个调制模块 (Mod-Demod) 和至少一个上采样 (Upsample) 层、至少一个常数 (Const) 层和至少一个噪声参数层。可选地, 调制模块的计算公式包括如下公式一、二、三:

公式一:

$$w'_{ijk} = s_i \cdot w_{ijk}$$

公式二:

$$\sigma_j = \sqrt{\sum_{i,k} w'_{ijk}{}^2}$$

公式三:

$$w''_{ijk} = w'_{ijk} / \sqrt{\sum_{i,k} w'_{ijk}{}^2 + \epsilon}$$

其中, i 、 j 、 k 表示调制模块中隐变量在不同维度上的编号, w_{ijk} 表示编号为 (i, j, k) 的隐变量, w'_{ijk} 表示公式一的计算结果, σ_j 表示公式二的计算结果, ϵ 表示极小数, w''_{ijk} 表示公式三的计算结果。

在一些实施例中, 第一变换模型包括第一映射网络 and 第一合成网络。

步骤 302, 将多个第一噪声样本分别输入第一映射网络, 得到多个第一噪声样本分别对应的隐变量。

在一些实施例中, 第一映射网络包括多个全连接层, 将多个第一噪声样本分别输入第一映射网络后, 经过多个全连接层, 得到多个第一噪声样本分别对应的隐变量。

步骤 303, 将多个第一噪声样本分别对应的隐变量分别输入第一合成网络, 得到多个第一噪声样本分别对应的第一训练图像。

在一些实施例中, 将多个第一噪声样本分别对应的隐变量分别输入第一合成网络, 多个第一噪声样本分别对应的隐变量在第一合成网络中经过仿射变换、调制、上采样之后, 得到多个第一噪声样本分别对应的第一训练图像。

步骤 304, 基于多个第一噪声样本分别对应的第一训练图像, 调整第一变换模型的权重参数。

可选地, 得到多个第一噪声样本分别对应的第一训练图像之后, 通过对多个第一噪声样本分别对应的第一训练图像进行判别的判别结果, 调整第一变换模型的权重参数。

在一些实施例中, 第一变换模型包括第一判别网络。该步骤 304 还包括如下步骤:

1、将多个第一噪声样本分别对应的第一训练图像分别输入第一判别网络, 得到多个第一噪声样本分别对应的第一判别损失;

2、基于多个第一噪声样本分别对应的第一判别损失, 调整第一变换模型的权重参数。

在该实现方式中, 采用第一变换模型中自带的第一判别网络, 对多个第一噪声样本分别对应的第一训练图像进行判定, 确定多个第一噪声样本分别对应的第一训练图像的生成质量, 从而确定多个第一噪声样本分别对应的第一判别损失, 再基于此调整变换模型的权重参数。可选地, 调整变换模型的权重参数时, 第一判别网络的权重参数也一同进行调整。可选地, 第一判别网络为 GAN (Generative Adversarial Networks, 生成对抗网络), 第一判别损失可表示为 GAN_loss (GAN 损失)。

在一个示例中, 第一训练图像为基于第一噪声样本模拟的人脸图像, 第一判别网络用于评判第一训练图像与真实人脸的图像相符合的概率。例如, 第一训练图像中人脸的完整程度越高, 第一判别损失越小; 又例如, 第一训练图像中人脸的五官比例越接近真实人脸中五官的比例, 第一判别损失越小。

在一些实施例中，第一变换模型在训练完成之后，将其第一映射网络舍去，将除第一映射网络的其他部分作为第一变换模型。

步骤 305，获取第二训练样本集，第二训练样本集包括多个原始图像样本。

可选地，原始图像样本为第一类风格的图像。在一些实施例中，原始图像样本为未经处理过的真实人脸的图像。

步骤 306，将多个原始图像样本分别输入重构模型，生成多个原始图像样本分别对应的隐变量。

在一些实施例中，将多个原始图像样本分别输入重构模型，重构模型对多个原始图像样本分别进行编码，生成多个原始图像样本分别对应的隐变量。

在一个示例中，如图 5 所示，重构模型包括输入层 51、多个残差网络（resblock）层 52 和全连接层 53。

步骤 307，将多个原始图像样本分别对应的隐变量输入第一变换模型，生成多个原始图像样本分别对应的重构图像。

在一些实施例中，将多个原始图像样本分别对应的隐变量，直接输入第一变换模型的第一合成网络中，生成多个原始图像样本分别对应的重构图像。可选地，多个原始图像样本以及多个原始图像样本分别对应的重构图像为第一类风格的图像。

需要说明的是，训练重构模型的过程中时，第一变换模型是已训练完成的模型，第一变换模型的参数不发生变化。

步骤 308，基于多个原始图像样本，以及多个原始图像样本分别对应的重构图像，确定多个原始图像样本分别对应的重构模型的损失。

在一些实施例中，基于多个原始图像样本，以及多个原始图像样本分别对应的重构图像分别在各个网络或函数中的输出，确定重构模型的损失中包括的多个子损失。

在一些实施例中，该步骤 308 还包括如下子步骤：

1、基于将多个原始图像样本分别对应的重构图像，分别输入第一判别网络得到的输出结果，确定第一子损失，第一子损失用于指示重构图像的第一特征性。

可选地，第一特征性为重构图像中与需要包含的特征的符合程度。在一个示例中，当重构图像为人脸图像时，第一特征性是指重构图像与人脸特征相符合的程度。重构图像中的人脸越接近真实人脸，第一子损失越小。

2、基于将多个原始图像样本以及多个原始图像样本分别对应的重构图像，分别输入感知网络得到的输出结果，确定第二子损失，第二子损失用于指示原始图像样本与原始图像样本对应的重构图像，在目标特征上的第一符合程度。

可选地，通过目标特征，能够判断重构图像是否是基于某个原始图像样本生成的图像。在一个示例中，目标特征为人脸标识特征，人脸标识特征用于区分不同的人脸。通过感知网络（Perceptual Network）比对原始图像样本中的人脸标识特征，与原始图像样本对应的重构图像中的人脸标识特征，确定原始图像样本与原始图像样本对应的重构图像为同一个人的图像的概率，从而确定第二子损失。可选地，第二子损失为感知损失（Perceptual loss）。可选地，第二子损失的计算方式可以参考如下公式四：

$$\text{公式四：Perceptual loss} = E((VGG(x) - VGG(G(x)))^2)$$

其中，Perceptual loss 表示第二子损失，x 表示原始图像样本，VGG(x) 表示将原始图像样本输入 VGG（Visual Geometry Group Network，视觉几何群网络）-16 得到的输出结果，G(x) 表示原始图像样本对应的重构图像，VGG(G(x))² 表示将原始图像样本对应的重构图像输入 VGG-16 网络得到的输出结果。

3、基于将多个原始图像样本以及多个原始图像样本分别对应的重构图像，分别输入回归函数得到的输出结果，确定第三子损失，第三子损失用于指示原始图像样本与原始图像样本对应的重构图像，在目标特征上的第二符合程度。

可选地，回归函数为 L1 函数、L2 函数等等。通过原始图像样本与原始图像样本对应的重构图像，分别对应于 L 函数的输出结果，确定第三子损失。可选地，第三子损失为 L1_loss。可选地，第三子损失的计算方式可以参考如下公式五：

$$\text{公式五： } L1_loss = E(x - G(x))$$

其中，L1_loss 表示第三子损失，x 表示原始图像样本，VGG(x) 表示将原始图像样本输入 VGG-16 得到的输出结果。

4、基于第一子损失、第二子损失和第三子损失，确定重构模型的损失。

在一个示例中，将第一子损失、第二子损失和第三子损失进行直接求和或加权求和，计算得到重构模型的损失。在一些实施例中，对第一子损失、第二子损失和第三子损失进行加权求和时，第一子损失、第二子损失和第三子损失分别对应的权重的值由相关技术人员根据实际情况进行设定，本申请实施例对此不作具体限定。

在一个示例中，重构模型的损失的计算方式参考如下公式六：

$$\text{公式六： } Loss = GAN_loss + \text{Perceptual loss} + L1_loss$$

其中，Loss 表示重构模型的损失，GAN_loss 表示第一子损失，Perceptual loss 表示第二子损失，L1_loss 表示第三子损失。

步骤 309，基于多个原始图像样本分别对应的重构模型的损失，调整重构模型的权重参数。

可选地，基于多个原始图像样本分别对应的重构模型的损失，调整重构模型的权重参数，使得重构模型的损失尽可能得小。在一些实施例中，当迭代训练的次数达到预设次数时，表示重构模型训练完成。在另一些实施例中，重构模型的损失连续小于重构损失阈值，表示重构模型训练完成。可选地，重构损失阈值的具体取值由相关技术人员根据实际情况进行设定，本申请实施例对此不作具体限定。

步骤 310，训练得到第二变换模型，第二变换模型用于基于第二噪声样本生成第二训练图像，第二训练图像为第二类风格的图像。

本步骤 310 的内容与本申请实施例步骤 301~304 的内容相同或相近，此处不再赘述。

步骤 311，将第一变换模型与第二变换模型进行嫁接，生成嫁接后的变换模型。

本步骤 311 的内容与上述图 1 实施例步骤 104 的内容相同或相近，此处不再赘述。

步骤 312，基于重构模型和嫁接后的变换模型，生成图像生成模型，图像生成模型用于将第一类风格的待变换图像，变换为第二类风格的目标图像。

本步骤 312 的内容与上述图 1 实施例步骤 105 的内容相同或相近，此处不再赘述。

综上所述，本申请实施例提供的技术方案中，通过先训练得到第一变换模型，在基于第一变换模型训练重构模型，减少了训练重构模型所需的时间，进一步降低了整体的模型训练的时间成本。

本申请实施例中，将训练完成的第一变换模型的权重参数，作为第二变换模型训练时的初始参数，使得第二变换模型能够保留第一变换模型的部分特性，便于提升第一变换模型和第二变换模型之间嫁接后的协调性。

下面，结合图 6 和图 7 对本申请一个实施例提供的图像生成方法进行介绍。在本实施例中，主要以该方法应用于上文介绍的计算机设备中来举例说明。如图 6 所示，该方法可以包括如下几个步骤（步骤 601~602）：

步骤 601，如图 7 所示，将第一类风格的待变换图像 71 输入重构模型 72，生成待变换图像 71 对应的隐变量 73。

可选地，待变换图像 71 为未经处理过的、真实人脸的图像。

步骤 602，基于待变换图像 71 对应的隐变量 73，采用嫁接后的变换模型 74 生成待变换图像 71 对应的目标图像 75，目标图像 75 为第二类风格的图像。

可选地,嫁接后的变换模型 74 是通过将第一变换模型与第二变换模型进行嫁接生成的模型;第一变换模型用于基于第一噪声样本生成第一训练图像,第一训练图像为第一类风格的图像;第二变换模型用于基于第二噪声样本生成第二训练图像,第二训练图像为第二类风格的图像。

在一些实施例中,嫁接后的变换模型 74 中来自于第二变换模型的权重网络,先基于待变换图像 71 对应的隐变量 73 将待变换图像 71 风格化,使得待变换图像 71 变换为第二类风格的图像,之后,嫁接后的变换模型 74 中来自于第一变换模型的权重网络,对风格化的待变换图像 71 中的目标特征进行调整,从而得到最终的目标图像 75。对目标特征的解释说明可参见上述图 3 实施例的步骤 308,此处不再赘述。

综上所述,本申请实施例提供的技术方案中,通过基于第一变换模型和第二变换模型嫁接得到的图像生成模型生成目标图像 75,由于图像生成模型(嫁接后的变换模型 74)同时具备第一变换模型良好地保持图像的目标特征的优势,以及第二变换模型生成风格较为强烈的第二类风格的图像的优势,能够在较为良好地保持待变换图像的目标特征的前提下,生成风格较为强烈的目标图像 75,提升了图像生成模型的变换效果。

下述为本申请装置实施例,可以用于执行本申请方法实施例。对于本申请装置实施例中未披露的细节,请参照本申请方法实施例。

请参考图 8,其示出了本申请一个实施例提供的图像生成模型的训练装置的框图。该装置具有实现上述图像生成模型的训练方法示例的功能,所述功能可以由硬件实现,也可以由硬件执行相应的软件实现。该装置 800 可以包括:模型训练模块 810 和模型生成模块 820。

所述模型训练模块 810,用于训练得到第一变换模型,所述第一变换模型用于基于第一噪声样本生成第一训练图像,所述第一训练图像为第一类风格的图像。

所述模型训练模块 810,还用于基于所述第一变换模型,训练得到重构模型,所述重构模型用于将原始图像样本映射为所述原始图像样本对应的隐变量。

所述模型训练模块 810,还用于训练得到第二变换模型,所述第二变换模型用于基于第二噪声样本生成第二训练图像,所述第二训练图像为第二类风格的图像。

所述模型生成模块 820,用于将所述第一变换模型与所述第二变换模型进行嫁接,生成嫁接后的变换模型。

所述模型生成模块 820,还用于基于所述重构模型和所述嫁接后的变换模型,生成图像生成模型,所述图像生成模型用于将所述第一类风格的待变换图像,变换为所述第二类风格的目标图像。

综上所述,本申请实施例提供的技术方案中,通过分别训练第一变换模型和第二变换模型,由于第一变换模型具备良好地保持图像的目标特征的优势,第二变换模型具备生成风格较为强烈的第二类风格的图像的优势,训练过程中,分别针对第一变换模型和第二变换模型各自的目标优势进行训练,再将训练得到的第一变换模型和第二变换模型进行嫁接,得到嫁接后的变换模型,基于重构模型和嫁接后的变换模型生成图像生成模型,较大程度地降低了模型训练的复杂程度,因而所需训练样本和训练迭代次数较少,节省了模型训练的时间成本。

在示例性实施例中,所述第一变换模型包括第一映射网络 and 第一合成网络;如图 9 所示,所述模型训练模块 810,包括:样本获取子模块 811、变量生成子模块 812、图像生成子模块 813 和参数调整子模块 814。

所述样本获取子模块 811,用于获取第一训练样本集,所述第一训练样本集包括多个第一噪声样本。

所述变量生成子模块 812,用于将所述多个第一噪声样本分别输入所述第一映射网络,得到所述多个第一噪声样本分别对应的隐变量。

所述图像生成子模块 813,用于将所述多个第一噪声样本分别对应的隐变量分别输入所

述第一合成网络,得到所述多个第一噪声样本分别对应的第一训练图像。

所述参数调整子模块 814,用于基于所述多个第一噪声样本分别对应的第一训练图像,调整所述第一变换模型的权重参数。

在示例性实施例中,所述第一变换模型包括第一判别网络;如图 9 所示,所述参数调整子模块 814,用于:

将所述多个第一噪声样本分别对应的第一训练图像分别输入所述第一判别网络,得到所述多个第一噪声样本分别对应的第一判别损失;

基于所述多个第一噪声样本分别对应的第一判别损失,调整所述第一变换模型的权重参数。

在示例性实施例中,如图 9 所示,所述模型训练模块 810,包括:损失确定子模块 815。

所述样本获取子模块 811,还用于获取第二训练样本集,所述第二训练样本集包括多个原始图像样本。

所述变量生成子模块 812,还用于将所述多个原始图像样本分别输入所述重构模型,生成所述多个原始图像样本分别对应的隐变量。

所述图像生成子模块 813,还用于将所述多个原始图像样本分别对应的隐变量输入所述第一变换模型,生成所述多个原始图像样本分别对应的重构图像,所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像为所述第一类风格的图像。

所述损失确定子模块 815,用于基于所述多个原始图像样本,以及所述多个原始图像样本分别对应的重构图像,确定所述多个原始图像样本分别对应的重构模型的损失。

所述参数调整子模块 814,还用于基于所述多个原始图像样本分别对应的重构模型的损失,调整所述重构模型的权重参数。

在示例性实施例中,所述第一变换模型包括第一判别网络;如图 9 所示,所述损失确定子模块 815,用于:

基于将所述多个原始图像样本分别对应的重构图像,分别输入所述第一判别网络得到的输出结果,确定第一子损失,所述第一子损失用于指示所述重构图像的第一特征性;

基于将所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像,分别输入感知网络得到的输出结果,确定第二子损失,所述第二子损失用于指示所述原始图像样本与所述原始图像样本对应的重构图像,在目标特征上的第一符合程度;

基于将所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像,分别输入回归函数得到的输出结果,确定第三子损失,所述第三子损失用于指示所述原始图像样本与所述原始图像样本对应的重构图像,在所述目标特征上的第二符合程度;

基于所述第一子损失、所述第二子损失和所述第三子损失,确定所述重构模型的损失。

在示例性实施例中,在训练过程中,所述第二变换模型的初始权重参数为所述第一变换模型的权重参数。

在示例性实施例中,所述模型生成模块 820,用于:将所述第一变换模型的多层权重网络中的 n 层权重网络,与所述第二变换模型的多层权重网络中的 m 层权重网络进行嫁接,生成所述嫁接后的变换模型;所述 n 层权重网络与所述 m 层权重网络的层次不相同,所述 n 为正整数,所述 m 为正整数;或者,将所述第一变换模型的多层权重网络中的权重参数,与所述第二变换模型的多层权重网络中对应的权重参数进行求和或平均或做差运算,生成所述嫁接后的变换模型。

在示例性实施例中,所述模型生成模块 820,用于:

将所述重构模型和所述嫁接后的变换模型进行组合,得到组合后的变换模型;

获取第四训练样本集,所述第四训练样本集包括至少一个原始图像样本,以及所述至少一个原始图像样本分别对应的第二类风格的图像;

采用所述第四训练样本集,对所述组合后的变换模型进行微调,生成所述图像生成模型。

请参考图 10，其示出了本申请一个实施例提供的图像生成装置的框图。该装置具有实现上述图像生成方法示例的功能，所述功能可以由硬件实现，也可以由硬件执行相应的软件实现。该装置 1000 可以包括：变量生成模块 1010 和图像生成模块 1020。

所述变量生成模块 1010，用于将第一类风格的待变换图像输入重构模型，生成所述待变换图像对应的隐变量。

所述图像生成模块 1020，用于基于所述待变换图像对应的隐变量，采用嫁接后的变换模型生成所述待变换图像对应的目标图像，所述目标图像为第二类风格的图像。其中，所述嫁接后的变换模型是通过将第一变换模型与第二变换模型进行嫁接生成的模型；所述第一变换模型用于基于第一噪声样本生成所述第一类风格的图像；所述第二变换模型用于基于第二噪声样本生成所述第二类风格的图像。

综上所述，本申请实施例提供的技术方案中，通过基于第一变换模型和第二变换模型嫁接得到的图像生成模型生成目标图像，由于图像生成模型（嫁接后的变换模型）同时具备第一变换模型良好地保持图像的目标特征的优势，以及第二变换模型生成风格较为强烈的第二类风格的图像的优势，能够在较为良好地保持待变换图像的目标特征的前提下，生成风格较为强烈的目标图像，提升了图像生成模型的变换效果。

需要说明的是，上述实施例提供的装置，在实现其功能时，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的装置与方法实施例属于同一构思，其具体实现过程详见方法实施例，这里不再赘述。

在示例性实施例中，提供了一种计算机设备，所述计算机设备包括处理器和存储器，所述存储器中存储有计算机程序，所述计算机程序由所述处理器加载并执行以实现上述图像生成模型的训练方法。

在示例性实施例中，提供了一种计算机设备，所述计算机设备包括处理器和存储器，所述存储器中存储有计算机程序，所述计算机程序由所述处理器加载并执行以实现上述图像生成方法。

在示例性实施例中，还提供了一种计算机可读存储介质，所述存储介质中存储有计算机程序，所述计算机程序在被处理器执行时以实现上述图像生成模型的训练方法。

在示例性实施例中，还提供了一种计算机可读存储介质，所述存储介质中存储有计算机程序，所述计算机程序在被处理器执行时以实现上述图像生成方法。

在示例性实施例中，还提供了一种计算机程序产品，当该计算机程序产品被处理器执行时，其用于实现上述图像生成模型的训练方法。

在示例性实施例中，还提供了一种计算机程序产品，当该计算机程序产品被处理器执行时，其用于实现上述图像生成方法。

应当理解的是，在本文中提及的“多个”是指两个或两个以上。字符“/”一般表示前后关联对象是一种“或”的关系。

以上所述仅为本申请的示例性实施例，并不用以限制本申请，凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权利要求书

1.一种图像生成模型的训练方法，其特征在于，所述方法包括：

训练得到第一变换模型，所述第一变换模型用于基于第一噪声样本生成第一训练图像，所述第一训练图像为第一类风格的图像；

基于所述第一变换模型，训练得到重构模型，所述重构模型用于将原始图像样本映射为所述原始图像样本对应的隐变量；

训练得到第二变换模型，所述第二变换模型用于基于第二噪声样本生成第二训练图像，所述第二训练图像为第二类风格的图像；

将所述第一变换模型与所述第二变换模型进行嫁接，生成嫁接后的变换模型；

基于所述重构模型和所述嫁接后的变换模型，生成图像生成模型，所述图像生成模型用于将所述第一类风格的待变换图像，变换为所述第二类风格的目标图像。

2.根据权利要求1所述的方法，其特征在于，所述第一变换模型包括第一映射网络 and 第一合成网络；

所述训练得到第一变换模型，包括：

获取第一训练样本集，所述第一训练样本集包括多个第一噪声样本；

将所述多个第一噪声样本分别输入所述第一映射网络，得到所述多个第一噪声样本分别对应的隐变量；

将所述多个第一噪声样本分别对应的隐变量分别输入所述第一合成网络，得到所述多个第一噪声样本分别对应的第一训练图像；

基于所述多个第一噪声样本分别对应的第一训练图像，调整所述第一变换模型的权重参数。

3.根据权利要求2所述的方法，其特征在于，所述第一变换模型包括第一判别网络；

所述基于所述多个第一噪声样本分别对应的第一训练图像，调整所述第一变换模型的权重参数，包括：

将所述多个第一噪声样本分别对应的第一训练图像分别输入所述第一判别网络，得到所述多个第一噪声样本分别对应的第一判别损失；

基于所述多个第一噪声样本分别对应的第一判别损失，调整所述第一变换模型的权重参数。

4.根据权利要求1所述的方法，其特征在于，所述基于所述第一变换模型，训练得到重构模型，包括：

获取第二训练样本集，所述第二训练样本集包括多个原始图像样本；

将所述多个原始图像样本分别输入所述重构模型，生成所述多个原始图像样本分别对应的隐变量；

将所述多个原始图像样本分别对应的隐变量输入所述第一变换模型，生成所述多个原始图像样本分别对应的重构图像，所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像为所述第一类风格的图像；

基于所述多个原始图像样本，以及所述多个原始图像样本分别对应的重构图像，确定所述多个原始图像样本分别对应的重构模型的损失；

基于所述多个原始图像样本分别对应的重构模型的损失，调整所述重构模型的权重参数。

5.根据权利要求4所述的方法，其特征在于，所述第一变换模型包括第一判别网络；

所述基于所述多个原始图像样本,以及所述多个原始图像样本分别对应的重构图像,确定所述多个原始图像样本分别对应的重构模型的损失,包括:

基于将所述多个原始图像样本分别对应的重构图像,分别输入所述第一判别网络得到的输出结果,确定第一子损失,所述第一子损失用于指示所述重构图像的第一特征性;

基于将所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像,分别输入感知网络得到的输出结果,确定第二子损失,所述第二子损失用于指示所述原始图像样本与所述原始图像样本对应的重构图像,在目标特征上的第一符合程度;

基于将所述多个原始图像样本以及所述多个原始图像样本分别对应的重构图像,分别输入回归函数得到的输出结果,确定第三子损失,所述第三子损失用于指示所述原始图像样本与所述原始图像样本对应的重构图像,在所述目标特征上的第二符合程度;

基于所述第一子损失、所述第二子损失和所述第三子损失,确定所述重构模型的损失。

6.根据权利要求1至5任一项所述的方法,其特征在于,在训练过程中,所述第二变换模型的初始权重参数为所述第一变换模型的权重参数。

7.根据权利要求1至5任一项所述的方法,其特征在于,所述将所述第一变换模型与所述第二变换模型进行嫁接,生成嫁接后的变换模型,包括:

将所述第一变换模型的多层权重网络中的n层权重网络,与所述第二变换模型的多层权重网络中的m层权重网络进行嫁接,生成所述嫁接后的变换模型;所述n层权重网络与所述m层权重网络的层次不相同,所述n为正整数,所述m为正整数;

或者,

将所述第一变换模型的多层权重网络中的权重参数,与所述第二变换模型的多层权重网络中对应的权重参数进行求和或平均或做差运算,生成所述嫁接后的变换模型。

8.根据权利要求1至5任一项所述的方法,其特征在于,所述基于所述重构模型和所述嫁接后的变换模型,生成图像生成模型,包括:

将所述重构模型和所述嫁接后的变换模型进行组合,得到组合后的变换模型;

获取第四训练样本集,所述第四训练样本集包括至少一个原始图像样本,以及所述至少一个原始图像样本分别对应的第二类风格的图像;

采用所述第四训练样本集,对所述组合后的变换模型进行微调,生成所述图像生成模型。

9.一种图像生成方法,其特征在于,所述方法包括:

将第一类风格的待变换图像输入重构模型,生成所述待变换图像对应的隐变量;

基于所述待变换图像对应的隐变量,采用嫁接后的变换模型生成所述待变换图像对应的目标图像,所述目标图像为第二类风格的图像;

其中,所述嫁接后的变换模型是通过将第一变换模型与第二变换模型进行嫁接生成的模型;所述第一变换模型用于基于第一噪声样本生成所述第一类风格的图像;所述第二变换模型用于基于第二噪声样本生成所述第二类风格的图像。

10.一种图像生成模型的训练装置,其特征在于,所述装置包括:

模型训练模块,用于训练得到第一变换模型,所述第一变换模型用于基于第一噪声样本生成第一训练图像,所述第一训练图像为第一类风格的图像;

所述模型训练模块,还用于基于所述第一变换模型,训练得到重构模型,所述重构模型用于将原始图像样本映射为所述原始图像样本对应的隐变量;

所述模型训练模块,还用于训练得到第二变换模型,所述第二变换模型用于基于第二噪声样本生成第二训练图像,所述第二训练图像为第二类风格的图像;

模型生成模块，用于将所述第一变换模型与所述第二变换模型进行嫁接，生成嫁接后的变换模型；

所述模型生成模块，还用于基于所述重构模型和所述嫁接后的变换模型，生成图像生成模型，所述图像生成模型用于将所述第一类风格的待变换图像，变换为所述第二类风格的目标图像。

11.一种图像生成装置，其特征在于，所述装置包括：

变量生成模块，用于将第一类风格的待变换图像输入重构模型，生成所述待变换图像对应的隐变量；

图像生成模块，用于基于所述待变换图像对应的隐变量，采用嫁接后的变换模型生成所述待变换图像对应的目标图像，所述目标图像为第二类风格的图像；

其中，所述嫁接后的变换模型是通过将第一变换模型与第二变换模型进行嫁接生成的模型；所述第一变换模型用于基于第一噪声样本生成所述第一类风格的图像；所述第二变换模型用于基于第二噪声样本生成所述第二类风格的图像。

12.一种计算机设备，其特征在于，所述计算机设备包括处理器和存储器，所述存储器中存储有计算机程序，所述计算机程序由所述处理器加载并执行以实现如上述权利要求 1 至 8 任一项所述的图像生成模型的训练方法，或实现如上述权利要求 9 所述的图像生成方法。

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中存储有计算机程序，所述计算机程序由处理器加载并执行以实现如上述权利要求 1 至 8 任一项所述的图像生成模型的训练方法，或实现如上述权利要求 9 所述的图像生成方法。

14.一种计算机程序产品，其特征在于，所述计算机程序产品由处理器加载并执行以实现如上述权利要求 1 至 8 任一项所述的图像生成模型的训练方法，或实现如上述权利要求 9 所述的图像生成方法。