



(12) 发明专利申请

(10) 申请公布号 CN 114648767 A

(43) 申请公布日 2022. 06. 21

(21) 申请号 202210323610.X

G06N 20/00 (2019.01)

(22) 申请日 2022.03.30

(71) 申请人 京东科技信息技术有限公司

地址 100176 北京市北京经济技术开发区

科创十一街18号院2号楼6层601

(72) 发明人 袁韶祖 陈蒙 刘瑞雪

(74) 专利代理机构 北京同立钧成知识产权代理

有限公司 11205

专利代理师 张娜 臧建明

(51) Int. Cl.

G06V 30/226 (2022.01)

G06V 30/19 (2022.01)

G06K 9/62 (2022.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

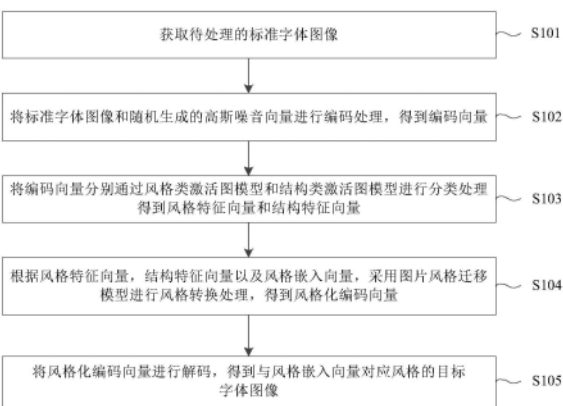
权利要求书3页 说明书13页 附图5页

(54) 发明名称

风格化书法的处理方法、装置、设备及介质

(57) 摘要

本发明提供一种风格化书法的处理方法、装置、设备及介质，在该方法中，在获取到待处理的标准字体图像后，将标准字体图像和随机生成的高斯噪音向量进行编码处理，得到编码向量。将编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量。然后根据风格特征向量，结构特征向量以及风格嵌入向量，采用图片风格迁移模型进行风格转换处理，得到风格化编码向量，再将风格化编码向量进行解码，得到与风格嵌入向量对应风格的目标字体图像。通过引用用户的风格嵌入向量使得生成的字体图像更贴近用户的风格，另外通过引入高斯噪音向量可以提高生成字体的多样性，更贴近用户撰写的实际情况。



1. 一种风格化书法的处理方法,其特征在于,包括:
获取待处理的标准字体图像;
将所述标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量;
将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量;
根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量;
将所述风格化编码向量进行解码,得到与所述风格嵌入向量对应风格的目标字体图像。
2. 根据权利要求1所述的方法,其特征在于,所述将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量,包括:
将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向量;
将所述编码向量通过结构类激活图模型进行结构分类处理,得到所述结构特征向量。
3. 根据权利要求1或2所述的方法,其特征在于,所述根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理之前,所述方法还包括:
从预先配置的风格嵌入隐藏向量列表中,获取所述风格嵌入向量,所述风格嵌入向量用于表征目标字体风格的手写字体的风格;其中,所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。
4. 根据权利要求3所述的方法,其特征在于,所述根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量,包括:
将所述风格特征向量和所述结构特征向量进行叠加相乘处理,得到处理后的向量;
将所述处理后的向量和所述风格嵌入向量输入所述图片风格迁移模型进行风格转换处理,得到所述风格化编码向量。
5. 根据权利要求3所述的方法,其特征在于,所述方法还包括:
获取多个不同字体风格的手写字体图像;
通过对每个字体风格的手写字体图像的风格进行机器学习,得到用于表征每个字体风格的手写字体风格的风格嵌入向量。
6. 根据权利要求3所述的方法,其特征在于,所述方法还包括:
通过风格判别器检测所述目标字体图像中的字体风格,获取风格检测分;
通过结构判别器检测所述目标字体图像中的字体结构,获取结构检测分;
根据所述风格检测分以及所述结构检测分,计算获取生成的所述目标字体图像的图像损失值;
根据所述图像损失值对所述图片风格迁移模型的参数进行调整,以使所述图像损失值低于预设损失值阈值。
7. 根据权利要求6所述的方法,其特征在于,所述方法还包括:
获取多次生成的多个所述目标字体图像;
根据多个所述目标字体图像,每次输入的高斯噪音向量,生成次数以及预设的平衡因

子,确定多样性损失;

根据所述多样性损失对所述图片风格迁移模型的参数进行调整。

8. 一种风格化书法的处理装置,其特征在于,包括:

获取模块,用于获取待处理的标准字体图像;

编码模块,用于将所述标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量;

分类模块,用于将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量;

转换模块,用于根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量;

解码模块,用于将所述风格化编码向量进行解码,得到与所述风格嵌入向量对应风格的目标字体图像。

9. 根据权利要求8所述的装置,其特征在于,所述分类模块具体用于:

将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向量;

将所述编码向量通过结构类激活图模型进行结构分类处理,得到所述结构特征向量。

10. 根据权利要求8或9所述的装置,其特征在于,所述转换模块还用于:

从预先配置的风格嵌入隐藏向量列表中,获取所述风格嵌入向量,所述风格嵌入向量用于表征目标字体风格的手写字体的风格;其中,所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。

11. 根据权利要求10所述的装置,其特征在于,所述转换模块具体用于:

将所述风格特征向量和所述结构特征向量进行叠加相乘处理,得到处理后的向量;

将所述处理后的向量和所述风格嵌入向量输入所述图片风格迁移模型进行风格转换处理,得到所述风格化编码向量。

12. 根据权利要求10所述的装置,其特征在于,所述装置还包括:处理模块,用于:

获取多个不同字体风格的手写字体图像;

通过对每个字体风格的手写字体图像的风格进行机器学习,得到用于表征每个字体风格的手写字体风格的风格嵌入向量。

13. 根据权利要求10所述的装置,其特征在于,所述装置还包括:训练模块,用于:

通过风格判别器检测所述目标字体图像中的字体风格,获取风格检测分;

通过结构判别器检测所述目标字体图像中的字体结构,获取结构检测分;

根据所述风格检测分以及所述结构检测分,计算获取生成的所述目标字体图像的图像损失值;

根据所述图像损失值对所述图片风格迁移模型的参数进行调整,以使所述图像损失值低于预设损失值阈值。

14. 根据权利要求13所述的装置,其特征在于,所述训练模块还用于:

获取多次生成的多个所述目标字体图像;

根据多个所述目标字体图像,每次输入的高斯噪音向量,生成次数以及预设的平衡因子,确定多样性损失;

根据所述多样性损失对所述图片风格迁移模型的参数进行调整。

15. 一种电子设备,其特征在于,包括:

编码器,处理器,存储器,解码器以及交互接口;

所述存储器用于存储所述处理器的可执行指令;

其中,所述处理器通过执行所述可执行指令使所述电子设备执行权利要求1至7任一项所述的风格化书法的处理方法。

16. 一种可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至7任一项所述的风格化书法的处理方法。

17. 一种计算机程序产品,其特征在于,包括计算机程序,所述计算机程序被处理器执行时用于实现权利要求1至7任一项所述的风格化书法的处理方法。

风格化书法的处理方法、装置、设备及介质

技术领域

[0001] 本发明涉及计算机领域,尤其涉及一种风格化书法的处理方法、装置、设备及介质。

背景技术

[0002] 随着生成式对抗网络(Generative Adversarial Networks,简称:GAN)的发展,它已经被广泛应用在了艺术创作的多个领域,例如流行音乐创作、诗歌生成,图片去噪,人脸风格转换等等。AI书法生成指的是计算机通过学习有标注的书法家的书法作品,模拟书法家风格自动生成艺术字体,解放人力的同时又能产生艺术和应用价值。

[0003] 目前,较成熟的字体生成模型有以下两种:(1) Zi2zi模型,模型来源于pix2pix模型并有所改进,通过条件生成网络,加入类别嵌入,得到能够同时处理多种风格的GAN;(2) CycleGan模型,引入了cycle loss,CycleGan模型可以使用成对的数据,来学习两个不同种类的字体的样式,但CycleGan模型只能由一个类别映射到另一个类别。

[0004] 然而,每个用户写出的同一个字体的字有所不同,同一个用户写的同一个字也不是每次都一样,目前的书法字体的生成模型在生成字体时,同一个字生成的书法字体一模一样,同一个字体生成也是一模一样,多样性比较差。

发明内容

[0005] 本发明实施例提供一种风格化书法的处理方法、装置、设备及介质,用于解决目前的书法字体的生成模型在生成字体时,同一个字生成的书法字体一模一样,同一个字体生成也是一模一样,多样性比较差的问题。

[0006] 第一方面,本发明实施例提供一种风格化书法的处理方法,包括:

[0007] 获取待处理的标准字体图像;

[0008] 将所述标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量;

[0009] 将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量;

[0010] 根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量;

[0011] 将所述风格化编码向量进行解码,得到与所述风格嵌入向量对应风格的目标字体图像。

[0012] 在一种具体的实施方式中,所述将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量,包括:

[0013] 将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向量;

[0014] 将所述编码向量通过结构类激活图模型进行结构分类处理,得到所述结构特征向量。

[0015] 在一种具体的实施方式中,所述根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理之前,所述方法还包括:

[0016] 从预先配置的风格嵌入隐藏向量列表中,获取所述风格嵌入向量,所述风格嵌入向量用于表征目标字体风格的手写字体的风格;其中,所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。

[0017] 在一种具体的实施方式中,所述根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量,包括:

[0018] 将所述风格特征向量和所述结构特征向量进行叠加相乘处理,得到处理后的向量;

[0019] 将所述处理后的向量和所述风格嵌入向量输入所述图片风格迁移模型进行风格转换处理,得到所述风格化编码向量。

[0020] 在一种具体的实施方式中,所述方法还包括:

[0021] 获取多个不同字体风格的手写字体图像;

[0022] 通过对每个字体风格的手写字体图像的风格进行机器学习,得到用于表征每个每个字体风格的手写字体风格的风格嵌入向量。

[0023] 在一种具体的实施方式中,所述方法还包括:

[0024] 通过风格判别器检测所述目标字体图像中的字体风格,获取风格检测分;

[0025] 通过结构判别器检测所述目标字体图像中的字体结构,获取结构检测分;

[0026] 根据所述风格检测分以及所述结构检测分,计算获取生成的所述目标字体图像的图像损失值;

[0027] 根据所述图像损失值对所述图片风格迁移模型的参数进行调整,以使所述图像损失值低于预设损失值阈值。

[0028] 在一种具体的实施方式中,所述方法还包括:

[0029] 获取多次生成的多个所述目标字体图像;

[0030] 根据多个所述目标字体图像,每次输入的高斯噪音向量,生成次数以及预设的平衡因子,确定多样性损失;

[0031] 根据所述多样性损失对所述图片风格迁移模型的参数进行调整。

[0032] 第二方面,本发明实施例提供一种风格化书法的处理装置,包括:

[0033] 获取模块,用于获取待处理的标准字体图像;

[0034] 编码模块,用于将所述标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量;

[0035] 分类模块,用于将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量;

[0036] 转换模块,用于根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量;

[0037] 解码模块,用于将所述风格化编码向量进行解码,得到与所述风格嵌入向量对应风格的目标字体图像。

[0038] 在一种具体的实施方式中,所述分类模块具体用于:

[0039] 将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向

量；

[0040] 将所述编码向量通过结构类激活图模型进行结构分类处理，得到所述结构特征向量。

[0041] 在一种具体的实施方式中，所述转换模块还用于：

[0042] 从预先配置的风格嵌入隐藏向量列表中，获取所述风格嵌入向量，所述风格嵌入向量用于表征所述目标字体风格的手写字体的风格；其中，所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。

[0043] 在一种具体的实施方式中，所述转换模块具体用于：

[0044] 将所述风格特征向量和所述结构特征向量进行叠加相乘处理，得到处理后的向量；

[0045] 将所述处理后的向量和所述风格嵌入向量输入所述图片风格迁移模型进行风格转换处理，得到所述风格化编码向量。

[0046] 在一种具体的实施方式中，所述装置还包括：处理模块，用于：

[0047] 获取多个不同字体风格的手写字体图像；

[0048] 通过对每个字体风格的手写字体图像的风格进行机器学习，得到用于表征每个字体风格的手写字体风格的风格嵌入向量。

[0049] 在一种具体的实施方式中，所述装置还包括：训练模块，用于：

[0050] 通过风格判别器检测所述目标字体图像中的字体风格，获取风格检测分；

[0051] 通过结构判别器检测所述目标字体图像中的字体结构，获取结构检测分；

[0052] 根据所述风格检测分以及所述结构检测分，计算获取生成的所述目标字体图像的图像损失值；

[0053] 根据所述图像损失值对所述图片风格迁移模型的参数进行调整，以使所述图像损失值低于预设损失值阈值。

[0054] 在一种具体的实施方式中，所述训练模块还用于：

[0055] 获取多次生成的多个所述目标字体图像；

[0056] 根据多个所述目标字体图像，每次输入的高斯噪音向量，生成次数以及预设的平衡因子，确定多样性损失；

[0057] 根据所述多样性损失对所述图片风格迁移模型的参数进行调整。

[0058] 第三方面，本发明实施例提供一种电子设备，包括：

[0059] 编码器，处理器，存储器，解码器以及交互接口；

[0060] 所述存储器用于存储所述处理器的可执行指令；

[0061] 其中，所述处理器通过执行所述可执行指令使所述电子设备执行第一方面任一项所述的风格化书法的处理方法。

[0062] 第四方面，本发明实施例提供一种可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现第一方面任一项所述的风格化书法的处理方法。

[0063] 第五方面，本发明实施例提供一种计算机程序产品，包括计算机程序，所述计算机程序被处理器执行时用于实现第一方面任一项所述的风格化书法的处理方法。

[0064] 本发明实施例提供的风格化书法的处理方法、装置、设备及介质，在该方法中，在获取到待处理的标准字体图像后，将标准字体图像和随机生成的高斯噪音向量进行编码处

理,得到编码向量。将编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量。然后根据风格特征向量,结构特征向量以及风格嵌入向量,采用图片风格迁移模型模型进行风格转换处理,得到风格化编码向量,再将风格化编码向量进行解码,得到与风格嵌入向量对应风格的目标字体图像。通过引用风格嵌入向量使得生成的字体图像更贴近风格嵌入向量所对应的用户的风格,另外通过引入高斯噪声向量可以提高生成字体的多样性,更贴近用户撰写的实际情况。

附图说明

[0065] 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一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0066] 图1为本发明实施例提供的风格化书法的处理方法实施例一的流程图;

[0067] 图2为本发明实施例提供的风格化书法的处理方法实施例二的流程图;

[0068] 图3为本发明实施例提供的风格化书法的处理方法实施例三的流程图;

[0069] 图4a为本发明一实例提供的风格化书法的生成装置的示意图;

[0070] 图4b为本发明一实例提供的风格化书法的生成效果示意图;

[0071] 图5为本发明实施例提供的风格化书法的处理装置实施例一的结构示意图;

[0072] 图6为本发明实施例提供的风格化书法的处理装置实施例二的结构示意图;

[0073] 图7为本发明实施例提供的风格化书法的处理装置实施例三的结构示意图;

[0074] 图8为本发明实施例提供的电子设备实施例的结构示意图。

具体实施方式

[0075] 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根据本实施例的启示下作出的所有其他实施例,都属于本发明保护的范围。

[0076] 本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”“第四”等(如果存在)是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本发明的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0077] 目前的几种成熟的字体生成模型主要存在以下一些问题:

[0078] 1),生成的字体质量有待提升,现有技术提供的Zi2zi模型,CycleGan模型以及U-GAT-IT模型等几种模型都是在较为规则的数据集上做的实验。而在比较难的手写字体上就会有明显的劣势。一般来说,zi2zi模型的效果是最差的,CycleGan次之。

[0079] 2),多样性是书法生成的一个基本要求,这三种模型都没考虑字体生成的多样性,

这在书法生成不符合用户手写字体的一般规律。

[0080] 3),不同的人写出的字是不同的,因此模型生成的字体也应该根据能适用多种风格的输出。CycleGan和U-GAT-IT都无法做到这一点,虽然zi2zi模型能够生成多风格的字体,但是必须使用成对的数据,在复杂手写字体生成的质量也不够高,这严重限制了其应用场景。

[0081] 综上所述,目前现有技术中的字体生成方案中,同一个标准字生成的书法字体一模一样,同一个字体生成的也是一模一样,多样性比较差,导致不能更贴近实际手写字体,限制了应用场景。

[0082] 针对上述问题,本发明提出一种风格化书法的生成方案,能够生成高质量,多样性较高,且贴近用户风格的书法字体。

[0083] 本发明技术方案的核心思路:发明人在对生成书法的模型的研究过程中发现,对于汉字来说两个主要的特征就是结构和风格,即文字的结构以及不同字体的风格。现在的生成模型中均未将这两个特征作为注意力进行考虑,另外,一般的生成模型如果只输入标准字体,那么经过处理运算,必然输出的是一样的字体,为了能够更贴近普通用户撰写的实际情况,可以对模型的输入引入一定的随机噪音,在不影响整体风格和结构的情况下,生成不完全一样的字体。同时考虑不同的用户撰写同一个字体也有一定的区别,可以通过机器学习的方式获取到不同的用户的风格化特征,也引入生成字体时的输入,进一步提高字体生成的多样性。

[0084] 本发明提供的风格化书法的处理方法,可以应用在能够进行数据处理功能的计算机设备,云服务器,服务器,笔记本电脑,智能手机,电脑等电子设备中,对于具体的设备形态,本方案不做限制。

[0085] 下面通过几个具体实施例对本发明提供的风格化书法的处理方法进行详细说明。

[0086] 图1为本发明实施例提供的风格化书法的处理方法实施例一的流程图,如图1所示,该风格化书法的处理方法包括以下步骤:

[0087] S101:获取待处理的标准字体图像。

[0088] 在本步骤中,如果需要生成某个类型的书法或者某个用户风格化的书法字体,首先需要获取到待生成的文字图像,也就是这里的标准字体图像。标准字体图像中可以包括一个或者多个预设的宋体,黑体等标准字体的文字。这里的标准字体的文字一般指汉字。

[0089] S102:将标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量。

[0090] 在本步骤中,为了能够更贴近用户撰写的实际情况,同一个用户撰写同一个文字的时候也会出现一定的差异,不太可能是一模一样的,因此在对标准字体图像进行编码时,本方案中引入了高斯噪音向量,该向量是随机生成的噪音向量,主要是用于对生成的字体产生多样性的作用。

[0091] 将标准字体图像和上述高斯噪音向量输入一个编码器,例如可以采用多层卷积神经网络(Convolutional Neural Networks,简称:CNN)进行编码,得到编码向量。

[0092] S103:将编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量。

[0093] 具体的,将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向量;将所述编码向量通过结构类激活图模型进行结构分类处理,得到所述结构特

征向量。

[0094] 在本方案中,为了更进一步的提高文字的风格化生成的准确度,提高生成字体的效果,在该方案中引入结构和风格两个注意力,对于输入的编码向量,可以通过风格类激活图模型进行分类处理得到风格特征向量,对编码向量通过结构类激活图模型进行分类处理得到结构特征向量。在进行分类处理的过程中,在风格类激活图模型中的分类器输出风格权重数据,也就是风格注意力热图(在该方案中,类激活图(CAM)也称为注意力热图)后,与输入的图像(即编码向量)进行叠加相乘得到对应的风格特征向量。同样的,在结构类激活图模型中的分类器输出结构权重数据,也就是结构注意力热图后,与输入的编码向量进行叠加相乘得到对应的结构风格特征向量。

[0095] 本方案中使用了汉字的风格注意力和结构注意力来提升生成效果是本案的核心。

[0096] S104:根据风格特征向量,结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量。

[0097] 其中,该图片风格迁移模型可以是AdaIN模型或CLIN模型。CLIN模型是在AdaIN基础上通过风格隐向量控制风格编码的处理模型。

[0098] 在本步骤中,在通过分类之后,得到了用来表征风格的风格特征向量和表征字体结构的结构特征向量之后,则基于这两个特征向量以及预先设置的用于表征字体风格的结构特征向量,采用预先训练好的图片风格迁移模型进行风格转换处理以得到最终的风格化编码向量。在该方案中,如果该图片风格迁移模型为CLIN模型,应理解该CLIN模型是在AdaIN基础上,通过一个隐向量 α 、 β 实现的风格编码,来控制风格转换的处理模型,这里引入的两个隐向量是用来体现用户撰写风格的向量,也就是前述的风格嵌入向量。在具体应用过程中,可以通过机器学习的方式对每个用户手写的字体进行学习,得到每个用户的不同字体风格对应的风格嵌入向量。

[0099] 在一种具体实现方式中,电子设备可以从预先配置的风格嵌入隐藏向量列表中,获取前述的风格嵌入向量,所述风格嵌入向量用于表征目标字体风格的手写字体的风格;其中,所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。

[0100] 因此在具体进行特定用户或者特定字体风格的结构化字体生成时,需要指定目标字体风格,以便在生成过程中,能够从风格嵌入隐藏向量列表中获取到该目标字体风格对应的风格嵌入向量,以输入图片风格迁移模型中进行处理。

[0101] 本方案中的风格嵌入隐藏向量列表是预先配置的,其中包括用于表征多个字体风格的结构特征向量,这些字体风格的结构特征向量是预先根据机器学习得到的。也就是说,需要预先在该电子设备中或者其他具有数据处理分析功能的设备中,获取多个不同字体风格的手写字体图像,并通过对每个字风格的手写字体图像的风格进行机器学习,得到用于表征每个字体风格的手写字体风格的结构特征向量,并将全部风格嵌入向量存储在风格嵌入隐藏向量列表中,并采用用户的身份标识或者字体风格的结构标识等唯一标识进行区别。即在本步骤执行之前,需要预先获取多个不同字体风格的手写字体图像,然后针对每个字体风格,将其手写字体图像采用机器学习的方式进行分析处理,得到用于表征该字体风格的手写风格的结构特征向量。

[0102] 在另一种具体实现中,将所述风格特征向量和所述结构特征向量进行叠加相乘处理,得到处理后的向量。然后将所述处理后的向量和所述目标用户的风格嵌入向量输入所

述图片风格迁移模型进行风格转换处理,得到所述风格化编码向量。

[0103] 在该实现中,分类处理完得到了风格特征向量和结构特征向量两个特征向量,后续处理过程中可以将两个向量相乘得到一个向量,然后再和目标用户的风格嵌入向量一并输入图片风格迁移模型进行风格化处理。

[0104] S105:将风格化编码向量进行解码,得到与风格嵌入向量对应风格的目标字体图像。

[0105] 最后,在引入了高斯随机向量,以及目标字体风格的风格嵌入向量进行了风格转换之后,得到风格化编码之后,输入解码器进行解码,从而得到最后转换后的与风格嵌入向量对应的目标字体图像。

[0106] 本实施例提供的风格化书法的处理方法,通过在生成字体的过程中,引入随机生成的高斯噪音向量,来增加生成的字体的多样性,并使用不同的类激活图模型,即风格化类激活图模型和结构类激活图模型,得到风格化的注意力和结构注意力,通过集成的方式显著提升模型效果。同时在具体的生成过程中,引入了不同用户的风格嵌入向量,增加个人风格化影响,进一步提高生成字体的多样性。

[0107] 图2为本发明实施例提供的风格化书法的处理方法实施例二的流程图,如图2所示,在上述实施例的基础上,该风格化书法的处理方法中CLIN模型的训练过程与应用过程相似,在训练过程中该风格化书法的处理方法还包括以下步骤:

[0108] S201:通过风格判别器检测目标字体图像中的字体风格,获取风格检测分。

[0109] 本方案主要应用在对模型进行训练的过程中,为了进一步提高图片风格迁移模型生成的字体的质量,需要在训练过程中,对于生成字体的结果进行判别,本方案中提供了两个判别器,分别用来对生成的字体的结构和风格进行判别,以确定生成的字体的质量,如果质量不佳则可以反馈至前端生成过程中,对图片风格迁移模型进行优化。

[0110] 在本步骤中,通过设置的风格判别器对生成的字体图像的字体风格进行检测,可以得到风格检测分,该风格检测分用来指示该字体图像中字体是否为真,即是否贴近目标用户或者目标字体风格真实撰写的字体风格。

[0111] S202:通过结构判别器检测目标字体图像中的字体结构,获取结构检测分。

[0112] 在本步骤中,通过设置的结构判别器检测生成的字体图像的字体结构,得到一个结构检测分,该结构检测分用来指示字体结构是否正确。

[0113] S203:根据风格检测分以及结构检测分,计算获取生成的目标字体图像的图像损失值。

[0114] 在本步骤中,基于检测出来的风格检测分和结构检测分可以通过预设的损失函数,计算出手写的真实字体与生成的字体之间的图像损失值,一般来说如果非常相似,那么图像损失值则接近或者等于0,若图像损失值远大于0,则可以确定出真实字体与生成的字体图像不够接近,需要对图片风格迁移模型的参数进行调整。

[0115] S204:根据图像损失值对图片风格迁移模型的参数进行调整,以使图像损失值低于预设损失值阈值。

[0116] 在本步骤中,在得到图像损失值之后,根据图像损失值的大小确定是否对图片风格迁移模型进行进一步调整,如果图像损失值等于0或者与0之间的差值小于预设范围,也就是说图像损失值低于预设损失值阈值,那么则可以确定为该图片风格迁移模型生成的字

体已经与手写的字体无限接近,可以直接采用该图片风格迁移模型进行对应的风格化字体生成。然而,如果图像损失值不为0或者与0之间的差值大于预设范围,也就是图像损失值大于预设损失值阈值,则可以确定为该图片风格迁移模型生成的字体已经与手写的字体之间还存在一定的差距,则需要对图片风格迁移模型的参数进行进一步调整,以使生成的字体更贴近真实字体。

[0117] 在实际训练过程中,每次对图片风格迁移模型的参数进行调整之后,可以再次重复前述实施例一和实施例二中的所有步骤,再次进行判断,直至图像损失值低于预设损失值阈值,则得到用于生成风格化书法的图片风格迁移模型。

[0118] 图3为本发明实施例提供的风格化书法的处理方法实施例三的流程,如图3所示,在上述实施例的基础上,该风格化书法的处理方法还包括以下步骤:

[0119] S301:获取多次生成的多个目标字体图像。

[0120] 在本步骤中,该多个目标字体图像指的是对同一个文字生成的多个字体图像。

[0121] S302:根据多个目标字体图像,每次输入的高斯噪音向量,生成次数以及预设的平衡因子,确定多样性损失。

[0122] S303:根据多样性损失对图片风格迁移模型的参数进行调整。

[0123] 在本方案中,为了能够使风格化书法生成的字体比较多样,即对于同一个文字生成的同一个风格的字体也有所不同,本方案还设置了用于计算多样性损失的损失函数,在该损失函数中需要输入多个目标字体图像,每次生成时候采用的高斯噪音向量,也就是多个高斯噪音向量,生成次数以及预设的平衡因子,计算得到一个多样性损失的数值。一般来说该多样性损失越小越好,可以根据实验情况预先设置门限值,在多样性损失没有小于预设的门限值之前,可以对图片风格迁移模型的参数进行调整直至多样性损失小于预设的门限值。

[0124] 综上所述,本发明提出基于类激活图的书法生成方案,能够生成高质量、多样性、风格化的书法。具体的整个书法生成方案的过程中,使用不同的类激活图模型,即风格化类激活图模型和结构类激活图模型,得到风格化的注意力热图(即风格特征向量)和结构注意力热图(即结构特征向量),通过集成的方式显著提升生成字体的效果,该部分是本发明技术方案的核心。为了该方案中采用的模型能使用非成对的数据训练学习多风格转换,提出了图片风格迁移模型(例如CLIN模型),能够生成多种不同的风格。为了解决图片生成多样性的问题,本方案还提出了一个控制多样性生成的损失函数,让模型每次都生成不一样的字体。

[0125] 在上述任一实施例的基础上,下面通过一具体实施例对本发明提供的风格化书法的处理方案进行说明。

[0126] 图4a为本发明一实例提供的风格化书法的生成装置的示意图,如图4a所示,本发明技术方案提供的风格化书法生成过程,输入的是标准字体(例如宋体或楷体,也就是前述的标准字体图像),如图中所示的(人...月)。输入的同时还混合了一个随机生成的噪音,即高斯噪音向量,该高斯噪音向量是为了产生多样性。输入的标准字体图像经过一个编码器(多层CNN)得到一个编码向量。这个编码向量会经过两个CAM(类激活图)模块,它相当于分类器,分别对编码向量进行风格分类和结构分类处理。将分类器输出的权重数据和原来的图片叠加相乘就得到一个和该类别相关的图片注意力热图,这个热图就是和类别最相关的

注意力(由于这个注意力是在分类的基础上得到的,因此这种注意力热图又叫类激活图,也就是前述方案中的风格特征向量和结构特征向量)。本方案中使用了汉字风格注意力和汉字结构的注意力两种集成的形式来提升效果,也是本发明技术方案的核心技术点。

[0127] 此外,本方案的风格化书法的生成装置中间会有一个CLIN的结构,这是在AdaIN和CIN基础上,通过一个隐向量 α 、 β 实现的风格编码,来控制风格的转换。接下来是一个解码器将这些CLIN出来的向量解码成一张图,解码器主要也是由CNN组成。

[0128] 随后会有两个判别器:结构判别会判断这个字的类别是否和目标结构一致,风格判别会判断该风格是否和目标风格一致。判别器的判断信号会反馈给前面的字体生成器(该生成器中包括编码器,两个分类器,CLIN模型以及解码器部分),这样生成器就会生成结构和风格更加接近目标字体风格的字体图像。

[0129] 综合上述,可知在本发明技术方案提供的风格化字体生成方案中,编码器和解码器是对抗生成网络的生成器最为基本的结构,编码器是将图像编码成一个向量,同理解码器是将一个向量还原回图像尺寸大小,这里编码器和解码器都是由4层的卷积神经网络组成。生成器主要用来生成质量更好的图像。在训练过程中,判别器也是对抗生成网络最基本的结构模块,这里也是有4层卷积神经网络组成;主要用来判别哪些图像是生成的,哪些是真实的,反过来生成器会尽量去生成让判别器无法判别的高质量图像,即无法判别生成的图像是生成的,确定为真实的时,整个生成器的过程的效果达到最佳。该生成器的训练过程汇总,通过这样不断的迭代学习,就形成了一种对抗式的网络结构。

[0130] 为了能够生成高质量、多样性、风格化的书法,本方案中的书法生成方案有以下几个创新点:

[0131] 1)、结构类激活图和风格类激活图(也称为结构类激活图模型和风格类激活图模型)

[0132] 汉字的基本元素可以分成风格和结构,因此从这两个方面入手,构建了两种注意力机制。

[0133] 由于模型训练时,模型会有一个反向生成的过程,因此几种类别的图像都可能经过生成器的编码器,因此这里就加入了一个风格CAM的判别,相应的也会得到一个风格的类激活图模型,对输入的图像进行处理可得到风格特征向量,这个激活图就代表模型捕捉到的当前风格的最重要或者需要学习的部分,从而提升生成的风格转换质量。

[0134] 结构上汉字可以分成5种结构,上下结构:要、星,左右结构:部、补,半包围结构:包、风,全包围结构:国、图,独立结构:一、上。文字输入到模型中时,通过自行的判别会得到一个结构类激活图模型,对输入的图像进行处理可结构特征向量),该激活图可以保证字的结构。

[0135] 为保证生成的字体必须具备清晰的字形结构,在整个生成器的后面还加了一个额外的结构判别器来保证字形。

[0136] 这样一来,通过两种类激活图,书法字形和风格的生成质量都能被保证。

[0137] 2)、CLIN模型控制生成风格

[0138] CLIN模型是本方案提出用于控制生成风格的模型,是图片风格迁移模型的一种具体实现模型,该CLIN模型是在AdaIN的基础上优化改进得来的,这是一种风格转换方法。AdaIN最初用一张目标图片来控制图片的生成风格,其原理就是目标图片压缩成一个 α 和 β

向量,将这两个向量输入到AdaIN模块中就可以生成特定风格的图片了。受此启发,本方案中初始化了多组 α 和 β 向量,每一组向量对代表一种特定的风格,也可以代表一个特征的用户撰写风格。AdaIN的主要在Instance Normalization基础上改进而来,本方案的CLIN模块也加入了这种Normalization的设计。

[0139] 3)、结构和风格判别器以及多样性损失函数

[0140] 如前面所述,生成器生成了具有风格化和结构清晰的字体图像,这里我们就相应加入两个判别器来判断生成器的质量,如果生成质量不佳就反馈给生成器继续学习。在该方案中,应理解,判别器会对生成的字体进行分类,判断是模型生成的还是真实撰写的,同样的生成器会尽可能使生成的字体图像以假乱真。这就形成了对抗生成网络。质量不佳就是判别器判断出图片是模型生成的而不是真实的。注:判别器输出的是对图片的一个打分,下面的几个判别器D分别代表了不同判别器。

[0141] 损失函数0 (对抗生成损失,每个gan网络的基本损失):

[0142] $L = \log D(y) + \log(1 - D(Gf(x, z_i, e)))$

[0143] 该损失函数0包含两部分,前一部分是将真实图像y输入到判别器,判别要判断出y是真实的,即输出的分数为1,后半部分是对生成图像D(Gf(x, z_i, e))的判别,即应该输出分数为0。为了理解这一公式,理想情况下,判别器打分都答对了, $\log 1 + \log(1 - 0) = 0$,这时候损失值L就是0,或者与0之间的差值小于一个预设值,也就是说图像损失值低于预设的损失值阈值,则对于生成器就不需要调整;反之,如果L不为0或者远大于0,也就是说图像损失值高于损失值阈值,就应该调整生成器,因为生成器生成的图像效果还不够好。

[0144] 损失函数1: $L = \log D(c | Gf(x, z_i, e))$ 。

[0145] 该损失函数1用于多风格损失和结构判别(损失都一样),要判断出生成的图像属于哪个类别(这里假设有c个类别),如果在本应该的类别上分数较低就要优化前面的生成器。

[0146] 损失函数2(多样性损失):

$$[0147] \quad L = \sum_{i=1}^K \frac{1}{K \|Gf(x, z_i, e) - Gf(x, z_j, e)\| + 1 + \epsilon}$$

[0148] 该损失函数2是为了生成多向性的字体提出的。 G_f 代表的是生成器, Z (i和j代表不同的噪音向量)是模型和图像一起输入的一个高斯噪音向量。由于输入的向量每次都不相同,因此这里相当于输入K个不同的高斯造影向量,K为表示数量的正整数,同样的输出也是K张字体图像。为了让这K张不同的字体图像尽量不同。整个字体生成装置就会将L损失朝着更小的方向优化。相应的,分母上生成的不同字体之间的距离也会越来越大,这样一来字体之间的就产生有多样性,即同一张输入的标准字体图像会产生不同的输出结果,即不同的生成的字体图像。该损失函数2的公式上加了一个 ϵ 作为平衡因子,是为了防止损失函数2分母为0的情况。

[0149] 在上述实施例中,应理解,生成器(也可以称为生成网络)至少包括图4a所示的装置中编码器,风格和结构的分类器,以及CLIN模型以及解码器。在根据上述任损失函数的值对生成器进行优化时,不仅可以调整CLIN模型中的参数,也可以对多层CNN网络进行梯度更新,也可以同时调整CLIN模型的参数和CNN网络梯度,以使损失降低,对此本方案不做限制。

前述的方法实施例中只是以对CLIN模型的参数调整为例进行举例说明,并不限定只能调整模型的参数。

[0150] 图4b为本发明一实例提供的风格化书法的生成效果示意图,如图4b所示,采用本发明实施例提供的技术方案中,同一个字“宿”可以生成多个风格化的字体图像。图中右侧的上面两个字体分别代表不同的风格,可以看到不同字体的风格差距明显。图中右侧下面两个字代表同一字体的多样性生成,可以看到确实产生了风格的改变。

[0151] 综上所述,本发明提供的风格化书法的处理方案,在进行风格化书法字体的生成过程中,引入随机生成的高斯噪音向量,来增加生成的字体的多样性,并使用不同的类激活图模型,即风格化类激活图模型和结构类激活图模型,得到风格化的注意力和结构注意力,通过集成的方式显著提升模型效果。同时在具体的生成过程中,引入了不同用户的风格嵌入向量,增加用户个人风格化影响,在提高了风格化,生成质量的同时,进一步提高生成字体的多样性。

[0152] 图5为本发明实施例提供的风格化书法的处理装置实施例一的结构示意图,如图5所示,该风格化书法的处理装置10包括:

[0153] 获取模块11,用于获取待处理的标准字体图像;

[0154] 编码模块12,用于将所述标准字体图像和随机生成的高斯噪音向量进行编码处理,得到编码向量;

[0155] 分类模块13,用于将所述编码向量分别通过风格类激活图模型和结构类激活图模型进行分类处理得到风格特征向量和结构特征向量;

[0156] 转换模块14,用于根据所述风格特征向量,所述结构特征向量以及风格嵌入向量,采用图片风格迁移模型进行风格转换处理,得到风格化编码向量;

[0157] 解码模块15,用于将所述风格化编码向量进行解码,得到与所述风格嵌入向量对应风格的目标字体图像。

[0158] 本实施例提供的风格化书法的处理装置,用于执行前述任一方法实施例提供的技术方案,其实现原理和技术效果类似,通过引用用户的风格嵌入向量使得生成的字体图像更贴近手写的风格,另外通过引入高斯噪音向量可以提高生成字体的多样性,更贴近用户撰写的实际情况。

[0159] 在上述实施例的基础上,所述分类模块13具体用于:

[0160] 将所述编码向量通过风格类激活图模型进行风格分类处理,得到所述风格特征向量;

[0161] 将所述编码向量通过结构类激活图模型进行结构分类处理,得到所述结构特征向量。

[0162] 在一种具体的实施方式中,所述转换模块14还用于:

[0163] 从预先配置的风格嵌入隐藏向量列表中,获取所述风格嵌入向量,所述风格嵌入向量用于表征目标字体风格的手写字体的风格;其中,所述风格嵌入隐藏向量列表中包括多种字体风格对应的风格嵌入向量。

[0164] 在一种具体的实施方式中,所述转换模块14具体用于:

[0165] 将所述风格特征向量和所述结构特征向量进行叠加相乘处理,得到处理后的向量;

[0166] 将所述处理后的向量和所述风格嵌入向量输入所述图片风格迁移模型进行风格转换处理,得到所述风格化编码向量。

[0167] 图6为本发明实施例提供的风格化书法的处理装置实施例二的结构示意图,如图6所示,在上述实施例的基础上,该风格化书法的处理装置10还包括:处理模块16,用于:

[0168] 获取多个不同字体风格的手写字体图像;

[0169] 通过对每个字体风格的手写字体图像的风格进行机器学习,得到用于表征每个每个字体风格的手写字体风格的风格嵌入向量。

[0170] 图7为本发明实施例提供的风格化书法的处理装置实施例三的结构示意图,如图7所示,在上述实施例的基础上,该风格化书法的处理装置10还包括:训练模块17,用于:

[0171] 通过风格判别器检测所述目标字体图像中的字体风格,获取风格检测分;

[0172] 通过结构判别器检测所述目标字体图像中的字体结构,获取结构检测分;

[0173] 根据所述风格检测分以及所述结构检测分,计算获取生成的所述目标字体图像的图像损失值;

[0174] 根据所述图像损失值对所述图片风格迁移模型的参数进行调整,以使所述图像损失值低于预设损失值阈值。

[0175] 进一步的,在另一种具体实现方式中,所述训练模块17还用于:

[0176] 获取多次生成的多个所述目标字体图像;

[0177] 根据多个所述目标字体图像,每次输入的高斯噪音向量,生成次数以及预设的平衡因子,确定多样性损失;

[0178] 根据所述多样性损失对所述图片风格迁移模型的参数进行调整。

[0179] 本实施例提供的风格化书法的处理装置,用于执行前述任一方法实施例提供的技术方案,其实现原理和技术效果类似,在此不再赘述。

[0180] 图8为本发明实施例提供的电子设备实施例的结构示意图,如图8所示,该电子设备100,包括:

[0181] 处理器111,存储器112,编码器113,解码器114以及交互接口115;

[0182] 所述存储器112用于存储所述处理器111的可执行指令;

[0183] 其中,所述处理器11通过执行所述可执行指令使所述电子设备100执行前述任一方法实施例提供的风格化书法的处理方法的技术方案。

[0184] 可选的,存储器112既可以是独立的,也可以跟处理器111集成在一起。

[0185] 可选的,当所述存储器112是独立于处理器111之外的器件时,所述电子设备100还可以包括:

[0186] 总线,用于将上述器件连接起来。

[0187] 该电子设备用于执行前述任一方法实施例中的技术方案,其实现原理和技术效果类似,在此不再赘述。

[0188] 本发明实施例还提供一种可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现前述任一方法实施例提供的风格化书法的处理方法的技术方案。

[0189] 本发明实施例还提供一种计算机程序产品,包括计算机程序,所述计算机程序被处理器执行时用于实现前述任一方法实施例提供的风格化书法的处理方法的技术方案。

[0190] 本领域普通技术人员可以理解:实现上述各方法实施例的全部或部分步骤可以通

过程序指令相关的硬件来完成。前述的程序可以存储于一计算机可读取存储介质中。该程序在执行时,执行包括上述各方法实施例的步骤;而前述的存储介质包括:ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0191] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或对其部分或全部技术特征进行等同替换;而这些修改或替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

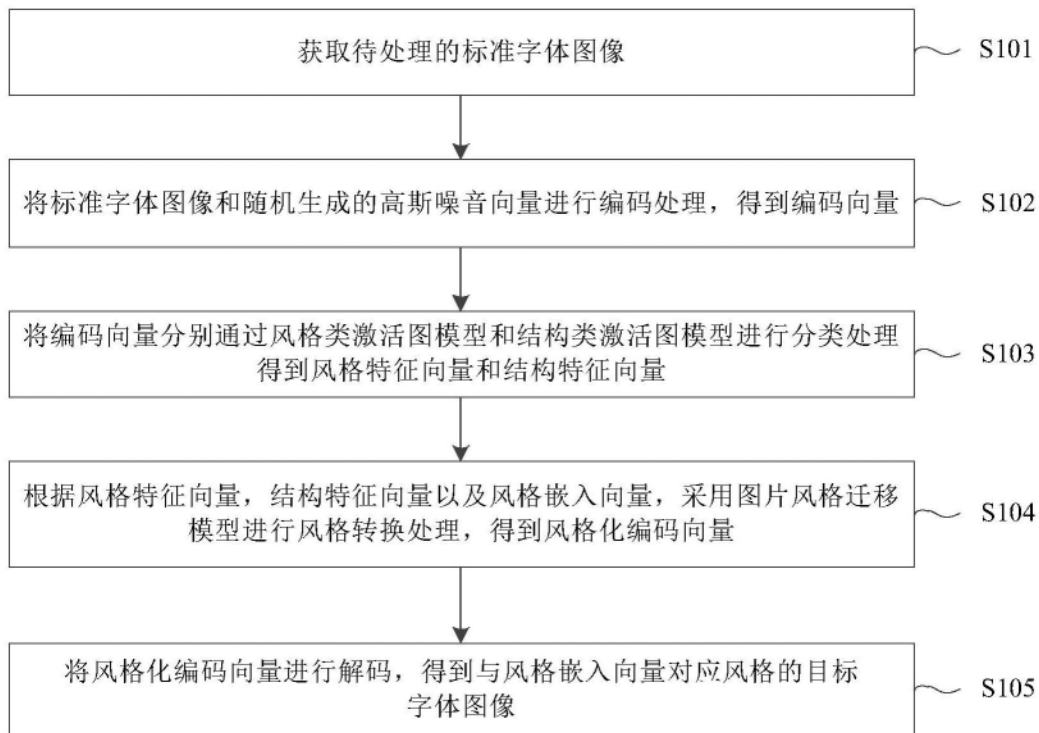


图1

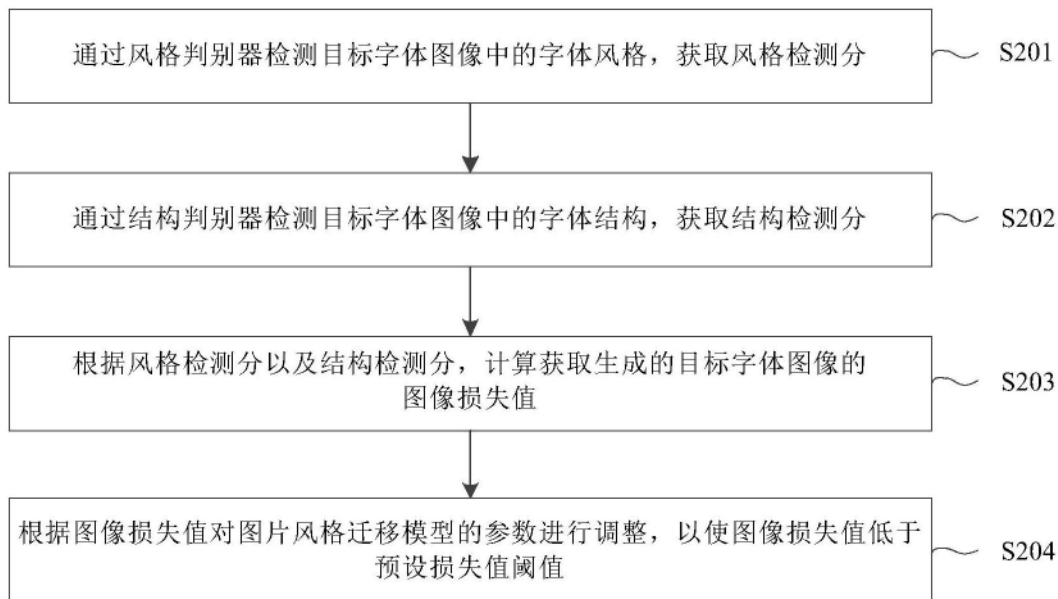


图2

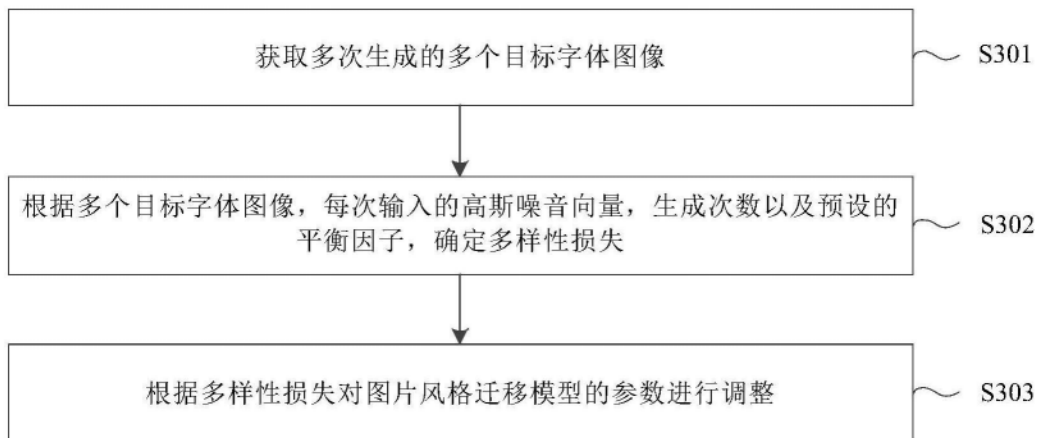


图3

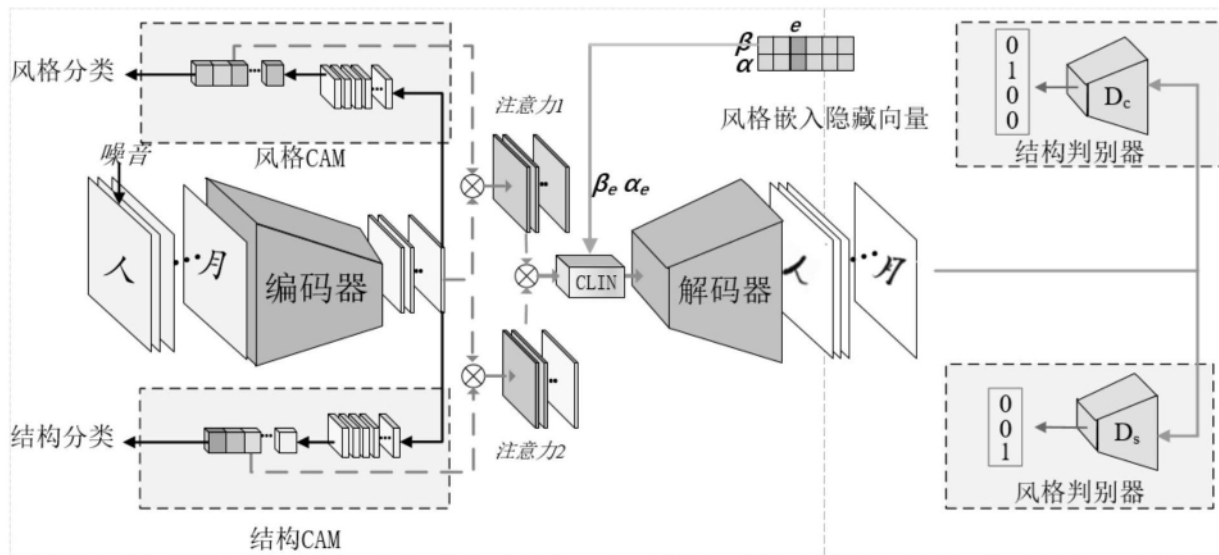


图4a

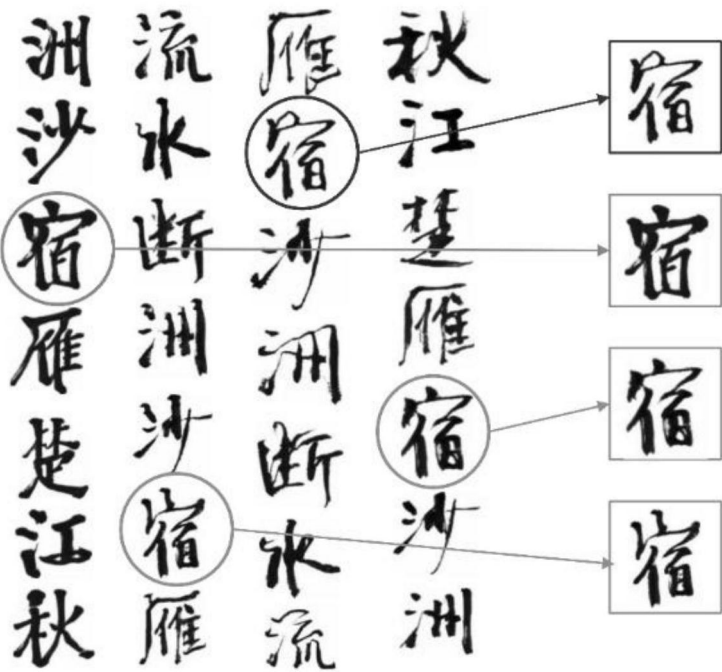


图4b

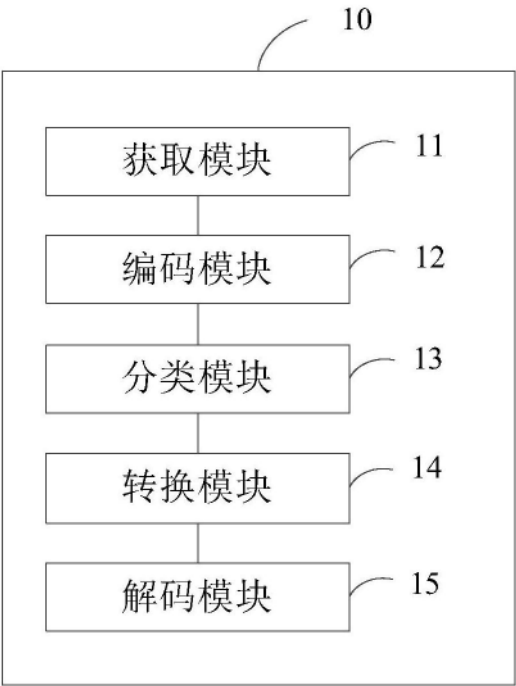


图5

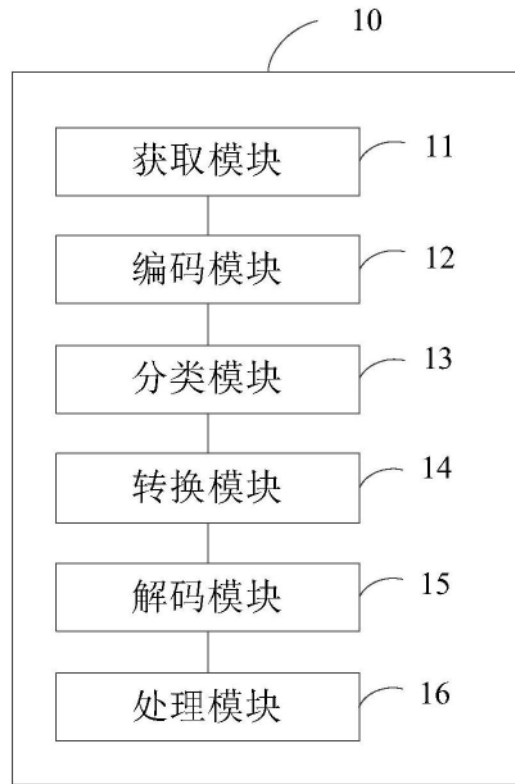


图6

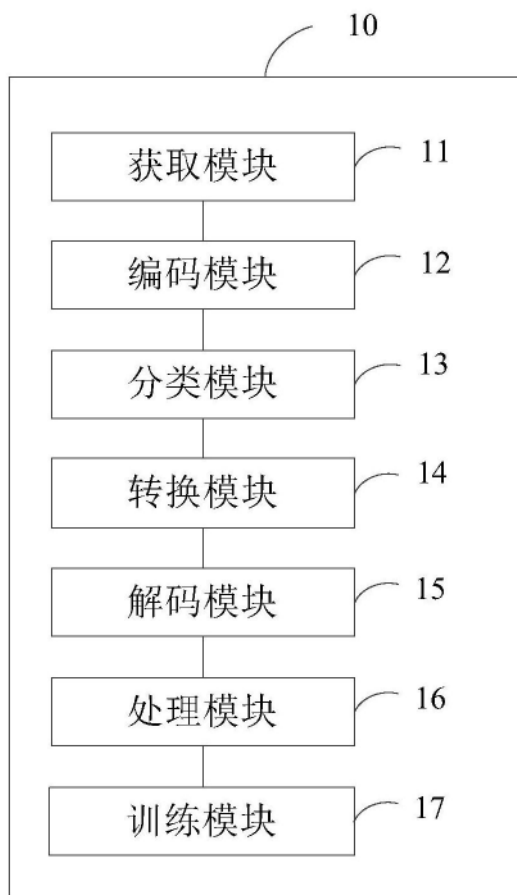


图7

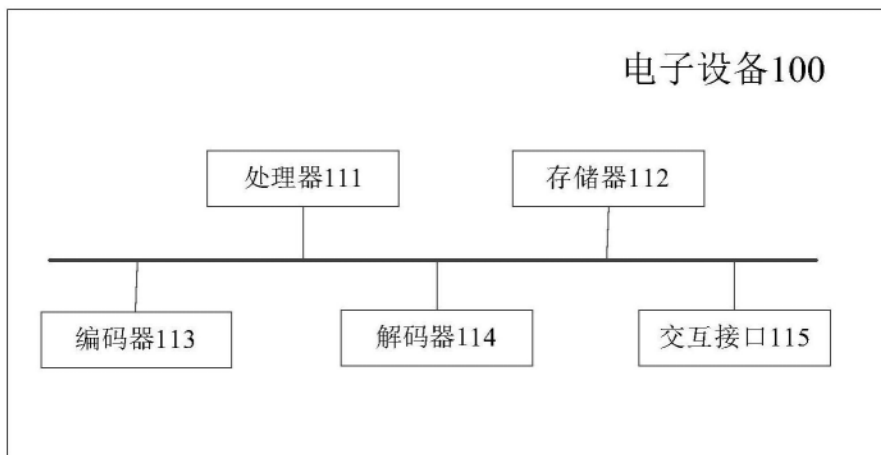


图8