



(12) 发明专利申请

(10) 申请公布号 CN 112861806 A

(43) 申请公布日 2021.05.28

(21) 申请号 202110288312.7

(22) 申请日 2021.03.17

(71) 申请人 网易(杭州)网络有限公司

地址 310052 浙江省杭州市滨江区长河街
道网商路599号4幢7层

(72) 发明人 温琦 袁焱 范长杰 胡志鹏

(74) 专利代理机构 北京超凡宏宇专利代理事务
所(特殊普通合伙) 11463

代理人 钟扬飞

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

G06N 20/00 (2019.01)

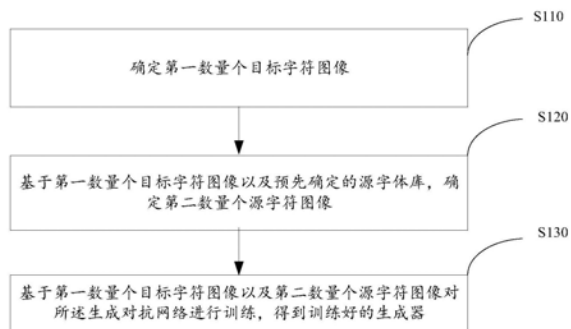
权利要求书3页 说明书12页 附图4页

(54) 发明名称

基于生成对抗网络的字体数据处理方法及装置

(57) 摘要

本发明提供了一种基于生成对抗网络的字体数据处理方法及装置。通过确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。以此可以实现使用少量的目标字符图像既可以使模型具备生成目标字体的能力,提升了模型的应用场景,提高了用户体验。



1. 一种基于生成对抗网络的字体数据处理方法,其特征在于,包括:

确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;

基于所述第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;

基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,包括:

基于第一数量个目标字符图像,从预先确定的源字体库中渲染得到与所述目标字符图像对应相同的字符的第三数量个源字符图像;

从预先确定的源字体库中随机渲染得到第四数量个源字符图像;所述第二数量个源字符图像由所述第三数量个源字符图像和所述第四数量个源字符图像构成。

3. 根据权利要求2所述的方法,其特征在于,所述第三数量与所述第四数量均等于所述第一数量;

所述第一数量个目标字符图像和所述第三数量个源字符图像构成第一图像集,所述第一图像集包括多个第一图像对,每个所述第一图像对包括同一字符对应的目标字符图像和源字符图像;

所述第一数量个目标字符图像和所述第四数量个源字符图像构成第二图像集,第二图像集包括多个第二图像对,每个所述第二图像对包括不同字符对应的目标字符图像和源字符图像。

4. 根据权利要求3所述的方法,其特征在于,所述基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,包括:

基于所述第一图像集和所述第二图像集对所述生成对抗网络进行训练,得到训练好的生成器。

5. 根据权利要求4所述的方法,其特征在于,所述基于所述第一图像集和所述第二图像集对所述生成对抗网络进行训练,得到训练好的生成器,包括:

基于所述第一图像集以及第一损失函数对所述生成对抗网络进行第一训练;

基于所述第二图像集以及第二损失函数对所述生成对抗网络进行第二训练;

迭代的交替进行所述第一训练以及第二训练直至符合终止条件,输出训练好的生成器。

6. 根据权利要求5所述的方法,其特征在于,所述第一损失函数基于对抗损失、一致性损失以及对齐损失确定;所述第二损失函数基于所述对抗损失、所述一致性损失以及风格损失确定;其中,所述对抗损失用于指示分布以及分类上的损失,所述一致性损失用于指示循环能力以及字符属性上的损失,所述对齐损失用于指示内容和特征上的损失,所述风格损失用于指示特征分布上的损失。

7. 根据权利要求6所述的方法,其特征在于,所述对抗损失基于主要部分损失和辅助部分损失确定;

所述主要部分损失基于如下公式确定：

$$L_{adv}^{x \rightarrow y} = \mathbb{E}_y [(D_t(y))^2] + \mathbb{E}_x [(1 - D_t(G_s(E_s(x))))^2]$$

所述辅助部分损失基于如下公式确定：

$$L_{cam}^{x \rightarrow y} = \mathbb{E}_y [(\eta_{D_t}(y))^2] + \mathbb{E}_x [(1 - \eta_{D_t}(G_s(E_s(x))))^2]$$

其中，L表示损失函数，x为源域，y为目标域，adv表示对抗，cam表示类激活图，E表示期望， D_t 表示判别器， E_s 表示从源域到目标域的编码器， G_s 表示从源域到目标域的解码器， $\eta_{D_t}(\cdot)$ 表示判别器中添加的辅助分类器， E_s 和 G_s 共同构成第一子生成器，所述训练好的生成器包括所述第一子生成器。

8. 根据权利要求6所述的方法，其特征在于，所述一致性损失包括循环损失和身份损失；

所述循环损失基于如下公式确定：

$$L_{cycle}^{x \rightarrow y} = \mathbb{E}_x [|x - G_t(E_t(G_s(E_s(x))))|_1]$$

所述身份损失基于如下公式确定：

$$L_{identity}^{x \rightarrow y} = \mathbb{E}_y [|y - (G_s(E_s(y)))|_1]$$

其中，L表示损失函数，x为源域，y为目标域，cycle表示循环，identity表示身份，E表示期望， E_s 表示从源域到目标域的编码器， G_s 表示从源域到目标域的解码器， E_t 表示从目标域到源域的编码器， G_t 表示从目标域到源域的解码器， E_t 和 G_t 共同构成第二子生成器，所述训练好的生成器包括所述第二子生成器。

9. 根据权利要求6所述的方法，其特征在于，所述对齐损失包括内容损失和特征损失；

所述内容损失基于如下公式确定：

$$L_{L1}^{x \rightarrow y} = \mathbb{E}_{x,y} [|y - G_s(E_s(x_p))|_1]$$

所述特征损失基于如下公式确定：

$$L_{constancy}^{x \rightarrow y} = \mathbb{E}_{x,y} [|E_t(y) - E_t(G_s(E_s(x_p)))|_2]$$

其中，L表示损失函数，x为源域，y为目标域， L_1 表示内容，constancy表示不变性，E表示期望， x_p 表示配对数据， E_s 表示从源域到目标域的编码器， G_s 表示从源域到目标域的解码器， E_t 表示从目标域到源域的编码器。

10. 根据权利要求6所述的方法，其特征在于，所述风格损失基于如下公式确定：

$$L_{style}^2 = \|\mathbb{E}_y [\phi(E_t(y))] - \mathbb{E}_x [\phi(E_t(G_s(E_s(x_r))))]\|_{H_k}^2$$

其中，L表示损失函数，x为源域，y为目标域，style表示风格，E表示期望， x_r 表示不配对数据， E_s 表示从源域到目标域的编码器， G_s 表示从源域到目标域的解码器， E_t 表示从目标域到源域的编码器， ϕ 表示特征图， H_k 表示具有特征内核k的再生核希尔伯特空间。

11. 根据权利要求1所述的方法，其特征在于，还包括：

基于所述预先确定的源字体库以及所述训练好的生成器生成所述目标字体的字体库。

12. 根据权利要求1所述的方法，其特征在于，还包括：

确定待处理图像，所述待处理图像中包括待替换字符的第三图像；

基于所述训练好的生成器生成所述待替换字符对应所述目标字体的第四图像；
将所述待处理图像中的第三图像替换为所述第四图像。

13. 一种基于生成对抗网络的字体数据处理装置,其特征在于,包括:

第一确定模块,用于确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;

第二确定模块,用于基于所述第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;

训练模块,用于基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。

14. 一种计算机设备,其特征在于,包括存储器以及处理器;所述存储器中存储有可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述权利要求1至12任一项所述的方法的步骤。

15. 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存储有机器可运行指令,所述计算机可运行指令在被处理器调用和运行时,所述计算机可运行指令促使所述处理器运行所述权利要求1至12任一项所述的方法。

基于生成对抗网络的字体数据处理方法及装置

技术领域

[0001] 本发明涉及计算机技术领域,尤其是涉及一种基于生成对抗网络的字体数据处理方法及装置。

背景技术

[0002] 汉字是古老而珍贵的文化遗产。自古以来,无数优秀的书法家留下了宝贵的书法作品,成为人类文明的辉煌成就。但是,在悠久的历史中,许多珍贵的书法作品丢失了。不同于英语只有26个字母,汉语覆盖了成千上万个字,每个字都有不同的字形并代表不同的含义,此外,不同的书法家都有自己的书写风格,具有特有的整体结构和笔触细节。

[0003] 对汉字生成的研究已经有很长时间,基于图像的方法拆分并重组数据集中的相应笔画和部首来生成所需的汉字,但是,目前的方式都需要较为庞大的训练样本。

发明内容

[0004] 本发明的目的在于提供一种基于生成对抗网络的字体数据处理方法及装置,以缓解了现有技术中存在的训练样本数据量需求大的技术问题。

[0005] 第一方面,本发明提供一种基于生成对抗网络的字体数据处理方法,包括:

[0006] 确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;

[0007] 基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;

[0008] 基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。

[0009] 在可选的实施方式中,所述基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,包括:

[0010] 基于第一数量个目标字符图像,从预先确定的源字体库中渲染得到与所述目标字符图像对应相同的字符的第三数量个源字符图像;

[0011] 从预先确定的源字体库中随机渲染得到第四数量个源字符图像;所述第二数量个源字符图像由所述第三数量个源字符图像和所述第四数量个源字符图像构成。

[0012] 在可选的实施方式中,所述第三数量与所述第四数量均等于所述第一数量;

[0013] 所述第一数量个目标字符图像和所述第三数量个源字符图像构成第一图像集,所述第一图像集包括多个第一图像对,每个所述第一图像对包括同一字符对应的目标字符图像和源字符图像;

[0014] 所述第一数量个目标字符图像和所述第四数量个源字符图像构成第二图像集,第二图像集包括多个第二图像对,每个所述第二图像对包括不同字符对应的目标字符图像和源字符图像。

[0015] 在可选的实施方式中,所述基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,包括:

[0016] 基于所述第一图像集和所述第二图像集对所述生成对抗网络进行训练,得到训练好的生成器。

[0017] 在可选的实施方式中,所述基于所述第一图像集和所述第二图像集对所述生成对抗网络进行训练,得到训练好的生成器,包括:

[0018] 基于所述第一图像集以及第一损失函数对所述生成对抗网络进行第一训练;

[0019] 基于所述第二图像集以及第二损失函数对所述生成对抗网络进行第二训练;

[0020] 迭代的交替进行所述第一训练以及第二训练直至符合终止条件,输出训练好的生成器。

[0021] 在可选的实施方式中,所述第一损失函数基于对抗损失、一致性损失以及对齐损失确定;所述第二损失函数基于所述对抗损失、所述一致性损失以及风格损失确定;其中,所述对抗损失用于指示分布以及分类上的损失,所述一致性损失用于指示循环能力以及字符属性上的损失,所述对齐损失用于指示内容和特征上的损失,所述风格损失用于指示特征分布上的损失。

[0022] 在可选的实施方式中,所述对抗损失基于主要部分损失和辅助部分损失确定;

[0023] 所述主要部分损失基于如下公式确定:

$$[0024] \quad L_{adv}^{x \rightarrow y} = \mathbb{E}_y[(D_t(y))^2] + \mathbb{E}_x[(1 - D_t(G_s(E_s(x))))^2]$$

[0025] 所述辅助部分损失基于如下公式确定:

$$[0026] \quad L_{cam}^{x \rightarrow y} = \mathbb{E}_y[(\eta_{D_t}(y))^2] + \mathbb{E}_x[(1 - \eta_{D_t}(G_s(E_s(x))))^2]$$

[0027] 其中,L表示损失函数,x为源域,y为目标域,adv表示对抗,cam表示类激活图,E表示期望, D_t 表示判别器, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, $\eta_{D_t}()$ 表示判别器中添加的辅助分类器, E_s 和 G_s 共同构成第一子生成器,所述训练好的生成器包括所述第一子生成器。

[0028] 在可选的实施方式中,所述一致性损失包括循环损失和身份损失;

[0029] 所述循环损失基于如下公式确定:

$$[0030] \quad L_{cycle}^{x \rightarrow y} = \mathbb{E}_x[|x - G_t(E_t(G_s(E_s(x))))|_1]$$

[0031] 所述身份损失基于如下公式确定:

$$[0032] \quad L_{identity}^{x \rightarrow y} = \mathbb{E}_y[|y - (G_s(E_s(y)))|_1]$$

[0033] 其中,L表示损失函数,x为源域,y为目标域,cycle表示循环,identity表示身份,E表示期望, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, G_t 表示从目标域到源域的解码器, E_t 和 G_t 共同构成第二子生成器,所述训练好的生成器包括所述第二子生成器。

[0034] 在可选的实施方式中,所述对齐损失包括内容损失和特征损失;

[0035] 所述内容损失基于如下公式确定:

$$[0036] \quad L_{L1}^{x \rightarrow y} = \mathbb{E}_{x,y}[|y - G_s(E_s(x_p))|_1]$$

[0037] 所述特征损失基于如下公式确定；

$$[0038] \quad L_{constancy}^{x \rightarrow y} = \mathbb{E}_{x,y}[|E_t(y) - E_t(G_s(E_s(x_p)))|_2]$$

[0039] 其中,L表示损失函数,x为源域,y为目标域, L_1 表示内容,constancy表示不变性,E表示期望, x_p 表示配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器。

[0040] 在可选的实施方式中,所述风格损失基于如下公式确定:

$$[0041] \quad L_{style}^2 = \|\mathbb{E}_y[\phi(E_t(y))] - \mathbb{E}_x[\phi(E_t(G_s(E_s(x_r))))]\|_{H_k}^2$$

[0042] 其中,L表示损失函数,x为源域,y为目标域,style表示风格,E表示期望, x_r 表示不配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, ϕ 表示特征图, H_k 表示具有特征内核k的再生核希尔伯特空间。

[0043] 在可选的实施方式中,还包括:

[0044] 基于所述预先确定的源字体库以及所述训练好的生成器生成所述目标字体的字体库。

[0045] 在可选的实施方式中,还包括:

[0046] 确定待处理图像,所述待处理图像中包括待替换字符的第三图像;

[0047] 基于所述训练好的生成器生成所述待替换字符对应所述目标字体的第四图像;

[0048] 将所述待处理图像中的第三图像替换为所述第四图像。

[0049] 第二方面,本发明提供一种基于生成对抗网络的字体数据处理装置,包括:

[0050] 第一确定模块,用于确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;

[0051] 第二确定模块,用于基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;

[0052] 训练模块,用于基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。

[0053] 第三方面,本发明提供一种计算机设备,包括存储器以及处理器;所述存储器中存储有可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现上述前述实施方式任一项所述的方法的步骤。

[0054] 第四方面,本发明提供一种计算机可读存储介质,所述计算机可读存储介质存储有机器可运行指令,所述计算机可运行指令在被处理器调用和运行时,所述计算机可运行指令促使所述处理器运行所述前述实施方式任一项所述的方法。

[0055] 本发明提供一种基于生成对抗网络的字体数据处理方法及装置。通过确定第一数量个目标字符图像,所述目标字符图像中字符的风格为目标字体;基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,所述第二数量大于所述第一数量,所述第二数量个源字符图像中的部分与所述目标字符图像对应相同的字符;基于所述第一数量个目标字符图像以及所述第二数量个源字符图像对所述生成对抗网络进行训练,得到训练好的生成器,所述训练好的生成器用于生成所述目标字体的字符图像。以

此可以实现使用少量的目标字符图像既可以使得模型具备生成目标字体的能力,提升了模型的应用场景,提升了用户体验。

附图说明

[0056] 为了更清楚地说明本申请实施例的技术方案,下面将对本申请实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,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他相关的附图。

[0057] 图1为本申请实施例提供的一种基于生成对抗网络的字体数据处理方法流程示意图;

[0058] 图2为本申请实施例提供的另一种基于生成对抗网络的字体数据处理方法流程示意图;

[0059] 图3为本申请实施例提供的另一种基于生成对抗网络的字体数据处理方法流程示意图;

[0060] 图4为本申请实施例提供的一种基于生成对抗网络的字体数据处理装置结构示意图;

[0061] 图5为本申请实施例提供的一种计算机设备结构示意图。

具体实施方式

[0062] 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。通常在此处附图中描述和示出的本发明实施例的组件可以以各种不同的配置来布置和设计。

[0063] 因此,以下对在附图中提供的本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围,而是仅仅表示本发明的选定实施例。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。

[0064] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释。

[0065] 字形图像的自动生成可以大大减少字体设计者的人工成本,同时,对于书法初学者来说,模仿大师的杰作将会对于书法学习非常有帮助,因此根据少量的参考字生成完整的个性化字库非常有意义且具有挑战性。

[0066] 下面结合附图,对本发明的一些实施方式作详细说明。在不冲突的情况下,下述的实施例及实施例中的特征可以相互组合。

[0067] 图1为本发明实施例提供的一种基于生成对抗网络的字体数据处理方法流程示意图。该方法可以应用于计算机设备,如图1所示,该方法可以包括如下步骤:

[0068] 步骤S110,确定第一数量个目标字符图像。

[0069] 其中,该目标字符图像中字符的风格为目标字体。该目标字体可以为待生成的字体。例如,该目标字符的图像可以为用户上传的图像也可以为通过其他方式来指定的特定

风格的目标字体的图像。例如,该目标字符图像可以由用户绘制得到,该图像可以包括有限多个,该有限多个目标字符图像可以作为用户绘制的该目标字体的风格,其中,用户绘制的方式可以通过电脑制图或者拍摄手绘画像。再例如,该目标字符的图像还可以根据一些书法作品确定,该目标字符的图像可以为拍摄的该书法作品的图像。具体可以根据实际需要确定。

[0070] 该第一数量为一个有限的数量,通常情况下该第一数量越大,得到的生成器所生成的图像越逼真,相应的训练所需要的资源也就越多,所以该第一数量可以根据实际掌握的目标字符图像的数量、对生成器能力的需求以及训练资源的需求等综合考虑来预先确定。

[0071] 步骤S120,基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像。

[0072] 其中,该第二数量大于第一数量,该第二数量个源字符图像中的部分与目标字符图像对应相同的字符。

[0073] 该源字体库可以为预先确定的标准字体库,例如该预先确定的标准字体库可以为宋体库,基于该源字体库可以生成大量的源字体的源字符图像,该源字体库的大小可以根据实际需要确定。

[0074] 基于该源字体库可以丰富训练生成器所需要的素材,基于目标字符图像以及源字体库确定一部分与目标字符图像对应相同的字符源字符图像和另一部分与目标字符图像对应不相同的字符源字符图像,可以更有针对性的确定用于训练生成器的训练样本。

[0075] 在一些实施例中,对于源字符图像的获取方式可以包括多种方式。作为一个示例,可以基于第一数量个目标字符图像,从预先确定的源字体库中渲染得到与目标字符图像对应相同的字符的第三数量个源字符图像;从预先确定的源字体库中随机渲染得到第四数量个源字符图像;该第二数量个源字符图像由第三数量个源字符图像和第四数量个源字符图像构成。

[0076] 其中,该第三数量与第四数量可以均等于第一数量;此时,该第一数量个目标字符图像和第三数量个源字符图像可以构成第一图像集;该第一图像集包括多个第一图像对,每个第一图像对包括同一字符对应的目标字符图像和源字符图像。该第一数量个目标字符图像和第四数量个源字符图像构成第二图像集;该第二图像集包括多个第二图像对,每个第二图像对包括不同字符对应的目标字符图像和源字符图像。以此通过第一图像集作为配对的数据,用以提升生成器的在不同字形下结构之间精细的相关性的能力,通过第二图像集作为不配对的数据,用以提升生成器的在字体结构的粗粒度区分的能力。

[0077] 步骤S130,基于第一数量个目标字符图像以及第二数量个源字符图像对生成对抗网络进行训练,得到训练好的生成器,训练好的生成器用于生成目标字体的字符图像。

[0078] 生成对抗网络(Generative Adversarial Networks, GAN)自从提出以来就引起了学术界和工业界的兴趣。它已成功应用于许多不同领域,并获得了令人印象深刻的结果,例如图像生成,图像修复,图像编辑,图像转换等。生成对抗网络成功的关键点在于,判别器试图将生成的图像与真实图像区分开,而生成器则试图混淆判别器的判断。

[0079] 图像转换旨在学习一种映射功能,该功能可以将图像从源域转换为目标域。它已被广泛用于许多应用中,例如,艺术风格的图像转换,语义分割,照片增强或对象替换。

[0080] 可以基于第一数量个目标字符图像以及第二数量个源字符图像生成训练样本,基于训练样本对生成对抗网络进行训练,基于训练好的生成对抗网络中的训练好的生成器来生成目标字体的字符图像。其中,该训练好的生成器的输入可以为其他字体的字符图像,输出为目标字体的字符图像。

[0081] 通过本发明实施例,基于少量的目标字符图像以及已知的源字体库来确定训练样本,基于生成式对抗网络对生成器进行训练,得到的生成器具备在不同字形下结构之间能够识别精细的相关性的能力,且在字体结构具有粗粒度区分的能力,使得生成器可以基于少量的目标字符图像进行训练,便可以具备生成目标字体的能力,且生成的目标字体逼真,由于对目标字符图像的数量需求较小,所以模型的应用场景更广,用户体验更高。

[0082] 在一些实施例中,结合步骤S120中将第一数量个目标字符图像以及第二数量个源字符图像划分为第一图像集和第二图像集的方式,可以基于第一图像集和第二图像集对生成对抗网络进行训练,得到训练好的生成器。

[0083] 具体可以通过如下步骤来实现:

[0084] 步骤1),基于第一图像集以及第一损失函数对生成对抗网络进行第一训练;

[0085] 步骤2),基于第二图像集以及第二损失函数对生成对抗网络进行第二训练;

[0086] 步骤3),迭代的交替进行第一训练以及第二训练直至符合终止条件,输出训练好的生成器。

[0087] 其中,应该知道的是,该生成对抗网络包括生成器和判别器,该第一训练和第二训练均需要对生成器和判别器分别进行优化。该终止条件可以根据实际需要确定,例如训练达到指定的次数,判别器的识别准确率符合预设条件,或者损失函数保持一段时间恒定等等。

[0088] 在一些实施例中,用于训练生成式对抗网络的损失函数可以包括多种。作为一个示例,上述第一损失函数可以基于对抗损失、一致性损失以及对齐损失确定;上述第二损失函数可以基于对抗损失、一致性损失以及风格损失确定。

[0089] 其中,该对抗损失用于指示分布以及分类上的损失;该一致性损失用于指示循环能力以及字符属性上的损失;该对齐损失用于指示内容和特征上的损失;该风格损失用于指示特征分布上的损失。基于上述方式确定的损失函数,能够强化模型在特定的方向上的能力。

[0090] 例如,该对抗损失可以包括主要部分损失和辅助部分损失,基于该主要部分损失和辅助部分损失可以确定该对抗损失。

[0091] 该主要部分损失可以基于如下公式(一)来确定:

$$L_{adv}^{x \rightarrow y} = \mathbb{E}_y [(D_t(y))^2] + \mathbb{E}_x [(1 - D_t(G_s(E_s(x))))^2] \text{ 式 (一)}$$

[0093] 该辅助部分损失可以基于如下公式(二)来确定:

$$L_{cam}^{x \rightarrow y} = \mathbb{E}_y [(\eta_{D_t}(y))^2] + \mathbb{E}_x [(1 - \eta_{D_t}(G_s(E_s(x))))^2] \text{ 式 (二)}$$

[0095] 其中,对于上述公式(一)和公式(二),L表示损失函数,x为源域(例如,该源域可以为第二数量个源字符图像),y为目标域(例如,该目标域可以为第一数量个目标字符图像),adv表示对抗,cam表示类激活图,E表示期望, D_t 表示判别器, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, $\eta_{D_t}(\cdot)$ 表示判别器中添加的辅助分类器。在本申请实

施例中,该 E_s 和 G_s 可以共同构成第一子生成器,该训练好的生成器可以包括训练好的第一子生成器。

[0096] 再例如,该一致性损失可以包括循环损失和身份损失:

[0097] 其中,该循环损失可以基于如下公式(三)来确定:

$$L_{cycle}^{x \rightarrow y} = \mathbb{E}_x [|x - G_t(E_t(G_s(E_s(x))))|_1] \text{式(三)}$$

[0099] 该身份损失可以基于如下公式(四)来确定:

$$L_{identity}^{x \rightarrow y} = \mathbb{E}_y [|y - (G_s(E_s(y)))|_1] \text{式(四)}$$

[0101] 其中,对于上述公式(三)和公式(四),L表示损失函数,x为源域(例如,该源域可以为第二数量个源字符图像),y为目标域(例如,该目标域可以为第一数量个目标字符图像),cycle表示循环,identity表示身份,E表示期望, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, G_t 表示从目标域到源域的解码器。在本申请实施例中,该 E_t 和 G_t 共同构成第二子生成器,训练好的生成器包括训练好的第二子生成器。

[0102] 再例如,该对齐损失包括内容损失和特征损失:

[0103] 其中,该内容损失可以基于如下公式(五)确定:

$$L_{L_1}^{x \rightarrow y} = \mathbb{E}_{x,y} [|y - G_s(E_s(x_p))|_1] \text{式(五)}$$

[0105] 该特征损失基于如下公式(六)确定:

$$L_{constancy}^{x \rightarrow y} = \mathbb{E}_{x,y} [|E_t(y) - E_t(G_s(E_s(x_p)))|_2] \text{式(六)}$$

[0107] 其中,对于上述公式(五)和公式(六),L表示损失函数,x为源域(例如,该源域可以为第二数量个源字符图像),y为目标域(例如,该目标域可以为第一数量个目标字符图像), L_1 表示内容,constancy表示不变性,E表示期望, x_p 表示配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器。

[0108] 再例如,该风格损失可以基于如下公式(七)确定:

$$L_{style}^2 = \|\mathbb{E}_y [\phi(E_t(y))] - \mathbb{E}_x [\phi(E_t(G_s(E_s(x_r))))]\|_{H_k}^2 \text{式(七)}$$

[0110] 其中,对于上述公式(七),L表示损失函数,x为源域(例如,该源域可以为第二数量个源字符图像),y为目标域(例如,该目标域可以为第一数量个目标字符图像),style表示风格,E表示期望, x_r 表示不配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, ϕ 表示特征图, H_k 表示具有特征内核k的再生核希尔伯特空间。

[0111] 综上所述,全部损失函数可以表示为(参见公式(八)):

$$L = \lambda_1 L_{GAN} + \lambda_2 L_{consistency} + \lambda_3 L_{alignment} + \lambda_4 L_{style} \text{式(八)}$$

[0113] 其中,在公式(八)中,该 L_{GAN} 表示对抗损失,该 λ_1 表示该对抗损失的权重;该 $L_{consistency}$ 表示一致性损失,该 λ_2 表示该一致性损失的权重;该 $L_{alignment}$ 表示对齐损失,该 λ_3 表示该对齐损失的权重;该 L_{style} 表示风格损失,该 λ_4 表示该对抗损失的权重。

[0114] 在一些实施例中,其中,在本申请实施例中,该训练好的生成器可以用于生成字体库。例如,可以基于预先确定的源字体库以及训练好的生成器生成目标字体的字体库。

[0115] 在具体实现时,可以将一些有限的真迹作为目标字符图像,基于该有限的真迹,训练生成器,并基于训练好的生成器,生成包含更多字符的字体库,该字体库中字符的风格很好的继承了该有限的真迹的风格。或者,用户可以绘制有限的目标字符图像,基于该绘制的图像,训练生成器,并基于训练好的生成器,生成包含更多字符的字体库,该字体库中字符的风格可以很好的继承用户绘制的目标字符图像的风格,以此可以极大的降低用户在生成字体库时所需要绘制的目标字符图像的数量,极大的降低了工作量,提升了工作效率,提升了用户体验。

[0116] 在一些实施例中,确定待处理图像,待处理图像中包括待替换字符的第三图像;基于训练好的生成器生成待替换字符对应目标字体的第四图像;将待处理图像中的第三图像替换为第四图像。例如,该待处理图像可以包括多种字体,可以基于该待处理图像中的多种字体统一为其中的任意一种字体,以此可以使得该待处理图像中的字体风格得到统一。例如,该待处理图像可以包括目标字符图像,该目标字符图像可以通过对待处理图像进行识别得到,该目标字符图像可以为用户在该待处理图像中选定的一些字符图像。确定该目标字符图像后,可以对生成器进行训练得到训练好的生成器,基于该训练好的生成器,将该待处理图像种的全部需要统一的字符图像进行替换,以此实现该待处理图像中字体风格的统一。

[0117] 下面结合一个具体示例对本发明实施例进行进一步地介绍。参见图如图2和图3所示。

[0118] 步骤1),得到少量的目标字体图片 y 。该目标字体图片 y 即为第一数量个目标字符图像。该目标字体图片 y 可以为真迹,也可以其他的字符的图像。参见图3所示,该目标字体图片 y 可以包括“皚、磅、蛀、徨、伎、晴”6个目标字体图片,也即图3中的“Ground truth y ”。

[0119] 步骤2),根据 y ,从标准字体宋体TTF中渲染出配对的字符图片 x_p 。其中,该配对的字符图片 x_p 即为第二数量个源字符图像中的一部分,也即第三数量个源字符图像。参见图3中的“input x ”中分别与“Ground truth y ”一一对应的6个源字符图像。

[0120] 步骤3),从宋体TTF中随机渲染出其他字符图片 x_r 。

[0121] 步骤4),对成对数据 $\{x_p, y\}$ 使用对抗损失、一致性损失、对齐损失进行训练,使用梯度下降法最小化损失。

[0122] 步骤5),对不成对数据 $\{x_r, y\}$ 使用对抗损失、一致性损失、风格损失进行训练,使用梯度下降法最小化损失。

[0123] 步骤6),对于步骤4)、步骤5)迭代地交替训练,直到损失函数收敛得到最终生成器,输入源字符(可以基于宋体TTF确定),可以得到效果很好的目标字符。

[0124] 只使用少量的目标字体进行训练,如图3所示,第一行为输入的宋体,第二行为生成的字体,第三行为真值。可以看到无论是风格还是内容上,模型都能完美地模仿出目标字体。

[0125] 通过本发明实施例,基于少量的目标字体图片 y 以及已知的宋体TTF来确定训练样本(成对数据 $\{x_p, y\}$ 和不成对数据 $\{x_r, y\}$),基于该训练样本对生成器进行训练,得到的生成器具备在不同字形下结构之间能够识别精细的相关性的能力,且在字体结构具有粗粒度区分的能力,使得生成器可以基于少量的目标字符图像进行训练,便可以具备生成目标字体的能力,且生成的目标字体逼真,由于对目标字符图像的数量需求较小,所以模型的应用场

景更广,用户体验更高。

[0126] 图4为本发明实施例提供的一种基于生成对抗网络的字体数据处理装置结构示意图,如图4所示,该装置包括:

[0127] 第一确定模块401,用于确定第一数量个目标字符图像,目标字符图像中字符的风格为目标字体;

[0128] 第二确定模块402,用于基于第一数量个目标字符图像以及预先确定的源字体库,确定第二数量个源字符图像,第二数量大于第一数量,第二数量个源字符图像中的部分与目标字符图像对应相同的字符;

[0129] 训练模块403,用于基于第一数量个目标字符图像以及第二数量个源字符图像对生成对抗网络进行训练,得到训练好的生成器,训练好的生成器用于生成目标字体的字符图像。

[0130] 在一些实施例中,第二确定模块402具体用于:

[0131] 基于第一数量个目标字符图像,从预先确定的源字体库中渲染得到与目标字符图像对应相同的字符的第三数量个源字符图像;

[0132] 从预先确定的源字体库中随机渲染得到第四数量个源字符图像;第二数量个源字符图像由第三数量个源字符图像和第四数量个源字符图像构成。

[0133] 在一些实施例中,第三数量与第四数量均等于第一数量;

[0134] 第一数量个目标字符图像和第三数量个源字符图像构成第一图像集,第一图像集包括多个第一图像对,每个第一图像对包括同一字符对应的目标字符图像和源字符图像;

[0135] 第一数量个目标字符图像和第四数量个源字符图像构成第二图像集,第二图像集包括多个第二图像对,每个第二图像对包括不同字符对应的目标字符图像和源字符图像。

[0136] 在一些实施例中,训练模块403具体用于:

[0137] 基于第一图像集和第二图像集对生成对抗网络进行训练,得到训练好的生成器。

[0138] 在一些实施例中,训练模块403具体用于:

[0139] 基于第一图像集以及第一损失函数对生成对抗网络进行第一训练;

[0140] 基于第二图像集以及第二损失函数对生成对抗网络进行第二训练;

[0141] 迭代的交替进行第一训练以及第二训练直至符合终止条件,输出训练好的生成器。

[0142] 在一些实施例中,第一损失函数基于对抗损失、一致性损失以及对齐损失确定;第二损失函数基于对抗损失、一致性损失以及风格损失确定;其中,对抗损失用于指示分布以及分类上的损失,一致性损失用于指示循环能力以及字符属性上的损失,对齐损失用于指示内容和特征上的损失,风格损失用于指示特征分布上的损失。

[0143] 在一些实施例中,对抗损失基于主要部分损失和辅助部分损失确定;

[0144] 主要部分损失基于如下公式确定:

$$L_{adv}^{x \rightarrow y} = \mathbb{E}_y [(D_t(y))]^2 + \mathbb{E}_x [(1 - D_t(G_s(E_s(x))))^2]$$

[0146] 辅助部分损失基于如下公式确定:

$$L_{cam}^{x \rightarrow y} = \mathbb{E}_y [(\eta_{D_t}(y))]^2 + \mathbb{E}_x [(1 - \eta_{D_t}(G_s(E_s(x))))^2]$$

[0148] 其中,L表示损失函数,x为源域,y为目标域,adv表示对抗,cam表示类激活图,E表

示期望, D_t 表示判别器, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, $\eta_{D_t}(\cdot)$ 表示判别器中添加的辅助分类器, E_s 和 G_s 共同构成第一子生成器, 训练好的生成器包括第一子生成器。

[0149] 在一些实施例中, 一致性损失包括循环损失和身份损失;

[0150] 循环损失基于如下公式确定:

$$[0151] \quad L_{cycycle}^{x \rightarrow y} = \mathbb{E}_x [|x - G_t(E_t(G_s(E_s(x))))|_1]$$

[0152] 身份损失基于如下公式确定:

$$[0153] \quad L_{identity}^{x \rightarrow y} = \mathbb{E}_y [|y - (G_s(E_s(y)))|_1]$$

[0154] 其中, L 表示损失函数, x 为源域, y 为目标域, $cycle$ 表示循环, $identity$ 表示身份, E 表示期望, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, G_t 表示从目标域到源域的解码器, E_t 和 G_t 共同构成第二子生成器, 训练好的生成器包括第二子生成器。

[0155] 在一些实施例中, 对齐损失包括内容损失和特征损失;

[0156] 内容损失基于如下公式确定:

$$[0157] \quad L_{L1}^{x \rightarrow y} = \mathbb{E}_{x,y} [|y - G_s(E_s(x_p))|_1]$$

[0158] 特征损失基于如下公式确定:

$$[0159] \quad L_{constancy}^{x \rightarrow y} = \mathbb{E}_{x,y} [|E_t(y) - E_t(G_s(E_s(x_p)))|_2]$$

[0160] 其中, L 表示损失函数, x 为源域, y 为目标域, L_1 表示内容, $constancy$ 表示不变性, E 表示期望, x_p 表示配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器。

[0161] 在一些实施例中, 风格损失基于如下公式确定:

$$[0162] \quad L_{style}^2 = \|\mathbb{E}_y [\phi(E_t(y))] - \mathbb{E}_x [\phi(E_t(G_s(E_s(x_r))))]\|_{H_k}^2$$

[0163] 其中, L 表示损失函数, x 为源域, y 为目标域, $style$ 表示风格, E 表示期望, x_r 表示不配对数据, E_s 表示从源域到目标域的编码器, G_s 表示从源域到目标域的解码器, E_t 表示从目标域到源域的编码器, ϕ 表示特征图, H_k 表示具有特征内核 k 的再生核希尔伯特空间。

[0164] 在一些实施例中, 还包括:

[0165] 生成模块, 用于基于预先确定的源字体库以及训练好的生成器生成目标字体的字体库。

[0166] 在一些实施例中, 还包括:

[0167] 第三确定模块, 用于确定待处理图像, 待处理图像中包括待替换字符的第三图像;

[0168] 生成模块, 用于基于训练好的生成器生成待替换字符对应目标字体的第四图像;

[0169] 替换模块, 用于将待处理图像中的第三图像替换为第四图像。

[0170] 本申请实施例提供的基于生成对抗网络的字体数据处理装置, 与上述实施例提供的基于生成对抗网络的字体数据处理方法具有相同的技术特征, 所以也能解决相同的技术问题, 达到相同的技术效果。

[0171] 如图5所示, 本申请实施例提供的一种计算机设备700, 例如, 该计算机设备700可

以为预处理服务器,包括:处理器701、存储器702和总线,存储器702存储有处理器701可执行的机器可读指令,当电子设备运行时,处理器701与存储器702之间通过总线通信,处理器701执行机器可读指令,以执行如上述基于生成对抗网络的字体数据处理方法的步骤。

[0172] 具体地,上述存储器702和处理器701能够为通用的存储器和处理器,这里不做具体限定,当处理器701运行存储器702存储的计算机程序时,能够执行上述基于生成对抗网络的字体数据处理方法。

[0173] 对应于上述基于生成对抗网络的字体数据处理方法,本申请实施例还提供了一种计算机可读存储介质,计算机可读存储介质存储有机器可运行指令,计算机可运行指令在被处理器调用和运行时,计算机可运行指令促使处理器运行上述基于生成对抗网络的字体数据处理方法的步骤。

[0174] 本申请实施例所提供的基于生成对抗网络的字体数据处理装置可以为设备上的特定硬件或者安装于设备上的软件或固件等。本申请实施例所提供的装置,其实现原理及产生的技术效果和前述方法实施例相同,为简要描述,装置实施例部分未提及之处,可参考前述方法实施例中相应内容。所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,前述描述的系统、装置和单元的具体工作过程,均可以参考上述方法实施例中的对应过程,在此不再赘述。

[0175] 在本申请所提供的实施例中,应该理解到,所揭露装置和方法,可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,又例如,多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0176] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0177] 在本申请所提供的几个实施例中,应该理解到,所揭露的装置和方法,也可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本申请的多个实施例的装置、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分,模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现方式中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0178] 另外,在本申请提供的实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0179] 功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本申请的技术方案本质上或者说对现

有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备执行本申请各个实施例移动控制方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read-Only Memory,简称ROM)、随机存取存储器(Random Access Memory,简称RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0180] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释,此外,术语“第一”、“第二”、“第三”等仅用于区分描述,而不能理解为指示或暗示相对重要性。

[0181] 最后应说明的是:以上实施例,仅为本申请的具体实施方式,用以说明本申请的技术方案,而非对其限制,本申请的保护范围并不局限于此,尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本申请实施例技术方案的范围。都应涵盖在本申请的保护范围之内。

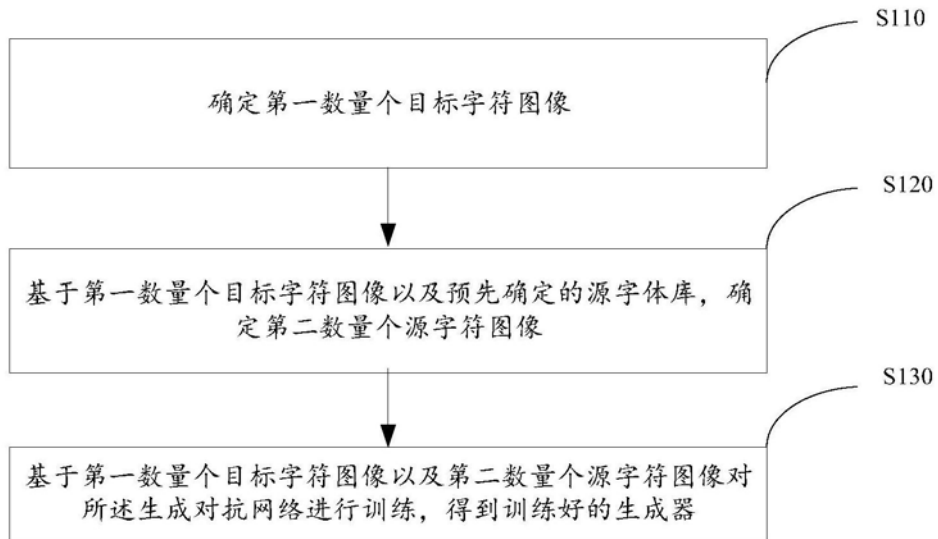


图1

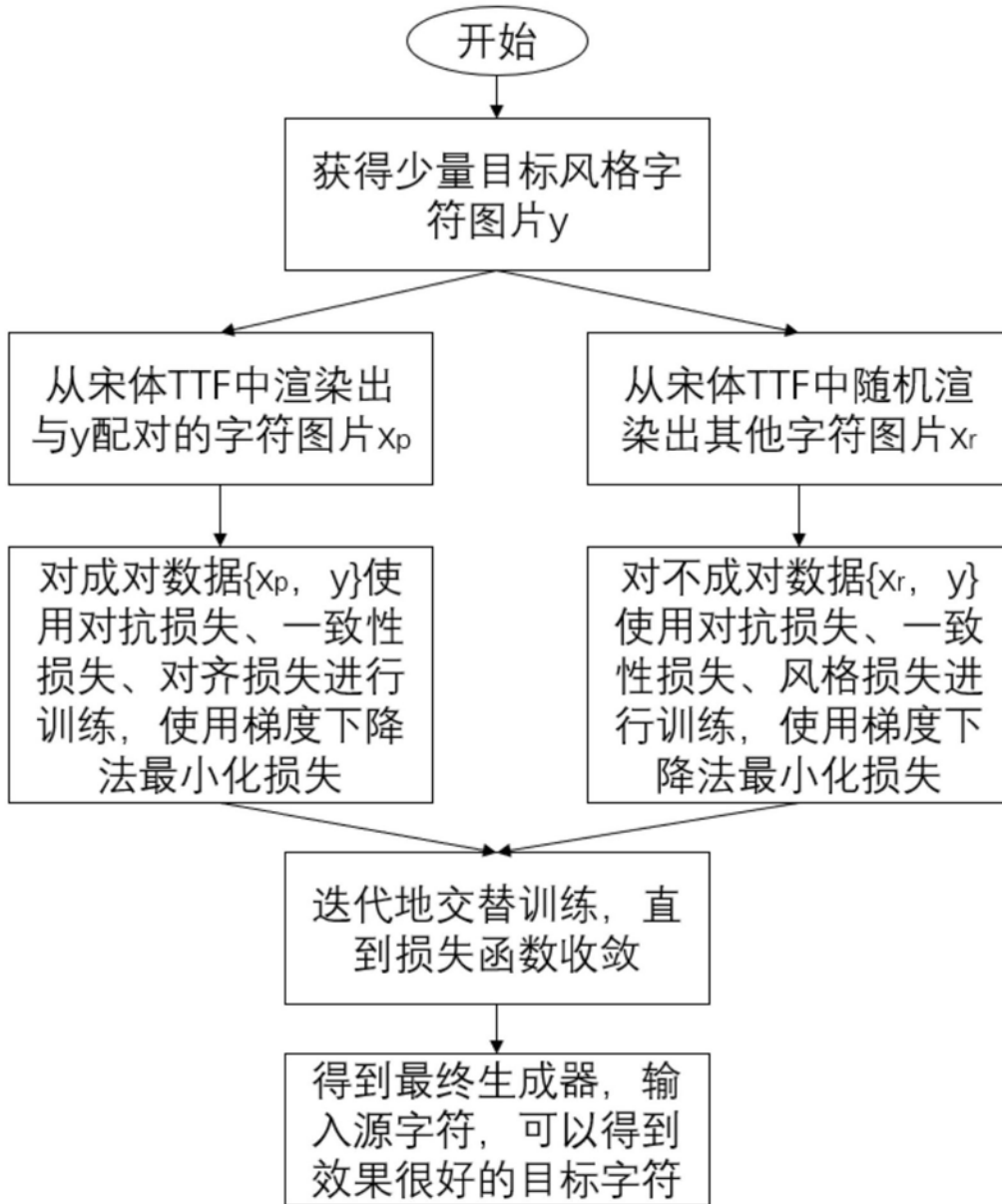


图2

Input x	皚	磅	蛀	徨	伎	睛
Ours	皚	磅	蛀	徨	伎	睛
Ground truth y	皚	磅	蛀	徨	伎	睛

图3

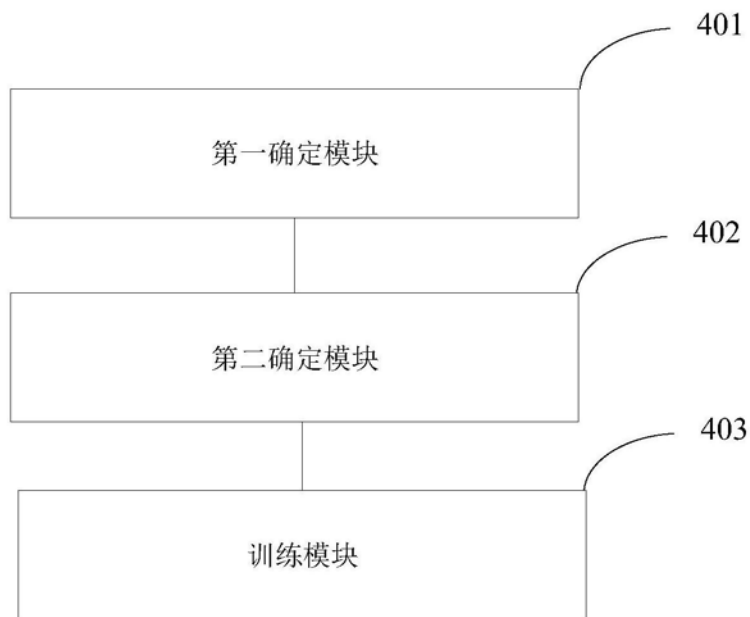


图4

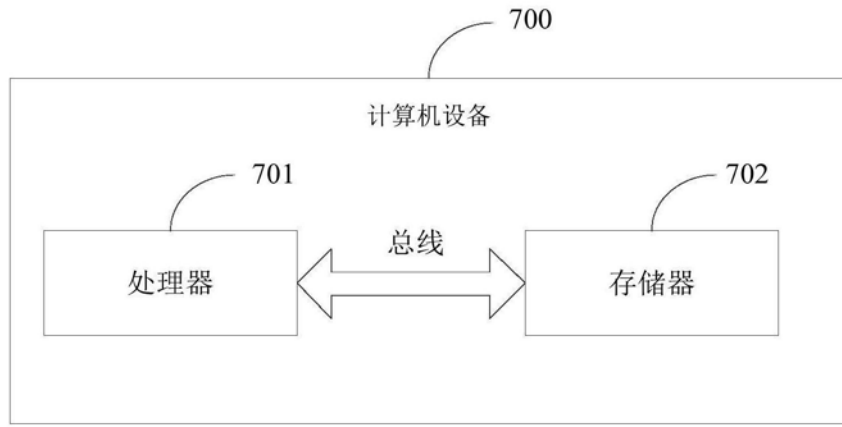


图5