

汉字结构引导下的大千字体生成方法研究

张 攀¹ 徐瑞娟²
ZHANG Pan XU Ruijuan

摘 要

针对搜集张大千书法作品,构建大千字体过程中存在的字库容量不足问题,文章提出了一种结构引导下的大千风格汉字生成方法。该方法将具备常见印刷体汉字风格迁移的 zi2zi 模型作为基本结构,融入可双向约束风格迁移过程的 CycleGAN 模型形成 zibyzi 网络,在汉字骨架结构引导下,分别实现汉字风格附着与笔画细节微调。在测试集中实验获得的结构相似度达到 0.83,结合现存的大千书法作品,最终扩展形成了包含 6 000 个汉字的大千字库。

关键词

大千字体; 结构引导; zi2zi; CycleGAN

doi: 10.3969/j.issn.1672-9528.2025.09.006

0 引言

中文汉字以其优美的空间笔画结构,形成了独具中华文化特色的书法体系,是世界文字史上独一无二的瑰宝。如飘逸潇洒的草书、方正端严的金体、遒劲雄浑的毛体等,都彰显出了汉字艺术的博大精深。纵观历史,从各朝各代无数书画大家,到当今的众多专业学者、业余爱好者等都积极投身于书法创作与临摹之中,汉字书法在其基本的信息传递作用基础之上,更是文化价值、个人魅力的重要载体。随着人工智能技术的不断发展,尤其是与文化艺术领域的不断融合,中文汉字书法的创新步伐持续提速。潘今一等人^[1]基于风格迁移和薄板样条技术实现了汉字笔画结构、布局等的微调,但其并非真正意义上的新的字体风格的转换。秦嘉霖等人^[2]则利用主观汉字构形原理,即中文汉字由典型的重复笔画结构组合而成,并结合改进的条件生成对抗网络,实现了较少训练样本情况下的汉字风格迁移,但实验结果表明其更适合笔画结构规整的汉字之间进行风格转换,对字形结构不规整场景的适用性有限。姚伟健等人^[3]针对繁体汉字特性,提出了一种自适应生成对抗网络,解决了其他模型生成繁体汉字时的伪影和模糊现象。陈二开等人^[4]在汉字风格迁移的研究中也发现神经网络模型生成的粉笔字存在笔画缺失和模糊等问题,进而在改进 CycleGAN 模型中增加了自注意力机制。整体上,这些已有研究都更擅长打印字体的风格迁移,而对非打印字体的风格转换能力则较弱。

张大千作为近代集书法与绘画成就于一身的艺术家,留存诸多作品。除了蜚声中外的仕女图系列、泼墨系列等绘画作品,其书法成就同样颇高,相关作品中的文字笔锋与结构都极具个性,当前研究多将其归纳为“大千体”。不同于过往结合人工智能技术研究的其他中文字体,初步统计已公开的张大千书迹相对常用 6 000 个汉字,仍然不足,还无法真正构成日常标准字体库^[5]。为此,基于前沿的生成式人工智能技术,本文利用 zi2zi 模型和 CycleGAN 模型构成 zibyzi 模型,实现了文字骨架结构引导下的张大千风格字体的生成,从而初步完成大千字库的构建。

1 方法原理

1.1 zi2zi 模型

zi2zi 模型^[6]是 Tian 等人开源的一个字体风格迁移项目,实现了众多字体文字之间的相互转换,主要基于 pix2pix^[7]、Ac-Gan 和 DC-GAN 等技术实现,其中尤其引入的 pix2pix 算法使得汉字生成笔画细节得到了大幅度的优化。整体架构如图 1 所示。核心模块包括用于文字图像信息提取与压缩的编码器、用于文字图像生成的解码器、字符码与字体码向量、进行字体分类的判决器。

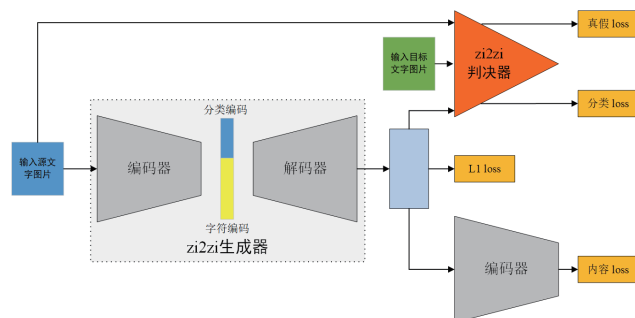


图 1 zi2zi 模型结构

1. 内江师范学院人工智能学院 四川内江 641100
2. 内江师范学院张大千美术学院 四川内江 641100
[基金项目] 四川省教育厅人文社会科学(张大千研究)项目
“大千字体数字化生成方法研究”(ZDQ2024-09)

训练该模型判别器和生成器的损失函数分别为:

$$d_{\text{loss}} = d_{\text{loss}}^{\text{real}} + d_{\text{loss}}^{\text{fake}} + \text{category}_{\text{loss}} / 2 \quad (1)$$

式中: $d_{\text{loss}}^{\text{real}}$ 表示真实图像对训练判别器的误差; $d_{\text{loss}}^{\text{fake}}$ 表示假的图像对训练判别器的误差, 由于此时的目标字体可能包含多类, 因此 $\text{category}_{\text{loss}}$ 表示分类误差。

$$g_{\text{loss}} = \text{cheat}_{\text{loss}} + l_{\text{loss}}^1 + \text{fake}_{\text{loss}}^{\text{category}} + \text{const}_{\text{loss}} \quad (2)$$

式中: $\text{cheat}_{\text{loss}}$ 表示假的图像对在判别器中产生的误差; l_{loss}^1 表示生成图像与目标图像之间对应像素点的误差; $\text{const}_{\text{loss}}$ 表示生成图像的编码误差, 由于此时总是每个字体单独训练, 因此 $\text{fake}_{\text{loss}}^{\text{category}}$ 表示字体错误分类误差。

1.2 CycleGAN 模型

CycleGAN 模型是 Zhu 等人^[8]提出的一种图像风格转换生成对抗网络, 其原理如图 2 所示。部分学者将训练数据替换为多种字体的中文汉字, 构建了 Handwritten-CycleGAN^[9], 取得了较好的效果。模型的核心原理是图像 A 通过生成器 G_{A2B} 转换为预测图像 B 后再通过生成器 G_{B2A} 可还原为图像 A , 同理, 图像 B 通过生成器 G_{B2A} 转换为预测图像 A 后再通过生成器 G_{A2B} 可还原为图像 B , 由此两个过程构成一种双向闭环。

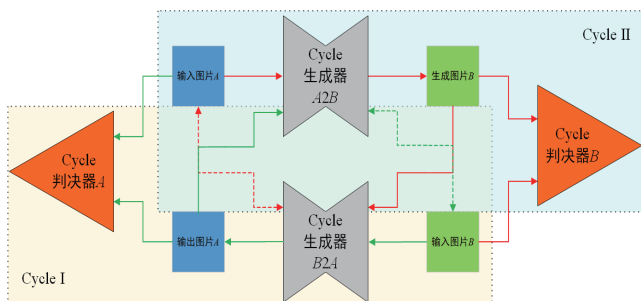


图 2 CycleGAN 网络结构

该模型训练时的损失函数用公式表示为:

$$\begin{aligned} \text{gan}_{\text{loss}}(G_{A2B}, G_{B2A}, D_A, D_B) = & \text{gan}_{\text{loss}}(G_{A2B}, D_B, A, B) \\ & + \text{gan}_{\text{loss}}(G_{B2A}, D_A, B, A) \quad (3) \\ & + \lambda_{\text{cyc}} \text{cyc}_{\text{loss}}(G_{A2B}, G_{B2A}) \\ & + \lambda_{\text{ide}} \text{ide}_{\text{loss}}(G_{A2B}, G_{B2A}) \end{aligned}$$

式中: $\text{gan}_{\text{loss}}()$ 表示生成对抗网络本身的损失函数, 即训常规生成对抗网络 G_{A2B} 和 G_{B2A} 所需的判别器损失; $\text{cyc}_{\text{loss}}()$ 表示循环一致性损失; $\text{ide}_{\text{loss}}()$ 表示身份映射损失, 用于共同保证双向训练过程中, 同域图像的一致映射, 如输入图像 A 与生成图像 A 、输入图像 B 与生成图像 B , 因此其具体损失函数直接计算两幅图像之间的像素差即可。

1.3 大千字体生成模型设计

参考 zi2zi 与 CycleGAN 两个模型, 本文使用 2 个 zi2zi 构成新的 CycleGAN, 并取名为 zibyzi。即执行 A 域图像与 B 域图像相互转换的 GAN, 由 zi2zi 中的生成器构成, 执行 A

域和 B 域图像判定的判别器则由 zi2zi 中的判别器构成。模型相应的总体损失函数用公式表示为:

$$\begin{aligned} \text{zibyzi}_{\text{loss}} = & g_{\text{loss}}^{A2B} + d_{\text{loss}}^{A2B} + g_{\text{loss}}^{B2A} + d_{\text{loss}}^{B2A} \\ & + \lambda_{\text{cyc}} \text{cyc}_{\text{loss}}(G_{A2B}, G_{B2A}) \\ & + \lambda_{\text{ide}} \text{ide}_{\text{loss}}(G_{A2B}, G_{B2A}) \quad (4) \end{aligned}$$

式中: λ_{cyc} 和 λ_{ide} 都是权重参数, 通过实验获得。

考虑到张大千字体具有明显的个性特征, 与常见打印字体相差极大, 因此 A 域图像与 B 域图像分别是原始张大千字体图像与对应文字的骨架图像, 即模型在测试使用时将读取文字骨架图像, 并在骨架信息的引导下, 生成带有特定笔画风格特征的大千字体文字。

2 实验分析

2.1 实验环境与评价指标

实验基于 Intel I7 处理器、RTX 2070S 显卡、16 GB RAM、1 TB 存储空间、Python 3.7、PyTorch 1.5、OpenCV 4.1 等软硬件。文字生成效果采用结构相似度 (structural similarity index measure, SSIM) 指标进行量化评价。

$$\text{SSIM}(y_i, S(x_i)) = \frac{(2u_a u_b + c_1)(2\delta_{ab} + c_2)}{(u_a^2 + u_b^2 + c_1)(\delta_a^2 + \delta_b^2 + c_2)} \quad (5)$$

式中: x_i 表示计算文字骨架, 通过网络模型后生成的文字图像 $S(x_i)$ 与标准图像 y_i 之间的相似程度; u_a 和 u_b 分别表示为标准图像和生成图像的均值; δ_a^2 表示标准图像的标准差; δ_{ab} 表示标准图像和生成图像之间的协方差; c_1 和 c_2 表示常量, 分别取 6.502 5、58.522 5。

2.2 数据集

训练网络模型的是张大千字体与对应骨架信息的图像对。从原始图像提取汉字估计时, 如图 3 所示, 分别执行腐蚀与膨胀结合的形态学变换、中值滤波等过程。

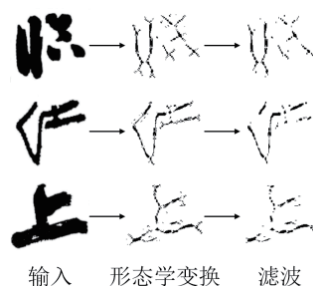


图 3 汉字骨架提取过程

最终如图 4 所示, 该数据集共含有 807 组有效数据, 每组图像对都是大小为 512 px×256 px 的 24 位图。训练时, 按照 6:1:3 的比例划分为训练集、验证集、测试集。

考虑到字体风格迁移模型容易产生过拟合问题, 本文同时构造了另一个相同容量的数据集 2, 如图 5 所示, 用于提升模型的泛化能力。



图4 训练数据集1



图5 训练数据集2

3 实验与讨论

针对模型实现效果，论文进行了对比实验。分别在只使用 zi2zi 模型、CycleGAN 模型、不使用数据集2训练的 zibyz¹ 模型，以及使用数据集1和数据集2混合训练的 zibyz² 模型等四种情况下生成的典型大千字体如图6所示。其中前4个输入的文字骨架图像来自数据集中的测试集，后4个则来自随机手写的文字骨架图像。

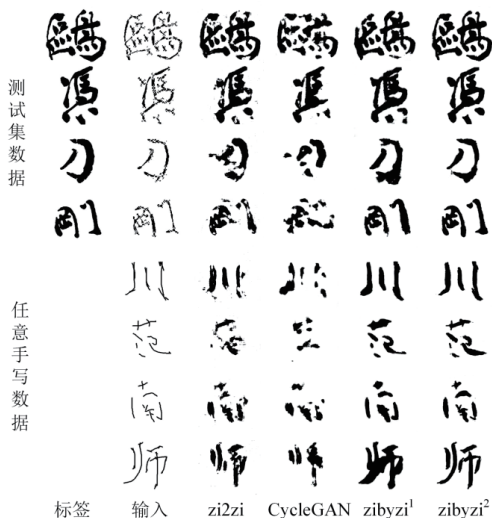
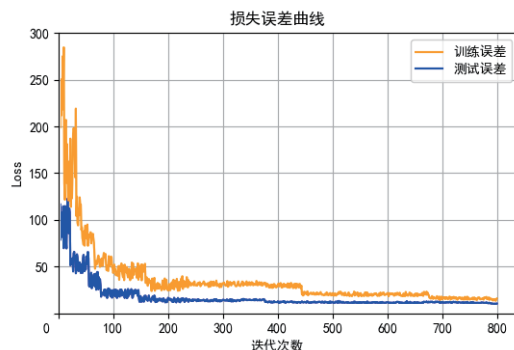


图6 实验对比结果

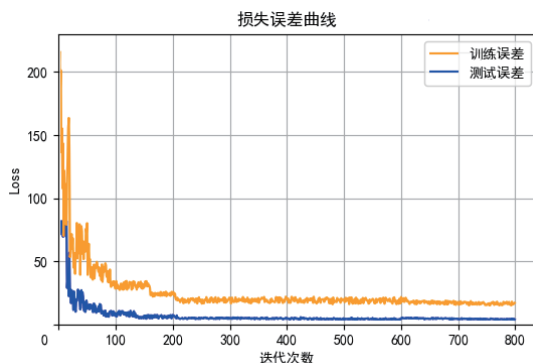
分析可知，zi2zi 模型能实现简单结构汉字的大千字体风格迁移，对结构较为复杂汉字的生成效果较差，训练时发现其分类损失函数能较好收敛，但内容损失函数收敛困难，潜在原因可能是本文的数据集容量对模型而言过低，结合其项目的其他实验结果分析可知，该模型更适合可简单构造大容量数据集的标准打印字体之间的风格转换。CycleGAN 模型的生成效果也欠佳，尽管其双向循环的模式，在一定程度上

扩展了训练数据集容量，但其损失函数中并不包含文字编码损失函数，使得其更适合具有连续纹理特征的图像间的风格迁移，而非具有明显间断空间结构的汉字。

图7所示为不同数据集下 zibyz¹ 模型训练时的损失函数，结合对任意手写文字骨架生成大千字体的效果可知，不使用数据集2训练的 zibyz¹ 模型尽管也能实现收敛，但泛化能力相对较差，最终使得测试误差较大。对于这种深度卷积神经网络模型中常见的过拟合问题，本文额外添加了数据集2进行混合训练后，降低了模型的训练误差，最终在测试集中获得的 SSIM 达到 0.83。



(a) 数据集1训练误差



(b) 数据集1和数据2联合训练误差

图7 文字前景提取效果

4 结论

基于汉字骨架信息，本文实现了字体风格的有效附着，并结合已有的张大千书法作品，形成了大千字库构造方法，文字统一风格附着时采用 zi2zi 网络进行汉字笔画的初生成，并以其为重复组件，借鉴 CycleGAN 网络思想，设计 zibyz¹ 网络后进行笔画细节的再次优化。在测试集中与任意手写汉字骨架的引导下测试中，获得的量化与非量化结果表明了所提出网络模型及使用方法的有效性。尽管基于该方法可以实现大千字体字库的构建，但所采用的大部分输入汉字骨架仍需手动书写完成，极具主观性，且除了笔画细节外，汉字结构也是体现各体书法特色与风格差异的重要载体，因此后续工作将着重研究大千风格汉字骨架的生成方法。

(下转第31页)

的文档相关性判断。实验结果表明,LLM4Search 在核动力知识文档搜索场景中表现出色,不仅提高了文档查找的效率,还加速了核电建设工作的整体推进。这一技术的成功应用,不仅扩展了大语言模型在专业领域应用的理论边界,也为核动力行业的知识管理和信息获取提供了创新的解决方案。

参考文献:

- [1] 李金忠,刘伟东,陈盛博.搜索结果多样化排序:新进展与展望[J/OL].计算机工程与科学,1-26[2025-01-17].http://kns.cnki.net/kcms/detail/43.1258.tp.20241011.1311.006.html.
- [2] AIZAWA A. An information-theoretic perspective of TF-IDF measures[J]. Information processing & management, 2003, 39(1): 45-65.
- [3] ZHAN J T, MAO J X, LIU Y Q, et al. Optimizing dense retrieval model training with hard negatives[EB/OL]. (2014-04-16)[2025-06-21].https://doi.org/10.48550/arXiv.2104.08051.
- [4] 刘华玲,张子龙,彭宏帅.面向闭源大语言模型的增强研究综述[J].计算机科学与探索,2025,19(5):1141-1156.
- [5] YANG A, YANG B S, HUI B Y, et al. Qwen2 technical report[EB/OL].(2024-09-10)[2025-03-23].https://doi.org/10.48550/arXiv.
- [6] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[EB/OL].(2019-07-26)[2025-06-23].10.48550/arXiv.1907.11692.
- [7] NOGUEIRA R, JIANG Z Y, LIN J. Document ranking with a pretrained sequence-to-sequence model[EB/OL]. (2023-03-

14)[2025-06-12].https://doi.org/10.48550/arXiv.

- [8] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[C]//NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 1877-1901.
- [9] 王合庆,魏杰,景红雨,等. Meta-RAG: 基于元数据驱动的电力领域检索增强生成框架[J/OL]. 计算机工程,1-11[2025-01-17].https://doi.org/10.19678/j.issn.1000-3428.0070415.
- [10] ACHIAM J, ADLER S, AGARWAL S, et al. GPT-4 technical report[EB/OL].(2023-03-15)[2025-04-12].https://www.semanticscholar.org/paper/GPT-4-Technical-Report-Achiam-Adler/163b4d6a79a5b19af88b8585456363340d9efd04.
- [11] JARVELIN K, KEKÄLÄINEN J. Cumulated gain-based evaluation of IR techniques[J]. ACM transactions on information systems (TOIS), 2002, 20(4): 422-446.

【作者简介】

徐浩然(1990—),男,四川南充人,硕士,馆员,研究方向:计算机在核动力方面的研究及应用。

曹杰铭(2000—),男,四川广安人,硕士研究生,研究方向:大语言模型, email:2416729460@qq.com。

张瑾昀(1996—),男,四川遂宁人,硕士,研究方向:数字化研发、知识图谱。

颜雄(1992—),男,湖北监利人,硕士,研究方向:计算机应用。

(收稿日期:2025-04-14 修回日期:2025-09-09)

(上接第26页)

参考文献:

- [1] 潘今一,王亚蒙,王伟,等.基于风格迁移和薄板样条的扩充汉字样本方法[J].浙江工业大学学报,2020,48(1):25-29.
- [2] 秦嘉霖,刘维尚.基于直观汉字构形原理的C³-GAN字体生成优化方法[J].包装工程,2023,44(10):193-201.
- [3] 姚伟健,赵征鹏,普国媛,等.稠密自适应生成对抗网络的篆体字风格迁移模型[J].计算机辅助设计与图形学学报,2023,35(6):915-924.
- [4] 陈二开,李成城,邬友,等.基于改进CycleGAN的粉笔字书写风格迁移研究[J].印刷与数字媒体技术研究,2024(6):100-109.
- [5] 张攀,周芳利,杨彪,等.大千字库智能化构建方法研究[J].内江师范学院学报,2024,39(6):55-59.
- [6] zi2zi[EB/OL].(2019-06-26)[2025-02-16].https://github.com/kaonashi-tyc/zi2zi.
- [7] ISOLA P, ZHU J Y, ZHOU T H, et al. Image-to-image translation with conditional adversarial networks[C]//2017 IEEE

Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE,2017: 1125-1134.

- [8] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV).Piscataway:IEEE,2017:2242-2251.
- [9] CHANG B, ZHANG Q, PAN S Y, et al. Generating handwritten chinese characters using CycleGAN[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway:IEEE,2018: 199-207.

【作者简介】

张攀(1989—),男,四川资阳人,硕士,讲师,研究方向:交叉学科、图像处理。

徐瑞娟(2004—),女,江西瑞昌人,本科在读,研究方向:产品设计。

(收稿日期:2025-05-08 修回日期:2025-09-16)