

A Multi-Style Chinese Characters Writing Intelligent Tool Based on Small-scale Training Data

Zhen Zeng, Jie Wang, Xi Lyu

Sichuan Fine Arts Institute
zengzhen@scfai.edu.cn, 2024010023@st.scfai.edu.cn, lucy@scfai.edu.cn

Abstract

Chinese characters are a unique blend of language and art, featuring diverse artistic styles. Mastering these styles requires extensive practice and limits public participation. To encourage broader participation, we developed a real-time, interactive tool that supports multiple Chinese character art styles. This tool uses a diffusion model and several LoRA models to capture the diversity of Chinese character art. It generates personalized, visually striking Chinese character artworks in real-time by utilizing handwritten input, allowing users to adjust various stylistic parameters.

Introduction

Chinese characters are an important part of daily life in China and a key medium for cultural exchange. Their diverse styles have evolved from ancient to modern forms, making them challenging to write, share, and promote.

The rise of GenAI has reshaped Chinese character writing and sharing experience. Since the advent of Generative Adversarial Networks (GAN) (Goodfellow et al. 2014), generative Chinese character art has flourished. Examples include *No Longer Write - Mochiji* (Cheung 2020) and *Cangjie's Poetry* (Zhang and Ren 2021). Additionally, Zeng et al. (2020) developed the *Dry Branches Style Font* using CycleGAN (Zhu et al. 2017). However, during this period, GAN's high data requirements and unpredictable outputs posed challenges in balancing creative freedom and stability.

With diffusion models (Nichol and Dhariwal 2021), tools like D·design enable the quick generation of Chinese character font designs. Users input text, select a font style, and use style prompts to modify the content image and customize texture effects. However, it's important to recognize that generating Chinese characters is a process of writing, not prompting. While text-to-image generation simplifies the creation of Chinese character art, it may detract from the authentic writing experience. Although D·design advances intelligent generation in this field, such tools are not well suited to fostering public engagement in the act of writing.

We developed an intelligent tool to encourage participation from non-professionals, including non-native Chinese speakers. By integrating handwriting and diverse styles, the tool invites public engagement in writing while highlighting the charm of Chinese character art.

Overview

Diverse Chinese Character Art Style Models Trained on Small-Scale Data

We provided 10 Chinese character art style LoRA models (Hu et al. 2021), and built the training dataset by scanning and processing 50 high-resolution images per category, sourced from the photo archives of old books and collections.

We used the WD14-tagger (a plug-in for Stable Diffusion WebUI) to generate tags. For example, when processing images featuring Loong and phoenix, the plug-in generated tags like “drawing”, “no humans”, “SOLO”, “leaf” and “birds”. However, due to the scarcity of traditional Chinese character art in the diffusion model’s training dataset, those tags were often inaccurate. Therefore, we manually adjusted the tags to ensure accurate training of the LoRA model for precise image generation. For the calligraphic style of Loong and phoenix, we organized tags with “HuaNiao fonts” and “colorful” as primary tags, and “Loong,” “phoenix,” “butterfly,” “flower,” “bird” as additional tags. “Chinese characters,” “white background,” “single element,” “no humans,” and “SOLO” described the overall image state.

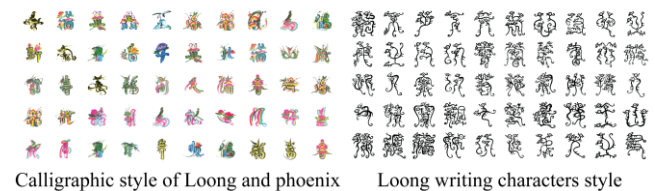


Figure 1: Examples of Chinese Character Art Dataset.



Figure 2: Diagram of Manual Adjustment of Tags

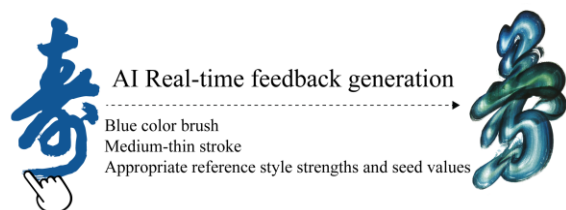


Figure 3: Diagram of Default Parameter Generation

After preparing small-scale training datasets, we trained the models using the LoRA script tool. Once completed, we applied the X/Y/Z Plot script in ComfyUI to determine the optimal model and intensity parameters. To account for distinct color tendencies in some art styles, we tested various brush colors and sizes to ensure the generated results closely matched the training data. We recorded parameters such as seed and strength values. The optimal LoRA model, along with corresponding trigger tags, recommended brush sizes and colors, and ideal strength and seed values, will be set as defaults. These will be automatically applied when users select a style, ensuring the best possible generation.

Comparative Assessment of Generation Effects

The tool's technical process converts the user's handwriting into Base64 format, generating a stylized image with stroke-based input using ComfyUI img2img and LoRA workflow. The final output preserves the user's writing style while faithfully replicating the trained Chinese character art style. The LoRA models trained on small-scale datasets, combined with handwriting aligned with creative actions, reflect the theme of Chinese character art creation. This approach encourages public participation in co-creative writing, enhancing the tool's reception and user experience.

For the generated effects, we compared current AIGC tools supporting img2img creation. Due to the limited performance of Midjourney and DALL·E in image-based generation, we selected a typical stable diffusion img2img workflow combined with the ControlNet plugin, excluding the LoRA models. The following tests used the same diffusion model with same positive prompts for comparison.

The results indicate that the img2img workflow, even when combined with ControlNet, preserves the input image's graphical form but fails to capture the unique stylistic essence of traditional Chinese character art. Similarly,



Figure 4: Comparative Assessment of Generation Effects

Midjourney and DALL·E cannot accurately reproduce Chinese character forms, and their generated styles deviate significantly from the desired effects.

Exhibition and User Feedback

Our work was exhibited at the Sichuan Fine Arts Institute, Chengdu Art Museum, and Chongqing International Expo Center, attracting 120,000 visitors. Most used the tool to create artworks featuring their names, and many non-native Chinese speakers showed great interest in it.

We surveyed 137 random visitors, mostly young and middle-aged non-professionals, who expressed strong interest in Chinese character creation and generative AI. Users operated the tool easily, producing satisfactory works, with most rating their creations 4-5 out of 5. Participants felt the tool deepened their understanding of Chinese character writing and sparked greater interest in the art form.

Conclusion

For general tasks, large models excel in efficiency and accuracy. However, in art, where individuality is key, they struggle to generate personalized content beyond their original training data, and text-based inputs often fall short of creative expectations. This work demonstrates that using handwritten Chinese character strokes with the diffusion model and 10 LoRA models trained on small-scale data, produces superior stylistic effects compared to large models. This approach better captures the aesthetics of Chinese character art while preserving personalized nuances. While this method offers better results, the time required for dataset creation and training limits the diversity of styles in our tool. Future work will expand style diversity to better reflect aesthetic variety and promote broader public participation in Chinese characters writing.

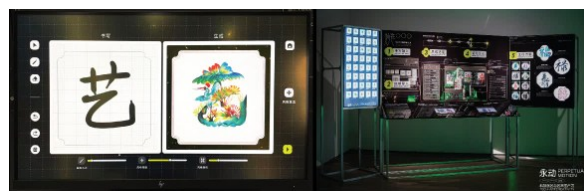


Figure 5: Interactive Display Effects on Site

Acknowledgments

This work was supported by the Science and Technology Program of the Chongqing Education Commission (grant no.KJZD-M202301002), Chongqing Social Science Planning Program (grant no. 2023NDYB195), Sichuan Fine Arts Institute Doctoral Research Program (grant no. 23BSQD003).

References

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. 2014. Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., & Wang, S., et al. 2021. LoRA: Low-Rank Adaptation of Large Language Models. arXiv:2106.09685.
- MANA. 2020. No Longer Write – Mochiji. <https://www.manamana.net/video/detail?id=125962#!zh>. Accessed: 2024-10-02.
- Nichol, A. and Dhariwal, P. 2021. Improved Denoising Diffusion Probabilistic Models. arXiv:2102.09672.
- Zeng, Z., Liao, X. and Lyu X. 2020. Exploration of Chinese Characters Font Design Based on Machine Learning. *PACKAGING ENGINEERING*. 41(18):48-52. doi: 10.19554/j.cnki.1001-3563.2020.18.006.
- Zhang, W., Ren, D. and Legrady, G. 2021. Cangjie's poetry: an interactive art experience of a semantic human-machine reality. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 4(2), pp.1-9.
- Zhu, J.Y., Park, T., Isola, P. and Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).