

AI3604 Computer Vision Project Proposal: Few-Shot Handwritten Chinese Content Generation via Style-Based Generative Adversarial Networks

Kailing Wang
521030910356

Zhiyuan Zhang
521030910377

Xiangyuan Xue
521030910387

Jiazi Bu
521030910395

Qirui Li
521030910397

1. Introduction

Chinese calligraphy is a form of art renowned for its intricate beauty. However, generating Chinese text has long been a challenging endeavor. Unlike Indo-European languages which apply alphabets, Chinese characters are significantly more complex. In this project, we aim to develop an end-to-end system that generates handwritten Chinese contents given style samples and content text. Previous works have tried multiple network architectures including W-Net[3], CycleGAN[2], and Transformer[1]. Style aggregation[6] and visual archetypes[5] are also explored to improve the quality of generated characters. In this project, we want to explore style-based methods[4] for handwritten Chinese character generation, which can enhance the robustness of few-shot learning.

2. Datasets

Plenty of open-source datasets are available for this project. Several datasets which we may use are listed below. In addition, visual archetypes can be rendered with fonts, which can be an extension of our datasets.

- **CASIA**: A widely used Chinese handwriting database. It contains 1,800,000 characters written by 1,020 people.
- **SCUT-EPT**: A non-commercial handwritten dataset. It contains 50,000 text line images selected from examination papers of 2,986 volunteers.
- **ICDAR13-HCTR**: A dataset of handwritten Chinese text recognition. It contains 3,432 text line images.

3. Pipeline

According to our problem description, the task can be decomposed into three stages.

1. Train a handwritten Chinese character generator which can imitate the style of sample characters within few shots. It will be implemented by a style-based GAN.
2. Based on the generator, we will further develop an end-to-end system that generates handwritten Chinese contents given style samples and content text.

3. Cascaded with large language models (LLM) or optical character recognition (OCR), the system can be used to generate creative contents or make text replacement.

4. Task Assignment

The datasets will be collected and preprocessed by Qirui Li and Zhiyuan Zhang. The code part will be cooperatively implemented by Xiangyuan Xue, Kailing Wang, Zhiyuan Zhang and Jiazi Bu. The experiments and fine-tuning will be done by Jiazi Bu and Qirui Li. The report will be mainly written by Kailing Wang and Xiangyuan Xue, but everyone will contribute. More details to be determined.

References

- [1] Ankan Kumar Bhunia, Salman Khan, Hisham Cholakkal, Rao Muhammad Anwer, Fahad Shahbaz Khan, and Mubarak Shah. Handwriting transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1086–1094, 2021. 1
- [2] Bo Chang, Qiong Zhang, Shenyi Pan, and Lili Meng. Generating handwritten chinese characters using cyclegan. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 199–207. IEEE, 2018. 1
- [3] Haochuan Jiang, Guanyu Yang, Kaizhu Huang, and Rui Zhang. W-net: one-shot arbitrary-style chinese character generation with deep neural networks. In *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part V 25*, pages 483–493. Springer, 2018. 1
- [4] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019. 1
- [5] Vittorio Pippi, Silvia Cascianelli, and Rita Cucchiara. Handwritten text generation from visual archetypes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22458–22467, 2023. 1
- [6] Licheng Tang, Yiyang Cai, Jiaming Liu, Zhibin Hong, Mingming Gong, Minhu Fan, Junyu Han, Jingtuo Liu, Errui Ding, and Jingdong Wang. Few-shot font generation by learning fine-grained local styles. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7895–7904, 2022. 1