MaLiang: An Emotion-driven Chinese Calligraphy Artwork Composition System

Ruixue Liu[*]
liuruixue@jd.com
JD AI, Beijing, China

Shaozu Yuan[*]
yuanshaozu@jd.com
JD AI, Beijing, China

Meng Chen
chenmeng20@jd.com
JD AI, Beijing, China

Baoyang Chen
chenbaoyang@cafa.edu.cn
CAFA, Beijing, China

Zhijie Qiu
qiuzhijie@cafa.edu.cn
CAFA, Beijing, China

Xiaodong He
xiaodong.he@jd.com
JD AI, Beijing, China

ABSTRACT

We present a novel Chinese calligraphy artwork composition system (MaLiang) which can generate **aesthetic**, **stylistic** and **diverse** calligraphy images based on the emotion status from the input text. Different from previous research, it's the first work to endow the calligraphy synthesis with the ability to express fickle emotions and composite a whole piece of discourse-level calligraphy artwork instead of single character images. The system consists of three modules: emotion detection, character image generation, and layout prediction. As a creative form of interactive art, MaLiang has been exhibited in several famous international art festivals[1].

CCS CONCEPTS

• **Applied computing** → **Fine arts**; • **Computing methodologies** → **Image representations**.

KEYWORDS

Interactive Art; Calligraphy Generation; GAN

1 INTRODUCTION

Calligraphy, the art of writing, is an indispensable part of Asian culture. Its composition is influenced by the inner psychology and emotion of calligrapher. Emotion matters to calligraphy composition. The stroke movement of character and the overall layout of artwork change vividly along the calligrapher's emotion transition (shown in Figure 1). Rich emotions make the calligraphy artwork vivid and impressive. It's essential to take emotion information into account for calligraphy synthesis.

However, previous research on Chinese calligraphy generation neglects the emotion's contribution. With Generative Adversarial Networks (GANs) [4], there are some attempts made to generate calligraphy images with pix2pix [6] or CycleGAN [1]. They only treat
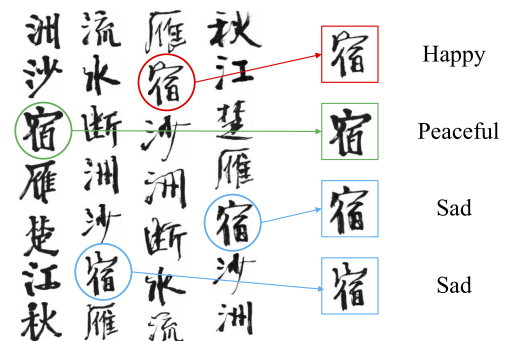

Figure 1: An example of generated calligraphy artwork with diverse emotions. The artwork reflects the emotion transition vividly (box in different colour) and also illustrates the diversity of generation (box in the same colour).

the Chinese calligraphy generation as a general-purpose image-to-image translation task and overlook the artistry and completeness of it as a piece of artwork, thus they can't produce *aesthetic*, *stylistic*, and *diverse* character images with different emotions. There are several challenges of generating emotion-driven Chinese calligraphy artworks: 1) how to catch the fickle emotions and connect them with the artistic styles of calligraphy? 2) how to arrange the layout of all character images into a complete piece of artwork and reflect the emotion changes? 3) how to generate diverse calligraphy images like human beings, as real calligrapher doesn't composite exactly the same image each time. To tackle above challenges, we present a novel system to finish the task in three steps. It firstly detects the emotion for each sentence in the input text based on an emotion classifier. Then, it generates the emotional stylistic image for each character with a modified GAN structure. Thirdly, by combining both linguistic and visual information of the character, it predicts the layout features with a Recurrent Neutral Network (RNN). Due to the lack of off-the-shelf sentimental image generation corpus, we collaborate with a famous calligrapher[2] in China and collect a large-scale Chinese calligraphy image dataset with rich emotions. To the best of our knowledge, MaLiang is the first system for emotion-driven discourse-level Chinese calligraphy composition.
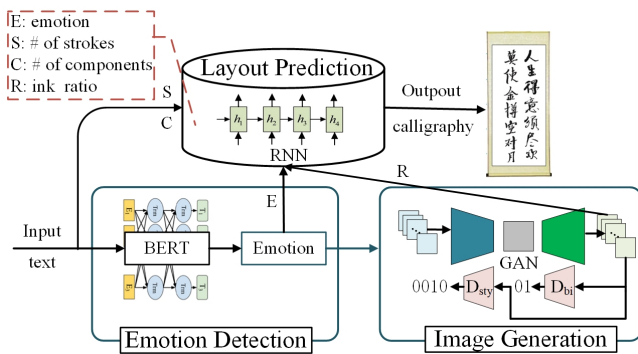

Figure 2: System Architecture

Figure 3: Model performance comparisons.

2 ARCHITECTURE

As shown in Figure 2, our system consists of three modules: 1) an emotion detection module which detects emotions from the input text, 2) an image generation module which transfers character image from standard font to a specific emotional style, 3) a layout prediction module which arranges all single character images into a complete piece of artwork.

Emotion Detection. To catch the fickle emotions from the input text, we first build a sentiment classifier based on state-of-the-art pre-trained language model BERT [3]. Then each line is analyzed and the emotion transition pattern is obtained by the classifier. As the text of Chinese calligraphy is usually from classical Chinese articles, we fine-tune the model with a manually-labelled fine-grained sentiment poetry corpus [2], which covers 3 sentiment classes, namely peaceful, happy, and sad.

Image Generation. Take the emotion label as input, we formulate the emotional Chinese calligraphy generation as an unsupervised image-to-image translation task. Inspired by U-GAT-IT [5], we propose a novel calligraphy generation model based on the popular GAN structure [4], which concentrates on the semantically more important areas with the help of attention mechanism. To promote the quality of generated images, we have two special designs. Firstly, to distinguish the style difference between various emotions, we propose a style discriminator D_{sty} to distinguish the generated images from multiple emotional styles. Secondly, to promote the diversity of generated images, we implement a multi-encoder ensemble modeling technique during the inference stage. During training, we train auxiliary models with different initialization of hyper-parameters. The best-performed model on test set is selected as the base model. During inference, we ensemble the base encoder and auxiliary encoders to create stochastic variations. Figure 1 shows a generated calligraphy artwork. Figure 3 illustrates the comparisons of the character images generated by our proposed model (ECalliGAN) and several baselines (zi2zi [6], CycleGAN [1], and U-GAT-IT [5]). We set the standard font (Liukai) as source image, and the three emotional styles as target image. These examples show the effectiveness of our strategy for generating stylistic and diverse calligraphy images.

Layout Prediction. The layout of calligraphy artwork includes the size of character, the character spacing, and the line spacing. It not only fo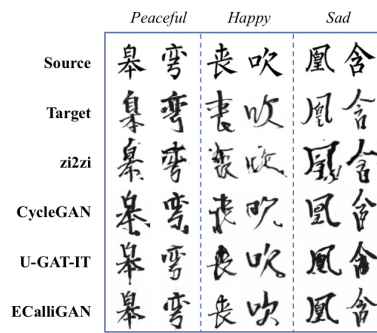llows some writing habits of the calligrapher, but is also influenced by the emotions. Here, we propose to model the character size prediction as sequential modeling problem with a vanilla recurrent neural network (RNN). We feed the following features into the RNN as input: 1) emotion label, 2) ink density of each character image (the ratio of black pixels to total pixels in a standard square box), 3) number of strokes in a character, 4) number of components of a character. For the character and line spacing, we adopt the rules from the professional calligrapher directly.

3 DEMONSTRATION

We implement a demo application of MaLiang, which allows the user to create Chinese calligraphy artworks interactively. It's developed with python and the models are trained with Tensorflow. The demo application needs GPU server (NVIDIA Tesla P40) for real-time running. The system is very easy to use. The user only needs to enter the text and click the button, then the whole piece of Chinese calligraphy artwork will be generated. Besides, more logging information such as the predicted emotion labels and layout features will also be printed out for better understanding the system. Please watch the demo video in the supplemental materials for more details.

4 CONCLUSION AND FUTURE WORK

In this paper, we present MaLiang, a Chinese calligraphy generation system, that can synthesis *aesthetic*, *stylistic*, and *diverse* discourse-level calligraphy artwork with rich emotions. MaLiang utilizes a sentiment classifier to analyze the given text. Then it adopts a multi-style discriminative GAN to integrate the emotion information into character image generation. We also define the dominant features for calligraphy layout and propose a layout network to sequentially model these features with linguistic and visual information from characters. In the future, we will enrich the emotional styles of the generated calligraphy such as *lonely, cheerful, fearful* and make the demo public available.

5 ACKNOWLEDGMENTS

REFERENCES

[1] Bo Chang, Qiong Zhang, Shenyi Pan, and Lili Meng. 2018. Generating handwritten chinese characters using cyclegan. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 199–207.

[2] Huimin Chen, Xiaoyuan Yi, Maosong Sun, Wenhao Li, Cheng Yang, and Zhipeng Guo. 2019. Sentiment-controllable Chinese poetry generation. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*.

[3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.

[5] Junho Kim, Minjae Kim, Hyeonwoo Kang, and Kwang Hee Lee. 2020. U-GAT-IT: Unsupervised Generative Attentional Networks with Adaptive Layer-Instance Normalization for Image-to-Image Translation. In *International Conference on Learning Representations*.

[6] Yuchen Tian. 2017. Master chinese calligraphy with conditional adversarial networks. https://github.com/kaonashi-tyc/zi2zi.