

Machine Learning Approach on AI Painter: Chinese Traditional Painting Classification and Creation

Kaiqi QIU ¹, kaiqiqiu@buaa.edu.cn

Feiru WANG ², Master of Materials Engineering, Jilin University, wangfr17@jlu.edu.cn

Yingxi TANG*, Zhicheng International Academy, yingxitang@buaa.edu.cn

Keywords: *Chinese traditional painting—AI painter—DALL•E—VGG*

CHNT Reference: Qiu, Wang, and Tang (2022). 'Machine learning approach on AI painter: Chinese traditional painting classification and creation', in CHNT Editorial board. Proceedings of the 27th International Conference on Cultural Heritage and New Technologies, November 2022. Heidelberg: Propylaeum.

Abstract

Chinese traditional landscape paintings, as one of the treasures of Chinese culture, are unique to human civilization. Among these art works, some incorporate black ink with Chinese art paper to present their story, while others, such as *A Panorama of Rivers and Mountains (QianLi JiangShan Tu)*, brought authentic hue in which sparkle brilliantly throughout Chinese history. Western oil paintings, which hold great differences with Chinese traditional painting in techniques and forms of expression, are also popular worldwide with their unique artistic value. With the extended application of AI, the original creativity of AI has been widely discussed. Nevertheless, the fusion of AI painting with traditional Chinese painting is still insufficient. In this paper, multiple mainstream machine learning models are applied to research to compare their efficiency in classification and creation. The first part of our research is the training of classification models between Chinese landscape painting and western oil painting and then apply them into accuracy experiments. The data set we applied within the training and the test is the painting data set labeled by team members based on former contributor's work. We applied VGG and Support Vector Machine as classification models. After training, we used test set and acquired the average recognition accuracy rates of 98.5% for VGG and 88.33% for SVM, respectively. The second part of research is devoted to the experiment of AI painting creation. Within this research, we decided to use SAP-GAN, DALL-E 2, together with Night Cafe to compare their ability to recreate Chinese landscape paintings. The result of our focus group research shows that the DALL-E 2 model combined with the style transfer of Night Cafe produced the most outstanding result.

Introduction

With its distinct characteristics and long tradition, Chinese traditional painting is a valuable cultural heritage of global significance. We mainly put our focus on Wang Ximeng, a young talented painter in the Song dynasty (Ruili Zheng, 2012), one of the masters in traditional landscape painting. His painting is a masterpiece of "blue-green landscape" (*qinglü shanshui*) (The Palace Museum, 2020).

Wang Ximeng did not use synthetic colors when he painted, but all mineral colors. The pigment that he applied are extracted from graphite, malachite, blue copper ore, lapis lazuli, cinnabar, estrogen, ochre and Tridacnina (Giant Clams). With these minerals, the pigments can be preserved for thousands of years.



Fig. 1. Natural mineral pigments: (a) malachite, (b) blue bronze ore, (c) lapis lazuli, (d) cinnabar, (e) estrogen, (f) ochre, and (g) giant clams (Rui Wang and Zunli Mo, 2022)



Fig. 2. A Panorama of Rivers and Mountains (Partial) (www.dpm.org.cn, 2022)

Though the masterpiece he had created, he left the public's sight without a precise record of his fate not long after he completed *A Panorama of Rivers and Mountains*. This masterpiece also became the only known painting that is preserved. The artistic value and historical value are all extraordinary. The former application of AI within art creation previously includes several applications for art creation. The first one is the Generative Adversarial Network (Sketch-and-Paint GAN), the first end-to-end framework capable of producing high-quality Chinese paintings with intelligible, edge-defined landscapes (Alice Xue, 2020). However, the GAN model is not advanced enough because the AI techniques are constantly improving and more algorithms are gradually applied. The second one is DALL-E and its improved version of DALL-E 2 (Aditya Ramesh et al., 2022). The third one is Imagen developed by Google also appeared recently (Chitwan Saharia et al., 2022). As AI techniques are rapidly developing, the Generative Adversarial Networks may underperform the existing AI systems, which we decided to introduce the DALL-E 2 model to perform the imitation of the original style of Chinese landscape painting.

Our experiment is divided into two steps, the first task that we are completing is the classification of the drawings with our classification modes and continued with the creation of new picture. We compared the efficiency and the outcomes of multiple models and finally decided to choose the VGG model with the greatest efficiency.

Method

To train our model, we constructed data sets that are consisted by paintings that falls into three categories, as shown in figure 3 and 4. The third dataset is created through cropping toward the original picture of *A Panorama of Rivers and Mountains*.

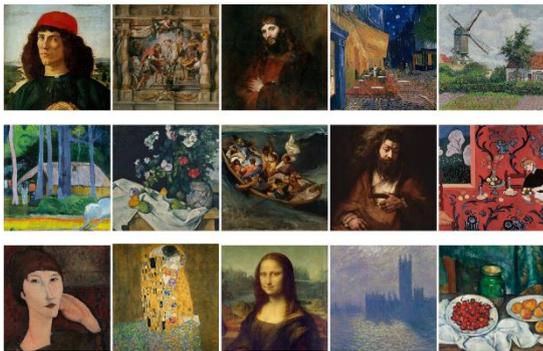


Fig. 3. Examples of western painting data set

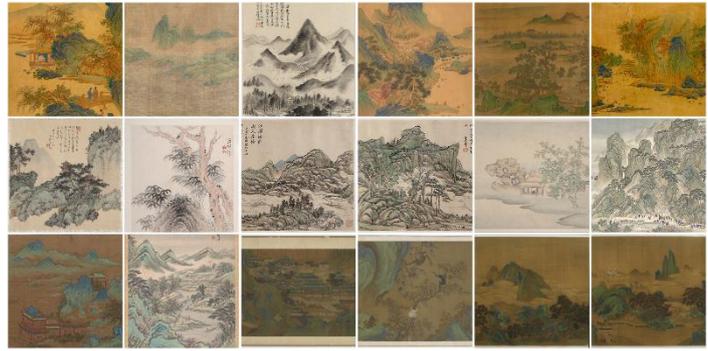


Fig. 4. Examples of traditional Chinese paintings data set by Alice Xue (2020)

Classification Models

To reach the best effect of our classification model, we researched and experimented with the following algorithms for classification and compared their effect.

LASSO is a feature selection approach based on a linear regression model with L1 -norm regularization. LASSO was first proposed by Robert Tibshirani in 1996 as the Least absolute shrinkage and selection operator and is a compression estimator. The feature set extracted in this paper better expresses the artistic style of ink painting works and improves the classification recognition performance and robustness of ink painting works.(Wang Min, et al., 2015)

Support vector machines (SVMs) are binary classification models whose basic model is a linear classifier defined by maximizing the interval on the feature space, which distinguishes it from a perceptron (Shuqiang Jiang, et al., 2003). In recent years SVM learning has been applied to a wide range of real-world applications where it has been found to offer superior performance. (Shuqiang Jiang, et al., 2003). VGG16 as a famous model was selected as well.

Creation model

For the next step, we continued to reinforce the effect of machine creation, and we further compared the following models for the creation effect to demonstrate the effect and choose the best one as our final application. Current GAN-based art generation methods produce unoriginal artwork due to their dependence on conditional input. Here, we propose Sketch-And-Paint GAN (SAP-GAN), the first model which generates Chinese landscape paintings from end to end, without conditional input (Alice

Xue, 2020). We describe a DALL-E approach for this task based on a transformer that autoregressively models the text and image tokens as a single stream of data. With sufficient data and scale, our approach is competitive with previous domain-specific models when evaluated in a zero-shot fashion (Aditya Ramesh, 2022). Imagen is a text-to-image diffusion model with an unprecedented degree of photorealism and a deep level of language understanding (Chitwan Saharia, 2022). In mid-2021, the online art scene exploded with the invention of an amazing text-to-image art generation method known as VQGAN+CLIP. It quickly became the app's most popular creation method (nightcafe.studio, 2022), and Night Cafe was popular tool as well.

Experimental results

For the classification model, we acquired a precision rate result of VGG 98.5% and SVM 88.33%. This result is satisfying for our expectation of realizing the classification between western paintings and Chinese traditional paintings.

After completing the training for the model that involved the DALL-E 2 and the GAN-based creation, we discovered that the ink landscape painting created by DALL-E 2 model is better than the GAN-based model.



Fig. 5. Chinese painting created by DALL-E 2 model



Fig. 6. Figure created by applying DALL-E to create ink landscape painting and further apply style transferring by Night Cafe

Afterward, we discovered that the training set need to be simple scenery composed of azurite, slate blue mineral yellow dataset, which produced satisfactory results as Fig.6 shows.

Finally, we further discovered that in the creation of Blue-green landscape (Qinglü Shanshui), the best choice was combining the style creation of night cafe with the creation of DALL-E 2. During our evaluation of the paintings that are processed by night cafes, the style of Wang Ximeng is instantly recognized by human expertise.

Focus group research

We invited twenty experts in Chinese traditional painting to evaluate the drawings. Participants are first shown with both the drawing from the person and the drawings composed of our model. Then the specialists are asked whether they believe the painting is created by human or computer models.

Table 1. Statistics of Scoring by Human Expertise

	Similarity to a human painting	Similarity to the style of <i>A Panorama of Rivers and Mountains</i>	Clarity
SAP-GAN	88%	60%	85%
DALL-E 2	92%	65%	90%
DALL+Night Cafe	96%	98%	96%

From the scoring results above, DALL-E created more delicate pictures and was more recognized by professional painters than the GAN network for creating the traditional Chinese black and white landscape paintings. What's more, after the creation of DALL-E accompanied by the style transferring produces the best result in the field of Blue-green landscape.

Conclusion

The result reveals that DALL-E 2 model was more accurate and outperformed the GAN-based art generation, which was very suitable for the identification and dating of Chinese ancient paintings. After training the machine, we can receive the results that are similar to the style of Chinese traditional landscape painting, which is classified as Qing Lyu Shan Shui, and successfully acquire the result that is similar to the original style when we integrated DALL-E 2 with Night Café as style transfer.

Transfer learning

We decided to apply the same method to the masterpieces of Huang Gongwang. His masterpiece *The Dwelling in Fuchun Mountain* is one of the most famous artworks in Chinese traditional painting. The interesting result shows the Ximeng Wang style of *The Dwelling in Fuchun Mountain*.



Figure 7: DALL E+Night Cafe integrated model applied on *The Dwelling in Fuchun Mountain*

References

Zheng, R. (2012). "An Analysis of the Color Language of Wang Ximeng's 'A Panorama of Rivers and Mountains'", *Beauty and Times: Journal of Fine Arts*, 000(008), pp. 61—62. [Online]. available at: <http://www.cnki.com.cn/Article/CJFDTotal-MSDZ201208018.htm> (Accessed: 29 May 2022).

- A Panorama of Rivers and Mountains[The Palace Museum. (2020). [Online]. available at: https://en.dpm.org.cn/collections/collections/2015-03-27/843.html?utm_medium=website&utm_source=archdaily.com (Accessed: 27 May 2022).
- Wang, R. and Mo, Z. (2022). Science Popularization Competition | Spring Festival Dance "The Journey of a Legendary Landscape Painting": The Chemical Code in the Heirloom Painting "Thousand Miles of Rivers and Mountains, Koushare.com. [Online]. available at: <https://www.koushare.com/post/postdetail/5998> (Accessed: 10 June 2022).
- Wang Ximeng's painting of a thousand miles of rivers and mountains. (no date). [Online]. available at: <https://www.dpm.org.cn/collection/paint/228354.html> (Accessed: 27 May 2022).
- Liang, X. (2022). View | "Only this green" Just this glance is a thousand years, Moment.rednet.cn. [Online]. available at: <https://moment.rednet.cn/pc/content/2022/02/11/10862313.html> (Accessed: 26 May 2022).
- Alice, X. (2020). "End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks", IEEE Winter Conference on Applications of Computer Vision (WACV) 2021, p. 9. [Online]. available at: <https://arxiv.org/abs/2011.05552> (Accessed: 15 June 2022).
- Ramesh, A. et al. (2022). "Hierarchical Text-Conditional Image Generation with CLIP Latents", p. 27. [Online]. available at: <https://arxiv.org/abs/2204.06125> (Accessed: 12 June 2022).
- Saharia, C. et al. (2022). "Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding", Imagen, p. 46. [Online]. available at: <https://arxiv.org/abs/2205.11487> (Accessed: 13 June 2022).
- Jiang, S., Gao, W. and Wang, W. (2003). "Classifying traditional Chinese painting images", Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint, 3, p. 5. [Online]. available at: https://www.researchgate.net/publication/4072550_Classifying_traditional_Chinese_painting_images (Accessed: 26 June 2022).
- Wang, M. et al. (2015). "Ink painting classification recognition by heterogeneous feature group selection", Information & Communications, 152, p. 2. [Online]. available at: <https://wenku.baidu.com/view/3dba621f80eb6294dc886c71.html?fr=xueshu> (Accessed: 25 June 2022).
- "About Nightcafe Studio". Nightcafe Creator, 2022, [online] <https://nightcafe.studio/pages/about-nightcafe>(Accessed 12 July 2022)

I. Conflict of Interests Disclosure

There are no financial or personal relationships with other individuals or organisations, such as sponsors, that could make this research work appear biased or influenced.