

A Method for Traditional Shadow Play Figure's Head Regeneration Based on Generative Adversarial Network

Xiaofang HUANG ^{a,1} and Jingzhou HUANG ^b

^a College of Art and Communication, China Jiliang University, Hangzhou 310018

^b School of Automation, Hangzhou Dianzi University, Hangzhou 310018

Abstract. The field of artistic style transfer and creation has introduced more and more intelligent methods. We apply the technology of generative adversarial network to sculpt creation for traditional shadow play figure's head, and suggest a recreation method based on generative adversarial network for traditional shadow play figure's head. Firstly, we collect and make data sets of side face of real people and tradition shadow play figure's head; Secondly, we train a CycleGAN model using former data sets and obtain transformed results from side face to shadow play figure's head. Lastly, we analyze experiment results, expose existing problems and present future work. Our application method provides thoughts for developing the art of shadow play with scientific approaches.

Keywords. Artistic style transfer, sculpt innovation of shadow play figure's head, generative adversarial network

1. Introduction

With the development of science and technology, the increasing popularity of new media has impacted the inheritance of some excellent traditional Chinese culture, and many ancient arts are on the verge of extinction. Shadow play is one of them. Shadow play is regarded as the epitome of folk culture, which can not only show the regional cultural characteristics, but also create economic value. Traditional shadow puppets can be divided into shadow play figures and others, among which heads are the most complicated. In this paper, shadow play figures' heads are selected as the research object, trying to carry out new artistic creation with the support of technology, so that it not only retains traditional characteristics, but also integrates the needs of the times, and radiates new artistic charm.

This article applies generative adversarial network technology in the field of modern intelligent image processing to traditional shadow play figure' head regeneration. It is one of the important exploration innovation researches of the traditional shadow play art, as culture spirit inheritance opened wider space, at the same time for other inheritance of in-tangible cultural heritage protection and innovation development provides reference.

¹ Corresponding Author.

The structure of this paper is as follows: Section 1 briefly introduces the research background and purpose; The second section introduces the research trend at home and abroad and determines the research scheme; Section 3 describes the implementation process of the research scheme; Section 4 analyzes the experimental results and points out the shortcomings of current research and the direction of improvement. Section 5 summarizes the content of this article.

2. Related Work

2.1. Research and development trend of shadow play figures

Since 1960s, there have been some studies on the origin, formation and analysis for characteristics of shadow play figure in different regions [1,2], which are of great significance and value. In the 21st century, digital shadow play figure is generated by combining shadow play art and digital technology [3,4]. With the help of digital technology, traditional shadow play figures are rejuvenated and better protected and promoted.

In contrast, there are relatively less researches on innovation of shadow play figure sculpt. From the perspective of shadow play prototype and dimension, innovation of shadow play figure sculpt can be roughly divided into four categories: (1) Virtual two-dimensional shadow play figure modeling. For example, computer graphics technology can be used to generate real-time animated interactive games [5], in which the sculptures come from physical shadow play puppets that are portrayed by artists or craftsmen. (2) Two-dimensional shadow puppet sculpt based on real people or objects. For instance, high-resolution digital images are obtained by shooting real objects, and then these images are processed for shadow play figure modeling [6]. (3) Three-dimensional virtual shadow puppet sculpt. For example, we can use Maya MEL animation programming language to sculpt and equip characters [7]. (4) Three-dimensional shadow puppet modeling based on real people or objects. For instance, 3D animation [8] can be generated by collecting human body structure and action with Kinect depth sensor, or shadow play character sculpt and animation [9] can be generated based on 3D scanning data of human body. Comparatively speaking, modeling methods based on real data can improve people's attention and interest, and the current rapid development of intelligent image processing methods provide support for it.

2.2. Research status of face art style transfer technology

Liang [10] divided face style transfer methods into three categories: face stylization based on facial features, matching synthesis and deep learning. The first two types of methods have very complicated process and the results are not real and natural. The third method emerged with the development of deep learning technology, which can achieve high-quality end-to-end image domain conversion, and has achieved remarkable results. Dong [11] classified the research on image style transfer based on deep learning technology into three categories: methods based on iterative optimization [12], transfer network [13] and Generative Adversarial Network (GAN) [14]. The core idea of GAN is to make the distribution of generated image domain approach the distribution of target image domain by using the game between generative network and discriminant network. From the perspective of using supervised learning or not, image transformation algorithms can be divided into two categories: GAN based on paired data and unpaired

data [15]. To the head regeneration task, we have to choose the unsupervised training method because the one-to-one corresponding data in two image domains cannot be obtained.

CycleGAN [16] is one typical representative of generation adversarial models for image domain transformation based on unsupervised training. The proposed cycle consistency loss restrains image reconstruction by bidirectional transformation of source-target-source domain and target-source-target domain, so that the features of images can be preserved and significant semantic differences can be avoided. DiscoGAN [17] and DualGAN [18] are similar works.

At present, GAN has been successfully applied in a large number of unsupervised image domain transformation tasks [19], which provides excellent references for the study of this paper. However, compared with the current research, the transformation in this paper still has the following challenges: (1) In the existing applications, the transformation is mostly completed by converting and enhancing features and the image content remains unchanged, while the transformation objective in this paper requires a great change of image content; (2) In existing applications, the transformation does not require the generation of new image content, while this paper requires the generation of non-existing image content (headwear); (3) The facial features of the figures are changeable, while the shadow play figures have few facial changes, which makes it difficult to show different features of the real face. The transformation objective for this paper can be represented in Fig. 1. Obviously, the transformation task in this paper is more complex. In the existing works, CycleGAN has better universality and wider application, and is also the solution of many successful cases, so we make it as the basic technology of this paper.



Figure.1 Transformation target diagram

3. Shadow play figure' head regeneration based on CycleGAN

The conversion process is shown in Fig. 2. Two data sets are used in the training process. Data set 'A' is a collection of real human side facial images, data set 'B' is a collection of traditional shadow play figures' heads. Also, two data sets are used in the testing process. Data set 'C' is a collection of real human side facial images, data set 'D' is a collection of newly generated images of traditional shadow play figures' heads.

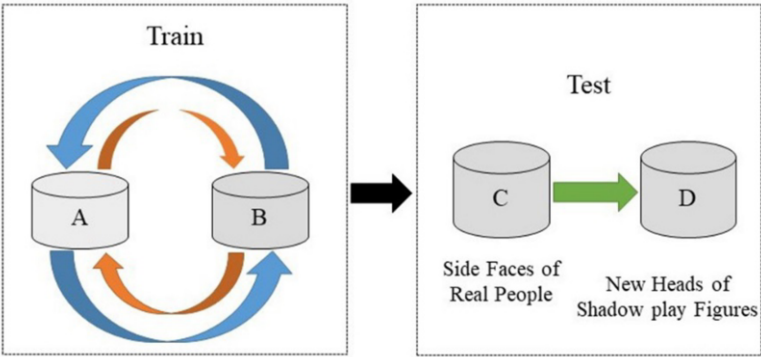


Figure 2. Flow chart

3.1. Data acquisition and preprocessing

This paper selects heads of traditional shadow play figure from various regional genres that are close to the human morphology, and establishes a data set of traditional shadow play figures' heads with a capacity of 500 labeled 'B'. To reduce the difficulty of the conversion problem from real face to traditional shadow play figure's head, and avoid possible problems caused by inconsistent view angles and inconsistent conversion subjects to image proportions, and improve the quality of transformation, we select real human facial images with the same view of the traditional shadow play figures' head images, and establish the data set of the facial images with a capacity of 500 labeled 'A'. An example of data set A and B is shown in Fig. 3. Left image is from dataset A, and right image is from dataset B.



Figure 3. Sample dataset

3.2. Real side face to traditional shadow play figure's head conversion

3.2.1. Principle of image transformation based on CycleGAN

CycleGAN can accomplish the transformation between any two image domains, and the training process does not require the one-to-one correspondence between the data of the two image domains. This unsupervised training method is especially suitable for the transformation between image sets that are unlabeled or difficult to label. Another advantage of CycleGAN is that the data sets to be transformed are of equal priority, which means that the training data sets can be specified in any order.

The principle of CycleGAN is as follows: Assume X and Y respectively stand for the two image domains to be converted; the transformation function from X to Y is G,

the transformation function from Y to X is F, and the discriminators for X and Y are respectively D_X and D_Y , as shown in Fig. 4. The transformation function G has the function of translating the data in X to Y, which is also called the generator from X to Y, and F is the corresponding reverse generator. D_X and D_Y are a set of discriminators, which are used to judge whether the features of transformed data are consistent or not with that of the current image domain.

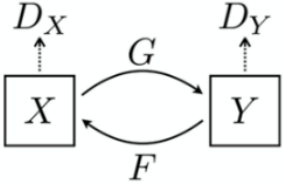


Figure 4. schematic diagram for CycleGAN framework

There are two steps in one training iteration of CycleGAN. The first step is to train two generators, which optimizes parameters based on the cycle-consistency loss in order to make the difference as small as possible between the data before and after two transformations. The second step is to train the two discriminators so that D_X and D_Y have the ability to accurately estimate if the current input is the data that are transformed to the corresponding image domain. This two-step training process embodies the idea of “adversarial”. The generator tries to capture the features of two image domains to achieve one-way conversion, and then uses the discriminator to estimate whether the features of transformed images are close to that of images in the target domain or not. Through alternate training, generators become more and more capable of transforming, and discriminators become more and more capable of estimating.

3.2.2. Experimental environment and settings

This paper adopts the open-source implementation of CycleGAN [16] to conduct experiments. The side face data set of real people is as image domain X, and the head data set of shadow play figure sculptures is as image domain Y, 50 images are randomly selected from the two data sets as the test set. We run the program with the option --NO_FLIP, and other parameters are consistent with the default ones.

4. Experimental Results and Analysis

Some experimental results are shown in Figure 5. As can be seen from Figure 4, some realistic shadow play figure sculptures can be generated based on the side view of a real person’s head, and the silhouette of the shadow play figure is basically consistent with the hair-style of the real character. However, in terms of facial curve features, it is difficult to find the corresponding relationship before and after transformation from shadow play figure sculpt. In addition, the generated shadow play figure headdress only has similar texture, while the modeling components are far different from the real shadow play figure, indicating that the model cannot distinguish the headdress composition from the real shadow play figure. Furthermore, across all transformations, there are few good outcomes and reverse transformations do not work well.



Figure 5. Partial results of the overall transformation

We selected 24 users to evaluate the newly generated images of shadow play figures’ heads. The evaluation results are shown in the Tab.1 and Fig.6.

Table 1. The statistical table of user evaluation

User Quantity	Average Degree of Satisfaction	Variance of Satisfaction	Standard Deviation of Satisfaction
24	52.7%	29.78%	5.46%

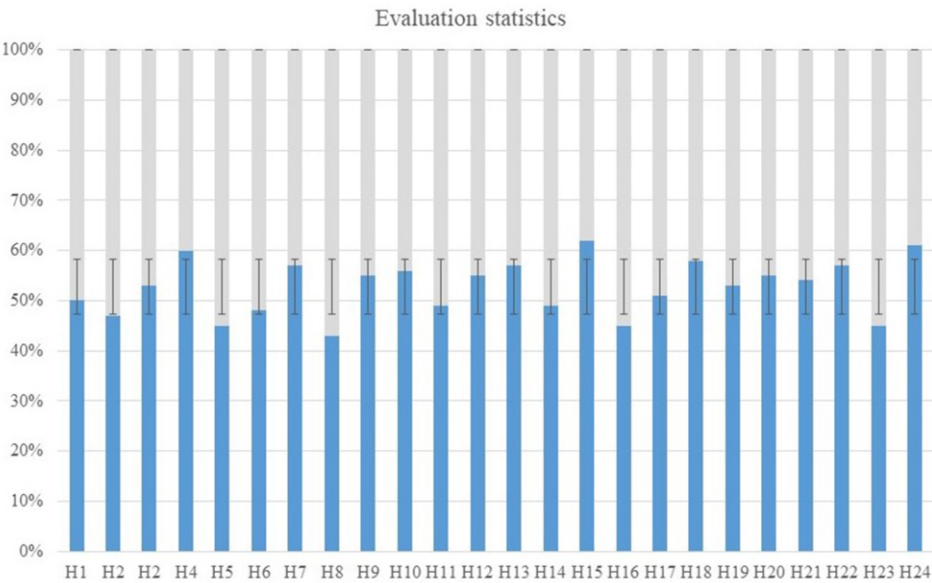


Figure 6. The statistical table of user evaluation

5. Conclusion

In this paper, CycleGAN is used to regenerate the head of shadow play figure, and some effects have been achieved. However, due to the complexity of the problem, there is still a big gap from the ideal sculpt. In the future, we can start from two aspects:

(1) For face transformation, considering the fusion of facial curve features between real person and shadow play figure, it can mainly retain the facial features of people, break through the limitation of lack of changes in traditional shadow play figure facial sculpt, and create personalized shadow play figure facial sculpt;

(2) For the transformation of hair accessories, the transformation of style and color can be separated, and the transformation relationship between modern hair styles and shadow play figure headdress is established, so that the generated shadow play figure headdress can conform to the characteristics of traditional structure.

Acknowledgement

This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ18E050008.

References

- [1] Wang I C. *Globalization and theater spectacles in Asia*. CLCWeb: Comparative Literature and Culture, **15** (2013), 2, 1-8
- [2] Sedana I N. Innovation of Wayang puppet theatre in Bali //Xing J, Ng P. *Indigenous Culture, Education and Globalization*, vol 1. 1rd ed. Berlin: Springer Berlin Heidelberg, 2016: 67-80
- [3] Peng Jianbo. *Characteristic resource building for the intangible cultural heritage - based on the shadow play digital museum*. Library Work and Study, **5** (2012), 5, 33-36(in Chinese)
- [4] Huang Xiaofang. *Personalized modelling research of shadow puppets based on traditional features*. Art & Design, **1** (2019), 2, 140-141(in Chinese)
- [5] Lam T K, Talib A Z B H, Mohd. Azam Osman. *Real-time visual simulation and interactive animation of shadow play puppets using OpenGL*. International Journal of Information, Control and Computer Sciences, **2** (2008), 9, 2934-2940
- [6] Zhu Y B, Li C J, Shen I F, et al. *A new form of traditional art: visual simulation of Chinese shadow play* //Proceedings of Special Interest Group on Computer Graphics and Interactive Techniques Conference. New York: Association for Computing Machinery, 2003
- [7] Ghani D A. *Digital puppetry: comparative visual studies between Javanese & Malaysian art*. International Journal of Applied Engineering Research, **13** (2018), 6, 3579-3589
- [8] Held R, Gupta A, Curless B, et al. *3D puppetry: a Kinect-based interface for 3D animation* //Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology. New York: Association for Computing Machinery, 2012, pp.423-434
- [9] Huang X, Sun S, Zhang K, et al. *A method of shadow puppet figure modeling and animation*. Frontiers of Information Technology & Electronic Engineering, **16** (2015), 5, 367-379
- [10] Liang Xiaojuan. *Research on Cartoon and Caricature Generation based on Style Transfer*. Chengdu: University of Electronic Science and Technology of China. The Library of the UESTC, 2019(in Chinese)
- [11] Dong Wei. *Generative adversarial network and its application in neural style transfer*. Ningbo: Ningbo University. The Library & Information Center of NBU, 2018(in Chinese)
- [12] Chen D, Yuan L, Liao J, et al. *Stereoscopic neural style transfer* //Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 6654-6663
- [13] Ulyanov D, Lebedev V, Vedaldi A, et al. *Texture networks: feed-forward synthesis of textures and stylized images* //Proceedings of the 33rd International Conference on Machine Learning. New York: JMLR.org, 2016: 1349-1357

- [14] Goodfellow I, Pouget-abadie J, Mirza M, et al. *Generative adversarial nets*. [2014-06-10]. <https://arxiv.org/abs/1406.2661>
- [15] Huang X, Liu M Y, Belongie S, et al. *Multimodal unsupervised image-to-image translation* //Proceedings of the European Conference on Computer Vision (ECCV). Switzerland: Springer Nature Switzerland AG, 2018: 179-196
- [16] Zhu J Y, Park T, Isola P, et al. *Unpaired image-to-image translation using cycle-consistent adversarial networks* //Proceedings of IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2242-2251
- [17] Kim T, Cha M, Kim H, et al. *Learning to discover cross-domain relations with generative adversarial networks* //Proceedings of the 34th International Conference on Machine Learning. Sydney: JMLR.org, 2017: 1857-1865
- [18] Yi Z, Zhang H, Tan P, et al. *DualGAN: unsupervised dual learning for image-to-image translation* //Proceedings of IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 2849-2857
- [19] Yi R, Liu Y J, Lai Y K, et al. *APDrawingGAN: generating artistic portrait drawings from face photos with hierarchical GANs* //Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Los Alamitos: IEEE Computer Society Press, 2019: 10735-10744