

Article

Interaction Between Public Art Sculpture and Urban Environment and Aesthetic Enhancement Effect Based on Deep Learning Model Evaluation

Shenghua Duan^{1,*}, Xi Zhao¹, and Chuxu Hu²

¹ School of Art, Zhejiang Shuren University, Hangzhou 310000, Zhejiang, China

² Division of Design, Dongseo University, 47011 Busan, South Korea

* **Correspondence:** ceecee123456@163.com

Abstract: The created public art sculpture is a material form that expresses the public spirit of the city. This paper proposes a deep model capable of enhancing the aesthetic quality of public art sculptures. The model uses the inverse mapping network of the augmented network to weaken the restriction of paired data sets required for training, and at the same time designs an effective loss function, that is, constructs the color and texture losses that are actively learned in training through generative adversarial rules, and enhances generative sculpture. The total variational loss of smoothness that improves the aesthetic quality of the sculpture to some extent. On this basis, this paper improves the design idea of content consistency loss. Experiments on the interaction between public art sculptures and the urban environment and the enhancement of aesthetics.

Keywords: Deep learning, Public art sculpture, Urban environment, Aesthetic enhancement

1. Introduction

The number of sculptures in cities has continued to increase, creating a good artistic atmosphere for the living environment of the masses [1]. The rise of modern cities has promoted the transformation of the functions of urban sculpture, that is, more traditional religious propaganda and landmark buildings have been transformed into environmental beautification functions. Buildings, streets, trees, and decorative objects are all constituent elements of different urban space environments [2]. The characteristics of sculpture in cities are not only the shape and color of the sculpture itself, but also whether it can become the focus of the environmental space [3, 4]. Therefore, the design concept and placement of urban sculptures should go hand in hand with the environment, considering both the integration of the sculptures and the environment, and the possibility of their separation, so as to show the “soul” of the city with its own space. The design of the shape, volume and color of the urban sculpture should be subject to the environmental atmosphere. According to different types of cities, sculpture should have its functional characteristics. Urban sculpture as public art must be people-oriented. When sculptors create urban sculptures, they must first understand the consciousness, spiritual outlook, yearning, pursuit or state of mind of the person where the work is located, and then it is possible to create a sculpture that has a distinct personality and can reflect the spiritual outlook of a person [5].

The use of cutting-edge technologies and raw materials in the ecological design of public sculptures is also an inevitable trend in the development of art [5, 6]. Aesthetic public art sculptures can

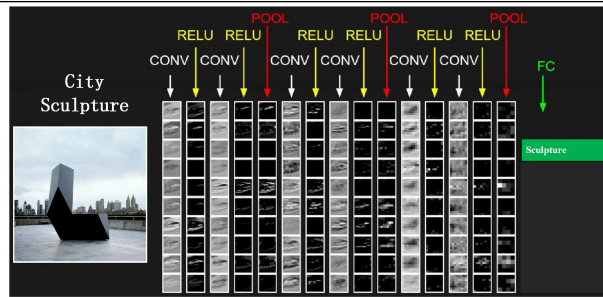


Figure 1. Convolutional Neural Network Structure and Working Diagram

not only attract users' attention, but also bring freshness and surprise to users, endow products with emotional meaning, and shorten the distance between users and products. However, at present, the aesthetics of such public art sculptures are mostly subjectively evaluated by designers and their supervisors [6,7]. Even if general design aesthetic principles (balance, contrast, symmetry, etc.), the answer is not obvious. With the development of computer vision technology, computers can gradually understand and calculate beauty, and then can assist designers in completing aesthetic and business decisions [8,9]. Aesthetic public art sculptures can not only attract users' attention, but also bring freshness and surprise to users, endow products with emotional meaning, and shorten the distance between users and products. However, at present, the aesthetics of such public art sculptures are mostly subjectively evaluated by designers and their supervisors. Even if general design aesthetic principles (balance, contrast, symmetry, etc.), the answer is not obvious. They can gradually understand and calculate beauty, and then can assist designers in completing aesthetic and business decisions.

2. Related Work

2.1. Deep Learning

Since the 1980s, artificial neural networks have emerged and attracted attention in the intersection of science and biology. It belongs to computer science that imitates the activity of biological neurons [10]. The activity of the neural network is abstracted and some kind of connection is established [11]. The human neural network includes many neurons, and the information spreads between the neurons through the synaptic structure.

CNN include 2D and 3D neural networks. Each layer of its convolution is composed of multiple filters. The filters have their own learnable parameters. These parameters are continuously updated during the convolution process to better fit the desired neural network. Output, a schematic diagram of the structure and working is given in the cs231n course, as shown below, as you can see, the convolutional neural network in Figure 1. These structures jointly convert the input of a picture into a feature vector, and finally map features and labels through the FC fully connected layer.

There are multiple convolutional layers in the convolutional neural network. The convolutional feature maps of different levels can be obtained after convolution operations are performed on the input sculpture through the convolutional layers.

The model design strictly uses 3×3 convolutional layers and a total of five 2×2 pooling layers connected after partial. The model structures of VGG16 and VGG19 are shown in Figure 2 and Figure 3. Using a 3×3 convolution kernel structure can not only reduce the number of parameters, but also stack more convolutional layers, deepen the depth of the network, and make training and testing more effective. For example, the size of the receptive field corresponding to the pixel obtained by the convolution operation with two layers of 3×3 convolution kernels is the same as that obtained by the convolution operation with a 5×5 convolution kernel, and the parameters of the former Fewer quantities. Similarly, the effect of using four layers is equivalent to a 9×9 convolutional layer. It can be seen that using a smaller convolution kernel has more nonlinear transformations than using a larger

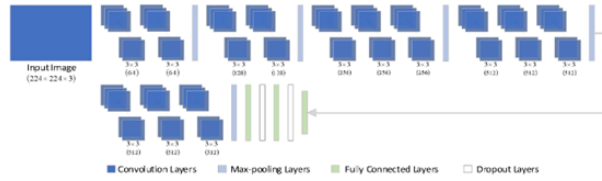


Figure 2. VGG16

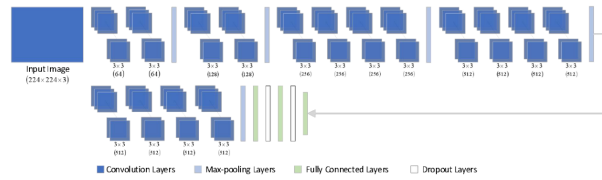


Figure 3. VGG19

convolution kernel, which makes the convolutional neural network. The network has stronger feature learning ability.

2.2. Aesthetic Evaluation of Deep Learning Models

This research created the "Domestic Mobile Sculpture Sculpture Aesthetics Dataset" to try to acquire oriental aesthetic thinking and assist domestic designers in their design practice. The development of diverse datasets is a trend that is critical to advancing this field of research(see Figure 4and 5.

3. Algorithm Design

3.1. System Structure

This chapter implements a weakly supervised end-to-end sculptural aesthetic enhancement model that can address the limitations of traditional image enhancement problems requiring supervised learning on paired datasets [12]. That is, this work optimizes a deep learning architecture that, through training, can map a given domain of low-quality sculptures to the domain of high-quality sculptures. It is worth mentioning that these two sculpture domains are mutually exclusive. independent, and there is no corresponding relationship [13]. In order to achieve the goal, this chapter studies the use of Generative Adversarial Network (GAN) [14, 15] and Convolutional Neural Network (CNN) to map

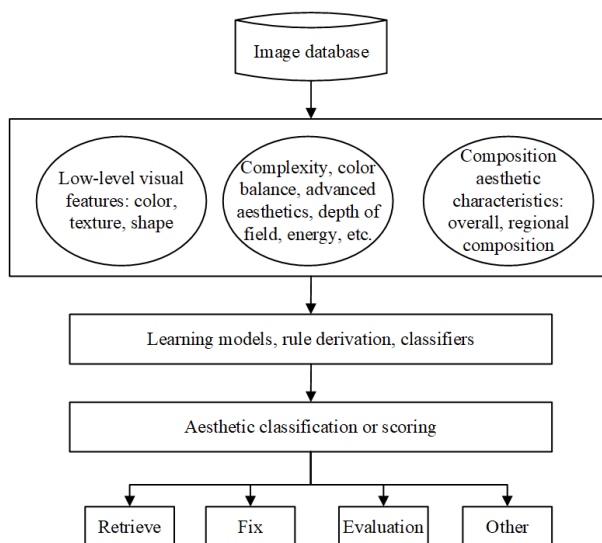


Figure 4. The First Aesthetic Quality Assessment Methods

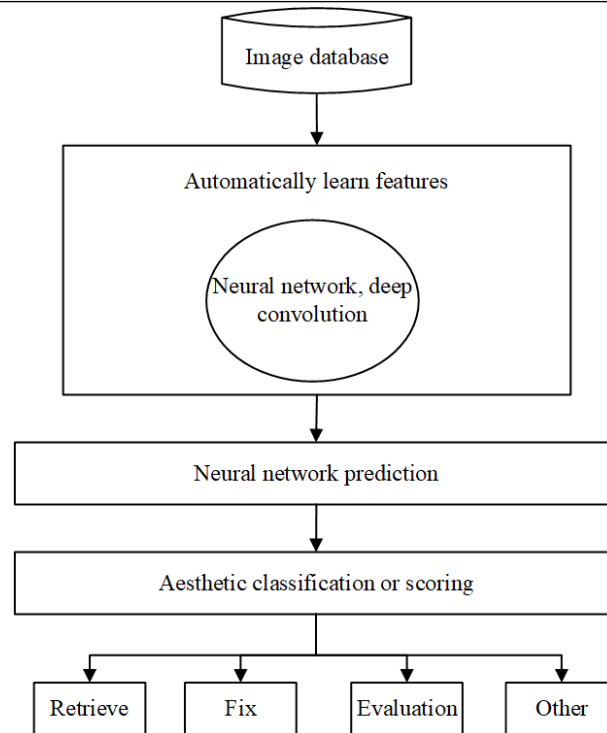


Figure 5. The Second Aesthetic Quality Assessment Methods

the enhanced sculpture original sculpture space to reduce the need for paired reference sculptures, texture, content semantic loss function to enhance the aesthetics of sculpture.

The entire model system contains a generative image map $G : X \rightarrow Y$ and an inverse generative map $V : Y \rightarrow X$. To ensure that the content of the final image $G(x)$ and the original image x after the generated mapping will not change, this chapter uses the U-Net reconstructed images to construct a content-aware loss. Using the content-aware loss function can effectively avoid the restriction of pairing input data before and after training. The model then complements the construction of the loss function with two discriminators D_{co} discriminator D_{te} , and a total variation (TV) loss. where the purpose of D_{co} is to distinguish the high-quality image y and the reconstructed enhanced image $\bar{y} = G(x)$ in terms of the color of the image. The discriminator D_{te} distinguishes the difference between the two from the perspective of image texture.

Therefore, the main experimental tasks of this work are: to ensure that the generator G can not change the content semantics of the original sculpture x , and maintain the content consistency of the enhanced sculpture; to ensure that the generated sculpture can better map the color and texture information of high-quality sculptures, improve the aesthetic quality of the sculpture; This chapter implements a weakly supervised end-to-end sculptural aesthetic enhancement model that can address the limitations of traditional image enhancement problems requiring supervised learning on paired datasets. That is, this work optimizes a deep learning architecture that, through training, can map a given domain of low-quality sculptures to the domain of high-quality sculptures. It is worth mentioning that these two sculpture domains are mutually exclusive. independent, and there is no corresponding relationship. In order to achieve the goal, this chapter studies the use of GAN and CNN to map the enhanced sculpture original sculpture space to reduce the need for paired reference sculptures, texture, content semantic loss function to enhance the aesthetics of sculpture use the TV loss function to ensure that the generated sculpture is smoother. Next, we introduce the network structure and different types of loss functions used in the experiments in detail.

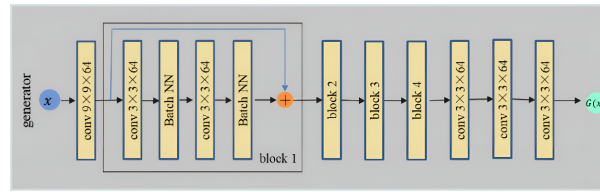


Figure 6. The First Network Structure of Generator and Discriminator

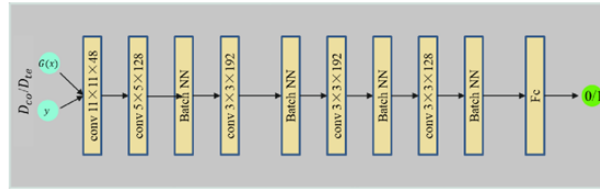


Figure 7. The Second Network Structure of Generator and Discriminator

3.2. Model Network Structure

The generation network G and the inverse generation network V use the same and will not be repeated here. Specifically, the network of G and V is constructed by first passing through a 9×9 convolutional layer, and then connecting 4 residual block structures, each residual block consists of 2 alternating with batch normalization (BN) consists of 3×3 convolutional layers. After the residual block, connect two 3×3 convolutional layers and one 9×9 convolutional layer, and finally get the output result. The number of channels in all layers of the network is 64. Except for the last output layer, the probability of whether the input sculpture is from the augmented sculpture or the original sculpture (see Figure 6 and 7,).

The curve is shown in the Figure 8:

3.3. Loss Function Design

3.3.1. Color Loss

In order to make the enhanced sculpture better fit the color characteristics of high-quality sculptures, this chapter directly adopts a discriminator network D_{oc} to narrow the distance between the two through dynamic learning. avoiding being influenced by texture and content information, in this chapter, we will enhance sculptures $G(x)$ and high-quality sculptures y . After Gaussian blurring, it is sent to the discriminator as the input part for training and learning. Sculpture blurring refers to the reconfiguration of each pixel value in the sculpture, a common method is to set the size of each pixel value to the average value of the surrounding pixel values. However, the simple and average method is too crude, does not take into account the continuity of the sculpture, and ignores the close

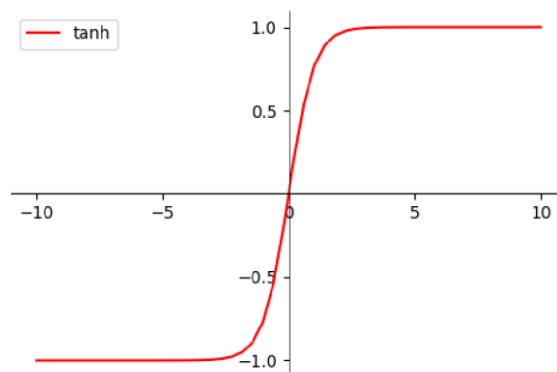


Figure 8. Tanh Activation Function

relationship between the close points of the sculpture and the alienation of the relationship between the distant points. Therefore, a reasonable weight should be used, the discriminator D_{oc} is to identify whether the color distribution belongs to an augmented sculpture or a high-quality sculpture, while the goal of the generator is to fool the discriminator to make the augmented sculpture fit the color distribution of the high-quality sculpture as closely as possible. This is consistent with the objective function of the original GAN network and is a minimization problem. Therefore, the color loss L_{color} can be defined as a standard GAN network loss function in the form:

$$\min_G \max_{D_{co}} V(D_{co}, G) = E_{y \sim p\gamma(y)} [\log D_{co}(y_g)] + E_{x \sim p\gamma(x)} [\log(1 - D_{co}(G(x)_g))]. \quad (1)$$

Among them, $G(x)_g$ are the real high-quality sculpture and the augmented sculpture after Gaussian blurring respectively. It can be seen that the purpose of the color loss is to make the enhanced sculpture have a similar color distribution to the high quality sculpture.

3.3.2. Texture Loss

Similar to the design idea of color loss, the texture quality of the sculpture before and after enhancement is directly learned and assessed by a discriminator D_{te} . Texture features can reflect the homogeneous phenomenon in sculpture. While describing the global characteristics, it can also reflect the surface characteristics of objects corresponding to sculpture or sculpture area. Unlike color features, textures are not based on pixel features, but represent the grayscale distribution of pixels and their surrounding adjacent areas, that is, local texture information. Global texture information refers to the repetition of local texture information to different degrees. Therefore, this regional locality property can get rid of the influence of local bias and has excellent experimental results. The rotational invariance and anti-noise ability of texture feature also make its application field more extensive. In the experiment, in order to remove the influence of other sculptural features such as color, the enhanced sculpture $G(x)$ and high-quality sculpture y are first taken as grayscale images, and then used as the input data of D_{te} for training.

$$\min_G \max_{D_{te}} V(D_{te}, G) = E_{y \sim p\gamma(y)} [\log D_{te}(y_t)] + E_{x \sim p\gamma(x)} [\log(1 - D_{te}(G(x)_t))]. \quad (2)$$

Among them, $G(x)_g$ are the grayscale images of the real high-quality sculpture and the enhanced sculpture, respectively. It can be seen that the texture loss enables the enhanced sculpture to have similar texture information to the high-quality sculpture

3.3.3. Content-aware Loss

To make the enhanced sculpture better fit the color characteristics of high-quality sculptures, this chapter directly adopts a discriminator network D_{oc} to narrow the distance between the two through dynamic learning. avoiding being influenced by texture and content information, in this chapter, we will enhance sculptures $G(x)$ and high-quality sculptures y . After Gaussian blurring, it is sent to the discriminator as the input part for training and learning. Sculpture blurring refers to the reconfiguration of each pixel value in the sculpture, a common method is to set the size of each pixel value to the average value of the surrounding pixel values.

The content loss $L_{content}$ can be understood as minimizing the feature difference of the sculpture before and after enhancement. L2 regularization is used to calculate this difference, and the expression is defined as follows:

$$L_{content} = \frac{1}{N_i H_i W_i} \|\psi_i(x) - \psi_i(\tilde{x})\|. \quad (3)$$

Using content loss can force the augmented sculpture to maintain the same content information as the original image.



Figure 9. The Simulation Diagram of the Experimental Result of the Feature Point Matching of the Ideal Matching Surface

4. Simulation Results and Analysis

Extensive experiments are conducted on publicly available large-scale standard sculpture datasets, and a series of common metrics for GAN and aesthetic evaluation are used to compare the proposed model with Differences between original GANs. Since there is a certain subjectivity in aesthetics, based on the intuitive visual effect comparison, we also constructed an aesthetic scoring program for generating sculptures, which gives qualitative visual results in the form of user scoring.

4.1. Training Strategy and Experimental Details

On the basis of designing an effective loss function, an excellent training strategy should be adopted in order to train an aesthetic GAN network structure that meets the requirements. Especially for deep learning models, appropriate training strategies and training methods play an important role in the process of model learning and fitting. Based on this, this subsection discusses the training strategy and parameter setting of the aesthetic GAN network. The experimental machine in this work is configured with a GeForce GTX Titan X GPU, the operating system version is Ubuntu 14.04. The number of iterations of the experiment is 20,000, and the training is performed in a mini-batch mode, where each batch size is 32. During training, reducing the learning rate reasonably after multiple iterations helps the model to converge. Therefore, the learning rate is reduced by half every 2000 iterations.

4.2. Experimental Results and Analysis

Carry out specific experiments on sculpture data containing the same information, sculpture data containing different information, and sculpture data containing the same information with different resolutions, and display and analyze the experimental results. The feature point matching experiment is performed on the sculpture data (ideal matching surface) containing the same information (only the matching results are displayed, and the feature point extraction is not displayed). The sculpture data is shown in Figure 9. The first row of sculpture data comes from the CMU-OxfordSculpture building Sculpture Sculpture Dataset (published dataset, including 143,000 statues and statues, 2197 sculptures from 242 artists, each sculpture in the dataset collects a large amount of sculpture data from different angles, the resolution is 640×640 , which conforms to the original sculpture data collection method based on the 3D reconstruction technology of multi-view sculpture), the second line of sculpture data is collected from the Honor30 mobile phone ordinary camera (1:1 mode), the resolution is 1920×1920 , and the ratio value of the experiment is the same Set to 0.6.

The number of sculptures in cities has continued to increase, creating a good artistic atmosphere for the living environment of the masses. The rise of modern cities has promoted the transformation of the functions of urban sculpture, that is, more traditional religious propaganda and landmark buildings have been transformed into environmental beautification functions. Buildings, streets, trees, and decorative objects are all constituent elements of different urban space environments. The characteristics of sculpture in cities are not only the shape and color of the sculpture itself, but also whether it can become the focus of the environmental space [16]. Therefore, the design concept and placement of

	CUHKPQ	AADB	AVA	Cityscapes
Original	5.02 ±0.23	5.09 ±0.43	5.42 ±0.43	4.94 ±0.45
Ref. [20]	6.17±0.28	6.23±0.40	6.53±0.43	6.05±0.44
Ours	6.29±0.27	6.29±0.34	6.41±0.46	6.13±0.42

Table 1. NINA Values Corresponding to Different Methods in Different Datasets

urban sculptures should go hand in hand with the environment, considering both the integration of the sculptures and the environment, and the possibility of their separation, so as to show the “soul” of the city with its own space. The design of the shape, volume and color of the urban sculpture should be subject to the environmental atmosphere. According to different types of cities, sculpture should have its functional characteristics. Urban sculpture as public art must be people-oriented. When sculptors create urban sculptures, they must first understand the consciousness, spiritual outlook, yearning, pursuit or state of mind of the person where the work is located, and then it is possible to create a sculpture that has a distinct personality and can reflect the spiritual outlook of a person [17].

Considering the complexity of sculpture aesthetics, in order to effectively evaluate the enhancement effect of sculpture aesthetics, this section conducts comparative experiments from three perspectives: qualitative, quantitative scores, and user learning.

1. Qualitative analysis: Figure 7 shows some of the results on the dataset, where the first, third, and fifth columns are the original low-quality sculptures to be enhanced, and the second, fourth, and sixth columns are the enhancement results of this method. A top-down observation is not difficult to find that the use of augmented models improves the visual effect of the original sculpture to a certain extent and enhances the aesthetic appeal. For example, in Figure 8, in the second, fourth, and sixth columns, the sculptures are brighter as a whole, the objects are clearer, the colors of the sculptures are more vivid, the textures are richer, and the contrast is more vivid. At the same time, the method of this work is more reasonable in color distribution than the result of literature [18], and the sculpture is more clearly visible. The experimental results on other datasets in Figure 9 further verify the aesthetic enhancement performance of our method.
2. Quantitative analysis: In the experiment, the NIMA and ACQUINE aesthetic scores consistent with Chapter 3 were used as the beauty scores of the enhanced sculptures are significantly improved. Except for a slight deficiency on the AVA dataset, the method in this work has similar or even better performance. The quantitative score results in Table 1 and 2 also prove the correctness of the observations, and verify the model’s ability to optimize the aesthetics of sculpture from another perspective.
3. User learning: quality of the sculpture to satisfy the user’s vision Therefore, the user learning method is adopted as the subjective evaluation index. This chapter designed an aesthetic evaluation procedure and distributed it to 30 experimenters. 60 sculptures were randomly selected in each dataset, and for each sculpture, the experimenters were provided with before and after augmentation results and asked to choose sculptures with higher aesthetic quality. The experiment is not limited by time and space, and the experimenter can submit the final result at any time. From Figure 10, it can be concluded that the aesthetic quality of the augmented sculptures in this work is significantly better than the unenhanced original sculptures, especially in the Cityscapes dataset, 91.67% of the experimenters chose the augmented sculptures. According to Figure 10, it can be concluded that the enhanced sculpture is more in line with the user’s aesthetic preference and has more aesthetic appeal.

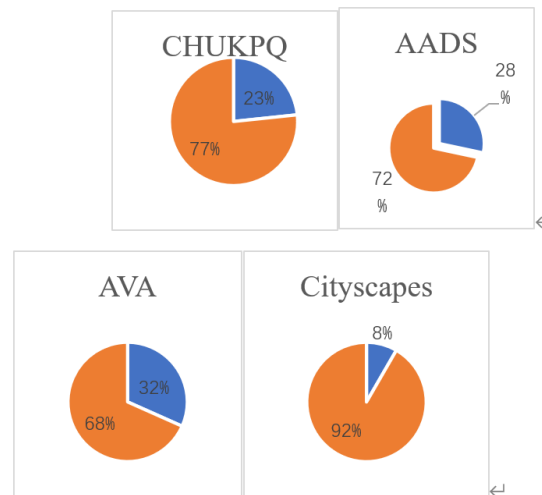


Figure 10. User Learning: Pairwise Comparison of the Original Image and the Enhanced Sculpture Yields Good Results

	CUHKPQ	AADB	AVA	Cityscapes
Original	7.76±0.25	7.63±0.44	7.95±0.36	7.12±0.33
Ref. [20]	8.64±0.27	8.79±0.42	9.13±0.39	8.57±0.29
Ours	8.69±0.25	8.87±0.39	9.04±0.41	8.60±0.31

Table 2. ACQUINE Values Corresponding to Different Methods in Different Datasets

5. Conclusion

This work combines the excellent effect of generative adversarial networks in learning the distribution of sculptures, and experimentally completes an end-to-end sculpture aesthetic enhancement model. That is, it is able to learn from two completely independent datasets representing low-quality and high-quality domains, respectively, enabling the transformation of low-quality sculptures into high-quality sculptures with higher aesthetic properties. This work utilizes several different types of loss functions to comprehensively constrain the generative capabilities of GANs in terms of color, content, texture, and smoothness. The model is validated on several publicly available datasets.

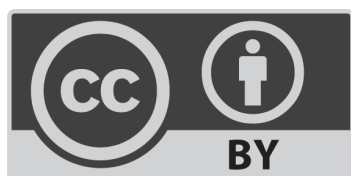
Funding Statement

This study received the following funding support: Suzhou University, Suzhou City, Anhui Province Project Fund Number: 2023BSK071.

References

1. Niu, Z., Zhong, G. and Yu, H., 2021. A review on the attention mechanism of deep learning. *Neurocomputing*, 452, pp.48-62.
2. Ranganathan, D.G., 2021. A study to find facts behind preprocessing on deep learning algorithms. *Journal of Innovative Image Processing*, 3(1), pp.66-74.
3. Pandian, A.P., 2021. Performance evaluation and comparison using deep learning techniques in sentiment analysis. *Journal of Soft Computing Paradigm (JSCP)*, 3(02), pp.123-134.
4. Zheng, J., 2021. Structuring artistic creativity for the production of a ‘creative city’: Urban sculpture planning in Shanghai. *International Journal of Urban and Regional Research*, 45(5), pp.795-813.

5. Liu, Y., 2021. Design and Research of Teaching Support System for Urban Landscape Sculpture Course. In *Application of Big Data, Blockchain, and Internet of Things for Education Informatization: First EAI International Conference, BigIoT-EDU 2021, Virtual Event, August 1–3, 2021, Proceedings, Part I 1* (pp. 96-104). Springer International Publishing.
6. Ma, J., Li, Y., Shen, D. and Zhang, R., 2021, February. Integration of regional culture and place environment—Taking the landscape sculpture design scheme of Yuanbaozi Wetland Park as an example. In *IOP Conference Series: Earth and Environmental Science* (Vol. 676, No. 1, p. 012022). IOP Publishing.
7. Datta, R. and Wang, J.Z., 2010, March. ACQUINE: aesthetic quality inference engine-real-time automatic rating of photo aesthetics. In *Proceedings of the International Conference on Multimedia Information Retrieval* (pp. 421-424).
8. Liao, Q., English Teaching Project Quality Evaluation Based on Deep Decision-Making and Rule Association Analysis. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 118, pp.119-127.
9. Xu, J., Optimizing English Education in the Information Era: A Multimodal Approach Based on BOPPPS Teaching Model. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 118, pp.33-48.
10. Liu, L., 2024. Application of Relying on 3D Modeling Technology in the Design of Urban Sculpture. *J. Electrical Systems*, 20(1), pp.176-191.
11. Gao, Y. and Shen, T., 2023. Construction of Evaluation Index of Nature-Relatedness Effect of Public Art Landscape Under the Background of Sustainable Development of Environment. *International Journal of Professional Business Review: Int. J. Prof. Bus. Rev.*, 8(4), p.50.
12. Luo, N., Ibrahim, R. and Abidin, S.Z., 2022. Transformation of children’s paintings into public art to improve public spaces and enhance people’s happiness. *International Journal of Environmental Research and Public Health*, 19(24), p.16780.
13. Cheung, M., Smith, N. and Craven, O., 2021. The impacts of public art on cities, places and people’s lives. *The Journal of arTs ManageMenT, law, and socieTy*, 52(1), pp.37-50.
14. Tang, J., Yan, L. and Xu, J., 2021. Tourists’ experience of iconic public art in Macau. *Journal of Tourism and Cultural Change*, 19(1), pp.79-96.
15. Wu, Y., Liu, Q., Hang, T., Yang, Y., Wang, Y. and Cao, L., 2024. Integrating restorative perception into urban street planning: A framework using street view images, deep learning, and space syntax. *Cities*, 147, p.104791.
16. Yang, X., 2024. Construction of teaching system of public art major using CMOS image sensor technology. *Scientific Reports*, 14(1), p.5497.
17. Liu, Z., 2022, March. Urban Public Art Communication Algorithm based on Smart Virtual Cloud Platform and Artificial Intelligence Algorithm. In *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1682-1685). IEEE.
18. Yang, Z., Zhang, L. and Wu, Z., 2022. Research on performance evaluation of urban furniture function design based on internet of things digitization. *IEEE Access*, 10, pp.72895-72906.



©2024 the Author(s), licensee Combinatorial Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)