Analysis of the Emotional Conveyance Ability of AI-Generated Artworks

Zhuo Yali, Hou Guanhua, Wang Huiwen

ORCID link of the first author: https://orcid.org/0009-0008-4315-943X

Abstract: This paper explores the ability of art generated by artificial intelligence (AI) to convey emotions and its potential impact on modern art theory and practice. Employing a mixed research method, this study first evaluates the differences in audience emotional responses between AI-generated art and traditional human art through quantitative analysis. Subsequently, qualitative interviews with art critics, artists, and ordinary viewers were conducted to delve into the uniqueness and limitations of AI art in emotional expression and perception. The study finds that while AI-generated artworks excel in visual complexity and novelty, they still have limitations in conveying deep emotions and personal experiences, affecting the depth of emotional resonance. Additionally, the study explores the application of AI technology in art creation and the potential changes it may bring to artistic creation methods and art market structures. This research not only provides a new perspective on the application of AI in the field of art but also lays an important foundation for the future development of art theory and practice. With the further development and popularization of AI technology, its role in art creation and aesthetic evaluation is expected to expand further, requiring the joint scrutiny and guidance of the art community, technology experts, and policymakers.

Key words: artificial intelligence, emotional communication, art

1. Introduction

Artificial intelligence (AI), as the latest generation of digital technology, has already demonstrated powerful technological advantages and extensive application prospects in various fields such as mechanical manufacturing, medical surgery, intelligent transportation, aerospace, and smart city. Since the portrait “Edmond de Belamy” was generated in 2018, AI-generated art has got extensive attention. After long-term development, the latest use of AI is to generate images based on natural language descriptions. This has to a great extent significantly improved the productivity of art creators.

In this regard, for artistic creation in the past, artists need adequate practice and accumulation of experience in order to convey what they are envisioning more accurately, whether they use traditional tools such as paper and brush or modern electronic tools such as digital panels. However, with the rise of AI-generated technology, people can now use AI to generate beautiful images according to their preferences on relevant platforms, whether they have received artistic training or not.

Imitation of brush strokes in AI-generated artworks

According to the current situation in the art auction market, there is some uncertainty regarding the acceptance of AI-generated artworks. In recent years, AI-generated artworks have received great attention in the art auctions held by famous museums. According to the competition information from the website of Colorado State Fair, the artwork “Théatre D’opera Spatial” generated by Midjourney won the first place among digital artworks.

AI is perceived as lacking emotion, and artistic creation involves emotional expression, so people tend to believe that these systems are unable to create artworks imbued with emotion like humans can. However, the reality is that artworks generated by the emotionless AI are not necessarily outright rejected by people (although we all know that there is not a complete positive correlation between the artistic value and the expression of emotion); instead, these artworks can even resonate with people to some extent. This prompts us to contemplate why artworks generated by emotionless AI can resonate with people in conveying emotions. In other words, to what extent do AI-generated artworks have the ability to convey emotions, and to what extent will our understanding of AI-generated artworks expand the boundaries of modern art theory? By exploring these questions, our understanding of modern art will also be further enhanced.

A conceptual scene generated by AI

1.1 Research objectives and problem definition

With the new phenomena and challenges occurring in artistic creation involving AI in the present society as a research background, this paper explores the ability of AI-generated artworks to convey emotions and analyzes the potential impact they may have on modern art theory and practice. The notion of "emotional conveyance" mentioned in this paper is not discussed on the premise that AI possesses emotions. In fact, assessing whether AI possesses emotions is quite complex. While this paper may touch upon aspects of AI emotions, it does not intend to delve into abstract concepts such as "consciousness" and "emotion" in AI. The emotional conveyance ability mentioned in this paper refers to the ability of AI-generated artworks to evoke emotional responses from audiences in experiments. In other words, it examines to what extent AI-generated artworks can provoke emotional reactions from audiences compared to traditional paintings.

2 Text

2.1 Literature review and theoretical basis

The first issue is about the legal ownership of AI creations. Samuel Asimo discussed in detail the copyright attribution of AI-generated artworks in his paper titled *The Creative Rights of the Machine: Challenges in AI and Copyright Law* published in 2019. Asimo pointed out that the current framework of copyright laws is inadequate for AI-generated works, and there are legal loopholes. He suggested revising existing laws to include AI as potential copyright holders. In contrast, Jeanette Wing and James Grimmelmann examined the issue concerning the authorship of AI-generated artworks from legal and philosophical perspectives in their paper titled *Who is the Author? AI and the Art of Creation* published in 2020. Their argument is that although AI is capable of performing creative tasks, it lacks the intent and autonomy for creation. Therefore, the authorship should be attributed to the developer or user of the AI.

In addition, in terms of the innovation and imitation between AI and human artists, Rebecca Fiebrink discussed the potentials and limits of AI in artistic creation in her paper titled *Artificial Imagination: The Limits and Potentials of AI in Creativity* published in 2018, with particular emphasis on AI's strong ability to mimic existing artistic styles and its lack of creativity. Fiebrink put forward that although AI can generate technically impressive works, its creativity still depends on human's input and guidance. Furthermore, Marcus Du Sautoy discussed whether machines can be considered creators in his paper titled *Can Machines Be Creative? AI and the Future of Art* published in 2021. Du Sautoy believed that although AI can generate stunning artworks, the so-called "creativity" of machines is actually backed up by the algorithms designed by humans. Therefore, true innovation still requires the intuition and emotions of human artists. Published in 2021 by Hannah Davis, the paper *Emotion and Expression in AI Art: The Limits of Artificial Creativity* examines AI-generated art from the perspective of emotional expression. Davis emphasized that although AI can generate visually striking artworks, these works often lack genuine emotional depth and personal expression. This is because AI cannot understand emotions and personal experiences in the same way humans can. Davis appealed to the public and the art world in his research to reevaluate the application of AI in artistic creation and its impact on the emotional quality of artworks.

In terms of the ethical and moral issues involved in artistic creation by AI, Margaret A. Boden discussed the ethical issues caused by AI in artistic creation in detail in her paper *Ethics of Artificial Creativity: Issues and Challenges* published in 2020. These issues include the moral rights of machine-generated artworks and the social inequalities possibly arising from the use of AI. Boden called for further research in her paper to ensure that the application of AI in the field of art is responsible and ethical. Arthur I. Miller discussed the mechanics and creativity of AI in artistic creation in his paper titled *Artificial Artists: The Mechanics of AI Creativity*. Miller pointed out that although AI can mimic and reproduce the styles and techniques of human artists, AI-generated works often lack profound cultural and emotional connotations. Therefore, controversy is often sparked in the evaluation of artistic values and creativity of AI-generated artworks. In his research, Miller questioned the creative ability of AI, namely, whether AI is merely a tool that can execute the preset rules and algorithms, or if it can truly "create" new forms of art. Robert Pepperell provided an in-depth analysis of how AI can alter the labor dynamics in artistic creation in his paper titled *The Role of Artificial Intelligence in the Future of Artistic Labor* published in 2021. Pepperell discussed the potential of AI to reduce the demand for human labor in artistic creation, particularly in technical and repetitive tasks. However, he also pointed out that such changes may redefine the role of artists and the nature of the artistic creation process, which caused widespread discussions about the applicability and impact of AI in the art world. In the paper titled *AI and the End of Artistic Labor* published in 2022, the author Elaine Herzberg proposed that the widespread application of AI could lead to the automation of artistic creation, which would not only change the way for creating artworks but also potentially affect the structure of the art market and people's assessment of the value of artworks. Herzberg provided a detailed analysis of how AI may cause artworks to be overly commercialized and homogeneous while reducing the labor cost in artistic creation, which raised concerns among artist and art enthusiasts.

The acceptance and evaluation criteria of AI art are also hot topics in academic discussions. In the paper *The Perception of AI Art in Contemporary Society* jointly published by Christoph Bartneck and Dominic DiFranzo in 2023, the authors discussed how the public evaluates AI-generated artworks, and made a comparison with the evaluation of traditional artworks. They found that while some audiences may appreciate the technical achievements of AI art, the majority feel that these works lack the emotion and depth of human artists, so they still have reservations towards AI-generated artworks. This study indicates that there are significant differences in the evaluation and acceptance of AI-generated artworks compared to those created by human artists.

It is worth noting that some scholars focus on how AI interprets artworks. In her paper titled *AI as a Tool for Understanding Art: Possibilities and Limitations* published in 2019, Linda D. Henderson discussed how AI is utilized in art history and criticism, particularly its application in analyzing the artworks for which big data are available. She believed that although AI can uncover the patterns and connections hidden within artworks, at the same time, she warned that it may overlook the cultural context and emotional connotations of artworks. Henderson's paper not only helps us better understand the role of AI in art analysis but also critiques the limitations of AI in interpreting artworks.

In other words, the application of AI in the realm of art has brought forth many complex academic and practical issues, including the automation of artistic creation, the authenticity of emotional expression, and the evaluation criteria for artworks. These issues involve multiple aspects such as technological, ethical, and cultural areas, thus requiring the involvement and in-depth discussion of experts from various fields including artists, technology developers, legal experts, and cultural theorists. Based on in-depth analysis of these literatures, we can have a comprehensive understanding of how AI can shape the forms of contemporary art, and predict its potential effects on the future art world.

2.2 History of application and current status of artificial intelligence in artistic creation

We reviewed the significant technological breakthroughs of AI in artistic creation in recent years as well as its history of application and current status. AI painting, as a hot research direction in the field of computer vision, is expanding its application scope in areas such as artistic creation, film and media, industrial design, and art education through techniques such as natural language processing, large-scale image-text pre-training models, and emerging diffusion models. The development of AI painting mainly revolves around two types of tasks: One is “Image-to-Image”, and the other is “Text-to-Image”. For the “Image-to-Image” approach, different development paths and trends have emerged based on two models, i.e., autoencoder-based models and generative adversarial network (GAN)-based models. For Text-to-Image, three models including the diffusion model-based ones with different architectures have emerged.

The history of AI painting can be traced back to the 1970s when traditional methods were predominantly used, namely, employing computer software or computer-controlled programs in conjunction with mechanical devices for painting. For example, in 1972, Professor Harold Cohen designed a computer drawing program called “Aaron”, which is one of the earliest and most sophisticated artwork generators. The program can draw on a fixed piece of paper by controlling its robotic arm. In 2006, Simon Colton developed “The Painting Fool”, a painting software capable of extracting color information from digital photographs and simulating painting materials found in the real world to create artworks. With the development of deep learning, hardware capabilities, and large-scale datasets, various generative models have emerged and been used in AI painting tasks. One of the most representative models is the Generative Adversarial Network (GAN) proposed by Goodfellow et al. in 2014, which combines a generator and a discriminator to generate images. A diffusion model based on Markov chain has also been gradually used in AI painting tasks. Such model can generate new images by gradually removing and restoring noise of the image. It has the advantages of stable training processes, high-quality generated images, and ease of control. Currently, researchers have proposed various AI painting models, which can be divided into two categories according to the types of input data: (1) Image-to-Image Model, which automatically generates new images by inputting existing images; and (2) Text-to-Image Model, which generates new images with features described in the input text. Currently, a new painting model which generates images based on both text and images has emerged.

Image-to-Image models are primarily composed of two types, namely, autoencoder-based models and generative adversarial network-based models. The latent space of the autoencoder-based Image-to-Image model is a space for representing compressed data. It can capture important features and models of data while removing redundancy and noise. In the Image-to-Image model, the latent space can reflect both the content and the style of the image. The autoencoder-based Image-to-Image model, with latent space representation at the core, measures the similarity between the generated images and input images through reconstruction error. Its advantage is that no data annotation is required, and only input of images is needed. It can also improve the reconstruction quality and robustness or diversity of latent representations. Its weakness is that the continuity and interpretability of latent representations cannot be guaranteed.

Image-to-image model based on generative adversarial network: Although the autoencoder is an important Image-to-Image model, it cannot effectively separate and control image content and style, and cannot handle tasks such as image style transfer. To address such issues, the generative adversarial network is introduced.

There are mainly two types of image style transfer algorithms: One is based on image iteration and the other is based on model iteration. The algorithm based on image iteration directly optimizes white noise images iteratively to achieve style transfer. The algorithm based on model iteration achieves style transfer by iteratively optimizing the neural network model. Compared to the method based on image iteration, the generative adversarial networks represented by GAN and their improved versions are more flexible and efficient.

Pix2pix model, which is proposed by Isola et al., is an Image-to-Image conversion model based on conditional Generative Adversarial Network (cGAN). It utilizes paired images for image conversion, where the generator employs a U-Net architecture to replace the traditional encoder-decoder architecture. The final images generated by this model closely resemble real images. However, this model is trained with a large amount of paired image data, and it is difficult to collect such paired image datasets.

CycleGAN, on the other hand, does not need paired data for model training. This model comprises two generators, G and F, and two discriminators, D and D. The introduction of model architecture and cycle consistency loss marks the first implementation of conversion between unpaired images. While widely applied in the field of style transfer, they also have drawbacks. Firstly, they tend more to alter the color style of input images rather than their geometric structure. Secondly, when input images contain images not present in the training process, mapping of the generator may lead to diverse changes. Thirdly, the model cannot generate diverse results and has weak data generalization capabilities. In response to the weak generalization capability of CycleGAN, Sanakoyeu et al. introduced a style-aware loss function, enabling the model to be trained on a broad category of artistic styles rather than specific ones. To address the issue of single results generated by CycleGAN, Zu et al. introduced the BicycleGAN model, enabling the generation of images with multiple styles. BigGAN, introduced by Brock et al., is a larger model composed of multiple CycleGANs. Compared to CycleGAN, it increases the model scale, improves the training efficiency, and avoids locally optimal solutions. In addition, it also enhances domain diversity, thus improving the model performance and robustness. The introduction of BigGAN has provided a method for large-scale training of models based on the original GAN architecture. It also helps to discover the instability inherent in large-scale GANs and proposes technical methods to mitigate such instability. Although it is characterized by highly vivid generated images, it also has drawbacks such as a large model size, numerous parameters, and high training costs. StyleGAN model, which is proposed by Karras et al., introduces style vectors and adaptive instance normalization to control the style and details of generated images. This model consists of a mapping network f and a generator network g. f can map the original latent vectors into a decoupled style space, so that style vectors from different dimensions correspond to image features at different levels, thus achieving finer-grained image editing. StyleGAN series have now been updated to StyleGAN3. It has addressed the issue of stickiness of details during image translation and rotation in previous versions, and significantly improved the quality of generated images. GANs have greatly promoted the development of AI painting in various aspects, but they still have some drawbacks. For instance, the stability and convergence of models are poor in the face of complex and diverse data, resulting in a weak control over the output result. As a result, random images can be easily generated. More importantly, the content generated based on the basic architecture of GANs closely resembles the existing content, which means that the output images are imitations of existing works rather than innovations. This does not align with people's expectations of AI painting.

Text-to-Image model: Such models can generate images according to natural language descriptions without being limited by existing image data, so they are highly valuable for application. The development course of Text-to-Image models can be divided into three stages according to their generation mechanism: GAN-based, VAE and pre-trained language model (PLM)-based, and diffusion model-based ones.

GAN-based Text-to-Image model: Such models utilize the generator and discriminator of GANs to learn the mapping relationship between text and images, thereby generating images related to the semantics of the text. Image generation based on conditional information: A piece of label information is added to cGAN on the basis of GAN, so that the generator can generate an image that matches the label information provided, while the discriminator can judge whether the image is real and consistent according to the label information.

The GAN-INT-CLS model, proposed by Reed et al. on the basis of cGAN, is the first GAN-based Text-to-Image model. The architecture includes two discriminators: GAN-CLS and GAN-INT. GAN-CLS not only needs to distinguish real images from fake images but also needs to determine if the images are generated according to the provided text. GAN-INT is used for data interpolation.

Although the model can only generate images with a resolution of 64×64 pixels, it provides a new method for Text-to-Image generation. Reed et al. introduced the attention mechanism and write-in network based on the GAN-INT-CLS model. The GAWWN model, based on Progressive Growing of GANs, supports the drawing of images in the desired region by inputting specific location descriptions based on conditional encoding, thus increasing the resolution of the images generated to 128×128 pixels. Han et al. proposed the StackGAN and StackGAN++ models, which optimize the generator, and introduce a stacking approach to stack multiple GANs in a serial manner. The first model in the stack generates a low-resolution image containing a rough shape and colors of the object, while subsequent GANs generate high-resolution images on these basis respectively. The SDGAN, proposed by Yin et al., designs a discriminator based on a twinning mechanism, which learns consistent high-level semantics based on contrastive loss, as well as a generator based on an attention mechanism, which integrates diverse low-level semantics based on conditional batch normalization of semantics.

Image generation based on the attention mechanism: XU et al. introduced the attention mechanism into GAN models and proposed the AttnGAN model, which avoids the loss of image details through the attention mechanism and multi-stage refinement, achieving high-quality, fine-grained image generation based on the textual descriptions. Qiao et al. proposed the MirrorGAN model based on AttnGAN. This model introduces a mirroring component and a global and local co-attention module, and gradually enhances the diversity and semantic consistency of generated images based on the local word attention and global sentence attention.

Text-to-Image model based on VAE and PLM: Such models learn the latent space between text and images mainly based on VAE, and can improve the text understanding capabilities as well as the accuracy and diversity of generated images in conjunction with PLM.

VAE and Transformer-based model: DALL-E is a Text-to-Image model based on VQ-VAE-2 and Transformer. It combines a high-capacity Transformer model as part of the encoder and decoder. It utilizes VQ-VAE-2 to compress images into a set of discrete codes, and then understand and generate these codes based on the Transformer model. Although it performs well in terms of semantic understanding and creative image generation, and is capable of fusion creation, scene understanding, and style transfer, it may not generate high-quality images in the field of zero-shot learning and professional niche field, and its resolution is limited.

VAE and contrastive learning-based model: Contrastive Language-Image Pre-Training (CLIP) model can effectively learn the visual concept from natural language supervision and predict the most relevant images based on text without directly optimizing the task. With contrastive learning, the model is pre-trained on a large number of text-image pairs so that it can distinguish between positive and negative samples. After CLP was introduced, people began to consider how to use its image-text matching validation mechanism to guide the image feature vector to match the specified conditional encoding vector of text, thus obtaining the images matching the textual descriptions.

Based on the idea of combining CLIP and other image generation models, Patashnik et al. proposed the hybrid model StyleCLIP, which combines CLIP with StyleGAN, and Rameen et al. proposed the VOGAN-CLIP model. DALL-E2 increases the image resolution to 1024×1024 pixels and ensures that the images better match the semantic information. The Midjourney painting platform combines VQGAN with CLIP to generate images based on the latent vectors of images. Its V5 version employs a Transformer-based pre-trained language model called OpenCLIP, which can handle both text and images simultaneously and learn the semantic alignment between them. It also utilizes an adaptive generation strategy, which can realize dynamic adjustment of the image resolution and style based on the complexity and details of the text description.

Diffusion model-based Text-to-Image model: The diffusion model archives modeling of data via a diffusion process from data distribution to Gaussian distribution and a denoising process from Gaussian distribution to data distribution. Diffusion models have the advantages such as high training stability, diverse samples, no need for log-likelihood or partition functions, and good mathematical properties. In particular, when combined with text conditions or other guiding techniques, they can generate realistic images matching semantics. However, the forward propagation and reverse diffusion processes take a long time, resulting in a slow model training and inference speed. As a result, they cannot generate images in a single step like the first two models mentioned above, so diffusion models did not receive widespread attention at the beginning.

The GLDE model represents a significant milestone in the field of Text-to-Image generation, but it is not the endpoint. Google has proposed a new Text-to-Image model called Imagen based on the idea of diffusion model. This model introduces a Threshold Diffusion Sampler which can use a larger classifier-free guidance weight, as well as a new U-Nt architecture which boasts higher computational efficiency, memory efficiency, and convergence speed. Imagen directly encodes text information using the T5-XXL large-scale language model and then generate images based on the encoded text using a conditional diffusion model directly. As it directly uses the T5-XXL model and need not learn a priori model, it has richer semantic knowledge than CLIP does. The Stable Diffusion model proposed by Rombach et al. is a Text-to-Image model with high resolution, rich details, and semantic consistency. The model adopts the LDM image generation paradigm. Unlike the previous diffusion process in a pixel space, LDM can realize diffusion process in a lower-dimensional latent space, which can greatly reduce the training and inference costs. Baidu, a search engine in China, released its large-scale Text-to-Image model in 2022, called ERNIE-ViLG2.0. This model has surpassed Stable Diffusion, DALL-E2, and other models in both zero-shot FID-30K and artificial blind evaluation in the MS-COCO validation dataset of the public dataset. The ERNIE-ViLG2.0 model adopts a hybrid denoising expert modeling approach based on the knowledge enhancement algorithm, which allows the model to select different "denoising expert" networks in different generation stages, thereby achieving more detailed denoising task modeling and consequently improving the quality of generated images.

3. Research Method

3.1 Research design

This research employs a mixed research method. First, it evaluated the differences in causing emotional responses of audiences between AI-generated artworks and traditional human artworks through quantitative analysis. Subsequently, qualitative interviews were conducted to gather opinions from art critics, artists, and common audiences, aiming to delve deeper into the uniqueness and limitations of AI art in emotional expression and perception.

In the research process, in order to better differentiate and quantify the type and intensity of emotions, our theoretical research in the field of emotions was based on categorical emotion models. Categorical emotion models, also known as discrete emotion models, are based on the fundamental idea that there are several important emotions that are universally recognized. These emotions are independent, and that means there are no correlations between them. These basic emotions include sadness, joy, fear, anger, disgust, and surprise. Categorical models of emotions categorize emotions into different categories or types. In this category, commonly used models include the Robert Plutchik model, Paul Ekman model, and OCC model. The Paul Ekman model distinguishes emotions based on six basic categories. According to his theory, the six basic emotions arise from different neural systems activated by the way an individual perceives a situation. Therefore, emotions are not interdependent. These basic emotions include sadness, joy, disgust, anger, surprise, and fear. However, in addition to these emotions, other complex emotions may also arise, such as greed, desire, shame, guilt, and pride. The Robert Plutchik model assumes that a few major emotions appear in an opposing set, and their combination gives rise to complex emotions. Plutchik categories the eight basic emotions into opposite pairs: joy vs. sadness, surprise vs. anticipation, anger vs. fear, and trust vs. disgust. Plutchik states that each emotion has a variable degree of intensities[1].

3.2 Quantitative analysis: experimental design

We identified a sample set of 120 AI-generated artworks and traditional paintings, including works obtained from art databases or on-line platforms. In this experiment, we categorized the major emotions conveyed by all the artworks into six basic categories according to Paul Ekman's point of view in the categorical emotion model, i.e., sadness, joy, disgust, anger, surprise, and fear (other complex emotions such as greed, desire, shame, guilt, and pride are not discussed in our preliminary research), which covered various styles and periods.

We made classification based on the combination of emotion types (sadness, joy, disgust, anger, surprise, and fear) and painting types (traditional paintings[2], and AI-generated works), and then divided the pre-selected paintings into 12 categories, with 10 paintings in each category. One of the ten paintings was randomly selected to be observed by the participants, and all categories were observed by participants (this can reduce the additional impact of a single specific painting on participants' emotions). Then, we collected participants' personal information, such as age, gender, and art background, to understand the impact of personal factors on emotional perception.

During the experiment, we did not inform the participants of AI-generated artworks. We only displayed the artworks to them and asked them to score the extent of their emotional touch (on a scale from one to five). Each participant needs to observe 12 artworks and answer 24 questions, with the session lasting for 40 minutes. Finally, we organized the questionnaire results into a data table and performed analysis using SPSS.

---

[1] However, the emotion theory is still controversial now. To facilitate the conduct of our research, we will temporarily refrain from delving into complex theoretical issues. The OCC model proposed by Ortony, Clore, and Collins opposes the concept of "basic emotions" introduced by Plutchik and Paul Ekman. The OCC model claims that emotions arise from people's perception of events, and these events are different in intensity. The OCC model divides emotions into twenty-two categories, and adds sixteen emotions, such as relief, reproach, envy, self-criticism, shame, admiration, disappointment, pity, fear confirmation, admiration, hope, sorrow, satisfaction, schadenfreude, liking, and disliking, to the basic emotions proposed by Paul Ekman. Therefore, it covers a broader range of emotional expressions. Researchers can use any categorical emotion model to describe emotions based on their preferences. However, due to its wider categories, the OCC model provides a more comprehensive range of emotional expressions.

[2] In a broad sense, they are non-AI-generated artworks.

We did not exclude artists and art critics when selecting samples. After the research was ended, we interviewed those who showed interest.

Due to copyright restrictions, only some of the AI work is shown in this article (some of this work is crawled from social media platforms and websites).

Diverse emotional expressions of AI-generated works

happiness

3.3 Qualitative analysis: interview and case study

When filling out the personal information, we noticed that among the 161 samples, artists and art critics were more active in the relevant tests. On the premise of not interfering with the overall experiment result, we conducted interviews on relevant topics, and sorted out expressions of the same type, as shown below.

The following are interviewees (in pseudonyms):

Tang Yuhan (young artist, once studied at the China Academy of Art), Zhang Yue (freelance artist, with works exhibited multiple times), Li Yuer (contributor to an art magazine), Wang Li (student), Li Mu (programmer), Wang Zihan (teacher)

Question 1:

In the whole process, did you realize that some of the artworks we displayed were drawn by AI?

Tang Yuhan: Yes. Actually, when I saw the fourth painting, I could tell that it was AI-generated.

Question 2:

What differences do you perceive in emotional expression between AI-generated paintings and traditional paintings?

Zhang Yue: At first glance, it may not be apparent, but upon closer observation, the traces of AI generation become quite evident. For instance, as to the emotional expressions such as bright colors or sharp lines in the painting, AI can mimic them very well, and sometimes even express stronger emotions. However, upon deeper reflection, you may find the painting somewhat strange, and you seem to be unable to comprehend the source of emotions conveyed by the painting.

Wang Li: I think the artworks generated by AI seem to look better, because they have very novel and complex visual effects. However, they don't leave a lasting impression after you view them because they don't touch your heart.

Li Yuer: Actually, a painting created by AI did impress me. The painting shows a very intense atmosphere with a very complete form. Although its theme is somewhat cliché, when I confirmed that this painting was created by AI, I felt very surprised, even a sense of being deceived.

Wang Zihan: My feeling is quite subtle. Some paintings give me the impression that a person doesn't understand what sadness is, and he can only mimic others shedding tears, but he does not understand why he cries. Some AI-generated artworks are excellent in form, but they lack the genuine emotional expression like human beings. When looking at these paintings, it seems like you can't grasp the emotions and information of the creator, and you can't resonate with him.

3.4 Analysis of quantitative data:

I collected 155 valid answers from 161 volunteers. Among the participants, the oldest is 55 years old, and the youngest is 25 years old, with an average age of 34.76 years. Among them, 52.67% are female, and 47.24% are male (non-binary or other genders are not considered here). The average age is 22.7 years. Among these participants, 61.33% have a bachelor's degree or below, while 37.68% have a master's degree or higher.

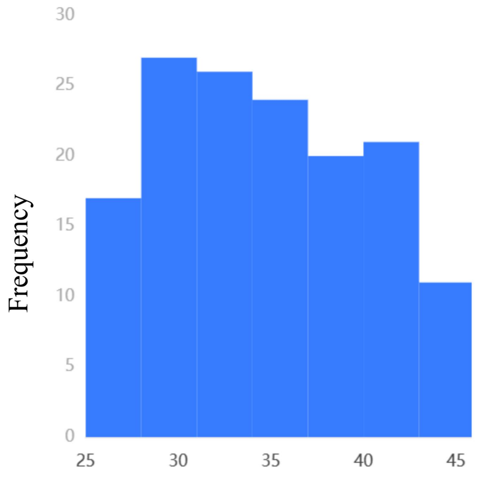

Figure 1: Age distribution of participants

Data source: Questionnaire survey

Table 1: Age distribution data of participants

Type	Quantification	Maximum	55
Sample size	151	Minimum	25
Missing value	1	Median	34
Deduplication	25	Coefficient of variation	0.172
Mean value	34.76	Variance	35.821
Standard deviation	5.985	S-W normality test	Satisfied (P=1.000)

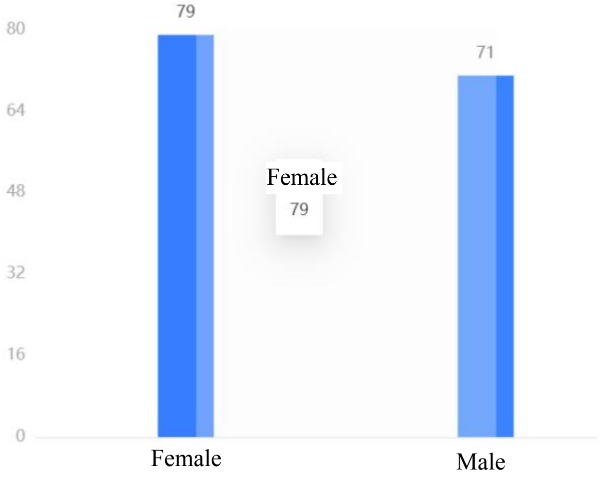

Figure 2: Statistics of gender of participants

Data source: Questionnaire survey

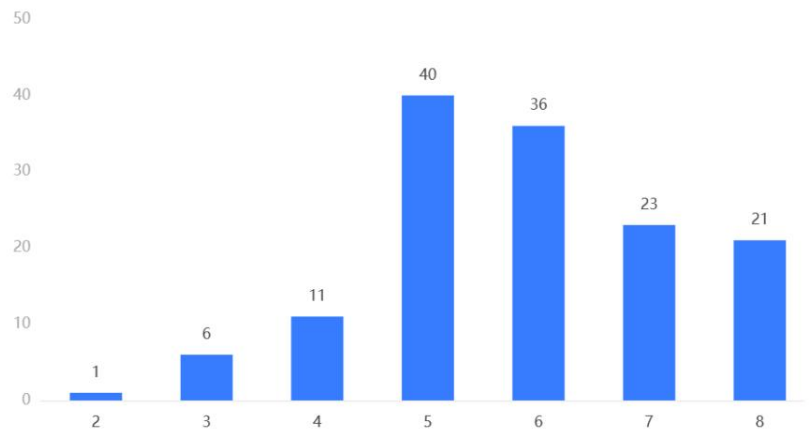

Figure 3: Statistical graph of scores on the perception of AI-generated artworks

Data source: Questionnaire survey

Table 2: Statistical data of scores on the perception of AI-generated artworks

Type	Quantification	Maximum	9
Sample size	151	Minimum	2
Missing value	2	Median	6
Deduplication	8	Coefficient of variation	0.255
Mean value	6.094	Variance	2.424
Standard deviation	1.557	S-W normality test	Satisfied (P=1.000)

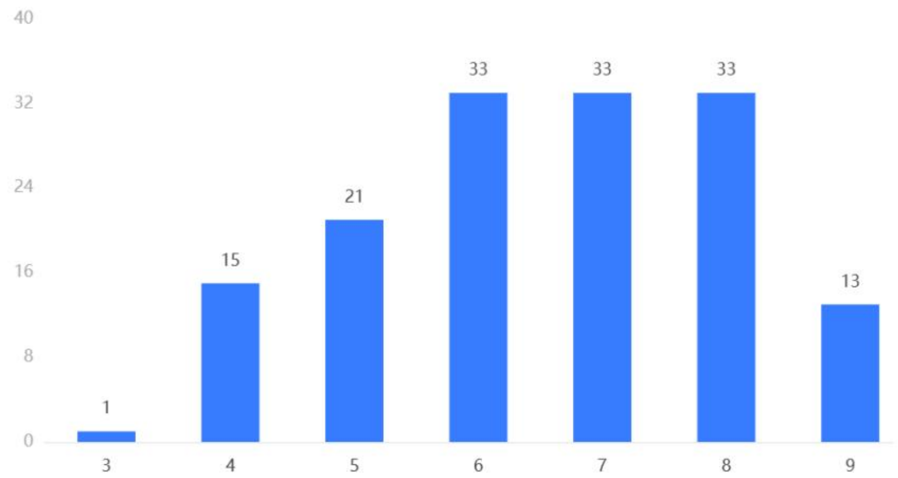

Figure 4: Statistical graph of scores on the perception of traditional paintings

Data source: Questionnaire survey

Table 3: Statistical data of scores on the perception of traditional paintings

Type	Quantification	Maximum	9
Sample size	151	Minimum	3

Missing value	2	Median	7
Deduplication	7	Coefficient of variation	0.226
Mean value	6.564	Variance	2.194
Standard deviation	1.481	S-W normality test	Satisfied (P=1.000)

We can see that the scores on the perception of AI-generated artworks are generally lower than those of traditional paintings. Among the scores on the perception of traditional paintings, there are 13 instances where the score is 9 or higher. However, among the scores on the perception of AI-generated artworks, there are no instances where the score is 9. It is worth noting that among the artworks with a score of 6, there are 36 AI-generated artworks and 33 traditional paintings, with a difference of only 3. In addition, traditional paintings mostly distribute in the sections of 6, 7, and 8 scores, while AI-generated artworks mostly distribute in the section of 5 and 6 scores.

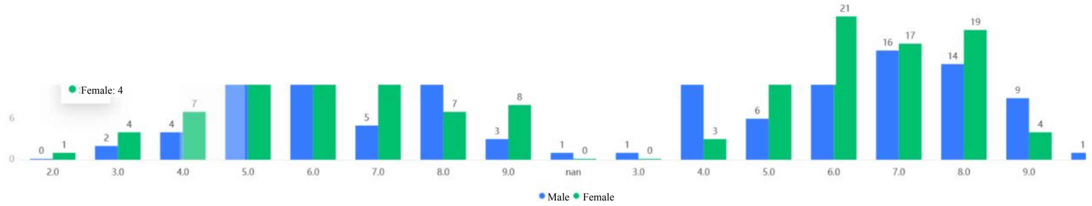

Figure 5: Distribution of males and females in the scores on the perception of AI-generated artworks and those of traditional paintings

Data source: Questionnaire survey

It is interesting that we have also discovered the subtle difference in perception scores based on the gender factor. For the scores on the perception of AI-generated artworks, 24 males gave a score of 5, and only 16 females gave the same score. Only 5 males gave a score of 7 for AI-generated artworks, while the number of female participants giving such a score reached 18. For traditional paintings, 15 and 17 females gave scores of 5 and 7, respectively, which is not much different from the number of females giving such scores to AI-generated artworks. This seems to indicate that, in the 5 and 7 score ranges, males are more likely to have different emotional responses to AI-generated artworks compared to traditional paintings.

4 Research Results

4.1 Quantitative research results

Throughout the entire experiment, we found that the perception scores for AI-generated artworks were generally lower than those for traditional paintings. This difference was particularly noticeable in the high-score range (9 scores), indicating that AI-generated artworks cannot match the level of emotional expression achieved by traditional paintings in conveying strongest emotions.

4.2 Qualitative research results

In qualitative analysis, we found that interviewees repeatedly mentioned the word “form” and tended to use words like “novel”, “complex”, “visually striking”, and “complete” to describe the AI-generated artworks. When discussing the limitations of AI-generated artworks, they used words like “resonate” and “understand” to express their thoughts. According to the interview, we can draw the conclusion that although AI-generated artworks are good in visual complexity and novelty, they still have limitations in conveying deep emotions and personal experiences, which affects the depth of emotional resonance.

At the same time, we discovered a very interesting phenomenon. When interviewees were informed that a painting was generated by AI, their reactions were generally divided into two types. Some had anticipated that, and others were shocked. Some even expressed the emotion of being deceived. In the interview with the latter, we found that once we told them the artwork was generated by AI, the topic of our conversation shifted from the artwork itself to the AI image generation technology. This is indeed an interesting phenomenon. When people realize that an artwork is created by AI, the emotions they had previously invested in it seem to vanish instantly. In other words, the artwork does not resonate as easily on an emotional level as it would if people believed it was created by a human.

4.3 Interpretation of the research findings

Through qualitative and quantitative analysis, we have gained a preliminary understanding of the differences between AI-generated artworks and traditional paintings in emotional conveyance. The greatest differences are in two aspects: One is that AI-generated artworks tend to receive perception scores that are concentrated in the 5 to 6 range; the second is that there were no AI-generated artworks in the highest score range. In response to this, we give interpretations from three aspects. First, we believe that the result is consistent with normal distribution. However, whether the result indicates that AI-generated artworks cannot achieve the highest emotional conveyance level cannot be demonstrated solely by this experiment. Secondly, from the perspective of the dataset, against the background of mass culture and the internet information age, images with relatively ordinary emotional conveyance abilities dominate, which directly affects the emotional conveyance quality of AI-generated artworks. Finally, compared to the traditional artistic creation process, the process of AI generation has its shortcomings. Currently, AI is unable to form a complete chain from emotional perception to emotional conveyance. Manifested in the emotional expression ability of AI, there is a lack of an innate emotional logic, resulting in the inability to resonate with viewers.

4.4 Interpretation of research results from the perspective of art theory analysis

From the perspective of art theory, the entire research can be considered as a test of art, where AI-generated artworks compete with traditional paintings. The reason for this is that AI is gradually permeating our lives, and now it has extended to the field of art. Art has long been regarded as a typical uncontrolled domain, and its complexity and insolvability will not diminish as algorithms are introduced. Therefore, the convergence of aesthetics and AI becomes crucial. Many people believe that art, aesthetics, and creativity represent the pinnacle of human abilities: They are not something that can be replicated simply as technology advances. Therefore, they are considered the final barrier that prevents the progress of AI. However, this barrier seems to be challenged. The research results show that AI is developing rapidly in the field of art, and the stronghold of human dominance in art appears to be at risk. It must be acknowledged that the development of AI has posed significant challenges to the research of modern art theory. At the same time, it has revealed some complex artistic phenomena, facilitating the modern art theory to develop further.

Puzzles of creativity, meanings, and emotional expression:

In our research, we found that although AI-generated artworks did not have an emotional conveyance ability higher than that of traditional paintings, AI can generate a creative image with a level rival that of human's creations. As we mentioned before that interviewees often used words like "novel" to describe the visual effects of AI-generated artworks, which is worth noting. According to mainstream societal views, creativity is considered to be a domain exclusive to humans, while machines are mainly responsible for repetitive mechanical tasks. However, when machines and computers advance to the stage of AI, we find that the creativity of AI-generated artworks seems to be on par with that of humans. This contradicts our traditional understanding, because according to the art theories in the past, people believed that emotional expression and creativity were both based on the concept of autonomy. We temporarily regard AI as a non-autonomous mechanical creation. Then where do the emotional conveyance ability and creativity of AI derive from?

Actually, the concept of pure creativity stems from the admiration of individual autonomy[3]. When a concept arises from the admiration of something, we must keep vigilant. This is because what we may grasp is not an objective concept, but a conceptual tool for storytelling. Thinking in this way, we may raise a question: Is there an absolute relationship between creativity and individual autonomy? Perhaps we can answer this question through a case study. Lisi Edwards and his colleagues trained a system by analyzing some large image databases and inferred specific evolutionary patterns. Then, they generated images with a new style that matched the previous predictions, with a goal to showcase the likelihood of predicting the development of artistic styles.

What does this have to do with our question? According to the authors, prediction of the system is surprisingly consistent with the actual evolution of such a style in the history of visual arts, showing that the development of specific styles is based on algorithms. This suggests that the evolution of style is not merely an accidental product of history or the spontaneous inspiration of a unique group of artists; instead, it is the inevitable development guided by internal rules. From this perspective, the development of style may not be deterministic, but it is still the product of a finite combination of a series of elements that can be detected and replicated by the data analysis system. In the combination of infinite data elements, a clear developmental path ultimately emerges by the adoption of a complex screening mechanism. This path not only represents the evolution of artistic styles but also reveals the crucial role of creativity within it. Furthermore, "creativity" is a label used by observers to describe the phenomena whose underlying processes they do not understand. In computers, what we call creativity is actually the arrangement and combination of different data elements.

We have another case to further illustrate this point. In 2016, Lee Sedol was defeated by AlphaGo, so he believed that this program could create innovative ways of playing chess. It reveals that some chess moves or strategies that were once thought to be creative are actually entirely predictable. In the second match, AlphaGo made a move that was considered highly creative by many critics, catching the opponent off guard and helping the computer win the game. Observers consider this move creative just because neither the players nor the experts truly understand the underlying strategy behind AlphaGo. In fact, from the perspective of the machine, this move is the product of optimization process evaluation, and other possible moves have also experienced the same optimization process to be selected. In this sense, when we consider something as "creative", it is because we lack of understanding of it. We are familiar with what we know and consider it ordinary, while regarding the unknown as unusual. In other words, if we believe that humans are creative but AI is not, it is because we have a deeper understanding of the working process of AI, whereas we do not have such a comprehensive understanding of humans. The development of technology seems to prove that the so-called unusual phenomena are actually the products of ordinary processes.

From this perspective, we might have reached a critical juncture. We may either completely abandon the concept of creativity, or continue to maintain our common understanding of creativity and agree to extend this concept to non-human phenomena. When it comes to judging certain specific cases in art,

---

[3] Of course, such admiration is widely recognized in modern society, whereas it was the opposite in ancient times. Plato's "Theory of Recollection" is a typical example. Art history is the process of reconstructing or replicating existing things. In this sense, artists are merely discoverers rather than creators; art is not pure invention but rather an expression through skillful imitation of reality. In ancient times and the Middle Ages, true creativity was understood as "creatio" or "ex-nihilo", considered to be the exclusive prerogative of God. People believe that the historical development of artistic styles is the unpredictable outcome of leaps in creativity. We can reconstruct it through retrospective analysis, but its emergence cannot be predicted in advance.

the situation is complex. In fact, in the previous discussion, we have already revealed some mechanical attributes of creativity. However, further speaking, can the combination of data elements alone fully encapsulate the creative connections in the human mind and people's recognition of such creativity? In reality, when computers replicate human works, they do not follow the same process that humans use to create these works in real life. No one would believe Piet Mondrian created his works following an algorithm similar to that used by the pseudo-Mondrian in 1966, even though the public prefers the artworks generated by AI over the original one. We cannot overlook the symbolic, historical, and conceptual significance behind the innovative styles of artists, nor can we ignore their roles in the development of abstract art, figurative art, expressionism, and minimalism. In other words, algorithms cannot replicate the cultural process through which Piet Mondrian created abstract paintings. On the contrary, the program merely imitates the final product superficially. We appreciate Mondrian's paintings because they represent the artistic journey of the artist during the creative process, as well as the cultural role he plays in the history of painting.

Representation of contemporary artistic styles in AI-generated artworks

In summary, it can be observed that the relationship between creativity and autonomy has been initially disentangled. However, even in this process of disentanglement, there still exists a hidden interconnected connection. In such a process of exploration, AI-generated artworks seem to serve as a control group for the convenience of our deeper understanding of them. If creativity is no longer solely contingent upon human subjectivity, then in this contrast, how will meaning and emotion manifest themselves?

In classical art, the unique thoughts, emotions, imagination, and creativity of humans as the creators, namely free will and reflective consciousness, are the core elements that define art as art. Correspondingly, the individual sensory experiences and emotional interests that artists draw upon during their creative process are often regarded as significant sources of interpretation for the meaning of art. By contrast, in AI art, art is generated based on "computation" (gaming); in other words, art is generated based on the aesthetic architecture of algorithms. In this way, the artistic skills reflected in classical artistic creations formed in the process of long-term practice are immediately replaced by the algorithmic logic of AI.

In such a contrast, we can see that the cognitive linguistic significance of AI art essentially originates from the formal syntactic relationship between language and symbols. If symbols can naturally convey emotions, then the human aesthetic logic of "emotion manifested by symbol" can be reversed into the AI art logic of "symbol evoking emotion". Therefore, we can directly explore the semiotic significance of AI art without the need to change existing artistic evaluation rules. If we must relate the emotional and social elements of symbols to understand them, then AI art becomes a quasi-art form similar to human art. The interpretation of the meaning of AI art is to place its algorithmic logic (including program modules, art samples, aesthetic principles, etc.) within the social context of art, much like how we place modern art within the context of "art conventions" after the time of Duchamp.

The factors that create meaning are also the factors that interpret artistic meaning. AI can create "art" reflecting conceptual instructions based on these instructions, but humans need to understand artworks based on practical perception and tangible objects. Art, as the crystallization of human wisdom, is the result of the evolution of human creativity over long periods of social life. Artworks integrate complex elements such as the artist's emotions, ethics, beliefs, and attitudes towards survival and life. From the perspective of human art, AI lacks the necessary social elements because it can produce results directly based on instructions solely without undergoing an evolutionary process. Therefore, even though AI can paint or produce beautiful "music" programmatically like elephants, chimpanzees, and other primates and dolphins, it is difficult to directly consider them as writers or artists, unless AI art works are supplemented with the "social elements" they lack. As mentioned earlier, the input and output in AI art creation are similar to the stimulus-response relationship in organisms. It differs from human artistic creation in its emphasis on process and procedural elements, thus forming the interpretation of the meaning of AI art. We should focus our analysis on the artistic realms created by AI algorithm systems that can mirror human emotions and imagination, namely, a "new nature" world with distinct literary or aesthetic qualities, but not the "bit world" constructed by AI algorithm systems. As AI algorithm systems are capable of understanding this embodied world, and comprehending and even predicting the artistic outcome when expressing "emotions" and "imagination", our admiration for AI art can extend beyond mere binary data. Instead, we appreciate the new visual world it crafts through transformations including language, lines, colors, structures, etc. Thus, the ultimate interpretation of AI art is not only a conceptual interpretation of the "intention world" and the world of "emotions" and "imagination" in AI art but also an embodied interpretation of all creative "subject" activities, including those of humans.

The ultimate significance of AI art not only depends on our understanding of the semantic relationship between AI art inputs and outputs, but also depends on our actual cognition of emotional communication and creativity between humans and nature. If all human life experiences can be expressed using literary and creative language and symbols, AI can also select appropriate information from the massive database of language and symbols for simulation expression. In other words, the essential distinction between human art and AI art does not solely lie in the complexity and uniqueness of expression. If AI cannot comprehend the symbolic meaning in its own works, even if we define the emotions based on expressive and reflective abilities as the fundamental characteristic of human art, art is not about "conveying" but about "expressing". The former is a natural physical process, while the latter signifies a spiritual process; the former conveys appearances, while the latter expresses ideas. The former stems from logical thinking, whereas the latter arises from imaginative thought. Without identifying the AI that rivals human intelligence, we cannot transcend the limitations of expression and reflection to unlock boundless opportunities for our own generation based on "emotions".

This indicates that, to achieve a comprehensive understanding of AI art, we must progress gradually from the understanding of sensory aesthetics and reflection to pure conceptual interpretation and concrete experiential understanding. This process mirrors the shift observed in the Western contemporary art in the 20th century, i.e., transitioning from sensory forms towards conceptual interpretation and embodied understanding. As elevated by hermeneutics, AI art transcends the realm of contemporary aesthetics (art of sensibility) and enters the domain of art philosophy (conceptual mind). This transformation shifts the essence of artistic aesthetics ("existence is perception") to the essence of artistic interpretation ("existence is interpretation"). Consequently, interpretation becomes a fundamental element for AI art to demonstrate its own existence and the generation of meaning.

5. Conclusion

5.1. Research conclusion

Although we adopted the data analysis method in our experiment, we found during our research that how to summarize fuzzy natural language and transform it into a clear data model is a highly complex task. In fact, in the classification of emotional types, we had to simplify visual emotions by discarding its other content. Moreover, during the sampling of artworks, a significant degree of randomness persists, so we need to further expand the scope of works to demonstrate the outcomes of the experiment more precisely .

Furthermore, various emotions may manifest different degrees of immediate impact, accompanied by diverse forms of expression. Whether AI has its strong point in conveying different emotions remains unexplored.

5.2 Recommendations for artistic creation

In the process of artistic creation, how artists establish a collaborative and progressive relationship with AI remains an important question to consider. Although currently, AI may not excel in higher-quality emotional creativity, does this suggest that in the future, when artists employ AI to express their emotions, it will still fall short in emotional expression?

In future artistic creation, this remains a direction worthy of our exploration.

References:

[1]Bartneck, C., & DiFranzo, D. (2023). The Perception of AI Art in Contemporary Society. Public Opinion Quarterly, 87(1), 159-178. doi: 10.1093/poq/nfz034

[2]Boden, M. A. (2020). Ethics of Artificial Creativity: Issues and Challenges. Ethics and Information Technology, 22(2), 123-136. doi: 10.1007/s10676-020-09542-6

[3]Colton, S. (2006). The Painting Fool: Stories from Building an Automated Painter. In: Creativity & Cognition, 4(2), 3-12. doi: 10.1145/3056662

[4]Davis, H. (2021). Emotion and Expression in AI Art: The Limits of Artificial Creativity. Journal of Aesthetic Studies, 5(2), 134-145. doi: 10.1017/jas.2021.12

[5]Du Sautoy, M. (2021). Can Machines Be Creative? AI and the Future of Art. Annals of Machine Creativity, 2(1), 45-59. doi: 10.1016/j.amcr.2021.01.003

[6]Ekman, P. (1992). An Argument for Basic Emotions. Cognition & Emotion, 6(3/4), 169-200. doi: 10.1080/02699939208411068

[7]Fiebrink, R. (2018). Artificial Imagination: The Limits and Potentials of AI in Creativity.

Creativity Research Journal, 30(2), 156-163. doi: 10.1080/10400419.2018.1446744

[8]Goodfellow, I. J., et al. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems, 27, 2672-2680. doi: 10.5555/2969033.2969125

[9]Henderson, L. D. (2019). AI as a Tool for Understanding Art: Possibilities and Limitations. Art History, 42(4), 720-735. doi: 10.1111/1467-8365.12449

[10]Herzberg, E. (2022). AI and the End of Artistic Labor. Arts & Technology, 9(4), 230-244. doi: 10.1016/j.artech.2022.03.005

[11]Isola, P., et al. (2017). Image-to-Image Translation with Conditional Adversarial Networks. CVPR, 5967-5976. doi: 10.1109/CVPR.2017.632

[12]Karras, T., et al. (2019). StyleGAN: A Generative Adversarial Network for Style-Based Generative Modeling. Proceedings of the Computer Vision and Pattern Recognition, 43(6), 8491-8500. doi: 10.1109/CVPR.2019.00873

[13]Miller, A. I. (2022). Artificial Artists: The Mechanics of AI Creativity. Journal of Creative Behavior, 56(1), 70-85. doi: 10.1002/jocb.455

[14]Patashnik, O., et al. (2021). StyleCLIP: Text-Driven Manipulation of Latent Styles for Realistic Image Generation. ACM Transactions on Graphics, 40(4), 1-14. doi: 10.1145/3450626.3459936

[15]Pepperell, R. (2021). The Role of Artificial Intelligence in the Future of Artistic Labor. Art & Labor, 12(3), 88-102. doi: 10.1080/10632921.2021.1902233

[16]Plutchik, R. (2001). The Nature of Emotions. American Scientist, 89(4), 344-350. doi: 10.1511/2001.4.344

[17]Qiao, T., et al. (2021). MirrorGAN: Learning Text-to-Image Generation by Redescription. CVPR, 1505-1514. doi: 10.1109/CVPR42600.2021.00153

[18]Reed, S. E., et al. (2016). Generative Adversarial Text to Image Synthesis. Proceedings of the 33rd International Conference on Machine Learning, 48, 1060-1069.

[19]Sanakoyeu, A., et al. (2018). Style-Aware Content Loss for Real-time HD Style Transfer. ECCV, 698-714. doi: 10.1007/978-3-030-01234-2_41

[20]Sedol, L. (2016). Reflections on My Match with AlphaGo. Journal of Artificial Intelligence Research, 55, 1-10. doi: 10.1613/jair.1.11168

[21]Wing, J., & Grimmelmann, J. (2020). Who is the Author? AI and the Art of Creation. Harvard Journal of Law & Technology, 33(1), 209-234. doi: 10.2139/ssrn.3539024

[22]Xu, et al. (2017). AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks. CVPR, 1316-1324. doi: 10.1109/CVPR.2017.147

[23]Zhu, J. Y., et al. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2242-2251. doi: 10.1109/ICCV.2017.244

[24]Zu, K., et al. (2019). BigGAN: Large Scale GAN Training for High Fidelity Natural Image Synthesis. ICLR, 1-17.