International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 4555-4563

https://doi.org/10.70517/ijhsa464299

A Hierarchical Guidance Design Method for Oil Painting Pictures Based on the Probability Distribution of Visual Attention

Baoqun Wang[1,*]
[1] School of Fine Arts and Design, Huainan Normal University, Huainan, Anhui, 232038, China
Corresponding authors: (e-mail: Wbqun2012@hnnu.edu.cn).

Abstract Traditional oil painting relies on the artist's subjective experience and skill accumulation, and the creation process is time-consuming and uncertain. In the field of modern digital art, although computer-aided painting technology provides convenient tools, it still has obvious deficiencies in simulating the texture, hierarchy and visual guidance effect of real oil paintings. This study proposes an oil painting picture hierarchical guidance design method based on the probability distribution of visual attention, and realizes intelligent oil painting generation by constructing Markov attention transfer model and hierarchical fusion generating adversarial neural network. In the method, the state distribution function is used to establish the attention transfer probability matrix, and the joint training framework of structural GAN and texture GAN is designed to train the model using a dataset of 700 oil painting images. The experimental results show that compared with the suboptimal algorithm Pix2PixHD, this method improves 7.024 in FID index, 6.38 in CFID index compared with PD-GAN, and the Manhattan distance is reduced to 36.19, which is significantly better than PD-GAN's 75.33. As verified by the visual effect evaluation of 41 subjects, the generated oil paintings are better in terms of the overall aesthetics, color fullness and compositional momentum, all of which show good hierarchical and visual guidance effects. This method provides a new technical path for digital art creation, and improves the creation efficiency while maintaining artistry.

Index Terms Visual Attention Probability Distribution, Generative Adversarial Network, Oil Painting Hierarchy Design, Markov Model, Image Generation, Art Creation

I. Introduction

In the creation of oil painting, the picture level is the most basic and important content to express the creator's painting emotion [1]. Picture hierarchy is manifested in the grouping, combination and configuration of creative contents in the works, and different creators need to deal with the synergistic relationship of picture hierarchy with the ideal picture effect at the early stage of painting [2], [3]. Similarly, in excellent artistic creation, the picture hierarchy is not only manifested in the appropriate layout of the painted content, but also highlights the creator's emotional expression of art [4], [5].

The picture hierarchy itself is a creative activity with a lot of thought, just like a wonderful movie, which not only needs to have a full and rich picture, but also needs to have a vivid plot, but also needs to be full of fun and artistic aesthetics [6]-[8]. However, the traditional picture level design of oil paintings basically relies on the artist's own experience, coupled with the insufficient application of technology, which leads to the significant limitations of picture level design [9]-[11]. In contrast, the design method based on the probability distribution of visual attention can achieve intelligent oil painting picture hierarchy design through attention probability modeling, hierarchy optimization, and style enhancement [12]-[14].

This study proposes a technical route that combines visual attention theory with generative adversarial networks. Firstly, a visual attention probability distribution model based on Markov process is constructed, which describes the dynamic change law of the audience's attention in the process of appreciating the oil painting through the state transfer matrix, and provides theoretical guidance for the subsequent picture hierarchy design. Then we design a hierarchical fusion generative adversarial neural network architecture, including two sub-networks, structural level GAN and texture level GAN, which are responsible for generating the line structure and texture details of the picture, respectively. Through the joint training mechanism, the effective fusion of structural and texture information is realized, and finally oil paintings with a good sense of hierarchy and visual guidance effect are generated.

II. Based on the visual attention probability distribution model

II. A. Construction of the Visual Attention Shift Model

II. A. 1) Setting of the state distribution function

It has been shown that in a relatively single picture condition and a short continuous time segment, the stable duration of attention in a particular state is exponentially distributed. Accordingly, the distribution function of T_i for the duration of attention in the i th state is $G_i(t) = 1 - e^{-\lambda_i t}$. where $\lambda_i = 1/t_i$, where $\overline{t_i}$ is the average duration of attention located in the i th state.

II. A. 2) State transfer analysis and determination of transfer probability matrices

The transfer probability a_{ij} in the transfer probability matrix can be determined in conjunction with the event probability operator relationship.

(1) Visual attention is focused on the 0-state at moment t and remains in the 0-state at moment $t + \Delta t$:

$$\begin{aligned} P_{00}(\Delta t) &= P\{X(t+\Delta t) = 0 \mid X(t) = 0\} \\ &= P\{T_0 > \Delta t\} = 1 - \lambda_0 \Delta t + o(\Delta t) \end{aligned} \tag{1}$$

(2) t moment attention focused on the 0 state, $t + \Delta t$ from the 0 state jump out of the case, analogous to the idea of great likelihood estimation can be considered, the probability of jumping from the 0 state to the i state and the average lag time of the attention is located in the state is proportional to the average lag time, and at the same time, due to the probability of the transfer between the probability of satisfying the $P_{01}(\Delta t) + P_{02}(\ Deltat) = 1 - P_{00}(\Delta t)$, i.e., the attention must jump to one of the two states 1 and 2 after jumping out of the O state, so $P_{01}(\Delta t), P_{02}(\Delta t)$ is:

$$P_{01}(\Delta t) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \lambda_0 \Delta t + o(\Delta t) \tag{2}$$

$$P_{02}(\Delta t) = \frac{\lambda_1}{\lambda_1 + \lambda_2} \lambda_0 \Delta t + o(\Delta t) \tag{3}$$

(3) Determination of the transfer probability matrix A is homogeneous due to the 0, 1, 2 three-state structure and symmetric homogeneity of positions:

$$\begin{cases} P_{11}(\Delta t) = 1 - \lambda_1 \Delta t + o(\Delta t) \\ P_{10}(\Delta t) = \dfrac{\lambda_2}{\lambda_0 + \lambda_2} \lambda_1 \Delta t + o(\Delta t) \\ P_{12}(\Delta t) = \dfrac{\lambda_0}{\lambda_0 + \lambda_2} \lambda_1 \Delta t + o(\Delta t) \\ P_{22}(\Delta t) = 1 - \lambda_2 \Delta t + o(\Delta t) \\ P_{20}(\Delta t) = \dfrac{\lambda_1}{\lambda_0 + \lambda_1} \lambda_2 \Delta t + o(\Delta t) \\ P_{21}(\Delta t) = \dfrac{\lambda_0}{\lambda_0 + \lambda_1} \lambda_2 \Delta t + o(\Delta t) \end{cases} \tag{4}$$

The above analysis leads to the transfer rate matrix.

II. A. 3) Calculation of the Attention State Distribution Vector

Assuming that the initial moment of attention is located in front of the oil painting, based on Markov correlation theory, a differential equation is established to solve the attention state vector.

$$\begin{cases} (p_0'(t), p_1'(t), p_2'(t)) = (p_0(t), p_1(t), p_2(t))A \\ (p_0(0), p_1(0), p_2(0)) = (1, 0, 0) \end{cases} \tag{5}$$

Firstly, we do Laplace transform to solve both ends of Eq. (5), and then we do Laplace inversion to get the attention state vector. Since the above inversion process requires solving the one-element higher order equation, which is more complicated, the generalized form of the moment state vector is not given. In practical applications, it can be solved by programming according to the actual parameter values. Since $\lim_{i}\to \infty p_i(t) = p_i, \lim_{t\to\infty} p_i^{'}(t) = p_i$ and also combining $p_0 + p_1 + p_2 = 1$, the probability distribution of attention in steady state can be found.

II. B.Approximate solution to the probability of attentional states

In this paper, we propose a probabilistic approximation of attentional state solution based on time-slice segmentation, the idea of which is to construct a model by segmenting the continuous long-time gaze time into time slots, and collecting the parameters in each segmented time slot, where the attentional state distribution $(p_0^{k-1}(15), p_1^{k-1}(15), p_2^{k-1}(15))$ at the end of the previous time slot is the initial value of the next one, and calibrating the attentional state distribution of the short continuous time slots within that time slot:

$$\begin{cases} (p_0^{'k}(0), p_1^{'k}(0), p_2^{'k}(0)) = (p_0^k(0), p_1^k(0), p_2^k(0))A_k \\ (p_0^k(0), p_1^k(0), p_2^k(0)) = (p_0^{k-1}(15), p_1^{k-1}(15), p_2^{k-1}(15)) \end{cases} \tag{6}$$

The above time period models are sequentially added into the set P to form a function set $P = \{(P_1^1(t), P_2^1(t), P_3^1(t)), (P_1^2(t),..P_2^2(t), P_3^2(t)), \cdots, (P_1^N(t), P_2^N(t), P_3^N(t))\}$ for the probability of the audience's attention state during the whole process of oil painting level, where N is the number of time periods.

III. Hierarchical Design Model of Oil Painting Picture Based on Visual Attention Guidance

III. A. Generating Adversarial Neural Networks

III. A. 1) Generator

The generator of a GAN produces simulated data from a set of random noise vectors that undergo a series of complex nonlinear transformations. The fundamental goal is to generate data samples that are as realistic as possible, enough to “trick” the discriminator. This means that the data distribution generated by the generator is as close as possible to the real data distribution. While generators can take many forms, neural networks are widely used because of their powerful generalized function approximation algorithms. It is worth noting that too deep a network structure may cause problems such as gradient vanishing or gradient explosion, which requires a careful balance between network depth and performance when designing the generator.

III. A. 2) Discriminators

In Generative Adversarial Networks [15], the data generated by the generator is labeled as “fake data”, while the data from the real dataset is treated as “real data”. The role of the discriminator is to handle both types of data, analyzing and identifying the authenticity of the input data through its built-in network structure, and outputting a decision value accordingly. This decision value is usually a probability value between 0 and 1, reflecting the probability that the discriminator will judge the input data to be true. If this value is close to 1, then the discriminator considers the input data to be true; conversely, if the value is close to 0, then the discriminator considers the input data to be false (i.e., created by the generator). The main task of the discriminator is to distinguish whether the input data is real or created by the generator, similar to a binary classification problem, and its network architecture usually employs a neural network, which is similar to that of the generator, in order to efficiently learn and model complex distributions of data.

III. A. 3) GAN training process

In GAN networks, the generator and discriminator each have different loss functions that are used to guide the training process.

Together, these two loss functions form the core mechanism of GAN training. The discriminator is a binary classification model whose loss function is usually designed as a binary cross-entropy loss, which is used to measure the discriminator's ability to correctly discriminate between real and generated data, i.e., the

$$J(D) = -\frac{1}{2}E_{x\sim P_{dace}(x)}[\log D(x)] - \frac{1}{2}E_{z\sim P_z(z)}[\log(1 - D(G(z)))] \tag{7}$$

where x is a sample of real data; z is a random noise vector with distribution P_z; The $data$ is the real data and its distribution is P_{data}; $G(\cdot), D(\cdot)$ denote the generating mapping function and the discriminant mapping function, respectively, $G(z)$ denotes the data generated by the generator based on the random noise z, and $D(x)$ denotes the probability that the discriminator outputs the data x as the real data.

The generator's loss function will vary depending on the form of the game. For the simplest binary zero-sum game, the generator's loss is what the discriminator gets:

$$J(G) = -J(D) \tag{8}$$

At the early stage of training, since the generator has not been sufficiently trained to generate realistic data, the discriminator can easily distinguish the generated data from the real data, resulting in the generator not getting enough gradient. Therefore, related literature proposes to train the generator by maximizing $\log D(G(z))$ instead of minimizing $\log(1-D(G(z)))$. In addition, the weak gradient problem of the generator can be solved by replacing the very, very small game with a non-saturated game to obtain the loss functions of the discriminator and generator as follows:

$$\begin{cases} J(D) = -\frac{1}{2}E_{x-P_{daa}(x)}[\log D(x)] - \frac{1}{2}E_{z-P_z(z)}[\log(1-D(G(z)))] \\ J(G) = -\frac{1}{2}E_{z-P_z(z)}[\log(D(G(z)))] \end{cases} \tag{9}$$

In this approach, the generator's loss function is no longer simply the negative of the discriminator's loss function, but is defined based on its camouflage ability, thus providing a more efficient gradient signal for learning optimization.

Based on the respective loss functions of the generator and the discriminator, the objective function of the GAN can be formulated as a minimization-maximization problem, defined as the following equation:

$$\min_G \max_D V(D,G) = E_{x-P_{dina}(x)}[\log D(x)] + E_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{10}$$

The above equation essentially contains two opposing optimization problems: the generator works to minimize a function, while the discriminator view maximizes it.

III. B. Hierarchical Modeling of Oil Painting Pictures

III. B. 1) Model Architecture

The paper proposes a hierarchical fusion based on image structure and texture to generate an adversarial neural network to realize the intelligent generation of drawing images, thus assisting the creators to create works with better results. Firstly, the drawn sketches are input into the system model and normalized preprocessing is performed to ensure that the image conforms to the model. Next, the image is sampled from the uniformly distributed noise and the sampled data is fed into the structural GAN, which in turn generates the line structure of the image, and then the texture GAN is utilized to fuse the structural and textural features of the image, so that a realistic image can be output in the end.

III. B. 2) Structural Hierarchy GAN

Structural GAN uses Deep Convolutional Generative Adversarial Network (DCGAN) to learn to generate structural line graphs. The input to the generative network G is taken from uniformly distributed noise and the output is a structural line drawing. In the paper, a 100-dimensional vector is used to represent the input noise z and the size of the output image is 72×72×3. The discriminative network D will classify the generated structural line drawings based on the real images obtained from deep learning. The generative network uses a 10-layer model and performs convolutional operations after passing through the fully connected layers to finally generate the structural line drawings. Batch normalization is used for each layer in the setup and ReLU activation function is used except for the last layer which uses Tanh activation function.

III. B. 3) Texture Hierarchy GAN

The generative network is modified into a conditional GAN, i.e., the conditional information is used as an additional input to the generator G and the discriminator D. The generative network is modified into a conditional GAN. The structural line drawings are used as additional, input to the discriminator D. The input will not only make the generated image more realistic, but will also require the generated image to match the structural line drawings, thus ensuring its controllability. In training this discriminator, only real images and their corresponding structural line drawings are considered as positive examples, so that higher resolution 128×128×3 images can be generated by texture GAN.

III. B. 4) Joint Modeling of Texture GAN and Structure GAN

After training the structural GAN and the texture GAN separately, this time, all network layers were also combined for training and learning. Images were generated with uniformly distributed noise as input and structural line drawings were generated from the structural GAN and then the final image was generated based on the texture GAN. As a result, the loss from the texture GAN is also passed to the structure GAN during the joint learning period. in addition it should be noted that the generated structure line graph needs to pass through the sampling layer with bilinear interpolation before it is fed to the texture GAN. The discriminator in the texture GAN treats the generated pattern line drawings and images as negative samples and the real structure line drawings and real images as positive samples.

For the structural GAN, the generator receives not only gradients from its own discriminator, but also gradients passed through the texture GAN generator. This kind of joint learning network can generate both realistic structural line drawings and better final images. The resulting generator network loss function for structural GAN can be expressed as:

$$L_{joint}^{G}(\hat{z},\tilde{z}) = L^{G}(\hat{z}) + \lambda \cdot L_{cond}^{G}(G(\hat{z}),\tilde{z}) \tag{11}$$

where $\hat{z}$ and $\tilde{z}$ represent two sets of uniformly distributed samples in structural GAN and texture GAN, respectively. The first term on the right side of the equation represents the adversarial loss of the structural GAN discriminator, while the second term on the right side of the equation represents the loss of the texture GAN. The value of λ is a hyperparameter, which in this experiment is set to 0.1, which is smaller than the learning rate of the texture GAN, thus avoiding the occurrence of overfitting.

III. C. Experimental results and analysis

III. C. 1) Experimental setup

(1) Dataset

The dataset contains 700 images of oil paintings collected from the web.In this chapter, the pixel image dataset is identified as $P=\{p_{i,}\, i=1,....,N\}$, identifying the cartoon image dataset as $C=\{c_{i,}\, j=1,...,M\}$. In order to train differentiation of texture features using a texture discriminator, this chapter generates a dataset of pixel images with inconspicuous texture styles for all pixel images in P $p_{i,}$ by the method described in $E=\{e_{i,}\, i=1,...,N\}$; In order to use the color discriminator to differentiate the training of color features, this chapter generates a Gaussian blurred cartoon image dataset using the conventional method for all cartoon images in C $c_{i,}$ $X=\{x_{i,}\, i=1,...,M\}$. All datasets are RGB images of size 256×256.

(2) Training Details

The experiments in this chapter are also conducted on an ubuntu 16.04 LTS Quadro P6000 GPU system, and the proposed network models for both texture discriminator and color discriminator are implemented based on Pytorch. The Adam optimization algorithm is used for the training process, and the initial learning rate is set to 0.0002, β_1 is set to 0.5, and β_2 is set to the default value of 0.999.

III. C. 2) Quantitative analysis

A quantitative analysis was performed on the oil painting dataset to compare the evaluation metrics of this paper's method with those of other models on the art image restoration task. We randomly selected a mask from the test set to occlude the test set images and then restored them using the trained model. Among the image restoration models include PConv, RFR, PIC, PD-GAN & RFR. We also tested the metrics of the two translation models, Pix2Pix and Pix2PixHD, on this task, and their metrics are shown in Table 1 and Table 2, respectively. The best and second best evaluation metrics are shown in bold and underlined fonts, respectively. Meanwhile, $\downarrow$ indicates that the smaller the index is, the better it is, and $\uparrow$ indicates that the larger the index is, the better it is.

It can be intuitively seen that this paper's method has obvious advantages over other repair or translation algorithms in quantitative indexes, and this paper's method is optimal compared to the rest of the algorithms in FID and CFID, and suboptimal and very close to the optimal algorithms in PSNR and SSIM. On FID, the improvement is 7.024 compared to the suboptimal algorithm Pix2PixHD, and on CFID, the improvement is 6.38 compared to the suboptimal algorithm PD-GAN.

On the oil painting dataset, this method is optimal in one metric and suboptimal in three metrics. On SIFID, the method achieves an improvement of 0.006 compared to the suboptimal algorithm PD-GAN, and on FID, the difference is only 1.335 compared to the optimal algorithm RFR.

Table 1: Data gathering maps are like the experimental indicators of repair

	FID ↓	CFID ↓	SIFID ↓	PSNR ↑	SSIM ↑
PConv	57.425	61.157	0.039	20.214	0.811
PIC	49.719	50.764	0.029	18.537	0.853
RFR	110.074	115.834	0.062	18.672	0.811
PD-GAN	46.842	47.291	0.025	8.143	0.437
Pix2Pix	61.573	62.753	0.051	15.409	0.739
Pix2PixHD	46.628	48.192	0.051	19.154	0.822
This algorithm	39.604	40.911	0.031	19.375	0.827

Table 2: Experimental index of painting image repair

	FID ↓	CFID ↓	SIFID ↓	PSNR ↑	SSIM ↑
PConv	107.141	103.814	0.057	22.194	0.774
PIC	129.207	132.381	0.197	19.573	0.732
RFR	85.791	82.573	0.045	19.794	0.785
PD-GAN	92.312	96.517	0.033	20.096	0.684
This algorithm	84.436	84.886	0.027	20.151	0.701

III. C. 3) Comparative analysis of experiments

This section focuses on comparing the color information of the methods in this paper and PD-GAN, because these two methods can generate texture features with more oil painting style compared to other methods. So this section further explores the learning ability of different models for color features when the texture style has been achieved. In this section, both color histograms and pixel RGB values are used to compare the color information of the input dataset and the generated oil paintings.

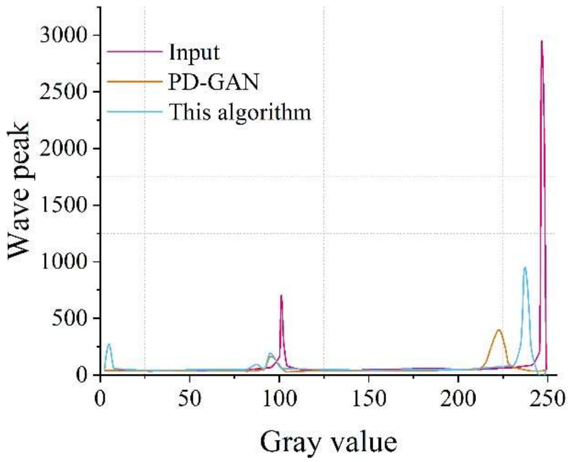
For further comparison, the corresponding RGB values are represented as parts 1,2,3,4 and statistics are performed as shown in Table 3, and the mean of Manhattan distance [16] is used to represent the color differences. Where r,g,b is the gray value, of the R,G,B channels.

From the table, it can be seen that the RGB values of both images of DPIGAN are closer to the RGB values of the input image, and the Manhattan distance obtained from the calculation is much smaller than that of PD-GAN. This indicates that the method in this paper does a much better job in preserving the color information of the image, has a much better ability to learn the color features, and produces oil painting images that are better than PD-GAN in terms of visual quality.

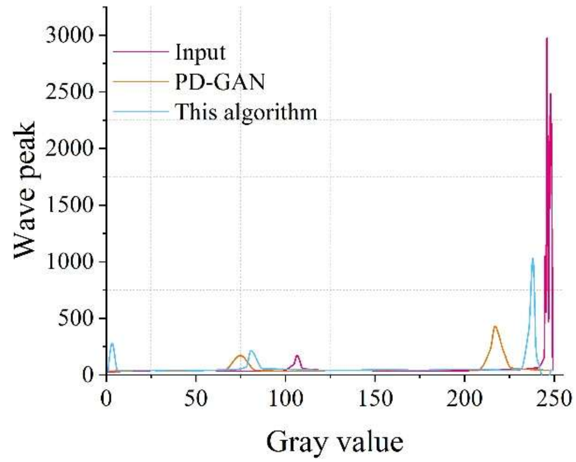
Table 3: Comparison method RGB

Contrast image	Part 1	Part 2	Part 3	Part 4	Manhattan distance
Enter image 1	(251,234,152)	(250,213,212)	(255,255,255)	(145,81,45)	-
PD-GAN	(229,193,150)	(221,187,195)	(222,219,225)	(121,91,72)	75.33
This algorithm	(251,224,180)	(247,212,208)	(237,248,250)	(140,101,77)	36.19
Enter image 2	(241,163,185)	(103,109,125)	(255,255,255)	(221,164,175)	-
PD-GAN	(205,155,164)	(93,77,87)	(226,225,226)	(193,154,152)	79.12
This algorithm	(225,170,172)	(92,80,88)	(241,241,242)	(221,167,167)	40.23

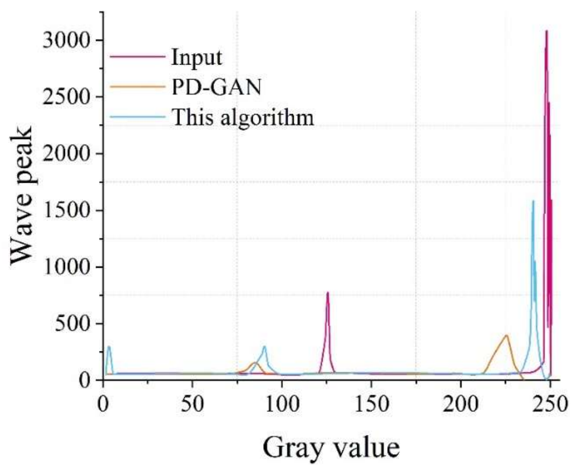
Figure 1 shows the color histogram comparison of the input image. As the color histogram represents the proportion of different colors in the whole image. From the above figure, it can be seen that in the three channels R, G and B, the peak part of the histogram of this paper's method is closer to the peak position of the oil painting image, which represents that the gray value of most of the pixels of this paper's method in the three channels is closer to the gray value of the oil painting image and the image brightness is more consistent. So this paper's method performs better than PD-GAN in the retention of color features, which proves the effectiveness of setting color discriminator in this paper's method.



(a) R channel



(b) B channel



(c) C channel

Figure 1: The RGB three channel color histogram comparison diagram

IV. Analysis of the visual effect of oil painting picture hierarchy

The subjects were university students in different majors (vehicle, mechanical design, finance, robotics, web media design, auditing, industrial design, etc.), 41 (23 female, 18 male) with an average age of 22 years old, were recruited to participate in this study. A questionnaire was designed for the participants in the experiment, according to the population to be tested to answer the questions, and then synthesize the professional, gender ratio, art experience and other conditions of the final confirmation of the 40 experimental personnel, which will be divided into three different groups of professional group, learning group, non-professional group, each group are 20 personnel, in order to ensure that statistically significant comprehensiveness and typicality.

In the questionnaire screened by the experimental population, intuition and a priori knowledge factors were included in the questionnaire as criteria influencing aesthetic judgment. Participants were asked whether they had seen the paintings, i.e., whether they had studied or attended an art history course, to distinguish whether they had a priori knowledge of oil painting as a means of dividing the professional and non-professional groups.

Each person scored each picture on a Likert scale for overall aesthetics, color fullness, richness of content, and compositional momentum, and then ranked the scores to obtain their respective mean, variance, and standard deviation data, and analyzed the scoring data using univariate means and variances to determine the complexity of each picture, arbitrarily truncating the data of three oil paintings, for example, as shown in Table 4.

Overall, low complexity M = 4.12~5.23, DS = 1.88~2.62; medium complexity M = 5.53~5.51 low complexity, DS = 1.12~1.10; high complexity M = 5.84~6.2, DS = 0.87~1.10; the three groups of values are positively correlated, and the main effect is obvious, which can show that aesthetic complexity is correlated with the evaluation index; And it is normally distributed, which can show the significance of complexity evaluation.

Table 4: Aesthetic complexity score data

Evaluation index	Oil painting 1			Oil painting 2			Oil painting 3		
	Mean	Variance	Standard Deviation	Mean	Variance	Standard Deviation	Mean	Variance	Standard Deviation
Overall aesthetic feeling	4.975	1.875	1.473	5.792	1.184	1.053	5.929	1.176	1.155
Color full	4.122	2.615	1.612	5.533	1.192	1.125	5.841	1.233	1.247
Content richness	4.902	2.341	1.527	5.533	1.371	1.166	6.2	1.2	0.873
Composition	5.234	1.9	1.355	5.514	1.202	1.102	5.973	1.147	1.103

V. Conclusion

In this study, a breakthrough in intelligent oil painting generation technology is successfully achieved by constructing a hierarchical guidance design method for oil paintings based on the probability distribution of visual attention. Experimental validation shows that the method outperforms the existing mainstream algorithms in a number of performance indexes, in which the SIFID index is improved by 0.006 compared with the suboptimal algorithm PD-GAN, which shows a good image quality advantage. The results of visual effect analysis show that the evaluation data of 40 subjects present normal distribution characteristics, with the average ratings of low-complexity works ranging from 4.12 to 5.23, and the average ratings of high-complexity works reaching 5.84 to 6.2, which proves the effectiveness of the method at different levels of complexity. The comparative analysis of color features further verifies the superiority of the method, and the Manhattan distance is significantly reduced in the comparison of RGB values, and the distance value of the second group of test images is only 40.23, which is much lower than that of the comparative method, which is 79.12.In terms of technological innovations, the combination of the Markovian attention-shifting model and the layered GAN architecture provides a new theoretical framework for the intelligent generation of oil paintings, and the joint hyperparameter set to 0.1 learning mechanism effectively avoids the overfitting problem. This research not only provides practical technical tools for the field of digital art creation, but also opens up a new research direction for the application of computer vision in the field of art, which is of great significance for promoting the deep integration of artificial intelligence and traditional art.

Funding

This work was supported by Anhui Province Philosophy and Social Sciences Planning Project: “Regional Characteristics of ‘New Anhui School Oil Painting’ and Its Aesthetic Research” (AHSKY2020D112).

References

[1] Zhang, Y. (2017, May). Personal Emotional Expression in the Creation of Oil Painting. In 2017 International Conference on Culture, Education and Financial Development of Modern Society (ICCESE 2017) (pp. 530-532). Atlantis Press.
[2] Li, Y. (2023). Research on the Influence of the Aesthetics of Traditional Chinese Painting on Contemporary Oil Painting Landscape Creation. Frontiers in Art Research, 5(8).
[3] Zeng, Y. (2021, December). The Art of Oil Paintings Creation with Computer Multimedia Technology. In 2021 International Conference on Forthcoming Networks and Sustainability in AIoT Era (FoNeS-AIoT) (pp. 238-243). IEEE.
[4] Wang, C. (2020, October). The Aestheticization of Oil Painting Teaching and the Construction of Multi-Dimensional Oil Painting Teaching System Based on Network Cloud Platform. In Journal of Physics: Conference Series (Vol. 1648, No. 2, p. 022018). IOP Publishing.
[5] Zhang, Y. (2017, June). Function and Manifestation of Decorative Color in Oil Painting Creation. In 2nd International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2017) (pp. 487-489). Atlantis Press.
[6] Xu, Y. (2016, May). The emotion analysis during oil painting creation process. In 2nd International Conference on Arts, Design and Contemporary Education (pp. 37-40). Atlantis Press.
[7] Shang, L., & Liu, X. (2017, April). Cross Boundary Phenomenon Analysis: Printmaking Artists' Creation of Oil Paintings. In 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017) (pp. 391-394). Atlantis Press.
[8] Shi, Y., & Tsetsegdelger, D. (2024). The Road to Change in Oil Painting Creation and Appreciation in the Age of Artificial Intelligence. International Journal of Finance and Investment, 1(1), 50-53.
[9] Qi, J. (2024). Research Report of the Creative Inspiration of Oil Painting Landscape in Northwest China. Art and Society, 3(3), 13-25.
[10] Yihan, H. (2024). A Probe into The Modernity of Chinese Oil Painting. International Journal of Advanced Culture Technology, 12(3), 199-205.
[11] Huang, F., Wu, B. H., & Huang, B. R. (2015). Synthesis of oil-style paintings. In Image and Video Technology (pp. 15-26). Cham: Springer International Publishing.
[12] Wang, W., & Shen, J. (2017). Deep visual attention prediction. IEEE Transactions on Image Processing, 27(5), 2368-2378.
[13] Yantis, S. (2016). Control of visual attention. In attention (pp. 223-256). Psychology Press.

[14] Mnih, V., Heess, N., Graves, A., & Kavukcuoglu, K. (2014). Recurrent models of visual attention. Advances in neural information processing systems, 27.

[15] Marziyeh Karimzadeh,Mohsen Alambardar Meybodi,Mohammadreza Shams & Zahra Goli Malekabadi. (2025). Microfluidic droplet detection using synthetic data generated by a Generative Adversarial Network. Engineering Applications of Artificial Intelligence,156(PA),111032-111032.

[16] Xiaowei Wang,Yang Yue,Fan Zhang,Yizhao Wang & Zhihua Zhang. (2025). Active Detection of Interphase Faults in Distribution Networks Based on Energy Relative Entropy and Manhattan Distance. Electric Power Systems Research,241,111397-111397.