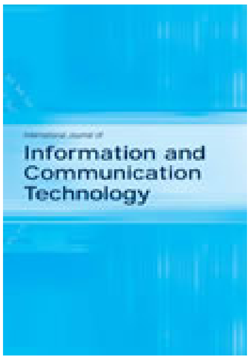


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Innovative design of ceramic products based on extension reasoning and generative adversarial networks

Zhijun Meng, Leiming Zhu, Gao Chen

DOI: 10.1504/IJICT.2025.10073895

Article History:

Received:	19 July 2025
Last revised:	01 September 2025
Accepted:	02 September 2025
Published online:	22 October 2025

Innovative design of ceramic products based on extension reasoning and generative adversarial networks

Zhijun Meng

Faculty of Humanities and Arts,
Macau University of Science and Technology,
Macao 999078, China
Email: mengzhijun@sina.com

Leiming Zhu*

Design College,
Wenzhou Polytechnic,
Wenzhou 325035, China
Email: ramiaweizhu@163.com
*Corresponding author

Gao Chen

Management Department,
University of Otago,
Dunedin 9016, New Zealand
Email: chech380@student.otago.ac.nz

Abstract: This paper proposes a novel framework integrating extension reasoning with generative adversarial networks to address rigid cultural element expression in ceramic product design. The methodology comprises three innovations: 1) a primitive model analysing object-element, relation-element, and feature-element for topological pattern transformation; 2) multimodal fusion of geometric features, chromatic attributes quantified by GIE-Lab values, and semantic associations; 3) a lightweight network optimised with leaky rectified linear unit activation and Adam optimiser, achieving 30% faster convergence than stochastic gradient descent. Experiments on 3,262 Jingdezhen ceramic images demonstrate superior performance with inception score 5.7 and Fréchet inception distance 2.13, outperforming autoencoder (2.3/5.61), variational autoencoder (3.8/4.54), and deep convolutional generative adversarial network (4.6/5.22). The framework enhances design efficiency by eight times compared to manual processes, validated through Kano model analysis with 89% user satisfaction. This work establishes a technical paradigm for intelligent modernisation of traditional craftsmanship.

Keywords: extension reasoning; generative adversarial network; ceramic cultural and creative product design.

Reference to this paper should be made as follows: Meng, Z., Zhu, L. and Chen, G. (2025) 'Innovative design of ceramic products based on extension reasoning and generative adversarial networks', *Int. J. Information and Communication Technology*, Vol. 26, No. 38, pp.75–90.

Biographical notes: Zhijun Meng is a PhD candidate in Digital Media at the Faculty of Humanities and Arts, Macau University of Science and Technology. His research interests include primarily on visual communication and digital media.

Leiming Zhu received his Master degree from Wuhan University of Technology in 2009. He is currently working in Wenzhou Polytechnic, as the Head of Visual Communication Major. His research interests include product information visualisation.

Gao Chen received his Master degree from University of Otago in 2017. Currently, he is studying as a PhD student at the University of Otago. His research interests include business model innovation, value co-creation theory, operations management, and quality process improvement.

1 Introduction

Ceramic cultural and creative works play an important role in people's lives and are a powerful carrier of cultural heritage. More and more designers are beginning to pay attention to the design of ceramic cultural and creative works. Due to the lack of technology in the past, the design of ceramic cultural and creative works only relied on experience and subjective extraction of cultural elements. The accuracy and fit of the extraction were greatly deviated from the actual situation (Wei and Cheng, 2022; Huang, 2024). There are 'element stacking' and 'modelling homogeneity' phenomena, lack of innovation, which give consumers a stereotype about the design of products, affecting the sales and popularity of ceramic cultural and creative products (Sun et al., 2021). In addition, it often takes a lot of time and energy for designers to design a work, which affects the efficiency of ceramic cultural (He, 2022).

In recent years, more and more computer technologies have been integrated into the design of ceramic cultural and creative works. People have begun to seek objective and diverse creative methods to present more and higher-quality works faster. The study pointed out that cultural and creative development can not only meet the needs of mass cultural consumption, but also alleviate the economic pressure of museums, but the problem of homogeneity needs to be solved. Through the case of the Forbidden City in Beijing, a strategy of using digital technology (such as AR) to enhance interactive experience is proposed, emphasising the balance between cultural connotation and practical design. It is proposed to activate the innovation ecology by establishing a ceramic design database and holding international competitions. The case shows that the 'cultural IP + modern design' model can increase the added value of products by 40%. At the same time, it is pointed out that the weak linkage between industry, academia and research is a key factor restricting the cultivation of creative talents. Ren (2022) developed a ceramic pattern generation system based on the GAN network. Experiments have shown that its design efficiency is 8 times higher than that of manual work.

Combined with the Kano model, a user preference prediction system is constructed, which increases product satisfaction from 62% to 89%. The study also explores the copyright issues of AI-generated content and recommends the use of blockchain technology to achieve design traceability. He and Tao (2025) developed an AI-assisted pattern innovation system for the intangible cultural heritage of blue printed cloth. Through deep learning, the semantics of traditional patterns are analysed to generate variant designs that conform to modern aesthetics. A digital process parameter library is established to increase production efficiency by 50% while retaining the characteristics of handmade texture. The study verifies the feasibility of AI in the living inheritance of intangible cultural heritage.

GAN achieves data generation through adversarial training between generators and discriminators without explicitly modelling probability distribution (Goodfellow et al., 2020). The core contributions include: proposing a minimax game framework, proving the existence of Nash equilibrium, and demonstrating the generation effect of MNIST and CIFAR-10. Its indirect modelling idea has inspired a large number of subsequent generative model studies. Liu et al. (2022) proposed to improve the DCGAN model by improving the generation quality through four key technologies: replacing the pooling layer with strided convolution, removing the fully connected layer, introducing batch normalisation, and optimising the activation function (ReLU/Tanh for the generator and leaky ReLU for the discriminator). Experiments show that it has significant effects in facial arithmetic operations and image completion tasks, and has become the basic template for subsequent CNN architecture GAN. When the data distribution is relatively simple, GAN, with its lighter framework compared to DCGAN, offers more efficient training. Furthermore, when the sample size is relatively small, GAN can avoid the overfitting issue of DCGAN. Mi et al. (2023) proposed a method combining Wasserstein distance and confidence loss for small sample scenarios. Wasserstein distance solves the gradient vanishing problem, while confidence loss optimises data screening by quantifying the reliability of generated samples. Experiments have shown that it can stably generate high-quality samples in scarce data fields such as medical imaging, and the FID index is better than traditional GAN. When the overlap between the true distribution and the generated distribution is small, the gradient of the Wasserstein distance may still be small, affecting the training effect. Small sample data may lead to inaccurate estimation of the true distribution, which in turn affects the effectiveness of the Wasserstein distance. In the case of extremely limited data, mode collapse may still occur, resulting in a lack of diversity in generated samples. In addition, Elasri et al. (2022) proposed a review paper that systematically sorted out the development of image generation technology based on deep learning. It focuses on comparing the advantages and disadvantages of paradigms such as GAN, VAE, and autoregressive models, analysing common challenges such as mode collapse and training instability, and summarising the application scenarios of evaluation indicators (such as IS and FID). Finally, it looks forward to future directions such as multimodal generation and controllable generation (Li et al., 2025).

In the field of image generation, Liu et al. (2024) proposed a text-to-image generation framework guided by prior knowledge, which improves the controllability of generation by incorporating domain knowledge (such as object relationship constraints or semantic rules). Typical applications include scene composition optimisation and fine-grained attribute matching. On the COCO dataset, the semantic consistency between text description and generated image is improved by about 15%. Lu et al. (2024) achieves multi-style tile image generation through classifier guidance. The pre-trained classifier is used to extract style features, and the conditional GAN is combined to control visual elements such as texture and ‘colour. This method verifies the generation diversity in the field of interior design, and supports users to interactively adjust the style strength through a slider. Cheema and Naeem (2025) propose a layout-guided GAN architecture for the task of generating social media cover. The first stage generates a semantic layout graph through text parsing, and the second stage synthesises visual elements based on the layout. User studies show that its generation results are 20% higher than the baseline model in terms of topic relevance and aesthetic score. Chen and Zhao (2025) design a multi-layer fusion mechanism to optimise the efficiency of text-image alignment. A cross-modal attention layer is introduced between the encoder and decoder to dynamically aggregate text features of different granularities. Compared with StackGAN, the inference speed is increased by 30%, and the SSIM score of SOTA is achieved on the CUB bird dataset. Yang et al. (2025) achieved fine-grained text control through a semantic fusion module. The key innovation is to construct a hierarchical semantic tree to map grammatical components such as noun phrases and adjectives to different generation stages.

In view of the above research, there are still the following two defects, one is the input of the generation model uses random noise and does not have element characteristics, and another is the model is highly complex, difficult to train, and not universal. This study proposes to combine extension reasoning and generative adversarial networks and apply them to the design of ceramic cultural and creative products. Through primitive modelling, the explicit and implicit characteristics of ceramic cultural symbols are analysed, and a three-dimensional model of ‘object element-relation element-feature element’ is established to achieve topological transformation and innovative reorganisation of traditional patterns. After that, a generator with strong robustness and stable performance is obtained by training the generative adversarial network.

2 Relevant technologies

2.1 Extension reasoning

Extension reasoning is one of the core methodologies of extension science. Its essence is to study the transformation mechanism of contradictory problems through formal models. It constructs a quantitative model of contradictory problems based on primitive theory and uses extension transformation to realise logical deduction from contradictory state to compatible state. Different from conventional reasoning, its significant feature is the effectiveness of reasoning under the premise of dealing with contradictions. It is defined as ‘studying the logic of transforming contradictory problems into non-contradictory problems’.

The research object of extension science is the contradictory problems in the objective world. The so-called contradictory problem refers to the problem that the purpose people want to achieve cannot be achieved under the existing conditions. In the process of research, researchers in extension science found that in many engineering fields, such as management, control, computer technology, artificial intelligence, machinery, electrical engineering, etc., various contradictory problems will be encountered. The starting point of extension science research is to establish a set of methods to deal with contradictory problems based on the laws and theories of contradictory problems. Computers have the advantages of large storage capacity and fast speed. Therefore, studying how to express problems in formal language, describe the purpose and conditions of problems, establish a set of reasoning methods, and finally, let computers help people propose strategies to solve contradictory problems is the destination of extension science. Mathematical models can handle a large number of problems that require precision, but they cannot handle problems where the goals and conditions are incompatible. The reasons are:

1. When solving contradiction problems, the object itself and the characteristics of the object is also be considered
2. The transformation to solve the contradiction problem has both quantitative and qualitative parts
3. Classical mathematics studies deterministic things, while solving contradiction problems requires considering the transformation of things (including quantitative and qualitative changes).

Therefore, mathematical models are difficult to describe the process of solving contradiction problems.

In order to use formal methods to deal with various contradiction problems in the objective world, we must first study how to describe the various things in the objective world. To this end, topology theory has established matter elements, event elements and relation elements, which are the logical cells of topology. Among them, the matter element model supports the quantitative expression of user needs in design problems. The event element model describes events through verbs and feature tuples, supporting similar extension reasoning of function trees. The relation element describes the association characteristics between objects and is used to construct a transmission network for contradiction problems. The formalised model of primitives describing information, knowledge, intelligence and various problems is called an extension model. With an extension model, we can propose strategies to solve various contradictory problems based on the extension of primitives, using extension theory and extension methods.

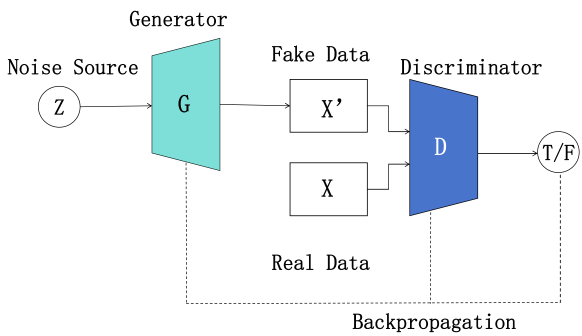
Extension reasoning breaks through the boundary rigidity of classical sets and dynamically describes the process of quantitative change to qualitative change through correlation functions. In case reasoning, distance calculation is used to achieve high-dimensional space similarity measurement and support interval domain retrieval. For the extension representation of cultural elements, it can be carried out from two aspects: semiotic hierarchical modelling and transformation of contradictory problems. For example, the theory of semiotic hierarchical modelling can be used to extract surface (morphology/'colour) and deep (auspicious meaning) material elements from ceramic patterns to establish a two-dimensional design knowledge base. Using the theory of transformation of contradictory problems, the extension set theory is used to solve the contradiction of 'modernisation of traditional patterns', such as realising the contemporary translation of pattern composition rules through conduction transformation.

At present, extension models have been widely used in management, control, artificial intelligence, etc. In this paper, based on extension models, the object-element, event-element, and relationship-element of ceramic cultural creation are analysed to establish a primitive model, which is used for subsequent training of generative adversarial networks to generate innovative ceramic cultural creations.

2.2 GAN

GAN is excellent model used in the image generation. The generator and the adversary repeatedly compete to generate a more ideal picture.

Figure 1 GAN model (see online version for colours)



In Figure 1, When processing image data, the discriminator usually uses convolutional layers or other architectures suitable for the data type. The optimisation goal of the GAN is a minimax game:

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

where z represents random noise, usually conforming to a certain distribution (such as Gaussian distribution), and x is the real data.

(3) $D(x) \in [0, 1]$ is the discriminator function, which represents the probability that the output sample is true.

For expectation (E):

$$E(x) = \int_{-\infty}^{+\infty} xp(x)dx \tag{2}$$

Therefore

$$E_{x \sim P_{data}(x)}[\log D(x)] = \int_{-\infty}^{+\infty} P_{data}(x) \log D(x)dx = \int_x P_{data}(x) \log D(x)dx \tag{3}$$

$$\begin{aligned} E_{z \sim P_z(z)}[\log(1 - D(G(z)))] &= \int_{-\infty}^{+\infty} E_{Z \sim p_z(z)}[\log(1 - D(G(z)))]dz \\ &= \int_x P_g(x) \log(1 - D(x))dx \end{aligned} \tag{4}$$

In summary, the adversarial function of the GAN can be expressed as:

$$\begin{aligned} V(D,G) &= \int_x P_{data}(x)\log(D(x))dx + \int_x P_g(x)\log(1-D(x))dx \\ &= \int_x P_{data}(x)\log(D(x)) + P_g(x)\log(1-D(x))dx \end{aligned} \tag{5}$$

For the *D*, the D's goal needs to be maximised, that is:

$$\max_D V(D) = \int_x P_{data}(x)\log(D(x)) + P_g(x)\log(1-D(x))dx \tag{6}$$

Draw the derivative of $D(x)$ and set the derivative to 0:

$$\frac{\partial V(D)}{\partial D(x)} = \frac{P_{data}(x)}{D(x)} - \frac{P_g(x)}{1-D(x)}a = 0 \tag{7}$$

The solution of the optimal discriminator is:

$$D^*(x) = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)} \tag{8}$$

Substitute into the original objective function:

$$\begin{aligned} \min_G V(G) &= \int_x P_{data}(x)\left(\log\left(\frac{2P_{data}(x)}{P_{data}(x)+P_g(x)}\right) - \log 2\right)dx + \int_x P_g(x)\left(\log\left(\frac{2P_g(x)}{P_{data}(x)+P_g(x)}\right) - \log 2\right)dx \\ &= \text{-}\log 2\int_{\mathrm{x}} P_{data}(x) + P_g(x)dx + \int_x P_{data}(x)\log\left(\frac{2P_{data}(x)}{P_{data}(x)+P_g(x)}\right)dx + \int_x P_g(x)\log\left(\frac{2P_g(x)}{P_{data}(x)+P_g(x)}\right)dx \end{aligned} \tag{9}$$

where

$$\text{-}\log 2\int_{\mathrm{x}} P_{data} + P_g dx = -2\log 2 = -\log 4 \tag{10}$$

JS divergence is defined as:

$$JSD(P \,\|\, Q) = \frac{1}{2}D_{KL}\left(P \,\|\, \frac{P+Q}{2}\right) + \frac{1}{2}D_{KL}\left(Q \,\|\, \frac{P+Q}{2}\right) \tag{11}$$

where

$$D_{KL}(P \,\|\, Q) = \sum_i P(i)\log\left(\frac{P(i)}{Q(i)}\right) = \int_x P(x)\log\left(\frac{P(x)}{Q(x)}\right)dx \tag{12}$$

Use *JS* divergence to replace the formula $\min_G \mathrm{V}(G)$ to obtain:

$$\min_G \mathrm{V}(G) = -\log 4 + 2 * JSD\left(P_{data} \,\|\, P_g\right) \tag{13}$$

2.3 The training process of GAN

The training steps have two stages:

1. Fix the discriminator and train the generator. When the training continues, the accuracy of the generator continues to rise, and eventually it deceives the

discriminator. At this time, the discriminator is basically in a blind guessing state, and the probability of judging whether it is fake data is 50%.

2 Fix the generator and train the discriminator. Through continuous training, it can accurately judge fake data.

Repeat the first and second stages, an optimal generator is obtained, which can be used to generate target data. The above process can be expressed by the following pseudo code:

The generated most robust network is trained using stochastic gradient descent steps. Where k is the number of steps of the D, and k =1 is used during the training process

for number of training iterations do

for k steps do

- Sample minibatch of m noise samples $\{z^{(1)}, \ldots, z^{(m)}\}$ from noise prior $P_g(z)$
- Sample minibatch of m samples $\{x^{(1)}, \ldots, x^{(m)}\}$ from data generating distribution $P_{data}(x)$
- Update the discrimination by ascending its stochastic gradient:

$$\nabla\theta_d \frac{1}{m}\sum_{i=1}^{m}\left[\log D\left(x^{(i)}\right)+\log\left(1-D\left(G\left(z^{(i)}\right)\right)\right)\right]$$

end for

- Sample minibatch of m noise samples $\{z^{(1)}, \ldots, z^{(m)}\}$ from noise prior $P_g(z)$
- Update the generator by descending its stochastic gradient:

$$\nabla\theta_g \frac{1}{m}\sum_{i=1}^{m}\log\left(1-D\left(G\left(z^{(i)}\right)\right)\right)$$

end for

In addition to the momentum method used in this process, there are other ways to update the gradient. During the training process, the optimal parameter update method can be found through comparison

---

GAN are considered to be one of the most promising and active models currently. They are currently mainly used in sample data generation, image generation, image restoration, image conversion, text generation, and other fields.

3 Introduction to Ex-GAN

3.1 Ex-GAN model

Ex-GAN is based on extension reasoning and GAN.

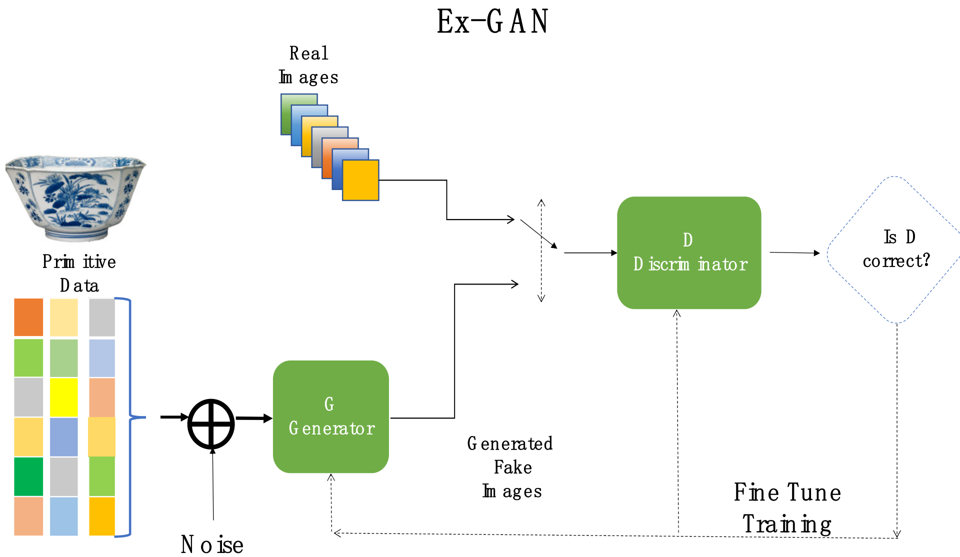
In Figure 2, the primitive model data mainly includes three parts:

1 Pattern gene deconstruction. Extract the geometric features, topological relationships, and process constraints of ceramics

2 ‘colour primitive mapping. Analyse the blue and white spectrum and use GIE-Lab values for quantification

3 Semantic association modelling. Establish a pattern-semantic mapping table.

Figure 2 Overall flow chart of the model (see online version for colours)



Cultural semantics are defined by triples (N, c, v). For example, the modelling example of the blue and white lotus pattern is as follows:

- Pattern gene deconstruction: pattern topological structure rules
- ‘colour primitive mapping: glaze chemical composition ratio
- Semantic association modelling: blue and white ‘lotus’ metaphor symbol library.

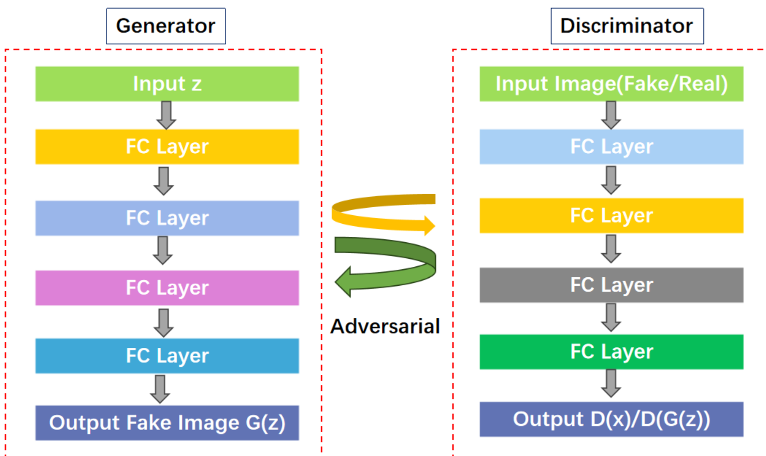
Multimodal fusion of object elements, relationship elements and feature elements, and the fused multi-features are normalised and noise data is added to avoid model overfitting. To ensure the convergence of the model, standardisation is performed before inputting it. In this paper, linear transformation is adopted to standardise the input ceramic photos. After that, the processed data is converted into a vector as the input of the G, and then the G outputs a fake picture of a cultural and creative work as the input of the discriminator. The fake picture of a cultural and creative work and the real picture are used as the input of the D respectively. The D calculates the probability of a real picture of a cultural and creative work and the probability of a generated picture respectively. By continuously optimising the loss value and adjusting the model parameters, the loss value is gradually converged. The accuracy and AUC value of the discriminator are calculated, and finally the optimal Ex-GAN is obtained.

3.2 GAN architecture

In Ex-GAN, a lightweight GAN is used, only a fully connected neural network is used as the basic network architecture, as shown in Figure 3.

The generative adversarial network model used in this paper consists of two parts: the generator and the adversary. The generator includes a four-layer fully connected neural network, the input is the multimodal feature fusion data of the primitive model, and the output is the image generated by the generator. The discriminator includes a four-layer fully connected neural network, the input is the real image or the image generated by the generator, and the output is the probability of judging it as a real image or a generator image. In the subsequent training, the effects of different functions on the model are fully compared, and finally Leaky ReLU is used. Through the game training of the generator and the adversary, a generative model with excellent performance is finally obtained, which is used to generate ceramic cultural and creative works with innovative significance.

Figure 3 GAN model architecture (see online version for colours)



4 Experimental results and analyses

4.1 Evaluation indicators

FID is used to measure the difference between the distribution of generated images and real images. It can be expressed as:

$$FID = \| \mu_r - \mu_g \|^2 + Tr\left(\sum\nolimits_r + \sum\nolimits_g - 2\left(\sum\nolimits_r \sum\nolimits_g\right)^{1/2}\right) \tag{14}$$

$-\mu_r, \sum_r$ The mean and covariance matrix of the real image features.

$-\mu_g, \sum_g$ The mean and covariance matrix of the generated image features.

$-T_r$ The trace of the matrix.

The calculation process is as follows:

Step1 Generate a batch of images of size N from the generator.

Step2 Use the pre-trained inception network to calculate the feature vector of each image.

Step3 Calculate the mean and covariance matrix of the feature vector of the generated image.

Step4 Calculate the FID.

It can reflect the distribution difference. By comparing the feature distribution, it is resistant to mode collapse. FID can detect whether the generated image covers all the modes of the real data distribution. In addition, it is sensitive to perceptual quality: the details and authenticity of the generated image directly affect the FID score. However, it has the following limitations:

1 Sensitive to noise: a small amount of noise in the generated image may significantly affect the FID score.

2 High computational cost: it is necessary to calculate the mean and covariance matrix.

IS can be expressed as:

$$IS = \exp\left(E_x\left[D_{KL}\left(p(y \mid x) \parallel p(y)\right)\right]\right) \tag{15}$$

The KL divergence, which is used to measure the difference between $p(y|x)$ and $p(y)$, and $p(y)$ represents the overall category distribution of the generated image, that is:

$$p(y) = \frac{1}{N}\sum_{i=1}^{N} p\left(y \mid x_i\right) \tag{16}$$

The calculation process is as follows:

Step 1 Generate a batch of images of size N from the generator.

Step 2 Use the pre-trained Inception network to calculate the category probability vector of each image.

Step 3 Calculate the average probability of each category over all images, and then calculate the entropy.

Step 4 Take the logarithm and calculate the expected value to get the final Inception Score.

The quality and diversity of images generated by GAN are positively correlated with IS values. However, it has the following limitations:

1 Ignoring the real data distribution: IS only evaluates the generated image itself and has nothing to do with the real data distribution.

2 Sensitive to category correlation: if the categories of the generated image distribution are unbalanced, the IS score may be too high or too low.

3 Not resistant to mode collapse: even if the generator only generates a few high-quality samples, the IS score may be very high.

In summary, The generation capability of EX-GAN is evaluated using a combination of FID and IS values to avoid the inaccurate evaluation problem caused by the limitations of the evaluation indicators.

4.2 Experimental results

4.2.1 Comparison of FID and IS values with other models

The dataset used in the training process of Ex-GAN comes from the Jingdezhen ceramic element set, which contains 3,262 images. The Ex-GAN proposed in this paper is compared with AE (Pinheiro Cinelli et al., 2021), VAE (Li et al., 2023), and DCGAN. AE is an unsupervised neural network model, which consists of an encoder and a decoder symmetrically. It is mainly used in data dimension reduction, feature extraction and image denoising tasks, but due to limited generation capabilities, the output sample fidelity is usually medium. VAE introduces the idea of probabilistic modelling based on AE. The latent layer of AE needs to be continuous to ensure that the objective function contains the reconstruction loss. This design supports latent space interpolation to generate new samples with smooth transitions, but the characteristics of optimising the lower bound of likelihood may cause the output image to be blurred. DCGAN adopts an adversarial training mechanism: the generator (G) forges images through transposed convolution upsampling, and the discriminator (D) uses convolutional networks. However, adversarial training has the risk of mode collapse (generated samples are single) and insufficient stability, and requires fine tuning of parameters. Although there are no explicit constraints on the latent space, semantic characteristics can be explored through inversion techniques.

Table 1 Comparison results with others.

Models no.	*Models*	*IS*	*FID*
1	AE	2.3	5.61
2	VAE	3.8	4.54
3	DCGAN	4.6	5.22
4	Ex-GAN	5.7	2.36

The IS value of Ex-GAN proposed in this paper can reach 5.7 and the FID is as low as 2.36. Both evaluation indicators are better than other models. This shows that the Ex-GAN has good generation ability and can generate a variety of more realistic ceramic cultural and creative works with innovative performance. The possible reasons are as follows:

1 The generated images are more detailed and realistic, which significantly improves FID and IS. AE/VAE relies on reconstruction loss and is prone to fuzzy output.

2 Semantic guidance of extension reasoning after introducing extension reasoning, the model can decompose and logicise input conditions (such as text or key points) to ensure that the generated content strictly complies with semantic constraints. This mechanism enhances the controllability and coherence of the image, avoids the mode collapse problem of DCGAN, and thus improves the IS score.

4.2.2 Effect of activation function on discriminator

Ex-GAN compares the effects of Sigmoid, Tanh, ReLU and Leaky ReLU activation functions on the model.

$$sig\,\mathrm{mod}(x) = \frac{1}{1+e^{-x}} \tag{17}$$

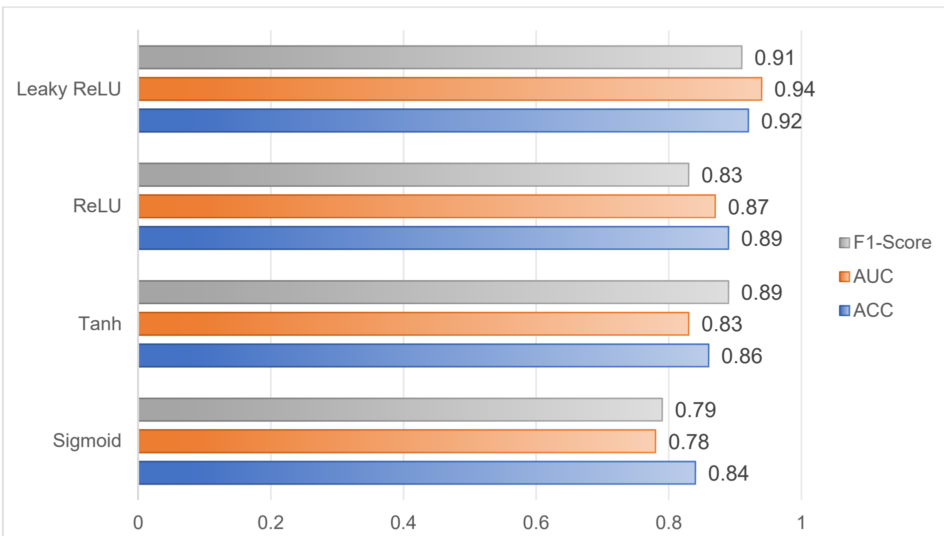
$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{18}$$

$$RELU = \max(0, x) \tag{19}$$

$$\mathrm{Leaky}RELU = \max(0.01\,x, x) \tag{20}$$

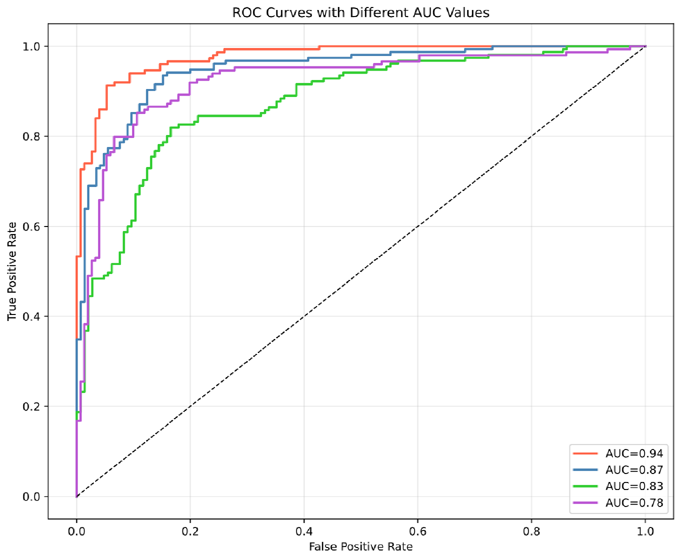
The effects of different activation functions on the discriminator are shown in Figures 4 and 5.

Figure 4 Effect of different activation functions on the discriminator (see online version for colours)



As can be seen from Figures 4 and 5, the discriminator can achieve an ACC of 0.92, AUC of 0.94, and F1-Score of 0.91 using Leaky ReLU, which are better than other activation functions. In the comparative experiment, the reasons why the leaky ReLU activation function performed best can be attributed to three points: First, compared with the standard ReLU, Leaky ReLU allows the gradient in the negative range to be propagated back, improving the efficiency of parameter updating. Secondly, its non-saturation characteristics effectively alleviate the gradient disappearance phenomenon and maintain a more stable gradient flow in deep networks. In addition, its computational complexity is comparable to that of ReLU, without significantly increasing computational overhead. These characteristics make it an ideal choice for balancing performance and efficiency. Therefore, Ex-GAN uses leaky ReLU as the activation function.

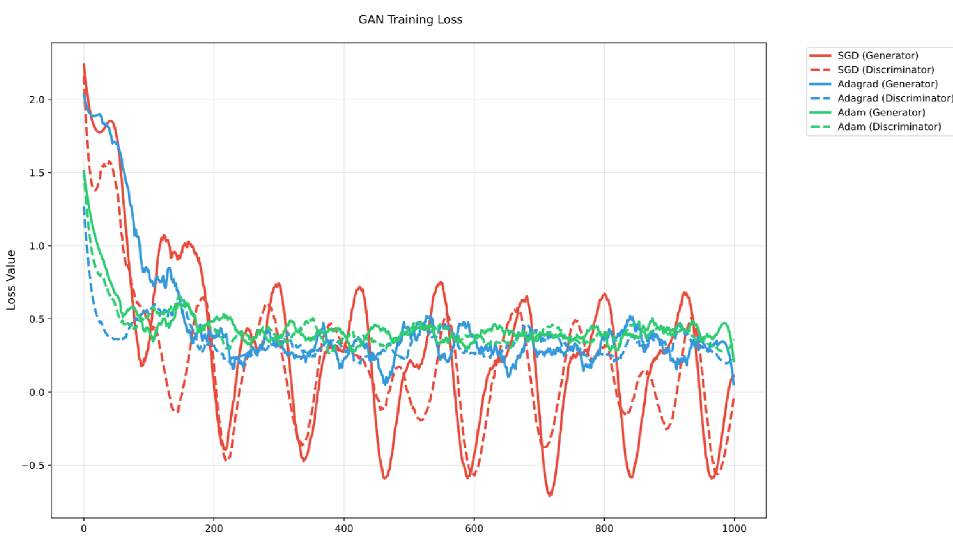
Figure 5 Effect of different activation functions on the discriminator roc curve (see online version for colours)



4.2.3 *Effect of different optimisers on the model loss function*

Ex-GAN compares the effects of different optimisers on the loss function.

Figure 6 Effect of different optimisers on the Ex-GAN loss function (see online version for colours)



In Figure 6, the model is trained for 1,000 epochs, and the loss function obtained by using the Adam optimiser converges faster and more stably. The reasons are as follows: Its built-in momentum mechanism helps accelerate convergence, while avoiding update bias in the initial stage through bias correction, making the loss value drop more stable. Finally, compared with traditional SGD or RMSProp, Adam is more robust to hyperparameters (such as initial learning rate), and experiments show that it can escape local optimality faster and achieve lower final loss values. These features make it a preferred solution for training complex deep models. Therefore, Adam is used as the optimiser in the training process of this paper.

5 Conclusions

The Ex-GAN model for the innovative design of ceramic cultural and creative products. First, multiple features of ceramic primitive data are integrated as the input of the generator, and then the generator and adversary game training is carried out, and finally the generation ability of the model is evaluated by indicators. The Ex-GAN IS index proposed in this paper can reach 5.7 and the FID index is as low as 2.36. The IS index can effectively illustrate that the model has the function of generating multimodal models, and the FID index verifies that the model can generate cultural and creative works that are closer to reality. Both indicators also show that the model has good generation ability. During the model training process, the effects of different activation functions and optimisers on the model are compared. The final model is trained using the leaky ReLU and the Adam, and the loss value is archived. The convergence process of the model is more intuitively observed using graphics. In addition, the Ex-GAN is compared with AE, VAE, and DCGAN. The comparison shows that the IS and FID performance indicators of the model are better than other models. This shows that Ex-GAN can generate realistic and practical cultural and creative works, and can be widely used.

Declarations

All authors declare that they have no conflicts of interest.

References

Cheema, A. and Naeem, M.A. (2025) 'CoverGAN: cover photo generation from text story using layout guided GAN', *Soft Computing*, pp.1–19.

Chen, W. and Zhao, H. (2025) 'EMF-GAN: efficient multilayer fusion GAN for text-to-image synthesis', *Computers and Graphics*, p.104219.

Elasri, M., Elharrouss, O., Al-Maadeed, S. and Tairi, H. (2022) 'Image generation: a review', *Neural Processing Letters*, Vol. 54, No. 5, pp.4609–4646.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2020) 'Generative adversarial networks', *Communications of the ACM*, Vol. 63, No. 11, pp.139–144.

He, J. and Tao, H. (2025) 'Applied research on innovation and development of blue calico of Chinese intangible cultural heritage based on artificial intelligence', *Scientific Reports*, Vol. 15, No. 1, p.12829.

He, Y. (2022) 'Research on innovative thinking of ceramic art design based on artificial intelligence', *Mobile Information Systems*, Vol. 2022, No. 1, p.3381042.

Huang, S. (2024) 'Investigating the relationship between ceramic art and AI-generated content: a study of aesthetics, creativity, and emotional responses in AI-generated ceramic art', *Pakistan Journal of Life and Social Sciences*, Vol. 22, No. 2.

Li, J., Zhang, C., Zhu, W. and Ren, Y. (2025) 'A comprehensive survey of image generation models based on deep learning', *Annals of Data Science*, Vol. 12, No. 1, pp.141–170.

Li, P., Pei, Y. and Li, J. (2023) 'A comprehensive survey on design and application of autoencoder in deep learning', *Applied Soft Computing*, Vol. 138, p.110176.

Liu, A-A., Sun, Z., Xu, N., Kang, R., Cao, J., Yang, F., Qin, W., Zhang, S., Zhang, J. and Li, X. (2024) 'Prior knowledge guided text to image generation', *Pattern Recognition Letters*, Vol. 177, pp. 89-95.

Liu, B., Lv, J., Fan, X., Luo, J. and Zou, T. (2022) 'Application of an improved DCGAN for image generation', *Mobile Information Systems*, Vol. 2022, No. 1, p.9005552.

Lu, J., Shi, M., Song, C., Zhao, W., Xi, L. and Emam, M. (2024) 'Classifier-guided multi-style tile image generation method', *Journal of King Saud University-Computer and Information Sciences*, Vol. 36, No. 1, p.101899.

Mi, J., Ma, C., Zheng, L., Zhang, M., Li, M. and Wang, M. (2023) 'WGAN-CL: a Wasserstein GAN with confidence loss for small-sample augmentation', *Expert Systems with Applications*, Vol. 233, p.120943.

Pinheiro Cinelli, L., Araújo Marins, M., Barros da Silva, E.A. and Lima Netto, S. (2021) 'Variational autoencoder', *Variational Methods for Machine Learning with Applications to Deep Networks*, pp.111–149, Springer.

Ren, H. (2022) 'Development and application of ceramic cultural and creative products based on artificial intelligence', *Wireless Communications and Mobile Computing*, Vol. 2022, No. 1, p.5733761.

Sun, X., Liu, X., Yang, X. and Song, B. (2021) 'Computer-aided three-dimensional ceramic product design', *Computer-Aided Design and Applications*, Vol. 19, No. S3, pp.97–107.

Wei, Z. and Cheng, X. (2022) 'Inheritance and innovation of traditional ceramics in Yuan River Basin', *Mobile Information Systems*, Vol. 2022, No. 1, p.1602312.

Yang, B., Xiang, X., Kong, W., Zhang, J. and Yao, J. (2025) 'SF-GAN: Semantic fusion generative adversarial networks for text-to-image synthesis', *Expert Systems with Applications*, Vol. 262, p.125583.