LayoutGAN for Automated Layout Design in Graphic Design: An Application of Generative Adversarial Networks

Sufen Liu[1, 2], Linzhi Zhou[1, 2,*]
[1]Digital media department, Guangzhou Institute of Software, Guangzhou 510990, China
[2]School of Art, CITl University of Mongolia, Urabato 999097, Mongolia
E-mail: yczhou8019@163.com
[*]Corresponding author

Keywords: artificial intelligence, graphic design, generative adversarial networks, layoutGAN, room plan layout

Received: September 6, 2024

The current field of visual communication design is particularly influenced by artificial intelligence, with automatic processing applications for images and color schemes having emerged, and a number of design models for intelligent layout have been proposed. However, the statistical learning framework model will face the limitation of black box, the process of graphic design cannot be separated from the real-time participation of designers, so intelligent graphic design is still essentially a semi-intelligent system. In this paper, LayoutGAN is used to study the layout generation problem in graphic design. The network structure is designed in accordance with the idea of LayoutGAN. Then experiments are conducted on MINST dataset to verify the effectiveness of the network structure. A network model was designed to address the limitations of insufficient vectorization results, difficulty in training and optimizing the pix2pix model in this task. Subsequently, a theoretical analysis was conducted on the annotation rules of the dataset, and a set of datasets for layout design tasks was annotated. Finally, the effectiveness of the proposed network model was demonstrated through experiments. The experimental results show that the wireframe rendering discriminator is better than the relationship-based discriminator, and then only the LayoutGAN with wireframe rendering discriminator is used to experiment on the room floor plan dataset, and finally the automatic floor plan generation is realized.

Povzetek: Opisano je LayoutGAN inteligentno orodje za avtomatizirano oblikovanje postavitev v grafičnem oblikovanju. S kombinacijo GAN in vektorske obdelave model izboljšuje kakovost in učinkovitost oblikovanja.

1 Introduction

Artificial intelligence has undergone three explosive developments since its inception. The first explosion in the 1960 s produced the first generation of intelligent programs, but was limited by the speed of computer computing, the complexity of the problems, and the lack of data. The second explosion in the 1980 s produced computer intelligence systems (expert systems) that hit the business model, but were defeated by Apple and IBM's general purpose computers [1]. Currently, artificial intelligence technology is playing an increasingly important role in helping humans to carry out information collection, information analysis work and carry out the process of decision-making, and intelligence will become the future development trend.

Of course, AI is still in the weak artificial intelligence stage, temporarily unable to have emotional feelings, cross-domain reasoning, abstract analogy and other subjective capabilities. But computers have three advantages over people: First, computing power, according to Moore's Law, computer processing power doubles about every two years, and can quickly complete complex operations. The second is the memory storage capacity, the computer can store data to call at any time. Third is the ability of objective accuracy, because there is no subjective emotion, only data processing problems, and therefore more objective than humans, and can repeat the same thing for a long time, will not be fatigue and ensure accuracy [2]. These advantages make AI superior to humans in solving certain purely rational problems, such as Alpha GO. However, design is not simply about solving problems, but also about understanding and creating beauty. This is a major problem that needs to be solved when AI penetrates into the design field, because weak AI has not yet been able to have subjective feelings and no aesthetic ability. In this regard, Fan Ling, director of Tongji Tezan Design and Artificial Intelligence Laboratory, believes that although the machine cannot yet solve the uncertainty of design, but can approach it. Machine cannot measure what is aesthetics, but can go through deep learning to produce products close to aesthetics.

Layout design for visual communication design is a relatively independent design art, the study is the visual language and artistic style of visual communication design. Standardized layout design is particularly important for graphic design because of its wide range of fields. Then the importance of layout design is also obvious. Layout distribution has a long history in China, and until today it has not been squeezed out of our lives. This is proof that our life and layout distribution cannot be separated from each other, and therefore cannot be separated from the layout design. Then the layout design also has its own characteristics, not arbitrary, layout design is its corresponding norms. Concepts and principles can be explained through lectures, but what makes an excellent layout design cannot be described by words alone. Based on this specification, the completed work of design can meet the aesthetic requirements of the public, and the utility of words can be prompted to be manifested [3]. At present, the combination of AI and visual elements in visual communication design has produced many intelligent applications, such as Logojoy, an intelligent logo design application that combines AI with textual elements, Prisma, an intelligent image processing application that combines AI with graphic elements, and Khroma, an intelligent color matching application that combines AI with color elements [4]. However, the combination of AI and layout is more complex than the combination with visual elements. This is because there are more elements involved, more ways to combine them, and a larger algorithmic model. This has become one of the important research directions of combining AI with visual communication design at present.

In this paper, we will explore the combination of AI and layout from the perspective of computer and designer, and propose a method that can use AI for layout design. This paper extends generative adversarial networks (GAN) to the field of graphic design, and tries to use the LayoutGAN model to conduct preliminary experiments on the MNIST dataset, debugging the network structure and hyperparameters that generate good results, and providing references for subsequent experiments. Then experiments are conducted on room layout maps in the floor plan dataset in order to achieve automatic generation of floor plan layouts.

2 Related works

2.1 Graphic design and layout

Graphic design, also known as visual communication design, is named after the medium of “vision” that enables communication between people and works. It is the artistic combination of symbols, images and text in a variety of forms to express the ideas and messages that designers want to convey to people. The graphic designer uses typography, layout, visual design, and technical support from computer software to complete the intended creation. The essence of graphic design is the process of design and the later completion of the work, which is the process of activities to communicate the plan, planning and vision through some visual form [5]. Graphic design covers a wide range of fields, mainly in the field of packaging (design of bags and packaging appearance), advertising (design of promotional posters), Internet (graphic elements of websites), logos (design of trademarks and brands), and publications (books and magazines, etc.). The implantation of graphic design can be seen in all industries, which shows the influence and importance of graphic design in today’s society.

Layout design, the so-called layout design, refers to the implementation of organic combination of various elements, such as text, images, colors, etc., in a limited space to achieve a reasonable arrangement. The implementation of typography for various elements is essentially a test of the designer’s professional skills and comprehensive ability, requiring the designer to have a good sense of innovation, but also to ensure the artistry and effectively meet the specific requirements of layout design [6]. The designer should think deeply and accurately about the design expression content, optimize the design expression methods, and enhance the artistry and coordination of the layout design with the help of diverse expression methods such as modeling and arrangement, to achieve good information transmission to the audience, enhance the visual communication effect, and reflect the designer’s exquisite design concept.

In the process of layout design, text, graphics and other elements are always indispensable, and with the role of these elements can easily achieve the transmission of information and sublimation of the effect, so that people have a different visual experience [7]. In the same way, the application of layout design in graphic design will also rely on the interaction and support of these elements.

As a special graphic design symbol, text is an effective carrier of information and an infectious design element in graphic design, and is therefore widely used. There are many reasons for using text, one of which is to increase readability. The text elements in layout design mainly undertake the task of conveying the designer’s intention and various effective information to the audience, so the text elements must have the basic readability. Only easy to recognize, easy to understand, in order to better serve the overall style of the work, with readable text to express the design theme and convey information [8]. The second is because the text has creativity. The text in layout design is not single and boring, but must present a personalized expression based on the theme of graphic design. In the external form and design tone to bring the audience a pleasant feeling, such as a reasonable deformation of the text strokes with, highlighting the beauty of the font structure and strokes, can make the graphic works more creative.

The value of graphics in graphic layout design is mainly reflected in expressing the design intention in the design, showing the design theme and creative ideas [9]. From this level, designers need to personalize the cut and deformation of graphics to highlight the detail effect when they are applied to their design [10]. In this process, the visual form of graphics should take the direction of human visual flow or the direction of movement trend as the main reference standard to maximize the existence value of graphics.

Graphic design works in the picture of the main priority, the theme is prominent, can effectively increase the layout of the visual hierarchy and rhythm. This point must attract the designer’s attention and be strictly observed in the specific application practice. A good layout can always let the audience receive information in a relaxed and pleasant atmosphere, sublimating the design effect in a subtle way [11]. Take “white space” as an example, the application of this technique can bring good visual aesthetics to graphic design works. At the same time, it does not cut off the integrity of the layout, and can create a virtual space to a certain extent, achieving artistic effects that cannot be achieved by intensive layout design [12].

2.2 Application of artificial intelligence in graphic design

The development of artificial intelligence and digital media has put forward new requirements for graphic design, and there are different requirements for graphics, text and color in graphic design. Graphic design in the new era is more often displayed on mobile terminals than just the old paper media, and digital media has become an important carrier of information dissemination [13]. The change in the way of media has accordingly changed the way people read, which has put forward new requirements for graphic design to adapt to the new media.

In February 2020, British artist Patrick Thomas and graphic designer Jonathan launched Open collab2.0, a digital platform for multiplayer visual collaboration [14]. It is an intelligent tool that uses random design elements to combine into new images. Participants first design a specific graphical language, which is then randomly combined by Open collab2.0 to generate different visual works. In this way, the resulting visuals are full of uncertainty and surprise. The experimentation of artificial intelligence in visual design goes far beyond this. In the previous concept of art, works of art were regarded as a vehicle for individual artists to express their emotions, an artistic activity unique to humans. However, in recent years, fine art works under the influence of artificial intelligence have begun to appear, where a large number of paintings are collected by computer algorithms, and then the computer breaks up and reconstructs the elements of the paintings to create new paintings.

Under the influence of technology, the development of art and design will produce new changes, and visual art is also groping for the most appropriate path to combine with artificial intelligence. In the past visual design, the presentation of information was mainly in the form of static graphics and text. However, with the development of science and technology and the advent of the information age, the previous static visual presentation form has been difficult to meet people’s aesthetic needs. Designers find that dynamic visual elements can often better attract people’s attention and convey information more effectively, and jumping and changing images can improve the efficiency of conveying information [15]. The previous static visual language can only guide the flow of information through the arrangement of layout, while dynamic visual language can better guide people’s visual attention and enhance the interaction between information and people.

On the basis of existing visual space models, a “fixed size visual perception binary experiment” is designed and conducted to verify the differences in geometric properties between physical space and visual space under complex environmental conditions in landscape environments, and to explore the geometric properties of visual space. Fit the experimental results using the current primary visual spatial model. Based on the fitting results, determine the constant parameters in each model and compare the applicability of different visual models in complex environments. Select the model with high fitting degree as the specific way to measure the transformation from physical space to visual space in the quantification process of this study.

Based on the imaging characteristics of visual senses and the related research and design issues and needs of landscape environment, sort out the quantitative elements and indicators of visual perception space for landscape environment, and determine the basic content of quantification. Construct a “Landscape Environment Visual Perception Spatial Index Measurement Framework” based on visual related principles and geometric properties of visual space, clarify the spatial characteristics and transformation logic of each stage of information transformation, and determine the measurement methods for various basic indicators according to the framework.

Based on the basic indicators and measurement of visual perception space, the basic characteristics of visual perception of landscape environment are determined according to research and practice related to landscape environment. Construct a “landscape motion element” model, using “landscape” as the description of visual features of the landscape environment, “motion” as the organization of features from local to overall, and “elements” as the medium of association between sensory perspectives and objective environmental elements. This enables the description of features from simple to complex, from local to global, from static to dynamic, and from sensory perspective to objective environment.

The main purpose of quantitative system design is to construct a three-dimensional visual perception spatial quantitative support system that meets the research and design needs of landscape environment related majors. The research on quantitative methods focuses on the quantitative implementation of visual perception spatial related indicators and features in landscape environments, using software commonly used in landscape environment practice as a platform for scene model construction and quantitative results display. In this process, it is necessary to consider the geometric and spatial characteristics of different quantification stages, meet the characteristics of landscape environment research, and overcome many problems in existing methods, in order to create a data model architecture and transformation method that meets the needs of visual perception spatial information quantification at different stages. On this basis, specific algorithms are designed based on the measurement methods of each indicator feature and the data structure of the spatial model, in order to achieve accurate quantification of spatial related indicator features of landscape environment visual perception [16].

The biggest difference between art design and fine art is that design works are not exactly the carrier of the creator's own emotional expression, but ultimately serve the public and meet people's physical and spiritual needs [17]. Art and design are disciplines that combine rationality and sensibility. In visual design, designers should establish a sense of interdisciplinary learning, think in multiple dimensions, and create visual design works not entirely based on personal preferences, but according to the specific situation, using psychology, ergonomics and other multidisciplinary knowledge. At the same time, art and design disciplines are not closed to each other, and in recent years, design disciplines have been making interdisciplinary attempts, using the concept of product design in environmental art design, and visual design in product design, while visual design also refers to the design ideas of environmental art design and product design. In addition, designers have to consciously improve aesthetic literacy and insist on learning aesthetics. For visual designers, drawing ability is still important. Although the use of artificial intelligence can create paintings, what is created in this way is often just a mechanized artwork that does not reflect human thinking; only artworks with human participation are humane and vivid. Therefore, visual designers have to learn from multiple perspectives and continuously enrich their expertise.

Artificial intelligence is now capable of simulation, learning and natural evolution in certain fields. This is not just a simple advancement in the accumulation of repetitive tasks, but is closer to a breakthrough that builds on the advanced modeling of occupations, just as there are certain rules that govern the work of any occupation. In a way, the emergence of artificial intelligence makes it possible to replace humans in their work.

Overall, the current research status of graphic design is mainly shown in 】 1.

Table 1: Research status of graphic design

Graphic Design and Layout	Visual Communication Design
	Layout Design
	Design themes and creativity
Artificial Intelligence and Digital Media	Design software
	Dynamic visual elements
	visual perception
	Visual features
	Feature Strategy and Extraction
	Quantitative System Design

3 Algorithm design

A good layout not only can reasonably use the space of graphic design, but also can make the graphic design present a certain degree of beauty, thus arousing the interest of users. In the past, layouts were usually designed by graphic designers, but the efficiency of layout generation would be greatly improved if the layout could be automatically generated. The problem of automatic layout generation has received increasing attention from industry and academia in recent years, and there have been many related studies, but there are fewer studies that use generative adversarial networks to solve the problem.

3.1 Generators

Represent the data in the model of this article as $z = \{(p_1, \theta_1), \cdots, (p_n, \theta_n)\}$, the adjusted category probability p' and the geometric parameter θ' are finally output, the output data form is represented as $G(z) = \{(p_1', \theta_1'), \cdots, (p_n', \theta_n')\}$,as shown in Figure 1 [18].

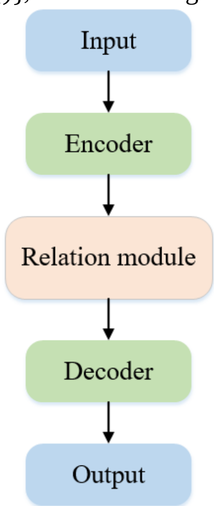


Figure 1: Schematic diagram of LayoutGAN generator structure

As shown in Figure 1, it is a network composed of an encoder, a relational module, and a decoder. The network structure of the encoder is a multi-layer perceptron, which first processes the class probability and geometric parameters of each element; The relationship module comprehensively considers the spatial and semantic relationships between elements, and further refines the features output by the encoder; The network structure of the decoder is also a multi-layer perceptron, which reconstructs the fine features formed by the relationship modules and finally outputs new class probabilities and geometric parameters. At this point, each element forms a good layout.

3.2 Relationship discriminator

Relationship-based discriminators have a structure similar to that of generators and consist of a coder, a relationship module, a classifier with the same structure of a multilayer perceptron is used to deliver a value of odds ratio for determining the true layout and the synthetic layout. The structure of the relational discriminator network is shown in Figure 2.

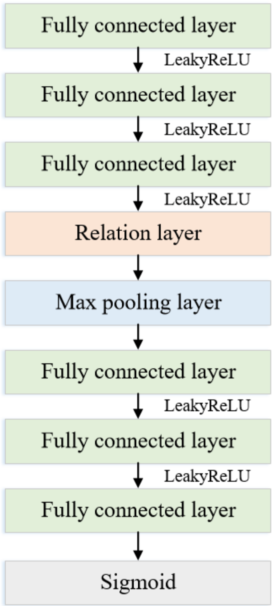


Figure 2: Structure of relational discriminator

The structure of a relation discriminator is somewhat similar to that of a generator, consisting of an encoder, relation module, and classifier. Based on the relationship discriminator, an encoder composed of multiple layers of perceptrons is first used to embed the class probability and geometric parameters of each element; Then use a simple relationship module to extract spatial relationships between elements and perform a max pooling operation. Finally, a classifier with a multi-layer perceptron structure is used to output a probability value for determining the true layout and synthetic layout.

3.3 Wireframe rendering discriminator

The wireframe rendering discriminator uses convolutional neural networks that are good at capturing spatial relationships. As convolutional neural networks usually only receive input of images, but the initial input is in the form of data consisting of category probabilities and geometric parameters, the wireframe rendering discriminator adds a wireframe rendering layer before the convolutional neural network, which converts the input data into the representation of images. Wireframe rendering process is shown in Figure 3 [19].

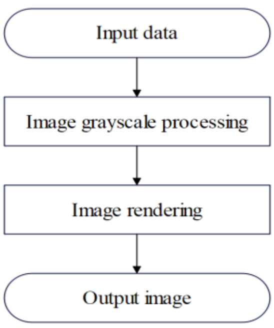


Figure 3: Wireframe rendering process

The number of elements input into the system model is defined as n, which can render the image. The image is represented as $F_\theta(x,y)$, and the size of the image is defined as W × H. After system processing, the resulting image is W × H × M. The element category is represented as M. Based on this, the relationships between images can be obtained (1) [20]:

$$I(x,y,c) = \max_{i\in[1,\cdots,n]} p_{i,c} F_{\theta_i}(x,y) \tag{1}$$

Where $p_{i,c}$ denotes the probability that element i belongs to category c.

When the element is a point, which is represented as $\theta_i = (x_i, y_i)$, the rendering of the point into a grayscale map is calculated as (2):

$$F_{\theta_i}(x,y) = k(x-x_i)k(y-y_i) \tag{2}$$

Where, $k(x-x_i) = max(0,1-|x-x_i|), k(y-y_i) = max(0,1-|y-y_i|)$.

When the element is a rectangle, which is represented as $\theta_i = (x_i', y_i^f, x_i^r, y_i^b)$, then rendering the rectangle into a grayscale map is calculated as (3):

$$F_{\theta_i}(x,y) = max\left(\begin{matrix} k(x-x_i^l)b(y-y_i^t)b(y_i^b-y), \\ k(x-x_i^r)b(y-y_i^t)b(y_i^b-y), \\ k(y-y_i^t)b(x-x_i^l)b(x_i^r-x), \\ k(y-y_i^b)b(x-x_i^l)b(x_i^r-x) \end{matrix}\right) \tag{3}$$

Where, $k(x-x_i^l) = max(0,1-\mathrm{I}x-x_i'\mathrm{I}), b(x-x_i^l) = min(max(0,x-x_i^l),1)$, Other representations are similar.

3.4 Based on generative adversarial network model

The layout design of the floor plan is the main content of indoor intelligent design tasks. Focusing on the task of using generative adversarial networks for floor plan design, a network model was designed to address the limitations of pix2pix model in this task, such as insufficient vectorization results, difficulty in training and optimization.

This section designs the structure of the pix2pix network model to address the issues in the layout design task of house layouts. The designed network model can achieve separate prediction of furniture, making it easier to extract numerical values during subsequent use. On the other hand, it also avoids the problem of equivalent furniture combinations to a certain extent. The original pix2pix generator model and the designed generator structure are shown in Figures 4 and 5.

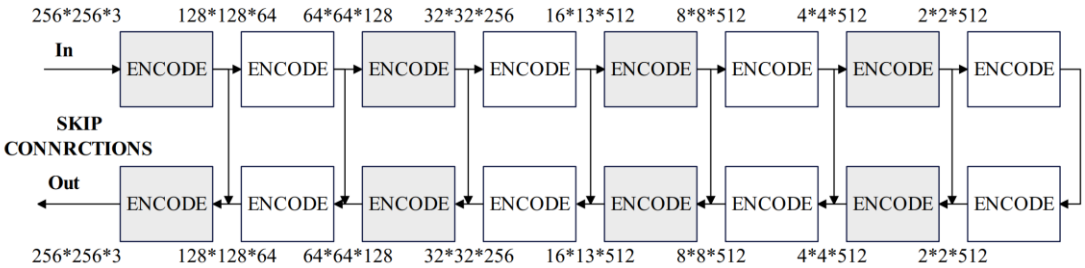


Figure 4: Pix2Pix generator model

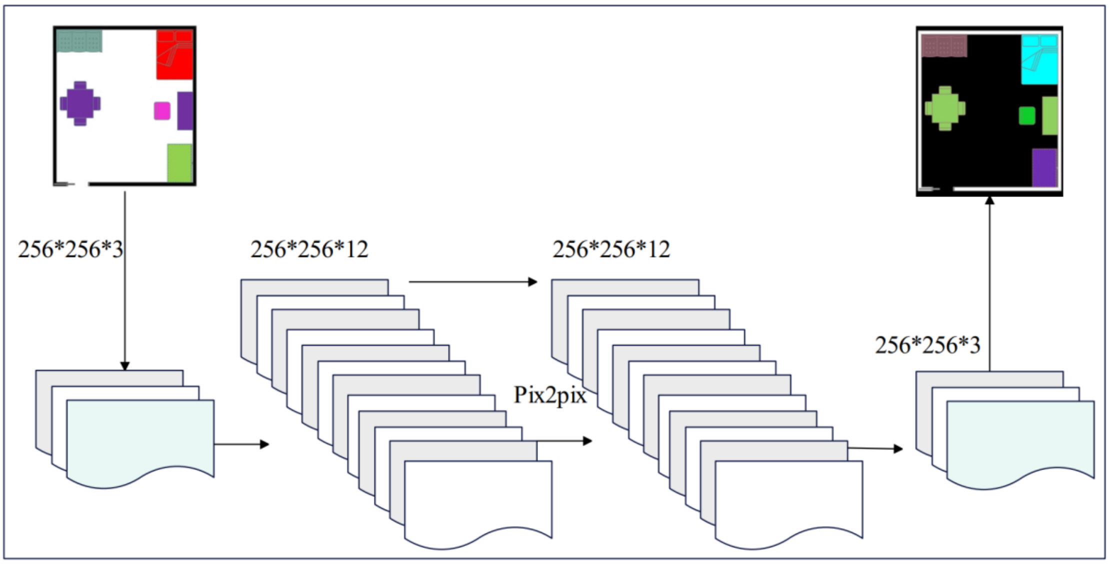


Figure 5: The designed generator model

The specific design scheme of the network model is as follows: (1) Add convolutional layers at the decoding end of the generator. There are two solutions to the problem of lacking numerical information in the pix2pix model. One method is to change the network infrastructure so that the network can directly output digital information such as furniture location and size. Another method is to optimize the output image to make it easier to calculate and extract numerical values. Directly using the network to output specific numerical information not only breaks the basic framework of image translation tasks, but also greatly increases the difficulty of dataset annotation. Therefore, this article chooses a scheme to optimize the output image. The main reason why the pix2pix model is difficult to extract parameters is that the output image only has one overall plane image, and the output layer of the generator needs to be adjusted first. In order to facilitate the extraction of subsequent values, the network needs to implement separate predictions for different furniture. Therefore, a convolutional layer is added to the decoder section of the generator to segment the information of different furniture. The size of the newly added convolutional layer is 256 * 256, which is the same as the output image. The number of channels in the convolutional layer is determined by the predicted number of furniture types. The main function of this layer is to achieve personalized generation of various types of furniture, and each channel in the convolutional layer corresponds to a probability prediction of a furniture layout.

(2) Normalize the newly added convolutional layers. Simply adding convolutional layers does not enable the network to achieve furniture separation prediction. In order to enable each channel of the newly added convolutional layer to predict the layout of furniture separately, it is necessary to normalize the convolutional layer in the channel dimension. Therefore, on the basis of adding convolutional layers at the decoding end of the generator, the softmax function is further used to normalize the values of each pixel in the image on each channel.

(3) Perform DW convolution operation on the newly added convolutional layer. By adding convolutional layers at the decoding end of the generator and normalizing the convolutional layers, the network has achieved preliminary separation prediction for furniture. But at this point, if the furniture probability map is directly used as the final output, excessive annotation work needs to be done on the data, and the discriminator model also needs to undergo significant changes. In order to maintain the simplicity of the model, it is necessary to convert the furniture probability maps of multiple channels into the final floor plan to eliminate the impact of adding convolutional layers on the model structure. Depth Separable Convolution Operation is shown in Figure 6.

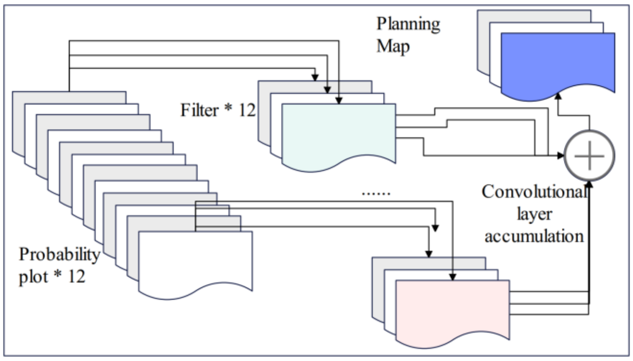


Figure 6: Depth separable convolution operation

(4) Add convolutional layers to the generator encoding segment. The pix2pix model uses the U-Net structure, which allows the input and output images to share underlying information. In the task of generating floor plan layouts, the positions of walls, doors, and windows in blank room layouts and floor plan layouts are the same, which greatly helps network models learn the correspondence between rooms and furniture. Adding a convolutional layer at the decoding end of the generator breaks this structure, so in order to maintain the invariance of the U-Net structure, a convolutional layer of the same size and dimension is finally added to the encoder part of the generator.

4 Experiments

4.1 Methods

In order to ensure the authenticity and effectiveness of the dataset, this article selects 500 sets of real floor plans with complete furniture layouts as the samples to be annotated. Due to the inability of neural networks to handle an indefinite number of rooms and furniture types, this article has made a trade-off between annotating rooms and furniture types. After statistical analysis of the sample data, this article selected the six most frequently occurring room types for annotation, namely bedroom, living room, dining room, bathroom, dining room, and study. These room types can meet people's basic needs for daily life, rest, entertainment, work, dining, kitchen and bathroom. At the same time, in order to ensure the integrity of the annotated image, this article uses a cluttered area to fill the annotated image. After confirming the annotated room types, this article annotated the high-frequency furniture in each room type, which is the main furniture in that room and the core of the furniture layout task. After determining the layout of the main furniture, the layout of the attached furniture can be confirmed through the dependency relationships between furniture, such as the dependency relationships between bedside tables and beds, chairs and tables.

Since the model input data are in the form of category probabilities and point coordinates, it is necessary to first convert the grayscale maps into point coordinates. The operation used is to first find out the coordinate information of each handwritten digital image with a grayscale value greater than 180, assuming that there are m coordinate points that meet this condition, and then generating 128 of them. Then the 128 coordinate points are normalized, and then the normalized coordinate points are combined with the category probabilities to finally obtain the input data of size 128×3. The experimental environment is shown in Table 2.

Table 2: Experimental environment

Name	Versions
Python	3.9.13
Tensorflow-gpu	NVIDIA GeForce GTX 1650
CUDA	12.1.1
CUDNN	V7.5.0
Opencv-python	4.4.1
Keras	2.3.0

In order to effectively utilize the existing training model, the experiment adopts transfer learning method. Firstly, the model is trained at a 0.5 bpc embedding rate. Then, the weights of the trained model at a 0.5 bpc embedding rate are initialized to the model at a 0.4 bpc embedding rate, and so on. Finally, the weights of the trained model at a 0.2 bpc embedding rate are used to initialize the model at a 0.1 bpc embedding rate When doing transfer learning, in order to avoid losing the information learned under the previous embedding rate due to a high learning rate, the learning rate is set to 0.000 01.

This article divides the sample dataset into a training set and a testing set in an 8:2 ratio. The training set is used for training the graphic design network model, optimizing the model's weights and bias parameters; The test set is used to evaluate the anti design capability of the model.

To conduct the experiment, this article collected images from the floor plan of building rooms, totaling over 45000 image samples, and set the layout through a simulation platform. In the experiment, the image size was set to 512 × 512, so all sample images were adjusted to the same size, as shown in Figure 7 as an example image. The object categories of the room and the corresponding color settings are shown in Table 3.

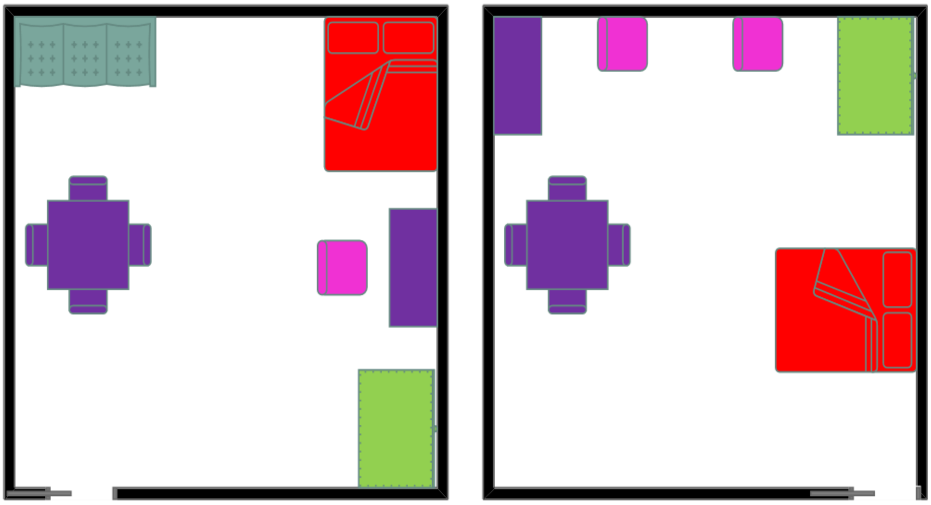


Figure 7: Schematic layout of the room plan

Table 3: Object categories of the room and corresponding color settings

Category	Color	RGB value
Wardrobe	White	(255,255,255)
Bedside table	Green	(0,255,0)
Double bed	Red	(255,0,0)
Table	Purple	(255,0,255)
Single bed	Blue	(0,0,255)
Dressing table	Yellow	(255,255,0)
Office chair	Pink	(255,192,203)
TV ark	Brown	(128,42,42)
Armchair	Orange	(255,128,0)
Floor lamp	Grass green	(107,142,35)
A footstool	Rose-bengal	(188,143,143)
Chair	Cyan	(0,255,255)

After the wireframe rendering layer, the output size is 64×64×12. In addition, the parameters of the convolutional layers are set differently. In this section, the number of output channels of the three convolutional layers is changed to 24, 48, and 96, respectively, and each convolutional layer is batch normalized.

4.2 Results

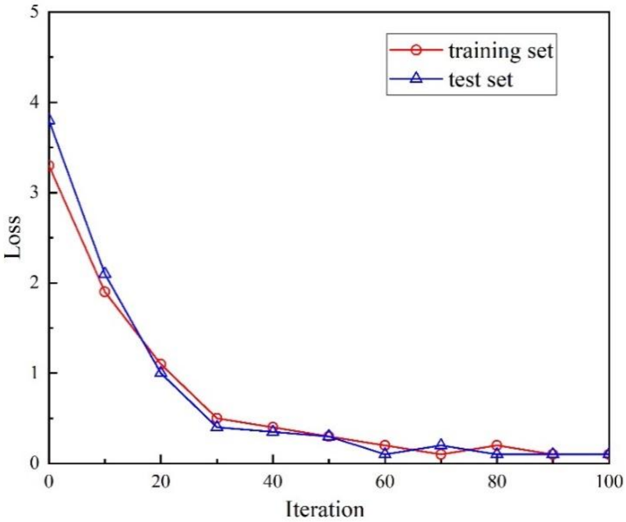
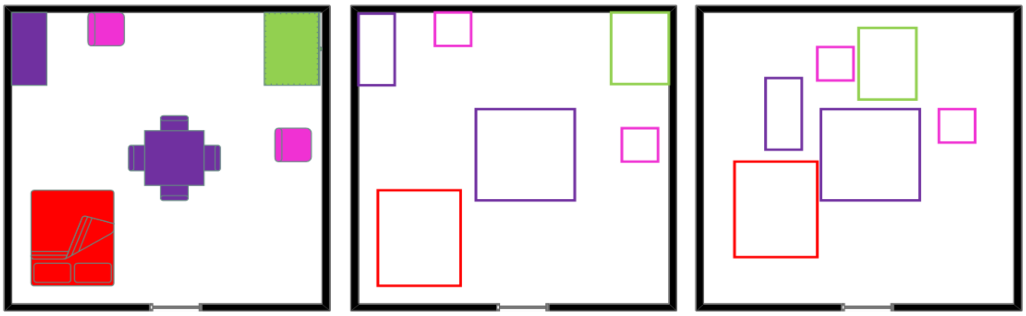


Figure 8: Loss curve during training

The stochastic gradient descent algorithm is chosen as the optimizer with a momentum of 0.9, a learning rate of 0.001, and 100 iterations. Figure 8 shows the loss curve of the training process. The model converges after approximately 60 iterations.

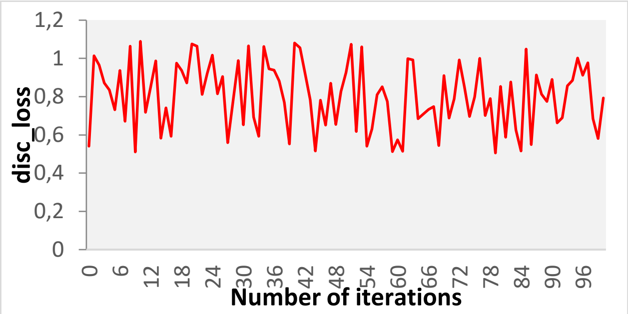
The experimental results based on the layout dataset are shown in Figure 9.



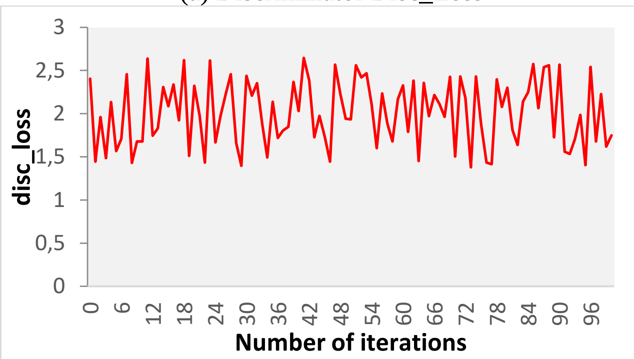
(a) (b) (c)

Figure 9: Comparison of experimental results of room plan layout

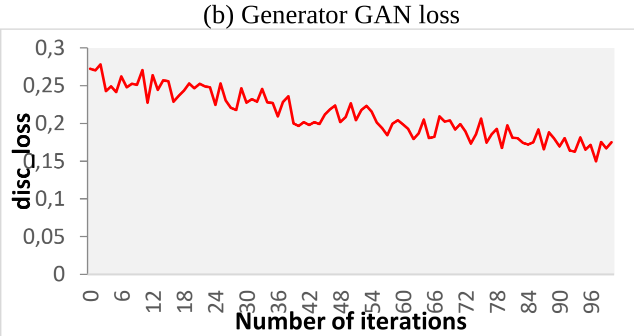
This experiment was iterated 20000 times, and the loss function of the model was recorded every 200 times. The specific recorded loss function is as follows Figure 10.



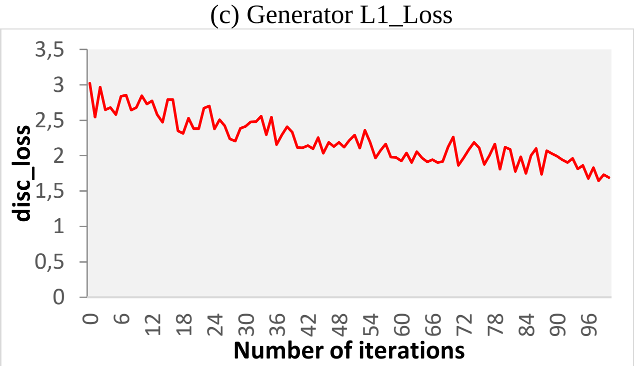
(a) Discriminator Disc_Loss



(b) Generator GAN loss



(c) Generator L1_Loss



(d) Generator total loss

Figure 10: Pix2Pix loss function curve

Through learning and experimental analysis of popular generative adversarial network models, the original Generative Adversarial Network (GAN), Deep Convolutional Generative Adversarial Network (DCGAN) models were selected for comparison with the model proposed in this paper. Compared to the original GAN model, the biggest feature of DCGAN is the use of convolutional layers, which improves the stability of model training. In addition to metrics such as convergence time and generation time, IS (initiation score) and FID (fr é chetin perception distance) were selected as key indicators to evaluate the performance of the model. IS and FID are widely used to measure the quality and diversity of output images from generative adversarial networks. The higher the IS score, the lower the FID value, and the higher the diversity and quality of the generated images. Among them, FID is closer to human judgment in evaluating the authenticity and variability of generated images, and is very suitable for evaluating subjective creative patterns. The comparative experimental data results are shown in Table 4.

Table 4: Comparison of experimental results between these model and typical models

Model	Convergence time/min	Generation time per 10000 sheets/s	FID	IS
The model in this article	15.84	69.79	28.81	3.31
The original GAN model	30.69	78.93	36.83	2.32
DCGAN model	18.81	75.36	29.80	2.41

4.3 Analysis and discussion

As shown in Figure 8, the model in this article begins to converge after 60 iterations, indicating that the model has good learning and convergence effects, and has certain practical effects.

Figure 9 (c) is the layout image with better generation effect selected from the results of batch generation. For visual comparison, the images with more similar layouts are selected from the original dataset, where Figure 9 (a) is the original image and Figure 9 (b) figure shows the objects of different categories in Figure 9 (a) figure with rectangular boxes of different colors. The Figure 9 (c) figure presents the layout of six indoor objects, where the generated bedside table is kept roughly on both sides of the bed, the dressing table and the chair are placed together, the TV stand is placed in front of the bed, and the floor lamp is placed next to the TV stand, which shows that LayoutGAN can indeed learn the layout pattern of the original dataset to some extent, but compared with the Figure 9 (b), the generated layout map has obvious misalignment and overlap and the objects are close to each other, which is an area for further improvement in future research.

As shown in Figure 10. Discriminator disc loss: the sum of the cross entropy loss between the real image and the sigmoid with element 1, and the cross entropy loss between the generated image and the sigmoid with element 0. Generator GAN loss, generator L1 loss, generator total loss. From the loss values of the models, it can be seen that neither the generator nor the discriminator model has achieved absolute victory and is still in a process of confrontation. But the GAN loss of the generator gradually increases over time, and the disc loss of the discriminator gradually increases over time, indicating that the discriminator is gradually gaining the upper hand, and the learning ability of the GAN generator is gradually unable to compete with the discriminator anymore. For the discriminator disc loss and generator GAN loss, when the value is log (2)=0.69, both models perform best because 0.69 indicates that the discriminator cannot make a judgment on whether the image is a real image or generated by the generator during this process, which is a manifestation of the average uncertainty of the two options. So comparing the discriminator disc loss with the generator GAN loss and 0.69, it can be seen that the discriminator has better discriminative ability than random numbers, while the generator performs worse than random numbers. And as the training progresses, the LI loss of the generator gradually decreases. This indicates that the difference between the generated image and the real image is becoming smaller and smaller, and the total loss of the generator, which is a combination of L1 and GAN losses, is also decreasing. The model training results are good.

In Table 4, compared with the original GAN model and DCGAN model, the constructed model exhibits better performance in innovative design of traditional calligraphy textile patterns.

5 Conclusions

After summarizing the common application areas of generative adversarial networks, this article further applies them to the field of graphic design to solve the problem of automatic generation of graphic layouts. Selecting LayoutGAN as the key research model, theoretical exploration was conducted on the generator, relationship discriminator, and wireframe rendering discriminator of the LayoutGAN model. Then, experiments were conducted on handwritten digital datasets using LayoutGAN models with different types of discriminators. The results show that the model proposed in this article has certain advantages in automatic layout of graphic design.

In the experiment of generating flat layouts, there is an obvious problem that the model training process is quite slow, which may be related to the amount of data, input data format, and network structure design. How to make LayoutGAN more efficient is a problem to be considered in the future. In addition, the experiment in this article only generated a rough random layout, and further exploration is needed to determine how to make the model generate a more sophisticated layout that meets specific conditions. Finally, LayoutGAN can also be experimented with on various types of datasets, such as magazine layouts, poster layouts, etc., to achieve multi-faceted applications and enhance the stability of LayoutGAN.

Availability of data and material

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

Declares that the author have no conflict of interest.

References

[1] Li, J., Yang, J., Zhang, J., Liu, C., Wang, C., Xu, T. Attribute-conditioned layout gan for automatic graphic design. IEEE Transactions on Visualization and Computer Graphics, 27(10), 4039-4048, 2020. https://doi.org/ 10.1109/TVCG.2020.2999335

[2] Hu, R., Huang, Z., Tang, Y., Van Kaick, O., Zhang, H., Huang, H. Graph2plan: Learning floorplan generation from layout graphs. ACM Transactions on Graphics (TOG), 39(4), 118, 2020. https://doi.org/ 10.1145/3386569.3392391

[3] Zheng, Y. Visual memory neural network for artistic graphic design. Scientific Programming, 2022(1), 2243891, 2022. https://doi.org/10.1155/2022/2243891

[4] McCartney, N., & Tynan, J. Fashioning contemporary art: a new interdisciplinary aesthetics in art-design collaborations. Journal of Visual Art Practice, 20(1-2), 143-162, 2021. https://doi.org/10.1080/14702029.2021.1940454

[5] Terzidis, K., Fabrocini, F., Lee, H. Unintentional intentionality: art and design in the age of artificial intelligence. AI & SOCIETY, 38(4), 1715-1724, 2023. https://doi.org/10.1007/s00146-021-01378-8

[6] Noble, K. 'Getting hands on with other creative minds': establishing a community of practice around primary art and design at the art museum. International Journal of Art & Design Education, 40(3), 615-629, 2021. https://doi.org/10.1111/jade.12371

[7] Zeller, S. Centering the periphery: reassessing Swiss graphic design through the prism of regional characteristics. Design Issues, 37(1), 64-75, 2021. https://doi.org/10.1162/desi_a_00625

[8] Lzicar, R. Swiss graphic design: A British invention? Design Issues, 37(1), 51-63, 2021. https://doi.org/10.1162/desi_a_00624

[9] Hashimi, S. A. A. Enhancing the creative learning experience through harnessing the creative potential of digital and social media platforms in art and design educational contexts. International Journal of Arts and Technology, 12(1), 84-101, 2020. https://doi.org/10.1504/IJART.2020.107681

[10] Lima, C., Alvelos, H., Barreto, S., Penedos-Santiago, E., Santos, R., Martins, N. Case studies of a trans-generational pedagogy of art and design. International Journal of Design in Society, 16(1), 2022. https://doi.org/10.18848/2325-1328/cgp/v16i01/113-129

[11] West, R., Burbano, A. AI, arts & design: Questioning learning machines. Artnodes, (26), 1-9,

2020. https://doi.org/10.7238/a.v0i26.3390

[12] Zhao, J., Zhao, X. Computer-aided graphic design for virtual reality-oriented 3D animation scenes. Computer-Aided Design and Applications, 19(1), 65-76, 2022. https://doi.org/10.14733/cadaps.2022.S5.65-76

[13] Methuku, H., Adel, I. Pioneering entrepreneurship using art & design pedagogy–a case study of scientific college of design, Oman. Journal of Arts and Humanities, 10(12), 72-79, 2021. https://doi.org/10.18533/jah.v10i12.2208

[14] Wang, C., Su-Lynn, G., Chen, A. From Chinese aesthetic to art and design: exploring Chinese aesthetics of Chinese ink painting to create contemporary art. The International Journal of Visual Design, 14(1), 11, 2020. https://doi.org/10.18848/2325-1581/CGP/v14i01/11-22

[15] Griffith, J. Cover letter: visual storytelling and empathy: an interview with graphic novelist and voices from the middle cover artist Sophie Escabasse. Voices from the Middle, 29(1), 10-13, 2021. https://doi.org/10.58680/vm202131423

[16] Salimbeni, G., Fol Leymarie, F., Latham, W. A machine learning application based on giorgio morandi still-life paintings to assist artists in the choice of 3D compositions. Leonardo, 55(1), 57-61, 2022. https://doi.org/10.1162/leon_a_02073

[17] Gulyamov, K. M., Ikramov, M. H. Development of artistic and creative abilities of future teachers of fine arts through computer graphics. JournalNX, 7(06), 95-99, 2021.

[18] Dzemyda, G., Sabaliauskas, M., Medvedev, V. Geometric mds performance for large data dimensionality reduction and visualization. Informatica, 33, 299-320, 2022. https://doi.org/10.15388/22-infor491

[19] Chen, H., Nedzvedz, O., Ye, S., Ablameyko, S. Crowd abnormal behaviour identification based on integral optical flow in video surveillance systems. Informatica, 29(2), 211-232, 2018. https://doi.org/10.15388/Informatica.2018.164

[20] Tiwari, P., Gupta, P. Entropy, distance and similarity measures under interval valued intuitionistic fuzzy environment. Informatica, 42(4), 617-627, 2018. https://doi.org/10.31449/inf.v42i4.1303]