Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/42/20230770

How does AI create and recommend corresponding wallpapers based on the games played by users?

Wenshu Hou[1,5,6,†], Zicheng Liu[2,†], Junjie Deng[3,†], Jiacheng Wang[4,†]

[1]University at Buffalo, State of New York,14260, USA
[2]Qingdao Academy, Qingdao, Shandong,266213, China
[3]Wuhan University of Science and Technology, Wuhan, 430070, Hubei, China
[4]Shanghai United International School Jiaoke Campus, Shanghai,201102, China

[5]khou257@gmail.com
[6]corresponding author
[†]These authors contributed equally to this work and should be considered co-second authors.

Abstract. The purpose of this paper is to give a comprehensive review of the related work that has everything to do with creating a wallpaper in artificial intelligence (AI) technology. Firstly, deep learning and neural network are summarized, especially generative adversarial networks that aim to effective generate images. Then, User interest modeling is summarized and analyzed, which is a key point to figure out the preference of the players. Further, main ideas are given about the trends and directions of wallpaper generation by AI.

Keywords: artificial intelligence, generative adversarial network, user interest modeling, deep learning.

1. Introduction

AI concept principles were proposed in the 1956 Dartmouth Institute. The invention of ChatGPT has led AI that become a popular technique that every field uses. ChatGPT has captured the world's attention since November 2022, demonstrating the extraordinary potential of artificial intelligence to millions of users around the world. Based on previous AI models, ChatGPT's strengths can be linked to its ability to respond to multiple languages that generate sophisticated and complex responses based on advanced modeling [1]. One can conclude that artificial intelligence will be a potential tool for social progress.

AI has different functions that can be used in different fields. It can draft a cover letter, write a few lines of poetry impersonating Shakespeare, and even craft a message for a dating app user to attract a mate [2]. It proves that AI has a powerful ability to satisfy people’s requirement.

Not only can AI generate text work for the client, but also AI can base client requirementsto generate wallpaper for potential customers. DeepArt is a famous AI graphic generator that allows users to generate various artworks and attracts many art lovers. Depending on Deep Art's good drawing skills, many game lovers use Deep Art to create their own game wallpapers. However, such game AI sites can be altered. AI can use big data-related crawler technology to find games or themes that users like.

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

This article has many techniques about AI and big data. First, the thesis will describe how AI can generate a high-resolution image. For example, AI uses GAN to repair images and improve resolution. Moreover, the training of the neural network is another important challenge faced by the project. Neural networks provide a powerful framework for modeling complex relationships and making predictions from large amounts of data. Their breakthroughs in image and speech recognition, natural language processing, and recommendation systems have revolutionized many areas of artificial intelligence. Finally, this essay will divide into 2 sessions, respectively, which is 1, "AI Technique" 2, "User Interest Modeling" to show the principle of AI to creat and optimize a wallpaper and publish this wallpaper to related people.

1.1. AI Technique

In the era of big data, such as collecting voice, text, image and other data is an extremely difficult problem to solve. Images are being generated all the time but the sharpness, type, size, etc. of the images are different. Therefore, it cannot be directly used in the research of artificial intelligence and deep learning. Human and material resources are required to collect images and process them. If image data that meet the requirements can be generated through technology, the process of collecting and processing images can be reduced, and manpower and material resources can be saved. In the next part, the article will discuss how AI creates images through Deep learning, Gans, and Application of Generative Adversarial Networks in Image Generation.

1.2. Deep learning and neural network

Machine learning plays an important role in modern society. In the domain such as financial risk management, medical image processing, transportation, and logistics management, people can utilize it to improve the efficiency of their work. Increasingly, these domains use a technique called deep learning. Representation learning guides the computer to extract raw data and to get the representations that are useful for the classification automatically. Deep learning is a kind of machine learning that has a number of levels of representation which are formed by composing non-linear modules and each transform the representation at one level into one at a more abstract level [3]. Hard and tough functions can be analyzed by composing enough such transformations. The concept of deep learning was officially proposed in 2006[4]. It helps humans to get a big improvement to close to the artificial intelligence community. By using deep learning, computers learn from experience and understand the world based on conceptual hierarchies. Thus, humans don't need to always manipulate computers directly to analyze information. Also, deep learning is advantageous to discover and analyze intricate structures and to apply them in different domains such as medicine and science.

1.2.1. Convolutional Neural Network. Convolutional Neural Network is a Feedforward Neural network with deep structure and convolutional computation. ConvNet is a deep-learning algorithm. There are three main layers compose ConvNets: the convolution layer, full connection layer, and pooling layer.

1.2.1.1. Convolution layer. The convolution layer uses a small matrix to learn image's features. It stores the spatial relationships between each pixels. In figure 1, we can see the operation of convolution. The matrix in which the filter is dragged onto the original image and the convolution operation is performed is called the feature map.

DOI: 10.54254/2755-2721/42/20230770

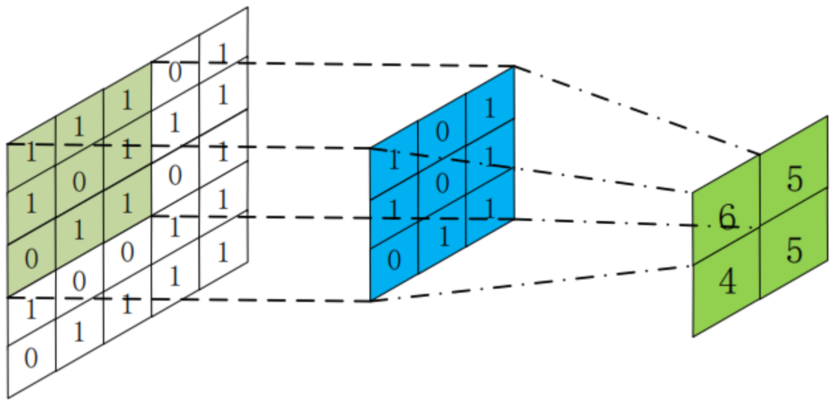

Figure 1. Operation of convolution [5].

In each feature map, the neuron uses the same weight parameter, which is known as filter. The filter is a feature detector for the original input picture. Different filters will create several feature maps to same picture. When the step size is increased, the image specifications continue to decrease. Zero-padding is a way that pad the input to border and an efficient way to give deeper manipulate to the output volumes's dimensionality.

1.2.1.2. Pooling layer. Pooling layers helps to decline the dimensional of the representation. Thus the number of parameters and computational complexity of the model are reduced. The pooling operation is performed by a fixed-size window that move on the feature map according to its step size and then calculates its output. There are two operations called Max pooling and average pooling. Max pooling method gets patches from the input feature maps, outputs the maximum values, and discards all the other values. Another pooling operation is average pooling, which outputs the average of the elements in the pooling window. Using average pooling can aim to reduce the number of learnable parameters. Also, it can help to ensure that CNN accepts inputs of variable size [6].

1.2.1.3. Fully Connected Layer. The fully connected layer is a classifier in the network convolutional neural network. If an operation such as the convolution layer maps the raw data to the hidden feature space, the full join layer maps the learned feature representation to the label space of the sample.In other words, the features are grouped together for the final classifier or regression.

1.3. Generative Adversarial Networks (GAN)

Generative adversarial networks let two neural networks play games with each other to learn. There is a good example that can illustrate the relationship between two neural networks. First, there is a man to fabricate a banknote, and he puts the fake banknote in the cash detector. When the detector finds the counterfeit banknote, it will generate a report that verifies the banknote feature, and this man will base the report on renewing the approach to make fake banknotes until the detector consider the fake banknote is real money. In the model, the generator and discriminator study the data through continuous adversarial games to falsify the data. The two have been perfecting their pictures through comparison in the game. Finally, make the samples obtained by the generator realistic. Meanwhile, the discriminator cannot divide between actual and generated samples. Thus, the result of AI training can reach the Nash equilibrium state.

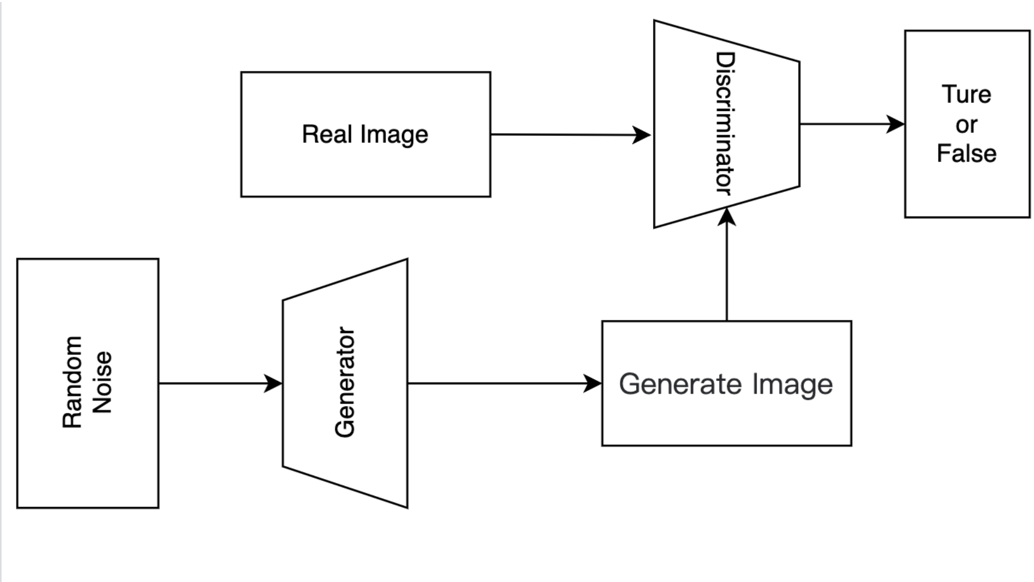

Figure 2. GAN Structure Diagram.

The most direct application of the GAN framework is to configure both the generator(G) and the discriminator(D) as a multi-layer perceptron, as shown in Figure 2.We define an a priori input noise variable p_z (z)and denote the mapping of noise variables to data space as differential function G(z;θg) .We also define a multilayer perceptron D(x;θd) that outputs a single scalar.D(x) represents the probability that x came from the training data instead of being generated by p_g.We train the discriminative model to maximize the difference between samples produced by the generative model G and the training samples, and at the same time train the generative model G to minimize log(1-D(G(z))).

D and G played a two-person maximin game with a value function V(G,D). D must determine the generated data as much as possible, and the gap between the generated data and the real data should be as small as possible. That is to say, in the process of circulation, the discrimination ability of the discriminator D and the ability of the generator G to generate closer to the real data are enhanced at the same time:

$$\min_{G}\max_{D} \mathrm{V(G,D)} = \mathbb{E}_{x \sim p_{data}(x)}[logD(x)] + \mathbb{E}_{z \sim p_z(z)} \log\left(1 - \mathrm{D}\big(\mathrm{G(z)}\big)\right) \tag{1}$$

Among them, $\mathbb{E}$ is the expectation, $p_{data}(x)$ is the real data distribution, and $p_z(z)$ is the distribution that the input of the generator obeys.

For Formula 1, first consider the optimal case of D for a given G. If G is given, then Formula 1 becomes the optimization problem of maximization, such as

$$\begin{aligned}\mathrm{V(G,D)} &= \int_{\mathrm{x}} \mathrm{p_{data}(x)log(D(x))dx} + \int_{\mathrm{z}} \mathrm{p_z(z)log(1-D(g(z)))dz} \\ &= \int_{\mathrm{x}} \mathrm{p_{data}(x)log(D(x))dx} + \mathrm{p_z(z)log(1-D(g(z)))dz}\end{aligned} \tag{2}$$

Further, formula 2 can be solved to obtain the optimal situation of D as:

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} \tag{3}$$

After obtaining the optimal situation of D when G is fixed, Formula 3 can be substituted into Formula 1, and then the solution of Formula 1 can be considered as

$$\min_{G} V(G,D_G^*) = E_{X\sim P_{data}}\left[\frac{p_{data}(x)}{p_{data}(x)+p_g(x)}\right] + E_{X\sim P_g}\left[\frac{p_{data}(x)}{p_{data}(x)+p_g(x)}\right]$$
$$= -2\log(2) + 2\mathrm{JSD}(p_{data} \parallel p_g) \quad (4)$$

Among them, JSD is Jensen–Shannon Divergence, which can quantify the difference between the two distributions. The smaller the value, the closer the two distributions are, and vice versa, the larger the difference between the two distributions. $\mathrm{JSD}(p_{data} \parallel p_g)$ represents the difference in distribution between real and generated data.

From Formula 4 and the meaning of Jensen–Shannon Divergence, it can be seen that the Jensen–Shannon Divergence of the entire network training should be smaller and smaller. By optimizing the objective function, the distribution of generated data is getting closer and closer to the distribution of real data, and there is no need to optimize parameters θ_g and θ_d itself. Until the distribution of real data and generated data is close to or even the same, at this time the value of the objective function is the lowest, and the model is in the optimal state. In actual training, it can be considered that the training of the model has reached the optimal situation.

The ability of Gans to create images lies in their ability to learn the underlying data distribution from the training data set. They learn the patterns and features of real images during training, and then use that knowledge to generate new composite images that resemble real data. The longer and more varied the training process, the more realistic and varied the images produced.

1.4. Application of Generative Adversarial Networks in Image Generation

GAN has powerful modeling ability and plays an important role in image generation. At present, it is applied in various academic research and engineering applications to solve some practical problems.

1.4.1. Image transfer. Image style transfer refers to the transformation of an input image into corresponding images with different styles while preserving the content of the original image. For example, an ink painting can be converted into a sketch. This technique of style transfer is applied in various visual image domains. Among Gans, CGAN uses pair-matched datasets to achieve style transfer but is limited by the scarcity of paired datasets, while CycleGAN can be trained with unpaired datasets. The MAM-CycleGAN method improves the accuracy of image transfer and the visual perception of stylized images. It makes the image retain the feature details as much as possible in the process of restoration and conversion and effectively eliminates or reduces the problem of missing content information in the source domain. The effect diagram is shown in figure 3.

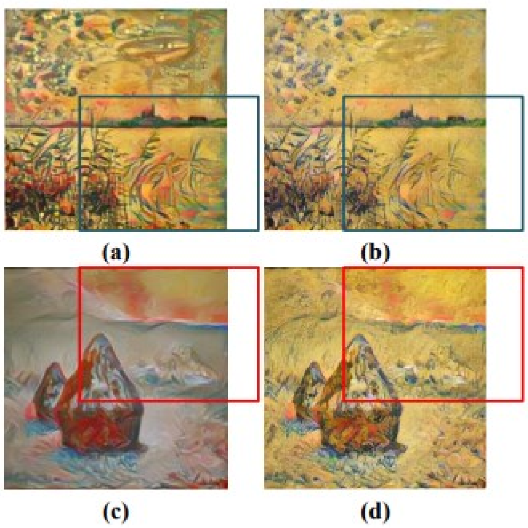

Figure 3. SRM ablation experiment [7].

DOI: 10.54254/2755-2721/42/20230770

1.4.2. Image super resolution. Image super-resolution is an approach of image which low resolution change to high resolution by increasing the number of pixels per inch. In 2017, Christian pioneered the introduction of generative adversarial networks into super-resolution algorithms and proposed SRGAN. When super-resolution reconstruction is carried out under large-scale scaling factors, the high-frequency details of the obtained image are improved, and the bias in visual perception is reduced. ESRGAN is optimized in adversarial loss and perceptual loss respectively, which improves the problem of detail illusion [7]. The Real-ESRGAN model is optimized on the ESRGAN model, which not only introduces the high-order degradation model, but also uses since filter to simulate the common ringing and overshoot phenomena in the image, making the low-resolution image more complete [8,9].

1.4.3. Image restoration. Image inpainting refers to restoring the lost part of the image and then restoring the original image, shown in figure 4. There are two image inpainting methods which use deep learning. Those are the image inpainting method of convolution neural network and the image inpainting method of generative adversarial network. Most of the image inpainting methods of generative adversarial networks are partially improved in the generative network of generative adversarial networks to improve the repair effect. For example, Iizuka et al. proposed an image inpainting method that integrates global semantic and local semantic consistency constraints in the generative adversarial network model [9]. It consists of an inpainting network and two discriminative networks, the global discriminative network inputs the whole repair result, and the local discriminative network inputs the repair result of the missing area. Kamyar et al. proposed a two-stage inpainting network [10]. In the first stage, the model first predicted the restored image texture, and in the second stage, the image was inpainted, which could improve the quality of image inpainting results.

Figure 4. Image restoration [5].

2. User interest modeling

The user interest model is used to describe and represent user interests and preferences. They are widely used in personalized recommendations, advertising targeting, search ranking, and other fields. The user interest model can help the system better understand the user and provide a personalized experience. User interest models can be constructed in a variety of ways. The display pattern is based on the user's explicit behavioral intent, such as the user's search history, favorites, or subscriptions. This pattern can directly reflect the user's interests and preferences but usually requires the active participation of the user. Implicit patterns, on the other hand, are based on the user's implied behaviors and actions, such as browsing history, click behavior, dwell time, etc. This data can be used to infer user interests and preferences by analyzing user behavior patterns and using machine learning algorithms. The implicit model, on the other hand, is to guess the user's personal interest information through the system analysis of the user's main behavior. For example, the system can guess whether a user is interested in something based on past Web services collected by the user and past search and optimization.

2.1. Classification of User Interest Models

In the classification of the user interest model, the individual user interest model and the group user interest model are discussed first.The individual user interest model refers to the model established for each user individually, including user background attributes, interests, behavioral habits, etc., to comprehensively reflect the user's interests. The group user interest model is based on the individual user interest model and developed based on clustering and classification theories. It can generate new user knowledge and can provide accurate information for group users through information push based on user interests in the later stage of the system. In short, the individual user interest model is a model established for individual users to reflect their interests comprehensively. The group user interest model is based on the individual user interest model and developed based on clustering and classification theories to provide accurate information for the group.

There are two categories of information acquisition methods for user interest models: the explicit mode and the implicit mode. Explicit models refer to obtaining user interests by filling out forms or extracting user preferences from user feedback, emphasizing user interaction. While explicit models are relatively easy, excessive user interaction may lead to dissatisfaction with the system. Implicit models refer to obtaining user interests through analyzing their browsing history, weblogs, and other information without requiring direct user interaction. The advantage of implicit models is that they do not interfere with users' normal browsing behavior. However, the disadvantage is that without the ability to identify IP addresses, the obtained interests cannot be accurately attributed. In summary, explicit models involve user interaction to obtain user interests, while implicit models analyze user browsing behavior to gather user interests.

Different user interest models be divided into three categories based on their time duration which is long-term interest models, the short-term interest models, and the period interest model. The long-term interest model refers to a model with long-term stability of user interest, which generally does not change significantly for a long period. This interest model is influenced by the user's fixed social attributes, such as education level, family environment, occupation, and work environment.

The short-term interest model refers to the model in that the user's interest has a short validity period and is easy to change. This model of interest may fluctuate greatly due to changes in an individual's short-term needs, emotions, or psychological factors. The period interest model refers to the model that interest behavior is closely related to a time point it is a kind of long-term interest model. This kind of interest model has strong regularity, so it can be divided into separate classes.

Different interest models are related to each other, and the quality of individual models directly affects the effect of group models. The use of explicit and implicit models, or long-term and short-term models alone, is not a complete capture of interest information. Therefore, the position of the interest model can be specifically determined by object dimension, time dimension, and mode dimension, where the object dimension corresponds to the object-oriented, the time dimension corresponds to the representation time, and the mode dimension corresponds to the acquisition mode.

2.2. Wallpaper recommendation based on user interest model

When a user plays a particular type of game, such as a science fiction game, this information is associated with the user's interest model in the recommendation system. The recommendation system will recommend AI wallpapers related to the game based on the user's game preferences, such as the wallpaper of the game character and the wallpaper of the game scene.

By connecting the user's game preferences and the need for AI wallpapers, the recommendation system can provide wallpaper choices that are more in line with the user's individual interests. Such personalized recommendation can increase user satisfaction and improve user experience of the recommendation system. These personalized recommendations are provided by the implicit pattern of the user interest model.

The implicit mode of user interest Model is to use VSM(Vector Space Model) to describe the user's feature information as a document, and convert the document information matching into a vector matching problem in vector space. Before judging the similarity between a document and a user's features, we must first extract the feature vector, define a document as a combination of a series of keywords, and each word is assigned a weight. Measure the similarity between a document vector representation and the user's interest features, that is, determine whether a document is in line with the user's interests. Since the user interest feature can also be represented as a user feature vector in the same space, it is usually represented by calculating the inner product or the cosine of the Angle between two vectors. The web service discovery model based on user interest mainly analyzes the correlation between web service description document and user feature document, and obtains the result that satisfies the user's interest characteristics. This includes three aspects of work.

（1）Document vector representation:Let document D be a collection of n WEB documents, D={d1,... ,di... dn}, i=1,2,... ,n, any document di in a document set can be represented as an M-dimensional vector di= (ti,... tij... tim), i=1,2,... ,n; J = 1, 2,... Where tij is the JTH feature keyword component of document di. The representation of document vector in this system is mainly based on the feature extraction of tf/idf (document frequency/anti-document frequency) : Dij=tij*log2(n/nj), where tij is the inversion frequency of the keyword ti in the document, n is the total number of documents in the document database, and nj is the number of documents containing the keyword tj in the document database.

（2）User feature representation:The description of user interest information is expressed in natural language, and the user feature vector can be represented by the method of representing document vector: P= ((y1,u1),(y2,u2),... ,(ym,um)), where (yi,ui) represents the weight ui of the subitem yi.

（3）Calculation of similarity coefficient between document and user feature.

$$\sin(D,P) = \cos\theta = \frac{\sum_{i=1}^{m} t_i \times p_i}{\sqrt{((\sum_{i=1}^{m} t_i^2)(\sum_{i=1}^{m} p_i^2))}} \tag{5}$$

Suppose D= (t1,... ,ti... ,tm) represents the document vector, P=(p1,p2,... pm) represents the user feature vector, the calculation of the similarity of the two vectors can carry out cosine normalization of the two vectors, and then use the inner product formula to calculate the cosine of the Angle between the two vectors. When the two vectors are the same, the similarity is 0. When the vectors do not have the same term, the greater the cosine value of the Angle between the two vectors, the more the search WEB service meets the needs of users.

（4）Determination of relative range: The user sets a correlation threshold, and the returned search results only contain documents above this threshold, that is, given a value user feature vector P and a correlation stop value R, it must be relevant for any returned document D:Sim(D,P)>R.

3. Conclusion

This article is mainly divided into two parts AI technique and User interest modeling. AI user preferences create wallpapers through GAN. Meanwhile, User interest modeling will recommend wallpapers according to user preferences. JSD plays an important role in GAN model. For instance, JSD can be used as an objective indicator to evaluate the quality of AI-generated images. A low JSD indicates that the distribution of AI-generated images is very similar to the distribution of real images, indicating that the generated images have high quality and realism.

Therefore, GAN has many functions in AI. GANs can upscale low-resolution images to higher resolution, a process called image super-resolution. By training on pairs of low-resolution and high-resolution images, the generator can produce a sharp and detailed version of the low-resolution action input image. Also, GANs can be used for image restoration, filling missing or damaged parts of an image with reasonable content. The generator learns to complete missing regions based on the surrounding context and patterns in the image.Meanwhile, User interest modeling plays an important role in AI

DOI: 10.54254/2755-2721/42/20230770

pushing wallpapers. Basing user's game record, User interest modeling will share the favorite wallpaper to user.

But Gans and User Interest Modeling can also cause a lot of social issues. For example,when applied to personalized recommendations, it involves collecting and analyzing user data to understand their preferences and behaviors. This may raise privacy concerns, as users may feel uncomfortable with their personal information being used without their explicit consent or being used to manipulate their choices. Moreover, Gans can be used to create content that resembles a copyrighted work, raising questions about the intellectual property and ownership of the generated content.

To address these societal issues, it is critical to implement responsible and ethical practices in the development and deployment of Gans and user interest modeling systems. This includes transparent data collection and use, fair sense algorithms, user control over data, and ongoing monitoring of bias and discrimination. In addition, industry standards, regulations, and public awareness campaigns can help mitigate potential negative impacts and promote responsible use of these technologies.

Acknowledgement
Zicheng Liu, Junjie Deng and Jiacheng Wang contributed equally to this work and should be considered co-second authors.

References

[1] Sallam, Malik. (2023). "ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns." Healthcare. Vol. 11. No. 6. MDPI.
[2] Jeff , Hancock. (2023). “How Will CHATGPT Change the Way We Think and Work?” Stanford News, news.stanford.edu/2023/02/13/will-chatgpt-change-way-think-work/.
[3] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. (2015)."Deep learning." nature 521.7553: 436-444.
[4] Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. (2006). "A fast learning algorithm for deep belief nets." Neural computation 18.7: 1527-1554.
[5] O'Shea, Keiron, and Ryan Nash. (2015)."An introduction to convolutional neural networks." arXiv preprint arXiv:1511.08458.
[6] Song. (2023)."Research on Style Transfer of Thangka Images Based on Generative Adversarial Networks". Northwest Minzu University, MA thesis.
[7] Wang, Xintao, et al. (2018)."Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European conference on computer vision (ECCV) workshops.
[8] Wang, Xintao, et al. (2022)."Realesrgan: Training real-world blind super-resolution with pure synthetic data supplementary material." Computer Vision Foundation open access 1: 2.
[9] Iizuka, Satoshi, Edgar Simo-Serra, and Hiroshi Ishikawa. (2017)."Globally and locally consistent image completion." ACM Transactions on Graphics (ToG) 36.4: 1-14.
[10] Nazeri, Kamyar, et al. (2019)."Edgeconnect: Generative image inpainting with adversarial edge learning." arXiv preprint arXiv:1901.00212.