Operational Research in Engineering Sciences: Theory and Applications
Vol. 7, Issue 1, 2024, pp. 158-175
ISSN: 2620-1607
eISSN: 2620-1747
DOI: https://doi.org/10.31181/oresta/070108

AI-DRIVEN BACKGROUND GENERATION FOR MANGA ILLUSTRATIONS: A DEEP GENERATIVE MODEL APPROACH

Zhang Yunqian[1,2*]

[1]Phd Candidate, Animation Major, Department of Performance, Film and Animation, Sejong University, Seoul, South Korea,05006.
[2]Research Interns, Department of General Education, Guangzhou Panyu Polytechnic , Guangzhou City, Guangdong Province 511483, China.

Received: 19 November, 2023
Accepted: 11 March 2024
First Online: 28 March 2024

Research Paper

Abstract: *This paper introduces the generation technique of manga illustration background, discusses the traditional background generation technique and stylized migration technique, and points out that the application of AI technology provides new possibilities for the creation of manga illustration backgrounds. Aiming at the limitations of traditional methods, the design concept and principle of conditional generative model based on deep learning are proposed, the implementation principles of convolutional neural network and generative adversarial network are introduced, and the conditional manga illustration background generation model based on deep learning is proposed by combining the two. The paper uses MindSpore software to train CNN models.CNN models are very good for processing data such as images and are able to reduce the number of parameters in the model and increase the speed of training the model.The model data has a wide range of applications and is capable of removing the influence of character factors from the image while providing high resolution.The model can effectively realize the conditional generation of manga illustration background. The practical effect of this technology is demonstrated through a case study and technical exploration of manga illustration background generation technology, which demonstrates the potential of the new technology and its application in creating richer and more vivid comic backgrounds. In the future, with the continuous improvement and promotion of this technology, more similar cases are expected to be seen, bringing more possibilities and innovations to the creation of comics and illustrations.*

Keywords: *Artificial Intelligence, Manga illustration, Background, Modelling, Deep learning*

*Corresponding Author: jiuxia00@163.com (Yunqian .Z)

1. Introduction

Comics are cultural and artistic products that use drawing as a means to reflect various things and information as well as human thoughts, ideas, and emotions (Grimaldi & Ehrler, 2023; Sanchez, 2023). The creation of comics is based on hand drawing and computerized drawing, and it is mainly disseminated through printed materials such as newspapers, magazines, books, and online media. In Figure 1, the proportion of comics in the cultural industry in five major cities in China from 2016 to 2020 is shown. Comics belong to the publication of "one content, multiple forms of expression" and have a strong pulling effect on the animation industry, game industry, symbolic image industry, and copyright-related industries (Ananthaswamy, 2023; Anthony et al., 2023). From the perspective of industrial operation, comics are the foundation and source of cultural and creative industries, and a huge industrial cluster can be formed by extending them with copyright as a link (Perkins, 2023; Zhuang et al., 2023).

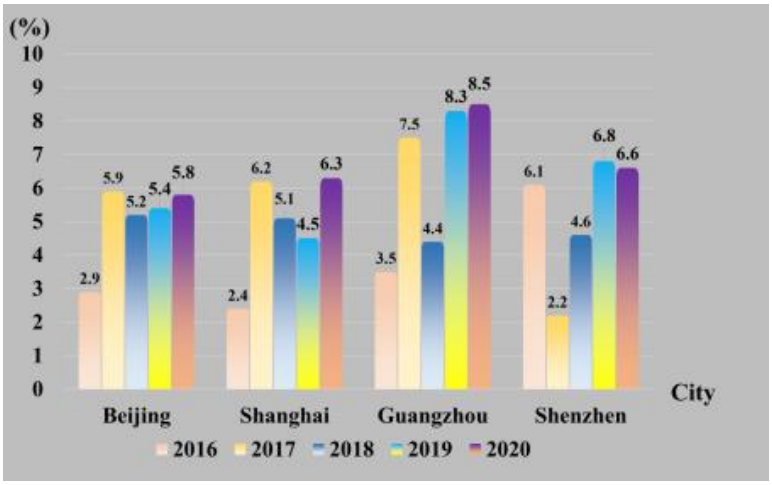

Figure 1 Proportion of the comic industry in the cultural industry in China's first-tier cities (Ananthaswamy, 2023; Anthony et al., 2023).

Generally speaking, a complete manga mainly consists of elements such as manga frames, dialog bubbles, illustrations, manga layout, and backgrounds (Gheisari et al., 2023; Zhou et al., 2022). Among these elements, the background in comics is usually used to depict the scene and environment in order to provide more visual information and atmosphere for the storyline (Zhan et al., 2023).

Manga illustration background is a two-dimensional visual static graphic art. In comics, the manga illustration background is used to interpret or decorate the plot of the picture. It can be used not only to complement the plot and dialog of a comic but also to convey emotion, atmosphere, and storyline through the images in order to enhance the presentation of the story (Abdelmigid et al., 2023; Yan et al., 2023). Manga illustration can also complete a narrative and can convey a complete story or concept. The composition (e.g., use of colour and line) of a manga illustration background is generally more streamlined than that of a fine art drawing (e.g., oil painting, watercolour), and text (e.g., dialogue, asides, onomatopoeia) can be added to form part

of the content (Ananthaswamy, 2023; Dhar et al., 2023; Grimaldi & Ehrler, 2023; Ma et al., 2022; Zhou et al., 2022). Manga illustration backgrounds are an integral part of comic creation.

In traditional comic production, the traditional background requires multiple processes to be carried out, which are mainly divided into two parts. The first part is the background design work. Creators need to design the script subplot according to the script situation combined with the story time, the character traits of the story characters, the storyline, the presentation form, and other requirements, and then enter the second phase of the drawing stage after determining that there is no error after the discussion in the meeting. In the drawing stage, there are traditional manual paper drawings and computer digital drawings two ways (Ananthaswamy, 2023; Chui et al., 2023; Li et al., 2023; Zhan et al., 2023; Zhuang et al., 2023). In the process of drawing the scripted subplot into a finished product, in addition to the need to refine the content of the script, it is also necessary to add details, such as light and shadow effects, simple, special effects, and details of flowers and plants. The creation of manga illustration backgrounds usually takes a lot of time, labour, and money. Usually, manga illustration backgrounds change with the storyline and have a low reuse rate (Huang et al., 2023; Yan et al., 2023).

In the era of rapid development of Internet technology, the mode of dissemination of manga illustrations is not limited to traditional printing and publishing but more in the form of electronic comics on the Internet. Readers prefer to attach corresponding background sound effects and appropriate dynamic backgrounds in the process of reading electronic comics. The traditional way of creating comics can no longer satisfy this kind of demand (Chui et al., 2023; Gheisari et al., 2023; Nugraha & Sutanto, 2020; Rahman et al., 2023; Sherratt et al., 2023).

Improving the production efficiency of manga illustration backgrounds, saving production costs, improving the utilization rate of manga illustration backgrounds, and enriching the dynamic elements of manga illustration backgrounds are urgent problems that the comic industry needs to solve. Fortunately, the rapid development of AI provides the possibility for the innovation of manga illustration background production (Ananthaswamy, 2023; Gheisari et al., 2023; Huang et al., 2023; Zhuang et al., 2023).

This article takes the manga illustration background as the object and discusses the application of AI technology in the production of manga illustration backgrounds. The article aims to improve the quality of background generation with deep generative modelling technology in AI. It aims to provide more possibilities and innovativeness for manga illustration background creation.

2. Related works

2.1 Existing background generation techniques

The techniques in traditional background manga illustration background generation are mainly two kinds of background generation based on traditional image processing techniques and stylized migration techniques.

Table 1: Comparison of comics with other media forms(Mukhutdinov et al., 2023; Taye, 2023)

Media Forms	Representation of visual information	Representation of visual information	Continuity of the picture
Japanese Manga	Cue in with a drawing	Expressing sounds in words or onomatopoeia	Time flows in each frame.
sculpture	Cue in with a drawing	not have	state of not working
picture book	Use paintings to prompt, plus article descriptions	Expressing sounds in words, hieroglyphs, and even essays	Static, illustrative form
fiction	Article descriptions (some will be illustrated)	Fiction article illustrations (some will be illustrated) express sound in words and articles, not pictorial graphics	None (or with static illustrations)
radio drama	No picture. Sound description.	Direct cues with sounds and soundtracks	not have
video	Directly prompted by a video screen	Direct cues with sounds and soundtracks	Dynamic, real-time flow

In Table 1, a comparison between comics and other media forms is given, from which it can be seen that comics are a static medium with temporal fluidity and are mainly presented in the form of drawings. Illustrations in comics are drawings used to decorate or supplement the storyline (Mukhutdinov et al., 2023; Taye, 2023). They can appear outside of dialog boxes or as whole-page drawings. There is no clear standard of distinction between manga illustrations and backgrounds in existing manga creations. Creators generally usually regard them as the same thing (Chui et al., 2023; Sun et al., 2023).

2.1.1 Based on traditional image processing techniques

Background generation based on traditional image processing techniques is the use of image processing techniques and image synthesis techniques to generate new backgrounds for manga illustrations by processing and synthesizing existing images (Mukhutdinov et al., 2023; Rahman et al., 2023). This may include synthesizing different elements (e.g., sky, ground, buildings, etc.) together to create new scenes. In Fig. 2, the principle of the traditional picture image reorganization technique is given, where the original data information is extracted and reorganized into a new picture by modelling (Li et al., 2023; Taye, 2023). Of course, the method can also apply various filters and special effects to change the appearance and style of the image to achieve the effect of a manga illustration background. For example, effects such as Gaussian blur, line strokes, fragmentation, exposure, etc., can be used to enhance the artistic sense and expressiveness of the image. In traditional image processing techniques, hand drawing (traditional drawing tools and digital painting software) is also required to create background elements in order to ensure that they are rich (Grimaldi & Ehrler, 2023; Zhou et al., 2022).

Traditional image processing techniques are relatively mature, and many creators are proficient in them. Using this technology, creators can personalize and create

unique manga illustration background effects according to specific needs (Li et al., 2023; Sun et al., 2023).

However, this technique still has not transcended the traditional way of making comic backgrounds, and it is difficult to create highly personalized and innovative background effects. Drawing or complex compositing requires a lot of time and effort, and the generated backgrounds may appear to be more traditional and conventional in comparison (Zhuang et al., 2023).

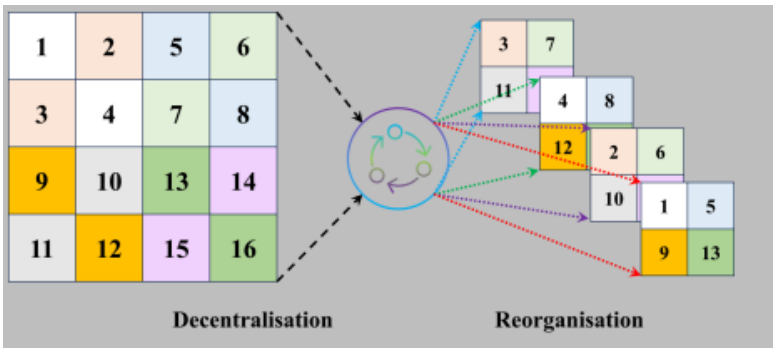

Figure 2: Conventional image processing techniques (Li et al., 2023; Taye, 2023)

2.2 Stylized migration techniques

The stylized migration technique is a deep learning-based image processing method. This technique is based on the application of basic AI techniques and is mainly based on the structure of a fully convolutional neural network with perceptual loss (Abdelmigid et al., 2023; Yan et al., 2023). The stylized migration technique essentially transfers the visual style of one image to another image, thus generating images with different styles. For example, converting a photographic photograph into an exaggerated secondary style or converting an image of a landscape figure into an ink style. In manga illustration background generation, the style migration technique can be used to apply images with different styles to the background generation, for example, converting a real-world photo style into a manga illustration style. A demonstration case of the stylized migration effect is given in Figure 3, from which it can be seen that the existing stylized migration effect is closer to the traditional filter effect and also provides simple background modification.

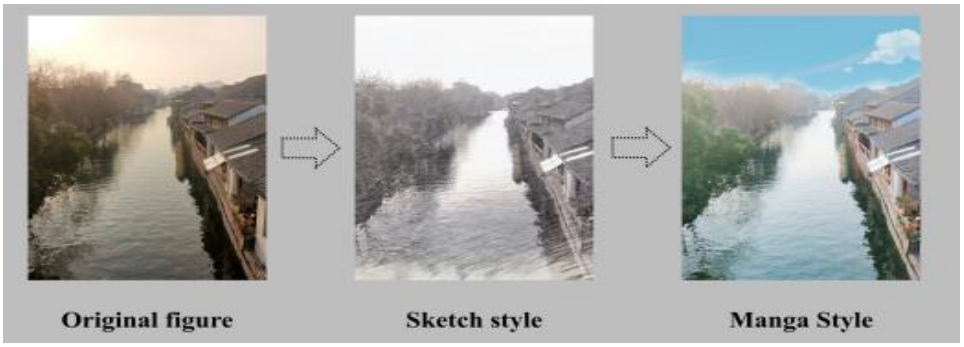

Figure 3: Stylized migration effect display (Anthony et al., 2023; Zhou et al., 2022)

The principle of stylized migration is given in Fig. 4. The principle can be briefly described as several links, such as image input, engine processing, style division, data iteration, and style output. In general, the stylized migration technique has certain application potential in manga illustration background generation; however, due to its insufficient depth of AI technology mining. Sometimes, the results of style migration may not be accurate enough, especially for complex manga illustration styles, which may not be able to fully meet the expectations, resulting in a certain difference between the generated background and the original style. At the same time, the technical depth is insufficient, and a large amount of computational resources and time are usually required in the style migration process, especially for high-resolution image processing.

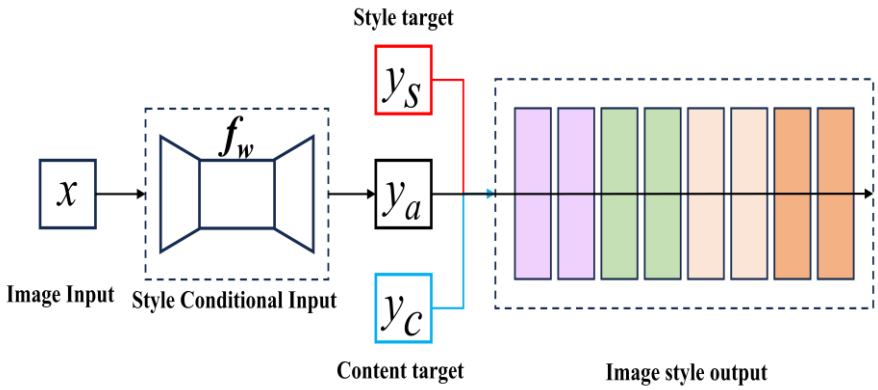

Figure 4: Principle of stylized migration (Nugraha & Sutanto, 2020; Zhan et al., 2023)

2.2 AI technology brings change.

If the transition of the manga illustration background from traditional hand drawing to digital drawing represents the progress of the production process, then the transition of manga illustration background generation technology from traditional image processing technology to stylized migration technology application is the beginning of technological innovation (Abdelmigid et al., 2023; Wenjun et al., 2023).

The rapid development of AI technology in recent years has brought unprecedented changes to various industries. In Table 2, the AI industry application map is given, and the comic production industry is also deeply affected by it. Intelligence, efficiency, and precision are the three main features of AI. In the field of manga illustration background production, stylized migration technology has initially used AI technology (Chui et al., 2023; Li et al., 2023) . Digging deep into AI technology can bring significant changes to the production of manga illustration backgrounds.

AI technology is able to develop modular systems through techniques such as deep learning and generative adversarial networks to enable more creative and personalized manga illustration background generation. Creators can use the system to generate backgrounds of various styles and themes, thus realizing more diverse creations.

Table 2: AI industry application landscape (Chui et al., 2023; Li et al., 2023)

	Sell individually or in small quantities	Financial	Medical care	Teach	Fabrication	Renewable energy	Movies and television
Research and development		√			√		
Give birth to a child.					√	√	√
Supply chains	√				√	√	
Market	√	√	√	√			
First aid	√	√	√	√			
Wind control		√					
Surety	√	√	√	√	√	√	

AI technology can speed up the process of creating backgrounds for manga illustrations and reduce the heavy workload of illustrators. By automating the generation of backgrounds, illustrators can focus more on storytelling and character design, reducing costs (Sun et al., 2023; Zhuang et al., 2023). At the same time, it is conducive to the rapid creation and release of comic works, enriching the industry and further promoting the innovation of comic production.

In this paper, a deep learning-based conditional generation model system for manga illustration backgrounds is produced by combining two deep learning models, a convolutional neural network and a generative adversarial network, with AI's deep learning of big data. Creators can generate corresponding manga illustration backgrounds in the form of descriptions or labels. For continuous comic creation, it is also possible to change part of the background by changing the local description.

3. Methodology

3.1 Design Concepts for Conditional Generative Models Based on Deep Learning

The background of a manga illustration requires rich details, the concept of rapid generation, and sometimes some customization, as well as a filter migration effect by modifying the relevant parameters. In order to fulfil these conditions and generate the appropriate system, the scientists coordinated with the comic background illustrator. A list of artistic, procedural, and intelligent principles that are required in the process of generating comic backgrounds have been outlined. At the same time, some scholars have also conducted research, and these principles are summarised as follows (Rahman et al., 2023; Raja & Kannimuthu, 2023).

1. Ability to learn the characteristics and context of the background of a manga illustration. Models need to be able to learn about the features of the context of a manga illustration, such as the characteristics of different scenes, colours, and compositional styles, and to understand how these features relate to the plot and content in the illustration.
2. Conditional generation models can be utilized. Creators can generate conditional background images by entering plot descriptions, dialog content, special requirements, etc.
3. Diversity and creativity can be considered. In generating the background of manga illustrations, the model needs to have a certain degree of diversity and creativity so that the generated background not only meets the conditions but also has a certain degree of change and innovation so that the background of the comic is more rich.
4. Achieving controllability and interpretability. Conditional generation models should have a degree of control and regulation; that is, after the model has been generated, it is possible for the user to know how the model has generated a particular image according to the conditions, and similarly, the user can change some of the conditions and thus refine the background image.
5. Combining reinforcement learning and transfer learning to improve background adaptation. Techniques such as combining reinforcement learning and transfer learning can be considered to improve the model's adaptability and effectiveness for the task of generating backgrounds for comic book illustrations so that it can better handle complex scenarios and details.

Based on the above design concept, a deep learning model specialized in background generation for manga illustration can be constructed, which can combine contextual information and conditions to generate diverse and creative background images that meet the conditions, thus assisting the comic creation process.

3.2 Principles of Conditional Generative Modeling Based on Deep Learning

In the above article, the design concept of a conditional generative model based on deep learning is described, and its main technical principle lies in the ability to learn the characteristics and context of the background of manga illustrations and the conditional generative model in two parts.

For the ability to characterize the background and context of the manga illustration is trained using convolutional neural network technique to extract image features and contextual information. For conditional generative models, conditional generative adversarial networks can be used for training.

3.3 Convolutional neural network

Convolutional neural networks are a class of deep learning neural networks used to process and analyze data with grid-like data structures. The structure of convolutional neural consists of four main parts, which are convolutional layer structure, fully connected layer structure, pooling layer structure and activation function structure. In the whole process of convolutional neural network operation,

the four parts play different roles. Among them, the convolutional layer structure acts like a data pool to extract data information and features from the input image. The role of the pooling layer structure is to dimensionalise the feature information of the picture to make it as low dimensional as possible for easy data transfer and classification. The fully connected layer structure then classifies the low dimensional data processed by the pooling layer (Ruan et al., 2023; Yan et al., 2023).

In a convolutional neural network, a convolutional operation is the operation of multiplying and then adding an element-by-element window of moveable data with an image. This small window is actually a fixed set of weights, which can be thought of as a specific filter (fter) or convolution kernel. The name of this operation, ""convolution,"" derives from this process of element-by-element multiplication and summation.

The core idea of CNN (Convolutional Neural Network) is to simplify complex problems, the source data with a large number of parameters will be dimensionalised into a small number of parameters with core value after convolutional mapping into the feature space. Moreover, in most of the scenarios, dimensionality reduction does not affect the results. Simple datamining of images can't preserve image features, but CNN preserves image features in a vision-like way, and it can effectively identify that it is a similar image when the image does flip, rotate or change position. Based on these two features it perfectly solves the two challenges of image recognition.

The principle of the convolutional neural network is given in Figure 5. The convolutional layer, as an important constituent structure in the convolutional neural network, utilizes a number of convolutional kernels that are much smaller than the size of the input image to carry out local, sparse operations in space, which can be viewed as the process of window-sliding computation of the image, and the weights of the convolutional kernels are shared, which can effectively save the amount of computation and storage space (Sherratt et al., 2023; Wenjun et al., 2023). The pooling layer is the operation of down sampling the convolution output feature map, the image after the convolution operation, the feature map obtained due to the high dimensionality, the amount of computation is too large, and it is easy to overfitting, pooling operation can reduce the dimension of the feature map, and reduce the overfitting situation. The use of nonlinear units (activation function) can realize the nonlinear mapping of the linear convolutional output to improve the ability of the network feature expression. Common activation functions include ReLU, Sigmoid, etc., and there are some variants, such as Leaky ReLU, ELU, etc., which can be used to choose different activation functions for different task requirements. The general convolutional neural network also contains a fully connected layer, which is located at the end of the whole convolutional neural network and is used to map the learned feature vector representations into the sample labelling space to achieve classification or regression for visual tasks.

The design of convolutional neural networks is inspired by the biological visual system, and its structure and parameter-sharing properties allow it to perform well in image processing and pattern recognition tasks.

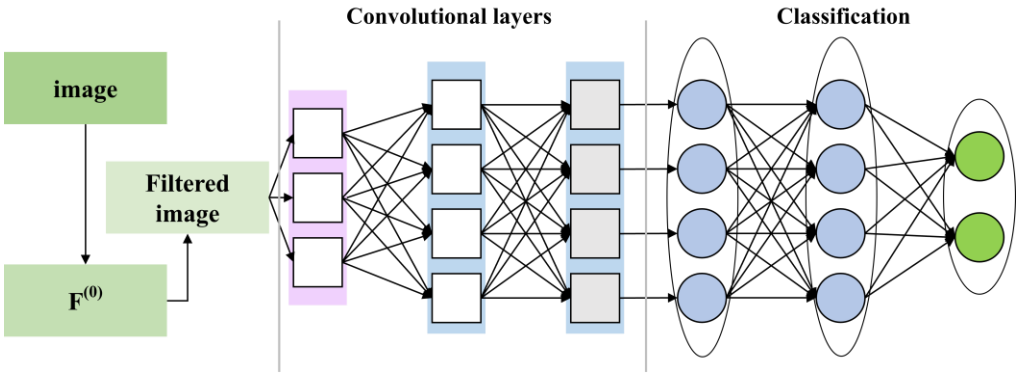

Figure 5 Principle of convolutional neural network (Sherratt et al., 2023; Wenjun et al., 2023).

3.4 Generating Adversarial Networks

Generative Adversarial Network is a deep learning model. The core idea of the generative adversarial network is to learn and improve together by letting two neural networks play against each other. A schematic diagram of the concept of a generative adversarial network is given in Fig. 6, from which it can be seen that a generative adversarial network consists of two neural networks that play against each other: generator G and discriminator (Nugraha & Sutanto, 2020; Rahman et al., 2023).

In this, the generator G is learning the distribution of real data, and the discriminator G is a classifier to discriminate whether the input data is real data or generated data. Then sampling is done from the hypothetical hidden space z, and after passing through the generative model G, the data x' = G(z) is generated. The real data is then fed into the discriminative model D along with the generated data to output the determined categories.

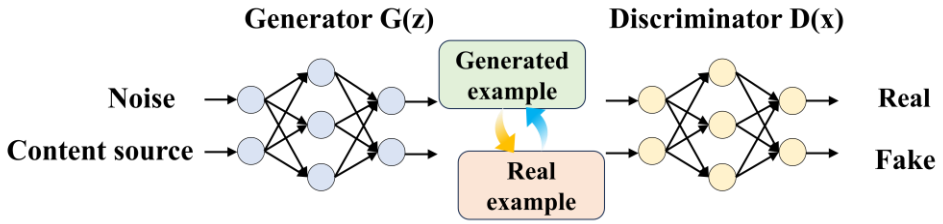

Figure 6 Schematic diagram of the concept of generating adversarial networks (Nugraha & Sutanto, 2020; Rahman et al., 2023)

The training process of generative adversarial networks consists of two main parts: adversarial training between the generator and the discriminator (Jiang et al., 2023; Yan et al., 2023). The process of generative adversarial network training is shown in Figure 7. From the figure, it can be seen that the generated data is gradually close to the real data distribution after broken discriminator iterations.

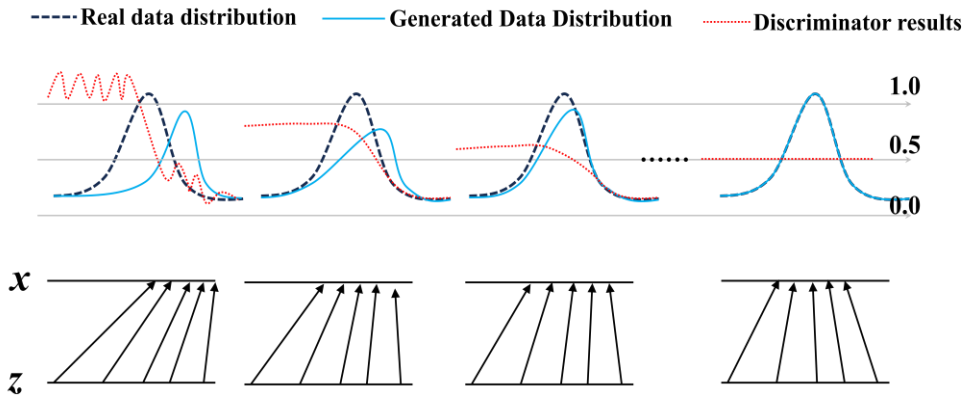

Figure 7 Generating adversarial network training process (Chui et al., 2023; Kim et al., 2023)

In Fig. 8 the framework of image generation based on adversarial ideas is given. From the figure, it can be seen that a generative adversarial network is a process of repeated training, iterative iteration, and eventually convergence to the real target. Throughout the iterative training process, the generator receives a random noise vector as input (also known as the background generation condition) and generates an image. The generated image is then transported along with the real image to the discriminator for discriminative confrontation.

Throughout the process, the generator expects the generated image to "fool or pass" the discriminator. The discriminator, on the other hand, correctly distinguishes the difference between the generated image and the input conditions according to the requirements of the goal. The two are constantly learning against each other, and the generator iteratively updates the generated image to optimize it. By iteratively updating the parameters, the generator gradually improves its ability to generate realistic images, while the discriminator also improves its ability to correctly identify authenticity.

Ultimately, the goal of the discriminator and generator is to find a balance where the generator is able to produce an image that meets the input conditions and the discriminator is able to correctly distinguish between the generated image and the true image.

In the workflow of GAN. Given a set of target samples, the generator will try to generate artificial samples that can trick the discriminator into seeing them as real target samples, achieving the goal of 'fake the real'. In turn, the discriminator will try to distinguish the real (target) samples from the fake (generated) samples. Through this cyclic training method, we can eventually get a generator that can generate samples similar to the target sample very well.

Compared with other models, GAN model has the characteristic of modelling and learning for all types of data distribution, and it has a wide range of application scenarios. Its basic principle is based on the three stages of "confrontation", "learning" and "correction", which can effectively reduce the data model error and infinitely close to the real data. At the same time, for the entire process of comic background

generation, GAN can be used to eliminate the human influence in the image, super-resolution, gesture migration, as well as any type of image transformation. GAN model has an unparalleled advantage.

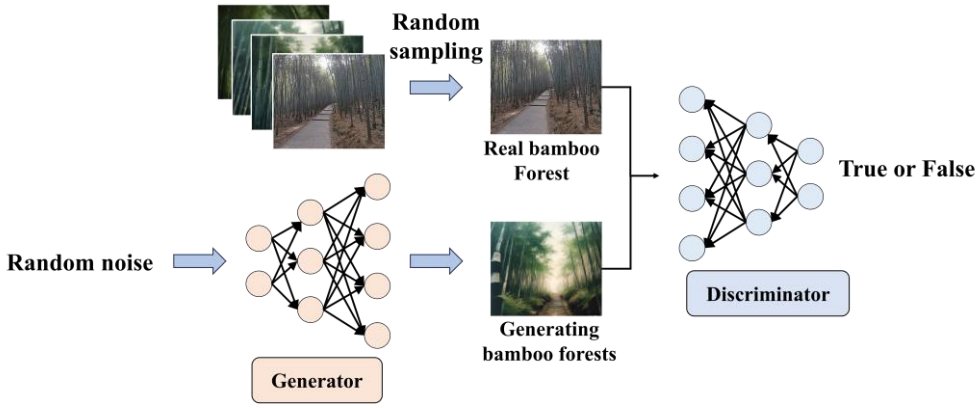

Figure 8 Image generation framework based on the idea of image confrontation (Abdelmigid et al., 2023; Yan et al., 2023)

3.5 Deep Learning-Based Background Generation Model for Conditional Manga Illustrations

Convolutional Neural Networks and Generative Adversarial Networks are two different types of neural networks, both of which are deep learning models that use a multi-layer neural network structure for learning and training.

In the process of generating backgrounds for manga illustrations, it is hoped that it needs to be realized by inputting specific parameters and descriptions. Among these two neural networks, convolutional neural is mainly used for image recognition, classification, and processing, which can be used to achieve accurate recognition of conditions through deep learning. When the data parameters are trained, the generative adversarial network focuses on generating image data close to the input conditions through user input. The combination of the two, the convolutional neural network through the rich big data deep learning training, can provide rich, accurate image classification and recognition data for the generative adversarial network, the generative adversarial network through the adversarial recognition of the image data generated closer to the generation of the conditions of the target.

The basic principle of combining the two is given in Fig. 9. As can be seen from the figure, the generative adversarial network can be used as a component of the convolutional network. The Generative Adversarial Network acts as a "discriminator" for the convolutional neural network and is responsible for determining the authenticity of the generated images. This combination allows the Generative Adversarial Network to produce more realistic images.

Convolutional neural networks can also be used as part of the "generator" component of a generative adversarial network. In this case, the adversarial network can be allowed to learn in-depth how to generate realistic images from random noise

and provide rich and accurate conditional generative data to the adversarial network.The two can each other and can achieve more powerful image generation, recognition, and processing capabilities.

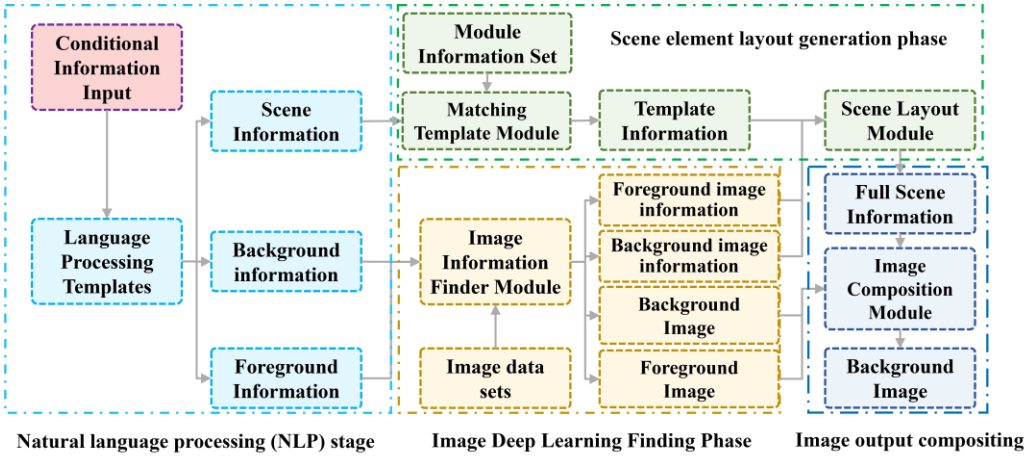

Figure 9 Flowchart of deep learning-based conditional manga illustration background generation model (Kim et al., 2023; Ma et al., 2022)

4. Results and discussion

4.1 Manga Illustration Background Generation Technology Case

According to the previous section, it is known that the background generation of comic illustrations relies on specific textual description information. As shown in Fig. 10, background generation is needed for two sketched cats. Firstly, the characteristic text information is given: sky, grass, colour, etc. After the system recognises the self-recognition language and processes it, it starts to look for the scene information in the database. Then according to the information content input by the user automatically discern the ease of generating the content gradually generates the relevant information. In Figure 10, the system judges that it is easiest to generate the colourless sky, and then generates the colourless sky data first.

In addition, image information, such as the size of the image and the content information it contains, needs to be extracted. In the scene layout stage, in addition to using the scene information output from the natural language processing stage, it is necessary to continue to supplement the scene information by defining templates and matching templates, as well as determining the position and size information of each image in the new image. In addition to this, the relationship between the foreground objects and the background area, as well as the positional relationship between the objects, need to be controlled to ensure that the final arrangement meets expectations. In the image synthesis stage, the discontinuity of the luminance transformations due to changes in illumination needs to be addressed. Image fusion methods can be used to ensure that the final synthesized result is a seamless image. Ultimately, the organic combination of these steps can help realize realistic and creative manga illustration background generation.

The background of the Manga illustration of a Manga cat running in the grass is given in Fig. 10. From the figure; it can be seen that the model in this paper is not only able to generate black and white line drawings but also able to realize the coloring function.

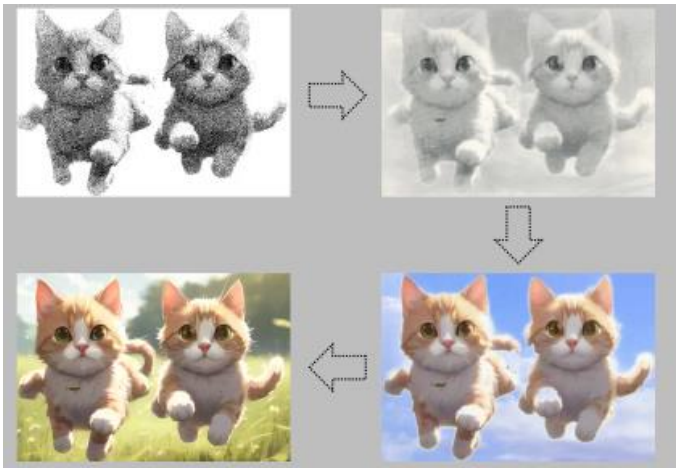

Figure 10 Manga illustration background generation display(Chen et al., 2023; Yan et al., 2023)

4.2 Exploring Background Generation Techniques for Manga Illustrations

4.2.1 Opportunities and challenges

Manga illustration background generation technology utilizes AI and computer vision technology to automate the generation of background scenes in Manga and illustrations by processing natural language descriptions and image input.

This technique has been explored in a variety of fields, such as natural language processing, image processing, computer graphics, and machine learning. Through in-depth research and continuous practice, scientists are committed to improving algorithms and models to enhance the quality and efficiency of background generation. Eventually, these efforts will hopefully lead to a more realistic and creative background generation technique for manga illustrations, bringing new possibilities for comic art and virtual scene creation.

The manga illustration background generation technique has some challenges and drawbacks at the current stage.

In Table 3 the analytical statistics of the comic illustration background generation technique proposed in this paper is presented, from the table it can be seen that the accuracy of the model is 76% accurate, 65% diverse, and 40% realistic as compared to the conditional equation entered. The model is to be further optimised.

Table 3 Model reliability statistics

Accuracy	Diversity	Authenticity
76%	65%	40%

5. Copyright and Ethics

AI technology is capable of reading large amounts of data for training and screening, and some cultural and copyright issues are often overlooked. Different countries and regions have different cultures and customs. Resulting in the same input conditions, some taboos may be violated, as well as copyright issues. There is no effective solution to these problems. Some scholars have set up proprietary filtering systems for each region to apply local customs. However, in some countries, customs vary from region to region, so there is no unanimous conclusion on how to effectively solve such problems. This requires more thought and application.

6. Research limitations and future research directions

6.1 Research limitations

AI technology is in its infancy, and the applications in various industries are not perfect enough to have enough supporting resources to refer to. This leads to misidentification of information during the generation process, resulting in inaccurate image content. In addition to the technology lacks modular input procedures to help the user to provide enough information for generation. Insufficient parameters provided by the user or the image information leads to a wide variety of generated content. In addition, the development of hardware is also an important constraint of the technology. Currently it takes a lot of time for the user to generate the appropriate information, and hardware with higher configurations has a higher price tag. The high price leads to many companies are not willing to make the cost investment. The last thing is that the need for technology in comic illustration background generation is not widely demanded. On the one hand, it is because the technology is not perfect at this stage, the node of industrial change has not yet arrived, on the other hand, it is because people are used to the traditional way, not willing to try to learn the emerging technology. The maturity and development of comic illustration background generation technology still need some time.

6.2 Future research directions

In future research, we will aim to optimise the data for the model to provide a larger sample size. Provide more generative conditional data, while using the model for more commercial scenarios. Only by continuously enriching the model and expanding the sample size can we make up for the shortcomings of the model and make the model more perfect.

7. Conclusion

This paper focuses on the application and development of background generation techniques for comic book illustration. The traditional background generation techniques are discussed. For the application and development of AI technology at this stage, a conditional comic illustration background generation model based on deep learning is proposed. The model consists of a combination of both convolutional neural network and generative adversarial network, which can effectively achieve the

conditional generation of comic illustration background. The paper demonstrates the practical use of the model through the case of generative techniques. The model can generate conditional backgrounds and provide colouring functions.

At the same time the background generation technique has some limitations. It is difficult to recognise complex information and lacks guidance for the user to enter comprehensive information. For complex conditions, the generated backgrounds are not effective. Due to the limitations of hardware and technology, it is more difficult to promote and apply the model. In the future, with the continuous improvement and promotion of this technology, it is expected to see more similar new models and cases, which will bring more possibilities and innovations for the creation of comics and illustrations.

References

Abdelmigid, H. M., Baz, M., AlZain, M. A., Al-Amri, J. F., Zaini, H. G., Morsi, M. M., Abualnaja, M., & Alhuthal, N. A. (2023). A novel generative adversarial network model based on GC-MS analysis for the classification of Taif Rose. *Applied Sciences*, *13*(5), 3052. https://doi.org/10.3390/app13053052

Ananthaswamy, A. (2023). In AI, is bigger always better? Nature, 615(7951), 202-205. https://doi.org/10.1038/d41586-023-00641-w

Anthony, C., Bechky, B. A., & Fayard, A.-L. (2023). "Collaborating" with AI: Taking a system view to explore the future of work. *Organization Science*, *34*(5), 1672-1694. https://doi.org/10.1287/orsc.2022.1651

Chen, H., Li, J., Zhang, J., Zhang, L., & Zeng, D. (2023). HSI-IPGAN: Hyperspectral Image Inpainting via Generative Adversarial Network. https://doi.org/10.21203/rs.3.rs-2461130/v1

Chui, K. T., Gupta, B. B., Jhaveri, R. H., Chi, H. R., Arya, V., Almomani, A., & Nauman, A. (2023). Multiround transfer learning and modified generative adversarial network for lung cancer detection. *International Journal of Intelligent Systems*, *2023*, 1-14. https://doi.org/10.1155/2023/6376275

Dhar, T., Dey, N., Borra, S., & Sherratt, R. S. (2023). Challenges of deep learning in medical image analysis—Improving explainability and trust. *IEEE Transactions on Technology and Society*, *4*(1), 68-75. https://doi.org/10.1109/TTS.2023.3234203

Gheisari, M., Ebrahimzadeh, F., Rahimi, M., Moazzamigodarzi, M., Liu, Y., Dutta Pramanik, P. K., Heravi, M. A., Mehbodniya, A., Ghaderzadeh, M., & Feylizadeh, M. R. (2023). Deep learning: Applications, architectures, models, tools, and frameworks: A comprehensive survey. *CAAI Transactions on Intelligence Technology*, *8*(3), 581-606. https://doi.org/10.1049/cit2.12180

Grimaldi, G., & Ehrler, B. (2023). AI et al.: Machines are about to change scientific publishing forever. *ACS energy letters*, *8*(1), 878-880. https://doi.org/10.1021/acsenergylett.2c02828

Huang, Z., Wang, J., Lu, X., Mohd Zain, A., & Yu, G. (2023). scGGAN: single-cell RNA-seq imputation by graph-based generative adversarial network. *Briefings in bioinformatics*, *24*(2), bbad040. https://doi.org/10.1093/bib/bbad040

Jiang, P., Xue, Y., & Neri, F. (2023). Convolutional neural network pruning based on multi-objective feature map selection for image classification. *Applied soft*

computing, *139*, 110229. https://doi.org/10.1016/j.asoc.2023.110229

Kim, I., Lee, M., & Seok, J. (2023). ICEGAN: Inverse covariance estimating generative adversarial network. *Machine Learning: Science and Technology*, *4*(2), 025008. https://doi.org/10.1088/2632-2153/acc638

Li, Q., Chen, G., & Luo, L. (2023). Mineral prospectivity mapping using attention-based convolutional neural network. *Ore Geology Reviews*, *156*, 105381. https://doi.org/10.1016/j.oregeorev.2023.105381

Ma, Z., Mei, G., & Piccialli, F. (2022). An attention-based cycle-consistent generative adversarial network for IoT data generation and its application in smart energy systems. *IEEE Transactions on Industrial Informatics*, *19*(4), 6170-6181. https://doi.org/10.1109/TII.2022.3204282

Mukhutdinov, D., Alex, A., Cavallaro, A., & Wang, L. (2023). Deep learning models for single-channel speech enhancement on drones. *IEEE Access*, *11*, 22993-23007. https://doi.org/10.1109/ACCESS.2023.3253719

Nugraha, B. W., & Sutanto, S. M. (2020). Perancangan Buku Untuk Membantu Dalam Menggambar Untuk Remaj. *Jurnal Vicidi*, *10*(2), 31-40. https://doi.org/10.37715/vicidi.v10i2.1928

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of university teaching & learning practice*, *20*(2), 07. https://doi.org/10.53761/1.20.02.07

Rahman, H., Bukht, T. F. N., Ahmad, R., Almadhor, A., & Javed, A. R. (2023). Efficient breast cancer diagnosis from complex mammographic images using deep convolutional neural network. *Computational intelligence and neuroscience*, *2023*. https://doi.org/10.1155/2023/7717712

Raja, K. C., & Kannimuthu, S. (2023). Conditional Generative Adversarial Network Approach for Autism Prediction. *Computer Systems Science & Engineering*, *44*(1). https://doi.org/10.32604/csse.2023.025331

Ruan, D., Chen, X., Gühmann, C., & Yan, J. (2023). Improvement of generative adversarial network and its application in bearing fault diagnosis: A review. *Lubricants*, *11*(2), 74. https://doi.org/10.3390/lubricants11020074

Sanchez, T. W. (2023). Planning on the Verge of AI, or AI on the Verge of Planning. *Urban Science*, *7*(3), 70. https://doi.org/10.3390/urbansci7030070

Sherratt, V., Pimbblet, K., & Dethlefs, N. (2023). Multi-channel Convolutional Neural Network for Precise Meme Classification. Proceedings of the 2023 ACM International Conference on Multimedia Retrieval, https://doi.org/10.1145/3591106.3592275

Sun, L., Shao, W., Zhu, Q., Wang, M., Li, G., & Zhang, D. (2023). Multi-scale multi-hierarchy attention convolutional neural network for fetal brain extraction. *Pattern Recognition*, *133*, 109029. https://doi.org/10.1016/j.patcog.2022.109029

Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, *11*(3), 52. https://doi.org/10.3390/computation11030052

Wenjun, Z., Benpeng, S., Ruiqi, F., Xihua, P., & Shanxiong, C. (2023). EA-GAN: restoration of text in ancient Chinese books based on an example attention generative adversarial network. *Heritage Science*, *11*(1), 42. https://doi.org/10.1186/s40494-023-00882-y

Yan, H., Zhang, H., Shi, J., Ma, J., & Xu, X. (2023). Inspiration transfer for intelligent

design: A generative adversarial network with fashion attributes disentanglement. *IEEE Transactions on Consumer Electronics*. https://doi.org/10.1109/TCE.2023.3255831

Zhan, C., Dai, Z., Yang, Z., Zhang, X., Ma, Z., Thanh, H. V., & Soltanian, M. R. (2023). Subsurface sedimentary structure identification using deep learning: A review. *Earth-Science Reviews*, *239*, 104370. https://doi.org/10.1016/j.earscirev.2023.104370

Zhou, X., Hu, Y., Wu, J., Liang, W., Ma, J., & Jin, Q. (2022). Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT. *IEEE Transactions on Industrial Informatics*, *19*(1), 570-580. https://doi.org/10.1109/TII.2022.3170149

Zhuang, L., Xu, A., & Wang, X.-L. (2023). A prognostic driven predictive maintenance framework based on Bayesian deep learning. *Reliability Engineering & System Safety*, *234*, 109181. https://doi.org/10.1016/j.ress.2023.109181