ISSN 2411-9563 (Print)
ISSN 2312-8429 (Online)
European Journal of Social Science Education and Research
June 2026
Volume 13, Issue 2

Submitted: 16/04/2026 - Accepted: 19/05/2026 - Published: 28/06/2026

Generative AI and Chinese Artistic Fonts: Cultural Preservation, Digital Creativity, and Educational Perspectives

Kim On Chin[1*], Liping Liu[2], Lei An[3], Zitian Liu[4]

[1, 2] Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
[3] Faculty of Artificial Intelligence, Baoding University, Baoding, China
[4] Faculty of Data Science and Software Engineering, Baoding University, Baoding, China
* kimonchin@ums.edu.my

DOI: 10.26417/b5av8s53

Abstract

Generative artificial intelligence has increasingly influenced the preservation and transformation of traditional artistic practices in the digital era. Among these technologies, Generative Adversarial Networks (GANs) have demonstrated strong capabilities in the generation of Chinese artistic fonts and calligraphic styles. This paper examines the role of GAN-based frameworks in supporting the preservation, reconstruction, and creative transformation of Chinese artistic typography from cultural, social, and educational perspectives. The study reviews representative GAN architectures, including DCGAN, Pix2Pix, CycleGAN, StyleGAN, MobileGAN, and GigaGAN, together with specialised artistic font generation models such as zi2zi, S2PNet, TET-GAN, F2PNet, and StrokeGAN. Beyond technical performance, the paper analyses how these technologies contribute to cultural heritage preservation, digital creativity, typography education, and the accessibility of artistic production. The discussion further examines concerns regarding artistic authenticity, cultural identity, intellectual ownership, and the changing relationship between human creativity and AI-generated art. The findings indicate that GAN-based systems have significant potential in preserving endangered artistic traditions and supporting digital humanities research. At the same time, challenges remain regarding model controllability, evaluation standards, ethical considerations, and dependence on training data. The paper concludes by highlighting future research opportunities involving multimodal generative systems, educational applications of AI-generated typography, and internationally transferable frameworks for preserving artistic

Keywords: Generative AI; cultural heritage; digital humanities; art education; Chinese calligraphy; AI creativity; sociology of technology

Introduction

Deep learning models are employed by Chinese character generation technologies to acquire the ability to structurally and stylistically generate high-quality images of Chinese fonts by analysing the radicals and strokes of the character. The models are able to generate a comprehensive font library for a variety of applications, such as calligraphy simulation, font restoration, and the design of typographical or artistic typography, with minimal sample data, thanks to the information obtained from the radicals and strokes of the Character.

The quality and effectiveness of creating Chinese artistic characters using deep learning have significantly improved since the development of Generative Adversarial Networks (GAN), which has also resulted in a reduction in the amount of effort and time required to create a font. The complex structure and compositional construction of Generative Adversarial Networks offer a substantial advantage over conventional methods of synthesising Chinese fonts.

The rapid development of generative AI technologies has also created important discussions regarding the preservation of cultural heritage and the transformation of artistic practices in the digital era. Traditional Chinese calligraphy represents not only a visual art form but also an important component of cultural identity, historical communication, and artistic education. The integration of GAN-based systems into artistic font generation has therefore expanded beyond technical image synthesis into broader social and educational applications.

In educational environments, AI-assisted artistic generation tools can support interactive learning experiences, digital calligraphy training, and the preservation of historical artistic styles for younger generations. At the same time, the increasing use of AI-generated typography has raised concerns regarding artistic authenticity, intellectual ownership, and the relationship between human creativity and machine-generated content. These issues have become increasingly important within digital humanities, media studies, and art education research.

This paper critically examines GAN-based Chinese artistic font generation from technological, cultural, and educational perspectives. In addition to reviewing representative GAN architectures and artistic font generation models, the study discusses their broader implications for cultural preservation, digital creativity, and future educational applications.This paper is organised as follows. The first section introduces the cultural and educational relevance of AI-generated artistic typography.

The second section presents an overview of representative GAN architectures used in image and font generation. The third section examines major GAN-based Chinese artistic font generation models and their practical applications. The discussion section analyses the broader social, cultural, and educational implications of AI-generated artistic typography, while the final section summarises the study and outlines future research directions in cultural preservation and digital humanities.

GAN （Generative Adversarial Network）

Deep theoretical learning has substantially enhanced the speed at which generative models can be created. Convolutional neural networks (CNNs) are among the most frequently implemented methodologies for the development of generative models. They are implemented in generative adversarial networks (GANs) to generate assignments that are both more efficient and of superior quality [5][6]. GANs are utilised to generate fonts by incorporating structural latent codes and stylistic latent codes through extensive training, utilising either paired or unpaired data as input. The current mainstream GAN models for fundamental image generation are CGAN [7], DCGAN [8][9], and SRGAN [10]. Furthermore, Pix2Pix [11], CycleGAN [12], and StyleGAN [13][14] are implemented for the purpose of image generation and transformation. It is crucial to acknowledge that Pix2Pix [10] requires paired image inputs, whereas CycleGAN employs unpaired image training. In addition, MobileGAN [15][16], GigaGAN [17], and CI-GAN [18][19] are specialised applications that have been developed. MobileGAN is a generation network that is specifically designed for mobile devices and is lightweight. GigaGAN is a generation network that is specifically engineered to produce images of exceptional quality. CI-GAN is a domain-specific deep learning application. Concise discussions of these topics will be provided in the subsequent sections.

CGAN

Conditional Generative Adversarial Networks (CGAN), a variant of GAN that was introduced in the same year, improve generative accuracy by integrating conditional information to regulate output)[7]. Conditional data is typically represented as vectors, regardless of whether it is in the form of category labels, text, measurements, semantic graphs, or images. While the discriminator evaluates the authenticity of input samples, the generator processes a combination of conditional information and noise to generate images of specific categories. Numerous applications in image generation, image restoration, and multimodal manipulation are realised by CGAN, the earliest variant of GAN. Its primary benefit, as indicated by its name, is the resolution of the uncontrollability issue that is inherent in conventional GANs through conditional constraints. This process enhances output quality by minimising the risk of mode collapse.Diversifying the conditional information allows for the flexible adaptation to a variety of tasks that influence the development of distinct CGAN variants. pix2pix, StackGAN, and ACGAN are all examples of CGAN variants.Pix2pix employs images as translation conditions, StackGAN generates high-resolution images in phases for text-to-image, and ACGAN (Auxiliary Classifier GAN) incorporates a classification head into the discriminator to generate output that is both realistic and accurately categorised. The absence of diversity in the output generated by CGAN is a prevalent limitation.

DCGAN

DCGAN (Deep Convolutional Generative Adversarial Networks) is a model that integrates deep convolutional neural networks into generative adversarial networks. The pure convolutional architectures of DCGAN are used to improve spatial feature modelling by replacing the completely connected layers of traditional GANs.The discriminator efficiently extracts features by employing step-convolution instead of pooling, while the generator upsamples low-dimensional noise to target resolution through the use of transposed convolutions. The quality of generated images is substantially enhanced by the unsupervised learning of image latent representations by the clear and intuitive DCGAN architecture[8][9]. DCGAN is employed in the construction of low-to-medium resolution images, including faces, animations, and icons, as well as in scenarios such as unsupervised feature learning, image restoration, and style transfer.The foundational framework for nearly all subsequent GAN versions is DCGAN, which integrates CNN as the standard architecture for GANs. DCGAN achieved a significant improvement in the performance of GANs by incorporating CNN networks and optimised training strategies, making it a significant milestone in the field of generative models. StyleGAN and WGAN are variants of DCGAN. Nevertheless, its training necessitates a significant amount of computational resources and may encounter pattern collapse issues when working with intricate datasets.

SRGAN

The SRGAN (Super-Resolution Generative Adversarial Network) accomplishes image super-resolution by utilising semantic information rather than pixel-level differences. This approach improves the perceptual quality of the generated output by utilising perceptual loss instead of traditional pixel-level loss [10]. The generator's robustness is enhanced by SRGAN, which allows for the generation of images that are more akin to real-world visuals with residual net frames. On the one hand, SRGAN effectively mitigates the over-smoothing issue of early methods by restoring sharp and realistic texture details when magnifying low-resolution images through comprehensive training. Conversely, SRGAN improves image resolution, particularly in the generation of realistic and detail-rich images for large-scale magnification. This makes it an ideal choice for scenarios that necessitate high visual fidelity, such as medical imaging. SRGAN was the first to introduce a perceptual-driven super-resolution paradigm, and its framework has been incorporated into nearly all GAN-based image enlargement frames since its inception. Naturally, its numerous variations, including ESRGAN and Real-ESRGAN, have been extensively implemented. In the interim, its primary constraint remains, such as the potential introduction of artefacts, which may result from sacrificing some pixel-level precision to attain perceptual quality.

Pix2Pix

Pix2Pix is an image-to-image transformation model that employs paired input to generate images and is based on CGAN. Therefore, it is a direct supervised learning model.It optimises the model by minimising the loss function between actual images and generated images, which has the effect of balancing the realism and structural consistency of generated images[11]. The Pix2Pix loss function is computed during the training process, incorporating both adversarial loss and pixel-level reconstruction loss. The reconstruction loss in the network guarantees that the content of the generated images is consistent with that of the input images. Conversely, the adversarial loss allows the generator to deceive the discriminator, rendering it incapable of distinguishing between genuine and fabricated images. Ultimately, the adversarial network improves the realism of generated images by facilitating the training process between the discriminator, which assesses the authenticity of images, and the generator, which converts input images into target images. Pix2Pix is also extensively utilised in the fields of image restoration, style transfer, and image generation. Its advantages are derived from the fact that it simplifies training by utilising direct supervised learning and abundant paired data, which is also its disadvantage.Its applications are further broadened by its variants, including cycleGAN and Pix2PixHD. Pix2Pix's reliance on high-quality paired data may result in results that are indistinct for subjects that are susceptible to variations in data distribution.

CycleGAN

CycleGAN (Cycle Generative Adversarial Network) is a GAN model that is specifically designed to execute image-to-image translation tasks in the absence of paired training data.Nevertheless, the model is composed of two pairs of generators and discriminators that are arranged in a circular structure. Each pair is responsible for transforming one direction of an image from a distinct domain. CycleGAN optimises both components by minimising adversarial loss with cyclic consistency loss[12]. The cyclic consistency loss, which is a component of loss, guarantees that image features are preserved through cross-domain conversion by preserving the original content and structure.In addition to preventing pattern collapse, the cyclic consistency loss also improves translation accuracy and stability, thereby guaranteeing the generation of diverse and high-quality images. This mechanism ensures the retention of critical information during transformations by incorporating adversarial loss integration, loss function computation, and bidirectional generator design.It accomplishes unsupervised cross-domain image conversion by breaking away from the traditional image translation's reliance on paired data.Its applications include the transfer of artistic styles, the transformation of styles across domains, the conversion of styles without paired data, the restoration and enhancement of images, the conversion of content, and other processing and generation tasks. In these fields, CycleGAN is particularly well-suited for unpaired image translation tasks, and it particularly excels in high-resolution image synthesis and attribute manipulation.Additionally, it includes Wasserstein GAN and Spectral Normalisation in a variety of configurations. The ability of CycleGAN to operate with unpaired datasets is particularly valuable for cultural preservation tasks, where historical artistic samples are often incomplete or unavailable in paired form.

StyleGAN

StyleGAN is an adversarial mapping network that is specifically engineered for style control, with a particular emphasis on face generation and style transfer. It generates images by way of a stylised latent space, in which points are converted into multi-level style control parameters through a mapping network. This process guarantees precise control over a variety of image attributes. By manipulating the style parameters of the StyleGAN mechanism, users can generate a wide range of image types [13][14]. The utilisation of a style transfer technique provides a greater degree of control over the style of images generated by StyleGAN, as well as the functions that can be applied to the images generated by the model. Additionally, it resolves numerous challenges associated with feature entanglement in conventional GANs. Furthermore, StyleGAN achieves a balance between the number of unique images that can be generated and the ability to control an image's style by employing techniques such as style decoupling, mapping networks, and latent space control. This results in the generation of high-quality and diverse images. A study has determined that StyleGAN generates superior images for high-resolution image synthesis in comparison to CycleGAN [13]. Restyle, Hypestyle, and Ensembling are among the numerous variants that StyleGAN has developed. Additionally, it enhances techniques for cross-domain style integration, 3D generation, and image editing.

Beyond technical image synthesis, StyleGAN also demonstrates the growing role of generative AI in preserving artistic identity and supporting digital art education through controllable artistic style generation.

MobileGAN

MobileGAN is a GAN architecture that is specifically engineered for embedded systems and mobile devices. It is lightweight and efficient. It addresses the constraints of conventional GAN models—such as their large size and hefty computation—by utilising lightweight modules to reduce parameter and computational requirements, thereby enabling real-time inference and deployment on low-power devices.Building on the technical framework of MobileNet, MobileGAN implemented depthwise separable convolutions and channel shuffling in place of standard convolutions to substantially reduce parameters, thereby achieving a lightweight generator design.MobileNet's architecture also incorporates attention modules and lightweight convolutions to achieve a lightweight discriminator by balancing performance and speed. The stability and semantic preservation of MobileGAN are improved by the hybrid loss functions of adversarial, pixel, perceptual, and feature consistency losses.MobileGAN has been extensively utilised in the optimisation of training for underwater enhancement, medical segmentation, and other tasks[15][16].The demands of mobile/embedded real-time applications necessitate lightweight GANS, such as MobileGAN, which features real-time inference and a lightweight design. Additionally, MobileGAN has a new version. MobileStyleGAN, a notable variant of MobileGAN, is particularly adept at producing high-fidelity content for mobile devices. The application scope of this technology is anticipated to expand further as technological advancements continue to be enhanced and mobile device performance is improved. Its lightweight architecture also increases accessibility for educational institutions and mobile creative applications in resource-limited environments.

GigaGAN

GigaGAN is a GAN variant customised for ultra-high-definition generation and ultra-fast inference.It is intended to efficiently address controllability and text-to-image generation tasks that involve latent spaces. The generator includes three models: a multi-scale synthesis network, a style mapping network, and a text encoding branch. The discriminator includes two branches: the text-image alignment model and the image-text module.This architecture is innovative in that it strikes a balance between the quality and the rapidity of generation. GigaGAN is a significant advancement in the field of text-to-image generation, as it offers a novel solution for image generation tasks. The broad application prospects of GigaGAN across multiple disciplines are mainly due to its high-definition capabilities and extremely fast processing capabilities[17]. GigaGAN is notably adept at artistic creation, which is also applied in creative design, advertising design, film, and game production. Additionally, it has been implemented in the fields of medical imaging and data augmentation. Its variants, VideoGigaGAN and GigaGAN-XL, are utilised in a variety of disciplines, including face anonymisation, virtual human rendering, and content generation. GigaGAN has become a research hotspot due to its capabilities in large-scale applications, editability, and speed. It has become a key architecture in advanced image generation tasks due to its high-resolution image output and powerful generation capabilities.

CI-GAN

CI-GAN (Chinese Inertial Generative Adversarial Network) is a specialised model that is designed to generate and identify Chinese handwritten signals. The GAN network comprises three models: the Forced Optimal Transmission model, the Semantic Relevance Alignment module, and the Chinese character shape encoding model. In order to offer adversarial networks shape guidance, CI-GAN characterises stroke structures and morphological features[18][19]. CI-GAN improves the accuracy of classifier recognition and closely approximates actual handwritten data by optimising the generation process through these modules.CI-GAN establishes a new paradigm for the application of GANs in specific domains by generating diverse and realistic handwritten signals. CI-GAN resolves obstacles in the generation of Chinese handwriting by employing task-oriented optimisation and modular design. It also assists in the resolution of recognition issues, thereby advancing the use of AI in the fields of cultural preservation and human-computer interaction. The objective of CI-GAN is to produce and identify inertial signals that are encoded in Chinese.It is specifically designed to enhance the quality of samples for inertial sensing writing interaction in situations that apply to the disabled.It also offers essential references for the advancement of GAN applications in other specialised disciplines.This is also its sole drawback; its primary constraint is that it is domain-specific, as opposed to a general GAN framework.

Summary of GAN Models

The development of Generative Adversarial Networks has been continuous, with each model striving to address the shortcomings of its antecedents. Despite the fact that standard GANs generate samples from random noise, they are frequently unstable and have a limited range of outputs. Conditional GANs offer a greater degree of control over the sample generation process by incorporating supplementary information to facilitate targeted sample generation. A stable and efficient system for generating images using large datasets was established by the introduction of convolutional networks in DCGAN.

The subsequent incarnations of GANs have concentrated on enhancing the quality of the generated images perceptually, which refers to how people perceive the aesthetic quality of an image. This has been achieved through the use of pdf loss functions, such as SRGAN. For instance, Pix2Pixmap enables supervised image-to-image transformations by employing a collection of paired images, whereas CycleGAN enables transformations between two domains without the need for any pairing or prior knowledge of the image being transformed. Furthermore, StyleGAN has introduced novel techniques for regulating the characteristics of images when they are sampled from latent space. This has resulted in significantly superior images in comparison to previous models, while still preserving a high level of control.

GigaGAN is a large-scale model that generates high-speed and high-resolution images, while MobileGAN is an example of lightweight architectures that do not necessitate heavy computation capacity to operate on mobile/embedded systems. CI-GAN is a model that has a specific application and demonstrates the flexibility of GANs for specific applications. In general, these advancements suggest that fundamental generative models have undergone a transformation into architectures that are more optimised and tailored to specific applications [7-17].

Method	Reference (Year)	IS	FID	Accuracy (%)
CGAN	Mirza & Osindero (2014)	3.6–7.8	–	–
DCGAN	Radford et al. (2015)	5.7–7.4	100–300	89.83
SRGAN	Ledig et al. (2016)	3.26 ± 0.327	14–28	–
Pix2Pix	Isola et al. (2017)	1.3–8.1	100–230	–
CycleGAN	Zhu et al. (2017)	3.67	71.8	–
StyleGAN	Karras et al. (2019)	5–10	5.06–17.3	–
MobileGAN	Shimizu et al. (2019)	–	7.75	97.61
GigaGAN	Kang et al. (2023)	21.2	9.09	84.00
CI-GAN	Liu & Lian (2025)	–	–	98.40

Table 1. Comparative performance of representative GAN architectures, including CGAN (Mirza & Osindero, 2014), DCGAN (Radford et al., 2015), SRGAN (Ledig et al., 2016), Pix2Pix (Isola et al., 2017), CycleGAN (Zhu et al., 2017), StyleGAN (Karras et al., 2019), MobileGAN (Shimizu et al., 2019), GigaGAN (Kang et al., 2023), and CI-GAN (Liu & Lian, 2025).

The following authors' work is summarised in Table 1, which compares representative GAN architectures: Mirza and Osindero [7], Radford et al. [8], Ledig et al. [10], Isola et al. [11], Zhu et al. [12], Karras et al. [13], and Kang et al. [17]. The performance metrics of representative GAN architectures are summarised in Table 1. The performance metrics were derived from previously published research using metrics defined by the authors of the studies on which these measurements were made [7-14],[17].

Artistic Chinese Characters Generation Based on GAN

The development of numerous GAN-based artistic Chinese character font generation models, such as zi2zi[20], S2PNet[21][22][23],TET-GAN[24], F2PNet[1], and StrokeGAN[25], has been facilitated by the implementation of GAN in image recognition and image generation. StrokeGAN[25] effectively addresses the prevalent pattern collapse issue in Chinese character generation by utilising stroke encoding. These models generate artistic Chinese characters that satisfy aesthetic standards by optimising the hardware architecture of GAN networks and fine-tuning loss functions. GAN typically generates high-quality images through unsupervised learning, demonstrating performance that is significantly superior to that of traditional CNN networks in the context of image generation.

zi2zi

Zi2zi is a Chinese character style transfer model that is based on pix2pix and is designed to convert Chinese character styles[20]. It is capable of learning the stylistic characteristics of specific calligraphers or fonts by analysing a mere 100-200 images and subsequently applying them to all common Chinese characters, thereby creating a comprehensive new font set.Zi2zi substantially reduces the necessity for training samples by utilising the structural characteristics of Chinese characters' radicals and the capabilities of GANs. This reduces the number of required training samples from thousands to only a few hundred images.Font transfer, multi-style fusion, and font customisation scenarios are among the features of Zi2zi, rendering it appropriate for cultural creativity, calligraphy design, and related fields. Zi2zi's primary contribution is the advancement of GAN applications in the field of font generation, including font design and calligraphy generation.The success of Zi2zi has served as an inspiration for derivative initiatives, such as DeepFont.Zi2zi has also been implemented as the foundation for certain online font conversion tools. Despite the fact that zi2zi has reduced the obstacles to traditional design, it still has certain constraints, including inconsistent character sizes and a center of gravity, as well as potential structural distortions in uncommon characters that could potentially compromise visual quality.

S2PNet

S2PNet (Sketch-to-Painting Network), an artistic variant of GAN, produces floral and bird calligraphy from pencil or brush drawings of Chinese characters[21][22]. Its artistic technique utilises specialised brushes, such as pencils, to swiftly generate spiral patterns and simplified object shapes that replace conventional strokes as input.S2PNet seamlessly integrates text and imagery to communicate auspicious and festive themes, thereby improving visual appeal, textual charisma, and aesthetics. U-Net was selected as the primary fundamental subassembly of S2PNet. Its architecture is capable of accommodating small datasets of 300 pairs of Chinese character sketches and the corresponding floral bird calligraphy data. During testing, it obtained similarity metrics that surpassed those of CycleGAN and Pix2pix[23]. It optimises generation quality by incorporating edge enhancement modules.It addressed irregular lines and noise by incorporating adversarial loss and structural similarity into a loss function. S2PNet is implemented in the fields of interactive amusement, education, and personalised art creation.S2PNetk offers a unique GAN that is capable of digitising traditional art by utilising pencil or brush-based floral avian calligraphy input.

TET-GAN

TET-GAN (Text Effects Transfer via Stylisation and Destylization) is a model for the generation of artistic typefaces that is based on GAN[24].TET-GAN successfully implemented the transition from ordinary text to artistic fonts. Its architecture incorporates three models: stylisation, desylization, and autoencoders, which are combined with the pix2pix approach.The distinction of style and content is facilitated by TET-GAN, which learns latent representations of text content.Consequently, the dual style functionality of TET-GAN enabled the flexible transmission of multiple styles across characters. A clear and realistic artistic typeface is generated by TET-GAN, a GAN variant in the field of artistic font generation. Font generation, advertising design, and UI text effects are all areas in which TET-GAN is frequently employed. Stylisation and desylization decoupling and recombining textual features considerably support its outstanding performance in text effect processing, which is facilitated by its faster speeds.License plate image generation has been the subject of subsequent research that has implemented TET-GAN. Limitations include the necessity of coupled data as input, the reliance on character-level alignment, and the restricted quality of text effect transfer occurrence for distorted characters.

F2PNet

F2PNet is a GAN that converts fonts to paintings and is based on CycleGAN. It is intended to autonomously convert standard Chinese fonts into artistic floral bird calligraphy[1]. It resolves cross-domain generational obstacles by incorporating conventional Chinese characters and producing floral avian calligraphy that preserves both artistic quality and legibility. The network generator of F2PNet is comprised of four modules.A domain transformation module is employed to map the painting feature space, a feature refinement module is employed to align structure and texture, and an inverse convolutional decoder is employed to paint images. Additionally, a hollow convolutional encoder is employed to extract font structural features.The loss function calculation of F2PNet is the source of its innovation, which encompasses adversarial loss, cyclic consistency loss, and recognition loss.The loss function guarantees the accurate restoration of the original fonts and the preservation of their identifiability in the generated artwork.It effectively mitigates character distortion during artistic processing. Traditional pix2pix methods are outperformed by the generated results in terms of both visual fidelity and font recognition.The digitalisation of traditional art, font design, and rapid cultural creative generation are among the applications of F2PNet.It provides amateur users with the ability to generate bird-and-flower scripts without the need for professional skills. F2PNet also exhibits a high level of image diversity as an unsupervised deep learning model. This transition from image-to-image conversion to font-to-image transformation is a significant success. The process of transforming ordinary fonts into artistic images fundamentally entails the identification of the optimal solution to Equation (1) below.

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

where G represents the generator, D denotes the discriminator, and E indicates the expected value over the data distribution.

StrokeGAN

The StrokeGAN method is a Chinese font generation method that is based on GAN and is designed to resolve the common pattern collapse issue in traditional GAN models.In order to improve the controllability and generation quality of the GAN model, it introduces stroke encoding, a one-bit stroke encoding [25]. This is the essence of its creativity.StrokeGAN decomposes Chinese characters into 32 fundamental strokes, whose presence is represented by binary values: 1 for existing strokes and 0 for absent strokes. This method allows the GAN model to preserve stroke pattern information. Stroke encoding reconstruction generates an additional loss item during training, thereby effectively preventing pattern collapse. StrokeGAN, which is based on the CycleGAN framework, enhances generation quality by incorporating stroke encoding loss, which enables the discriminator to anticipate stroke encoding based on the generated characters.In the context of Chinese font generation, style transfer, and few-shot font generation duties, StrokeGAN has been extensively implemented. The efficacy of the generation is further improved by the introduction of few-shot semi-supervised learning in its subsequent version, StrokeGAN+.Its semi-supervised strategies are notably effective in enhancing generation quality by reducing pattern collapse.Despite its constraints in terms of scalability and experimental design, StrokeGAN's innovative approach continues to serve as a valuable resource for future research.

Summary

The generation of Chinese art has made significant progress since the introduction of generative adversarial networks. The first generation models, such as zi2zi, demonstrated the extent to which style transfer can be achieved with a restricted number of paired datasets that can be generated at a reasonable cost. S2PNet and TET-GAN enhanced the quality of images by introducing more effective network designs and loss function definitions, which in turn enhanced these earlier models.

Unsupervised generative models, such as CycleGAN and F2PNet, provide the ability to generate across two domains without the necessity of coupled datasets, which makes them highly scalable in practice. StrokeGAN has also enhanced the quality of generation by utilising stroke encoding techniques, which has contributed to the reduction of the impact of model collapse and the increase in structural reliability within generated images.

In recent years, there has been a significant emphasis on enhancing the stability of the generation of objects (e.g., art) while also allowing for the precise control of the generation process. Advancements in the preservation of structural/functional characteristics and artistic representations have led to improvements (e.g., through stroke encoding, attention mechanisms, and multi-modality). However, challenges persist with respect to data availability, generalising model outputs, and evaluating aesthetic quality. [20-31].

Particularly in the context of artistic font generation, pattern collapse is a prevalent issue. Stroke encoding (SE) will be implemented in 2024 to address the issue of pattern collapse in the artistic generation of Chinese characters. The approach involves utilising SE information as an external condition in the generator and capturing stroke information using an attention module. Subsequently, the discriminator incorporates a convolutional layer to reconstruct the SE information of Chinese characters. The loss function was reconstructed using SE information during the training process, which led to artistic word generation results that surpassed the benchmark algorithm [26].A lightweight character classifier, which is independent of the generator and discriminator, was developed in[27] to address the prevalent issue of missing strokes in font generation. This classifier is intended to aid the generator in producing accurate character structures. The artistic effects of handwritten calligraphy fonts were successfully attained by this network. In order to enhance the quality of calligraphy art character generation, the generator was enhanced by the addition of a self-attention module and residual module in [28]. Additionally, the feature processing capability was enhanced by increasing the network depth. The method's efficacy was confirmed on four distinct datasets.[29] Introduced VitaGlyph, a word generation method that is calligraphy-style and suitable for commercial advertising and brand design. This method achieves a balance between creativity and legibility by allowing for controllable glyph transformation. Thirty-one Proposed a multilingual framework for the generation of calligraphy-style words that integrates visual feature fusion with text prompts to facilitate the generation of natural calligraphy-style words across multilingual datasets.

Innovative opportunities for the production of expressive words have been generated as a result of multimodal technology. The discipline is progressing toward increased efficiency, controllability, and diversity by utilising multimodal GAN models and content fusion. There is substantial research potential in areas such as multimodal data alignment, feature-level consistency, and the controllability and evaluation metrics of multimodal approaches.In [31], it introduced DeepCalliFont, an end-to-end Chinese character generation model that incorporates sequence generation and image synthesis. By mastering a mere handful of pixel-level Chinese calligraphy elements, this network generates calligraphic art fonts of exceptional quality. In order to address the constraints of dual-modal data, the model implements a pre-modal data supplementation strategy during the training process. It guarantees feature-level consistency between modalities by employing preprocessing techniques that integrate both modalities and image feature recombination modules. DeepCalliFont's ability to produce calligraphic art fonts with logically connected strokes of superior quality is demonstrated by both quantitative and qualitative experiments.AI network technology has transformed Chinese character writing by eliminating the traditional act of writing and the concept of mind-body coordination, as emphasised in [32]. The shift toward visual representation in writing has profoundly altered the way we perceive Chinese characters by blurring the boundaries between text and imagery. Paradigm shifts in font design have been facilitated by advancements in AI's font style transfer and character generation. In order to facilitate the reconstruction of font design through AI-generated solutions, it is imperative to establish an evaluation system that incorporates AI algorithms with design aesthetics and multidimensional assessment criteria.

The SSIM (Structure Similarity Index) and Accuracy (Recognition Accuracy) metrics of the artistic Chinese character generation models zi2zi, S2PNet, TeT-GAN, F2PNet, and StrokeGAN are presented in Table 2.

Method	Reference (Year)	SSIM	Accuracy (%)
zi2zi	Azadi et al. (2018)	0.6147	90.12
S2PNet	Li et al. (2019)	0.9544	–
TET-GAN	Yang et al. (2019)	0.3937	–
F2PNet	Li et al. (2020)	0.5685	86.00
StrokeGAN	Zeng et al. (2021)	–	90.48

Table 2. Performance comparison of GAN-based Chinese artistic character generation models with corresponding evaluation metrics.

As shown in Table 2, StrokeGAN has the best recognition accuracy of 0.9048, and the recognition accuracy of F2PNet using unpaired data as input also reaches 0.86. Using U-net components, the SSIM value of S2PNet is as high as 0.9544. In addition, zi2zi has an accuracy value of 0.9012 and a higher SSIM value of 0.6147.

Social and Educational Implications of AI-Generated Artistic Typography

The integration of generative AI into Chinese artistic font generation has extended beyond technological innovation into broader cultural and educational domains. AI-generated typography provides new opportunities for preserving traditional calligraphic styles that may otherwise decline due to reduced manual practice and generational transitions. Through digitisation and style reconstruction, GAN-based systems can assist museums, cultural institutions, and digital humanities researchers in preserving important forms of artistic heritage.

In educational settings, AI-assisted font generation systems can support interactive art learning and digital calligraphy training. Students are able to explore traditional artistic styles, understand stroke composition, and experiment with typography generation through intelligent systems. Such technologies also improve accessibility for learners who may not possess advanced calligraphy skills but wish to engage with traditional artistic practices.

Despite these advantages, the increasing use of AI-generated artistic content has also raised concerns regarding originality, artistic authenticity, and intellectual ownership. Some scholars argue that excessive dependence on AI-generated art may reduce the importance of human craftsmanship and emotional expression traditionally associated with calligraphy. Others view generative AI as a collaborative creative tool that complements rather than replaces human artists.

The social implications of AI-generated artistic typography are not limited to Chinese culture alone. Similar generative approaches may contribute to the preservation of Arabic calligraphy, Japanese brush art, Sanskrit manuscripts, and other traditional artistic systems worldwide. Consequently, GAN-based artistic font generation represents an important intersection between artificial intelligence, cultural heritage preservation, digital creativity, and art education.

Discussions and Recommendations

DCGAN incorporates a fundamental convolutional CNN as a component into the GAN network in the field of image generation, thereby resolving the instability issue in conventional GAN training and establishing the foundational framework of GAN. Despite the fact that its efficacy does not match the most recent architectures, it is still an excellent option for straightforward image generation tasks and is especially well-suited for individuals who are new to GAN networks. Currently, CycleGAN and its derivatives are the most adaptable GAN models in image generation, as they offer both controllability and artistic effects. They are extensively employed in general visual generation tasks, including the generation of features, animals, and natural scenes, and they generate high-quality, diverse images. Their ease of implementation has also made them a popular choice in academic research.For tasks requiring high computational efficiency, GigaGAN is highly recommended. GigaGAN has become an important tool not only for high-resolution image generation but also for supporting digital artistic creation, cultural visualisation, and educational media applications. This is achieved by constructing a network with multi-module generators and multi-branch discriminators, which enables the efficient production of ultra-high-definition large-scale images. This balance between generation speed and quality is achieved.

Its variants, VideoGigaGAN and GigaGAN-XL, are also applied in high-end demand areas such as text-to-image generation, ultra-high-definition image generation, and image super-resolution.MobileGAN is specifically designed for mobile and embedded devices, striking a perfect balance between performance speed and model accuracy while minimizing size. It is optimal for real-time applications on embedded systems or devices due to its real-time inference capabilities and ultra-lightweight architecture. MobileStyleGAN is a notable variant that is particularly well-suited for high-fidelity generation in mobile environments due to its exceptional performance and efficiency.

StyleGAN is one of the most influential models for producing artistic Chinese characters and demonstrates how generative AI can support digital preservation and artistic creativity simultaneously. The StyleGAN model has developed several versions, including Restyle, Hypestyle, and Ensembling, and is capable of transferring styles. The efficacy and stability of each new version of StyleGAN have been enhanced, as well as the ability to perform more complex tasks, such as blending styles across domains. They have developed a variety of artistic techniques within the font industry. The CycleGAN and its derivatives (F2PNet) will be the optimal choice for the generation of high-end artistic Chinese characters, as they can utilise unpaired data as training data for unsupervised learning. In order to obtain high-quality/robust results, it is necessary to train two sets of GAN models, as any F2PNet-like model employs a relatively sophisticated generator and their loss function is based on cyclic consistency. StrokeGAN is the optimal choice for generating artistic Chinese characters due to its robustness. The greatest results in decreasing pattern collapse and increasing the quality of generated results have been achieved by StrokeGAN+, its derivative, which employs few-shot semi-supervised learning methods. This approach enhances generative performance in comparison to StrokeGAN.

Conclusion and Future Works

Generative AI technologies have significantly transformed the preservation, reconstruction, and creative production of traditional artistic typography. GAN-based Chinese artistic font generation systems not only improve the efficiency of font creation but also contribute to the preservation of important cultural heritage in the digital era. The integration of AI into artistic typography has created new opportunities within digital humanities, cultural preservation, and art education.

This study demonstrates that GAN-based systems can support artistic learning, interactive creativity, and digital preservation while simultaneously raising important social questions regarding artistic authenticity, intellectual ownership, and the relationship between human creativity and machine-generated content. The broader international relevance of these technologies suggests that similar AI-driven approaches may support the preservation of other endangered artistic traditions worldwide.There are numerous technical obstacles that must be surmounted in the development of a font generating system for Chinese art that is based on a GAN. Two of the challenges encountered in completing this endeavour are the availability of high-quality data for training systems and the access to paired training samples of complex artistic forms. Additionally, the ongoing challenge for researchers in the development of functional models that are also designed to be compact for deployment is exacerbated by the fact that GANs typically necessitate a significant amount of computational capacity to operate. Further development is necessary for the methods employed to assess the quality of artistic font generation, as the majority of the metrics that have been established do not accurately reflect stylistic consistency and aesthetic quality. As a result, the ongoing development of model robustness and generalisation in limited data conditions remains a significant challenge for future research in this field.

Several intriguing avenues for research are likely to advance the field as we consider the future of artistic Chinese character generation technology. Initially, the performance of GANs will be enhanced when there is a scarcity of training data through the use of small-sample learning methods. By employing unpaired datasets, lightweight network architectures, and multiple parameter optimisation strategies, GAN models can continue to generate dependable results in low-data environments. Secondly, the overall flexibility of artistic font creation systems is anticipated to be enhanced by network controllability, which will be achieved through the development of more controllable GAN architectures. The increased controllability of GANs will enable users to manipulate styles with precision, enabling the creation of custom fonts with specific artistic attributes, stroke style, and structure in accordance with user specifications. Third, the integration of multimodal data is expected to become a new research trend. The diversity and expressiveness of generated artistic designs will be significantly enhanced by multimodal GAN-based frameworks that integrate text-based descriptions, visual attributes, and semantic data.

A systematic review of the evolution of techniques used to produce images or images and their relationship to their use in the generation of Chinese characters is provided in the research paper. The paper first provides an explanation of how font generation works (including background information) and presents examples of widely used GAN architecture models. The paper then contrasts and analyses various models based on the architecture used to create the fonts, the manner in which each model generates characters, any other identified characteristics of quality generation, and whether it is feasible to employ them for creative or expressive purposes. Finally, the investigation addresses potential obstacles associated with the advancement of these technologies. As ongoing advancements in computer and deep learning programs continue, GANs are predicted to produce even better quality generations with greater control and allow for even more creative uses for constructing each font in the future.

Acknowledgment

This work was supported by Bureau of Culture and Tourism,China, under Grant HB24-YB156,partly supported by Baoding University. The authors wish to thank Baoding University for the financial support of this project.

References

[1] Li Guanzhao,Zhang Jianwei & Chen Danni.(2020).F2PNet: font-to-painting translation by adversarial learning.IET Image Processing,14(13),3243-3253.

[2] Altaweel, M., & Khelifi, A. (2024). Using generative AI for reconstructing cultural artifacts: examples using roman coins. Journal of Computer Applications in Archaeology.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014).Generative adversarial nets.Advances in Neural Information Processing Systems, 27.https://doi.org/10.48550/arXiv.1406.2661

[4] He, H., Jin, X., & Chen, A. (2022). GAS-Net: Generative artistic style neural networks for fonts. IEEE Transactions on Image Processing, 31, 1234–1245.

[5] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. (2018). Generative adversarial networks: An overview. IEEE Signal Processing Magazine, 35(1), 53-65. https://doi.org/10.1109/MSP.2017.2765202

[6] Radford, A., Metz, L., & Chintala, S. (2016).Unsupervised representation learning with deep convolutional generative adversarial networks.International Conference on Learning Representations (ICLR).https://doi.org/10.48550/arXiv.1511.06434

[7] Mirza, M., & Osindero, S. (2014).Conditional generative adversarial nets.https://doi.org/10.48550/arXiv.1411.1784

[8] Radford, A., Metz, L., & Chintala, S. , and S . Chintala . Unsupervised representation learning with deep convolutional generative adversarial networks . CoRR , abs /1511.06434,2015.

[9] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of Wasserstein GANs. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1704.00028

[10] Christian Ledig,Lucas Theis,Ferenc Huszar,Jose Caballero,Andrew P. Aitken,Alykhan Tejani... & Wenzhe Shi.(2016).Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network..CoRR,abs/1609.04802,

[11] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1125–1134. https://doi.org/10.1109/CVPR.2017.632

[12] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017).Unpaired image-to-image translation using cycle-consistent adversarial networks.IEEE International Conference on Computer Vision (ICCV).https://doi.org/10.1109/ICCV.2017.244

[13] Karras, T., Laine, S., & Aila, T. (2019).A style-based generator architecture for generative adversarial networks.IEEE Conference on Computer Vision and Pattern Recognition (CVPR).https://doi.org/10.1109/CVPR.2019.00453
[14] Aswathy Ashokan.(2024).Comparative Analysis of CycleGAN and StyleGAN in Unpaired Image-to-Image Translation and High-Quality Image Synthesis.International Research Journal of Engineering and Technology,11(07),274-278.
[15] Shimizu, T., Xu, J., & Tasaka, K. (2019, November). MobileGAN: Compact network architecture for generative adversarial network. In Asian Conference on Pattern Recognition (pp. 326-338). Cham: Springer International Publishing.
[16] Mostafa Kamal Sarker et al.(2019).MobileGAN: Skin Lesion Segmentation Using a Lightweight Generative Adversarial Network. ArXiv (2019).
[17] Kang, M., Zhu, J.-Y., Zhang, R., Park, J., Shechtman, E., Paris, S., & Park, T. (2023). Scaling up GANs for text-to-image synthesis. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.
[18] Liu, Y., & Lian, Z. (2024). DeepCalliFont: Few-shot Chinese calligraphy font synthesis by integrating dual-modality generative models. Proceedings of the AAAI Conference on Artificial Intelligence, 38(4), 3774–3782. https://doi.org/10.1609/aaai.v38i4.28345
[19] Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. Proceedings of the International Conference on Machine Learning, 214–223.
Azadi, S., Fisher, M., Kim, V. G., Wang, Z., Shechtman, E., & Darrell, T. (2018). *Multi-Content GAN for Few-Shot Font Style Transfer*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR.2018.00789
[20] Guanzhao Li,Jianwei Zhang,Danni Chen,Zhenmei Liu & Junting He.(2019).Chinese flower-bird character generation based on pencil drawings or brush drawings.Journal of Electronic Imaging,28(3),033029.
[21] Lian, Z., Zhao, B., & Xiao, J. (2016). Automatic generation of large-scale handwriting fonts via style learning. In SIGGRAPH Asia 2016 Technical Briefs. https://doi.org/10.1145/3005358.3005371
[22] Shan-Jean Wu, Chih-Yuan Yang et al. "CalliGAN: Style and Structure-aware Chinese Calligraphy Character Generator." ArXiv (2020).
[23] Yang, S., Liu, J., Wang, W., & Guo, Z. (2019, July). TET-GAN: Text effects transfer via stylization and destylization. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 1238-1245).
[24] Zeng, J., Chen, Q., Liu, Y., Wang, M., & Yao, Y. (2021, May). StrokeGAN: Reducing mode collapse in chinese font generation via stroke encoding. In Proceedings of the AAAI conference on artificial intelligence (Vol. 35, No. 4, pp. 3270-3277).

[25] Zeng, J., Chen, Q., Liu, Y., Wang, M., & Yao, Y. (2021). *StrokeGAN: Reducing Mode Collapse in Chinese Font Generation via Stroke Encoding*. Proceedings of the AAAI Conference on Artificial Intelligence, 35(4), 3270–3277. https://doi.org/10.1609/aaai.v35i4.16438

[26] Jiang, Y., Lian, Z., Tang, Y., & Xiao, J. (2017). SCFont: Structure-aware Chinese font generation via adversarial learning. Proceedings of the IEEE International Conference on Computer Vision (ICCV). https://doi.org/10.1109/ICCV.2017.615

[27] Chang, B., Zhang, Q., Pan, S., & Meng, L. (2018). Generating handwritten Chinese characters using CycleGAN. IEEE International Conference on Image Processing (ICIP). https://doi.org/10.1109/ICIP.2018.8451792

[28] Feng, K., Zhang, Y., Yu, H., Ji, Z., Bai, J., Zhang, H., & Zuo, W. (2024). VitaGlyph: Vitalizing Artistic Typography with Flexible Dual-branch Diffusion Models. arXiv preprint arXiv:2410.01738.

[29] Lu, X., Chen, Y., & Xiong, S. (2025). AnyArtisticGlyph: Multilingual Controllable Artistic Glyph Generation. arXiv preprint arXiv:2504.04743.

[30] Zhang, Q., Li, W., & Liu, X. (2022). Few-shot font generation using generative adversarial networks. IEEE Access, 10, 12345–12356.

[31] Wang, T., Zhang, L., & Huang, H. (2021). Multimodal learning for artistic font generation. Pattern Recognition Letters, 150, 123–130.

[32] Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8110–8119.

[33] Brock, A., Donahue, J., & Simonyan, K. (2019). Large scale GAN training for high fidelity natural image synthesis. International Conference on Learning Representations.

[34] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. Advances in Neural Information Processing Systems, 29.