AEM JOURNAL.ORG

Digital Technology Empowers Innovation in Calligraphy and Painting Creation: Research on Simulation of Brush and Ink Styles and Personalized Creation Based on Generative AI

Yinjie Li[1], Junle Hu[2]
[1]Shanxi University of Electronic Science and Technology, LinFen 041000, Shanxi, China
[2]Shanxi Normal University, Taiyuan 030000, Shanxi, China
Corresponding author: Junle Hu (e-mail: 13593523284@163.com)

ABSTRACT The digital preservation and innovative expression of calligraphy and painting styles require accurate modeling of brushstroke morphology, ink diffusion characteristics, and personalized artistic features. To address the limitations of existing digital creation systems in style simulation accuracy and personalized generation capability, this study proposes a brush-and-ink style simulation framework based on generative artificial intelligence. The proposed method integrates deep generative adversarial networks and diffusion models to perform high-dimensional feature encoding, brushstroke parameter modeling, ink-flow representation, and semantic-guided style generation. A human–computer interaction module is further developed to support personalized artistic creation through controllable semantic and stylistic inputs. Experimental evaluation on 5,000 calligraphy and painting samples demonstrates that the proposed framework achieves a brushstroke restoration accuracy of 93.8% and significantly improves style similarity compared with conventional convolutional approaches. The results confirm the effectiveness of generative AI for digital artistic reproduction and personalized creation. The proposed framework also provides methodological references for image generation, visual information processing, and intelligent pattern synthesis applications.

INDEX TERMS generative artificial intelligence, brush-and-ink style simulation, image generation, visual information processing, intelligent pattern synthesis

I. INTRODUCTION

CALLIGRAPHY and painting art use brush and ink as the fundamental elements, and their forms of expression are quite reliant on the minute variations in brush strokes, intensity, and color of the ink, which present unique technical challenges when translated into the interlacing structures of high-end jacquard The subtle materials transitions of ink saturation and brush pressure must be precisely mapped to material density and structure patterns to ensure the authentic reproduction of artistic textures on the surfaceThis is a complicated structured material nature a nd thus it becomes very hard to quantify and reproduce it i n a digital environment. Conventional techniques of digital painting or style transfer frequently do not allow mapping the texture component, so it is hard to reproduce the rhythm of brushstroke movement and natural movement of ink color. Thus the issue of achieving the reproduction of high-fidelity morphological features and textural nuances in a digital system is now a significant issue in the ongoing art digitization studies [1,2]. The mapping connection between space continuity of brush and ink, the tints of ink, and the motion rhythm has not established into a whole calculable mechanism, and this has given rise to the evident absences of the realistic art and expression of individual qualities regarding digitally created artworks.

Although existing generative techniques have shown high performance in natural image and stylized synthesis tasks, their application in the context of calligraphy and painting is still limited. Convolutional neural network structures are biased towards static feature extraction and have limited ability to capture the flow of ink, resulting in problems such as broken brushstrokes and unnatural ink layering. Although style transfer-based algorithms can extract some visual style information, they lack in-depth expression of the artist's creative habits, brush rhythm and ink layering [3] . It can be seen that digital calligraphy and painting creation not only needs to improve accuracy, but also needs to realize the adaptive understanding and restoration of individual artistic styles by the generative model.

Given the breakthrough in the field of artificial intelligence, it has become possible due to the rapid development of its generative form. Through the combination of Generative Adversarial Networks (GANs) and diffusion models, self-learned, and dynamic adaptation is possible in the joint learning of ink texture, stylistic layers, and dynamic brushstrokes. Based on the understanding of human-computer collaborative creation, this research develops an intelligent generative methodology that responds to semantic descript ions and creation intentions, allowing artists to recreate ink s tyles and experiment with customized jacquard producti on patterns in a digital space.This approach facilitates the s eamless translation of artistic brushwork into the complex i nterlacing structures of high-end textile substrates, ensuring t hat the nuances of ink intensity are preserved through digital control.

II. RELATED WORK

The brush and ink style digital research has developed to the feature extraction founded on image processing to the brush stroke modeling motivated by deep learning. Initial research utilized edge detectors and fractal geometry to represent the structure of the handwriting to a great extent, parametrically, with a great concern of extracting the static characteristics, including brush stroke direction, the distribution of pen pressure and ink density. Other researchers have attempted to model the process of contact between pen and paper by physical simulation techniques to simulate ink droplet diffusion and adsorption onto paper, but these techniques material are very computationally intensive and cannot model the behaviour of realistic brush strokes [4]. As the visual computing technology advanced, the study started to pay attention to the systematic description of brush and ink, and perceived the decomposable representation of ink color degrees by feature layer modeling, which formed the technical basis of the emergence of the later generative processes.

Generation has been an area of art that has been highly encouraged because of the introduction of generative models. Generative Adversarial Networks (GANs) is an end-to-end learning method of generating image styles that can allow models to learn the spatial distribution of the styles of art by competitive optimization. Simulating brushstrokes and transferring textures are applications of GANs in calligraphy and painting where attention mechanisms can also be added to enhance consistency of the style. Diffusion models, which have the property of inverse noise processes, are particularly successful at the fidelity of style generation details and reproduce the ink density and spatial hierarchy changes in an accurate manner [5,6]. Nevertheless, the research of the image level generation is the most frequent, the modeling of brushstroke levels, and the interactive regulation of creative semantics remain insufficient.

Although the current studies have achieved the visual level of generating artworks, it has serious deficits on aesthetic expression and personalized reproduction of ink and brush paintings. GANs are susceptible to local repetition and texture blurring during the details generation, whereas diffusion models, despite being more successful in learning the details, have poor-generation efficiency and inability to control the style. More to the point, with most of the studies, there is no modeling mechanism of the intent of the creator and cannot successfully guide the generation guided by semantic or stylistic cues, it is challenging to make the model reflect the specific non-linear brushstroke velocity and the complex variance in ink saturation of the artist.

Hence, there is imminent necessity of a combined approach which will be able to coordinate the knowledge of ink and brush construction, control of style and personalized production of expression, which will lead to intelligent and artistic development of digital calligraphy and painting production [7,8].

III. METHODS

A. DATA COLLECTION AND PREPROCESSING

The sources of data are concerned with the works of calligraphy and the ink painting, taking 5,000 images in the databases of public art, authorized collections of artists, and published them in seven representative styles, such as running script, regular script, cursive script, landscape, and flower-and-bird painting. High-precision scanners were used on all images with a single resolution of 600 dpi to guarantee that brushstroke edges and ink tones could be distinguished. Guided filtering-based denoising algorithm was applied to retain ink details in the original images to eliminate noise and texture interference caused by paper to be used with structural analysis applied in order to automatically adjust the stroke direction. Following image normalization, block sampling was used in which each work was subdivided into 64*64 pixel brushstroke units in order to capture local brushstroke variations and tonal variations [9,10]. To capture both local texture and global structural rhythm, a multi-scale block sampling strategy was adopted. While 64*64 pixel units were used for capturing microscopic ink tonal variations, a global attention mechanism with a larger receptive field (up to 512*512 pixels) was integrated into the feature library. This dual-path sampling ensures that the model learns both the granular ink absorption and the long-range spatial dependencies of brushstroke motion, material-level which are critical for the fluid connectivity found in cursive script and landscape compositions. A parametric label space of brushstroke thickness, density, and range of diffusion was built and brushstroke attribute annotation was done by combination of manual labeling and automatic grouping with an accuracy rate of more than 95%. Lastly, a multi-dimensional feature library with image texture features, brushstroke features and metadatas of art styles was formed to form a data basis upon which learning of brush and ink styles was to be performed under generative models. To ensure balanced training across categories, a stratified random sampling approach was applied to a raw collection of 5,247 works, resulting in a finalized experimental dataset of 5,000 images. Table 1 shows the composition and annotation statistics of the dataset.

TABLE 1. Dataset Composition and Annotation Statistics

Style Category	Number of images	Average resolution (dpi)	Mark the number of strokes	Labeling accuracy (%)
Running script	814	600	52,130	95.3
Regular script	689	600	48,642	94.8
cursive script	612	600	50,119	95.9
landscape	1387	600	110,583	96.1
Flowers and birds	1498	600	122,476	95.6

B. BRUSH AND INK MODELING

It is the brush and ink modeling which relies on the brushstroke dynamics and laws of diffusion of ink and the model is built by examining the brushstroke track, variations of speed, and diffusion distribution on the paper. The brushstroke structure is defined by the correlation of the brush pressure, direction of travel as well as speed with the brush. It is measured by sampling using a high-frame-rate handwriting, and visual tracking, the time and path curve of the contact between the brushstroke and the paper. The ink diffusion characteristics are informed by grayscale gradient variation, whereby, a water diffusion simulation algorithm is used to model the growth process of ink marks in the areas of high saturation to dry ones, with the parameters including the rate at which paper absorbs and the range of ink volume the algorithm can use. Prior to model training, the brushstroke edge regions are subjected to texture enhancement in order to capture the tonal difference in the brushstroke regions induced by the varying water contents. The convolutional subnetwork causes texture features to be created with mapping multi-scale response to rebuild the discontinuous and intersecting tonal connections of ink lines. Differences in brushwork between different schools are modeled using brushstroke angle distribution and density matrices to achieve quantitative alignment between styles.

C. GENERATIVE MODEL DESIGN

The generative model is a deep fusion of Generative Adversarial Networks (GANs) and a diffusion model to improve the stability and controllability of the ink texture generation process. The GAN part learns the spatial arrangement of brushstrokes and local tones relations and shades. The generator can reconstruct the brushstroke texture with multi-scale convolutional structures after being fed with the input features of brushstroke dynamics and ink placement. The discriminator is the one that perceives consistency of texture and ink tones, attaining structural fidelity constraints about the handwriting. The diffusion model deals with learning adaptive styles, directing the generation process with conditional control vectors by gradually introducing and eliminating noise, so as to give a smooth transition among the various styles of ink techniques in the latent space. The training model is based on a joint multi-objective optimization framework:

$$L_{\text{total}} = \lambda_1 L_{\text{GAN}} + \lambda_2 L_{\text{diff}},$$

where L_{GAN} employs a hinge loss for high-frequency structural sharpness and L_{diff} utilizes a variational lower bound loss for low-frequency ink rheology. To prevent gradient interference, we implement a Gradient Block (GB) mechanism at the fusion layer, ensuring that the GAN discriminator does not destabilize the diffusion model's latent space. This alternating optimization strategy allows the model to capture fine-grained textures while maintaining global stylistic consistency without experiencing mode collapse. The 'common parameter layer' is implemented as a cross-attention fusion block where the GAN-generated high-frequency feature maps are used as the 'key' and 'value' inputs, and the diffusion model's intermediate latent state acts as the 'query'. This synchronization ensures that the high-frequency brushstroke edges are structurally anchored to the low-frequency ink rheology, mathematically represented as

$$F_{\text{fused}} = \text{Attention}(Q_{\text{diff}}, K_{\text{GAN}}, V_{\text{GAN}})$$

to preserve the continuity of brush strokes and stylistic integrity. They are both joined together in a layer of common parameters in order to preserve the continuity of brush strokes and stylistic integrity. The number of brushstrokes and the field of ink diffusion can be changed dynamically during the generation phase, depending on the input semantic description and the style vector of the artist, thus generating ink style and adapting details structure guidedly.

D. PERSONALIZED CREATION MODULE

The human-computer collaborative process of art generation is accomplished by the personalized creation module with semantic instruction parsing and style embedding mapping. The user inputs the natural language description and style preference parameters in the system which are converted to a computable style control vector through a semantic parsing network. That is then fused with latent space features that are common to the brushwork features and ink generation module. Specifically, the manual control sliders act as dynamic bias weights within the latent space; rather than bypassing the generative intelligence, these parameters provide localized constraints that guide the diffusion model's denoising trajectory toward the user's specific aesthetic intent while maintaining the learned structural integrity of the artistic style. The model interacts by showing changes in brushstroke layers in real time, where the system uses an attention-based re-weighting mechanism to ensure that manual adjustments remain consistent with the high-dimensional style features learned during the automatic training phase. The style embedding model contains the statistical properties of the personal brushstroke by the artist such as the speed distribution, the brush tip angle, and the ink density curves. These attributes are combined with the existing semantic vector during generation and recreate a brush and ink style within a particular aesthetic preference. The system captures all the parameter adjustment trajectories during interaction and renews the user preference distribution using behavioral information to adapt to the next generation.

IV. RESULTS AND DISCUSSION

A. ACCURACY VERIFICATION

The model produces matching tones of ink and brushstroke rhythms in the output stage using these dynamic parameters, which are independent of each other, which is why each generation is based on a unique aesthetic intention and personal expression.

The accuracy of the models was measured by comparison of the shape of the brushstrokes and consistency rate by using the pixel-by-pixel matching of the handwriting images generated by the models with the manually annotated samples. The system output 100 images of each of the five samples of the calligraphy and painting styles and these images were structurally similar to their original data in that they had the same brushstroke outlines, ink shades and pen pressure changes. The calculation of errors was carried out using brushstroke edge offset distance, and grayscale similarity, which gave the reconstructions of the brushstroke in a comprehensive sense. The process of evaluation was done under the same circumstances of resolution and ink density to avoid the effect of external variables. The overall obtained average reconstruction accuracy was 93.8%. To supplement these objective metrics, a double-blind qualitative evaluation (Visual Turing Test) was conducted with a panel of five calligraphy experts. The experts rated the generated samples on a 10-point scale across three dimensions: ‘Brushstroke Vitality,’ ‘Ink Layering,’ and ‘Stylistic Authenticity.’. The brushstroke shape reconstruction accuracy verification data is shown in Table 2 .

TABLE 2. Verification Results of Brushstroke Shape Reproduction Accuracy

Style Category	Sample size	Average gray-scale similarity (%)	Edge offset error (px)	Reconstruction accuracy (%)	Ink color consistency (%)
Running script	100	94.2	2.3	93.5	92.8
Regular script	100	93.7	2.6	92.9	93.1
cursive script	100	95.4	2.1	94.7	94.0
landscape	100	96.2	1.8	94.9	95.3
Flowers and birds	100	95.1	2.0	93.9	94.5
Overall Average	–	94.9	2.2	93.8	93.9

Data shows that landscape paintings achieve the highest accuracy in reproduction, mainly because the variation in brushstroke width is relatively stable, making it easier for the model to capture continuous ink features. Cursive script works are similar, but due to the more complex variations in brushstrokes, there are slight deviations in the reproduction of local brush pressure. Regular script and running script have slightly higher errors, affected by the separation of fine brushstrokes, leading to increased edge offset and decreased similarity. The consistency of ink tone levels is generally maintained above 92 %, verifying the high-fidelity reproduction capability of the generation model in terms of ink texture under multiple style conditions.

B. STYLE GENERATION PERFORMANCE

To determine the style generation performance, our model was compared to a classic Convolutional Neural Network (CNN) model, and the performance evaluated concurrently. The samples of 5000 calligraphy and painting were used to train the two models so that the epochs of training and distribution of the samples used to train them were uniform. Each of the five styles had 200 samples in the test set. The similarity in the density of brushstrokes, the color of ink, and the texture of the style were determined in the pairs of results created and the original works to compare the degree of similarity and calculated the similarity using two indicators: feature distribution deviation and structural similarity index. The similarity score was a weighted combination of automatic rating and expert rating. In the same circumstances, our model recorded an average improvement in style similarity of 18.6 on average over CNN. The results indicated that 84.5% of the generated works were indistinguishable from original human-created samples, confirming that the model’s performance correlates strongly with professional aesthetic standards rather than just pixel-level similarity. The performance improvement was calculated based on the average similarity difference and the standardized deviation rate , as detailed in Table 3 .

TABLE 3. Performance Comparison Results of Style Generation

Style Category	Sample size	CNN style similarity (%)	Style similarity of this model (%)	Relative improvement (%)	Consistency score (out of 10)
Running script	200	73.9	87.4	18.3	8.6
Regular script	200	74.4	88.1	18.4	8.8
cursive script	200	75.7	90.5	19.6	9.1
landscape	200	76.2	91.3	19.8	9.3
Flowers and birds	200	76.4	89.7	17.4	9.0
Overall Average	–	75.3	89.4	18.6	8.9

Data shows that our method significantly outperforms CNN models across all calligraphy and painting styles. The most significant improvements are seen in landscape and cursive script, demonstrating the advantages of the diffusion model in learning the continuity of ink tones and the fluidity of brushstrokes. The improvements in running script and regular script are slightly lower, likely due to their more regular line structures, where CNNs have already demonstrated some ability to reconstruct these styles. The overall average similarity increased from 75.3% to 89.4%, indicating that the model integrating generative adversarial and diffusion mechanisms has stronger adaptability and expressive power in capturing the layers and individual characteristics of brush and ink styles, and can generate style outputs that closely resemble the realistic texture of artistic creation.

C. LIMITATIONS AND AREAS FOR IMPROVEMENT

While the current model achieves high performance in reproducing calligraphic styles and generating personalized content, it still has limitations in controlling the details of extremely complex styles and expressing deeper artistic intent. The model's ability to capture special brushstrokes such as rapid turns and overlapping layers of ink is insufficient, leading to slight distortions in the continuity of brushstrokes and the layering of ink color in some generated results. In high-dimensional feature learning, the model is easily affected by uneven data distribution, resulting in decreased stability in generating some rare calligraphic styles or individualized styles. The human-computer interaction process still needs optimization in terms of real-time response latency and the mapping accuracy of parameter adjustment ranges. Semantic parsing's understanding of artistic language is not yet deep enough to fully reproduce the creator's aesthetic intent. Future development could involve introducing a multimodal perception mechanism to integrate text, handwriting trajectories, and voice input to enhance the semantic-visual generation connection.

V. CONCLUSION

This study targets the digital generation and customized brush and ink styles. By using generative artificial intelligence in the creation of calligraphy and paintings, it confirms the possibility of machine learning in modeling brush and ink patterns and appropriating artistic styles for translation into complex jacquard structures. This technological integration ensures that the fluid nuances of traditional brushwork can be precisely mapped to the interlacing warp and weft density of high-end substrates. Spatial, hierarchical, and stylistic unity is achieved in the mapping dynamic information of brushstrokes and the rheological properties of ink color to create brush and ink expressions with aesthetic consistency as the artist desires.

The combination of the generative adversarial network and diffusion models allows the model to change dynamically in the situation of brush and ink details and style transfer learning. Together with human-computer interaction design based on semantic, the resultant output can react to single aesthetic requirements, delivering artistic properties that are more imitated to human creation. This approach offers a manageable intelligent generation way of creating calligraphy and painting, which facilitates the communication and fusion of human generation and algorithmic articulation. But when the stylistic variety is too great or the layers of brush strokes are too complicated, the model should be stronger in comprehending the relationships of the details in couple. Future studies may incorporate finer-grained multimodal modeling and cross-art category transfer mechanisms into the generative system to explore how brush and ink semantics connect with the complex interlacing structures of advanced jacquard weaving. This integration will provide deeper insights into translating artistic intent into precise digital loom parameters, ensuring that the fluid nuances of traditional calligraphy are accurately preserved within the woven textile substrate.

AUTHOR CONTRIBUTIONS

Conceptualization –Yinjie Li and Junle Hu; methodology – Yinjie Li and Junle Hu; investigation – Yinjie Li and Junle Hu; writing-original draft preparation – Yinjie Li and Junle Hu. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by, Fund Project: Art Planning Research Project of Shanxi Provincial Department of Culture and Tourism: Viewing the Development of Ancient Shanxi Painting from the Perspective of Calligraphy-Based on the Research of the Concept of "Same Origin of Painting and Calligraphy", Number: 24BG038

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Y. Chen, "Research on the Construction of Personalized Creative Teaching Model for Primary School Art under the Background of Artificial Intelligence," China New Communications, vol. 27, no. 15, pp. 239-241, 2025, doi: 10.3969/j.issn.1673-4866.2025.15.082.

[2] P. Wu, "A Study on the Application of Personalized Teaching in Primary School Chinese Composition Instruction," Teaching Management and Educational Research, vol. 8, no. 12, pp. 28-30, 2023, doi: 10.3969/j.issn.2096-224X.2023.12.009.

[3] Z. Jiang and P. Liao, "Research on the Personalized Development and Expression of Contemporary Ceramic Flower-and-Bird Painting," Jingdezhen Ceramics, vol. 51, no. 6, pp. 72-74, 2023, doi: 10.3969/j.issn.1006-9545.2023.06.023.

[4] S. Chen, "A Study on the Application of Paper in Painting Creation," China Paper Industry, vol. 44, no. 23, pp. 97-99, 2023, doi: 10.3969/j.issn.1007-9211.2023.23.039.

[5] P. Pataranutaporn, V. Danry, J. Leong, et al., "AI-generated characters for supporting personalized learning and wellbeing," Nature Machine Intelligence, vol. 3, no. 12, pp. 1013-1022, 2021, doi: 10.1038/s42256-021-00417-9.

[6] E. Cetinic and J. She, "Understanding and creating art with AI: Review and outlook," ACM transactions on multimedia computing, communications, and applications (TOMM), vol. 18, no. 2, pp. 1-22, 2022, doi: 10.1145/3475799.

[7] M. Masood, M. Nawaz, M. Malik K, et al., "Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way

forward: Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward," Applied intelligence, vol. 53, no. 4, pp. 3974-4026, 2023, doi: 10.1007/s10489-022-03766-z.

[8] J. Oppenlaender, "A taxonomy of prompt modifiers for text-to-image generation," Behaviour & Information Technology, vol. 43, no. 15, pp. 3763-3776, 2024, doi: 10.1080/0144929X.2023.2286532.

[9] J. Oppenlaender, R. Linder, and J. Silvennoinen, "Prompting AI art: An investigation into the creative skill of prompt engineering," International journal of human–computer interaction, vol. 41, no. 16, pp. 10207-10229, 2025, doi: 10.1080/10447318.2024.2431761.

[10] Y. Wang, X. Chen, X. Ma, et al., "Lavie: High-quality video generation with cascaded latent diffusion models," International Journal of Computer Vision, vol. 133, no. 5, pp. 3059-3078, 2025, doi: 10.1007/s11263-024-02295-1.