Automatic generation of Chinese mural line drawings via enhanced edge detection and CycleGAN-based denoising

Check for updates

Haixia Feng[1], Qingwu Hu[2] ✉, Pengcheng Zhao[1], Daoyuan Zheng[1], Mingyao Ai[1], Siliang Chen[3] & Xiyu Hu[4]

Murals are a vital form of traditional Chinese art, rich in historical and cultural content. Line drawing, as a core technique, is widely used but still depends on manual tracing. This paper proposes an automatic method for generating mural line drawings by integrating edge enhancement, neural edge detection, and denoising. Enhance edges using image processing techniques, followed by a neural network (MLineNet) for line extraction. A cycle-consistent generative adversarial network (CycleGAN) refines the output by removing noise while preserving structural clarity. The model was evaluated using four metrics: structural similarity index (SSIM), texture complexity (TC), line connectivity index (LCI), and a comprehensive score (Q). On Dunhuang murals, it achieved scores of 89.54%, 93.77%, 88.14%, and 90.48%, respectively, and showed generalization to Baisha murals (Q = 89.29%). Results demonstrate the method's reliability in producing complete, clean, and continuous mural line drawings.

Murals, as an essential form of traditional Chinese painting, contain rich historical, cultural, and artistic information, vividly recording the social customs and folk beliefs of multiple ethnic groups across different dynasties[1–3]. As illustrated in Fig. 1d, mural creation typically starts with outlining shapes through line drawing, followed by color rendering and further artistic refinement. Mural line drawings, characterized by their simple ink lines without embellishment or color rendering, utilize variations in line length, connectivity, and smoothness to express the texture and spirit of depicted subjects, thus forming distinct visual rhythms and becoming an independent artistic form. Therefore, line drawings serve as a critical representational technique in Chinese painting, indispensable in Chinese traditional painting, mural art, and folk painting[4,5]. Currently, line drawings of murals are primarily produced by manual tracing. Artists usually perform field mapping with simple tools under limited lighting conditions, which is time-consuming and requires extensive training. Although the introduction of digital technologies such as 3D laser scanning and digital orthophotography imagery has improved efficiency, the drawing process still largely depends on manual work by skilled professionals. Despite gains in productivity, substantial labor and time investments remain necessary. Furthermore, variations in artistic perception and technical skills among artists result in inconsistent drawing quality, and the limited number of qualified professionals further constrains productivity. Therefore, there is an urgent need for advanced technological solutions to significantly enhance the efficiency and consistency of mural line drawing generation.

Line drawing generation involves extracting line information from color murals to generate black-and-white images composed solely of lines and blank spaces. Mathematically, line drawing generation can be formulated as a pixel-level classification problem[6], where each pixel of a two-dimensional image is classified according to its probability of belonging to line structures. Similar to edge images, mural line drawings depict object contours through black-and-white pixels, emphasizing line and shape representation. However, unlike traditional edge detection tasks[7–9], which typically focus on high-contrast edges in natural images and consequently lack detailed line features, mural line drawing generation emphasizes the details of lines and textures. Therefore, there is a significant difference between mural line drawing generation and natural image edge detection. Edge detection in natural images relies heavily on large public datasets for training and mainly captures object contours. In contrast, mural data are scarce, and the extraction of mural line drawings requires capturing richer structural features and artistic details, such as intricate clothing patterns (garments, headwear, etc.) and detailed facial expressions (eyebrows, eyes, etc.). Particularly, mural line drawing generation demands more precise and expressive lines aligned with traditional artistic styles. As depicted in Fig. 1a–c, deep learning-based edge detection methods can be categorized into three types: (1) general edge detection[10,11], (2) object contour detection[12,13],

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China. [2]School of Remote Sensing and Information Engineering, Hubei Luojia Laboratory, Wuhan University, Wuhan, China. [3]School of Architecture, Chang'an University, Xi'an, China. [4]College of Art, University of Edinburgh, Edinburgh, UK. ✉e-mail: huqw@whu.edu.cn
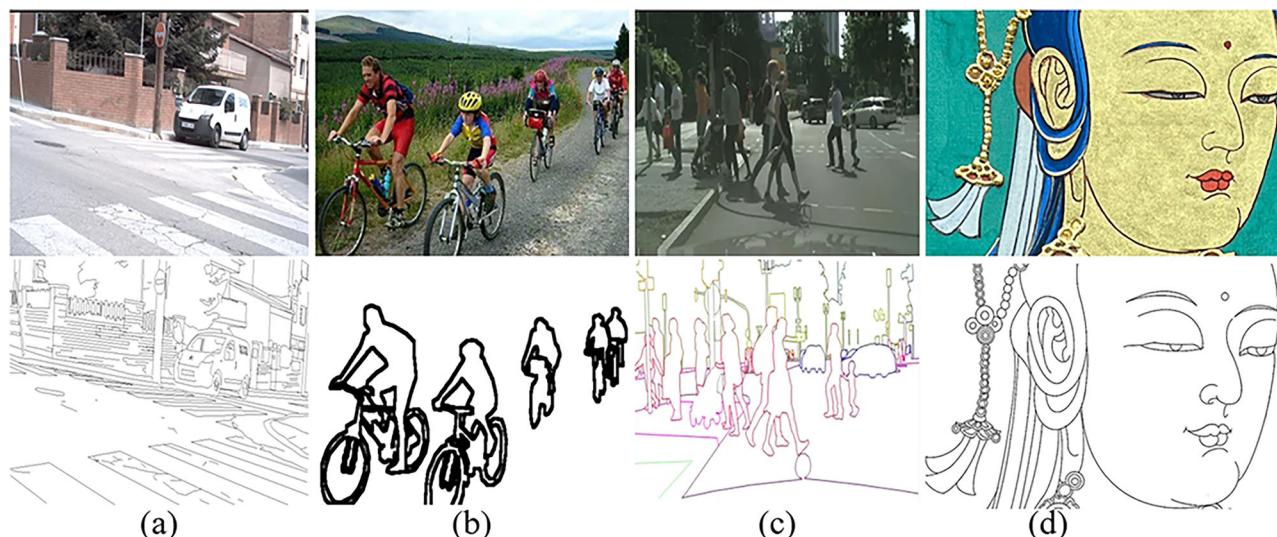
Fig. 1 | Edge detection and mural line drawings based on deep learning. a General edge detection, **b** object contour detection, **c** semantic edge detection, and **d** mural line drawings.

and (3) semantic edge detection[14,15]. The general edge detection is to identify significant pixel-value changes in images, resulting in detection outputs that contain considerable noise due to the inclusion of all edges. Object contour detection typically simplifies results and excludes internal texture or detail. Semantic edge detection yields concise edges by filtering out irrelevant details. While edge detection in natural images emphasizes contour clarity and object recognition, mural line drawings require finer, artistically meaningful line generation, as demonstrated in Fig. 1d.
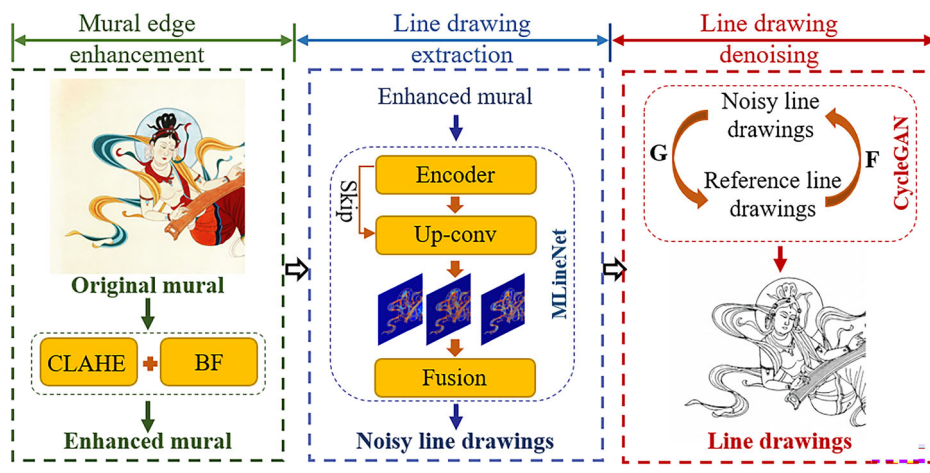
In recent years, edge detection techniques have significantly advanced in the fields of computer vision and image processing. Traditional approaches, such as Canny[16] and Sobel[17], remain widely used and have been optimized across various applications[18,19]. With the rise of deep learning, CNN-based edge detection methods have become mainstream[20], considerably improving accuracy and robustness[21,22]. Recently, Transformer structures have also been introduced into edge detection tasks, showing significant advantages over CNN-based methods in capturing long-range dependencies[23]. Moreover, edge detection methods based on generative adversarial networks (GANs) have been developed to generate clearer edge maps and effectively suppress background noise. Additionally, edge detection integrated with hardware optimization has become an important trend, showing great practical value in fields such as intelligent monitoring and autonomous driving[24,25].

For line drawing generation, early studies primarily relied on hand-crafted features, interactive methods, and various edge detection operators[26–28], such as the Canny and Sobel operators, to identify adjacent regions in images, forming continuous lines that outline the structural curves of murals. Liu J et al.[29] focused on the Dunhuang Mogao Grotto murals and employed computer-aided interactive techniques to first trace the mural boundaries and then refine the line drawings using a virtual brush. Similarly, He et al.[30] also studied Dunhuang Mogao Grotto murals and adopted a layered interactive approach to progressively generate stroke-based line drawings. In the study by Sun et al.[31], mural image data were sourced from true-color mural images provided by the Dunhuang Academy and publicly available mural artifact images, and 12 real line drawings were manually created. Their method initially applied heuristic routing to detect stroke outlines, followed by high-frequency enhancement filtering to extract internal stroke details, and finally, cooperative representation to generate complete strokes. Although these methods can produce line drawings that meet certain artistic requirements, they heavily rely on manual interaction, resulting in low efficiency and difficulty in handling large-scale data or automated generation tasks.

Benefiting from the powerful feature representation capabilities of deep learning, recent research has focused on designing sophisticated network architectures to learn high-level semantic representations[32–35]. Deng et al.[36] proposed a deep structural contour detection framework that integrates edge location and orientation information into a novel loss function, achieving improved accuracy in extracting structural contours from images. In the study by Wang et al.[37], the research focused on Thangka murals, where an edge detection model trained on the natural image BIPED dataset was directly applied to generate Thangka line drawings. Similarly, Wang et al.[38] focused on Dunhuang murals and generated mural line drawings using an edge detection model trained on the BIPED dataset. Their approach introduced a focal Tversky loss function to suppress background pixels near edge pixels and incorporated dilated convolution and spatial attention modules to enhance hierarchical features and enrich scale information. Liu et al.[39] also studied Dunhuang murals, using an edge detection model trained on a natural image dataset to generate mural line drawings. Their approach proposed a residual self-attention and convolution hybrid module to integrate local and global features, thereby improving the extraction of mural line drawings. Peng et al.[40] focused on cultural heritage murals and used deep learning models for line drawing extraction. Their approach introduced a detail-aware bidirectional cascaded network combined with a multi-scale U-Net framework to enhance feature extraction capabilities. In the study by Yu et al.[41], simulated mural samples were created using real mural production techniques to generate realistic line drawings. Their method enhanced mural images and employed Laplacian edge detection combined with a fine noise removal module to extract optimized line drawings. It is evident that due to the lack of publicly available datasets for mural line drawings, most current studies on mural line drawing generation rely on models trained and evaluated using natural image datasets. While this approach is feasible, it has inherent limitations in capturing the rich structural details and artistic features of murals. Mural line drawings require careful consideration of the smoothness and detailed representation of each stroke, imposing higher demands on line extraction and expression. Most existing edge detection methods are not specifically designed for mural line drawing generation, making them insufficient to fully capture mural-specific details and artistic styles. Therefore, effectively adapting existing methods for mural line drawing generation remains a pressing challenge that needs to be addressed.

To address the above challenges, this study proposes an automatic mural line drawing generation method, incorporating edge enhancement, line drawing extraction, and denoising technologies, enabling efficient and accurate extraction of mural lines. Specifically, this method first enhances the mural images' edge information using contrast-limited adaptive histogram equalization (CLAHE) and bilateral filtering (BF), providing high-

Fig. 2 | The overall workflow of mural line drawing generation.

quality inputs for subsequent line extraction. Secondly, a neural network for mural line drawings (MLineNet) is employed to extract mural edge features. Finally, a cycle-consistent generative adversarial network (CycleGAN) is introduced to refine and denoise the extracted lines, resulting in high-quality line drawings with clear, natural appearance and rich structural details. Using Dunhuang murals as a primary research subject, a mural line-drawing generation model was constructed based on the proposed method. The reliability and effectiveness of this model were quantitatively verified using multiple evaluation indicators, including structural similarity index (SSIM), texture complexity (TC), line connectivity index (LCI), and a comprehensive score (Q). The final results demonstrate that the proposed method achieves balanced performance in structural similarity, cleanliness, and connectivity of the image, showing good overall effectiveness.

The structure of this paper is as follows: The "Methods" section introduces the overall process of mural line drawing generation and details three key components, including mural edge enhancement, line drawing extraction, and line drawing denoising. The "Results" section presents the generation outcomes of Dunhuang murals, analyzes the performance of different structures, compares the proposed method with other methods, and evaluates our method's generalization capability through line drawing generation from Baisha murals. Finally, the "Discussion" section summarizes the findings and limitations of this study.

Methods
Overview
The Dunhuang Mogao Grottoes[42,43], located in Dunhuang, Gansu Province, China, are one of the largest and best-preserved repositories of Buddhist art in the world. Established in the 4th century Anno Domini, they have undergone continuous development over a millennium and currently consist of 492 caves, with mural paintings covering an area of over 50,000 square meters. The murals encompass a wide range of themes, including Buddhist figure paintings, transformation paintings, narrative paintings, and donor portraits, reflecting not only the spread and evolution of Buddhist beliefs but also the social and artistic styles of different historical periods. Buddhist figure paintings depict various representations of Buddhas, Bodhisattvas, and celestial kings, characterized by smooth lines and vibrant colors, showcasing artistic features unique to each dynasty. Transformation paintings illustrate Buddhist scriptures in visual form, making the teachings more accessible and easier to disseminate. Narrative paintings depict stories from the Buddha's past lives and karmic tales, promoting Buddhist doctrines and moral principles. Donor portraits portray the patrons who financed the construction of the caves, providing insight into the social structure and clothing culture of the time. These murals are not only treasures of religious art but also serve as invaluable historical, cultural, and social records, holding immense artistic and academic significance. DhMurals1714[44] is a dataset specifically designed for the study of Dunhuang

Mogao Grottoes murals. It comprises 1714 images along with incompletely annotated reference line drawings, aiming to address the limitation of available data caused by natural weathering and human-induced damage to the murals.

This paper proposes an automated method for generating mural line drawings, integrating mural edge enhancement, deep learning-based line extraction, and generative adversarial network-based denoising techniques to efficiently and accurately generate mural line drawings. Based on the DhMurals1714 dataset, this method seeks to mitigate issues of detail loss and blurring in reference line drawings of Dunhuang murals, enabling precise and high-quality extraction of mural line features. As shown in Fig. 2, the proposed approach consists of three key steps: mural edge enhancement, line drawing extraction, and line drawing denoising. First, to improve the quality of mural image edges, CLAHE and BF are applied[45], enhancing edge details in the initial images and providing high-quality inputs for subsequent line drawing extraction. Then, in the line drawing extraction stage, the deep learning network MLineNet is employed to extract mural line drawings. This network adopts an encoder-decoder structure, incorporating multi-scale feature fusion and edge-aware optimization strategies to capture lines at various scales while preventing excessive sharpening or information loss. MLineNet generated line drawings effectively reproduce the primary line drawings of murals; however, due to annotation errors and the incompleteness of reference line drawings, noise and false edges still exist, affecting the artistic quality and usability of the extracted lines. To further refine the line quality, this method introduces CycleGAN for unsupervised denoising and correction[46]. Given the lack of complete and precisely paired ground-truth line drawings, CycleGAN effectively removes noise and restores line smoothness and integrity through unsupervised learning, using only noisy line drawings and incomplete reference line drawings as input. As a result, the CycleGAN refined line drawings not only preserve the structural information of complete lines but also eliminate redundant noise and false edges, making the generated lines more detailed, clear, and natural, aligning with the esthetic standards of mural artwork.

Mural edge enhancement
Although the reference line drawings in the Dunhuang dataset display the edge structures of images, they suffer from partial detail loss or blurring due to inaccurate annotations. To address this issue, this study applies CLAHE and BF for edge enhancement before line drawing extraction, as shown in Fig. 3. CLAHE and BF are important image enhancement techniques widely used in tasks that improve image details and edge features. In mural image processing, the combination of CLAHE and BF achieves superior edge enhancement effects. CLAHE enhances local brightness differences to make edges more prominent, while BF smooths noise and removes artifacts introduced during enhancement, while preserving the edge features enhanced by CLAHE. The synergy of these techniques demonstrates
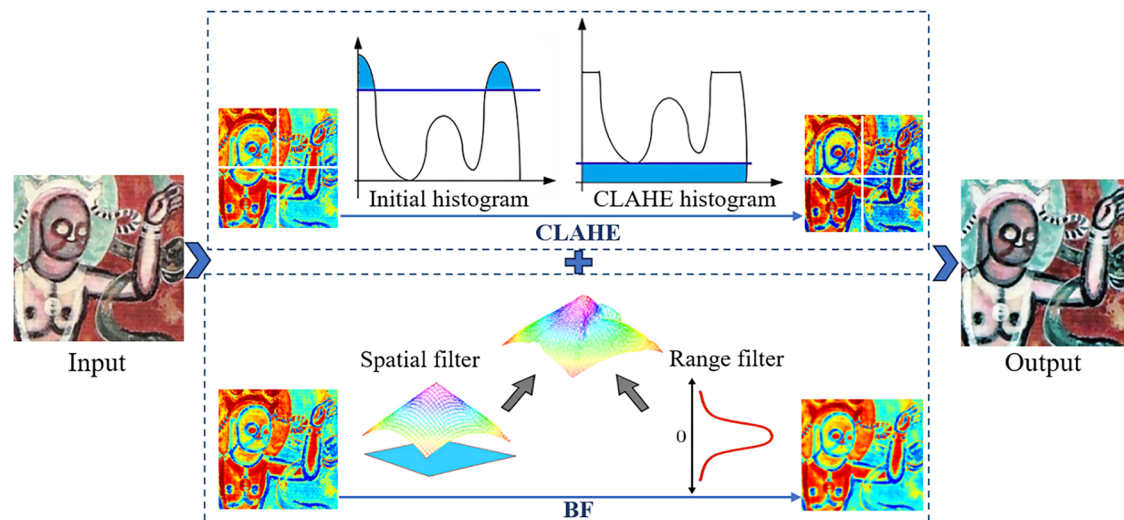
Fig. 3 | Visualization of mural edge enhancement.

outstanding performance in mural image edge enhancement, providing a solid foundation for subsequent line drawing extraction.

In mural image processing, CLAHE enhances edge features, making edge information clearer while preventing excessive enhancement in high-contrast areas. CLAHE is an improved version of the traditional histogram equalization method, addressing the issue where HE may introduce excessive enhancement and artifacts in high-contrast regions. The key feature of CLAHE is its localized approach, dividing the image into small contextual regions (sub-blocks) for histogram equalization while applying a contrast threshold to each region's pixels. This prevents the unevenness that global enhancement may introduce. Its formula is as follows:

$$H'(i) = \min(H(i), T) + \frac{\sum_j \max(H(j) - T, 0)}{N} \quad (1)$$

where $H'(i)$ represents the adjusted pixel value distribution in the histogram, T is the contrast threshold, and N is the total number of pixels.

BF preserves edge structures while smoothing image noise by integrating geometric distance and pixel intensity similarity, preventing the edge blurring often caused by traditional linear filtering methods. As a nonlinear filtering technique, BF combines information from the spatial domain and the intensity domain to smooth images while effectively preserving edge features. The core idea is to compute a weighted average of each pixel and its neighboring pixels, where the weight is determined by both spatial distance and intensity differences. The mathematical formulation of bilateral filtering is as follows:

$$I'(x) = \frac{\sum_{y \in \mathcal{N}(x)} G_s(\| x - y \|) G_r(|I(x) - I(y)|) I(y)}{\sum_{y \in \mathcal{N}(x)} G_s(\| x - y \|) G_r(|I(x) - I(y)|)} \quad (2)$$

where $I(x)$ is the pixel intensity at location x in the input image, $\mathcal{N}(x)$ represents the neighborhood of x, $G_s(\| x - y \|)$ is the Gaussian function in the spatial domain, which measures geometric distance between pixels, and $G_r(|I(x) - I(y)|)$ is the range filter, which evaluates intensity or color similarity between pixels. This ensures that neighboring pixels with similar intensities have a greater influence on the central pixel, while regions with significant intensity differences have less influence, thereby preserving edge information.

Line drawing extraction

This study employs MLineNet (as shown in Fig. 4) as the deep learning network for mural line drawing extraction, aiming to automatically produce high-quality line drawings to support cultural heritage digitization, artistic style transfer, computer-assisted painting, and heritage preservation applications. Mural line drawings are not only a form of artistic representation but also essential data for mural restoration, analysis, and research. Existing deep learning-based edge detection models[10,35,47] have demonstrated good performance in general edge detection tasks for natural images. These methods often rely on global context modeling, denoising processes based on generative models, uncertainty modeling, or high-precision operator design to improve accuracy and robustness in edge detection. However, these models are mainly designed for general object boundary detection tasks, with a focus on edge localization accuracy and model robustness. Compared with existing deep learning-based line drawing extraction methods[37–42], MLineNet is designed with the specific characteristics of mural line drawing tasks. It uses an encoder-decoder architecture combined with a multi-scale branch design (H_1, H_2, H_3) and feature fusion to separately extract coarse contours, local details, and global structural features. MLineNet places more emphasis on the clean, connectivity, and structural integrity of lines, rather than solely on edge detection accuracy.

Specifically, MLineNet includes three feature extraction stages within the encoder–decoder architecture, each responsible for capturing coarse contours, enhancing local details, and optimizing overall structural integrity. It incorporates a multi-scale feature fusion strategy to enhance the network's capacity to represent details at different scales. The network takes an input mural image of size $720 \times 720 \times 3$ and processes it through three stages to progressively extract features at different scales. In the feature extraction phase, the input first passes through two 3×3 convolution layers to extract local features and capture the basic edge structure of the mural. Max pooling (3×3) is then used to gradually reduce the spatial resolution of the feature maps and capture higher-level global features. For instance, the first stage increases the number of channels to 16 and reduces the feature map size to 360×360, while deeper layers further reduce the size to 180×180 with an increase in channels, ensuring the model can learn line information at different scales. Since mural line drawings contain rich fine details, such as facial features, clothing patterns, and decorative backgrounds, simple pooling operations may lead to a loss of detail. Therefore, 1×1 convolutions are introduced in certain layers for channel compression to preserve detailed features while controlling computational cost. In addition, residual connections are used within the encoder–decoder structure to establish information pathways, allowing lower-level features to be directly transmitted to subsequent layers. This helps mitigate feature loss in deep networks and enhances the retention of detailed lines.

In the decoding phase, MLineNet employs upsampling layers (UpConv) to gradually restore the feature maps to the original input resolution (720×720), ensuring that the output line drawing aligns spatially
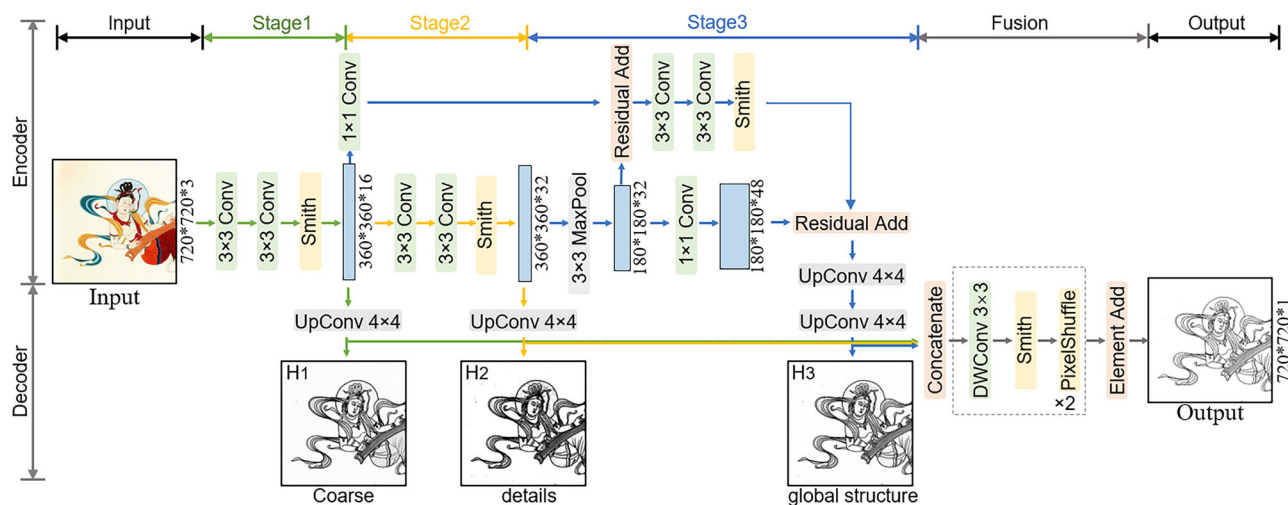
Fig. 4 | The architecture of the mural line drawing extraction network, named MLineNet.

with the input image and retains as much detail as possible. However, simple upsampling is not sufficient to produce high-quality line drawings. To further enhance line quality, MLineNet introduces a feature fusion module during decoding. This module concatenates the feature maps from the different scale branches (H1, H2, H3) as input, where H1 captures coarse contours, H2 enhances local details, and H3 specifically optimizes overall structural integrity, supplementing global framework information and reducing line discontinuities and local deviations. The fusion module consists of depthwise separable convolutions (DWConv, 3×3), Smish activation, and PixelShuffle operations. DWConv performs per-channel convolutions to extract local features with low computational cost, while PixelShuffle increases resolution and rearranges the feature structure to preserve detail distribution. Finally, these three features are integrated through a fusion layer to ensure that the generated line drawing maintains global coherence while preserving local detail. The fusion layer is defined as follows:

$$X = \text{Concatenate}(H_1, H_2, H_3) \qquad (3)$$

$$X' = \text{PS}(W_d * \text{Smith}(X)) \qquad (4)$$

$$Y = X' + \text{PS}(W_d * \text{Smith}(X')) \qquad (5)$$

where X represents the concatenated result, H_1, H_2, H_3 are features extracted at different scales, W_d denotes the depthwise separable convolution kernel, Smish is the activation function, and PS represents PixelShuffle, which performs pixel rearrangement. X' is the preliminary processed feature map, and Y is the final fused feature map.

Line drawing denoising

In the automatic generation of mural line drawings, the reference line drawings often exhibit missing or incomplete lines in certain areas. Directly relying on these incomplete annotations for training may lead to a decline in the quality of the generated line drawings, affecting their final visual effect and subsequent applications. To address this issue, we first employed image enhancement techniques and MLineNet during training to extract preliminary line drawings that compensate for the missing information in the reference data. However, although MLineNet can extract relatively complete line structures using deep learning techniques, the imperfections in the reference line drawings may still introduce noise during the extraction process. This can lead to blurred lines in some areas, discontinuous edges, or excessive pseudo-edges in the background. The presence of noise not only affects the overall clarity of the line drawings but may also interfere with subsequent mural studies, artistic style analysis, or digital archiving.

Therefore, to further optimize the quality of the generated line drawings, this paper introduces the CycleGAN model to denoise and refine the initially generated noisy line drawings, producing clearer, smoother, and structurally complete mural line drawings.

The advantage of CycleGAN lies in its ability to learn the mapping between domains through unsupervised learning from unpaired data. Traditional image denoising methods typically rely on strictly paired datasets, where each noisy image must correspond to a clean target image for the model to learn the denoising mapping. However, in the task of generating mural line drawings, due to the complexity and uniqueness of historical artifacts, it is often nearly impossible to obtain perfectly matched, clean line drawings as supervision signals, making paired data nearly infeasible. Against this backdrop, CycleGAN introduces a cycle consistency loss, constructing a closed-loop structure that maps from the noisy domain to the clean domain and then back again, ensuring that the input image can be as close as possible to the original after two mappings. This mechanism allows the model to effectively learn denoising and structural restoration capabilities even in the absence of paired supervision, making it particularly suitable for practical applications like mural line drawings, where obtaining ideal training data is challenging. Specifically, as shown in Fig. 5, CycleGAN consists of two generators and two discriminators. The input noisy line drawing (source domain X) and the reference line drawing (target domain Y) are not strictly paired. The generator G maps the source domain X to the target domain Y, while the inverse generator F maps the generated target domain Y back to the original source domain X. The training objectives for generators and discriminators are defined by adversarial loss, where $\mathcal{L}_{\text{GAN}}(G, D_Y)$ represents the adversarial loss between generator G and discriminator D_Y measuring the realism of the generated image $G(x)$. Similarly, $\mathcal{L}_{\text{GAN}}(F, D_X)$ represents the adversarial loss between the inverse generator F and discriminator D_X, evaluating the authenticity of the generated image $F(y)$. The adversarial loss function is defined as follows:

$$\mathcal{L}_{GAN}(G, D_Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x)))] \qquad (6)$$

$$\mathcal{L}_{GAN}(F, D_X) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D_X(x)] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log(1 - D_X(F(y)))] \qquad (7)$$

where x is the source domain image (noisy line drawing), y is the target domain image (reference line drawing), and $p_{\text{data}}(x)$ represents the data distribution of source domain images, and $\mathbb{E}$ is the expectation operator, which denotes the expected loss over the images. $\mathbb{E}_{x \sim p_{\text{data}}(x)}$ represents the expectation calculation over samples x drawn from the probability
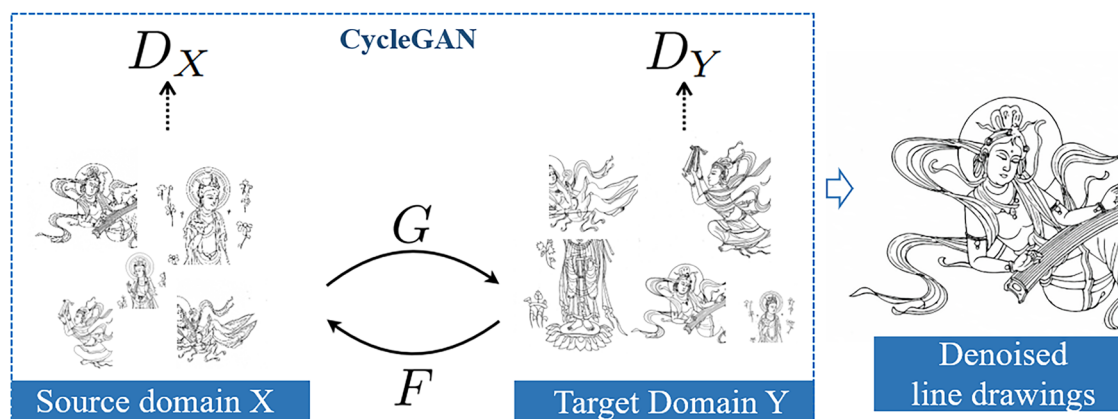
Fig. 5 | Line drawing denoising using CycleGAN.

distribution $p_{data}(x)$, and $\mathbb{E}_{y \sim p_{data}(y)}$ follows the same principle. D_X is the discriminator responsible for determining whether an image belongs to the source domain X, and D_Y serves the same function for the target domain Y.

To ensure a reasonable transformation from the source domain to the target domain, CycleGAN introduces cycle consistency loss. This loss ensures that the generated image maintains structural and content consistency with the input image by mapping the transformed image back to its original domain using the inverse generator F, then computing the difference between the original and reconstructed images. The cycle consistency loss function is defined as follows:

$$\mathcal{L}_{cyc}(G,F) = \mathbb{E}_{x \sim p_{data}(x)}[\| F(G(x)) - x \|_1] + \mathbb{E}_{y \sim p_{data}(y)}[\| G(F(y)) - y \|_1]$$

(8)

where $F(G(x))$ represents the source domain image x mapped through generator G and then back through inverse generator F, and $G(F(y))$ represents the target domain image y mapped through F and then back through G. The L_1 norm ($\| \cdot \|_1$) is used to calculate the pixel-wise difference between images. The final loss function consists of both adversarial loss and cycle consistency loss:

$$\mathcal{L}(G,F,D_X,D_Y) = \mathscr{L}_{GAN}(G,D_Y) + \mathscr{L}_{GAN}(F,D_X) + \mathscr{L}_{cyc}(G,F) \quad (9)$$

By leveraging CycleGAN, this study restores clear and complete mural line drawings from noisy and incomplete images. Through adversarial training and cycle consistency loss, the CycleGAN model effectively removes noise while preserving the structural integrity of the images, ultimately generating high-quality denoised and restored mural line drawings.

Results

In this part of the study, we conducted comprehensive experiments on the DhMurals1714 dataset and further performed generalization validation on the Baisha mural dataset to evaluate the reliability of the proposed automated mural line drawing generation method. The experiments include both qualitative analysis and quantitative comparisons to thoroughly assess the effectiveness of the approach. This section provides a description of datasets, implementation details, and results analysis.

Datasets description

DhMurals1714 is a dataset specifically designed for the study of murals from the Mogao Caves in Dunhuang. It contains a total of 1714 images, each with a resolution of 720×720 pixels[44]. The dataset was created to address the scarcity of usable mural data caused by natural weathering and human damage. Following a random sampling strategy, the dataset was divided into a 7:3 ratio for training and testing, with 1200 images used for training and 514 for testing.

The Baisha mural dataset originates from the Dading Pavilion murals located in Baisha Town, Yulong Naxi Autonomous County, Lijiang City, Yunnan Province[48,49]. These murals serve as an important source for the study of religious art in Southwest China during the Ming and Qing dynasties and have recently undergone digital preservation and scholarly analysis. The original mural image has a resolution of 5251×7406 pixels. By applying a sliding window of 720 pixels with a stride of 680, 70 sub-images of size 720×720 pixels were extracted, all of which were used for testing purposes.

Implementation details

The experiments were conducted using an NVIDIA A40 GPU, with PyTorch as the deep learning framework. For training the MLineNet model, we adopted the Adam optimizer to ensure training stability and convergence efficiency. The initial learning rate was set to 0.005 and was dynamically adjusted using a cosine annealing schedule, allowing the learning rate to gradually decrease during training to accommodate multi-scale feature optimization. To improve the model's generalization ability on the DhMurals1714 dataset, we designed a comprehensive data augmentation scheme. Each training image was augmented with three types of rotation (90°, 180°, and 270°), three scaling factors (2.0, 1.2, and 0.4), and two gamma correction values (0.3030 and 0.6060), resulting in 8 augmented samples per original image. Based on 1200 training images, a total of approximately 10,800 samples (720×720 pixels) were generated to simulate the diverse styles of mural line drawings, thereby enhancing the model's adaptability to unseen data. The training process involved pretraining the model on the BIPED dataset, followed by fine-tuning on the DhMurals1714 dataset to better capture the structural features of Dunhuang mural line art. A batch size of 8 was used, with a total of 60 epochs. An early stopping strategy was applied after 30 epochs to prevent overfitting.

In the training of the CycleGAN model, modeling was also based on the aforementioned 1200 training images with a size of 720×720 pixels. To clearly define the source domain (noisy domain) and the target domain (clean domain), we consider the preliminary line drawings as the source domain images (i.e., extracted line drawings with more noise, broken lines, or redundant elements), and the reference line drawings as the target domain images (i.e., reference line drawings) used to guide the generator in learning the denoising mapping. In terms of network structure, the generator adopts a ResNet architecture, while the discriminator uses a Patch-GAN network. For loss function design, CycleGAN incorporates both adversarial loss and cycle consistency loss, with the adversarial loss weight set to 1.0 and the cycle consistency loss weight set to 10.0. The optimizer is Adam with an initial learning rate of 0.0002 and a batch size of 2. Training is conducted over a total of 100 epochs, with the learning rate remaining constant for the first 50 epochs and then linearly decaying to 0 over the remaining 50 epochs, in order to stabilize the training process and improve final performance.

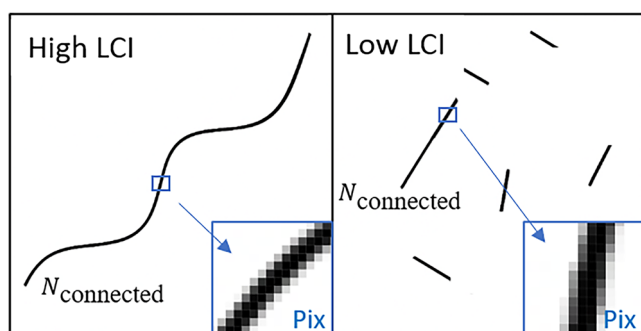**Fig. 6 |** Differences between high and low TC.

Fig. 7 | Differences between high and low LCI.

Evaluation metrics

To objectively and accurately evaluate the quality of the generated mural line drawings, multiple quantitative evaluation metrics are selected in this study. These metrics comprehensively assess various aspects of the image, including structural similarity, texture complexity and line connectivity, providing a strong basis for model optimization. Specifically, the selected metrics include SSIM, TC, LCI, and Q, which integrates all three indicators.

SSIM[50] evaluates the fidelity of the generated image to the original mural image in terms of luminance, contrast, and structural information, reflecting the model's accuracy in overall structural restoration. A higher SSIM value indicates greater structural similarity and perceptual quality. The SSIM is calculated as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (10)$$

where μ_x and μ_y are the mean intensities of images x and y, σ_x and σ_y are their standard deviations, and σ_{xy} is their covariance. Constants c_1 and c_2 are used to avoid division by zero. SSIM values typically range from 0 to 1, with values closer to 1 indicating higher structural and perceptual consistency.

TC evaluates the simplicity of line patterns in the image. This metric is derived from the contrast feature of the gray-level co-occurrence matrix (GLCM) proposed by Haralick et al. [51]. Higher contrast indicates greater gray-level variation and more complex textures. Figure 6 shows two 4×4 GLCM examples illustrating structural differences between high and low TC cases. The high-TC matrix (left) is concentrated near the diagonal, indicating smooth transitions and orderly structures, while the low-TC matrix (right) shows more values far from the diagonal, reflecting abrupt gray changes and noisy patterns. To ensure TC increases with visual simplicity, we apply an inverse normalization to the contrast, making the TC value

positively correlated with cleanliness:

$$TC = \frac{1}{1 + \text{Contrast}} \times 100\% \quad (11)$$

$$\text{Contrast} = \log\left(1 + \sum_{i,j}(i - j)^2 P(i, j)\right) \quad (12)$$

where $P(i, j)$ represents the gray-level co-occurrence matrix, and i and j are the grayscale values of the image pixels.

LCI measures the connectivity and completeness of lines in the image. Based on the theory of connected regions in image processing, it assesses whether the lines form a single or multiple complete structures[52]. The core idea is to calculate the ratio of the largest connected line region to the total number of line pixels. Figure 7 illustrates two examples of different LCI levels: the high-LCI image has continuous lines forming one main connected region with minimal gaps or noise; the low-LCI image contains fragmented lines with multiple disconnected segments. To better capture local structural consistency in large images, the image is divided into a 9×9 grid and the local LCI is computed in each region, with the final LCI being their average:

$$\text{LCI} = \frac{N_{\text{connected}}}{N_{\text{total}}} \quad (13)$$

where $N_{\text{connected}}$ is the number of pixels in the largest connected region and N_{total} is the total number of line pixels. The LCI ranges from 0 to 1, with higher values indicating better connectivity and clearer line structures.

Finally, the comprehensive evaluation Q is the average of the three aforementioned metrics, providing an overall quality assessment. The closer Q is to 1, the better the model's overall performance across all indicators. The calculation formula is as follows:

$$Q = \frac{\text{SSMI} + \text{TC} + \text{LCI}}{3} \times 100\% \quad (14)$$

Mural edge enhancement

This study applies edge enhancement processing to Dunhuang mural data prior to line drawing extraction using CLAHE and BF, as shown in Fig. 8. Specifically, (a) represents the original mural, while (b) is the corresponding heatmap, where colors indicate the magnitude of gradient variations—red regions denote higher gradient changes, while blue regions also indicate lower gradient variations. (c) presents the heatmap after CLAHE processing, which enhances local contrast and highlights edge details. (d) shows the result of applying BF filtering after CLAHE, where bilateral filtering effectively reduces texture noise, leading to smoother and more continuous edge information. Finally, (e) is the enhanced mural, in which the primary structural features become more distinct, and edge contrast is improved, providing a more precise input for subsequent line drawing extraction tasks. It is evident that the original heatmap (b) exhibits lower gradient edge structures of the mural, particularly in finer details such as garment textures, ornaments, and background patterns corresponding to hand gestures, where gradient variations are weak, making edge extraction challenging. After CLAHE processing (c), edge details are significantly enhanced, with more pronounced gradient variations in clothing patterns and character outlines, facilitating better feature distinction. Further integration of BF filtering (d) reduces edge noise to some extent, while maintaining relatively smooth and continuous contours, mitigating the texture interference caused by excessive enhancement. The final enhanced mural (e) not only preserves the original artistic characteristics of the mural but also effectively enhances the representation of primary structural elements, making line features more prominent and laying a solid foundation for high-precision line drawing extraction.

Fig. 8 | Comparison of mural edge enhancement effects. a Original mural, **b** mural heatmap, **c** heatmap after CLAHE, **d** heatmap after CLAHE and BF, and, **e** enhanced mural.
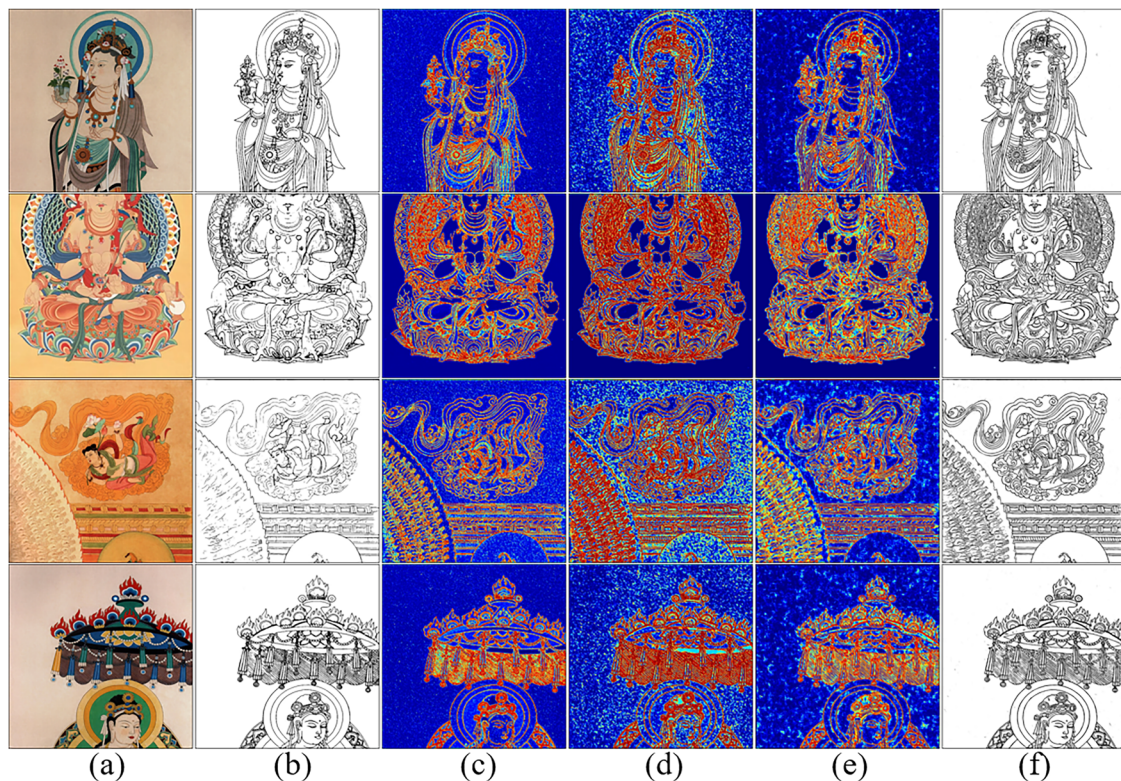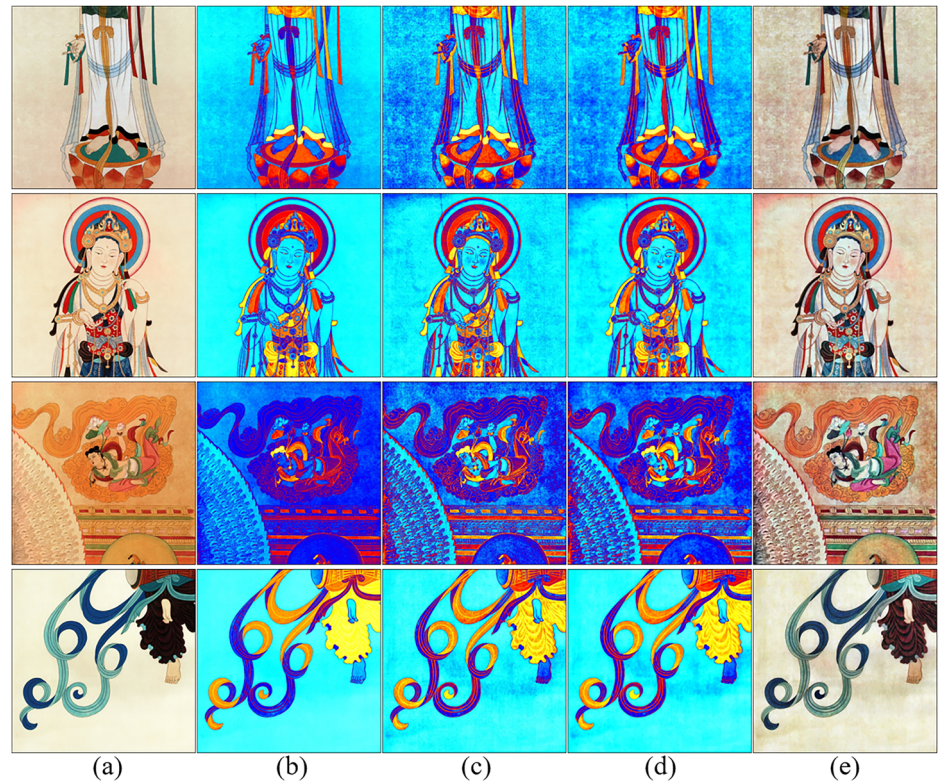

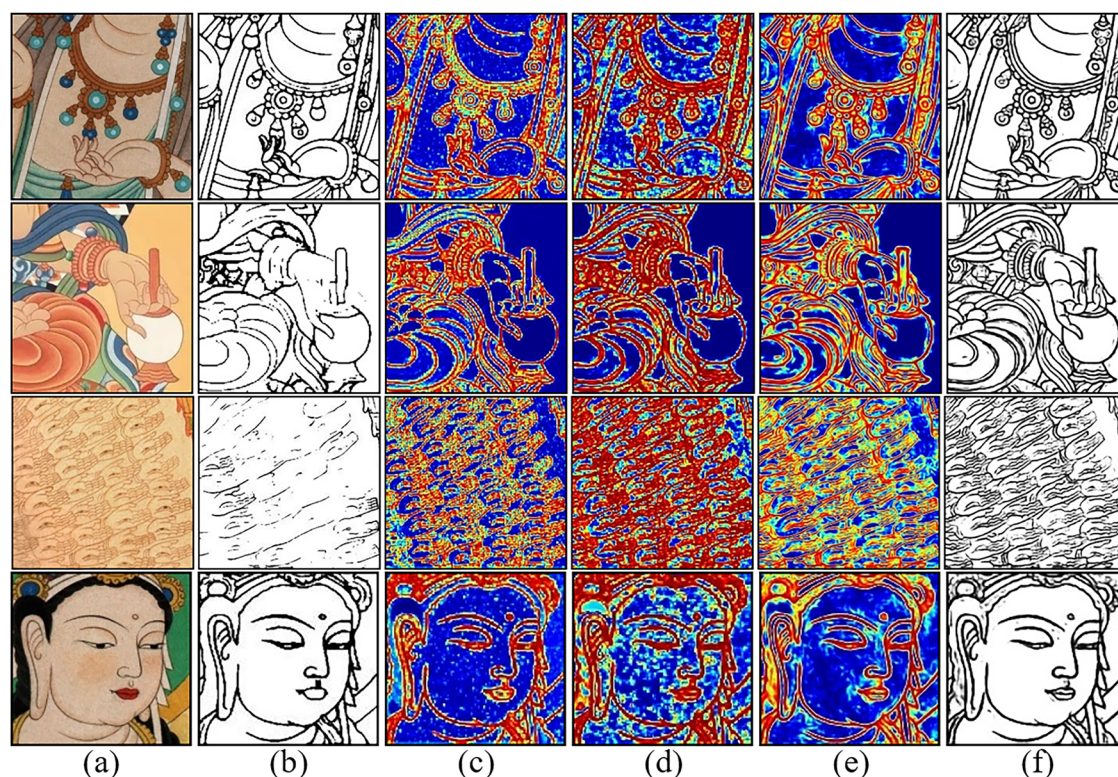**Fig. 9 | Intermediate feature layers and final fusion result of MLineNet. a** Mural, **b** reference line drawings, **c** H_1 heatmap, **d** H_2 heatmap, **e** H_3 heatmap, and **f** fusion result.

Line drawing extraction

To verify the effectiveness of MLineNet in the task of mural line drawing extraction, this study extracted line drawings based on the enhanced Dunhuang mural dataset. As shown in Fig. 9, MLineNet's performance in extracting mural line drawings is illustrated at different feature levels (H1–H3), as well as the final fused output (Y), while Fig. 10 presents the zoomed-in details. Through the encoder–decoder structure, the network first extracts features from low to high levels, and then utilizes multi-scale

Fig. 10 | Local comparison of Intermediate feature layers and final fusion result of MLineNet. a Mural, **b** reference line drawings, **c** H_1 heatmap, **d** H_2 heatmap, **e** H_3 heatmap, and **f** fusion result.

feature fusion to ensure that details at each scale are fully preserved. Each feature layer corresponds to different image features and details, and after fusion, it ensures a balance between the artistic expression and stability of mural line drawings. Specifically, Fig. 10c presents low-level features, which primarily capture rough outlines and shapes of the mural; Fig.10d reveals more details, particularly in the finer parts of the mural's lines; Fig. 10e further enhances the overall contour clarity while retaining more detailed information. By fusing these different scales, the final result in Fig. 10f achieves a well-balanced representation of line detail and connectivity. The experimental results show that MLineNet successfully realizes both fine expression and smooth connectivity of lines. Through multi-scale feature fusion, the model not only captures the delicate brushstrokes in murals but also preserves the overall artistic contours of the image. Notably, with the integration of the Fusion layer, the model effectively enhances the hierarchical structure and stability of lines, ensuring that each stroke is clearly presented within complex backgrounds. However, as observed from the results, although details are well preserved, some regions still contain noise.

Line drawing denoising

This study adopts CycleGAN to perform denoising tasks for mural line drawings, aiming to recover images with clearer lines and more complete structures from preliminary line drawings with noise. We use the noisy line drawings extracted in the previous steps (Fig. 11c) as source domain images and the reference line drawings (Fig. 11b) as target domain images, training the model in an unsupervised learning manner to learn the mapping from noisy to clean images, thereby achieving effective denoising and structural restoration. Figure 11 presents a visual comparison of several methods, including CycleGAN, as well as three image denoising methods: Non-Local Means (NL-Means)[53], Wiener filter[54], and Pix2pix[55]. All methods operate on the same preliminary line drawings, and the comparative results of different denoising methods and local approaches are shown in Figs. 11 and 12. Specifically, (b) shows the reference line drawings, which are handcrafted line drawings extracted from murals, demonstrating the edge structure of

the images. However, due to the incomplete accuracy of the annotations in the Dunhuang data, there are some missing or blurred details; (c) shows the preliminary noisy line drawings extracted, containing substantial noise, broken lines, and blurriness, which serve as the input source domain images for the model; (d) presents the NL-Means denoising results, which achieve denoising based on the similarity of image patches but lack in detail preservation; (e) shows the Wiener filter denoising results, where frequency-domain filtering smooths the noise but tends to blur the edges; (f) presents the Pix2pix denoising results, which, as a supervised method based on conditional Generative Adversarial Networks, depend on paired noisy and clean images for training, and while they can remove some noise, the generated results are very close to the reference line drawings and lack detail expression; (g) shows the CycleGAN denoising results, which, as an unsupervised learning method, learn the mapping between the two domains without relying on strictly paired samples, thereby better restoring the structural features and edge information of the images, resulting in clearer lines, smoother edges, and a significant improvement in overall visual quality, especially in maintaining structural similarity, edge clarity, and line connectivity.

Quantitative analysis of different stages

This study adopts multiple quantitative evaluation metrics to assess the quality of mural line drawings generated by different models for Dunhuang murals. These metrics include the SSIM, TC, LCI, and Q which integrates the three. To systematically analyze the impact of each stage on the quality of line drawing extraction, we conducted stepwise ablation experiments and individually evaluated the contributions of CLAHE + BF, MLineNet, and CycleGAN. These components were ultimately integrated into a complete method (Ours). Specifically, CLAHE + BF serves as a preprocessing step to enhance edge information; MLineNet is responsible for extracting preliminary line structures; and CycleGAN further denoises and improves the naturalness. Through these independent experiments, we analyzed the role of each stage and verified their synergy in the full pipeline. Table 1 presents the quantitative evaluation results for each stage in the Dunhuang mural line
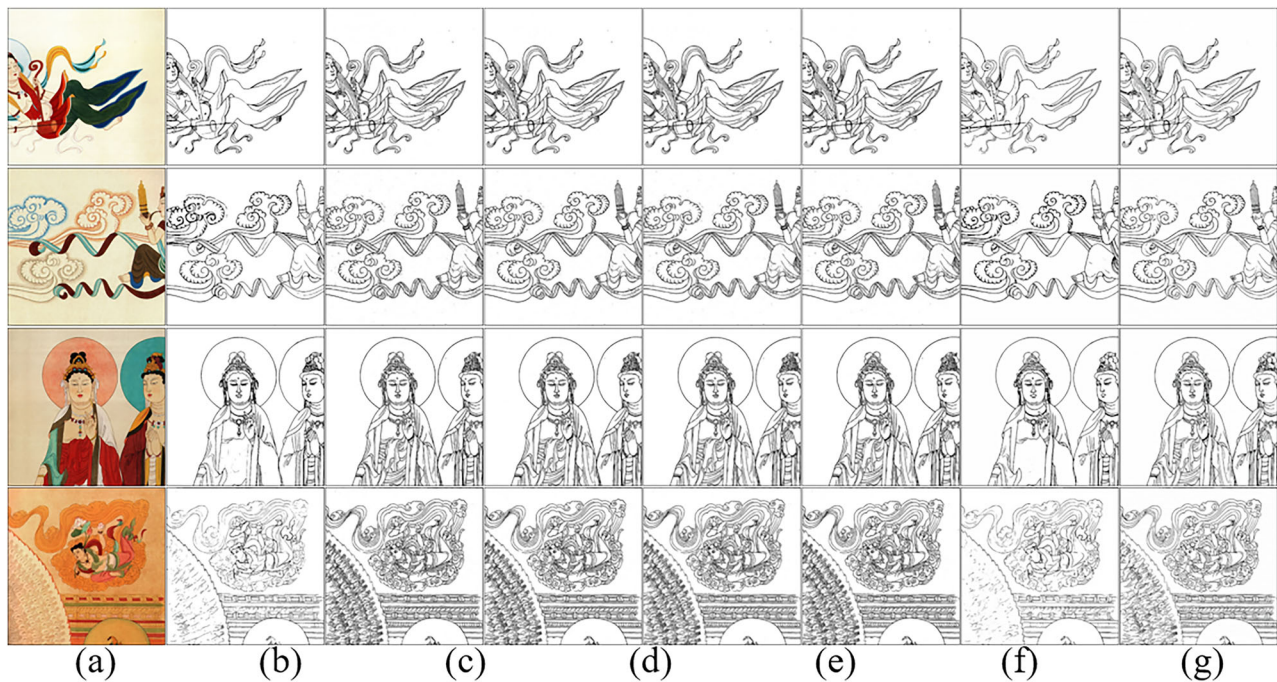
Fig. 11 | Comparison of line drawings denoising. a Mural, **b** reference line drawings, **c** noisy line drawings, **d** NL-Means, **e** Wiener, **f** Pix2pix, and **g** CycleGAN.
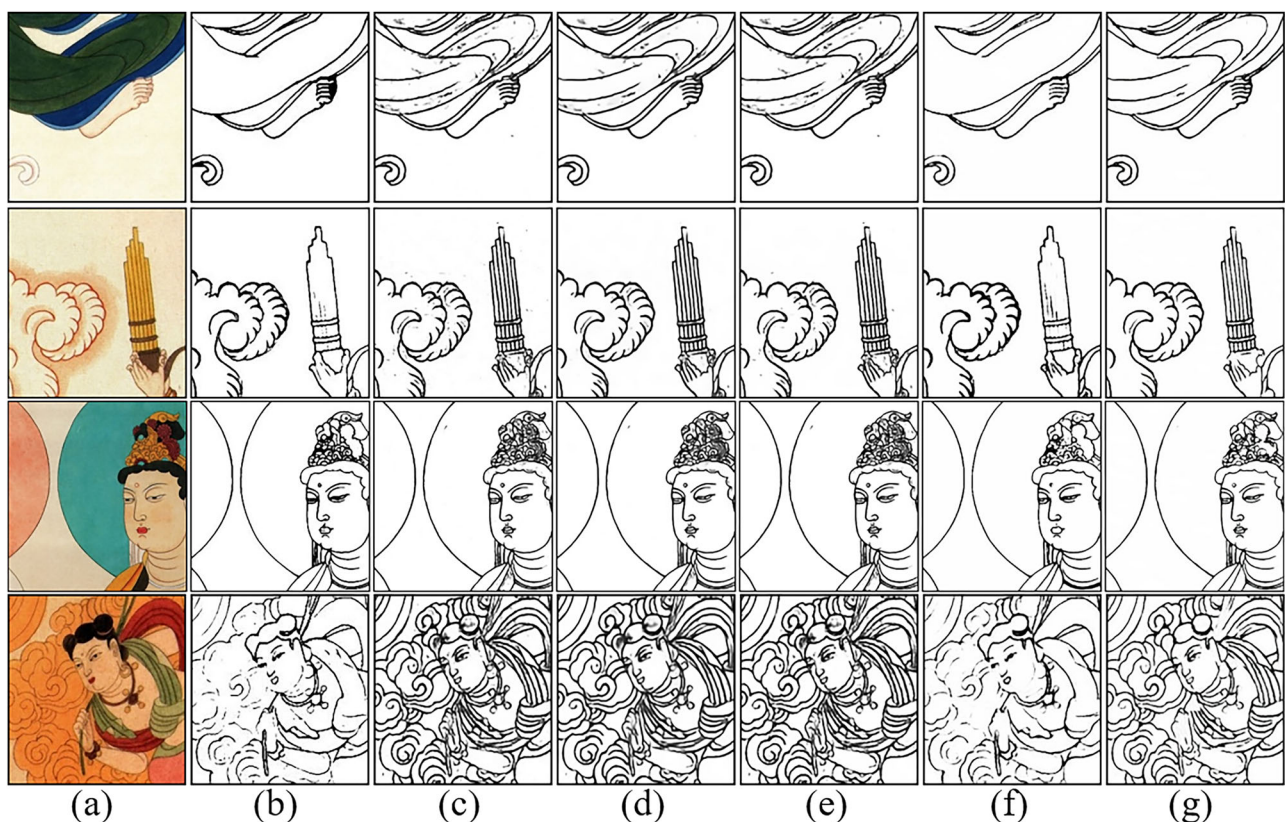

Fig. 12 | Local comparison of line drawings denoising. a Mural, **b** reference line drawings, **c** noisy line drawings, **d** NL-Means, **e** Wiener, **f** Pix2pix, and **g** CycleGAN.

drawing task, and Fig. 13 illustrates the radar chart comparison across all metrics.

As the base model, MLineNet demonstrated relatively stable performance, achieving an SSIM of 86.58%, a TC of 92.84%, an LCI of 85.37%, and a Q score of 88.26%, indicating its capability to extract complete and coherent mural line drawings. Building upon this, we introduced data augmentation (CLAHE + BF) for improvement, which resulted in an increase in SSIM to 88.19%, LCI to 86.97%, and Q to 89.19%, suggesting that enhanced input helps the model better preserve edge information and improve line connectivity. However, TC slightly decreased to 92.42%,

Fig. 13 | Radar chart comparison of various evaluation metrics for different stages in Dunhuang mural line drawing generation.
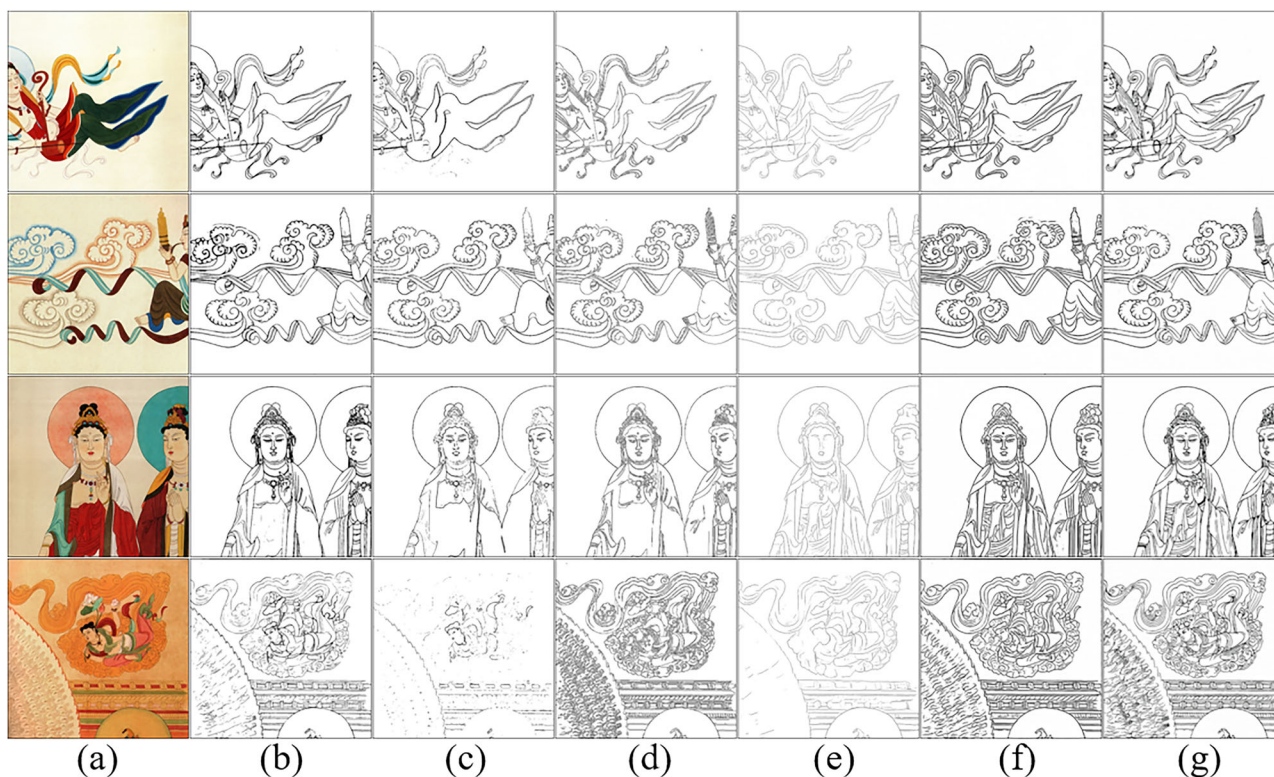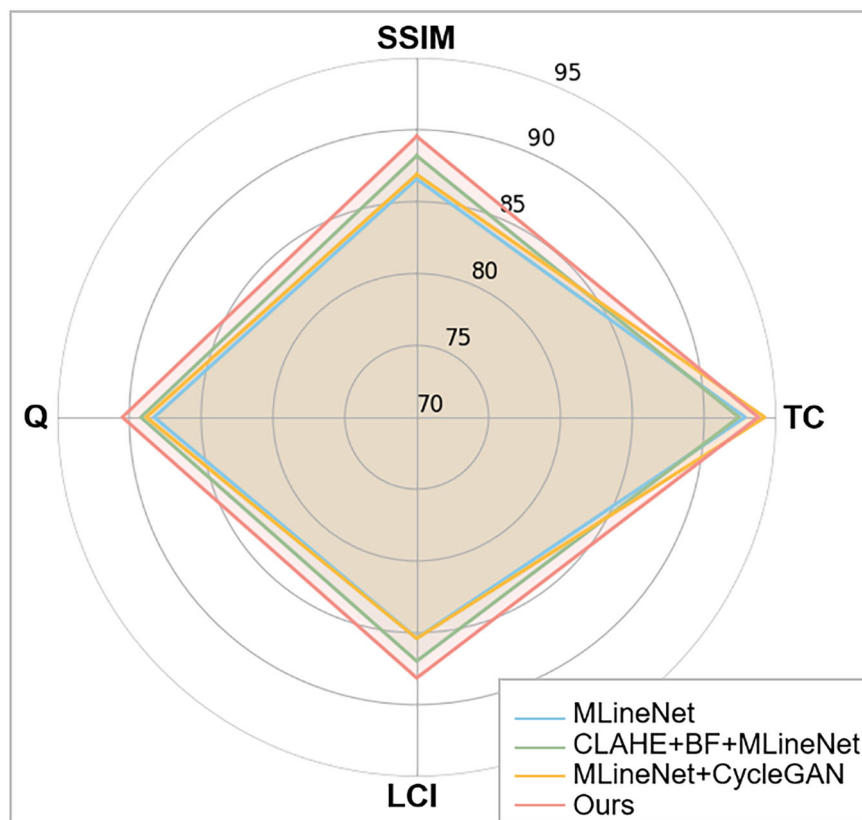

Fig. 14 | Comparison of different methods in Dunhuang mural line drawings generation. a Mural, **b** reference line drawings, **c** Sobel, **d** Canny, **e** DiffusionNet, **f** TEED, and **g** ours.

indicating that the augmentation process may introduce pseudo edges and increase texture complexity. After further incorporating CycleGAN, TC improved from 92.84% to 94.21%, confirming its effectiveness in reducing redundant textures. Finally, the complete method (Ours), combining data augmentation and denoising, achieved strong performance across all four metrics, with an SSIM of 89.54%, TC of 93.77%, LCI of 88.14%, and a Q score of 90.48%, demonstrating the overall effectiveness and synergy of the three-stage framework in generating complete, clean, and coherent mural line drawings.

Comparative analysis with other methods

To comprehensively evaluate the performance of different methods in generating Dunhuang mural line drawings, this study selects Sobel and Canny as classical edge detection operators, and DiffusionNet (diffusion probabilistic model for crisp edge detection)[10] and TEED (tiny and efficient model for edge detection generalization)[11] as deep learning-based comparison methods. Like traditional algorithms, Sobel and Canny can extract the main outlines of murals to some extent, while DiffusionNet and TEED, being deep learning models, offer stronger adaptability in terms of stability

and expressiveness in edge detection. The evaluation is conducted from three perspectives: qualitative analysis, quantitative analysis, and human perceptual assessment.

First, we perform a qualitative analysis of the line drawing results. Figure 14 shows the comparison of different methods on Dunhuang murals, and Fig. 15 presents the corresponding enlarged local regions. From the results, it can be observed that the Canny method performs relatively stably in extracting the main contours. However, due to its reliance on fixed gradient thresholds, it often results in the "double-edge problem," where both sides of a single line are detected as edges, producing redundant double-line structures. The Sobel operator is more sensitive to gradient changes and can capture strong edge signals, but its poor robustness to noise leads to fragmented lines, compromising the overall connectivity of the murals. In contrast, DiffusionNet applies a diffusion mechanism for edge detection, which helps simplify the lines to some extent, though it still suffers from detail loss in complex structures. TEED, as an efficient edge detection model, produces relatively clear overall contours and retains the main edge information well, but some disconnections still occur in fine line structures. The method proposed in this paper demonstrates a balanced performance in terms of structural similarity, simplicity, and connectivity, resulting in a well-rounded overall quality.

Secondly, the performance of different methods in generating Dunhuang mural line drawings was quantitatively compared, as summarized in Table 2. Figure 16 presents radar charts of various models across the evaluation metrics, providing a visual comparison of their performance differences. The results indicate that different methods show varying levels of effectiveness across the evaluated indicators. Among traditional edge detection methods, Sobel exhibits noticeable edge omissions in certain regions, leading to a relatively high TC of 94.56%, but the lowest LCI of only 82.19%. Canny achieves SSIM and LCI scores of 84.31% and 84.49%, respectively, yet suffers from the "double-edge problem," resulting in redundant lines and reduced overall clarity. In contrast, deep learning-based methods generally perform better across multiple indicators. DiffusionEdge, due

Table 1 | Quantitative evaluation results of different stages in Dunhuang mural line drawings generation

Stage			SSIM	TC	LCI	Q
CLAHE + BF	MLineNet	CycleGAN				
	√		86.58	92.84	85.37	88.26
√	√		88.19	92.42	86.97	89.19
	√	√	86.89	**94.21**	85.39	88.83
Ours			**89.54**	93.77	**88.14**	**90.48**

bold values indicate the maximum value for each metric.

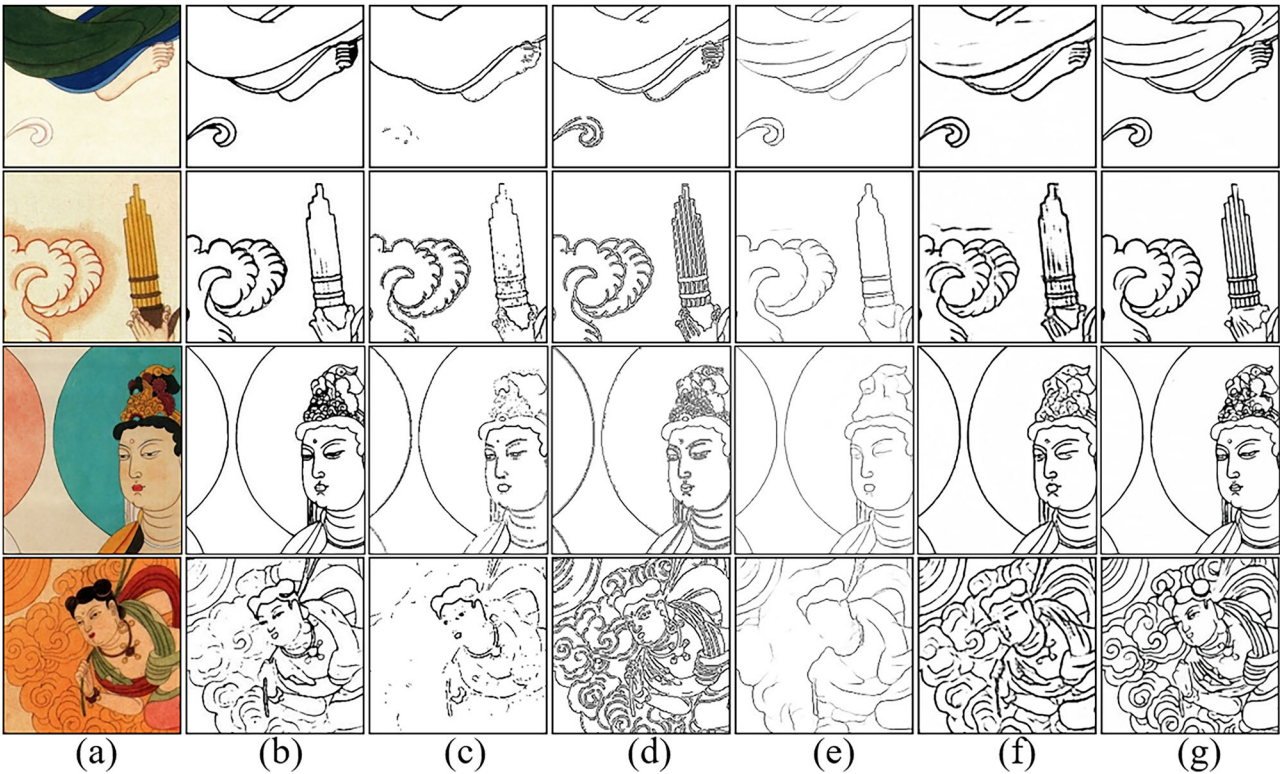**Fig. 15 | Local comparison of different methods in Dunhuang mural line drawings generation. a** Mural, **b** reference line drawings, **c** Sobel, **d** Canny, **e** DiffusionNet, **f** TEED, and **g** ours.

to its diffusion mechanism, excels in texture complexity with a TC of 96.55%, indicating cleaner lines with less noise. However, similar to Sobel, it has lower SSIM and LCI values of 82.79% and 83.02%, respectively, reflecting losses in detail and structural connectivity. The TEED method shows balanced performance, with an SSIM of 86.13%, TC of 93.37%, and LCI of 87.14%. The method proposed in this study achieves strong overall performance across all metrics, including an SSIM of 89.54%, TC of 93.77%, LCI of 88.14%, and a comprehensive score (Q) of 90.48%. These results indicate that the proposed method maintains a good balance between edge feature similarity, noise control and line connectivity.

Finally, to verify the effectiveness of the evaluation metrics (SSIM, TC, LCI) used in the Dunhuang mural line drawing task, a subjective experiment based on human perceptual judgment was conducted. Eight participants were invited to evaluate the line drawings generated by each method in three aspects: similarity, cleanliness, and connectivity, corresponding to SSIM, TC, LCI. A three-point scale was adopted (1 = poor, 2 = fair, 3 = good). The median (MED) and interquartile range (IQR) of the scores for each method were calculated to reflect the overall evaluation level and consistency of judgment. The subjective evaluation results are shown in Table 3. Additionally, the average of the three scores was computed as the overall perceptual score (OPS).

Table 3 presents the detailed subjective evaluation outcomes. Among traditional methods, Sobel received a similarity and connectivity score of 1 and a cleanliness score of 2, with an OPS of 1.17, the lowest among all methods, indicating limited structure recovery and poor line connectivity. Canny achieved a similarity and connectivity score of 2 and a cleanliness score of 1, yielding an OPS of 1.71, slightly outperforming Sobel. DiffusionEdge scored highest in cleanliness (MED = 3) but only 1 in similarity and connectivity, also resulting in an OPS of 1.71. This suggests effective noise removal but a lack of structural preservation and line connectivity. TEED scored 2 in similarity and connectivity and 1 in cleanliness, with an OPS of 1.92, reflecting moderate and stable subjective acceptance. The proposed method (Ours) achieved the highest OPS of 2.83, with a perfect score of 3 in similarity and connectivity and 2 in cleanliness. These results indicate that the method received consistent recognition from participants for preserving

Table 2 | Quantitative evaluation results of different methods in Dunhuang Mural line drawing generation

Method	SSIM	TC	LCI	Q
Sobel	81.14	94.56	82.19	85.96
Canny	84.31	89.43	84.49	86.08
DiffusionEdge	82.79	**96.55**	83.25	87.53
Teed	86.13	93.37	87.14	88.88
Ours	**89.54**	93.77	**88.14**	**90.48**

bold values indicate the maximum value for each metric.

Table 3 | Subjective evaluation results (MED, IQR, and OPS) for each method

Method	Similarity		Cleanliness		Connectivity		OPS
	MED	IQR	MED	IQR	MED	IQR	
Sobel	1	0	2	0.25	1	0	1.17
Canny	2	0	1	0.25	2	0	1.71
DiffusionEdge	1	0	3	0.25	1	0	1.71
Teed	3	0.25	1	0.25	2	0	2.13
Ours	3	0	2	0.25	3	0	**2.83**

bold values indicate the maximum value for OPS.

Fig. 16 | Radar chart comparison of various evaluation metrics for different models in Dunhuang mural line drawing generation.
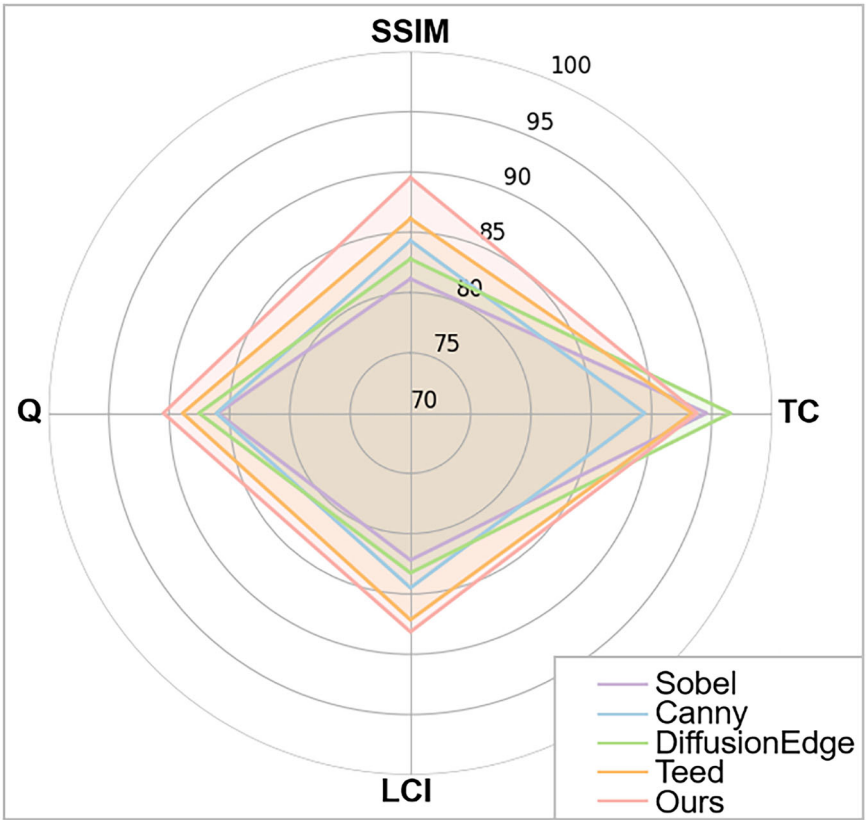
Fig. 17 | Comparative details of mural styles in Dunhuang and Baisha murals. a Facial features, **b** hand gestures, and **c** background decorations.
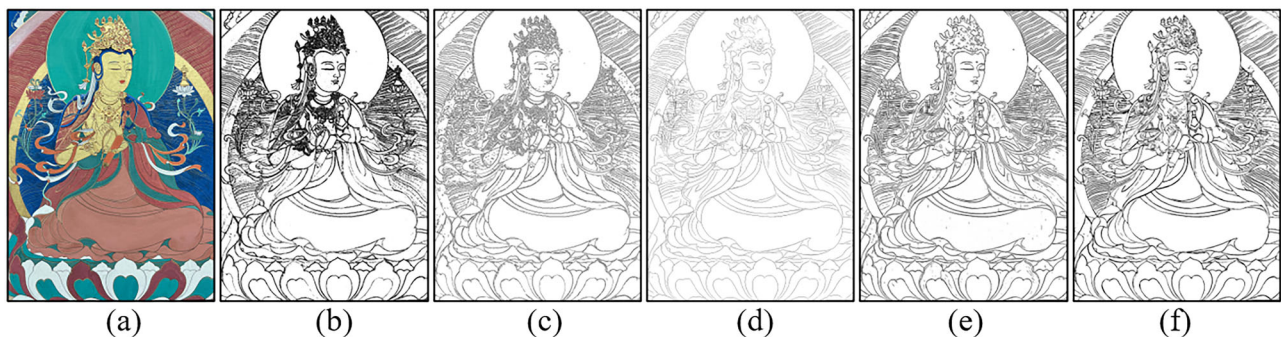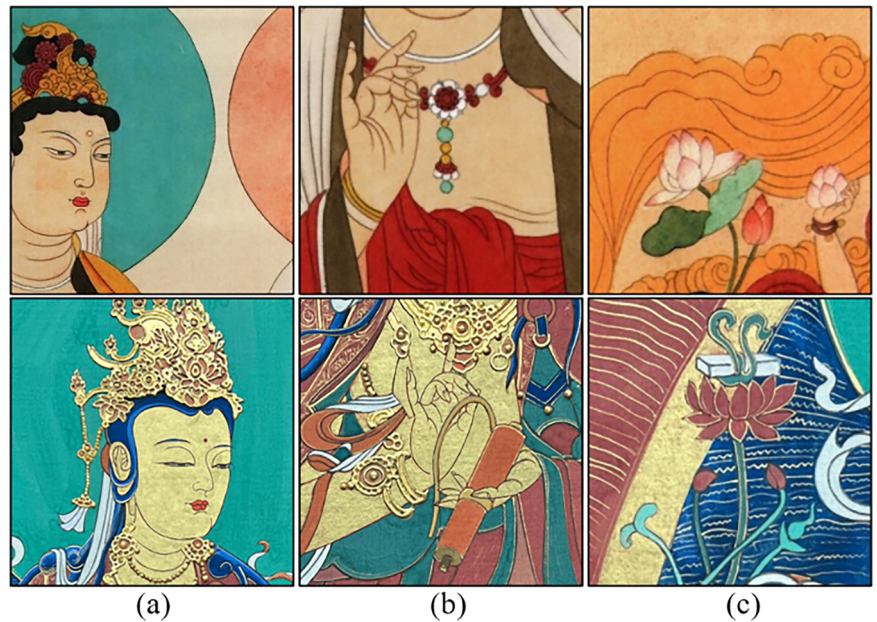
(a) (b) (c)

(a) (b) (c) (d) (e) (f)

Fig. 18 | Comparison of different methods in Baisha mural line drawings generation. a Mural, **b** Sobel, **c** Canny, **d** DiffusionNet, **e** TEED, and **f** ours.
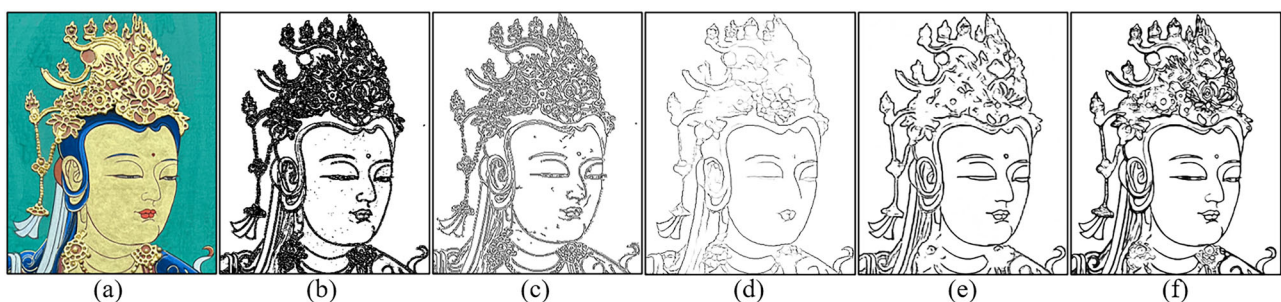

(a) (b) (c) (d) (e) (f)

Fig. 19 | Local comparison of different methods in Baisha mural line drawings generation. a Mural, **b** Sobel, **c** Canny, **d** DiffusionNet, **e** TEED, and **f** ours.

structural details, suppressing noise, and maintaining line connectivity. Furthermore, the subjective evaluation results align well with the objective metrics (SSIM, TC, LCI), suggesting that the adopted evaluation system is valid for assessing mural line drawing quality.

Line drawing generation and generalization verification for Baisha murals

The Baisha murals[48,49], located in Baisha, Yulong Naxi Autonomous County, Lijiang, Yunnan Province, are an important cultural heritage of the Naxi people, with the Dading Pavilion being one of the core temples. The Baisha

murals incorporate diverse cultural elements, including Tibetan Buddhism, Taoism, and Han Buddhism. Their content depicts Buddhist figures, transformation stories, and donor portraits, characterized by bold, flowing lines and rich, vibrant colors. As shown in Fig. 17, compared to the Dunhuang murals, the Baisha murals exhibit a more free and unrestrained overall style, with greater variation in brushstroke thickness, stronger color contrasts, and more intricate background decorations. These stylistic differences are reflected not only in visual features, such as more expressive facial depictions, simplified and dynamic hand gestures, and more complex decorative patterns. The model needs to handle variations in brushstroke thickness and

color layers, adapt to interference from intricate background decorations, and ensure the integrity of multi-scale structural information. Therefore, the Baisha murals serve as important historical materials for studying the religious art of southwestern China during the Ming and Qing dynasties, and in recent years, they have been increasingly digitized and studied in depth.

First, a qualitative analysis was conducted on the performance of different methods in generating line drawings for Baisha murals. Figure 18 shows a comparison of different methods, and Fig. 19 presents zoomed-in results. From the visual comparison, it can be observed that the Sobel and Canny methods exhibit significant limitations in line extraction. The Canny method, which is highly sensitive to local gradient variations, generates noise in both bright and dark regions. It also suffers from the "double-edge problem," leading to blurred and ambiguous line contours. Although the Sobel method captures the overall structure, it introduces considerable noise, resulting in cluttered line drawings that affect mural clarity and readability. Deep learning-based methods, such as DiffusionNet and TEED, improve line connectivity and structural integrity to some extent. DiffusionNet reduces background interference but lacks detail in fine line extraction. TEED performs better in maintaining global coherence and produces relatively stable lines. However, it still suffers from detail loss and

broken lines in complex decorative regions. In contrast, the proposed method demonstrates superior performance on the Baisha mural dataset by effectively preserving line structures while suppressing noise.
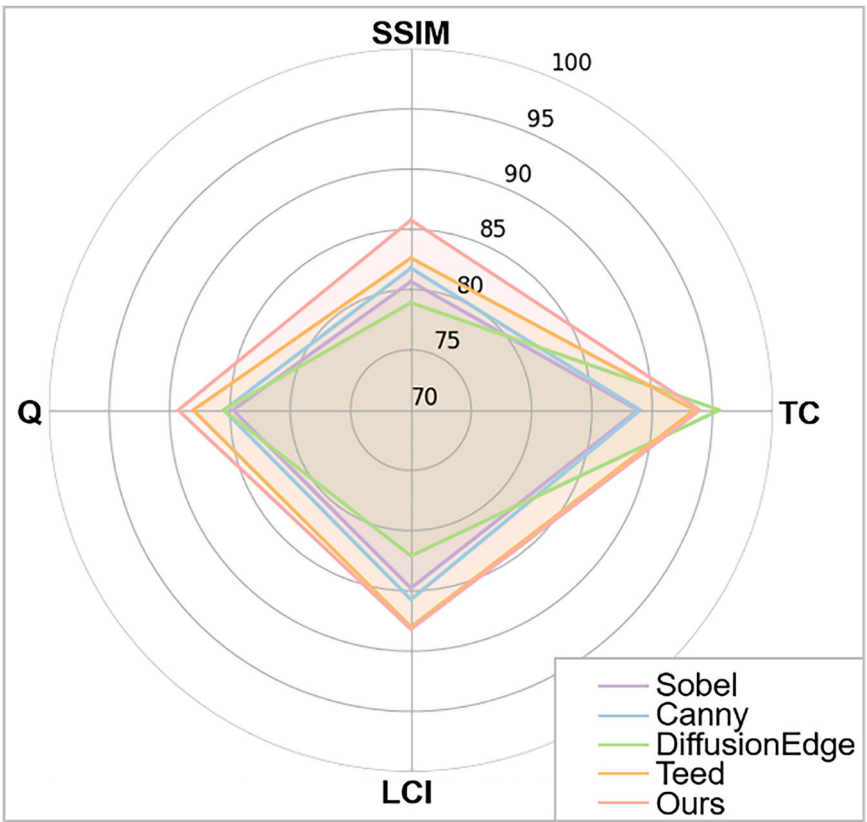
In addition, this study conducted a quantitative analysis of the performance of different methods in generating line drawings of Baisha murals. The experimental results are summarized in Table 4. Figure 20 presents a radar chart comparison of the models across various evaluation metrics, visually illustrating their differences. According to the results, among traditional methods, Sobel achieved an SSIM of 80.66%, TC of 88.81%, LCI of 84.77%, and a Q of 84.75%. Canny performed slightly better, with an SSIM of 81.78%, TC of 89.01%, LCI of 85.72%, and Q of 85.50%. Both methods showed relatively lower performance in TC and LCI, indicating limitations in texture simplification and line connectivity, though they retained some capability in structural restoration. DiffusionEdge achieved the highest TC at 95.58%, demonstrating strong suppression of redundant details. However, its SSIM and LCI were relatively lower among deep learning models— 78.92% and 82.10%, respectively—resulting in a Q score of 85.53%, indicating overall performance inferior to other deep models. TEED maintained relatively high scores across all metrics, with an SSIM of 82.62%, TC of 93.52%, LCI of 87.99%, and a Q of 88.04%, reflecting a well-balanced overall performance. The proposed method achieved consistently high scores in all four metrics: SSIM of 85.77%, TC of 93.93%, LCI of 88.18%, and Q of 89.29%. These results indicate a good balance between structural preservation, denoising, and line connectivity. Although the model is trained on Dunhuang mural data, it still demonstrates a certain degree of generalization ability when applied to Baisha murals. However, it is important to note that the current model is not specifically designed to accommodate the stylistic differences between different types of murals, such as variations in brushstroke thickness, color complexity, and decorative background styles. As a result, some decorative regions may exhibit issues such as local detail loss or incomplete structural representation. Overall, the proposed method achieved relatively stable results across all metrics, demonstrating strong overall performance in mural line drawing generation.

Table 4 | Quantitative evaluation results of different methods in Baisha Mural line drawing generation

Method	SSIM	TC	LCI	Q
Sobel	80.66	88.81	84.77	84.75
Canny	81.78	89.01	85.72	85.50
DiffusionEdge	78.92	**95.58**	82.10	85.53
Teed	82.62	93.52	87.99	88.04
Ours	**85.77**	93.93	**88.18**	**89.29**

bold values indicate the maximum value for each metric.

Fig. 20 | Radar chart comparison of various evaluation metrics for different models in Baisha mural line drawing generation.

Discussion

Line drawing is an essential form of artistic expression in Chinese painting, including traditional ink paintings, murals, and folk art. This study proposes an automated approach for generating mural line drawings by integrating mural edge enhancement, line drawing extraction, and line drawing denoising to achieve efficient and accurate line drawing generation. The main conclusions of this paper are as follows: (1) We propose an automated mural line drawing generation method that leverages CLAHE and BF for edge enhancement, followed by a specially designed neural network for mural line drawings, named MLineNet, for extracting mural line features. MLineNet adopts an encoder-decoder architecture with multi-scale feature fusion to effectively capture features across different scales. A CycleGAN is employed for denoising based on line drawing extraction results, generating high-quality line drawings that accurately capture rich structural details while maintaining clarity and naturalness. (2) Using Dunhuang murals as research subjects, a mural line drawing generation model was constructed based on the proposed method. The reliability of this model was evaluated using SSIM, TC, LCI, and Q. Experimental results showed that the proposed method achieved SSIM, TC, LCI, and Q of 89.54%, 93.77%, 88.14%, and 90.48%, respectively, demonstrating excellent performance in structural similarity, cleanliness, and connectivity. (3) To further validate the robustness and effectiveness of the proposed mural line drawing generation approach, we applied it to Baisha murals, where it achieved a Q value of 89.29%, confirming its generalization ability and adaptability to different mural datasets.

Although the experimental results demonstrate that the proposed method has a certain degree of adaptability and robustness, several limitations remain. First, the model training still relies on annotated line drawing data, which limits its scalability to large-scale, unlabeled datasets. Future work could explore the integration of semi-supervised or unsupervised learning strategies to reduce the dependence on annotated data. Second, for murals with diverse styles or complex structures, the model still encounters issues such as broken lines or missing details in certain regions. Future research could incorporate style-aware modules and adaptive feature modeling methods to enhance the model's ability to adapt to different mural styles. Additionally, the current validation experiments primarily focus on Dunhuang and Baisha murals, with a relatively limited range of data sources. Future research will aim to extend evaluations to mural datasets with greater geographical and historical diversity, such as the Yungang Grottoes and the Longmen Grottoes, to comprehensively assess the generalization ability and robustness of the method.

Data availability

Code availability

References

1. Ren, Y. & Liu, F. A model for inversion of hyperspectral characteristics of phosphate content in mural plaster based on fractional-order differential algorithm. *Sci. Rep.* **14**, 17898 (2024).
2. Lyu, Q., Zhao, N., Song, J., Yang, Y. & Gong, Y. Mural inpainting via two-stage generative adversarial network. *npj Herit. Sci.* **13**, 188 (2025).
3. Shen, J. et al. An algorithm based on lightweight semantic features for ancient mural element object detection. *npj Herit. Sci.* **13**, 70 (2025).
4. Guo, Q. The expression of lines in mural painting in Chinese meticulous figure painting. *Front. Art. Res.* **5**, 69–73 (2023).
5. Cheng, Y., Huang, M. & Sun, W. VR-based line drawing methods in Chinese painting. In *Proc. 2023 9th Int. Conf. Virtual Reality (ICVR)* (eds Huang, X.) pp. 604–610 (IEEE, Xianyang, CN, 2023).
6. Zhang, J. J., Chen, J. N., Meng, D. Y. & Wang, X. C. Exploring the uncertainty principle in neural networks through binary classification. *Sci. Rep.* **14**, 28402 (2024).
7. Jing, J., Liu, S., Wang, G., Zhang, W. & Sun, C. Recent advances on image edge detection: a comprehensive review. *Neurocomputing* **503**, 259–271 (2022).
8. Amer, G. & Abushaala, A. M. Edge detection methods. In *Proc. 2015 2nd World Symp. Web Appl. Netw.* (WSWAN) (ed. Tagoug, N.) pp. 1–7 (ITI, Piscataway, NJ, USA, 2015).
9. Çapkan, Y., Altun, H. & Fidan, C. B. Edge detection method driven by knowledge-based neighborhood rules. *Int. J. Eng. Technol. Innov.* **13**, 01–13 (2023).
10. Ye, Y., Xu, K., Huang, Y., Yi, R. & Cai, Z. DiffusionEdge: diffusion probabilistic model for crisp edge detection. *Proc. AAAI Conf. Artif. Intell.* **38**, 6675–6683 (2024).
11. Soria, X., Li, Y., Rouhani, M. & Sappa, A. D. Tiny and efficient model for the edge detection generalization. In *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)* (eds. Agapito, L., Berg, T., Kosecka, J., & Zelnik-Manor, L.) pp. 1356–1365 (IEEE, Paris, FR, 2023).
12. Yang, J., Price, B., Cohen, S., Lee, H. & Yang, M. H. Object contour detection with a fully convolutional encoder-decoder network. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)* (eds Agapito, L., Berg, T., Kosecka, J., Zelnik-Manor, L.) pp. 193–202 (IEEE Comput. Soc., Las Vegas, NV, USA, 2016).
13. Lisowska, A. Efficient edge detection method for focused images. *Appl. Sci.* **12**, 11668 (2022).
14. Pu, M., Huang, Y., Guan, Q. & Ling, H. B. RINDNet: Edge Detection for Discontinuity in Reflectance, Illumination, Normal and Depth. In *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)* (eds Agapito, L., Berg, T., Kosecka, J., Zelnik-Manor, L.) pp. 6879–6888 (IEEE Comput. Soc., Virtual Conf., 2021).
15. Leng, K. et al. SuperEdge: towards a generalization model for self-supervised edge detection. In *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)* (eds. Agapito, L., Berg, T., Kosecka, J., & Zelnik-Manor, L.) pp. 6675–6683 (IEEE, Paris, FR, 2024).
16. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986).
17. Kanopoulos, N., Vasanthavada, N. & Baker, R. L. Design of an image edge detection filter using the sobel operator. *IEEE J. Solid-state Circuits* **23**, 358–367 (1988).
18. Xu, C., Su, J., Wang, T., Zhang, Y. & Yu, A. Improved Canny edge detection operator. In *Proc. 2023 7th Int. Conf. Commun. Inf. Syst. (ICCIS)* (ed. Zhang, Y.) pp. 135–140 (IEEE Comput. Soc., Chongqing, CN, 2023).
19. Zhang, C. & Fang, J. D. Edge detection based on improved Sobel operator. In *Proc. 2016 Int. Conf. Comput. Eng. Inf. Syst. (CEIS)* (ed. Zhang, Y.) pp. 129–132 (Atlantis Press, Shanghai, CN, 2016).
20. Zangana, H. M., Mohammed, A. K. & Mustafa, F. M. Advancements in edge detection techniques for image enhancement: a comprehensive review. *Int. J. Artif. Intell. Robot.* **6**, 29–39 (2024).
21. Ji, S., Yuan, X. & Bao, J. Efficient stage features for edge detection. In *Proc. 2024 9th Int. Conf. Signal Image Process. (ICSIP)* (eds Li, B., Chen, L.) pp. 628–632 (IEEE, Nanjing, CN, 2024).
22. Jing, D., Li, B., Wang, S., Lin, J., & Jiao, Y. Edge detection in dark industrial environments. In *Proc. 2023 Int. Conf. Comput. Commun. (ICCC)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 1689–1693 (IEEE, Chengdu, CN, 2023).
23. Cui, S., Gui, Z., Hou, X., Peng, S., & Zhang, Q. Multi-scale generation and discrimination for sketch extraction of painted cultural relics. In *Proc. 2023 2nd Int. Conf. Image Process. Media Comput. (ICIPMC)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 95–100 (IEEE, Xi'an, CN, 2023).

24. Luo, W., & Yang, K. Target detection method for edge computing based on the improved YOLOv5 algorithm. In *Proc. 2024 4th Int. Conf. Electron. Circuits Inf. Eng. (ECIE)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 640–644 (IEEE, Hangzhou, CN, 2024).

25. Makka, A., Pateraki, M., Betsas, T., & Georgopoulos, A. 3D edge detection based on normal vectors. In *Proc. 2024 Int. Workshop 3D Virtual Reconstruct. Vis. Complex Archit. (3D-ARCH)* (eds. Campana, S., Fassi, F., Remondino, F.) pp. 295–300 (ISPRS, Siena, IT, 2024).

26. Hu, J., & Fang, W. Intelligent clean line drawing. In *Proc. 2016 Int. Conf. Robots Intell. Syst. (ICRIS)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 217–220 (IEEE, Zhangjiajie, CN, 2016).

27. Zhang, J., Wang, R.Z., & Xu, D. Automatic generation of sketch-like pencil drawing from image. In *Proc. 2017 IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 261–266 (IEEE, Hong Kong, CN, 2017).

28. Wang, T., Cao, C., Wang, C., Wang, Z. & Feng, Y. Coherence-enhancing line drawing for color images. *Sci. China Inf. Sci.* **56**, 1–15 (2013).

29. Liu, J., Lu, D., & Shi, X. Interactive sketch generation for Dunhuang frescoes. In *Proc. 2006 Int. Conf. Technol. E-Learn. Digit. Entertain. (Edutainment)* (eds. Pan, Z., Aylett, R., Diener, H., Jin, X., Göbel, S., Li, L.) pp. 943–946 (Springer, Hangzhou, CN, 2006).

30. He, J., Wang, S., Zhang, Y., & Zhang, J. A computational fresco sketch generation framework. In *Proc. 2013 IEEE Int. Conf. Multimedia Expo (ICME)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 1–6 (IEEE, San Jose, CA, USA, 2013).

31. Sun, D., Zhang, J., Pan, G., & Zhan, R. Mural2Sketch: A combined line drawing generation method for ancient mural painting. In *Proc. 2018 IEEE Int. Conf. Multimedia Expo (ICME)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 1–6 (IEEE, San Diego, CA, USA, 2018).

32. Huang, Z., Peng, Y., Xie, H., Fukusato, T., & Miyata, K. One-shot line extraction from color illustrations. In *Proc. 2021 Nicograph Int. (NicoInt)* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 19–22 (IEEE, Virtual Conf., 2021).

33. Lee, G. et al. Unpaired sketch-to-line translation via synthesis of sketches. In *Proc. SIGGRAPH Asia 2019 Tech. Briefs* (eds. Ma, J., Guan, S., Ishibashi, Y.) pp. 45–48 (ACM, Brisbane, AU, 2019).

34. Zhang, Q. et al. Deep learning for the extraction of sketches from spectral images of historical paintings. In *Proc. SPIE 11784, Optics for Arts, Architecture, and Archaeology VIII* (eds. Liang, H., & Groves, R.) pp. 11–20 (SPIE, Washington, DC, USA, 2021).

35. Xu, P. et al. Deep learning for free-hand sketch: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 285–312 (2023).

36. Deng, Y. & Liu, Y. Deep structural contour detection. *ACM Trans. Multimed. Comput., Commun., Appl. ((TOMM))* **16**, 1–19 (2020).

37. Wang, N., Wang, W. & Hu, W. Thangka mural line drawing based on cross dense residual architecture and hard pixel balancing. *IEEE Access* **9**, 48841–48850 (2021).

38. Wang, J., Liu, B., Li, J., Liu, W. & Du, S. Multi-stage feature fusion network for edge detection of Dunhuang murals. In *Proc. 2023 35th Chin. Control Decis. Conf. (CCDC)* (eds. Zhang, J. F., Wang, F., & Ma, H.) pp. 4684–4689 (IEEE, Yichang, CN, 2023).

39. Liu, B., He, F., Du, S., Zhang, K. & Wang, J. Dunhuang murals contour generation network based on convolution and self-attention fusion. *Appl. Intell.* **53**, 22073–22085 (2023).

40. Peng, J. et al. A relic sketch extraction framework based on detail-aware hierarchical deep network. *Signal Process.* **183**, 108008 (2021).

41. Yu, Z., Lyu, S., Hou, M., Sun, Y. & Li, L. A new method for extracting refined sketches of ancient murals. *Sensors* **24**, 2213 (2024).

42. Zhang, X. The Dunhuang caves: showcasing the artistic development and social interactions of Chinese Buddhism between the 4th and the 14th centuries. *J. Educ. Humanit. Soc. Sci.* **21**, 266–279 (2023).

43. Jin, J. & Sharudin, S. A. B. A Study on the Influence of the 8th Century A.D. Image of Feitian in the Mogao Caves on the character design of modern Chinese films. *Art. Soc.* **2**, 76–80 (2023).

44. Li, L. et al. Line drawing guided progressive inpainting of mural damage. arXiv Preprint:2211.06649 (2022).

45. Kilic, U. et al. Exploring the effect of image enhancement techniques with deep neural networks on direct urinary system x-ray (dusx) images for automated kidney stone detection. *Int. J. Intell. Syst.* **2023**, 3801485 (2023).

46. Liu, C. et al. The Medical image denoising method based on the CycleGAN and the complex shearlet transform. *Int. J. Adv. Comput. Sci. Appl.* **14**, 118−125 (2023).

47. Pu, M., Huang, Y., Liu, Y., Guan, Q. & Ling, H. Edter: Edge detection with transformer. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)* (eds. Agapito, L., Berg, T., Kosecka, J., & Zelnik-Manor, L.) pp. 1402–1412 (IEEE, New Orleans, LA, USA, 2022).

48. Liu, Z. Searching for a lost aura: a Naxi Dongba's spatial practices and space remaking in touristic commoditization. *J. Tour. Cult. Change* **16**, 348–364 (2018).

49. Yang, L., & Yu, Y. Restoration algorithm of Lijiang Baisha murals based on structure tensor and gradient priority calculation. In *Proc. 2023 2nd Int. Conf. Big Data, Inf. Comput. Netw. (BDICN)* (eds. Ma, J., Guan, S., & Ishibashi, Y.) pp. 305–311 (IEEE, Xishuangbanna, CN, 2023).

50. Silva, A. G. & Lotufo, R.d.A. Detection of lines using hierarchical region based representation. In *Proc. 17th Brazilian Symp. Comput. Graph. Image Process. (SIBGRAPI)* (eds. Silva, F. G. M., & Gomes, A. J. P.) pp. 58–64 (IEEE Comput. Soc., Curitiba, BR, 2004).

51. Haralick, R. M., Shanmugam, K. & Dinstein, I. H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **6**, 610–621 (2007).

52. Zhou Wang, U., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

53. Buades, A., Coll, B. & Morel, J.-M. A non-local algorithm for image denoising. In *Proc. 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)* (eds. Schmid, C., Soatto, S., & Tomasi, C.) pp. 60–65 (IEEE Comput. Soc., San Diego, CA, USA, 2005).

54. Metri, S. & Asha, T. Patch based wiener filter for image denoising. In *Proc. 2018 Int. Conf. Comput. Tech., Electron. Mech. Syst. (CTEMS)* (eds. Deshpande, S., & Kenchannavar, H.) pp. 392–396 (IEEE, Belagavi, IN, 2018).

55. Henry, J., Natalie, T. & Madsen, D. Pix2Pix GAN for image-to-image translation. In *Proc. 2021 Int. Conf. Comput. Vis. Appl. (ICVA)* (eds. Smith, A., & Johnson, B.) pp. 1–5 (ResearchGate, Virtual Conf., 2021).

Acknowledgements

Author contributions

H.X.F. and Q.W.H. designed the study and wrote the manuscript. P.C.Z., D.Y.Z., and M.Y.A. conducted data analysis and contributed to the writing and editing of the manuscript. S.L.C. and X.Y.H. provided resources and participated in the manuscript review. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Qingwu Hu.

Reprints and permissions information is available at
http://www.nature.com/reprints