applied sciences

MDPI

Article

Research on Digital Restoration and Innovative Utilization of Taohuawu Woodblock New Year Prints Based on Edge Detection and Color Clustering

Yingluo Dai [1], Fei Ju [1,*,†] and Yuhang Wen [2]

[1] College of Art Design, Nanjing Forestry University, Nanjing 210037, China; 2251310603@njfu.edu.cn
[2] College of Information Science and Technology & Artificial Intelligence, Nanjing Forestry University, Nanjing 210037, China
* Correspondence: jufei1986@njfu.edu.cn
† Fei Ju is recognized as co–first author of this work.

Abstract

Taohuawu woodblock New Year prints are one of the most representative traditional multicolor woodblock print forms from the Jiangnan region of China and are recognized as an Intangible Cultural Heritage at the provincial level in Jiangsu. However, the development of mechanized and high-tech production methods, combined with the declining role of traditional festive customs in modern society, has posed significant challenges to the preservation and transmission of this art form. Existing digital preservation efforts mainly focus on two-dimensional scanning and archival storage, largely neglecting the essential processes of color separation and multicolor overprinting. In this study, a digital restoration method is proposed that integrates image processing, color clustering, and edge detection techniques for the efficient reconstruction of the traditional multicolor woodblock overprinting process. The approach applies the K-means++ clustering algorithm to extract the dominant colors and reconstruct individual color layers, in combination with CIELAB color space transformation to enhance color difference perception and improve segmentation accuracy. To address the uncertainty in determining the number of color layers, the elbow method, silhouette coefficient, and Calinski-Harabasz index are employed as clustering evaluation methods to identify the optimal number of clusters. The proposed approach enables the generation of complete, standardized digital color separations, providing a practical pathway for efficient reproduction and intelligent application of TWNY Prints, contributing to the digital preservation and innovative revitalization of intangible cultural heritage.

Keywords: Taohuawu woodblock New Year prints; digital restoration; K-means++ clustering; Canny edge detection; cyclic U-Net

check for updates

Academic Editors: Marcello La Guardia, Valentina Bonora and Andrea Masiero

Received: 15 July 2025
Revised: 11 August 2025
Accepted: 15 August 2025
Published: 18 August 2025

Citation: Dai, Y.; Ju, F.; Wen, Y. Research on Digital Restoration and Innovative Utilization of Taohuawu Woodblock New Year Prints Based on Edge Detection and Color Clustering. *Appl. Sci.* **2025**, *15*, 9081. https://doi.org/10.3390/app15169081

1. Introduction

1.1. Introduction to Taohuawu Woodblock New Year Prints

1.1.1. History and Development

Taohuawu woodblock New Year prints (TWNY Prints) are an essential component of China's traditional woodblock art. Their origins can be traced back to the Song and Ming Dynasties, with the art form reaching its peak during the Qing Dynasty. Together with the Yangliuqing prints from Tianjin, the Yangjiabu prints from Shandong, and the Mianzhu prints from Sichuan, the TWNY Prints are collectively known as one of the "Four

Great Schools of Chinese New Year Prints" [1]. This unique art form not only embodies a profound historical and cultural heritage but also reflects the aesthetic preferences and spiritual aspirations of people across different historical periods.

The history of Chinese woodblock printing is long and distinguished, with its origins dating back to the Tang Dynasty (618–907 AD). The earliest known surviving example of woodblock printing is a Buddhist document known as the Jietie (ordination certificate), dated to the 29th year of the Kaiyuan era (741 AD), which is currently preserved in Saint Petersburg. This artifact features printed images of three Buddhas and serves as tangible evidence of China's early mastery of woodblock printing techniques [2]. This indicates that woodblock printing technology was applied to religious and cultural dissemination at a very early stage, laying the technical foundation for the later development of New Year prints. During the Song and Yuan Dynasties, woodblock printing experienced significant growth, while the Ming and Qing Dynasties witnessed the diversification and flourishing of this artistic tradition.

The Qing Dynasty marked the peak of TWNY Prints. During this period, TWNY Prints were characterized by full and balanced compositions, smooth and elegant lines, and bright, vivid colors, earning widespread popularity among the general public. Beyond serving as festive decorations, these prints carried rich cultural symbolism, often expressing wishes for good fortune and happiness. For example, the well-known work Yituan Heqi (Harmony and Unity) conveys people's aspirations for social harmony and personal well-being.

The production process for the TWNY Prints is highly sophisticated, encompassing three major stages: painting (designing the image), carving (engraving the woodblocks), and printing (layered color application). Through generations of practice, experienced artisans have developed a unique and refined set of craftsmanship techniques. The artistic style of TWNY Prints integrates elements of folk traditions and literati painting, with particularly distinctive approaches to color application. The imagery within these prints often embodies profound cultural symbolism. For instance, patterns such as the swastika (wan symbol), the Chinese character for "blessing" (fu福), and the character for "longevity" (shou寿) appearing on double-handled vases symbolize wishes for good fortune and long life. Similarly, motifs like plum blossoms, magpies, and bamboo on teapots convey the auspicious message of "good news arrives with the plum and bamboo" (Zhumei Bao Xi) [3].

1.1.2. Production Process

Characterized by multicolor designs, TWNY Prints employ the "one block, one color" overprinting technique, which involves carving multiple woodblocks, each corresponding to a specific color layer. These color layers are printed successively, one on top of another, to create the final image.

During this process, artisans must rely on their experience to precisely align the different color plates, ensuring the accuracy and consistency of the overprinting. This highly intricate and labor-intensive procedure demands both technical skill and artistic sensitivity.

Characterized by multicolor designs, the traditional production process of TWNY Prints consists of three major stages: painting (designing the image), carving (engraving individual woodblocks for each color layer), and printing (layer-by-layer application of colors) [4]. Every step of the way, the old craftsmen's unique experience and exquisite skills are united.

The production of TWNY Prints begins with the creation of the initial drawing, known as the design draft. The subject matter of these drafts is broad and diverse, often depicting folk traditions, auspicious symbols, and scenes from traditional operas, vividly reflecting people's aspirations for a better life. These drafts serve not only as the starting point for

artistic creation but also as carriers of the rich folk wisdom and artistic essence of the Jiangnan region.

The drawing process represents the first essential step in the production workflow, requiring artists to possess considerable experience and technical sensitivity. If the linework is unsuitable for carving or if the color arrangements are impractical for printing, the design draft will be returned or revised by the workshop before proceeding. The quality of the carving and printing stages ultimately determines the overall quality of the final TWNY prints.

TWNY Prints adopt a relief woodblock printing method characterized by overlapping brush-applied color layers. During the carving process, appropriate wood materials are selected, with pear wood being the preferred choice, and the design draft is carefully adhered to the wooden surface using glue. Specialized tools are then used to engrave the image onto the block with precision.

The multicolor overprinting process follows a strict sequence. First, the black ink outlines (key lines) are printed, establishing the framework of the image. Subsequently, each color layer is applied one by one, with each color corresponding to a single, dedicated woodblock.

1.1.3. Traditional Color Options

The use of color in TWNY Prints is renowned for its vibrant, eye-catching style and strong contrasts, fully reflecting the distinctive aesthetics of Jiangnan folk art. The color palette primarily features highly saturated hues, including red, peach pink, green, and bright yellow. Traditionally, the dominant colors consist of black, white, red, green, and yellow. The edges of color blocks are sharp, well-defined, and visually separated, with color combinations following the principle of "clear hierarchy and strong contrast." Black ink outlines are used to delineate contours, creating a visual structure that emphasizes color layers without relying heavily on internal linework. This is a characteristic known as the "color without lines" effect. Within this traditional color paradigm, red symbolizes good fortune and festivity, green represents vitality, and yellow conveys wealth and splendor, together forming a highly recognizable and culturally symbolic color system.

In addition to global initiatives such as UNESCO and Europeana, recent technical studies have further shaped the methodology of digitally recording traditional craftsmanship. Zabulis (2022) proposed a structured framework for identifying and digitally representing the core data, processes, and tacit knowledge embedded in craft practices, resonating with this study's objective of reconstructing the color layers from TWNY Prints [5]. Similarly, restoration of intricate craft textures (e.g., Miao embroidery) has been advanced using GAN-enhanced U-Net architectures with spatial channel attention, achieving superior reconstruction performance in metrics such as PSNR and SSIM [6]. Broader AI-driven restoration efforts for cultural heritage artifacts, including image enhancement, denoising, inpainting, and colorization tasks, have been systematically reviewed, demonstrating the versatility of intelligent image processing in heritage conservation [7,8]. Furthermore, a recent bibliometric analysis revealed a rapid growth of deep learning applications in cultural heritage image recognition and restoration between 1995 and 2024 [9]. Complementing these 2D approaches, advanced 3D technologies have substantially contributed to the preservation of intangible cultural heritage through faithful digital documentation and presentation [8,10].

1.2. Existing Issues

As an element of intangible cultural heritage, TWNY Prints face significant challenges in heritage preservation and transmission. Firstly, the inefficiency of manual production

is a prominent problem, as a piece of work needs to go through more than five overlay printing processes, and the craftsmen rely on their manual alignment experience to easily produce millimeter-level errors, resulting in a high rate of misalignment of colors. Secondly, as shown in the research conducted by scholars Yuan Xia Liu [11] and Yongling Huang [12], the current methods of digital preservation are still rather rudimentary, and related research in this area is still at the initial stage of high-definition scanning and recording. Most of these efforts are limited to two-dimensional image preservation and neglect the most critical aspects of the art form, namely the craftsmanship of color plate production and the layered overprinting techniques.

Thirdly, the cultural context of traditional New Year prints has undergone significant changes. With the acceleration of urbanization, these prints have gradually become detached from their original role within the folk traditions of the Chinese New Year. As a result, younger generations lack awareness and understanding of this cultural heritage, leading to a shrinking market and further endangering its survival.

Beyond technical innovation in image processing, preserving traditional craftsmanship through digital technologies has become a central theme in global cultural heritage discourse. UNESCO, through its 2003 Convention for the Safeguarding of the Intangible Cultural Heritage, categorizes traditional craftsmanship as one of five domains of intangible heritage, emphasizing the transmission of artisan skills and knowledge, which are considered fundamental to cultural continuity, rather than merely focusing on material objects [13]. UNESCO further advocates for community-led documentation efforts and the use of media production to record and revitalize these living traditions.

At the European level, Europeana operates as a comprehensive digital aggregator, bringing together over 50 million cultural objects, including craft-related artifacts and narratives, from more than 3000 institutions [14]. Within this ecosystem, the CRAFTED project has actively promoted the transfer of European crafts by enriching and sharing both tangible heritage and the intangible skills behind them. It features exhibitions, blogs, galleries, and video stories showcasing artisans and their techniques [15].

Recent research further underscores the value of digital approaches in systematically capturing traditional craft knowledge. A notable example is Zabulis (2022), who introduced a methodical framework for identifying and digitally representing core data, information, and procedural knowledge inherent in traditional craft processes [5]. This approach aligns closely with our goal of reconstructing color layers in TWNY Prints, emphasizing the necessity of structured digital documentation to preserve the tacit and procedural dimensions of artisanal heritage.

Collectively, these initiatives reflect a growing global commitment to digitally document, preserve, and circulate the intangible dimensions of traditional craftsmanship. They provide an enriching comparative framework and underscore the cultural and methodological relevance of our study on digitally reconstructing color layers within TWNY Prints.

2. Materials and Methods

2.1. Image Processing Related Methods

Color clustering algorithms in image processing, especially those based on K-means and its variants, have experienced a long period of development and evolution, and play a key role in the fields of image segmentation, color quantization, and target recognition.

K-means as a distance-based clustering algorithm has been rapidly introduced for color clustering due to its simplicity and efficiency. Songul Albayrak et al. improved the K-means algorithm for color quantization. The method determines the center of each color cluster by calculating a weighted average using histogram values and employs an average

distortion optimization strategy to improve the perceptual quality of the quantized image. This study also conducted experiments in two color spaces, RGB and CIELAB, to investigate the effect of color space on clustering effects [16]. A real-time color image segmentation method based on K-means clustering was proposed and implemented by Takashi Saegusa and Tsutomu Maruyama [17]. They recognized the computationally intensive and time-consuming problem of K-means in processing large images and a large number of clusters and worked to improve its performance by optimizing the distance computation for real-time processing on FPGAs. Y-C Hu and B-H Su also addressed the computational cost of the K-means algorithm for palette design by proposing two test conditions to accelerate the K-means algorithm [18]. The experimental results show that the method significantly reduces the computational effort. Md. Rakib Hassan and Romana Rahman Ema explored image segmentation using automated K-means clustering in RGB and HSV color spaces [19]. They pointed out that despite the plethora of image segmentation algorithms, evaluating their accuracy remains challenging. Sangeeta Yadav and Mantosh Biswas proposed an improved color-based K-means algorithm for clustering of satellite images [20]. The method is carried out in two phases, where initial clustering centers are first selected and computed through an interactive selection process, and then clustering is performed on this basis, aiming to improve the recognition accuracy of image clustering.

One of the main drawbacks of the K-means algorithm is that it is highly sensitive to the location of the initial clustering centers, which may cause the algorithm to converge to a local optimum solution [21]. In order to solve the problem of uncertainty in the number of clusters, Abd Rasid Mamat et al. investigated the application of K-means algorithm in three color models (RGB, HSV, and CIELAB) for determining the optimal number of clusters for an image with different color models, and evaluated the clustering effect by using the Silhouette index [22]. Ting Tu et al. proposed a K-means clustering algorithm based on multi-color space, which solves the problem that the parameters and initial centers of K-means need to be input manually in color image segmentation. Their study found that HSV and CIELAB color spaces perform better in color segmentation [23].

In recent years, with the development of deep learning, it has also been attempted to combine the traditional K-means clustering with deep learning. Sadia Basar et al. proposed a new adaptive initialization method for initializing the K-means clustering of RGB histograms to determine the number of clusters and find the initial centroids, which further optimizes the application of K-means in unsupervised color image segmentation [24]. Curtis Klein et al. proposed a method for automated UAV labeling of training data for online differentiation between water and land. The method consists of converting images to HSV color space, followed by image segmentation using K-means clustering, and a combination of morphological operations and contour analysis to select key points [25]. Maxamillian A. N. Moss et al. conducted a comparative study of clustering techniques such as K-means, hierarchical clustering algorithm (HCA), and GenieClust, and explored the integration of an auto-encoder (AE) [26]. It was found that K-means and HCA show good agreement in terms of cluster profile and size, proving their effectiveness in distinguishing particle types.

Edge detection remains a cornerstone technique in cultural heritage image processing. Feng et al. developed an automated generation method for mural line drawings by integrating edge enhancement, neural edge detection, and denoising techniques [27]. Similarly, Li et al. demonstrated the efficacy of edge detection combined with semantically driven segmentation for extracting architectural features in historical urban contexts [28]. The core objective is to suppress noise interference while maximizing edge detail and ensuring precise positioning of edges.

The Roberts operator is one of the simplest and fastest gradient-based edge detection operators. It uses a pair of 2×2 convolution kernels to compute a gradient approximation of the image intensities, and is mainly used for detecting diagonally oriented edges [29]. The advantage is that it is computationally inexpensive, but it may not work well with noisy images and the detected edges are usually thin [30]. The Prewitt algorithm, similar to the Sobel algorithm, is a gradient-based method that uses a pair of 3×3 convolutional kernels to compute the horizontal and vertical gradients. The Prewitt algorithm strikes a balance between edge detection performance and computational complexity [31]. It is effective at highlighting boundaries in an image, but its edge detection results can be coarser than more accurate algorithms such as the Canny algorithm. The Sobel algorithm is one of the most commonly used edge detection operators, and it detects edges by calculating a gradient approximation to the image intensity function. Similar to the Prewitt operator, the Sobel operator uses a 3×3 convolutional kernel, but it assigns a greater weight to the center pixel, making it relatively more effective at suppressing noise. However, the traditional Sobel algorithm has low edge localization accuracy and a limited ability to handle noise and edge continuity. To address these problems, some studies have proposed image fusion algorithms based on improved Sobel algorithms, combining Canny and LoG algorithms to optimize the edge detection results. Yuan et al. proposed a high-precision edge detection algorithm based on improved Sobel algorithm-assisted HED (ISAHED) [32], The detection performance is improved by increasing the gradient direction. Zhou et al. also proposed an improved Sobel operator edge detection method based on FPGAs, which utilizes the parallel processing capability of FPGAs to improve efficiency [33]. The LoG operator is a second-order derivative operator that first smooths the image using a Gaussian filter to reduce noise and then applies the Laplace operator to find zero crossings, which correspond to the edges of the image. The LoG operator is sensitive to rapid changes in the image intensity, and is able to detect fine edge details, but is very sensitive to noise. Therefore, Gaussian smoothing is usually required first.

The Canny edge detection algorithm plays a pivotal role in the field of image processing. The success of the Canny algorithm lies in the fact that it takes into account the three major criteria of detection (identifying as many real edges as possible), localization (determining the edge position precisely), and suppressing false response (reducing false edges due to noise). The Canny algorithm has become one of the most popular edge detection tools in image processing due to its excellent performance. It achieves this by employing a multi-stage process that includes noise reduction, gradient finding, non-maximum suppression, and hysteresis thresholding to reliably detect a wide range of edges. However, with the complexity of image application scenarios and higher requirements for real-time and robustness, researchers have continuously improved and extended the Canny algorithm. CaiXia Zhang et al. proposed an improved Canny algorithm, which combines adaptive median filtering to enhance the noise reduction ability of the image, and utilizes local adaptive thresholding for edge detection to solve the problem in the traditional Canny algorithm of the salt-and-pepper noise and the poor resistance and poor adaptability of threshold [34]. In their study, Yibo Li and Bailu Liu also proposed an improved algorithm for the effect of salt-and-pepper noise on the Canny algorithm, which designs a novel filter to replace the Gaussian filter in the traditional algorithm, aiming at removing the salt-and-pepper noise and extracting the edge information of the region of interest [35]. Shigang Wang et al. proposed an algorithm that fuses the improved Canny operator and morphological edge detection to efficiently deal with noise in images through hybrid filters [36]. Ruiyuan Liu and Jian Mao suggested using a statistical algorithm for denoising and combined it with genetic algorithm to determine the optimal high and low thresholds to solve

the problem of poor noise robustness and possible false edges or edge loss in the traditional Canny algorithm [37].

In addition to the improvement in noise removal, researchers have also optimized Canny itself for adaptive thresholding. Jun Kong et al. proposed an adaptive edge detection model based on an improved Canny algorithm, replacing Gaussian smoothing in the standard Canny algorithm with a morphological approach to highlight edge information and reduce noise [38]. In addition, they utilized fractional order differential theory to compute the gradient values. Ziqi Xu et al. improved the Canny operator by Otsu's algorithm and the double threshold detection approach [39], enhancing its capability in medical image edge detection. Baoju Zhang et al. proposed an improved Canny algorithm for the drawbacks of the traditional Canny operator that require manual intervention in Gaussian filter variance [40]. The algorithm performs a hybrid enhancement operation after Gaussian filtering of multispectral images, aiming to avoid losing edge details while denoising.

2.2. Key Technologies

2.2.1. Swatch Clustering Based on K-Means++

The RGB color space commonly used in computers is not "perceptually homogeneous", and the same mathematical distance is not equivalent to the same color difference in human vision. To overcome this shortcoming, the CIELAB color space is introduced in this study. Its components l, a, and b represent luminance, green–red, and blue–yellow, respectively. The biggest advantage of this space is that the Euclidean distance between two color points can be highly approximated by the "color difference" perceived by the human visual system, which provides a mathematical basis for machine learning algorithms to make judgments more in line with artistic intuition.

Determining the appropriate number of clusters, K, is crucial before performing a clustering operation. In this study, an introduction and a comparison of the following three commonly used evaluation methods are provided.

1. Elbow method;
2. Contour coefficient;
3. CH index.

The elbow method is evaluated by calculating the within-cluster sum of squares ($WCSS$) for different values of K. The $WCSS$ measures the sum of the squares of the distances of all points within each cluster to its center of mass. Its calculation formula is as follows:

$$WCSS = \sum_{j=1}^{K} \sum_{i \in C_j} ||x_i - \mu_j||^2 \tag{1}$$

where K is the number of clusters; C_j is the first clustering; x is the sample point in the cluster; and c_i is the center of the first cluster. Theoretically, as the K value increases, the $WCSS$ value decreases. By plotting the $WCSS$ versus K value, the "inflection point" (i.e., the "elbow point") where the slope of the curve changes dramatically is determined to be the recommended optimal K value.

The contour coefficient combines the cohesion and separation of the clusters. For a single sample point i, The profile factor $s(i)$ is calculated as follows:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{2}$$

where $a(i)$ is the average distance between the point and all other points in the cluster to which it belongs (this is a measure of cohesion) and $b(i)$ is the average distance between the point and all other points in the cluster to which it is closest (this is a measure of sepa-

rateness). $s(i)$ has a value in the range of $[-1, 1]$, and the closer the value is to 1, the better the clustering is. The average profile coefficient is calculated for all samples, with the value of K yielding the highest score being selected as the optimal choice.

The Calinski-Harabasz (CH) index, also known as the variance ratio criterion, assesses the quality of clustering by calculating the ratio of inter-cluster scatter to intra-cluster scatter. Its score is defined as follows:

$$s_{CH} = \frac{Tr(B_k)}{Tr(W_k)} \times \frac{N-k}{k-1} \tag{3}$$

The total number of samples, the number of clusters, the trace of the inter-cluster scatter matrix (which indicates the degree of separation between clusters), and the trace of the intra-cluster scatter matrix (which indicates the degree of closeness within clusters) are related concepts. A higher score on the CH index implies that the clusters themselves are more tightly grouped, and that the clusters are further apart, and the clustering will be more effective.

The specific procedure of K-means++ is as follows: first, a sample point from the dataset is randomly selected as the first clustering center. Next, for each sample point in the dataset, the shortest distance from that point to the most recently selected clustering center is calculated:

$$D(x_i) = \min_{1 \le j \le k} (d(x_i, c_j)) \tag{4}$$

The shortest distance $D(x_i)$ is calculated, where $d(x_i, c_j)$ denotes the Euclidean distance between sample point x_i and cluster center c_j, as defined in Formula (5):

$$d(x_i, c_j) = \sqrt{\sum_{m=1}^{n} (x_{i,m} - c_{j,m})^2} \tag{5}$$

Based on the distance calculated above, the probability of selecting the next clustering center is proportional to the square of the distance:

$$P(x_i) = \frac{D(x_i)^2}{\sum_j D(x_j)^2} \tag{6}$$

The formula suggests that the further a point is from an already selected cluster center, the higher the probability that it will be selected as the next cluster center. With this strategy, newly selected clustering centers are moved away from the already existing centers, thus increasing the dispersion of clustering centers throughout the data space.

2.2.2. Canny Edge Detection Technology

In order to accurately extract the contours of individual color blocks from the K-means++ clustering processed layered image of the TWNY Prints (referred to as $I_{layered}$ here), we apply the Canny edge detection algorithm. The algorithm is widely used because it seeks to achieve the best balance between the three core goals of accurately labeling the true edges, precisely locating the edge positions, and ensuring that the edges are single-pixel thin lines through a series of rigorous mathematical steps. The resulting edge image E, which is the key input for the subsequent optimization steps, can be summarized as follows:

$$E = Canny(I_{layered}) \tag{7}$$

The whole Canny algorithm process begins with Gaussian filtering. After the color clustering on $I_{layered}$ of the TWNY Prints, although distinct color blocks are formed at the

macro level, there may still be a small amount of noise or unsmooth areas on the micro level. To prevent these imperfections from being misclassified as edges, the algorithm first smooths the image. This step is accomplished by convolving the image with a 2D Gaussian kernel with the core formula:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{8}$$

Here σ (standard deviation) controls the strength of the smoothing and choosing the right σ value can strike a balance between effective noise suppression and preservation of the original details of the print. The smoothed image $I_{smoothed}$ is obtained by convolving the original layered image with a Gaussian kernel:

$$I_{smoothed} = I_{layered} * G \tag{9}$$

Next, the algorithm enters the gradient computation phase, which aims to locate the boundaries between the different color blocks in the TWNY Prints image. The nature of the edges is a dramatic change in the intensity of the image, and the gradient is a mathematical tool to measure this change. The algorithm usually employs the Sobel operator, which approximates the intensity changes G_x and G_y of the smoothed image in the horizontal and vertical directions by means of convolution kernels in both directions. Subsequently, based on these two gradient components, the gradient intensity G and gradient direction Θ can be computed for each pixel.

The gradient strength is calculated by the following formula:

$$G = \sqrt{G_x^2 + G_y^2} \tag{10}$$

The value G reflects the "strength" of the edge. In the TWNY Prints, this means that the greater the difference in color between the two blocks, the higher the G value.

The gradient direction is given by the following:

$$\Theta = arctan\left(\frac{G_y}{G_x}\right) \tag{11}$$

However, a gradient intensity map alone is not enough, as the edges thus produced are blurred and have width. In order to distill these rough boundaries into clear single-pixel lines, the algorithm performs non-maximal value suppression. This step examines the gradient direction Θ of each pixel and compares the gradient intensity of that pixel with the two neighboring pixels before and after it along this direction. A pixel is retained only if its gradient strength is a local maximum in the neighborhood of its gradient direction; otherwise, the pixel is suppressed. This process is like fine-tuning the blurred color block boundaries of a print, eliminating all non-centered pixels to produce clear, slim candidate edge lines.

The final step is dual thresholding with hysteresis linking, which is a key decision-making process to determine which candidate edge lines end up as definitive edges. The algorithm sets a high threshold, T_{high}, and a low threshold, T_{low}. Pixels with gradient strengths higher than T_{high} are considered strong edges, which are reliable building blocks for the main contours of the TWNY Prints. Pixels with gradient strength lower than T_{low} are considered as noise and are rejected, and pixels with gradient strengths in between are defined as weak edges. Then, the algorithm finds and connects all the weak edge pixels connected to it along the path from all the strong edge pixels through a lagged connection strategy. Eventually, only the weak edges that can connect to the strong edges are retained to form a complete contour line. This mechanism allows the algorithm to preserve continu-

ous but varying strength lines in the prints (e.g., folds in clothing), while effectively filtering out isolated false edges caused by subtle textures, thus providing high-quality edge maps "E" for subsequent contour optimization.

2.2.3. U-Net Model and Cyclic U-Net Model

U-Net is an encoder–decoder architecture for image segmentation, named for its symmetrical U-shaped structure. U-Net was proposed by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in 2015 [41]. It is mainly designed to address the problem of scarce training data and the need for accurate pixel-level localization in biomedical image segmentation. The encoder (left path) consists of multiple convolutional and pooling layers that are progressively downsampled to capture the global contextual information of the image; the decoder (right path) progressively recovers the spatial resolution by upsampling (e.g., transposed convolution) and fuses the feature maps of the same layer with the encoder through jump connections to preserve the detail information. The jump connection combines shallow high-resolution features with deep semantic features to resolve the conflict between localization accuracy and semantic understanding in segmentation tasks. The output layer is convolved to generate pixel-level segmentation masks. U-Net and its variants have demonstrated excellent performance in medical image segmentation, and have become one of the most mainstream deep learning architectures in this field [42], These variants further optimize feature fusion through dense jump connections and nested structures, and improve segmentation accuracy for complex textures (e.g., gradients; halos). The high accuracy makes U-Net suitable for the needs of processing hierarchical works such as TWNY Prints. The encoder part uses a multi-stage downsampling module, each stage contains two convolutional layers (with ReLU activation function) and a maximum pooling layer, whose mathematical expression is.

$$E_l = MaxPool(ConvBlock(E_{l-1})) \tag{12}$$

In this, the deep semantic features of the image are gradually extracted by stacking convolutional kernels, while the pooling operation compresses the spatial dimensions to enhance the generalization ability of the model. The decoder part implements up-sampling by transposed convolution with the kernel formula:

$$\mathrm{D}_l^{\mathrm{up}} = \mathrm{ConvTranspose}(\mathrm{D}_{l+1}), \mathrm{D}_l = \mathrm{ConvBlock}(\mathrm{Concat}(\mathrm{D}_l^{\mathrm{up}}, \mathrm{E}_l)) \tag{13}$$

Here, the transposed convolution (step size) learns the spatial extension of the feature map by backpropagation, and the computation in the equation relies on the reconstruction of the input features by the convolution kernel. The hopping connection between the encoder and decoder fuses the shallow high-resolution details with the deeper semantic information through channel splicing (concatenation), by which the problems of edge blurring and detail loss are effectively mitigated. The output layer uses convolution to map the final features of the decoder to a pixel-level categorization space and combines them with *Softmax* functions to generate a multiclass segmentation mask:

$$Mask(i,j) = Softmax(\sum_{c=1}^{C} D_1(i,j,c) \cdot W_c) \tag{14}$$

This project proposes the Cyclic U-Net architecture, an innovative model designed for hierarchical image segmentation, as shown in Figure 1, to address the unique characteristics of TWNY Prints.

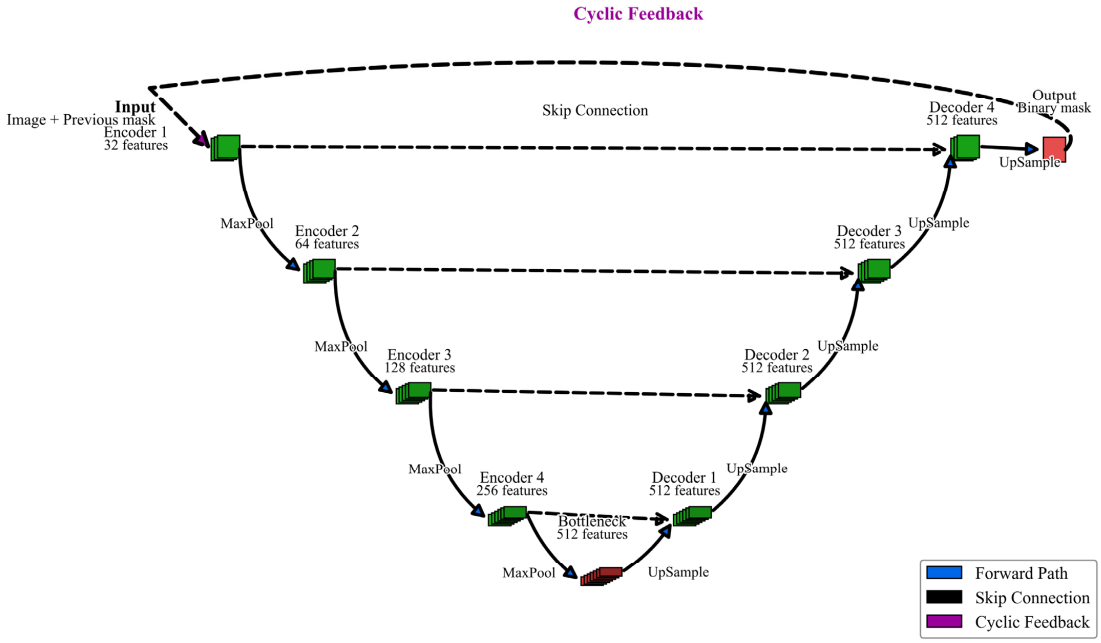


Figure 1. The architecture of the proposed Cyclic U-Net. The network is comprised of a symmetric encoder–decoder structure. The encoder path contains four downsampling blocks that sequentially reduce spatial dimensions while increasing feature channels from 32 to 256. A bottleneck layer with 512 features connects the encoder to the decoder. The decoder path then uses four upsampling blocks to progressively restore the spatial resolution, with feature channels decreasing from 512 back to 32. Skip connections are utilized to feed feature maps from each encoder block to its corresponding block in the decoder. The final component is a cyclic feedback path, which routes the output segmentation mask back to be concatenated with the input image.

While traditional U-Net extracts all target categories in a single forward propagation, Cyclic U-Net adopts a recursive processing strategy to extract different color palette layers in the image layer by layer through multiple iterations, which is more suitable for processing TWNY prints with a clear hierarchical structure. Different from traditional U-Net, Cyclic U-Net introduces a recursive processing mechanism. The input of the model includes not only the original image, but also the cumulative mask of the previously extracted layers, which can be expressed as follows:

$$X_t = [I, M_{t-1}] \tag{15}$$

where I is the original image and M_{t-1} is the combination of all layer masks extracted in the previous $t-1$ iterations. This design allows the model to "remember" what has been extracted and focus on finding new layers that have not yet been extracted. Cyclic U-Net introduces an adaptive stopping mechanism based on the area ratio and layer overlap, which automatically stops the iteration when the area of the newly extracted layer is smaller than a preset threshold or the overlap with the extracted layer exceeds a threshold:

$$Stop = \begin{cases} True, & if \frac{Area(M_t)}{Area(I)} < \tau_{area} \\ True, & if \frac{|M_t \cap M_{t-1}|}{|M_t \cup M_{t-1}|} > \tau_{iou} \\ False, & otherwise \end{cases} \tag{16}$$

2.3. Research Route

In this study, 200 TWNY Prints were selected from the *Complete Collection of Chinese Woodblock New Year Prints* as the sample set [43]. Since the dataset consists of scanned

images from books, the overall colors tend to be dull and grayish, with extraneous background tones. To address this, batch preprocessing was performed using computer-based image enhancement techniques to restore the original color fidelity of the images and remove the unnecessary background. No dedicated ICC color correction profile was applied due to the dataset being derived from published book scans; instead, a consistent preprocessing pipeline was applied to all samples to maintain relative color fidelity across the dataset.

To ensure accurate input data for subsequent processing, all TWNY Print images first undergo preprocessing to restore color fidelity and remove extraneous background tones. This step involves white balance correction, contrast enhancement, mild saturation adjustment, background normalization, and light edge sharpening, effectively compensating for the dull, grayish tones and color deviations introduced during book scanning.

Following this, the proposed technical framework—comprising four core stages—is applied to achieve precise digital layer separation and standardized color reconstruction. The method automatically determines the necessary number of color layers and performs color clustering within a perceptual color space aligned with human visual perception, ultimately generating pure, discrete color layers suitable for standardized digital production. To illustrate the preprocessing stage of this study, two representative images of TWNY Prints, shown in Figure 2a,b, were selected from the sample set to present a side-by-side comparison of the original scanned images and the results after color restoration and background removal.

(**a**) (**b**)

Figure 2. Preprocessing examples of TWNY Prints used throughout this study, showing the original scanned image (**left**) and the processed image after color restoration and background removal (**right**): (**a**) Blossoms bring wealth and honor; (**b**) Fruit bowl.

In addition, a Cyclic U-Net model is introduced to improve the efficiency of layer separation for TWNY Prints. This model is designed to learn the interrelationships between different layers, enabling direct output of the separated color layers.

2.3.1. Contour Extraction Using the Canny Operator

Given the artistic characteristics of TWNY Prints, which feature clear black ink outlines as the structural framework, preprocessing of the input images was first performed using the Canny edge detection algorithm. The purpose of this step is to extract the "black outline layer," which serves as the foundation for the print composition. This approach not

only aligns with the traditional "one block, one color" overprinting process but also provides precise layer boundaries for the subsequent extraction of pure color plates, thereby improving the accuracy of color segmentation.

To optimize the performance of the Canny edge detection algorithm, systematic comparative experiments were conducted, focusing on two key parameters: the low threshold and high threshold. The Canny algorithm uses a dual-threshold strategy for edge detection, where the high threshold identifies strong edges, while the low threshold helps connect these edges to form complete and continuous contours.

Figures 3 and 4 present the Canny edge detection results for selected representative images of TWNY Prints.

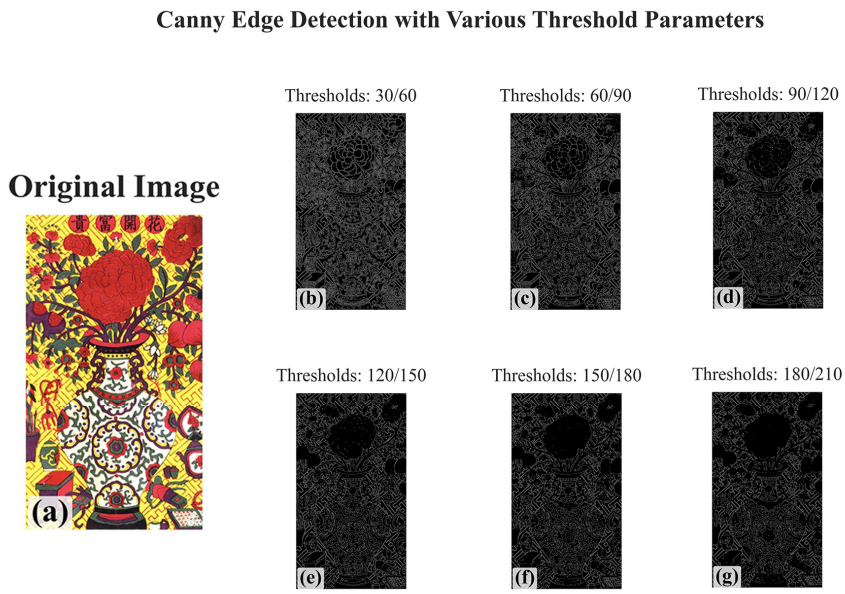


Figure 3. Canny edge detection threshold test 1 and parameter settings. (**a**) The original input image. (**b**–**g**) The resulting edge maps produced by the Canny algorithm with low/high thresholds set to 30/60, 60/90, 90/120, 120/150, 150/180, and 180/210, respectively. The figure visually demonstrates the algorithm's sensitivity to threshold selection, showing the progression from an excessively noisy output at low thresholds (e.g., (**b**)) to the loss of fine details at high thresholds (e.g., (**g**)).

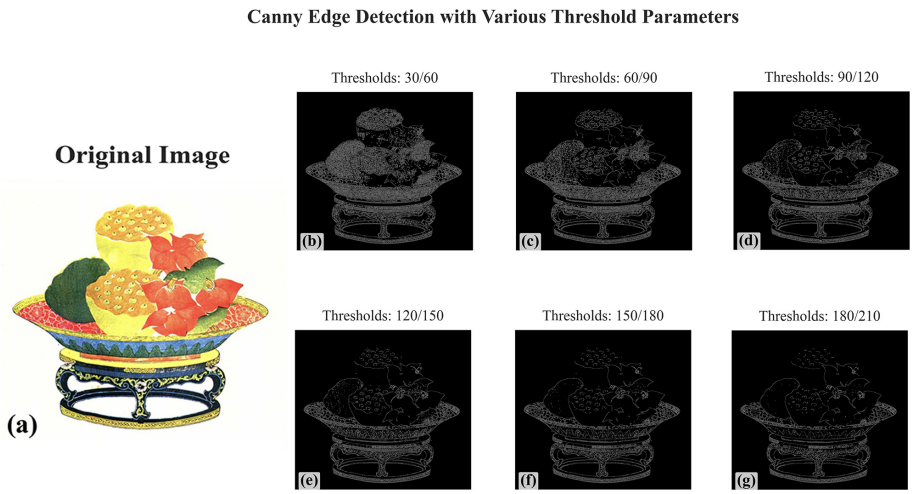


Figure 4. Canny edge detection threshold test 2 and parameter settings. (**a**) The original image for this case (fruit bowl). (**b**–**g**) show the Canny edge detection results using low/high thresholds of 30/60, 60/90, 90/120, 120/150, 150/180, and 180/210, respectively. This series of images also visually demonstrates the progression from an output with co-existing detail and noise at low thresholds (**b**) to one with only major contours and a loss of detail at high thresholds (**g**).

Testing successive threshold intervals of 30–60–90–120–150–180–210 showed that the combination of a low threshold set to 90 and a high threshold set to 120, as shown in Figures 3d and 4d, most accurately captures the ink line characteristics of the TWNY Prints. The advantage of this parameter combination lies in the fact that the lower low threshold (60) can capture enough potential edge information to avoid line breakage, while the moderate high threshold (90) effectively filters out the false edge generated by the color transition region and texture details, preserving the integrity and continuity of the ink lines. Excessively high threshold combinations (e.g., 150–180 or 180–210) result in the loss of fine ink lines, while excessively low threshold combinations (e.g., 30–60) introduce too much noise and non-ink line structures that interfere with subsequent color clustering.

2.3.2. Minimum Number of Clusters (K_min) Based on Color Dominance Analysis

To enhance the specificity of clustering analysis and prevent the omission of key dominant colors, the design of a preprocessing mechanism was implemented to determine the minimum number of clusters (K_min) based on color dominance analysis. This mechanism adaptively identifies K_min by performing statistical analysis of the image's color space and detecting visually significant colors.

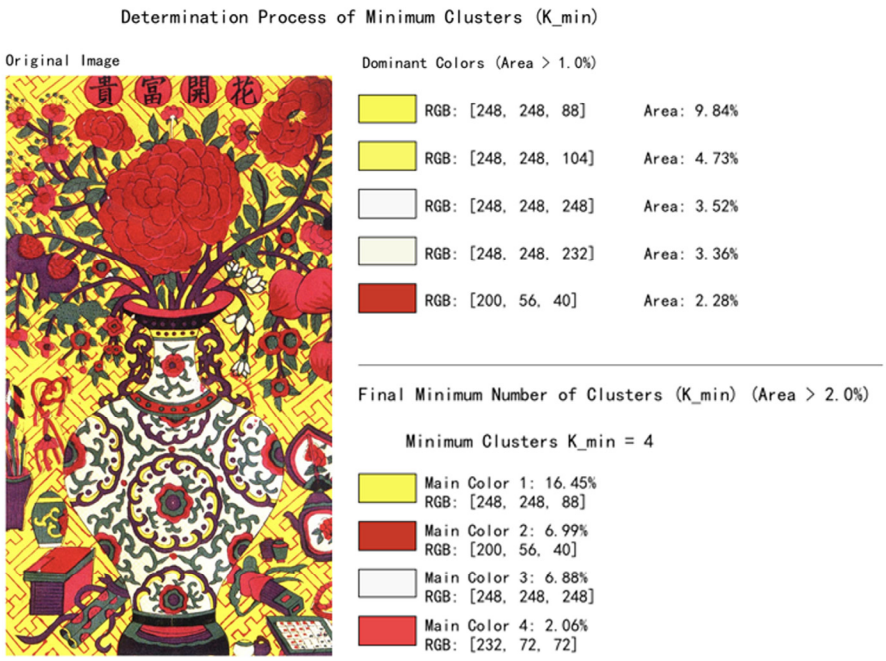
First, a color histogram is constructed in the RGB color space by discretizing the color value range into a $16 \times 16 \times 16$ three-dimensional grid, with each dimension corresponding to the R, G, and B channels. This partitioning preserves sufficient color detail while avoiding excessive subdivision that could lead to increased computational cost. The color area distribution is then derived by calculating the percentage of pixels in each non-empty grid cell relative to the total number of pixels.

Next, an area threshold of 1% (during color extraction) is applied to filter out all colors whose area ratio exceeds this value, defining them as "dominant colors." These colors play a key role in visual perception and must be retained during clustering, with their pairwise Euclidean distances subsequently computed. If the distance between two colors is below a predefined merging threshold (set to 20.0), they are considered visually similar and are merged into a single dominant color. During merging, smaller area colors are absorbed by larger area colors to ensure that the resulting colors accurately reflect the visual characteristics of the original image.

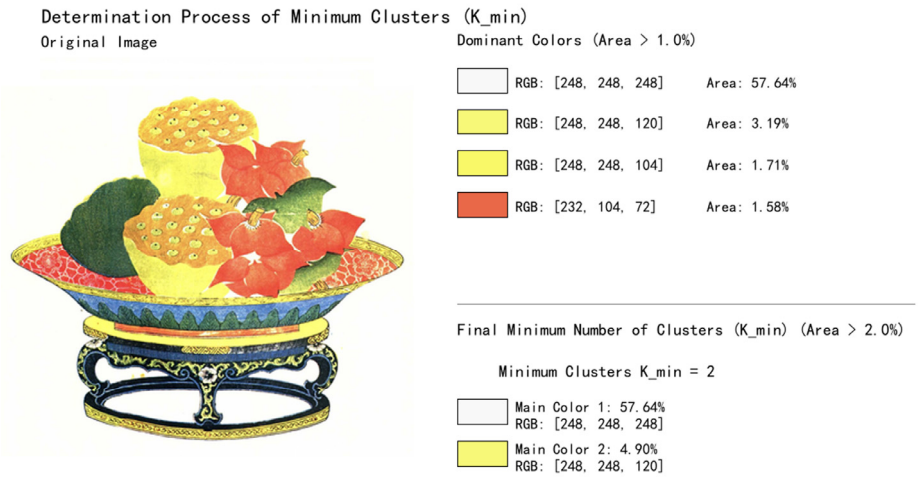
After the merging process, the remaining colors are re-ranked by area size, and a second area threshold (2% during color extraction) is applied to finalize the dominant color set. The number of remaining dominant colors is defined as K_min, representing the minimum number of clusters required to capture the primary color information in the image.

K_min serves as the lower bound for subsequent cluster number determination using methods such as the elbow method, silhouette coefficient, and Calinski-Harabasz (CH) index, with the search range set as [K_min, K_max]. In our experiments, K_max is set to 12, considering the typical color characteristics of TWNY Prints and balancing computational precision and efficiency.

The core advantage of this preprocessing mechanism lies in its adaptability to dynamically adjust clustering parameters according to the image's inherent color complexity. This approach prevents both under-clustering, which could result in the loss of key dominant colors, and over-clustering, which increases computational burden and introduces noise. Experimental results demonstrate that this mechanism is particularly effective for images with distinct color hierarchies, such as TWNY Prints, accurately capturing essential color layers and providing a reliable foundation for subsequent digital restoration and color reconstruction. Figure 5 illustrates the color merging process applied to representative sample images.



(**a**)



(**b**)

Figure 5. The process of determining the minimum number of clusters (K_min) through area share sorting and merging. This figure illustrates the method used to determine the number of main color clusters (K_min) for the sample images previously defined in the text. The process involves two steps: first, extracting dominant colors with an area share greater than 1.0%, and subsequently, determining the final K_min by merging similar colors and applying a stricter area threshold of 2.0%. (**a**) Shows the process applied to the "fruit bowl" sample image, for which the calculation results in a final K_min = 2. (**b**) Shows the result of applying the same process to the "vase" sample image from the "TWNY Prints" collection. Due to its higher complexity in color and composition, the method adaptively determines its minimum number of clusters to be K_min = 4.

2.3.3. Optimal Number of Clusters (K) and Method Selection Based on Multi-Indicator Assessment

To ensure that color difference calculations align more closely with human visual perception, the conversion of image color data from the RGB color space to the CIELAB color space was performed. The perceptual uniformity of the CIELAB space provides a more

reliable mathematical foundation for subsequent clustering evaluations. During preprocessing, standardization was applied, and a weighted strategy was introduced across different channels: the a* and b* channels in the CIELAB space, which carry the chromatic information, were assigned a weight factor of 1.5 to enhance color distinction capability.

To balance local color block integrity with global color consistency, the feature construction process incorporated spatial coordinate information, assigning a weight ratio of 5.0 to color information and 0.3 to spatial positioning. This design enables the clustering algorithm to effectively distinguish color differences while maintaining spatial continuity.

Within a search range not lower than the initial minimum cluster number K_min (as determined by the dominant color area analysis described previously), the determination of the optimal number of clusters K was performed through a comprehensive evaluation combining three classic clustering validation methods:

Elbow Method: This method analyzes the within-cluster sum of squares (WSSs) curve with respect to varying K values, identifying the point of maximum curvature as the optimal K. The implementation of an improved algorithm based on second-order difference and slope change analysis, with the introduction of a positional weighting factor, was carried out to avoid the selection of excessively high cluster numbers.

Silhouette Score: This metric evaluates the compactness and separation of clusters, with values ranging from −1 to 1, where scores closer to 1 indicate higher clustering quality. To balance computational efficiency with accuracy for large datasets, the adoption of a random sampling strategy was combined with silhouette score calculation based on 20,000 sampled pixels.

Calinski-Harabasz Index (CH-Index): This index assesses the clustering performance based on the ratio of between-cluster dispersion to within-cluster cohesion, with higher values indicating better clustering quality. The CH-Index is particularly sensitive to data distribution characteristics and performs well on TWNY prints with clearly separated color regions.

A weighted majority voting decision mechanism was developed; when two or more evaluation methods recommend the same number of clusters, that value is adopted as the final K. In cases where all three methods yield inconsistent results, priority is given to the value suggested by the silhouette score. Any recommended K values below the minimum acceptable threshold (typically set to 3) are automatically adjusted to the minimum to ensure sufficient color detail is preserved.

The layer separation results of representative sample images obtained using this method are shown in Figure 6.

In practical applications, the clustering performance of three color spaces (RGB, HSV, and CIELAB) was compared through simultaneous experiments, and the results were used to determine the entire technical process.

2.3.4. Image Segmentation of TWNY Prints Based on the Cyclic U-Net Model

To simplify the overall workflow and reduce the time required for layer separation, training of a Cyclic U-Net model was conducted to perform automated layer extraction for TWNY Prints. The process begins with systematic mask segmentation of the prints, where binary masks are generated for each individual color layer within every print. Subsequently, these layer-wise masks are sequentially merged to create an accumulated mask dataset tailored for Cyclic U-Net training.

The Cyclic U-Net model adopts the classical U-Net architecture but is specifically designed for progressive, layer-by-layer color plate extraction. The input to the network consists of a four-channel dataset (RGB image plus the cumulative extracted layer mask), while the output is a single-channel binary mask representing the next layer to be extracted.

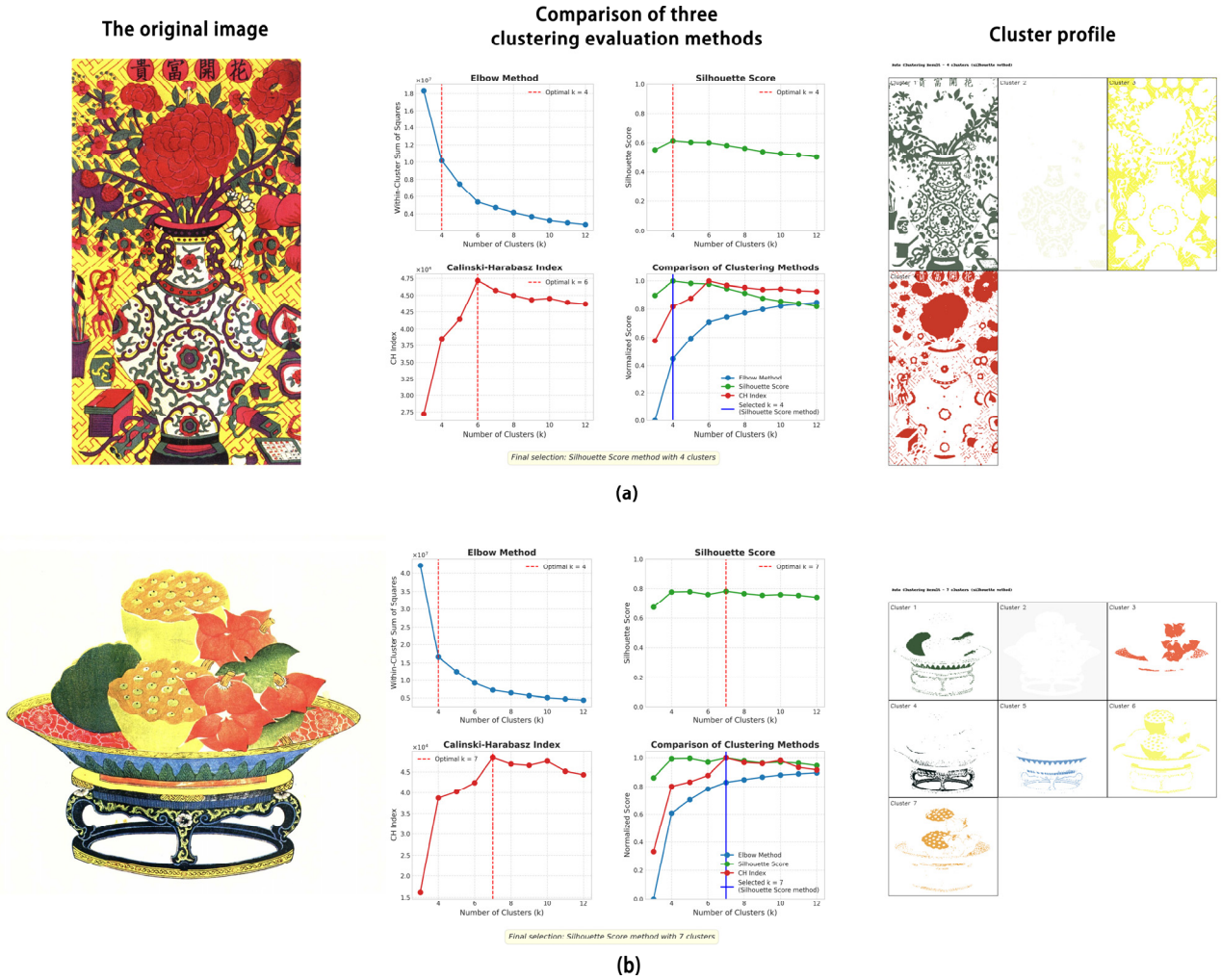


Figure 6. The workflow of clustering analysis applied to the sample images. (**a**) Blossoms bring wealth and honor; (**b**) Fruit bowl. The central column compares three evaluation methods (e.g., elbow method, silhouette score) to determine the optimal number of clusters, which is indicated by the red vertical line. The right column displays the final cluster profiles, illustrating the segmentation of the original images into distinct components.

Structurally, the model comprises an encoder–decoder framework. The encoder contains four downsampling stages, each consisting of two 3×3 convolutional layers followed by max-pooling, progressively increasing the number of feature channels ($64 \rightarrow 128 \rightarrow 256 \rightarrow 512$) while halving the spatial resolution. The decoder uses transposed convolutions for upsampling and integrates skip connections from corresponding encoder stages to preserve spatial details. Each convolutional block is equipped with batch normalization and ReLU activation functions, with dropout layers applied in deeper stages to mitigate overfitting.

To ensure the reproducibility of our experimental setup, the Cyclic U-Net was trained using a progressive training strategy. The network employed an encoder depth of five layers with decoder channel configuration [512, 256, 128, 64, 32], with attention mechanisms enabled and a dropout rate of 0.3. The AdamW optimizer was used with a weight decay of 1×10^{-5}, and the learning rate was gradually reduced from 1×10^{-3} to 1×10^{-4} during progressive training, dynamically adjusted via a CosineAnnealingWarmRestarts scheduler. A decreasing batch size strategy was adopted, starting from 4 in the initial stage and reducing to 1 in the final stage, over a total of 190 training epochs (split into three phases: 80 + 60 + 50).

Data augmentation was customized for TWNY Prints, including ±15° rotation, translation and scaling within a 0.1 range, brightness adjustment in the range [0.8, 1.2], horizontal flipping, Gaussian noise with a strength of 0.02, and elastic deformation. Vertical flipping was disabled to preserve the directional characteristics of the prints. Gradient clipping (maximum norm = 1.0) and early stopping were applied to ensure training stability and prevent overfitting.

The Cyclic U-Net operates in a recursive extraction manner: the initial input consists of the original image combined with an empty mask, yielding the binary mask for the first color layer. In subsequent iterations, the input is updated by combining the original image with the accumulated extracted masks, allowing the model to predict the next color layer.

To ensure the quality and continuity of the extracted color layers, multiple validation mechanisms are incorporated, including confidence thresholding, area proportion filtering, and clustering-based evaluation. This recursive design enables a single model to handle the extraction of an arbitrary number of color layers, significantly simplifying the traditional layer separation process, which often requires repeated manual parameter tuning and training of multiple models.

Further improvements are achieved through clustering optimization in the CIELAB color space and morphological post-processing, including edge contour preservation and connected region merging, which enhance mask precision and ensure accurate boundary separation between different printing color plates.

This end-to-end automated layer separation approach greatly improves the efficiency and accuracy of the digital processing workflow for TWNY Prints. Figures 7 and 8 demonstrate the iterative mask extraction process applied to representative sample images.

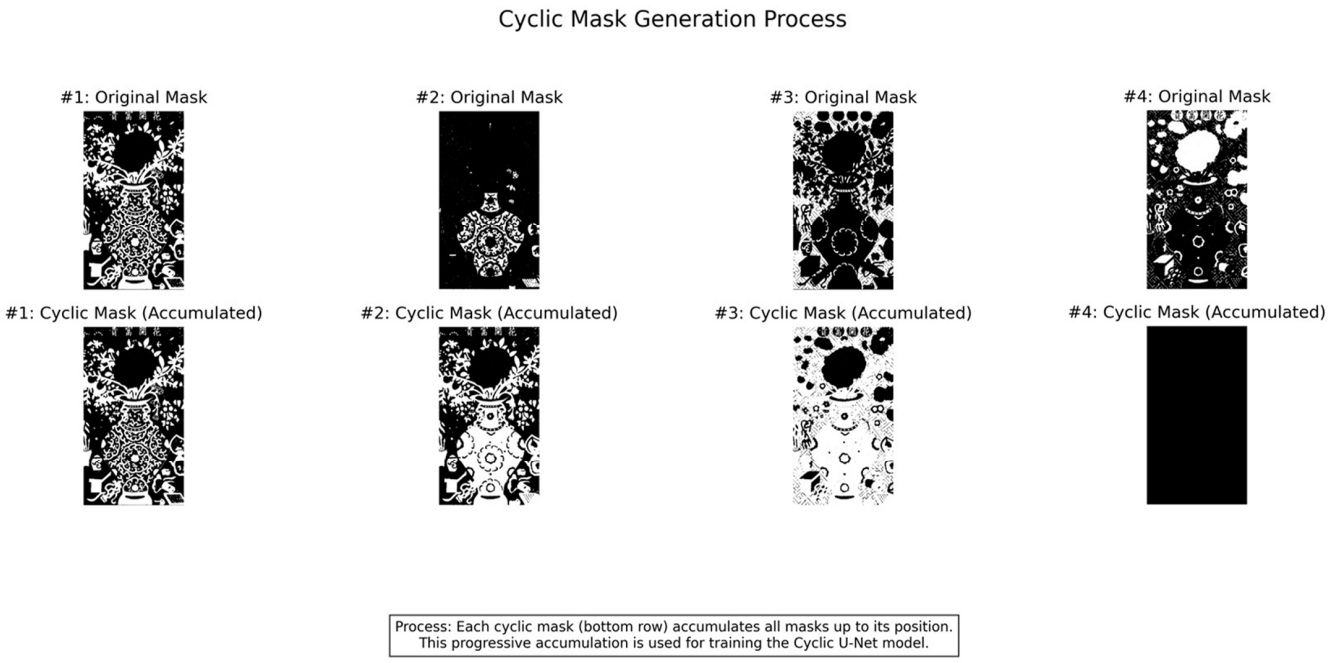


Figure 7. The cyclic mask generation process 1. This process involves the sequential accumulation of original masks (**top row**) to create progressively denser cyclic masks (**bottom row**). This sequence of accumulated masks serves as the training data for the cyclic in-fill model.

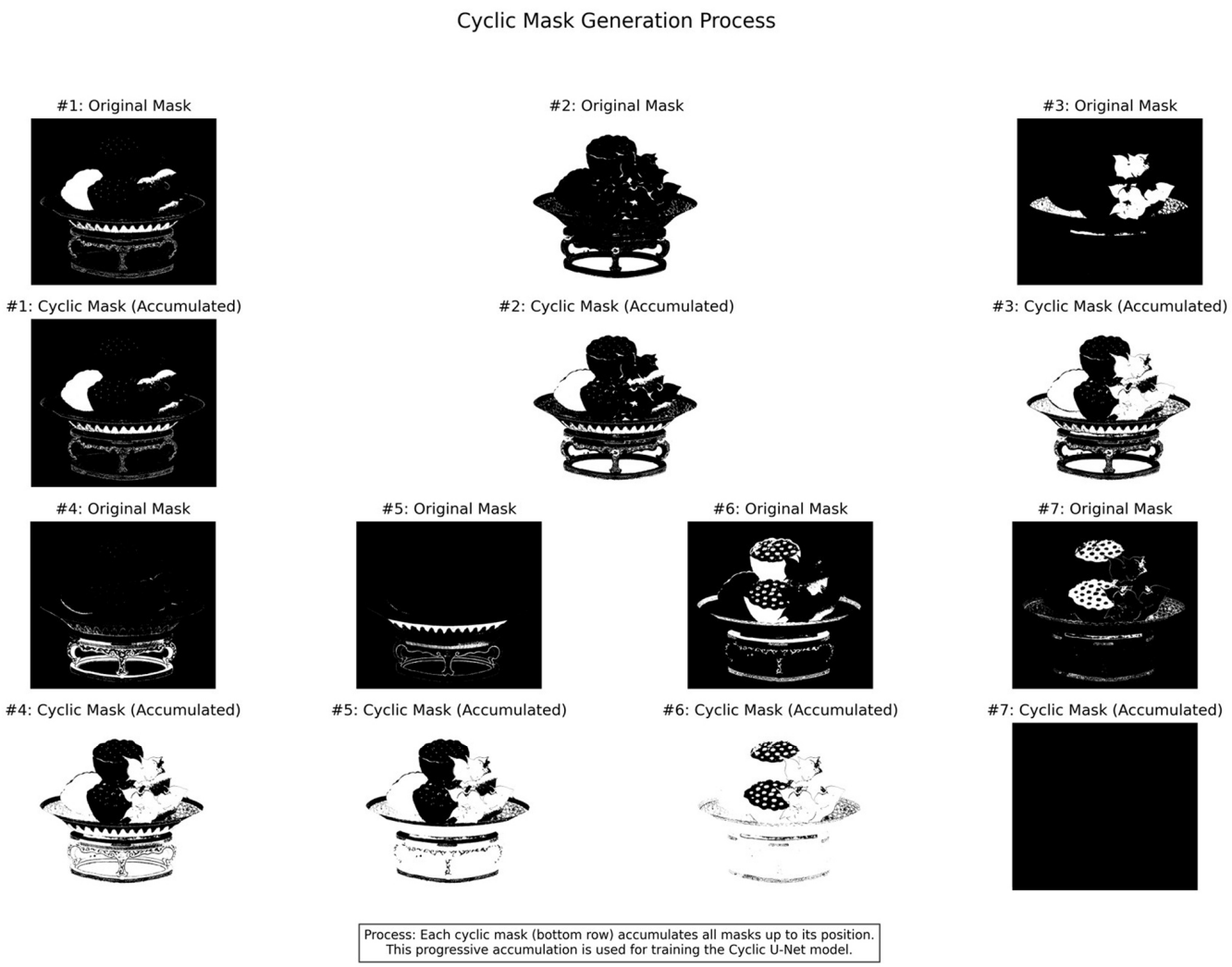


Figure 8. The cyclic mask generation process 2. This figure extends the process shown in Figure 7, illustrating subsequent sequential accumulations of original masks (**top row**, continuing from mask #4) to generate further progressively denser cyclic masks (**bottom row**). This extended sequence of accumulated masks continues to serve as the training data for the cyclic in-fill model.

2.4. Computational Resources and Runtime Efficiency

This study's computational resource and runtime efficiency tests were conducted in the following hardware environment: a single NVIDIA RTX 4090 GPU (24 GB VRAM), a 12-core Intel Xeon Platinum 8352V processor (2.10 GHz), 90 GB system memory, running Ubuntu 20.04, with PyTorch 2.0.0 and CUDA 11.8 acceleration.

Performance testing under this hardware configuration showed that the K-means++ algorithm processed 512 $\times$ 512 resolution images in an average of 1.8 $\pm$ 0.3 ms, whereas the full Cyclic U-Net pipeline required 127 $\pm$ 15 ms end-to-end. K-means++ was thus approximately 67 times faster than Cyclic U-Net. Specifically, the Cyclic U-Net's processing time was distributed as follows: encoder forward pass–23 ms, six-stage cyclic decoding–89 ms, and output generation–15 ms, with peak GPU memory usage of 19.2 GB. In contrast, K-means++ clustering optimization required only 0.8 GB VRAM, achieving a throughput of 526 images/s over a 1000-image test set, compared to 7.8 images/s for Cyclic U-Net.

Despite the speed advantage of K-means++, when used jointly with Cyclic U-Net in practical applications, the extra time cost of K-means++ accounted for only 1.4% of the overall pipeline, exerting minimal impact on total runtime while significantly improving seg-

mentation boundary precision and continuity. This demonstrates that the hybrid architecture achieves superior segmentation quality without sacrificing computational efficiency.

3. Results

3.1. Canny Contour Extraction

As shown in Figure 9, the Canny edge detection procedure completes the first stage of contour extraction. The low and high thresholds are set to 90 and 120, respectively, as discussed earlier. Taking Image 1 in the figure as an example, the application of Canny edge detection with a low threshold of 90 and a high threshold of 120 reveals the complex, multi-layered structure of the original image in detail.

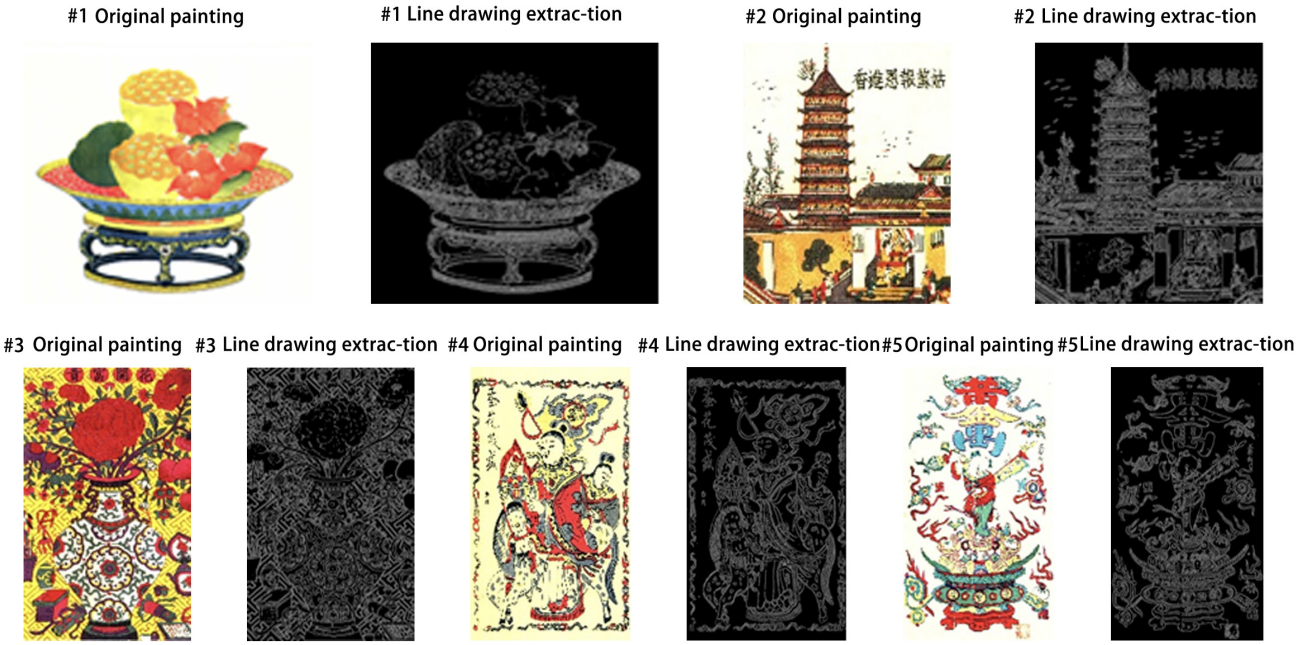


Figure 9. Demonstration of the line drawing extraction process on five different TWNY Prints. For each numbered pair, the original print (**left**) is processed to generate a corresponding line drawing (**right**), which effectively captures the essential contours and edges of the artwork.

First, the central circular motif of the image is clearly outlined, with both the primary contour and the smaller concentric circular edges inside it being well preserved. Surrounding the central motif, the edges of multiple curled, symmetrical patterns resembling petals or leaves are also successfully detected, collectively showcasing the organic and ornate nature of the design.

In addition to the central elements, the outer framing structures and corner decorations are also distinctly visible. The algorithm effectively captures the square frame lines around the periphery of the image, as well as the decorative patterns located at each corner. Furthermore, this parameter configuration preserves a significant amount of fine, grid-like or dotted textures used to enrich the background. While these dense edges increase the visual complexity of the result, they also faithfully reflect the highly intricate and textural style of the original print.

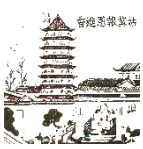
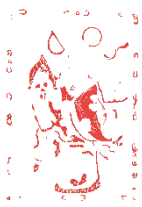
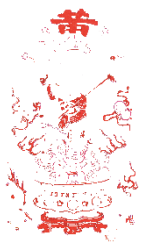
Thus, the threshold combination of a low threshold of 90 and a high threshold of 120 is identified as the most suitable parameter setting for Canny edge detection in the context of TWNY Prints.

3.2. K-Means Color Clustering Extraction

K-means++ clustering was performed without removing the edge contours extracted by the Canny algorithm, ensuring that the structural outlines remained intact during the color segmentation process. Meanwhile, three different evaluation methods were employed to determine the optimal number of clusters for each Taohuawu woodblock print.

As shown in Table 1, the specific layer separation results for each print are presented, illustrating the effectiveness of the proposed approach.

Table 1. This table displays the hierarchical results of K-means++ clustering on five samples of TWNY Prints. The 'Original painting' column shows the original images, 'color x' denotes the result of a particular color layer's segmentation, and 'Clustering results' indicates the final merged segmentation results after further combining the layers.

Serial Number Name	Original Painting	Color 1	Color 2	Color 3	Color 4	Color 5	Clustering Results
1						None	
2					None	None	
3					None	None	
4					None	None	
5						None	

The reconstruction of the clustering results in Figure 10 for each of the three color spaces (HSV, CIELAB, and RGB) using the three methods will also be shown. For simplicity, only one reconstructed and merged layer will be presented.

The optimal number of clusters (k) for each color space was determined using the silhouette method, yielding k = 4 for RGB, k = 7 for HSV, and k = 4 for LAB. The figure

visualizes the color segmentation results at both k = 4 (middle column) and k = 7 (right column), with the optimal outcome for each color space explicitly labeled.

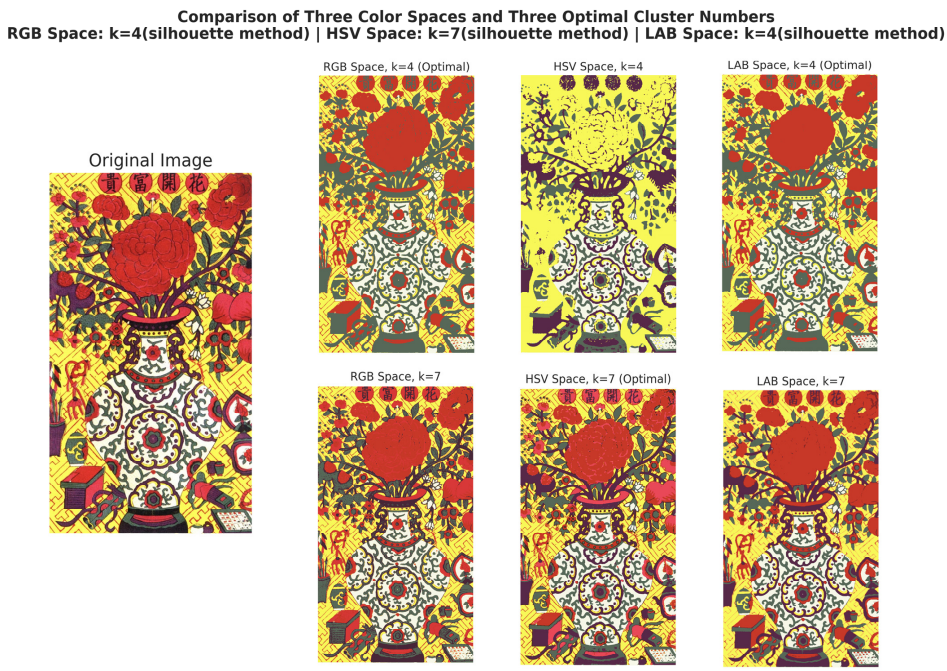


Figure 10. Comparison of image clustering results using three different color spaces (RGB, HSV, and LAB) on the vase.

The processing of TWNY Prints involved K-Means clustering in the CIELAB color space, followed by layer optimization in the RGB space. The primary advantage of selecting CIELAB lies in its spatial perceptual uniformity. This space is designed so that the Euclidean distances between values better correspond to the color differences perceived by the human eye. This means that when the algorithm searches for and aggregates pixels with similar colors in this space, the results align more closely with human intuition about "similar colors," effectively avoiding the incorrect separation of colors that appear visually similar but have significant differences in RGB values.

At the same time, the use of HSV (hue, saturation, value) space for this task was evaluated and ruled out. Although HSV intuitively separates color (hue), purity (saturation), and luminance (value), it has significant drawbacks in the field of image segmentation, especially when dealing with those images that contain a large number of black, white, and low-saturation colors. The main problem with the HSV space lies in the instability of its hue component. For colors that are close to black, white, or gray (i.e., colors with low saturation or extreme brightness), the hue value becomes extremely unstable, and small changes in noise or pixel values can cause the hue value to jump dramatically. As shown in the clustering comparison diagram in Figure 10, there are more black lines, and this instability of the hue component can severely affect the clustering algorithm, resulting in contour lines that cannot be stabilized into a class, thus destroying the structural integrity of the image.

The superiority of the CIELAB method can be clearly verified from our processing results (as shown in the "CIELAB k = 12" and "CIELAB k = 17" figures). The algorithm succeeds in accurately separating key visual elements in the image—the red halo at the top, the skin color of the character, the green area of the robe, and the yellow background-into separate color layers. Crucially, the CIELAB space stabilizes the low-saturation black contours so that they are preserved intact, maintaining the structure of the original artwork.

In summary, while the HSV space is useful in some scenarios, its instability in near-colorless regions makes it unsuitable for complex art image segmentation that requires precise preservation of black contours. On the other hand, the perceived homogeneity of the CIELAB space provides a more stable and reliable basis for segmentation, ensuring that the results both simplify the colors and remain faithful to the artistic character of the original artwork.

3.3. Cyclic U-Net Results Presentation

The output of Cyclic U-Net is the same black and white binary mask as the output of general U-Net. In the previous section, the actual processing and generation of cyclic masks for simulating the Cyclic U-Net were presented. Instead of presenting the output process, the output results are shown directly here in Table 2.

From the resulting images, it can be observed that the original K-means clustering method effectively segments the images, demonstrating excellent performance in handling edge regions and noise suppression. Although the Cyclic U-Net model does not achieve the same level of accuracy as K-means clustering, it significantly reduces the processing time required for a single TWNY Print.

Table 2. This table showcases a comparison of the layering results between K-means++ and Cyclic U-Net, using the vase as an example. K-means++ decomposes the image into four layers, while Cyclic U-Net breaks it down into five layers. The table presents the mask results, original color results, and the reconstruction results based on the layered segmentation for both methods.

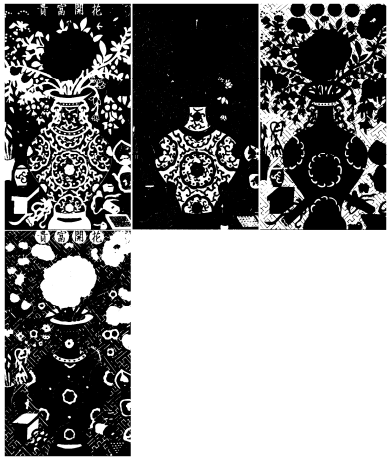
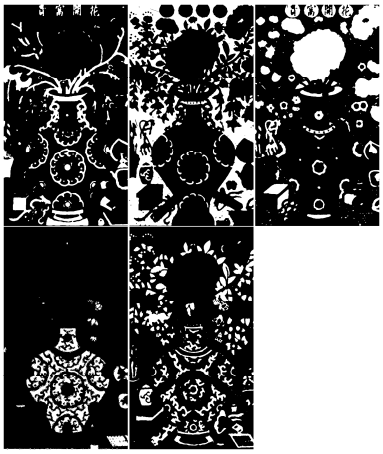
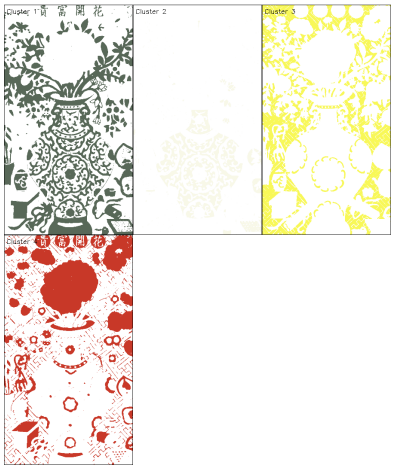
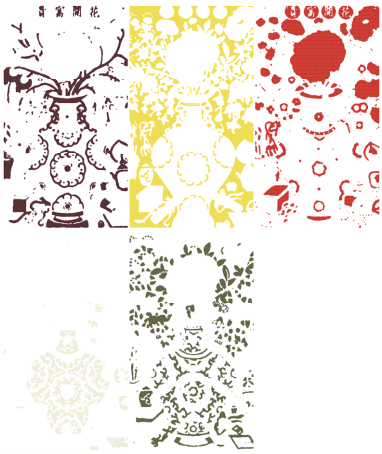
	K-Means Clustering Results	**Cyclic U-Net Results**
Masking Result Chart		
Layered results map		

Table 2. *Cont.*

	K-Means Clustering Results	Cyclic U-Net Results
Comparison chart of reconstruction results		

Performance evaluation of the Cyclic U-Net model in the hierarchical segmentation of TWNY Prints is shown in Figure 11. The results demonstrate that the model excels in extracting primary artistic elements, with IoU, Dice, and F1 scores for the first layer reaching 0.612, 0.759, and 0.762, respectively, confirming its reliability in core pattern recognition. Performance systematically declines with increasing layer depth, with the second layer achieving 0.387, 0.558, and 0.551 for the same metrics, still maintaining moderate segmentation accuracy. From the third layer onward, performance drops markedly, reflecting the inherent difficulty in extracting fine-grained details and background elements.

Precision–recall analysis shows a consistently conservative prediction behavior across all layers, with precision exceeding recall at all stages—from 0.834/0.701 in the first layer down to 0.198/0.045 in the sixth layer. This characteristic is of particular value in the digitization of artworks, as high precision ensures the accuracy of extracted patterns and minimizes false detections that could interfere with art analysis. Layer prediction accuracy analysis reveals the dataset's natural imbalance: simple compositions (1–2 layers) dominate, while complex multi-layer compositions are relatively rare, with prediction accuracy decreasing from 87.3% for single-layer prints to 21.7% for six-layer prints.

Performance decay analysis shows the steepest drop (26.5%) from the first to the second layer, followed by a relatively stable decline (28.1–28.4%) until the final layer, which shows the largest drop (52.9%). The training convergence curve demonstrates the effectiveness of the three-stage progressive learning strategy, with final training and validation losses stabilizing at 0.143 and 0.196, respectively, confirming the model's good generalization capability. These results highlight the practical potential of Cyclic U-Net in the digital preservation of cultural heritage, providing a reliable technical foundation for the artistic study and digital restoration of TWNY Prints.

Moreover, this work provides a foundational framework for applying deep learning techniques to the layer separation of TWNY Prints, paving the way for future advancements in digital restoration and automated processing. Historical or modern prints whose original woodblocks have been preserved are selected, and the algorithm's results are compared with the actual printing samples of each woodblock (Figure 12).

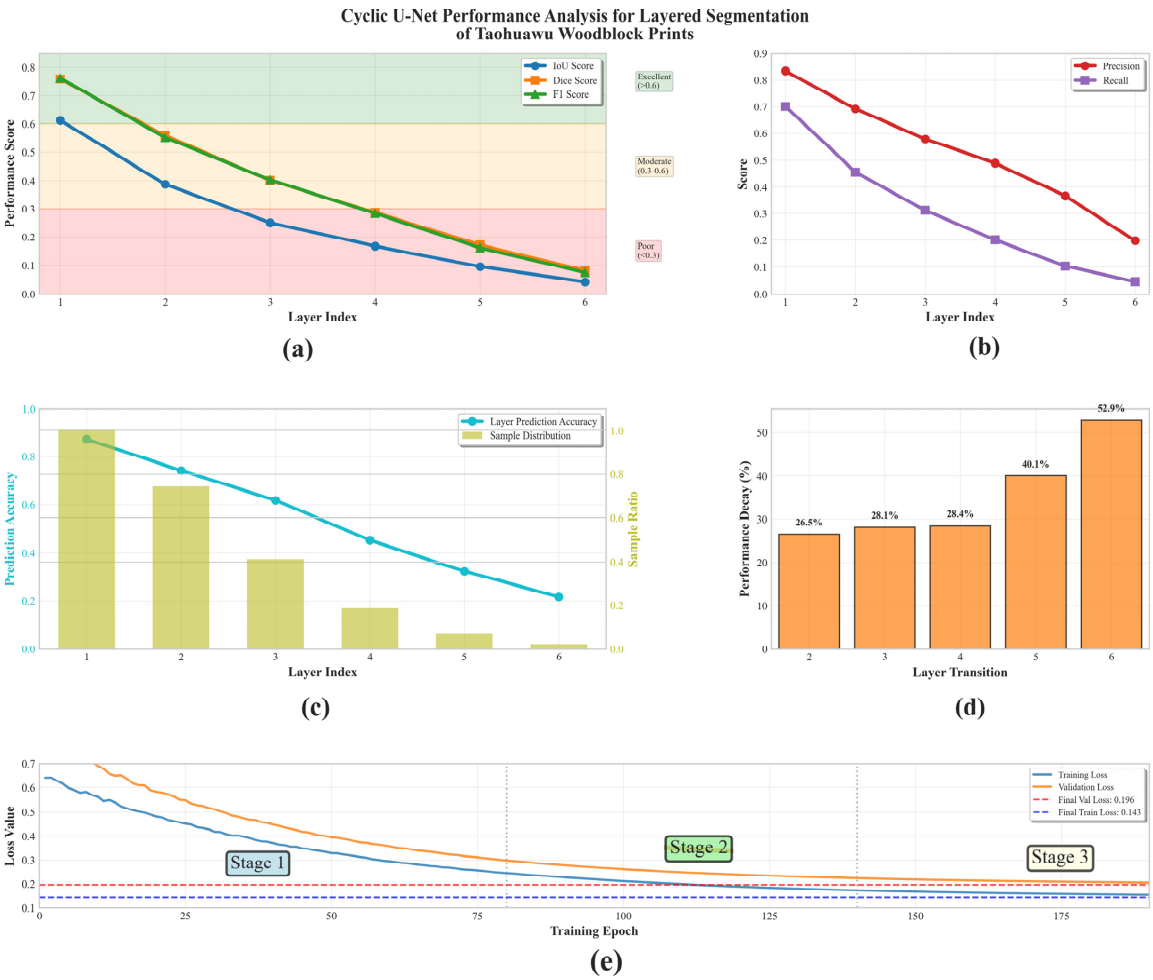


Figure 11. Cyclic U-Net performance analysis for layered segmentation of Taohuawu woodblock New Year prints (TWNY Prints): (**a**) segmentation performance by layer; (**b**) precision vs. recall; (**c**) layer prediction accuracy and sample distribution; (**d**) layer-wise performance decay; (**e**) training convergence.

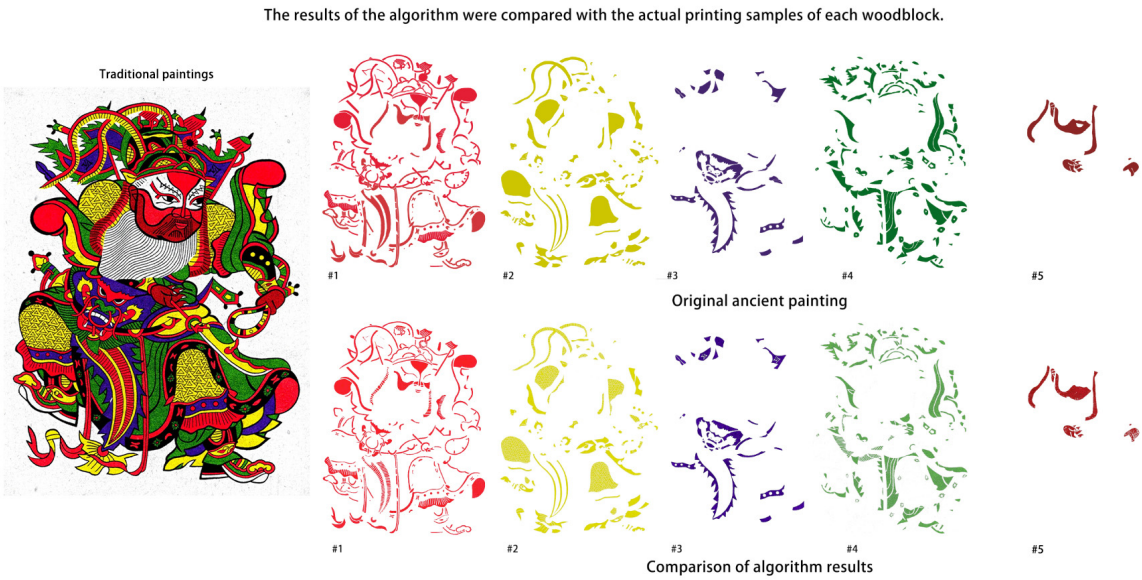


Figure 12. The traditional woodblock New Year painting "Door God", comparing the traditional color-printed woodblocks with the results obtained through algorithms. #1–#5 are the five over-printed color layers of the traditional color-printed woodblocks.

4. Discussion

4.1. Technical Advantages

This study's technical solution fully leverages the unique value of computer vision algorithms in the digital preservation of intangible cultural heritage. Targeting the distinctive characteristics of TWNY Prints, which are marked by prominent dominant colors and smooth color gradients, a K-means++ clustering algorithm optimized within the CIELAB color space was employed. The optimal number of color gradients was determined using contour compactness constraints combined with an automatic elbow method. The clustering results demonstrated a high degree of consistency with manually crafted gradient effects, effectively addressing the subjective bias inherent in traditional manual color identification.

In the edge detection phase, a multi-scale Canny operator system was innovatively developed by setting adaptive Gaussian filter parameters (σ) to eliminate noise caused by wood grain texture, alongside a dual-threshold scheme (low threshold 90/high threshold 120) that preserves fine details while suppressing color cluster spurious edges.

A Cyclic U-Net was adopted as the primary segmentation engine, showcasing significant advantages in processing TWNY Prints. While the conventional U-Net architecture excels in medical imaging, it faces limitations when handling the unique layered color structure of TWNY prints. The Cyclic U-Net introduces a progressive iterative mechanism: at each iteration, the original image combined with previously extracted layer masks is input to simulate the traditional artisan "one block, one color" overprinting technique.

This model embodies three core technical innovations. First, an adaptive stopping mechanism based on area proportion and inter-layer overlap ensures precise termination of the extraction process. Second, an optimization algorithm integrating cluster validation and morphological post-processing markedly enhances the accuracy of color layer separation, particularly excelling at managing characteristic woodblock print features such as ink diffusion and gradient areas. Third, through recursive learning, the model effectively learns critical features of TWNY Prints even with a relatively limited training dataset.

This deep learning–based layer separation method, together with the aforementioned K-means++ clustering and Canny edge detection techniques, forms a comprehensive technical pipeline that provides a systematic solution for the digital preservation and restoration of TWNY Prints.

4.2. Limitations

Several critical challenges remain to be addressed in the technical implementation of this study. First, unique traditional craft features of TWNY Prints, such as cloud-pattern gradients and ink diffusion effects, present significant challenges to existing color segmentation models. The K-means algorithm, when applied in the RGB color space, struggles to accurately capture the nonlinear tonal transitions created by the overlapping of mineral pigments, often resulting in abrupt color block segmentation or blurred boundaries in gradient regions. Meanwhile, although the Canny edge detector effectively identifies sharp edges, it lacks sufficient sensitivity to the subtle overlaps of brushstroke smudging and texture variations, frequently misclassifying these as noise and filtering them out.

While this study successfully demonstrates a novel framework combining traditional image processing with deep learning for the digital restoration of TWNY Prints, several important limitations remain, framing the context of the results and guiding future research.

First, the primary constraint of this research is the dataset. Currently, there is no publicly available, large-scale, and expertly annotated dataset for TWNY Prints. Our dataset was compiled from various online sources, which introduces variability in image resolution, color fidelity, and artifacting. This heterogeneity poses a significant challenge for

training deep learning models and for rigorous quantitative evaluation. The ground-truth masks used for the Cyclic U-Net were generated via a semi-automated process, which, while suitable for this preliminary exploration, does not possess the pixel-perfect accuracy required for calculating definitive performance metrics.

This dataset issue leads to a crucial distinction in how our two proposed methods can be evaluated:

For the K-means++ algorithm, performance is guided by established internal clustering validation metrics. In the methodology, the elbow method, silhouette score, and Calinski-Harabasz index are used to determine the optimal number of color layers in a data-driven manner, providing a quantifiable and reproducible basis for the color separation component of the framework.

For the Cyclic U-Net model, however, providing external evaluation metrics such as intersection over union (IoU) is challenging and could be misleading at this stage. The performance of such metrics is highly sensitive to the quality of the ground-truth annotations. Given the limitations of our current dataset, an IoU score would reflect the dataset's inconsistencies as much as the model's true capability. Furthermore, the Cyclic U-Net is framed not as a final, precise segmentation engine, but as a rapid, assistive tool intended to accelerate the workflow by generating a strong initial proposal for the layers. Its value lies in its efficiency and the capacity to learn the layering concept, a contribution that IoU alone cannot capture.

Therefore, the evaluation in this paper is intentionally bifurcated: a quantitative-driven assessment for the K-means++ process and a qualitative, proof-of-concept assessment for the Cyclic U-Net. This is considered the most intellectually honest approach given the current resource constraints.

Additionally, the selection of Canny edge detection thresholds in this study was performed manually, based on expert evaluation of traditional woodblock ink line characteristics. While this ensured that the extracted contours aligned closely with artisanal standards of clarity and continuity, it also introduces a degree of subjectivity. Future work will incorporate objective, data-driven parameter optimization methods to enhance reproducibility.

4.3. Future Directions

Our future work will be a direct response to these limitations. Our foremost priority is the development of a high-resolution, consistently captured, and expertly annotated benchmark dataset for TWNY Prints. This will not only enable the training of a more accurate Cyclic U-Net model but also allow for a fair and direct quantitative comparison between it and other methods using standard metrics like IoU. This will transition the model from an assistive tool to a verifiable, high-precision restoration solution.

Second, the future work includes the development of an "intelligent assistance plus manual refinement" collaborative design platform, featuring a visual color plate editing system. This platform will enable artisans to adjust parameters such as gradient tolerance and edge feathering in real time using a stylus, while the system simultaneously generates AI-optimized suggestions (e.g., color palette recommendations based on style transfer models). At the same time, it preserves the traditional craft control of manual fine-tuning for layer overlay opacity. This hybrid intelligent approach leverages convolutional neural networks for batch processing of repetitive color plate separation tasks, while utilizing human–computer interaction to pass down artisans' tacit knowledge. Ultimately, it forms a virtuous cycle of "algorithmic inference–experiential correction–data feedback," providing a digital preservation solution for intangible cultural heritage that balances technical rigor with artisanal craftsmanship.

Furthermore, in response to the need for broader applicability across heritage, museology, and art history communities, future research will incorporate an interdisciplinary discussion on integrating the proposed methodology into museum workflows, heritage center operations, and citizen participation initiatives. This will explore strategies for lowering technical entry barriers, enabling institutions with limited technological infrastructure to benefit from the framework, thereby enhancing its real-world adoption and cultural impact.

5. Conclusions

This study employs image processing and machine learning techniques to digitally restore the woodblock printing technology of TWNY Prints. Compared with traditional two-dimensional scanning and archiving methods, the developed algorithm intelligently identifies and extracts color blocks, simulating the multi-layer printing technique characteristic of TWNY prints. The resulting data can also be utilized for standardized production.

Canny edge detection effectively extracts the edge features of TWNY Prints. The optimized K-means++ algorithm for TWNY prints can automatically determine the optimal number of clusters and perform layered clustering. However, in practical applications, methods such as the elbow method, silhouette coefficient, and Calinski-Harabasz (CH) index must first be applied in the CIELAB color space to determine the optimal cluster number. Afterward, the data need to be converted back to the RGB color space for layer separation and reconstruction. Notably, the clustering performance of K-means/K-means++ in the CIELAB or HSV color spaces for actual color layering is suboptimal, indicating a need for further optimization of these algorithms tailored to specific color spaces.

Additionally, the introduction of a Cyclic U-Net model enables the learning of pixel-level color distinctions for layer segmentation of TWNY Prints, with the aim of reducing the complexity of the combined Canny and K-means++ segmentation steps. However, the current Cyclic U-Net model falls short of pixel-level segmentation accuracy. The observed results are hypothesized to stem from the U-Net architecture's limited capacity to fully capture features in color images, in combination with the sparse features in mask maps and the relatively low precision of dataset annotations, which primarily consist of manual fine-tuning assisted by algorithmic generation.

Based on the current experimental results, the introduction of a color attention mechanism into the Cyclic U-Net is proposed. Each color layer would be trained individually while ensuring that each model focuses on segmenting the most similar color layers (considering color similarity, area, and other factors), followed by transfer learning after training individual layers. This approach could mitigate cumulative errors and excessive loss, thereby improving overall segmentation accuracy.

The solutions provided herein offer a feasible pathway for the inheritance, protection, and innovative application of intangible cultural heritage. On a dynamic inheritance level, the system supports designers in rapidly generating cultural creative products that blend traditional charm with modern aesthetics, including constellation-themed trendy toys produced with screen printing techniques and smart New Year print folding screens incorporating dynamic light and shadow effects.

Furthermore, AR digital collectibles and intangible cultural heritage research applications have been leveraged to establish a three-dimensional inheritance ecosystem that integrates digital twin technology, virtual reality interaction, and scenario activation. This ecosystem not only preserves the craftsmanship genes of traditional New Year prints in their entirety but also integrates them into contemporary living spaces through new forms such as digital ink animation and immersive art installations. This approach pioneers a full-

chain protection model for intangible cultural heritage, encompassing "digital analysis–intelligent production–scenario regeneration."

The system can be extended to similar intangible cultural heritage projects, such as blue calico printing and Yangliuqing woodblock prints, providing replicable technical pathways and business models for the modern transformation of traditional crafts. Future work will focus on further optimizing algorithmic stability, expanding application scenarios, and promoting the deep integration of traditional culture and modern technology.

Author Contributions: Conceptualization, F.J.; methodology, F.J.; software, Y.W.; validation, Y.D., F.J. and Y.W.; formal analysis, Y.D.; investigation, Y.D.; resources, Y.D.; data curation, Y.D.; writing—original draft preparation, Y.D.; writing—review and editing, F.J.; visualization, Y.D.; supervision, F.J.; project administration, F.J.; funding acquisition, Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China National College Students' Innovation and Entrepreneurship Training Program (2024), grant number 202410298079Z.

Acknowledgments: The authors are thankful for the support from Nanjing Forestry University to conduct this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pan, S.; Dong, B.; Fu, R. Product Design for Yangliuqing Woodblock New Year Paintings Based on Eye Movement Experiment. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Virtually, 25–29 July 2021; pp. 375–383.
2. Rong, X. The Earliest Extant Example of Woodblock Printing: The Precept Certificate of the 29th Year of Kaiyuan (741 AD). *Pis' Mennye Pamiat. Vostoka* **2021**, *18*, 118–126.
3. Liu, F.; Zhang, M.; Zheng, B.; Cui, S.; Ma, W.; Liu, Z. Feature fusion via multi-target learning for ancient artwork captioning. *Inf. Fusion* **2023**, *97*, 101811. [CrossRef]
4. Qian, W.; Sharudin, S. Research on animation design of woodblock print characters based on audience aesthetic differences. *Insights Media J.* **2024**, *1*, 6–14. [CrossRef]
5. Zabulis, X.; Meghini, C.; Dubois, A.; Doulgeraki, P.; Partarakis, N.; Adami, I.; Karuzaki, E.; Carre, A.-L.; Patsiouras, N.; Kaplanidi, D. Digitisation of traditional craft processes. *J. Comput. Cult. Herit. (JOCCH)* **2022**, *15*, 1–24. [CrossRef]
6. Zhong, C.; Yu, X.; Xia, H.; Xie, R.; Xu, Q. Restoring intricate Miao embroidery patterns: A GAN-based U-Net with spatial-channel attention. *Vis. Comput.* **2025**, *41*, 7521–7533. [CrossRef]
7. Münster, S.; Maiwald, F.; di Lenardo, I.; Henriksson, J.; Isaac, A.; Graf, M.M.; Beck, C.; Oomen, J. Artificial intelligence for digital heritage innovation: Setting up a r&d agenda for europe. *Heritage* **2024**, *7*, 794–816. [CrossRef]
8. Skublewska-Paszkowska, M.; Milosz, M.; Powroznik, P.; Lukasik, E. 3D technologies for intangible cultural heritage preservation—Literature review for selected databases. *Herit. Sci.* **2022**, *10*, 3. [CrossRef]
9. Liu, E. Research on image recognition of intangible cultural heritage based on CNN and wireless network. *EURASIP J. Wirel. Commun. Netw.* **2020**, *2020*, 240. [CrossRef]
10. Hou, Y.; Kenderdine, S.; Picca, D.; Egloff, M.; Adamou, A. Digitizing intangible cultural heritage embodied: State of the art. *J. Comput. Cult. Herit.* **2022**, *15*, 1–20. [CrossRef]
11. Liu, Y.; Yin, J. Digital Cultural Design for Taohuawu Woodblock New Year Paintings Based on Cultural Translation. *Packag. Eng. Art Ed.* **2022**, *43*, 326–334.
12. Huang, Y.; Tan, G. Research on digital protection and development of China's Intangible Cultural Heritage. *J. Cent. China Norm. Univ. Humanit. Soc. Sci.* **2012**, *51*, 49–55.
13. Unesco. Available online: https://www.unesco.org/en (accessed on 10 June 2025).
14. Schotte, L. Sharing Stories of European Crafts and Artisanship. Available online: https://pro.europeana.eu/post/sharing-stories-of-european-crafts-and-artisanship (accessed on 18 May 2025).
15. CRAFTED: Enrich and Promote Traditional and Contemporary Crafts. Available online: https://pro.europeana.eu/project/crafted (accessed on 7 May 2025).
16. Albayrak, S. Color quantization by modified k-means algorithm. *J. Appl. Sci.* **2001**, *1*, 508–511. [CrossRef]

17. Saegusa, T.; Maruyama, T. Real-time segmentation of color images based on the K-means CLUSTERING on FPGA. In Proceedings of the 2007 International Conference on Field-Programmable Technology, Kitakyushu, Japan, 12–14 December 2007; pp. 329–332.
18. Hu, Y.C.; Su, B.H. Accelerated k-means clustering algorithm for colour image quantization. *Imaging Sci. J.* **2008**, *56*, 29–40. [CrossRef]
19. Hassan, R.; Ema, R.R.; Islam, T. Color image segmentation using automated K-means clustering with RGB and HSV color spaces. *Glob. J. Comput. Sci. Technol. F Graph. Vis.* **2017**, *17*, 25–33.
20. Yadav, S.; Biswas, M. Improved color-based K-mean algorithm for clustering of satellite image. In Proceedings of the 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2–3 February 2017; pp. 468–472.
21. Vardakas, G.; Likas, A. Global k-means++: An effective relaxation of the global k-means clustering algorithm. *Appl. Intell.* **2024**, *54*, 8876–8888. [CrossRef]
22. Mamat, A.R.; Mohamed, F.S.; Mohamed, M.A.; Rawi, N.M.; Awang, M.I. Silhouette index for determining optimal k-means clustering on images in different color models. *Int. J. Eng. Technol.* **2018**, *7*, 105–109. [CrossRef]
23. Tu, T.; Zhou, Z.; Xiao, P. Clustering color segmentation in multi-color space. In Proceedings of the 2018 2nd International Conference on Video and Image Processing, Hong Kong, China, 29–31 December 2018; pp. 118–122.
24. Basar, S.; Ali, M.; Ochoa-Ruiz, G.; Zareei, M.; Waheed, A.; Adnan, A.; Raza, M. Unsupervised color image segmentation: A case of RGB histogram based K-means clustering initialization. *PLoS ONE* **2020**, *15*, e0240015. [CrossRef]
25. Klein, C.; Speckman, T.; Medeiros, T.; Eells, D.; Basha, E. UAV-based automated labeling of training data for online water and land differentiation. In Proceedings of the 2018 International Symposium on Experimental Robotics, Buenos Aires, Argentina, 5–8 November 2018; pp. 106–116.
26. Moss, M.A.N.; Hughes, D.D.; Crawford, I.; Gallagher, M.W.; Flynn, M.J.; Topping, D.O. Comparative Analysis of Traditional and Advanced Clustering Techniques in Bioaerosol Data: Evaluating the Efficacy of K-Means, HCA, and GenieClust with and without Autoencoder Integration. *Atmosphere* **2023**, *14*, 1416. [CrossRef]
27. Feng, H.; Hu, Q.; Zhao, P.; Zheng, D.; Ai, M.; Chen, S.; Hu, X. Automatic generation of Chinese mural line drawings via enhanced edge detection and CycleGAN-based denoising. *npj Herit. Sci.* **2025**, *13*, 345. [CrossRef]
28. Li, D.; Huang, Y.; Inoue, T.; Inoue, K.; Zhang, Z. Image processing in the conservation of historic urban areas: The case of Dujiangyan, China. *Built Herit.* **2025**, *9*, 7. [CrossRef]
29. Yao, G.; Sun, A. Multi-guided-based image matting via boundary detection. *Comput. Vis. Image Underst.* **2024**, *243*, 103998. [CrossRef]
30. Zhou, X.; Chen, Y.; Lin, Z.; Su, Z.; Chai, Z.; Wang, R.; Hu, C. Non-spherical Janus microparticles localization using equivalent geometric center and image processing. *Opt. Commun.* **2024**, *560*, 130494. [CrossRef]
31. Dixit, K.; Gupta, P.; Kamle, S.; Sinha, N. Structural analysis of porous bioactive glass scaffolds using micro-computed tomographic images. *J. Mater. Sci.* **2020**, *55*, 12705–12724. [CrossRef]
32. Yuan, Y.; Chen, W.; Tang, J.; Yang, J. A High-Precision Edge Detection Algorithm Based on Improved Sobel Operator-Assisted HED. In Proceedings of the 2024 IEEE 14th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Copenhagen, Denmark, 16–19 July 2024; pp. 83–88.
33. Zhou, G.; Guo, S.; Chen, Z. Fpga-based improved sobel operator edge detection. *Front. Comput. Intell. Syst.* **2023**, *5*, 6–11. [CrossRef]
34. Zhang, C.; Zhang, N.; Yu, W.; Hu, S.; Wang, X.; Liang, H. Improved Canny-based algorithm for image edge detection. In Proceedings of the 2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Nanchang, China, 28–30 May 2021; pp. 678–683.
35. Li, Y.; Liu, B. Improved edge detection algorithm for canny operator. In Proceedings of the 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 17–19 June 2022; pp. 1–5.
36. Wang, S.; Ma, K.; Wu, G. Edge detection of noisy images based on improved canny and morphology. In Proceedings of the 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 29–31 October 2021; pp. 247–251.
37. Liu, R.; Mao, J. Research on improved Canny edge detection algorithm. *Proc. MATEC Web Conf.* **2018**, *232*, 03053. [CrossRef]
38. Kong, J.; Hou, J.; Liu, T.; Jiang, M. Adaptive image edge detection model using improved canny algorithm. In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 1–3 November 2018; pp. 539–545.
39. Xu, Z.; Ji, X.; Wang, M.; Sun, X. Edge detection algorithm of medical image based on Canny operator. *Proc. J. Phys. Conf. Ser.* **2021**, *1955*, 012080. [CrossRef]
40. Zhang, B.; Wang, F.; Li, G.; Zhang, C.; Zhang, C. A multispectral image edge detection algorithm based on improved canny operator. In *Communications, Signal Processing, and Systems: Proceedings of the 8th International Conference on Communications, Signal Processing, and Systems, Urumqi, China, 20–22 July 2019*; Springer: Singapore, 2020; pp. 1298–1307.

41. Taghanaki, S.A.; Abhishek, K.; Cohen, J.P.; Cohen-Adad, J.; Hamarneh, G. Deep semantic segmentation of natural and medical images: A review. *Artif. Intell. Rev.* **2021**, *54*, 137–178. [CrossRef]
42. Hussain, T.; Shouno, H. MAGRes-UNet: Improved medical image segmentation through a deep learning paradigm of multi-attention gated residual U-Net. *IEEE Access* **2024**, *12*, 40290–40310. [CrossRef]
43. Feng, J. *China Woodblock New Year Pictures Integration: Taohuawu*; Zhonghua Book Company: Beijing, China, 2011.