IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 February 2023, accepted 1 March 2023, date of publication 6 March 2023, date of current version 22 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3252892

RESEARCH ARTICLE

Structure-Texture Consistent Painting Completion for Artworks

IRAWATI NURMALA SARI AND WEIWEI DU, (Member, IEEE)
Department of Information Science, Kyoto Institute of Technology, Kyoto 606-8585, Japan
Corresponding author: Weiwei Du (duweiwei@kit.ac.jp)

ABSTRACT Image completion techniques have made rapid and impressive progress due to advancements in deep learning and traditional patch-based approaches. The surrounding regions of a hole played a crucial role in repairing missing areas during the restoration process. However, large holes could result in suboptimal restoration outcomes due to complex textures causing significant changes in color gradations. As a result, they led to errors such as color discrepancies, blurriness, artifacts, and unnatural colors. Additionally, recent image completion approaches focused mainly on scenery and face images with fewer textures. Given these observations, we present a structure-texture consistent completion approach for filling large holes with detailed textures. Our method focuses on improving image completion in the context of artworks, which are expressions of creativity and often have more diverse structures and textures from applying paint to a surface using brush strokes. To handle the unique challenges posed by artwork, we use Cohesive Laplacian Fusion that segments non-homogeneous areas based on structure diffusion and then applies texture synthesis to complete the remaining texture of the missing segmented area. This technique involves detecting changes in base structures and textures using multiple matched patches to achieve more consistent results. The experimental results show that our proposed method is competitive and outperforms state-of-the-art methods in missing regions and color gradations of art paintings.

INDEX TERMS Painting completion, structure-texture consistent completion, artworks, image inpainting.

I. INTRODUCTION

Image completion, also known as image inpainting, is a sophisticated algorithm used to recover missing or hole areas of an image by employing known pixels to restore the image to its original form with a visually appealing technique [1]. In recent years, image completion has seen rapid development and has found many practical applications, particularly in image editing, context-based image transmission, and Image-Based Rendering (IBR) [2]. Despite its practical usefulness, image completion still has limitations, such as producing artifacts, color discrepancies, and blurriness that require further investigation.

The majority of existing image completion methods have focused on restoring damaged photographs captured by cameras, such as human faces and landscapes. These images often feature complex structures, but lack in detailed textures. On the other hand, art painting presents a unique challenge due to the intricate brush strokes that create diverse structures and textures. It holds a significant place in museums and can captivate viewers with its artistic sense and atmosphere, similar to photography. Museum curators have preserved art paintings for hundreds of years with proper protection, but over time, particularly for ancient paintings, the paintings can become worn and damaged. In the past, special measures were taken to fill the damaged or hole parts of a painting by attempting to match similar colors to the surrounding areas. However, these efforts did not always result in optimal results. Additionally, graphic editing software like Adobe Photoshop and Inkscapes offer convenient options, but they did not be able to repair missing parts of a painting with the same skill as the artist feelings.

To address the issue of repairing damaged art paintings, several researchers have proposed alternative solutions using

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

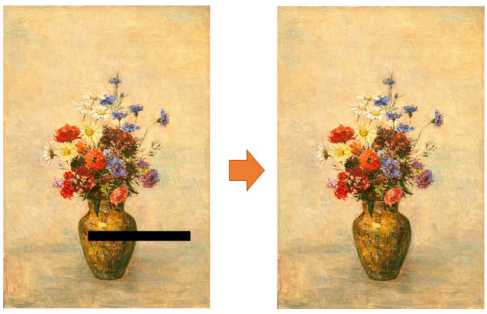

FIGURE 1. (a) Input: An art painting with a missing region, (b) Output: A restored art painting.

computer-based methods that were faster and more effective. These approaches had the potential to restore torn art paintings to their original state. However, limited available art painting datasets and copyright problems have resulted in limited research in this area, particularly for traditional techniques [3], [4]. Nevertheless, traditional techniques are still a viable option for art painting completion because they only require a single damaged painting for restoration, unlike deep learning models, which require at least 1000 images for training to achieve accurate results.

Traditional image completion methods can be categorized into two groups: pixel-based and exemplar-based approaches. Pixel-based techniques [5] have produced impressive results in detail, but only for small missing regions such as blind spots or lines. Exemplar-based approaches, on the other hand, identify compatible and similar known patches in the surrounding hole area that can be replicated and applied to the receptive fields [6], [7], [8], [9], [10], [11], [12], [13], [14]. However, there is no guarantee that a accurate patch match will always exist in the known space, and errors such as color discrepancies may occur in the restored region if the appropriate patch is not found.

Recently, deep learning has advanced the field of image completion with various innovative techniques [15], [16], [17]. For instance, the LaMa algorithm [18] used a feed-forward inpainting network and Fast Fourier Convolutions (FFC) to reconstruct complex regular structures in large missing regions. Similarly, the SC-FEGAN method [19] employed an end-to-end convolutional network with style loss, resulting in realistic outcomes, even in large areas. However, these methods could be time-consuming due to the extensive training required and the need for a high-performance computer. Additionally, long-term memory loss could lead to blurred and unnatural restoration results.

In addition, most traditional and deep learning approaches, like the techniques mentioned above, have focused only on natural or authentic images rather than artificial ones, such as art paintings. The complexity of brush strokes and color gradation in art paintings presents unique features that represent artistic languages. To address these challenges, we utilize a traditional method based on an exemplar-based approach. Our method proposes a structure-texture consistent completion approach for art paintings, prioritizing "structure first, texture next". Inspired by the process artists use to create beautiful paintings, starting with a sketch to define the structural plan, our method applies structure diffusion to identify distinct regions within holes. Then, to fill in the remaining area intensities, we use texture synthesis to create a surface quality with a sense of depth. In summary, there are two points of our contributions:

- We present a pioneering approach to repair the damages of art paintings that have torn and worn-out areas with both regular and irregular holes. Our approach yields satisfying results with only a single input image.
- We introduce a state-of-the-art algorithm called Cohesive Laplacian Fusion, which generates consistent structure and texture while reconstructing the missing region by considering human sensibility in performing fine details.

The rest of this paper is organized as follows. After the description of related works (Section II), we present our proposed method (Section III). Then, we demonstrate the experimental results (Section IV) and discuss implementation (Section V). Eventually, we explain the conclusion and future works (Section VI).

II. RELATED WORK

In recent years, image completion has increasingly relied on sophisticated techniques such as deep learning and traditional methods. However, both approaches faced similar challenges and required strong generalization ability to effectively restore damaged images. Researchers have explored image repair using both deep neural networks [15], [16], [17], [20], [21], [22], [23], [24] and traditional techniques [5], [6], [7], [8], [25], [26], [27], [28].

Adversarial learning and deep neural networks determined semantic preceding and important hidden information in an end-to-end model. Liu et al. [17] employed a partial convolutional layer in composing a mask and re-normalized it by a mask-update process for only valid pixels. This method produced blurriness while attempting to repair the large missing regions and failed for simple structured images. Pathak et al. [29] proposed context encoder-decoder networks that worked independently. The encoding process recovered the nearest neighbor patches, including segmentation and object detection. Moreover, the decoder was in charge of filling realistic pixel contexts using up-convolutions and non-linearities from the information of encoders. However, color discrepancy and blurriness often appeared in the receptive fields because of congestion in fully connected layers. Several researchers [30], [31], [32], [33] proposed image enhancement to reduce the blurriness of images by many learning strategies, loss functions, and network structures. These methods only improved the perception of the environment with poor illuminations, which was the unsolved open issue of image inpainting due to scattering

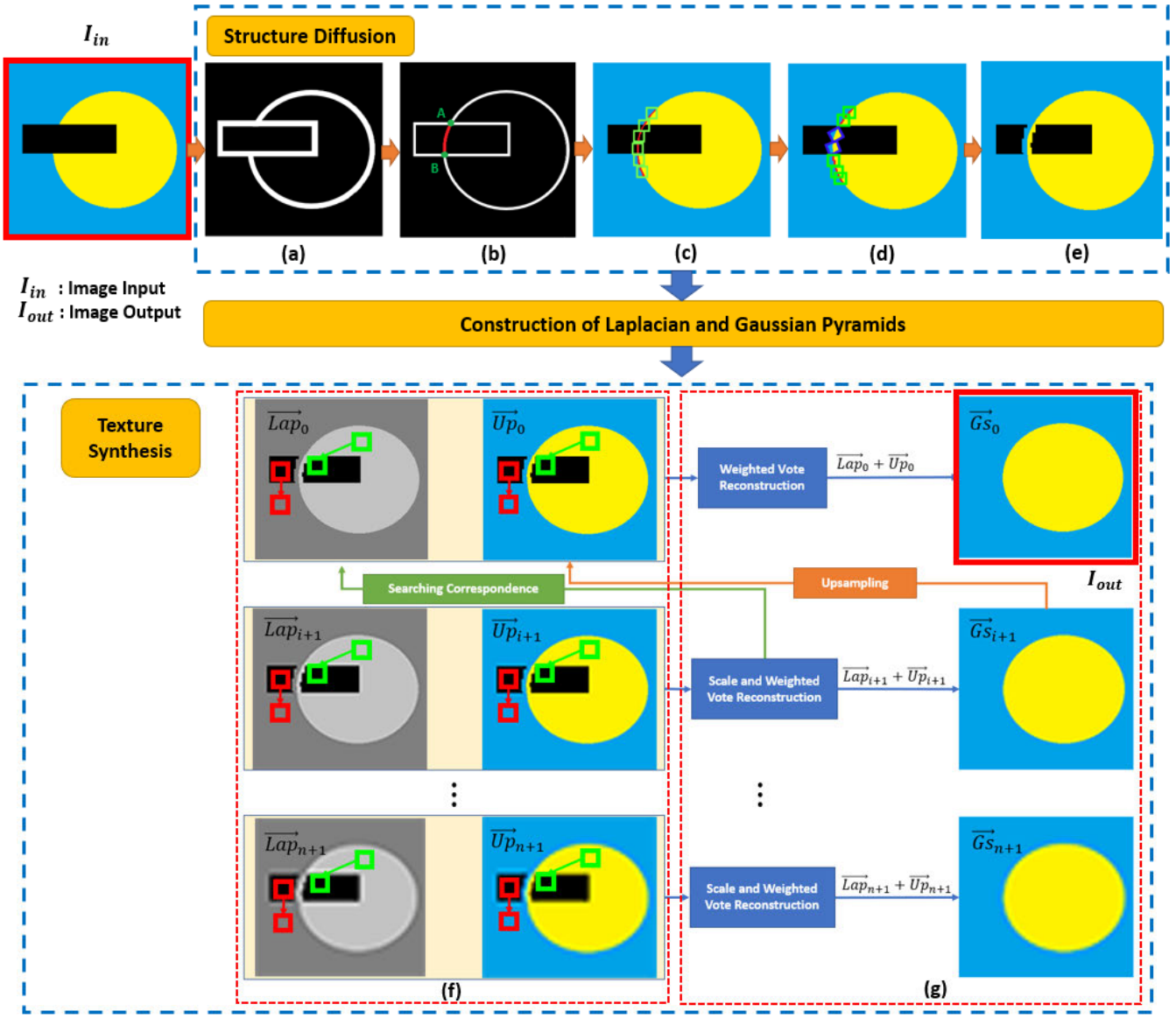

FIGURE 2. Cohesive Laplacian Fusion. (a) Edge detection, (b) Skeletonization and connecting gap of endpoints (red curve), (c) Patch detection along a guided curve (green patches), (d) Determining dynamic patches (blue patches), and (e) Patch-based propagation.

and matching patch errors of restored regions. State-of-the-art stable diffusion method [34] attempted to synthesize image inpainting using cross-attention layers of diffusion models into an adjustable generator to the impressive convolutional appearance. This method relied on the text prompt of their extensive general dataset, which users inputted depending on their desires instead of repairing as the original one. A recent learning method called Generative Adversarial Networks (GAN) was popular and has made significant progress in generating plausible completion of complex images. For instance, Zheng et al. [35] presented CM-GAN built by an image encoder using fourier convolution blocks and a global-spatial modulation-based decoder in constructing the comprehensive and details structures. Masaoka et al. [36] introduced missing region restoration based on perspective views by edge-enhanced GAN models with vanishing points detection. Nazer et al. [37] presented two-phase adversarial models composed of an edge generator and completion networks. Liu et al. [38] employed Probabilistic Diverse GAN (PD-GAN) for developing the context information and deep features of random noise in multiple scales. However, despite all of these deep learning methods generating powerful models, the training time was significantly increased due to maintaining the model's consistency at every stage of the process. In addition, the long-range dependence of all stages caused a loss of ability hole restoration for complicated structures and textures, such as in art paintings.

R_1 $(x-1, y+1)$	R_2 $(x, y+1)$	R_3 $(x+1, y+1)$
R_4 $(x-1, y)$	R_5 (x, y)	R_6 $(x+1, y)$
R_7 $(x-1, y-1)$	R_8 $(x, y-1)$	R_9 $(x+1, y-1)$

FIGURE 3. Skeletonization. Description of 9 pixels in a 3 × 3 window. There are two deleting sub iterations of a center point R_5, which are in R_1 and R_3, then R_7 and R_9. (x, y) is a coordinate.

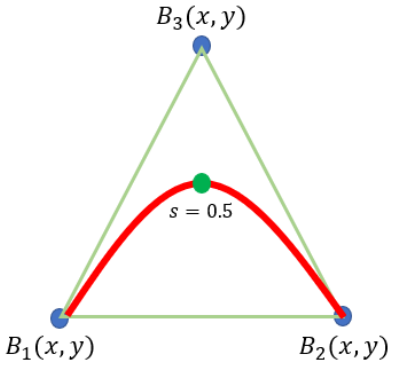

FIGURE 4. 3-point Bezier curve (red curve). A curve is always inside the convex hull of 3 points that are controlled by time (s). Time equals 0.5 representing 50% of the segment length.

In response to the weakness of deep learning approaches, the traditional technique was a faster and more straightforward method of repairing the holes by a single damaged image. Criminisi et al. [9] merged the advantages of structure and texture synthesis and relied on exemplar-based techniques to determine the fill order of the target area. Although the priority of filling patches was easy to define, this method still produced unreasonable results and demanded the handling of curved structures. Currently, researchers improved Criminisi method [9] to achieve more competence and robustness in completing damaged images [10], [11]. These approaches only generated good performances with the condition that they had simple curves and only if there were sufficient samples of structure and texture. Sun et al. [12] attempted to recover the large-scale missing region using two stages. The first step was to segment a large area into smaller parts by analyzing patches along user-specified curves in the unknown region, adopting selected patches around curves in the known area. The second step, the Criminisi method [9], was applied to accomplish the remaining smaller areas. Irawati et al. [8] proposed interactive structure propagation to segment the non-homogenous region of a large hole into small parts. Then, the remaining areas will be filled by Fast Marching Method (FMM). Methods [12] and [8] generated impressive results but required user guidance to create a pre-processing curve inside the missing region. Urano et al. [39] attempted to upgrade structure propagation automatically by adopting auxiliary line construction containing penetrator and instructor lines separately to synthesize the intersection inside holes. Irawati et al. [6] and Horikawa et al. [7] also struggled in automatic structure propagation by applying orthogonal viewpoints and clustered structure guidance of different perspective planes. Due to the implemented vanishing points of planar regions, they solved mainly straight lines, which did not support art painting based on brush strokes. Lee et al. [13] presented Laplacian pyramid to analyze and find the correspondence textures by applying a patch-based synthesis that considered edge structures. This technique did not handle depth ambiguities in diverse locations between foreground and background areas, which produced unsatisfactory results in art painting completion of various patterns and object designs.

Many researchers employed deep learning and traditional methods to develop art painting completion. Chen et al. [40] applied deep learning and manual sliding windows to find matched patches of damaged paintings. Due to the art painting containing detailed textures and a unique artistic sense of each stroke with different styles, this method failed to produce a robust convolution network model, resulting in blurred restored regions. In addition, inspired by the physical restoration of damaged holes in paintings, Wang et al. [41] presented user line drawings such as sketches to guide the inpainting algorithm as pre-processing. This technique was time-consuming and only achieved visually satisfactory results in simple paintings instead of abstract or landscape paintings with more complex structures and brushstroke patterns. Therefore, depending on observations, our paper proposed an approach that consistently captures structural and textural completion in a one-stage algorithm, as shown in Figure 1. Our method handles various paintings with different structures and textures in regular and irregular missing areas, such as frequent painting breakdowns.

III. METHODOLOGY

As mentioned, a large missing region was sensitive to blurriness and artifacts during restoration. Deep learning techniques attempted to train models of numerous datasets with more efforts. Despite being considered trusted and reliable algorithms, there need to be more satisfactory ways to obtain the desired results, such as disappearing color discrepancies and decreasing time-consuming.

According to investigations, although some researchers have disregarded the traditional inpainting method, it can be considered as an alternative way to enhance the restoration technique by considering a fast and uncomplicated mechanism like the proposed approach in this paper. Our method offers the restoration of large missing regions and arbitrary line holes in damaged art paintings. The Cohesive Laplacian Fusion consistently repairs holes by a one-stage structural and textural completion, generating the original paintings in accordance with the artist's intent. In this section, we describe the concept of our approach: (1) Cohesive Laplacian Fusion, and (2) A detailed explanation of our proposed algorithm.

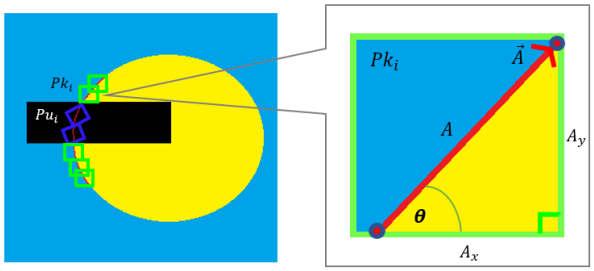

FIGURE 5. Direction of patch.

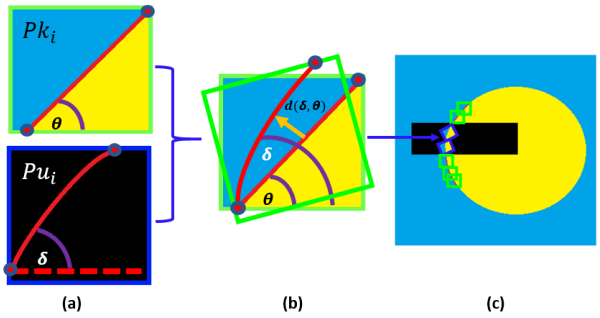

FIGURE 6. Patch-based propagation. (a) The closer direction of patches, (b) Adjusting curves using rotation, and (c) Patch stamping.

A. COHESIVE LAPLACIAN FUSION

A large missing region is a challenging case in image completion. Many recent kinds of research attempted to solve this issue [6], [8], [9], [15], [37]. However, blurriness and color discrepancy appeared as burdensome. In addition, arbitrary line holes in the boundary of objects raise the extensive impact of generating unsatisfactory results due to significant color divergences, especially in complex art paintings. Based on this issue, our approach proposes proficient problem-solving by consistently collaborating on structure and texture completion.

Figure 2 illustrates a mechanism of our collaboration system called Cohesive Laplacian Fusion. First, in the structure diffusion, we adopt Holistically-Nested Edge Detection (HED) [42] as pre-processing to achieve boundary edge. This method develops an end-to-end performance that resolves the ambiguity of the human ability to distinguish object boundary detection in natural images by learning well-to-do types of hierarchical features at different scales. Due to thick generated edges having uncertainty values, the thinning process is required by implementing skeletonization [43], applying a fast parallel algorithm using two sub-iterations deleting systems, as shown in Figure 3. In solving the gap caused by an unknown area, 3 points of Bezier [44] are defined to maintain the curve position fully, as shown in Figure 4. We adopt De Castelijau's algorithm [45] in building the 3-point Bezier curve (red curve) of $B_1(x, y)$, $B_2(x, y)$, $B_3(x, y)$:

$$\begin{aligned} C(x) &= (1-s)^2 B_1(x) + 2(1-s)sB_2(x) + s^2 B_3(x) \\ C(y) &= (1-s)^2 B_1(y) + 2(1-s)sB_2(y) + s^2 B_3(y) \end{aligned} \tag{1}$$

where s is a time that moves from 0 to 1, $C(x, y)$ represents a coordinate on a 3-point Bezier curve (red curve), $B_1(x, y)$ and $B_2(x, y)$ are points defining the boundary of a hole, and $B_3(x, y)$ is a supporting point.

After the connected curve is built as a structure sketch, we apply patch-based propagation in spreading identical patches based on dynamic patch correspondence. As shown in Figure 5, given Pu_i and Pk_i are patches inside the unknown and known region, respectively, which $i = 1, 2, 3, \ldots, n$. Compute the degree of curves Θ:

$$tan(\Theta) = \frac{A_y}{A_x} \tag{2}$$

$$\Theta = tan^{-1}(\frac{A_y}{A_x}) \tag{3}$$

Since structure continuity is critical within unknown regions, dynamic patching prioritizes correlations of similar or closer direction patches between known and unknown regions. For instance, Figure 6 presents adjusting curves of Pu_i and Pk_i by rotating the corresponding patch based on the distance of degrees:

$$d(\delta, \theta) = \delta - \theta \tag{4}$$

where δ is a degree of Pu_i that considers the designated curve is close to the straight line and θ is a degree of Pk_i. Then, the adjusted patch of the known region is stamped to the original central point location along the curve in an unknown region.

Furthermore, due to structure diffusion segmenting a non-homogeneous missing region of the object boundary into smaller homogeneous regions, texture synthesis is employed to generate a quality surface with similar intensities by estimating Laplacian of a Gaussian (LoG). According to [13], the computation of LoG is expensive due to its complex function. However, the Laplacian of a Gaussian works similarly to the Difference of Gaussian (DoG), which is used to consider edge structures in the convolution process. Therefore, we choose to apply the Laplacian of a Gaussian pyramid at different levels to preserve consistent performance in edge awareness and base textures with more straightforward computations.

In the texture synthesis, first, we need to build a Gaussian pyramid Gs to generate the pixel intensities at each level for analyzing the base structure. Then, this construction will establish the Laplacian pyramid Lap for enhanced edge awareness:

$$\begin{aligned} Gs_{i+1} &= downsample(Gs_i) \\ Up_i &= upsample(Gs_{i+1}) \\ Lap_i &= Gs_i - Up_i \end{aligned} \tag{5}$$

where Gs_{i+1} is a downsampled Gaussian of Gs_i, Up_i is an upsampled Gaussian of Gs_{i+1}, and Lap_i is transformed to Laplacian image from Gs_i and Up_i based on [13]. The total number of levels is $n + 1$.

After constructing Laplacian and Gaussian pyramids, we locate the matching patches at each level by enhancing a nearest neighbor search algorithm that approximates the most similar areas between the source S and target T based on the minimum normalized distance:

$$SP(S,T) = \sum_{s\in T} min_{t\in S}\left[\frac{1}{N_{Up_i^{s,t}}}(\alpha D(Up_i^s, Up_i^t) + \frac{1}{N_{Lap_i^{s,t}}}(\beta D(Lap_i^s, Lap_i^t))\right] \quad (6)$$

where i is the current level of pyramid, s and t are pixel locations of S and T, Up_i^s and Lap_i^s indicate patch of pyramid Up and Lap at level i and location s, Up_i^t and Lap_i^t exhibit patch of pyramid Up and Lap at level i and location t, $N_{Up_i^{s,t}}$ and $N_{Lap_i^{s,t}}$ are number of pixels in Up and Lap, respectively, $\alpha+\beta = 1$ stands for the frequency ratio of Up and Lap, and D is a distance metric with Sum of Square Distances (SSD).

Determining the optimal pixel value to fill in the missing region is an essential factor that affects the performance of image completion. Lee et al. [13] proposed a weighted blending of scales between the upsampled Gaussian Up_i and Laplacian image Lap_i to establish the optimal similarity between the target and source areas. However, this method applied voting system that had a sensitivity to detect similarity. It is assumed that all pixels located outside the target area at the current level will be considered. As a result, the current target region might not be able to be completely propagated from the nearest neighbor of the most recently restored area, causing a color discrepancy that did not blend smoothly with adjacent patches. To address this issue, our method improves the voting similarity function by covering the possibilities of nearest neighbor pixels that overlap in color even in the different levels using the enhanced weighted vote w_q:

$$w_q = \Psi(p,q,l)\Lambda(q), \qquad \Psi(p,q,l) = e^{-\frac{\frac{1}{2}d(p,q,l)}{(\mu+Z\sigma)^2}} \quad (7)$$

$$c_q = \frac{\sum_{\tilde{q}\epsilon Q} w_{\tilde{q}}\overline{N}_{\tilde{q}}(q-\tilde{q})}{\sum_{\tilde{q}\epsilon Q} w_{\tilde{q}}} \quad (8)$$

where $\Psi(p,q,l)$ and $d(p,q,l)$ describe the similarity and distance between source pixel p and target pixel q in level l, respectively, μ is mean of correspondence pixels, Z denotes qnorm of probability (75^{th} percentile) for normal distribution, σ is a standard deviation, $\Lambda(q)$ indicates a confidence weight at target pixel q, $\tilde{q}$ is overlapping colors from its nearest neighbor field $\overline{N}_{\tilde{q}}$, and Q is the total number of $\tilde{q}$.

B. OUR PROPOSED ALGORITHM

In this section, we discuss the details of our proposed algorithm as shown in Figure 2 and Algorithm 1. The input is a single art painting, I_{in}, with damaged regions such as large holes, scratches, or fold marks. To simplify the filling process, we utilize structure diffusion to create a structure sketch plan inside the missing region, which serves as guidance for segmenting non-homogeneous regions. First, we build an automatic environment by using edge detection [42] instead of relying on user-specified edges (Figure 2(a)). However, the detected edges are often too wide, so we use skeletonization to address this issue by accurately determining the location of single-pixel values. Then, we use the Bezier curve method, relying on three points around a missing region, to create a connected curve along the gap between the endpoints in the boundary of the hole (Figure 2(b)). After that, our method dynamically stamps identical patches from known to unknown regions along the guidance curve (red curve) while considering the consistency of related patches (Figure 2(c) and 2(d)).

Since the non-homogeneous missing region is segmented into smaller homogeneous regions using a prior technique (Figure 2(e)), we implement a state-of-the-art texture synthesis approach to complete the remaining regions convincingly. This method builds Laplacian and Gaussian pyramids from the coarsest to finest images to establish edge awareness and base textures, respectively. Then, to identify the patch similarity at each level, we apply a nearest neighbor searching algorithm as shown in Equation 6 (Figure 2(f)). Finally, we overlap multiple matched patches to estimate the optimal single-pixel value (Figure 2(g)) using Equation 7 to compute a weighted vote and relate it to the voting similarity function in Equation 8.

Algorithm 1 Cohesive Laplacian Fusion

Data: A single art painting with missing region I_{in}.
Result: A restored art painting I_{out}.
Structure Diffusion
Detect edges of art painting (Figure 2(a));
Apply skeletonization to thin the edges (Figure 2(b));
Connect the gap between the endpoints around a missing region (Figure 2(b));
Dynamic patch **do**
  Stamp identical patches along the curve by using patch-based propagation (Figure 2(c), 2(d), and 2(e));
while *unknownpatch* == *knownpatch*;
Laplacian and Gaussian Pyramid Construction
Texture Synthesis
Patch Matching **do**
  Find the matched patches based on a nearest neighbor search algorithm by using Equation 6 (Figure 2(f));
  Estimate the optimal pixel value by overlapping colors in each selected patch through a weighted vote in Equation 7 and the voting similarity function in Equation 8 (Figure 2(g));
while *missingpatch*! = *empty*;

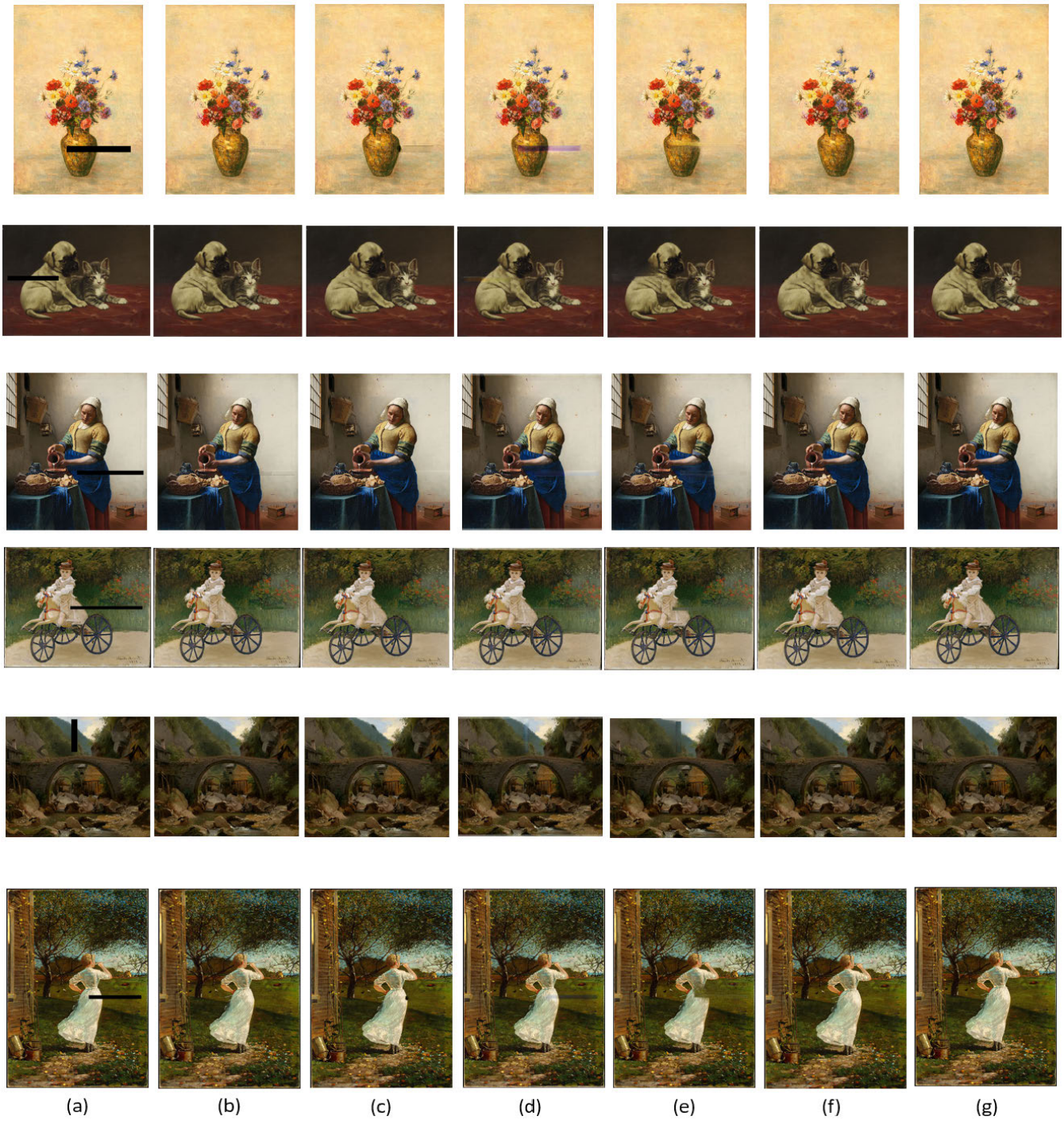

FIGURE 7. Art painting completion of large block missing regions (regular). Labeling images in order from top to bottom: Image 1 (Vase), Image 2 (Pets), Image 3 (Woman), Image 4 (Child), Image 5 (Scenery), and Image 6 (Girl). (a) Input, (b) Manual Structure Propagation (MSP) [12], (c) Laplacian Pyramid (LP) [13], (d) EdgeConnect GAN (EC-GAN) [37], (e) SketchEdit (SE) [46], (f) Ours, and (g) Ground Truth.

IV. EXPERIMENTAL RESULTS

According to observations made at art museums, various issues with damaged paintings occurred, such as tears and stains. To resolve these problems, our proposed approach has been successful in various scenarios, as outlined in this section. The first scenario deals with large, regular block holes in the painting that significantly impact the restored texture. The second scenario addresses arbitrary, irregular line holes that often occur randomly and result from long-term storage in museums. Our evaluations are based on the available online painting dataset from an art gallery in diverse museums, located at https://useum.org/.

FIGURE 8. Art painting completion of arbitrary missing regions (irregular). Labeling images in order from top to bottom: Image 1 (Vase), Image 2 (Pets), Image 3 (Woman), Image 4 (Child), Image 5 (Scenery), and Image 6 (Girl). (a) Input, (b) Manual Structure Propagation (MSP) [12], (c) Laplacian Pyramid (LP) [13], (d) EdgeConnect GAN (EC-GAN) [37], (e) SketchEdit (SE) [46], (f) Ours, and (g) Ground Truth.

Each painting in the dataset is roughly 600 × 800 pixels in size. To compare our approach with previous methods, we include traditional techniques like Manual Structure Propagation (MSP) [12] and Laplacian Pyramid (LP) [13], as well as deep learning methods like EdgeConnect GAN (EC-GAN) [37] and SketchEdit (SE) [46].

A. QUALITATIVE COMPARISON

To validate our results with human visual perception, we employ qualitative comparison to evaluate the details of structures and textures between previous and our works. Figures 7 and 8 show that the MSP method [12] generates the boundary edge of objects smoothly because it

segments a large unknown region into smaller areas using a user-specified curve as guidance. However, the texture area has color discrepancies, especially in Images 1, 2, 3, and 4 (Figure 7). Moreover, even when recovering the unknown irregular area (Figure 8), all results of MSP [12] also show line marks universally. It occurred because of insufficient patch-based synthesis after structure propagation was built, which entirely relied on image isophotes that affect filling priority.

LP approach [13] proposed the building of Laplacian pyramid to repair the unknown area by examining consistent patch similarity in each level and solving by a weighted vote of scales. However, due to only considering the target area of the current level to fill the missing region, this method produces color inconsistencies and black marks, as seen in all images (Figure 7). In addition, it is unable to fill arbitrary holes, such as in Images 1, 2, 3, and 6 (Figure 8), as there is less information available for finding identical patches at each level.

In contrast, the deep learning method EC-GAN [37] presented the prediction of a model based on edge-map guidance to recover missing regions. Unfortunately, bright colors of miscalculation and opaqueness are visible in all results of Figure 7 and 8 due to long-term memorization and an inadequate model obtained from a less available dataset of painting images.

Another compared deep learning method is SketchEdit (SE) [46]. This method conducted GAN models using an interactive image inpainting based on conditioned sketches of users and a structure-style vector. However, a model often fails when training warping sketches and delivering a significant blurriness impact to art paintings with many structural lines, as shown in all results images Figure 7 and 8.

As shown in Figures 7 and 8, our approach generates results that are closer to the Ground Truth (GT). Although there is some slight blurring, such as in Images 2 and 3 (Figure 7), the object boundaries are smoother and there are no artifact effects due to our approach segments large non-homogeneous blocks into smaller homogeneous missing regions by utilizing previously obtained curves as a guide. Furthermore, our proposed method performs better in irregular cases, as demonstrated in Figure 8. Despite irregular holes being more dispersed and smaller in comparison to regular holes shown in Figure 7, our method still yields improved results. It occurs because the normalized distance metric and enhanced weighted vote have a greater impact when filling homogeneous holes.

B. QUANTITATIVE COMPARISON

The most commonly used perceptual metrics are the Structural Similarity Index Measure (SSIM) [47] and Peak Signal to Noise Ratio (PSNR) [48] due to their simple application. However, our work evaluates results using the Learned Perceptual Image Patch Similarity (LPIPS) [49] metric, which applies a deep feature network based on human perceptual similarity judgments. This metric employs the learning of linear weights (perceptual calibration), tuning configuration, and Gaussian weights:

$$d(x, x_0) = \sum_l \frac{1}{H^l W^l} \sum_{h,w} \left\| wt^l \odot (\hat{y}^l_{h,w} - \hat{y}^l_{0h,w}) \right\|_2^2 \quad (9)$$

where $d(x, x_0)$ denotes a distance between ground truth x and restored images x_0. $\hat{y}^l, \hat{y}^l_0 \epsilon \mathbb{R}^{H^l \times W^l \times C^l}$ are a feature stack from l layers and unit-normalize in the channel dimension, respectively. H, W, and C designate the number of height h, weight w, and channel dimension c of layers, respectively. $wt^l \epsilon \mathbb{R}^{C^l}$ is an activation channel-wise.

In addition, our evaluation also relies on Deep Image Structure and Texture Similarity (DISTS) [50]. Based on a Conventional Neural Network (CNN), DISTS integrated spatial texture averages and feature structure maps to synthesize diverse texture patterns. The model combined quality assessments between convolution layers of texture using global means $l\left(\tilde{x}^i_j, \tilde{y}^i_j\right)$ and structure using global correlations $s\left(\tilde{x}^i_j, \tilde{y}^i_j\right)$ by computing the weighted sum of different convolution layers $D(x, y; \alpha, \beta)$:

$$l(\tilde{x}^i_j, \tilde{y}^i_j) = \frac{2\mu_{\tilde{x}^i_j}\mu_{\tilde{y}^i_j} + c_1}{\left(\mu_{\tilde{x}^i_j}\right)^2 + \left(\mu_{\tilde{y}^i_j}\right)^2 + c_1} \quad (10)$$

$$s(\tilde{x}^i_j, \tilde{y}^i_j) = \frac{2\sigma_{\tilde{x}^i_j\tilde{y}^i_j} + c_2}{\left(\sigma_{\tilde{x}^i_j}\right)^2 + \left(\sigma_{\tilde{y}^i_j}\right)^2 + c_2} \quad (11)$$

$$D(x, y; \alpha, \beta) = 1 - \sum_{i=0}^{m} \sum_{j=1}^{n_i} \left(\alpha_{ij} l\left(\tilde{x}^i_j, \tilde{y}^i_j\right)\right) + \left(\beta_{ij} s\left(\tilde{x}^i_j, \tilde{y}^i_j\right)\right) \quad (12)$$

where i and j are convolution layer of ground truth x and restored images y, respectively. $\tilde{x}^i_j$ and $\tilde{y}^i_j$ represent the convolution channel of x and y. m and n_i are number of convolution layer x and y in i-th layer,respectively. α and β indicate learned weights that satisfying by $\sum_{i=0}^{m} \sum_{j=1}^{n_i} \alpha_{ij} + \beta_{ij} = 1$. $\mu_{\tilde{x}^i_j}, \mu_{\tilde{y}^i_j}, \sigma^2_{\tilde{x}^i_j}, \sigma^2_{\tilde{y}^i_j}$ denote the global means and variances of $\tilde{x}^i_j$ and $\tilde{y}^i_j$, respectively. $\sigma_{\tilde{x}^i_j\tilde{y}^i_j}$ is a global covariance between $\tilde{x}^i_j$ and $\tilde{y}^i_j$. c_1 and c_2 are constants to numerical stability.

Table 1 demonstrates the accuracy of the results obtained through the use of the LPIPS metric (Equation 9). The table shows that when evaluating both regular and irregular missing regions, SE [46] has the lowest scores, and there is a noticeable color discrepancy in all the results, as depicted in Figures 7 and 8. In comparison, our approach exhibits higher scores compared to LP [13], MSP [12], EC-GAN [37], and SE [46]. Despite lower scores for MSP [12] in regular cases, as indicated in Table 1 and Figure 7, subjectively, MSP [12] outperforms LP [13] in certain parts. The metric has a significant impact on the evaluation of the smoothness of the entire restored region, as opposed to just specific areas, due to

TABLE 1. Accuracy of Learned Perceptual Image Patch Similarity (LPIPS).

Name	MSP [12]		LP [13]		EC-GAN [37]		SE [46]		Ours	
	Regular	Irregular	Regular	Irregular	Regular	Irregular	Regular	Irregular	Regular	Irregular
Vase	0.6700	0.6700	0.6970	0.7080	0.6670	0.6090	0.5690	0.3910	**0.8730**	**0.8760**
Pets	0.7430	0.7880	0.8390	0.8460	0.7180	0.7390	0.6700	0.4200	**0.9410**	**0.9410**
Woman	0.7340	0.7470	0.7470	0.7510	0.7390	0.7380	0.6200	0.4470	**0.8860**	**0.8910**
Child	0.6919	0.6600	0.7120	0.7190	0.6350	0.6120	0.5250	0.4320	**0.8190**	**0.8070**
Scenery	0.5944	0.5968	0.6720	0.6790	0.5780	0.5830	0.4500	0.4720	**0.8700**	**0.8760**
Girl	0.5951	0.6300	0.6760	0.6860	0.5880	0.6290	0.4540	0.4680	**0.7660**	**0.7680**

TABLE 2. Deep Image Structure and Texture Similarity (DISTS).

Name	MSP [12]		LP [13]		EC-GAN [37]		SE [46]		Ours	
	Regular	Irregular	Regular	Irregular	Regular	Irregular	Regular	Irregular	Regular	Irregular
Vase	0.9140	0.9168	0.9303	0.9264	0.9062	0.9153	0.7100	0.7535	**0.9556**	**0.9494**
Pets	0.9050	0.9529	0.9500	0.9623	0.8841	0.9510	0.6541	0.7354	**0.9749**	**0.9806**
Woman	0.9359	0.9480	0.9600	0.9643	0.9273	0.9478	0.7300	0.7692	**0.9670**	**0.9748**
Child	0.9085	0.9126	0.9113	0.9182	0.8937	0.9102	0.6503	0.7180	**0.9398**	**0.9248**
Scenery	0.8553	0.8720	0.9009	0.9084	0.8222	0.8636	0.5300	0.7815	**0.9404**	**0.9167**
Girl	0.9062	0.8862	0.9108	0.9148	0.9016	0.8769	0.6530	0.7245	**0.9410**	**0.9186**

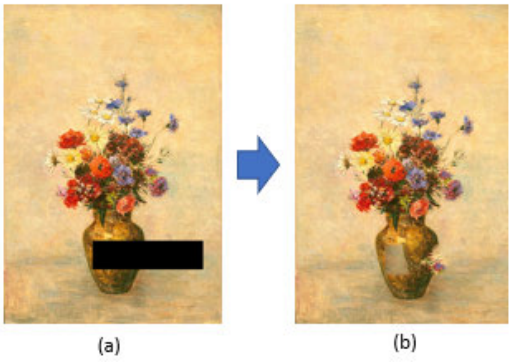

FIGURE 9. Failed Completion. (a) Input: An art painting with a large missing region, (b) Failed output contains large artifacts and blurriness.

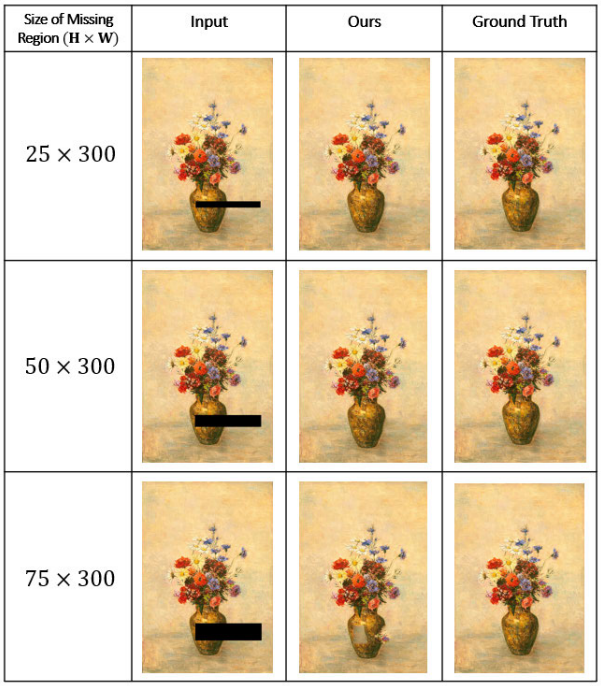

FIGURE 10. Comparison of different regular sizes.

the similar tuning configurations and perceptual calibrations of patch matching between unknown and known areas.

Additionally, Table 2 displays the accuracy results using the DISTS metric (Equation 12). While this comparison mainly focuses on structure and texture similarity, it still indicates that our algorithm has the highest score and SE [46] generates the lowest scores due to the presence of blurriness and significant color disparities in both regular and irregular cases.

V. DISCUSSION

As a human art that requires hand strokes to express art sense, painting has a significant context in every element. Based on our reliable approach, we attempted to tackle the damaged problems such as holes and scratches widely due to storage for a very long time in the museum. Integration methods of structure-texture completion emphasized an excellent technique of filling regular and irregular holes represented in the experimental result (Section IV). The massive hole of object boundaries was a challenging case that caused missing information among different regions with non-homogenous pixels. Before recovering the large unknown area, segmenting to smaller regions was an effective way to repair quickly using a more straightforward technique that only further focused on homogeneous pixels. Our structure diffusion executed this purpose well. We employed the Bezier curve method to help make the curve automatically in guiding our proposed dynamic propagation. As shown in our results, holes in the boundary area could be repaired smoothly like satisfying structures of the original images. The connected curve between endpoints around the missing region in dynamically

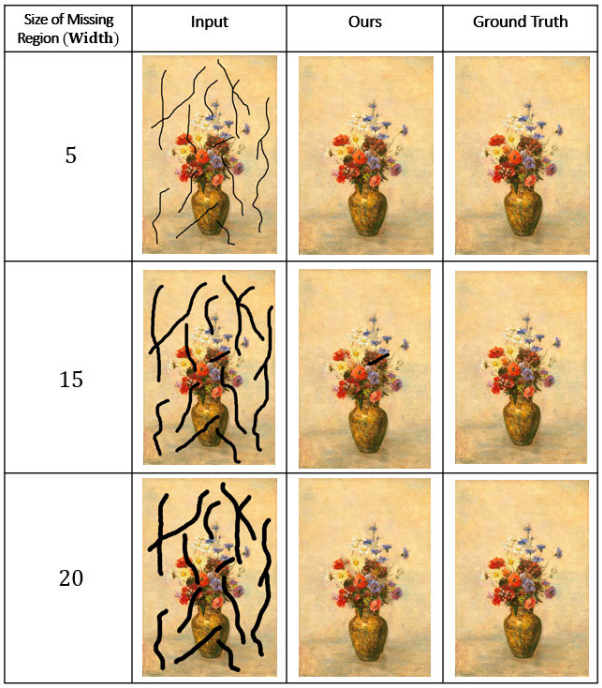

FIGURE 11. Comparison of different irregular sizes.

TABLE 3. Accuracy of Different Regular Sizes.

Size of Missing Region (H × W)	LPIPS	DISTS
25 × 300	**0.8730**	**0.9556**
50 × 300	0.700	0.9148
75 × 300	0.688	0.8977

guiding the patch-based propagation from knowns to an unknown area. Stamping patches considered the closer direction of patches by rotation was an important technique. However, when the missing region was too large and produced fewer similar pixel information in the known region, the patch-based propagation could not correctly spread suitable patches to repair the holes along a curve. This error generated some significant artifacts and blurriness, as shown in Figure 9.

Depending on our limitations, we also evaluated how well our method handles holes of different sizes. As shown in Figures 10 and 11, we tested regular holes of sizes 25 × 300, 50 × 300, and 75 × 300 pixels and irregular holes with widths of 5, 15, and 20 pixels. The results displayed that the larger holes resulted in lower quality restoration, as indicated by the accuracy metrics LPIPS and DISTS in Tables 3 and 4. According to the perceptual and scoring results, the optimal size for the regular hole was 25 × 300 pixels, and for the irregular hole was a width of 5 pixels. We applied these appropriate hole sizes to our scenarios of experimental results in Figures 7 and 8.

The smaller homogeneous regions could be effectively recovered by our superior approach, Cohesive Laplacian Fusion. A combination of Laplacian and Gaussian pyramids had the right portion in plausibly creating the restored texture image. The proposed minimum normalized distance metric of approximating the most identical patches between unknown and known areas enhanced the search algorithm. Moreover, our advanced weighted vote of overlapping colors covered all potential nearest neighbor pixels and comprised the restored pixel of the current unknown patch. This intention produced a proficient way to generate magnificent restoration that outperformed other compared methods.

TABLE 4. Accuracy of Different Irregular Sizes.

Size of Missing Region (Width)	LPIPS	DISTS
5	**0.8760**	**0.9494**
15	0.676	0.9203
20	0.501	0.9033

In addition, we selected several related approaches that considered edge as prior guidance for fair evaluations, such as MSP [12] employed a user-specified curve to segment the large hole manually, LP [13] utilized Laplacian pyramid in awareness of edge structures, EC-GAN [37] applied an edge generator followed by an image inpainting network, and SE [46] presented the sketch of users to help manipulation tasks. Although these methods completed the missing region in various damaged cases, compared results showed that our method generated more excellent restored fields and was closer to the original images. MSP [12] method looked similar to our method, which detected structural plan and texture propagation in accomplishing inpainting tasks. However, we improved their method to get better results. For instance, we applied dynamic curve automation of structure diffusion as guidance instead of user-defined curves as was applied in MSP [12]. Then we employed voting for the similarity function of nearest neighbor pixels overlapping in Laplacian and Gaussian pyramid systems.

According to the comparison between regular and irregular holes, such as based on human visual perception in Figure 7 and 8, and accuracy scores in Table 1 and Table 2, our results were better in the irregular case than regular. Despite the missing region of the large block (regular hole) was not much in total compared to an arbitrary issue (irregular holes), it had a larger area. Arbitrary scratches spread universally in an image, but the width of the holes was not too spacious. Our method utilized similar proposed methods in restoring both kinds of holes. However, the capability of patch-matching in pyramid systems for larger areas is more in demand, so the risk of failed voting is likely to occur.

VI. CONCLUSION

This paper introduced a novel structure-texture art painting completion that used a Cohesive Laplacian Fusion algorithm. In segmenting the non-homogeneous into smaller homogeneous unknown regions, structure diffusion worked well in satisfying propagated structures dynamically corresponding to patches along with a guided curve. After segmented homogeneous holes were generated, an enhanced method by

texture synthesis was applied to restore the remaining small missing regions. Before constructing the vote of overlapping pixels in each patch, we applied the minimum normalized distance of searching patch similarity that emphasized the consistency of each pyramid layer. Then, we determined the optimal pixels by employing the weighted similarity function of all potential nearest neighbor pixels. Finally, we have conducted experiments comparing state-of-the-art approaches to image completion with various kinds of damaged painting. Our method outperformed and was more competitive than compared methods based on human visual perception and accuracy metrics.

In the future, we plan to design a system when repairing the missing region that considers the art senses, such as strokes, lines, and brush movements. It will be challenging because integrating feeling habits and painting techniques are a complicated issue. In addition, we will attempt to solve painting completion with more large missing regions as input and present exciting opportunities.

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. 27th Annu. Conf. Comput. Graph. Interact. Techn.-SIGGRAPH*, 2000, pp. 417–424, doi: 10.1145/344779.344972.

[2] S. Masnou and J.-M. Morel, "Level lines based disocclusion," in *Proc. Int. Conf. Image Process.*, 1998, pp. 259–263, doi: 10.1109/ICIP.1998.999016.

[3] T. K. Shih, R.-C. Chang, L.-C. Lu, and H.-C. Huang, "Multi-layer inpainting on Chinese artwork," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jun. 2004, pp. 21–24, doi: 10.1109/ICME.2004.1394115.

[4] H. Wei and W. Shuwen, "Dunhuang murals inpainting based on image decomposition," in *Proc. 3rd Int. Conf. Comput. Sci. Inf. Technol.*, Jul. 2010, pp. 397–400, doi: 10.1109/ICCSIT.2010.5564944.

[5] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, no. 1, pp. 23–34, 2004, doi: 10.1080/10867651.2004.10487596.

[6] I. N. Sari, Y. Urano, and W. Du, "Image inpainting using orthogonal viewpoints and structure consistency in Manhattan world," in *Proc. the 8th Int. Virtual Conf. Appl. Comput. Inf. Technol.*, Jun. 2021, pp. 93–98, doi: 10.1145/3468081.3471129.

[7] E. Horikawa, I. N. Sari, and W. Du, "Image inpainting using clustered planar structure guidance," in *Proc. 8th Int. Virtual Conf. Appl. Comput. Inf. Technol.*, Jun. 2021, pp. 118–123, doi: 10.1145/3468081.3471133.

[8] I. N. Sari, E. Horikawa, and W. Du, "Interactive image inpainting of large-scale missing region," *IEEE Access*, vol. 9, pp. 56430–56442, 2021, doi: 10.1109/ACCESS.2021.3072366.

[9] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004, doi: 10.1109/TIP.2004.833105.

[10] S. Z. Siadati, F. Yaghmaee, and P. Mahdavi, "A new exemplar-based image inpainting algorithm using image structure tensors," in *Proc. 24th Iranian Conf. Electr. Eng. (ICEE)*, May 2016, pp. 995–1001, doi: 10.1109/IRANIANCEE.2016.7585666.

[11] D. J. Tuptewar and A. Pinjarkar, "Robust exemplar based image and video inpainting for object removal and region filling," in *Proc. Int. Conf. Intell. Comput. Control (I2C2)*, Jun. 2017, pp. 1–4, doi: 10.1109/I2C2.2017.8321964.

[12] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, "Image completion with structure propagation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 861–868, Jul. 2005, doi: 10.1145/1073204.1073274.

[13] J. H. Lee, I. Choi, and M. H. Kim, "Laplacian patch-based image synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2727–2735, doi: 10.1109/CVPR.2016.298.

[14] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–11, Jul. 2009, doi: 10.1145/1531326.1531330.

[15] Y. Ma, X. Liu, S. Bai, L. Wang, A. Liu, D. Tao, and E. Hancock, "Region-wise generative adversarial ImageInpainting for large missing areas," 2019, *arXiv:1909.12507*.

[16] J. Li, N. Wang, L. Zhang, B. Du, and D. Tao, "Recurrent feature reasoning for image inpainting," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7760–7768, doi: 10.1109/CVPR42600.2020.00778.

[17] G. Liu, F. A. Reda, K. J. Shih, T. C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 85–100.

[18] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with Fourier convolutions," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2022, pp. 2149–2159, doi: 10.1109/WACV51458.2022.00323.

[19] J. Youngjoo and P. Jongyoul, "SC-FEGAN: Face editing generative adversarial network with user's sketch and color," in *Proc. ICCV*, 2019, pp. 1745–1753.

[20] Y. Zeng, Z. Lin, J. Yang, J. Zhang, E. Shechtman, and H. Lu, "High-resolution image inpainting with iterative confidence feedback and guided upsampling," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 1–17.

[21] Y. Song, C. Yang, Y. Shen, P. Wang, Q. Huang, and C.-C. J. Kuo, "SPG-Net: Segmentation prediction and guidance network for image inpainting," 2018, *arXiv:1805.03356*.

[22] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5505–5514.

[23] Z. Yi, Q. Tang, S. Azizi, D. Jang, and Z. Xu, "Contextual residual aggregation for ultra high-resolution image inpainting," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7508–7517.

[24] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–14, Aug. 2017, doi: 10.1145/3072959.3073659.

[25] W. Zuo and Z. Lin, "A generalized accelerated proximal gradient approach for total-variation-based image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2748–2759, Oct. 2011, doi: 10.1109/TIP.2011.2131665.

[26] J. Dahl, P. C. Hansen, S. H. Jensen, and T. L. Jensen, "Algorithms and software for total variation image reconstruction via first-order methods," *Numer. Algorithms*, vol. 53, no. 1, pp. 67–92, Jan. 2010, doi: 10.1007/S11075-009-9310-3.

[27] T. F. Chan and J. Shen, "Nontexture inpainting by curvature-driven diffusions," *J. Vis. Commun. Image Represent.*, vol. 12, no. 4, pp. 436–449, 2001, doi: 10.1006/JVCI.2001.0487.

[28] T. F. Chan, S. H. Kang, and J. Shen, "Total variation denoising and enhancement of color images based on the CB and HSV color models," *J. Vis. Commun. Image Represent.*, vol. 12, no. 4, pp. 422–435, Dec. 2001, doi: 10.1006/JVCI.2001.0491.

[29] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2536–2544, doi: 10.1109/CVPR.2016.278.

[30] W. Zhang, P. Zhuang, H.-H. Sun, G. Li, S. Kwong, and C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," *IEEE Trans. Image Process.*, vol. 31, pp. 3997–4010, 2022, doi: 10.1109/TIP.2022.3177129.

[31] C. Li, C. Guo, L. Han, J. Jiang, M.-M. Cheng, J. Gu, and C. C. Loy, "Low-light image and video enhancement using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 9396–9416, Dec. 2022, doi: 10.1109/TPAMI.2021.3126387.

[32] W. Zhang, Y. Wang, and C. Li, "Underwater image enhancement by attenuated color channel correction and detail preserved contrast enhancement," *IEEE J. Ocean. Eng.*, vol. 47, no. 3, pp. 718–735, Jul. 2022, doi: 10.1109/JOE.2022.3140563.

[33] C. Li, C. Guo, and C. L. Chen, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4225–4238, Aug. 2021, doi: 10.1109/TPAMI.2021.3063604.

[34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 10684–10695.

[35] H. Zheng, Z. Lin, J. Lu, S. Cohen, E. Shechtman, C. Barnes, J. Zhang, N. Xu, S. Amirghodsi, and J. Luo, "CM-GAN: Image inpainting with cascaded modulation GAN and object-aware training," 2022, *arXiv:2203.11947*.
[36] K. Masaoka, I. N. Sari, and W. Du, "Edge-enhanced GAN with vanishing points for image inpainting," in *Proc. 23rd ACIS Int. Summer Virtual Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput.*, Jul. 2022, doi: 10.1109/SNPD-Summer57817.2022.00027.
[37] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, "EdgeConnect: Generative image inpainting with adversarial edge learning," 2019, *arXiv:1901.00212*.
[38] H. Liu, Z. Wan, W. Huang, Y. Song, X. Han, and J. Liao, "PD-GAN: Probabilistic diverse GAN for image inpainting," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 9371–9381.
[39] Y. Urano, I. N. Sari, and W. Du, "Image inpainting using automatic structure propagation with auxiliary line construction," in *Proc. 23rd ACIS Int. Summer Virtual Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput.*, Jul. 2022, doi: 10.1109/SNPD-Summer57817.2022.00026.
[40] M. Chen, X. Zhao, and D. Xu, "Image inpainting for digital Dunhuang murals using partial convolutions and sliding window method," *J. Phys., Conf. Ser.*, vol. 1302, no. 3, Aug. 2019, Art. no. 032040, doi: 10.1088/1742-6596/1302/3/032040.
[41] H. Wang, Q. Li, and Q. Zou, "Inpainting of Dunhuang murals by sparsely modeling the texture similarity and structure continuity," *J. Comput. Cultural Heritage*, vol. 12, no. 3, pp. 1–21, Oct. 2019, doi: 10.1145/3280790.
[42] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CPVR)*, Jul. 2015, pp. 1395–1403.
[43] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Commun. ACM*, vol. 27, no. 3, pp. 236–239, 1984, doi: 10.1145/357994.358023.
[44] D. Páleš and J. Rédl, "Derivations of the Bezier curve," *Math. Educ., Res. Appl.*, vol. 2, no. 1, pp. 1–7, Oct. 2016, doi: 10.15414/MERAA.2016.02.01.01-07.
[45] G. Farin, "The de Casteljau algorithm," in *Curves and Surfaces for CAGD* (The Morgan Kaufmann Series in Computer Graphics), 5th ed. USA: Elsevier, 2002, pp. 43–55, doi: 10.1016/B978-155860737-8/50004-1.
[46] Y. Zeng, Z. Lin, and V. M. Patel, "SketchEdit: Mask-free local image manipulation with partial sketches," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5951–5961, doi: 10.1109/CVPR52688.2022.00586.
[47] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?" in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, May 2002, pp. 3313–3316, doi: 10.1109/ICASSP.2002.5745362.
[48] M. A. Qureshi, M. Deriche, A. Beghdadi, and A. Amin, "A critical survey of state-of-the-art image inpainting quality assessment metrics," *J. Vis. Commun. Image Represent.*, vol. 49, pp. 177–191, Nov. 2017, doi: 10.1016/J.JVCIR.2017.09.006.
[49] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 586–595.
[50] K. Ding, K. Ma, S. Wang, and E. E. Simoncell, "Image quality assessment: Unifying structure and texture similarity," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2020, pp. 2567–2581.

IRAWATI NURMALA SARI received the B.S. degree from the Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 2012, and the M.S. degree from the National Taiwan University of Science and Technology (NTUST), Taiwan, in 2015. She is currently pursuing the Ph.D. degree with the Department of Information Science, Kyoto Institute of Technology (KIT), Japan. She was a Research Student with the Visual Information Laboratory, Kyoto Institute of Technology (KIT), in 2019. Her research interests include image inpainting, image processing, computer vision, pattern recognition, and machine learning.

WEIWEI DU (Member, IEEE) received the B.S. degree from the Tianjin Institute of Urban Construction, China, in 2002, the master's degree from The University of Aizu, Japan, in 2005, and the Ph.D. degree from Kyushu University, Japan, in 2008. In 2008, she joined the Kyoto Institute of Technology as an Assistant Professor. She has been a Full Associate Professor of information and human science with the Kyoto Institute of Technology, since 2019. Her research interests include pattern recognition, machine learning, neural networks, image processing, and computer vision.