Research Article

Images Inpainting Quality Evaluation Using Structural Features and Visual Saliency

Shuang Ma [1] **and Jinhe Liu**[2]

[1]*School of Automation, Beijing Information Science and Technology University, Beijing 100192, China*
[2]*Beijing Shiny Tech., Co, Ltd., Beijing 100143, China*

Correspondence should be addressed to Shuang Ma; lanxi1279@126.com

Despite the extensive research on developing robust image inpainting algorithms in recent years, there are almost no objective metrics for the quality assessment of inpainted images currently. Inspired by the feature coherence in the inpainted image and the human visual perception mechanism, this paper proposes an image inpainting quality assessment (IIQA) that takes into account both visual saliency and structural features. First, the quality issues associated with image inpainting are categorized into three aspects: incoherent structure, unreasonable texture, and other results that are inconsistent with human visual perception. These quality problems are further expressed as "regions of interest" and extracted by the visual saliency method using the natural statistics model. Subsequently, the structural features are computed based on the nonlinear diffusion of the horizontal and vertical gradient field of the inpainted image. Finally, the IIQA metric incorporates brightness, gradient similarity, structural similarity, and visual saliency is established. The quality evaluation process is conducted by comparing each patch within the inpainted region with its best match from the known region. The quantitative experimental results demonstrate the effectiveness of the proposed method, especially for images with structural discontinuity. A comparative study also shows that the Spearman rank order correlation coefficient of our method achieves 0.875 on certain databases, which outperforms existing IIQA metrics.

1. Introduction

The primary goal of image inpainting is to carefully select appropriate patches from the known region and subsequently utilize them to fill in the repaired area, thereby achieving a comprehensive restoration of the entire image. It involves estimating and speculating about the missing information. With the development of science and modern technology, methods such as super-resolution [1], attention mechanism [2], and deep learning [3] have been widely applied to image inpainting and achieved satisfactory results. However, inpainted image quality assessment (IIQA) currently depends on observers' subjective expertise. For observers, as long as there are no noticeable visual imperfections in the repaired image, the image inpainting task is considered to be successful. Despite the extensive research conducted on the development of robust image inpainting algorithms, there has been a lack of focus on creating effective quality assessment metrics to evaluate the performance of image inpainting

techniques. Due to the absence of reference images for comparison, the commonly used metrics for assessing image quality, such as mean squared error and peak signal-to-noise ratio, cannot be directly applied.

In recent years, efforts have been made by scholars to tackle these concerns using techniques such as feature analysis and machine learning. Wang et al. [4] believed that human eyes are highly adaptive to the changes in the structure information in images. Therefore, an indicator named structural similarity image measurement (SSIM) was proposed to measure the similarity between the reference image and the distorted image. Feie [5] proposed an image evaluation indicator, Borsal, to estimate the attention density of the sight in a narrow band around the target boundary. Meanwhile, they also proposed the StructBorSal index, which combined SSIM with visual saliency to evaluate the quality of structural regions in the inpainted image. Liu et al. [6] also proposed an image quality assessment metric based on the variant of SSIM. When it comes to the quality evaluation of

inpainted images, SSIM may not be reliable for images containing large inpainted regions as these areas can significantly deviate from the actual ones. Qureshi et al. [7] classified IIQA methods into three categories: structure-based methods, saliency-based methods, and machine learning-based methods. The performances of typical inpainting approaches were also tested in [7]. Hu et al. [8] proposed a nonreference quality evaluation metric for the Thangka image. The structural features of the Thangka image were extracted, and the differences between the original region and the repaired region were further analyzed. However, this metric is only available for Thangka images. Liu et al. [9] introduced the preattention theory to emulate visual perception by refining luminance-channel data and developed an image quality assessment metric of distorted images. The study in [10] utilized sample-based image completion methods to generate a set of 100 inpainted images. Subsequently, an evaluation metric based on the saliency map and repaired region was proposed and discussed. Unfortunately, only part of the image features were considered in [9, 10], and they performed well only in some special cases. Additionally, a significant limitation shared by the aforementioned metrics is the need for a reference image, which is not available for image inpainting.

Image quality evaluation approaches based on machine learning offer effective functional approximations of image features and evaluation scores [11]. Voronin et al. [12] proposed to extract the local binary pattern features of images and collect quality rating scores from subjective experiments. These data are utilized to train a support vector regression network for image quality prediction. The database consisted of 300 images, and a total of 10 participants were engaged in conducting subjective assessment. The findings demonstrated that the objective evaluation results have a strong connection with human perception. Isogawa et al. [13] proposed a learning-based sorting method to automatically estimate the optimal parameters for image inpainting algorithms. The relative ranking of different inpainted images was determined by employing a learning-to-rank approach. Meng et al. [14] introduced a novel approach for assessing the performance of nonreference image inpainting algorithms using deep rank learning. A pair of inpainted images were taken as input, and their ranking order was predicted. This method can be applied for the evaluation of both inpainted images and inpainting algorithms. Madhusudana et al. [15] tried to solve the problem of image quality assessment by using a deep convolutional neural network. Images containing synthetic and realistic distortions were taken as the database for the prediction of distortion type and degree. This method is more applicable for the quality assessment of images affected by synthetic or authentic distortions. Chen et al. [16] first applied unlabeled data to conduct self-supervised pretraining for blind image quality assessment. Distorted images are generated from high-quality samples and taken as the database. Additionally, a contrastive loss function is introduced to capture information that is sensitive to image quality. For machine learning-based image quality assessment methods, it is crucial to have subjectively and manually annotated rating scores for the training of regression models. In addition,

the current no reference-based image quality assessment algorithms rely on artificially created distorted image features. The learning phase needs a substantial amount of labeled samples; however, there are currently no public databases available for different machine learning-based IIQA methods.

The challenge in assessing the quality of image inpainting lies in the fact that the inpainting procedures might result in visible anomalies within and surrounding the inpainted regions. IIQA is essentially based on the consistency of the repaired region and the known region. Although reference images do not exist, available information can still be extracted from the known region and used for quality evaluation, which also makes feature extraction and feature comparison key factors for image quality assessment. Hence, we propose a new inpainted image quality assessment metric that reformulates visual inconsistency problems as human attention alterations and quantifies the similarity between inpainted regions and their best-match patches in the known region. The main contributions of the present work are as follows:

(i) We propose to extract the "regions of interest" in the inpainted image by applying an improved visual saliency method using natural statistics model (ISUN).

(ii) The structure map of the inpainted image is extracted through the nonlinear diffusion of the horizontal and vertical gradient field.

(iii) The IIQA metric involves brightness, gradient similarity, structure similarity, and visual saliency is proposed and tested on different inpainted images.

The rest of this paper is organized as follows: the visual saliency of inpainted images is calculated in Section 2, and the IIQA metric is established in Section 3. Experimental results and discussion are presented in Section 4, and the conclusion is drawn in Section 5.

2. The Extraction of "Regions of Interest" in the Inpainted Image

The reason why human eyes can recognize various image quality problems is that the inpainted regions often contain features that are not similar to their surroundings or even the entire image. As shown in Figure 1, we summarized some typical problems in the inpainted image. The main factors that affect the quality of image inpainting are structural inconsistencies, unreasonable texture, and other issues that do not align with human visual perception.

In the inpainted image, it is expected that the inpainted region will exhibit a coherent structure and texture to the known region. Two structural inconsistency problems are illustrated in the red squares in Figure 1. The edge of the water is discontinuous in Figure 1(a), and there is an unknown structure, a special circle, in Figure 1(b). According to human perception, the interlacing between different textures, the repetition of the same texture in a large area, and the block effect should be avoided. Examples of typical texture problems are given in Figures 1(c) and 1(d).
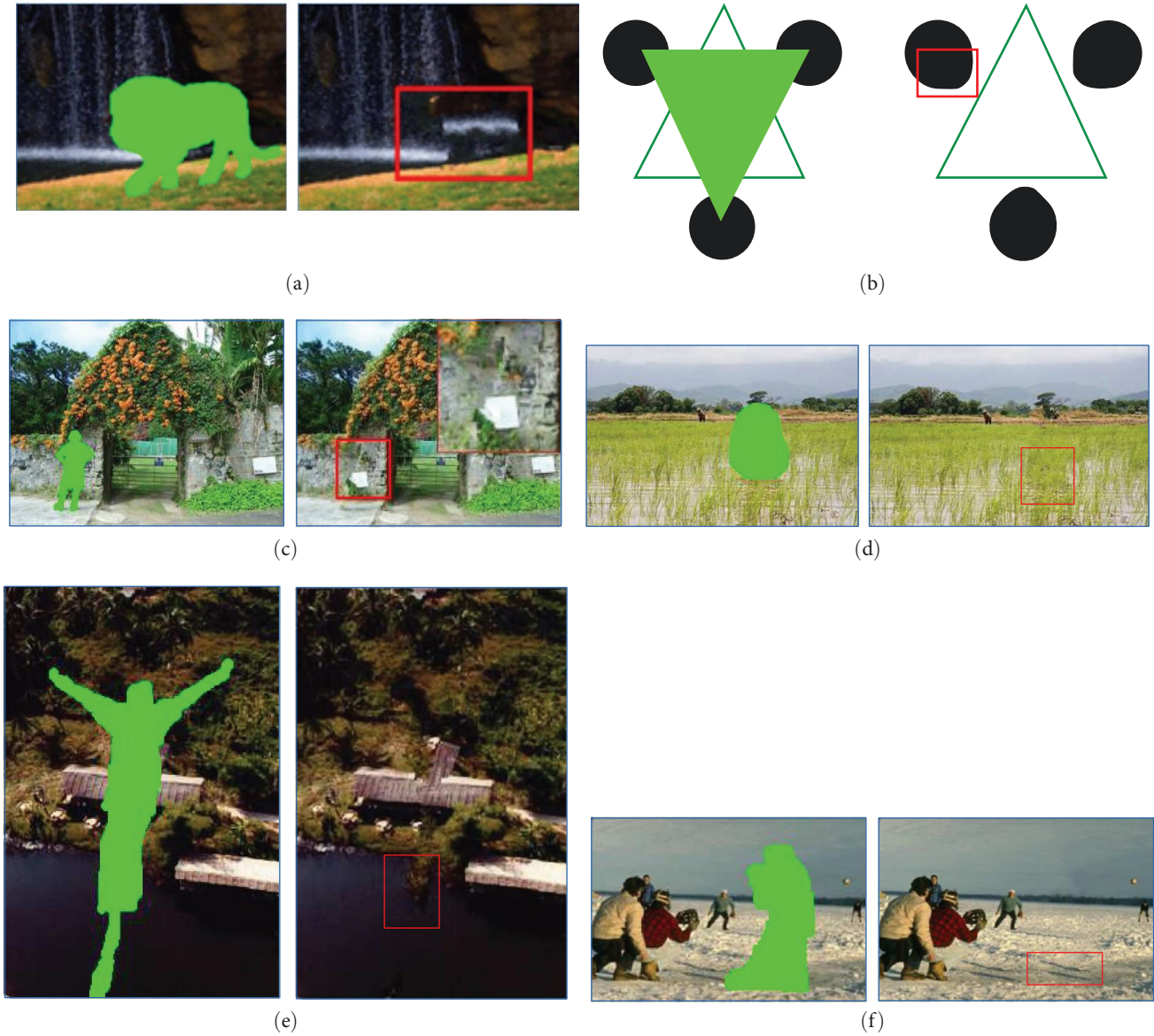
FIGURE 1: Examples of typical problems in image inpainting: (a) incoherent structure; (b) structure inconsistent with human cognition; (c) staggered texture; (d) overly repetitive texture; (e) abnormal result 1; (f) abnormal result 2.

Although texture irrationality is not as obvious as structure problems for human perception, it still has a great impact on the image quality. In addition, there are still some other problems, as shown in Figures 1(e) and 1(f). Although there is no structural discontinuity or texture irrationality in these images, we still believe that the plants in Figure 1(e) should not multiply infinitely in the water; also, no objects can cast a shadow at the given position in Figure 1(f).

It is not difficult to find that these areas in Figure 1 are "regions of interest" that will attract human attention easily. Therefore, the human visual perception model is applied to extract "regions of interest" in the inpainted image. In this way, traditional image evaluation approaches, which include feature extraction, feature comparison, and other processes, can be transformed into human visual attention alterations in the inpainted region. Image features are always used to

measure whether the region can attract human attention or not. Typically, regions with sparser features or more distinctive features compared to their surroundings exhibit higher saliency. In this paper, we introduced the ISUN statistics model [17], which utilizes the inherent statistical properties of images to extract the distorted region in inpainted images. In ISUN, massive natural images in a specific image library are taken as sample sets. Subsequently, the unsupervised learning of sample images is used to obtain basic independent component analysis (ICA) filters, which can represent image characteristics. Based on the probability distribution of the ICA feature set, we can get the saliency map of an image.

In ISUN, image features are acquired by applying ICA and principal component analysis (PCA) to a group of pre-processed images from a database [17]. For each image, 100

patches with size $b \times b \times 3$ from random locations are extracted, and each patch is treated as a $3b^2$ dimensional vector. The dimensionality of the patch collection is first reduced using PCA, and then the number of image features is limited to d. Afterward, ICA filters are applied to these patches to get the learned filters. In this way, features can be extracted from an image by filtering it with each of the learned ICA filters. As discussed in the ISUN model, the visual saliency of an image is defined as $P(X)^{-1}$ [9], where X represents the obtained ICA feature set. The components of X are statistically independent after ICA processing, so $P(X)^{-1}$ can be described as a product of a series of 1D distributions, as shown in Equation (1).

$$P(X = x) = \prod_i P(x_i), \tag{1}$$

where x_i is the ith element in vector x, i.e., the ith feature. The generalized Gauss distribution (GGD) given in Equation (2) is applied to simulate each 1D distribution shown in Equation (1).

$$P(x_i) = \frac{\theta_i}{2\sigma_i \Gamma(\theta_i^{-1})} \exp\left(-\left|\frac{x_i}{\sigma_i}\right|^{\theta_i}\right), \tag{2}$$

where $\theta_i > 0$ is the shape function, $\sigma_i > 0$ is the scale function and Γ is the gamma function. For each ICA filter, select a suitable 1D GGD function and estimate the parameters of the GGD function using the algorithm mentioned in [5]. Each feature is assigned a weight based on its frequency of occurrence to enhance the visual saliency of rare features. The flowchart of saliency map extraction is shown in Figure 2.

Taking the image to be impaired in Figure 3(a) as an example, the methods proposed in [18, 19] are applied to repair this image, and the results are shown in Figure 3(b). The visual saliency map of each inpainted image is extracted using the ISUN model, and the results are shown in Figure 3(c). Obviously, the incoherent structure and unreasonable contents in the inpainted region have notable saliency. The visual saliency map of the inpainted region is directly related to image quality. Both the saliency region size and the corresponding saliency degree indicate where image inpainting has not been effectively performed. In addition, not all regions in an image are equally important to human eyes, according to the visual cognitive theory. Take Figure 3(c), for example; human visual perception exhibits greater sensitivity toward structural breaks, shields, and discontinuities, while displaying relatively lower sensitivity toward texture details. Thus, structural features should be emphasized in image quality assessment.

3. IIQA Metric Using Structural Features and Visual Saliency

The SSIM proposed by Wang et al. [4] provided an effective way for image quality evaluation from the perspective of structural distortion. In SSIM, the mean, standard deviation, and covariance of an image are utilized as estimations for
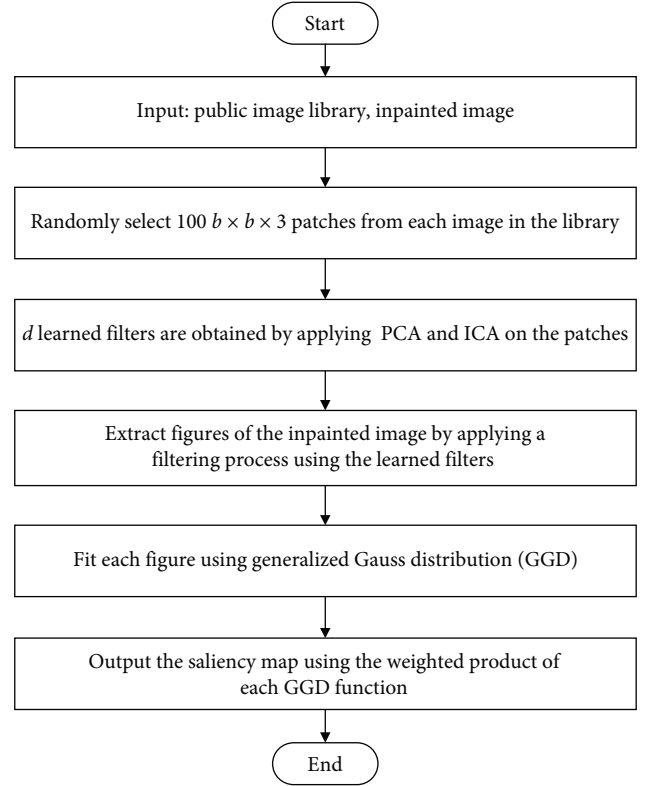

FIGURE 2: The flowchart of saliency map extraction.

brightness, contrast, and structural similarity, respectively. The SSIM metric is defined in Equation (3).

$$\text{SSIM}(f, f') = [l(f, f')]^\alpha \times [c(f, f')]^\beta \times [s(f, f')]^\gamma, \tag{3}$$

where f and f' represent the reference image and the image under evaluation. The parameters α, β, and γ serve to calibrate the weight of different indicators. In general cases, $\alpha = \beta = \gamma = 1$. The brightness function $l(f, f')$, contrast function $c(f, f')$, and structure function $s(f, f')$ are illustrated in Equations (4)–(6).

$$l(f, f') = \frac{2\mu_f \mu_{f'} + C_1}{\mu_f^2 + \mu_{f'}^2 + C_1}, \tag{4}$$

$$c(f, f') = \frac{2\sigma_f \sigma_{f'} + C_2}{\sigma_f^2 + \sigma_{f'}^2 + C_2}, \tag{5}$$

$$s(f, f') = \frac{\sigma_{ff'} + C_3}{\sigma_f \sigma_{f'} + C_3}, \tag{6}$$

where $\sigma_{ff'} = \frac{1}{N-1} \sum_{i=1}^{N} (f_i - \mu_f)(f_i' - \mu_{f'})$, C_1, C_2, and C_3 are constant, μ_f and $\mu_{f'}$ represent the average brightness of two images, σ_f and $\sigma_{f'}$ denote the standard variance of two images. According to Equations (3)–(6), structural similarity remains invariant to changes in contrast and brightness. However, SSIM computes the brightness, contrast, and structural similarity indiscriminately for all image regions. As
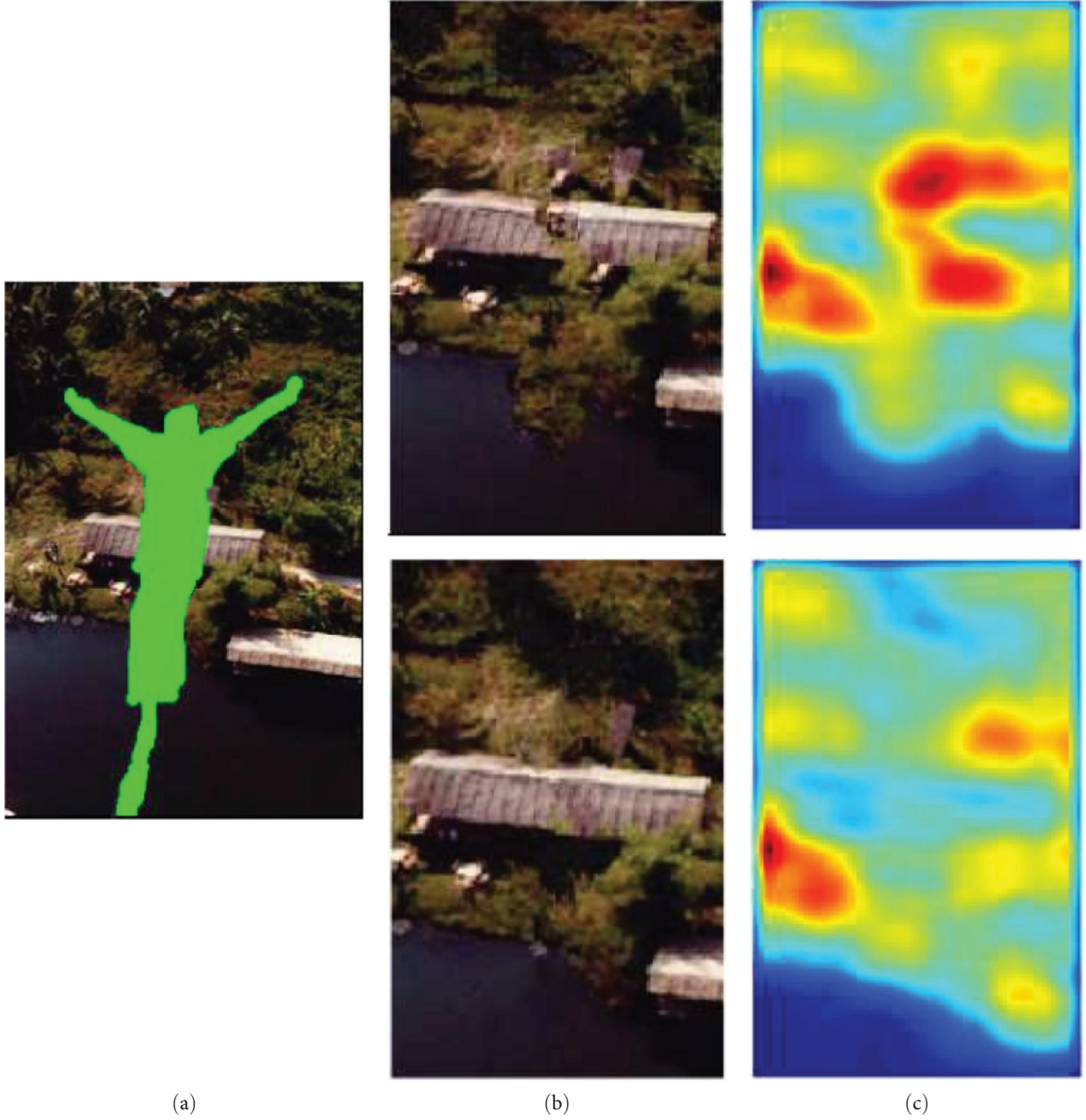
(a) (b) (c)

FIGURE 3: The saliency map of an inpainted image: (a) original image; (b) inpainted image; (c) saliency map.

discussed in Section 2, the subjective sensitivity of human eyes to texture regions and structure regions is quite different. In this paper, the spatial variation of image gradient features extracted by the structure tensor method is considered to replace the contrast function and structure function in SSIM.

The anisotropic diffusion-based structure tensor method [20] effectively enhances structure contours by employing nonlinear diffusion in color space, as well as horizontal, vertical, and 45° gradient spaces, thereby reducing feature redundancy and computational complexity. For an image I (x, y), the feature space extracted by the structural tensor method is given in Equation (7).

$$J_\sigma = K_\sigma \times (\nabla I \nabla I^T) = \begin{pmatrix} K_\sigma \times I_x^2 & K_\sigma \times I_x I_y \\ K_\sigma \times I_x I_y & K_\sigma \times I_y^2 \end{pmatrix}, \quad (7)$$

where K_σ is a Gaussian kernel with variance σ, I_x is the horizontal gradient, and I_y is the vertical gradient of an image. The convolution of K_σ with different gradient fields can smooth the gradient features in different directions. Unfortunately, the convolution with the Gaussian kernel often leads to the misalignment of target boundaries in the feature space. In order to solve the problem caused by the linear filter, the Gaussian convolution can be replaced by the nonlinear diffusion [21], as given in Equation (8).

$$\partial_t u_i = \operatorname{div}\left(g\left(\sum_{k=1}^{N_1} |\nabla u_k|^2 \right) \nabla u_i \right), \forall i, \tag{8}$$

where N_1 represents the number of gradient channels, we set $N_1 = 2$ in this paper, which means only the gradient features in the horizontal and vertical directions are considered. g is the diffusion coefficient with $g(|\nabla u|) = (|\nabla u| + \xi)^{-1}$, ξ is a small positive number and we set $\xi = e^{-10}$ here. Additive operator separation [22] is used to solve Equation (8). In the initial moment, $t = 0$, u_1^0 and u_2^0 represents the horizontal gradient $I_x^2 = [I(i, j) - I(i + 1, j)]^2$ and the vertical gradient $I_y^2 = [I(i, j) - I(i, j + 1)]^2$. When $t = K$, iteration stops, and features are recorded as u_1^K and u_2^K. Gradient feature U of image $I(x, y)$ can be obtained by adding the horizontal gradient and vertical gradient, as given in Equation (9).

$$U = \frac{1}{\nabla I}(u_1^K + u_2^K), \tag{9}$$

where $(\nabla I)^{-1}$ is a coefficient used to adjust the value range of U. Based on the gradient feature, the gradient similarity of images f and f' is defined in Equation (10).

$$d(f, f') = \frac{2\varphi_f \varphi_{f'} + C_4}{\varphi_f^2 + \varphi_{f'}^2 + C_4}, \tag{10}$$

where the value of φ_f and $\varphi_{f'}$ represent the average gradient features U_f and $U_{f'}$, C_4 is a smaller constant. The structural similarity function based on gradient features is defined in Equation (11).

$$s'(f, f') = \frac{\varphi_{ff'} + C_5}{\varphi_f \varphi_{f'} + C_5}, \tag{11}$$

where $\varphi_{ff'} = \frac{1}{N-1} \sum_{i=1}^{N} (U_{f,i} - \varphi_f)(U_{f',i} - \varphi_{f'})$. Accordingly, the proposed IIQA metric is defined as follows:

$$\mathrm{IIQA}(f, f') = [l(f, f')]^\alpha \times [d(f, f')]^\beta \times [s'(f, f')]^\gamma. \tag{12}$$

It is worth noting that in image inpainting, the reference image f does not exist for quality evaluation. We propose to measure the quality of image inpainting by comparing the inpainted region with the known region. For each pixel p in the repaired region, a patch $M(p)$ centered at p with size $w \times w$ is selected. Subsequently, the proposed IIQA is employed to search for the best match M of patch $M(p)$ within the known region. The similarity between $M(p)$ and the known region is defined as follows:

$$\mathrm{IIQA}_p = \min_{X \in \Phi_C, X(p) \in \Phi_S} \mathrm{IIQA}(M(p), M), \tag{13}$$

where Φ_C represents the known region, and Φ_S represents the inpainted region. As discussed in Section 2, human eyes are more sensitive to the fracture, occlusion, and discontinuity of the structure in the image rather than texture or other details.

Input: Image $I(x, y)$
Output: Q
$\quad u_1^0 = I_x, \ u_2^0 = I_y, \ k = 0$
if $k < \mathrm{K}$
$\quad u_i \leftarrow u_i^k, \ v_i \leftarrow K_\sigma \times u_i, \ g \leftarrow (\sum_{k=1}^{2} |\nabla v_i^k|^2 + \xi)^{-1}$
$\quad u_i \leftarrow v_{ix} + v_{iy}, \ u_i^{k+1} \leftarrow u_i, \ k \leftarrow k + 1$
else $U \leftarrow \frac{1}{\nabla I}(u_1^K + u_2^K)$
for all $M(p)\varepsilon\Phi_S$ **do**
$\quad M' \leftarrow argmin_{X \in \Phi_C, X(p) \in \Phi_S} \mathrm{IIQA}(M(p), M)$
$\quad \mathrm{IIQA}_p \leftarrow \mathrm{IIQA}(M(p), M')$
end for
$\quad Q \leftarrow \sum_{p=1}^{p} S_p \cdot \mathrm{IIQA}_p$

ALGORITHM 1: Quality Assessment of Inpainted Images.

The objective quality evaluation of image inpainting is thus defined as the linear weighted sum of saliency degree and IIQA_p for all patches within the inpainted region. Let S_p be the normalized saliency degree of pixel p in the inpainted region, and the quality Q of an inpainted image considering visual saliency and structural features is given in Equation (14).

$$Q = \sum_{p=1}^{P} S_p \cdot \mathrm{IIQA}_p, \tag{14}$$

where P is the number of pixels in the inpainted region. The IIQA algorithm is described in detail as follows:

4. Results and Discussion

Typical images shown in Figure 4(a) are taken as examples and repaired by the Criminisi algorithm [23], PAMSRIC algorithm [24], priority-BP algorithm [25], and DLIC algorithm [26], respectively. Sixteen inpainted images are obtained for quality evaluation, as shown in Figures 4(b) and 4(c). For subjective rating, 18 observers were selected to score the inpainted image according to the criteria given in Table 1. The proposed approach was employed for objective evaluation. The Pearson linear correlation coefficient (PLCC) and the Spearman rank order correlation coefficient (SROCC) [27] are applied to test the performance of IIQA.

As shown in Figure 4, inpainted images are classified into group (b) and group (c). The average subjective rating scores from 18 observers are shown in Figure 5(a). The objective scores using the proposed IIOA method are given in Figure 5(b). The value of PLCC and SROCC between the subjective scores and objective scores is shown in Table 2.

As shown in Table 2, the correlation between subjective and objective evaluation scores is very high, indicating that the objective evaluation results are basically consistent with the subjective rating. These results verified the effectiveness of the proposed IIQA method. It is also worth noting that the PLCC and SROCC coefficients of images in group (c) are lower than those of images in group (b). This is because some

(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

Image 1

(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

Image 2

(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

Image 3

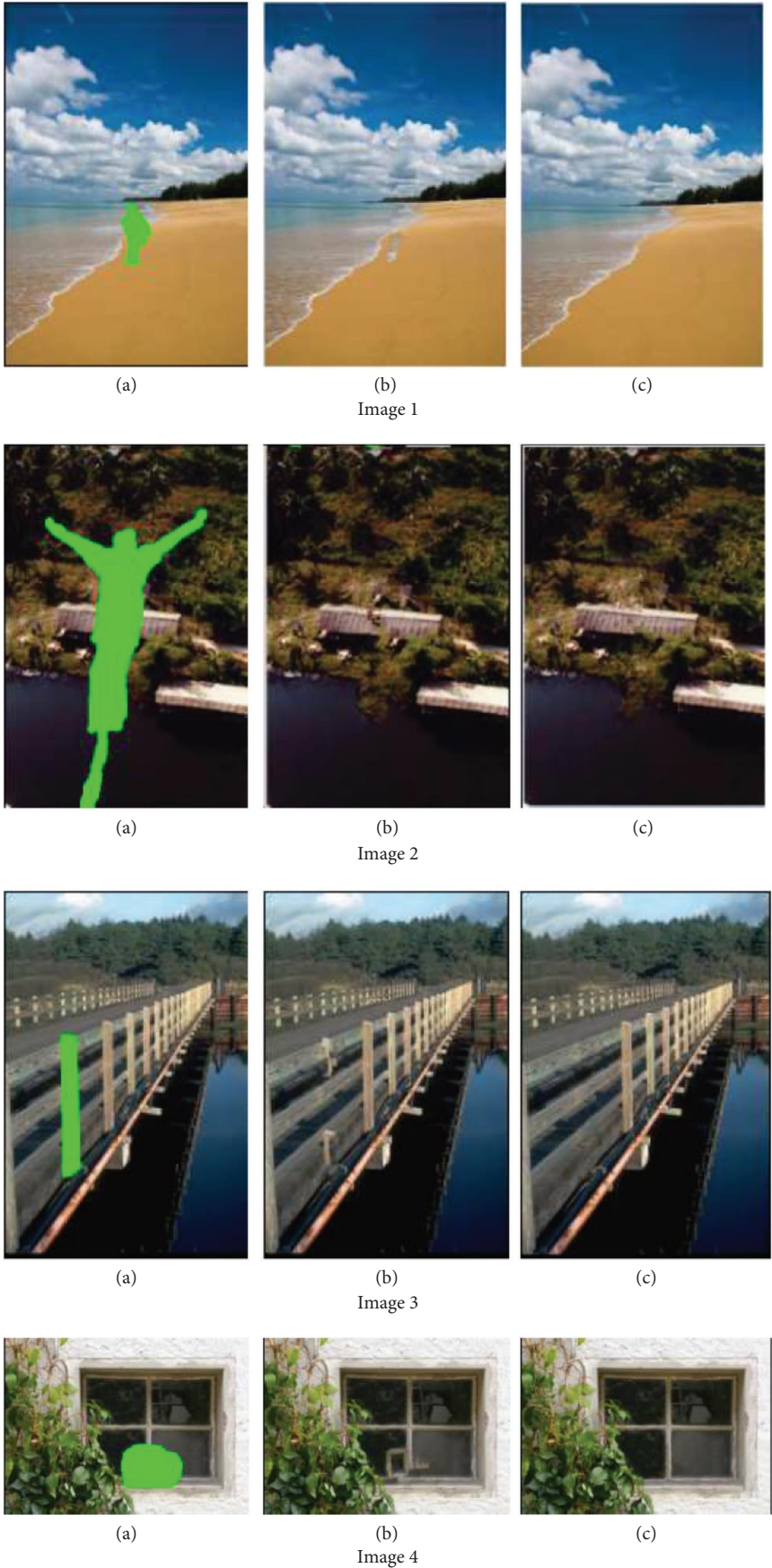
(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)
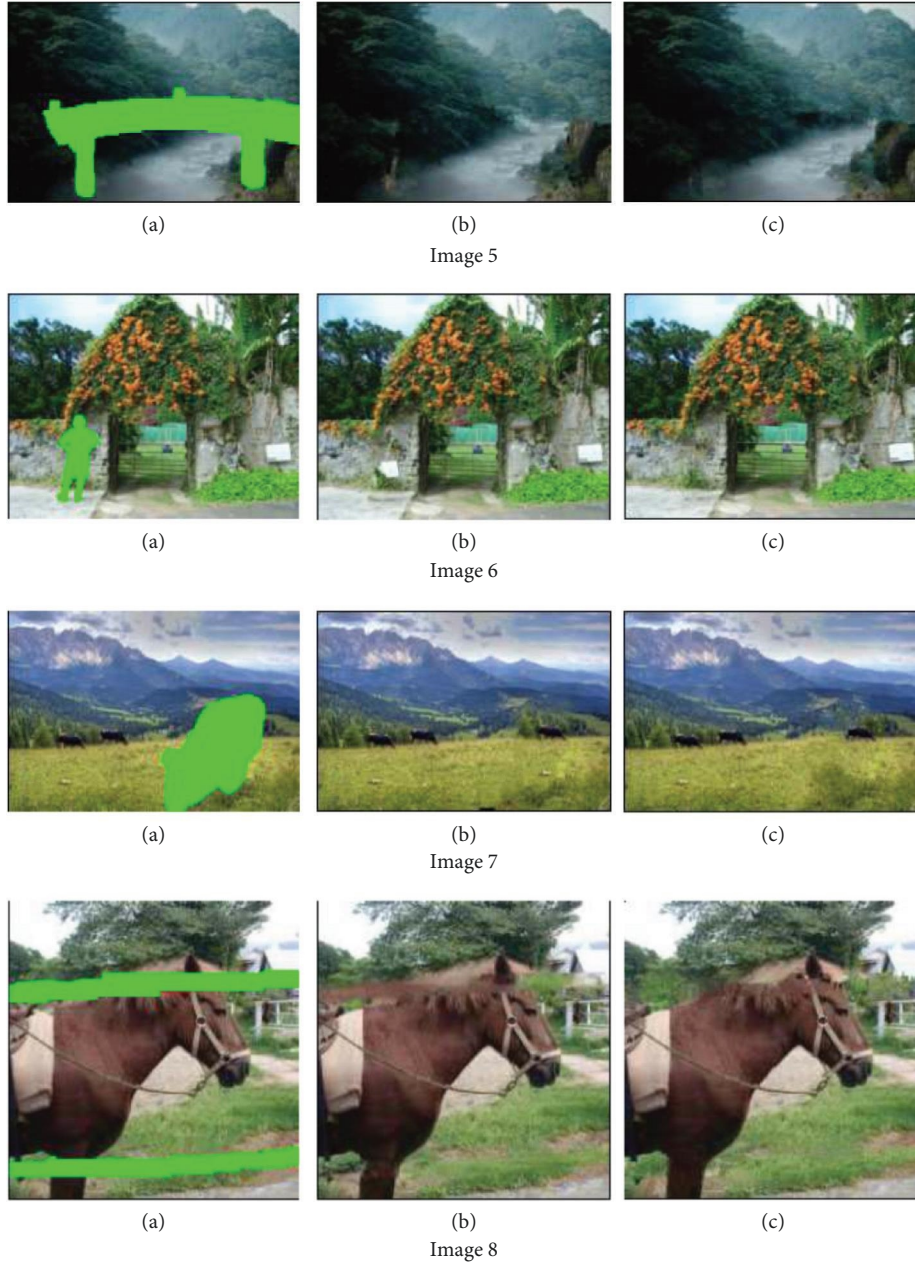
Image 4

Figure 4: Continued.

FIGURE 4: Collection of inpainted images. Image 1: (a) original image, (b) Criminisi's, and (c) PAMSRIC; Image 2: (a) original image, (b) Criminisi's, and (c) PAMSRIC; Image 3: (a) original image, (b) Criminisi's, and (c) DLIC; Image 4: (a) original image, (b) Criminisi's, and (c) PAMSRIC; Image 5: (a) original image, (b) Criminisi's, and (c) DLIC; Image 6: (a) original image, (b) Criminisi's, and (c) DLIC; Image 7: (a) original image, (b) priority-BP, and (c) DLIC; Image 8: (a) original image, (b) priority-BP, and (c) DLIC.

TABLE 1: The criteria for subjective rating.

Subjective perception	Score
Successful	5
Preferably	4
Common	3
Unsatisfying	2
Failed	1

inpainted images in group (b) are obviously not in line with the subjective perception of human eyes and are easier to be noticed by observers. However, the inpainted images in group (c) had no obvious structure or texture problems, which resulted in a certain difference between the subjective rating and objective evaluation. This is because, rather than texture issues, our approach exhibits a higher sensitivity toward structural defects that tend to capture human attention more easily.
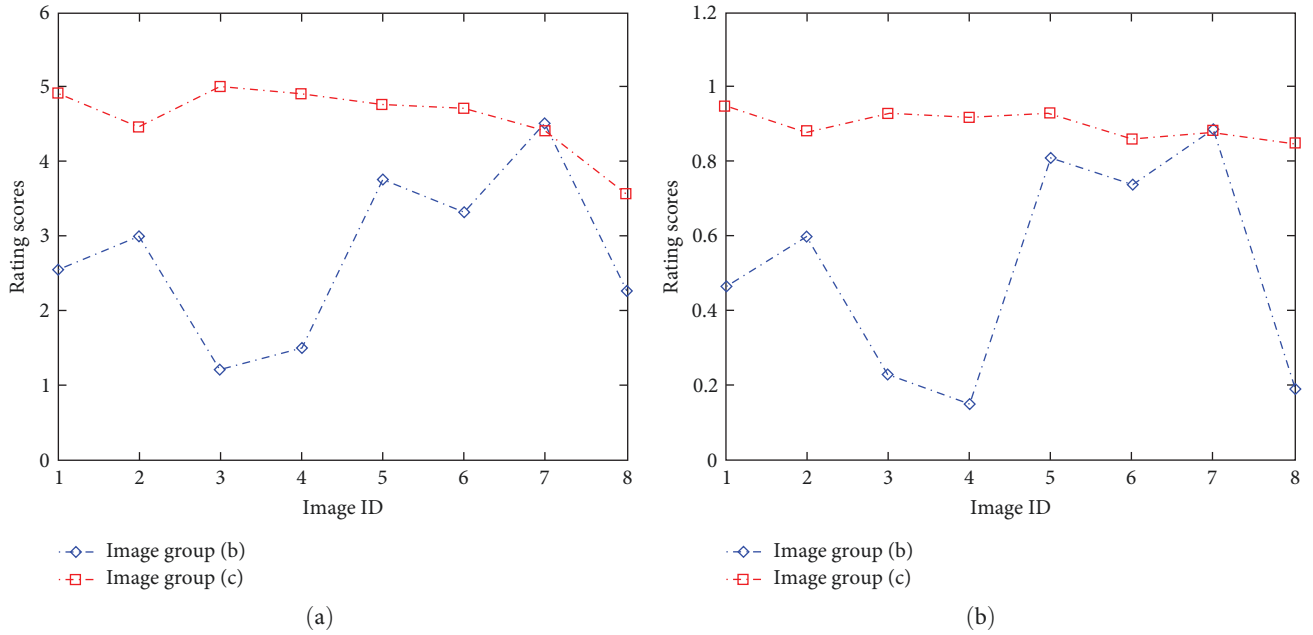
FIGURE 5: The rating scores of inpainted images: (a) subjective rating scores; (b) objective rating scores.

TABLE 2: Correlation coefficient between subjective and objective scores.

Image group	PLCC	SROCC
(b)	0.9432	0.9286
(c)	0.7662	0.8121

TABLE 3: Performance comparison of IIQA metrics.

IIQA metrics	SROCC
StrucBorSal	0.740
OIM	0.834
B-IIQA	0.840
Our method	0.875

The results in Figure 5 also indicate that the proposed approach does not directly depend on the size of the inpainted regions. For example, the inpainted regions are large in image 5 and image 7 compared to others; however, both the subjective and objective rating scores of these two images are relatively high. In contrast, the inpainted regions in image 1 and image 4 are smaller, but the rating scores of images 1(b) and 4(b) are much lower. The reason for this lies in the fact that our IIQA metric incorporates both visual saliency and structural features, thereby yielding consistent results with subjective evaluation, particularly for images with structural problems.

To further verify the performance of the proposed approach, we tried to compare our method to the related works. Unfortunately, IQA methods based on machine learning were trained using datasets generated locally, making it challenging to compare other methods with those metrics. Therefore, we chose StrucBorSal [5], OIM [10], and B-IIQA [28] for the comparative study. All these methods were applied on a public database TUM-IID which contains 272 inpainted images. The obtained SROCC values of these approaches are summarized in Table 3. Obviously the SROCC value of our method is higher than other metrics, indicating the effectiveness and superior performance of the proposed IIQA metric.

5. Conclusions

The absence of a reference image has consistently posed a significant challenge for the IIQA. In this paper, we proposed to address the IIQA problem by employing image feature extraction techniques and visual saliency analysis. The texture and structure defects in the inpainted image were converted into "regions of interest" and analyzed through the extraction of a visual saliency map. Based on the structure feature obtained by the nonlinear diffusion of gradient space, the IIQA metric, including brightness, gradient similarity, and structural similarity, was proposed and tested on different types of images. Experimental results showed that the rating scores by our method are consistent with the subjective rating results, especially for images with structural defects. The proposed metric also performed better than the existing ones. The limitation of our method is that when there are no obvious structure or texture problems in the inpainted image, the result showed a lower correlation with objective evaluation. In the future, our focus will be on IIQA metrics that integrate both human perception and deep learning to enhance its performance.

Data Availability

The data used to support the study are openly available in a public repository.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Chen, R. Xia, K. Yang, and K. Zou, "MICU: image super-resolution via multi-level information compensation and U-net," *Expert Systems with Applications*, vol. 245, Article ID 123111, 2024.

[2] Y. Chen, R. Xia, K. Yang, and K. Zou, "MFMAM: image inpainting via multi-scale feature module with attention module," *Computer Vision and Image Understanding*, vol. 238, Article ID 103883, 2024.

[3] Y. Chen, R. Xia, K. Yang, and K. Zou, "DGCA: high resolution image inpainting via DR-GAN and contextual attention," *Multimedia Tools and Applications*, vol. 82, no. 30, pp. 47751–47771, 2023.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[5] A. I. Feie, *Digital inpainting for artwork restoration: algorithms and evaluation*, Ph. D. Thesis, Gjøvik University College, Norway, 2012.

[6] T.-J. Liu, Y.-C. Lin, W. Lin, and C.-C. J. Kuo, "Visual quality assessment: recent developments, coding applications and future trends," *APSIPA Transactions on Signal and Information Processing*, vol. 2, no. 1, pp. 1–20, 2013.

[7] M. A. Qureshi, M. Deriche, A. Beghdadi, and A. Amin, "A critical survey of state-of-the-art image inpainting quality assessment metrics," *Journal of Visual Communication and Image Representation*, vol. 49, no. 11, pp. 177–191, 2017.

[8] W. Hu, Y. Ye, F. Zeng, and J. Meng, "A new method of Thangka image inpainting quality assessment," *Journal of Visual Communication and Image Representation*, vol. 59, pp. 292–299, 2019.

[9] L. Liu, T. Wang, and H. Huang, "Pre-attention and spatial dependency driven no-reference image quality assessment," *IEEE Transactions on Multimedia*, vol. 21, no. 9, pp. 2305–2318, 2019.

[10] D. Amirkhani and A. Bastanfard, "An objective method to evaluate exemplar-based inpainted images quality using Jaccard index," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26199–26212, 2021.

[11] V. A. Frantc, V. V. Voronin, V. I. Marchuk, A. I. Sherstobitov, S. Agaian, and K. Egiazarian, "Machine learning approach for objective inpainting quality assessment," *Mobile Multimedia/Image Processing Security, and Applications*, vol. 9120, pp. 229–237, 2014.

[12] V. V. Voronin, V. A. Frantc, V. I. Marchuk, A. I. Sherstobitov, and K. Egiazarian, "No-reference visual quality assessment for image inpainting," *Image Processing: Algorithms and Systems XIII*, vol. 9399, pp. 234–241, 2015.

[13] M. Isogawa, D. Mikami, K. Takahashi, and H. Kimata, "Image quality assessment for inpainted images via learning to rank," *Multimedia Tools and Applications*, vol. 78, no. 2, pp. 1399–1418, 2019.

[14] X. Meng, W. Ma, C. Li, and Q. Mi, "Siamese CNN-based rank learning for quality assessment of inpainted images," *Journal of Visual Communication and Image Representation*, vol. 78, Article ID 103176, 2021.

[15] P. C. Madhusudana, N. Birkbeck, Y. Wang, B. Adsumilli, and A. C. Bovik, "Image quality assessment using contrastive learning," *IEEE Transactions on Image Processing*, vol. 31, pp. 4149–4161, 2022.

[16] P. Chen, L. Li, Q. Wu, and J. Wu, "SPIQ: a self-supervised pre-trained model for image quality assessment," *IEEE Signal Processing Letters*, vol. 29, pp. 513–517, 2022.

[17] C. Kanan and G. Cottrell, "Robust classification of objects, faces, and flowers using natural image statistics," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2472–2479, IEEE, San Francisco, CA, USA, 2010.

[18] T. T. Dang, A. Beghdadi, and M. C. Larabi, "Perceptual evaluation of digital image completion quality," in *21st European Signal Processing Conference (EUSIPCO 2013)*, pp. 1–5, IEEE, Marrakech, Morocco, 2013.

[19] T. T. Dang, A. Beghdadi, and M. C. Larabi, "Visual coherence metric for evaluation of color image restoration," in *2013 Colour and Visual Computing Symposium (CVCS)*, pp. 1–6, IEEE, Gjovik, Norway, 2013.

[20] D. Chen, J. Li, S. Di et al., "Critical points extraction from building façades by analyzing gradient structure tensor," *Remote Sensing*, vol. 13, no. 16, Article ID 3146, 2021.

[21] S. Arridge and A. Hauptmann, "Networks for nonlinear diffusion problems in imaging," *Journal of Mathematical Imaging and Vision*, vol. 62, no. 3, pp. 471–487, 2020.

[22] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 398–410, 1998.

[23] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[24] G. Xu and S. Ma, "Image completion using dynamic-scale patch matching and layer-wise chain optimization," *Acta Electronica Sinica*, vol. 43, no. 3, pp. 529–535, 2015.

[25] N. Komodakis and G. Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2649–2661, 2007.

[26] S. Ma, Y. P. Tan, and G. Xu, "Image completion based on fusion of patch associated matching and low-rank matrix super resolution," *Journal of Computer-Aided Design & Computer Graphics*, vol. 27, no. 2, pp. 271–278, 2015.

[27] X. L. Chen, *Research on image quality evaluation technology based on visual features*, Ph.D. Thesis, Shanghai Jiaotong University, Shanghai, 2012.

[28] A. M. Rezki, A. Serir, and A. Beghdadi, "Blind image inpainting quality assessment using local features continuity," *Multimedia Tools and Applications*, vol. 81, no. 7, pp. 9225–9244, 2022.